STUDIA UNIV. BABEŞ–BOLYAI, INFORMATICA, Volume **XLVII**, Number 2, 2002

# ON THE CONVERGENCE OF ASYNCHRONOUS BLOCK NEWTON METHODS FOR NONLINEAR SYSTEMS OF EQUATIONS

IOAN LAZĂR

Abstract. Convergence of asynchronous block Newton methods for solving nonlinear systems of equations of the form $F(x) = 0$ are studied. Sufficient conditions to guarantee their local convergence are given. Our analysis emphasizes the connection between the conditions on $F$ involved in local convergence theorems for sequential and synchronous block Newton's method, and our settings for asynchronous block Newton methods. Our results are similar to the results of Szyld and Xu, obtained in an asynchronous nonlinear multisplitting context.

**Keywords**: numerical analysis, iterative methods, nonlinear system of equations, Newton methods

## 1. Introduction

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

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

where $F = (f_1, \ldots, f_n) : \Omega \subseteq \mathbf{R}^n \to \mathbf{R}^n$ is a nonlinear operator. Newton's method is based on the approximation $F(x) \approx F(x^k) + F'(x^k)(x - x^k)$, and is given by the iteration

$$x^{k+1} = x^k - F'(x^k)^{-1}F(x^k), \tag{2}$$

for $k = 0, 1, \ldots$, where $x^0$ is an initial guess.

Each linear system (2) can be solved in parallel using some kind of block iterative methods [10]. Block iterative methods are studied using the concept of multisplittings [10] and the application of block iterative methods for solving the systems (2) at each Newton step $k$ was considered in [14].

The application of the concept of multisplittings directly to the nonlinear system (1) were considered in [6] and [2] These methods are called parallel synchronous nonlinear multisplitting methods. The methods are called synchronous in the sense

---

2000 *Mathematics Subject Classification.* 65H10, 65Y05.

1998 *CR Categories and Descriptors.* G.1.0 [**Numerical analysis**] General: *parallel algorithms*; G.1.5 [**Numerical analysis**] Roots of Nonlinear Equations: *Iterative methods, Convergence, Systems of equations*.

that all processors have to wait at some synchronization point before proceeding to the next iteration.

The asynchronous nonlinear multisplitting methods were considered in [1] and [12], i.e. methods where no synchronization barrier is present (see [5, 3, 7] for some general discussions on asynchronous methods). Bahi et all [1] studied asynchronous nonlinear multisplitting methods in a general context for nonlinear fixed point problems, while Szyld and Xu [12] studied these methods for problems of the form (1), and extended the study to the case of overlapping blocks, i.e., certain variables are updated by more than one processors.

Our framework presented here is similar to the framework used by Xu [15] for the study of asynchronous block quasi-Newton methods. Our analysis emphasizes the connection between the conditions on $F$ involved in local convergence theorems for sequential and synchronous block Newton's method, and those used for asynchronous block Newton methods.

This paper is organized as follows: in section 2 we give a brief review of block Newton methods, a computational model for asynchronous block (Newton) methods and a corresponding mathematical model. The main result is presented in section 3, after a brief review of the tools used or study both synchronous and asynchronous cases. Finally some connections with different asynchronous block Newton type methods are discussed.

## 2. Asynchronous Block Newton Methods

Suppose $F$ and $x$ are conformally partitioned as follows $F = (F_1, \ldots, F_L)$, $x = (x_1, \ldots, x_L)$, where $x_i = (x_{i_1}, \ldots, x_{i_{n_i}}) \in \mathbf{R}^{n_i}$ and $F_i : \mathbf{R}^n \to \mathbf{R}^{n_i}$, $i = 1, \ldots, L$. Suppose the partition $S_i = \{i_1, \ldots, i_{n_i}\}$, $i = 1, \ldots, L$ is chosen such that $\cup_{i=1}^{L} S_i = \{1, \ldots, n\}$ and $S_i \cap S_j = \emptyset$ for $i \neq j$, $i, j = 1, \ldots, L$.

