On asynchronous Newton-Schulz type methods

Matthias Pott *

Abstract

The disadvantage of most known asynchronous iteration methods for linear and nonlinear equations seems to be, that the system to be solved, has to be related to an M-system. In this paper we will introduce an asynchronous method for nonlinear equations $F(x) = 0$, which is locally convergent, independent of the structures of the Jacobian of F . Admittedly, it seems easy to develop asynchronous variants of Newton's method, which converge even superlinearly, but the trouble of Newton's method cannot be avoided by asynchronous iterations, it is to solve again and again linear systems, which is on the whole a synchronized task. Sometimes it is also a hard task to compute the partial derivatives of F . Using a variation of the iterative method of Schulz (s. [9]) for inverting matrices, our method allows to solve nonlinear equations totally asynchronously, e.g. the iterations, the partial derivatives and the inverses of the Jacobians can be computed in parallel and without synchronization. Moreover, it converges locally superlinearly.

Key words: Asynchronous iterations, nonlinear equations, Newton's method, Schulz' method.

AMS(MOS) subject classification. 65J15.

1 Introduction

In [1] Baudet introduced asynchronous iteration methods for nonlinear fixed point problems and the more general *asynchronous iteration methods with memory*. In the latter case, there is a given operator

$$G : D_0^m \subset \mathbb{R}^{nm} \to D_0 \subset \mathbb{R}^n \tag{1}$$

for which a *fixed point* $\xi \in D_0$, e.g.

$$G(\xi, ..., \xi) = \xi, \tag{2}$$

is searched.

*Research supported by Sonderforschungsbereich 343: „Diskrete Strukturen in der Mathematik“, Universität Bielefeld. Postfach 100131, 33501 Bielefeld. Electronic mail: pott@mathematik.uni-bielefeld.de

Definition 1.1 Let G be a real operator as in (1), $\mathcal{S}$ be $n \cdot m$ sequences of natural numbers, denoted by

$$\begin{array}{lll} & s_i^l : \mathbb{N}_0 \to \mathbb{N}_0 \quad i = 1, ..., n \ , \ l = 1, ..., m \ , & \\ \text{with} & s_i^l(j) \leq j \quad \forall \ i, l, j \ , & (3) \end{array}$$

$\mathcal{J}$