The system (1) can be rewritten

$$F_l(x_1, \ldots, x_l, \ldots, x_L) = 0, l = 1, \ldots, L. \tag{3}$$

and we consider the following nonlinear block method. Given initial values $x = (x_1, \ldots, x_L)$, repeat the following procedure until convergence

For $l = 1, \ldots, L$

$$\begin{cases} \text{Solve for } y \text{ in } F_l(x_1, \ldots, x_{l-1}, y, x_{l+1}, \ldots, x_L) = 0, \\ \text{Set } x_l = y. \end{cases} \tag{4}$$

In (4) the order in which the block are updated could be arbitrary. The classical nonlinear block-Jacobi method and block-Gauss-Seidel method [4, 11] are special cases of such methods. For the purpose of parallel processing the nonlinear block-Jacobi method is nearly ideal, since up to $L$ processors can each perform one of the iterations in (4). Such iterations are synchronous in the sense that to begin the computation of the next iterate, each processor has to wait until all processors have completed their current iteration. By removing the synchronization and letting the pocessors continue their calculations according to the information currently available, we obtain asynchronous parallel methods.

Let the Jacobian of $F$ be partitioned conformally with $x$, and define

$$F'(x) = \begin{pmatrix} \frac{\partial F_1(x)}{\partial x_1} & \dots & \frac{\partial F_1(x)}{\partial x_L} \\ \dots & \dots & \dots \\ \frac{\partial F_L(x)}{\partial x_1} & \dots & \frac{\partial F_L(x)}{\partial x_L} \end{pmatrix}, \frac{\partial F_i(x)}{\partial x_j} = \begin{pmatrix} \frac{\partial f_{i_1}(x)}{\partial x_{j_1}} & \dots & \frac{\partial f_{i_1}(x)}{\partial x_{j_{n_j}}} \\ \dots & \dots & \dots \\ \frac{\partial f_{i_{n_i}}(x)}{\partial x_{j_1}} & \dots & \frac{\partial f_{i_{n_i}}(x)}{\partial x_{j_{n_j}}} \end{pmatrix},$$

The block diagonal matrix of $F'(x)$ is denoted by

$$D(x) = diag(\frac{\partial F_1(x)}{\partial x_1}, \dots, \frac{\partial F_L(x)}{\partial x_L}).$$

Using the above notations, one Newton step applied to the system (4) and starting from the initial value $x$ is

$$y_l = x_l - \left(\frac{\partial F_l(x)}{\partial x_l}\right)^{-1} F_l(x). \tag{5}$$

Note that when solving (5) we are only interested in the components $x_l$ corresponding to $S_l$. This means we work with a system of dimension $n_l$, although the initial system (1) is of dimension $n$. The evaluation of $F_l(x)$ in (5) is dependent on the entire vector $x$, that is the processor solving the equation (5) needs the components evaluated by other processors.

2.1. **Computational Model.** Denote the (approximate) solution of $F_l(x_1, \dots, x_{l-1}, y, x_{l+1}, \dots, x_L) = 0$ by $y_l = G_l(x)$, $l = 1, \dots, L$. Applying one step of Newton method gives the operator defined by (5).

Assume we are working with a (shared memory) parallel computer with $L$ processors and associate a block of components with each processor. Then a parallel variant of (3) can be implemented as in Algorithm 1. If the processors would wait for each other to complete each run through the loop we would get a parallel synchronous implementation of the procedure (3).

Here the processors continue the loop by collecting the needed vectors computed by the other processors according to the information available at the moment. A computational model for the asynchronous block method can be written as the pseudocod of Algorithm 1 shows.

Since the processors do not wait for each other, the processors get out of phase due to different run times for each loop. At a given time point, different processors will have achieved different number of iterations. In this context, the iteration number $k$ in (2) looses its meaning.

Using a direct linear solver for step 4 in Algorithm 1, for exemple an $LU$ factorization of $F'$, we obtain the *asynchronous block Newton method*

4'a: Factor $F_l'(x) = LU$
4'b: Solve $LUs = -F_l(x)$
4'c: $y_l := x_l + s$

Any other appropriate factorization such as $QR$ or Cholesky could be used as well.

**Algorithm 1** Pseudocode for the $l$th processor ($l = 1, \dots, L$). $x$ represents the initial guess $x_j$, $j = 1, \dots, L$. $x$ and *convergence* are global variables written in common memory.

```
read(converge)
while not converge do
    read(x)
    y_l = G_l(x)
    x_l := y_l; overwrite(x_l)
    read(converge);
end while
```

2.2. **Mathematical Model.** In order to analyse the asynchronous computational model presented in Algorithm 1 we consider a counter $k$ which is updated every time a new vector is computed by some processor and let $x_l^0 = x^0$, $l = 1, \dots, L$.

Let $I^k \subseteq \{1, \dots, L\}$ denotes all updated block components, then the asynchronous block Newton iteration is defined by

$$x_i^{k+1} = \begin{cases} x_i^{s_i(k)} - \left(\frac{\partial F_i(u)}{\partial x_i}\right)^{-1} F_i(u) & \text{for } i \in I^k, \\ x_i^k & \text{for } i \notin I^k, \end{cases} \tag{6}$$

for $i \in \{1, \dots, L\}$, $k = 0, 1, \dots$, where $u = (x_1^{s_1(k)}, \dots, x_L^{s_L(k)})$.

The iteration counts $s_i(k)$, $i = 1, \dots, L$ indicate the iteration, prior to $k$, when the $i$th block component was computed.

Let $S = \{(s_1(k), \ldots, s_L(k)) \in \mathbf{N}^L\}_{k \in \mathbf{N}}$ where $\mathbf{N} = 0, 1, \dots$ denotes the set of natural numbers. The *standard assumptions* for $I = \{I^k\}_{k \in \mathbf{N}}$ and $S$ are:

$$\forall i \in \{1, \dots, L\}, \forall k \in \mathbf{N}, s_i(k) \leq k, \tag{7}$$

$$\forall i \in \{1, \dots, L\}, \lim_{k \to \infty} s_i(k) = \infty, \tag{8}$$

$$\forall i \in \{1, \dots, L\}, \text{ the set } \{k \in \mathbf{N} | i \in I^k\} \text{ is infinite.} \tag{9}$$

The next definitions are similar to those considered by El Tarazi in [13] and will be used in our proofs.

We define the sequence $\{s(k)\}_{k \in \mathbf{N}} \subset \mathbf{N}$ by

$$s(k) = \min_i s_i(k). \tag{10}$$

We obtain immediately from (7) and (8)

$$s(k) \leq k \text{ and } \lim_{k \to \infty} s(k) = \infty. \tag{11}$$

Suppose that (7)–(9) are satisfied, then we can define an increasing sequence $\{k_l\}_{l\in\mathbf{N}}$ having the properties

$$\bigcup_{0\le s(k)\le k<k_0} I^k = \{1,\ldots,L\}, \tag{12}$$

$$\bigcup_{k_l\le s(k)\le k<k_{l+1}} I^k = \{1,\ldots,L\}. \tag{13}$$

The proofs given by Baudet [3] and El Tarazi [13] for general asynchronous iterations use the sequence $\{k_l\}$ defined above. This sequence says that the asynchronous iteration (6) updates all block components at least once at the steps $k_0, k_1, \ldots$.

If $k_{l+1} - k_l = L$ for all $l$, we get a synchronous block Gauss-Seidel iteration, and if the sequence of differences $\{k_{l+1} - k_l\}$ is bounded then we get a partially asynchronous algorithm.

## 3. Local Convergence

3.1. **Synchronous Newton Methods.** The standard assumptions on $F$ in synchronous (or sequential) case are:

- **(C1):** Equation (1) has a solution $x^*$.
- **(C2):** $F' : \Omega \to \mathbf{R}^{n\times n}$ is Lipschitz continuous on $\Omega$, with Lipschitz constant $\gamma$, i.e., $\|F'(x) - F'(y)\| \le \gamma\|x-y\|,$ for all $x, y \in \Omega$.
- **(C3):** $F'(x^*)$ is nonsingular.

These assumptions can be weakened without sacrificing convergence results presented here. However the classical result on quadratic convergence of Newton's method requires them.

The main result concerning the local convergence of Newton's method is presented in the next theorem.

**Theorem 3.1.** [11] *Let the standard assumptions (C1)–(C3) hold. Then there are $K > 0$ and $\delta > 0$ such that if $\|x^0 - x^*\| < \delta$ then the Newton iterates $\{x^k\}$ defined by (2) converge q-quadratically to the solution $x^*$ of (1).*

The convergence results on Newton's method follow from the basic results given in Lemma 3.2 and 3.3 (a variant of Banach lemma).

**Lemma 3.2.** [11] *Assume $F$ satisfies (C2). Then for all $x, y \in \Omega$,*

$$\|F(y) - F(x) - F'(x)(y-x)\| \le \frac{\gamma}{2}\|x-y\|^2. \tag{14}$$

In the context of Theorem 3.1, the inequality (14) is used to obtain the estimates

$$\|x^{k+1} - x^*\| \le K\|x^k - x^*\|^2,\ k = 0, 1, \ldots. \tag{15}$$

**Lemma 3.3.** [11] *Let* $A, C \in \boldsymbol{R}^{n\times n}$*, A nonsingular and* $\|A^{-1}\| \leq \alpha_1$*,* $\|C - A\| \leq \alpha_2$ *with* $\alpha_1\alpha_2 < 1$*. Then C is a nonsingular matrix, and*

$$\|C^{-1}\| \leq \frac{\alpha_1}{1 - \alpha_1\alpha_2}.$$

The hypothesis of theorem 3.1 does not give sufficient conditions for solving subsystems $F_l(x) = 0$, $l = 1, \dots, L$ of the system $F(x) = 0$, since the subsystem $F_l(x) = 0$ is solved only in respect to the components of block $l$.

The asynchronous iteration (6) is more close related to other Newton type methods which consider some splitting of the Jacobian,

$$F'(x) = B(x) - C(x). \tag{16}$$

and iterative processes

$$x^{k+1} = x^k - B(x^k)^{-1}F(x^k), k = 0, 1, \dots. \tag{17}$$

The Newton-SOR and Newton-Jacobi belong to this family of iterative processes. Ortega and Rheinboldt establish the following result concerning the iteration (17).

**Theorem 3.4.** [11] *Let the standard assumptions (C1)–(C3) hold, and suppose* $B : \Omega \to L(\boldsymbol{R}^n)$ *is continuous in* $x^*$*,* $B(x^*)$ *is nonsingular and* $\rho(B(x^*)^{-1}(F'(x^*) - B(x^*))) < 1$*. Then* $\{x^k\}$ *defined by (17) and (16) converges q-linearly to* $x^*$ *with q-order* $\rho(B(x^*)^{-1}(F'(x^*) - B(x^*)))$*.*

3.2. **Weighted maximum norms.** The assumptions (C1) and (C2) are also naturally for asynchronous block Newton method. The condition (C3) will be replaced by the following sufficient conditions which guarantee the existence of solutions of the subsystems and local convergence of the asynchronous method:

**(C3'):** All the matrices $\frac{\partial F_i(x^*)}{\partial x_i}$, $i = 1, \dots, L$ are nonsingular, and

$$\rho(|D(x^*)^{-1}(F'(x^*) - D(x^*))|) < 1.$$

**Remarks.** (a) (see also [15]) Conditions (C1) and (C2) are standard for Newton methods, and (C3') is natural for the convergence of asynchronous methods. Consider the linear case, where $F(x) = Ax - b$ and $F'(x) = A$. If there exists $A^{-1}$ then (C3') is necessary and sufficient for the convergence of the asynchronous block methods for the linear system $F(x) = 0$.

(b) Condition (C3') is also similar to the main requirement for the convergence given in the theorem 3.4, $\rho(B(x^*)^{-1}(F'(x^*) - B(x^*))) < 1$.

(c) (see also [15]) Condition (C3') holds when the Jacobian matrix $F'(x^*)$ is an $H$-matrix, since $F'(x^*) = D(x^*) - (D(x^*) - F'(x^*))$ is an $H$-splitting of $F'(x^*)$. Moreover, (C3') is equivalent to $F'(x^*)$ being an $H$-matrix if each block has only one component.

By the theory of nonnegative matrix, condition (C3') is quivalent to

**(C3"):** All matrices $\frac{\partial F_i(x^*)}{\partial x_i}$, $i = 1, \dots, L$ are nosingular, and there exists $\rho_0 < 1$ and a vector $w > 0$ such that

$$\|D(x^*)^{-1}(F'(x^*) - D(x^*))\|_w < \rho_0.$$

The weighted maximum norms used in (C3") are defined as follows. Let $w \in R^n$, $w > 0$ and $A \in \mathbf{R}^{n\times n}$ be partitioned conformally with $x$, then we define

$$\|x\|_w = \max\{\|x_i\|_{w_i}, 1 \le i \le L\} = \max\{\frac{|x_{i_j}|}{w_{i_j}}, 1 \le j \le n_i, 1 \le i \le L\}. \tag{18}$$

and the induced matrix norm $\|A\|_w = \max\{\frac{\|Ax\|_w}{\|x\|_w} : x \in \mathbf{R}^n \setminus \{0\}\}$.

We return to the subsystems (5). Starting from $x$ and applying one step of the Newton method for the $i$th block, we are interested to estimate

$$\begin{aligned} \|y_i - x_i^*\|_{w_i} &= \|x_i - x_i^* - \left(\tfrac{\partial F_i(x)}{\partial x_i}\right)^{-1} F_i(x)\|_{w_i} \\ &= \| \left(\tfrac{\partial F_i(x)}{\partial x_i}\right)^{-1} \left[F_i(x^*) - F_i(x) - \tfrac{\partial F_i(x)}{\partial x_i}(x_i^* - x_i)\right] \|_{w_i} \end{aligned} \tag{19}$$

As we can see from the right hand side of (19), the Lemma 3.2 cannot be applyed directly as for the sequential Newton method.

In order to obtain a similar lemma we can extend the weighted matrix norms for rectangular matrices as follows,

$$\|(A_{i1}, \dots, A_{iL})\|_{w_i} = \max\{\tfrac{\|(A_{i1},\dots,A_{iL})x\|_{w_i}}{\|x\|_w} : x \in \mathbf{R}^n \setminus \{0\}\},$$
$$\|A_{ij}\|_{w_i} = \max\{\tfrac{\|A_{ij}x_j\|_{w_i}}{\|x_j\|_{w_j}} : x_j \in \mathbf{R}^{n_j} \setminus \{0\}\}.$$

We immediately have,

$$\|A_{ii}(A_{i1}, \dots, A_{iL})\|_{w_i} \le \|A_{ii}\|_{w_i} \cdot \|(A_{i1}, \dots, A_{iL})\|_{w_i},$$
$$\||A|\|_w = \|A\|_w, \||(A_{i1}, \dots, A_{iL})|\|_{w_i} = \|(A_{i1}, \dots, A_{iL})\|_{w_i},$$
$$\|A\|_w = \max\{\|(A_{i1}, \dots, A_{iL})\|_{w_i}, 1 \le i \le L\},$$
$$\|A_{ij}\|_{w_i} \le \|(A_{i1}, \dots, A_{iL})\|_{w_i}, 1 \le i \le L.$$

These extensions were considered by Xu [15]. The following lemma will play a similar role for asynchronous block Newton methods as the lemma 3.2 for sequential Newton methods.

Because of the norm equivalence in finite dimensional spaces we can consider that the norm used in (C2) is the weighted norm $\|\cdot\|_w$, where $w$ is the vector defined in (C3").

**Lemma 3.5.** [15] *Under the conditions (C1), (C2) and (C3') we have*

$$\|\frac{\partial F_i(x)}{\partial x_i} - \frac{\partial F_i(x^*)}{\partial x_i}\|_{w_i} \le \gamma\|x - x^*\|_w, \quad \forall x \in S(x^*, \epsilon), \tag{20}$$

*and there exists* $\epsilon > 0$ *such that* $S(x^*, \epsilon) \subset \Omega$ *and*

$$\| \left(\frac{\partial F_i(x^*)}{\partial x_i}\right)^{-1} \left(\frac{\partial F_i(x)}{\partial x_1}, \dots, \frac{\partial F_i(x)}{\partial x_{i-1}}, 0, \frac{\partial F_i(x)}{\partial x_{i+1}}, \dots, \frac{\partial F_i(x)}{\partial x_L}\right) \|_{w_i} \le \rho_0, \tag{21}$$

$$\begin{aligned} (22) \qquad & \|x_i - x_i^* - \left(\tfrac{\partial F_i(x^*)}{\partial x_i}\right)^{-1} F_i(x)\|_{w_i} \\ & \leq \rho_0 \|x - x^*\|_w + \tfrac{\gamma}{2} \| \left(\tfrac{\partial F_i(x^*)}{\partial x_i}\right)^{-1} \|_{w_i} \cdot \|x - x^*\|_w^2, \end{aligned}$$

*for all* $i = 1, \dots, L, x \in S(x^*, \epsilon)$.

3.3. **Asynchronous Newton Method.** The next theorem represents the main result of the paper.

**Theorem 3.6.** *Let the assumptions (C1), (C2) and (C3'), and also the conditions (7)–(9) hold. Then there exists* $\delta > 0$ *such that if* $x^0 \in S(x^*, \delta)$ *then the sequence generated by asynchronous block Newton method converges to* $x^*$.

*Moreover, for* $l = 0, 1, \dots,$ *we have*

$$\|x^k - x^*\|_w \leq r^l \|x^0 - x^*\|_w, \forall k \geq k_l, \tag{23}$$

*where* $K > 0$, $r := \rho_0 + K\delta < 1$, *and the sequence* $\{k_l\}$*is defined by (12)–(13).*

*Proof.* We proceed in two steps: first we show that the sequence generated by asynchronous block Newton method is well defined and then it converges.

We consider $\beta > 0$ such that

$$\|F'(x^*)\|_w \leq \beta, \quad \|(\frac{\partial F_i(x^*)}{\partial x_i})^{-1}\|_{w_i} \leq \beta, \ i = 1, \dots, L.$$

*First part.* We choose $\delta > 0$ such that the matrices $\frac{\partial F_i(x)}{\partial x_i}$, $i = 1, \dots, L$ are nonsingular for all $x \in S(x^*, \delta)$. From (20),

$$\|\frac{\partial F_i(x)}{\partial x_i} - \frac{\partial F_i(x^*)}{\partial x_i}\|_{w_i} \leq \gamma \|x - x^*\|_w, \quad \forall x \in \Omega,$$

and by (C3') there exists $\frac{\partial F_i(x^*)}{\partial x_i}^{-1}$. Let $\delta > 0$ be such that the hypothesis of Banach lemma 3.3 hold, so there exists $\frac{\partial F_i(x)}{\partial x_i}^{-1}$, for all $x \in S(x^*, \delta)$. Now we choose $\delta$ small enough such that the assumptions of Lemma 3.5 also hold.

We show that if $x^0 \in S(x^*, \delta)$ then the sequence $\{x^k\}$ remains in $S(x^*, \delta)$. Suppose that for all $j$, $0 \leq j \leq k$, $\|x^j - x^*\|_w \leq \|x^0 - x^*\|_w$.

Let $u = (x_1^{s_1(k)}, \dots, x_L^{s_L(k)})$. For $i \in I^k$,

$$\begin{aligned} \|x_i^{k+1} - x_i^*\|_{w_i} &= \|x_i^{s_i(k)} - x_i^* - \tfrac{\partial F_i(u)}{\partial x_i}^{-1} F_i(u)\|_{w_i} \\ &= \|x_i^{s_i(k)} - x_i^* - \tfrac{\partial F_i(x^*)}{\partial x_i}^{-1} F_i(u)\|_{w_i} + \\ & \quad \|\tfrac{\partial F_i(x^*)}{\partial x_i}^{-1} - \tfrac{\partial F_i(u)}{\partial x_i}^{-1}\|_{w_i} \cdot \|F_i(u)\|_{w_i} \end{aligned} \tag{24}$$

Since

$$\begin{aligned} \|F_i(u)\|_{w_i} &= \|F_i(u) - F_i(x^*)\|_{w_i} \leq \|F(u) - F(x^*)\|_w \\ &\leq \|F(u) - F(x^*) - F'(x^*)(u - x^*)\|_w + \|F'(x^*)(u - x^*)\|_w \\ &\leq \tfrac{\gamma}{2} \|u - x^*\|_w^2 + \beta \|u - x^*\|_w, \end{aligned} \tag{25}$$

$$(26)\quad \begin{array}{rcl} \|\frac{\partial F_i(x^*)}{\partial x_i}^{-1} - \frac{\partial F_i(u)}{\partial x_i}^{-1}\|_{w_i} & = & \|\frac{\partial F_i(x^*)}{\partial x_i}^{-1}(\frac{\partial F_i(x^*)}{\partial x_i} - \frac{\partial F_i(u)}{\partial x_i})\frac{\partial F_i(u)}{\partial x_i}^{-1}\|_{w_i} \\ & \leq & \beta^2\gamma\|u - x^*\|_w \end{array}$$

we get

$$(27)\quad \begin{array}{rcl} \|x_i^{k+1} - x_i^*\|_{w_i} & \leq & \left[\rho_0 + \left(\beta\gamma(\frac{1}{2} + \beta^2) + \beta^2\gamma^2\|u - x^*\|_w\right)\|u - x^*\|_w\right]\|u - x^*\|_w \\ & = & (\rho_0 + K\|u - x^*\|_w)\|u - x^*\|_w, \end{array}$$

where $K = \beta\gamma(\frac{1}{2}+\beta^2)+\beta^2\gamma^2\delta$. Again, if necessary, we choose $\delta$ small enough such that $r := \rho_0 + K\delta < 1$, then (27) gives $\|x_i^{k+1} - x_i^*\|_{w_i} \leq \|x^0 - x^*\|_w$, for $i \in I^k$.

On the other hand, if $i \notin I^k$ then the $i$th component is not modified, $x_i^{k+1} = x_i^k$, and from the induction hypothesis it follows $\|x_i^{k+1}-x_i^*\|_{w_i} \leq \|x^0-x^*\|_w$, for $i \notin I^k$. The last two inequalities together with the norm definitions implies

$$\|x^{k+1} - x^*\|_w \leq \|x^0 - x^*\|_w,$$

that means the sequence $\{x^k\}$ generated by the asynchronous method is well defined and remains in $S(x^*, \delta)$ if $x^0 \in S(x^*, \delta)$.

*Second part.* Let $\{k_l\}$ be the sequence defined by (12) and (13). We show by mathematical induction that for all $l \in \mathbf{N}$

$$(28)\qquad \|x^k - x^*\|_w \leq r^l\|x^0 - x^*\|_w, \forall k \geq k_l.$$

hence $\{x^k\}$ is convergent, since $r = \rho_0 + K\delta < 1$.

Let $l = 0$. From the definition of $\{k_l\}$ it follows

$$\begin{array}{l} \forall k \geq k_0,\ \forall i \in \{1, \ldots, L\},\ \text{there exists } j : 0 \leq s(j) \leq j < k \\ \text{ such that } x_i^k = x_i^{j+1} \text{ and } i \in I^j. \end{array}$$

Using (27) we get $\|x_i^k - x_i^*\|_{w_i} = \|x_i^{j+1} - x_i^*\|_{w_i} \leq r\|x^0 - x^*\|_w$, for $i \in \{1, \ldots, L\}$, and by the definition of weighted norms, $\|x^k - x^*\|_w \leq r^1\|x^0 - x^*\|_w \leq, \forall k \geq k_0$.

Now, suppose for fixed $l \in \mathbf{N}$ we have

$$\|x^k - x^*\|_w \leq r^l\|x^0 - x^*\|_w, \forall k \geq k_l.$$

Using again the definition of $\{k_l\}$ we get

$$\begin{array}{l} \forall k \geq k_{l+1},\ \forall i \in \{1, \ldots, L\},\ \text{there exists } j : k_l \leq s(j) \leq j < k \\ \text{ such that } x_i^k = x_i^{j+1} \text{ and } i \in I^j. \end{array}$$

Let $u = (x_1^{s_1(j)}, \ldots, x_L^{s_L(j)})$. Using again (27), $\|x_i^k - x_i^*\|_{w_i} = \|x_i^{j+1} - x_i^*\|_{w_i} \leq r\|u - x^*\|_w = r\|x_p^{s_p(j)} - x^*\|_{w_p}$, where $p$ is an index for which the last equality holds (according to the definition of the norm $\|\cdot\|_w$). Since $s_p(j) \geq k_l$, it follows that $\|x_p^{s_p(j)} - x^*\|_{w_p} \leq r^l\|x^0 - x^*\|_w$ and the proof is complete. □

**Remark** Under the assumptions of Theorem 3.6, the asynchronous block Newton method converges with a rate of convergence (see [3]).

$$R = \liminf_{k\to\infty}[(-\log\|x^k - x^*\|)/k] \geq \rho_0.$$

If $F$ and $F'$ are computed innacurately then the asynchronous iteration (6) becomes, for $i \in I^k$,

$$x_i^{k+1} = x_i^{s_i(k)} - \left( \frac{\partial F_i(u)}{\partial x_i} + \Delta(u) \right)^{-1} (F_i(u) + \epsilon(u)), \tag{29}$$

where $u = (x_1^{s_1(k)}, \ldots, x_L^{s_L(k)})$. A similar local convergence theorem can be shown for an asynchronous block Newton perturbed method defined by (29) (see [9]). As in the sequential case [8], one can use Newton perturbed method to derive local convergence results for other Newton methods (e.g. chord method).

## References

[1] Jacques Bahi, Jean-Claude Miellou, and Karim Rhofir. Asynchronous multisplitting methods for nonlinear fixed point problems. *Numerical Algorithms*, 15:315–345, 1997.

[2] Zhong-Zhi Bai, Violeta Migallon, Jose Penades, and Daniel B. Szyld. Block and Asynchronous Two-Stage Methods for Midly Nonlinear Systems. *Num. Math.*, 82:1–21, 1999.

[3] Gérard M. Baudet. Asynchronous iterative methods for multiprocessors. *J. Association for Computing Machinery*, 25:226–244, 1978.

[4] Dimitri P. Bertsekas. *Distributed Asynchronous Computation of Fixed Points.* Prentice Hall, Englewood Cliffs, New Jersey, 1989.

[5] D. Chazan and W. Miranker. Chaotic relaxation. *Linear Algebra and its Applications*, 2:199–222, 1969.

[6] Andreas Frommer. Parallel Nonlinear Multisplitting Methods. *Numerische Mathematik*, 56:269–282, 1989.

[7] Andreas Frommer and Daniel B. Szyld. On Asynchronous Iterations. *J. Computational and Applied Mathematics*, 123:201–216, 2000.

[8] C.T. Kelley. *Iterative Methods for Linear and Nonlinear Equations.* SIAM Publications, 1995.

[9] Ioan Lazăr. On Asynchronous Two-Stage Newton Iterative Methods. *Rev. Anal. Numér. Théor. Approx.* submitted.

[10] D.P. O'Leary and R.E. White. Multi-Splitting of Matrices and Parallel Solution of Linear Systems. *SIAM J. Alg. Disc. Meth.*, 6:630–640, 1985.

[11] J.M. Ortega and W.G. Rheinboldt. *Iterative Solutions of Nonlinear Equations in Several variables.* Academic Press, New York, 1970.

[12] Daniel B. Szyld and Jian-Jun Xu. Convergence of Some Asynchronous Nonlinear Multisplitting Methods. *Numerical Algorithms*, 25:347–361, 2000.

[13] Mouhamed Nabih El Tarazi. Some Convergence Results for Asynchronous Algorithms. *Num. Math.*, 39:325–340, 1982.

[14] R.E. White. Parallel Algorithms for Nonlinear Problems. *SIAM J. Alg. Disc. Meth.*, 7:137–149, 1986.

[15] Jian-Jun Xu. Convergence of Partially Asynchronous Block Quasi-Newton Methods for Nonlinear Systems of Equations. *J. Computational and Applied Mathematics*, 103:307–321, 1999.

Babeş-Bolyai University, Faculty of Mathematics and Informatics, Cluj-Napoca, str. M. Kogălniceanu 1

*E-mail address*: ilazar@cs.ubbcluj.ro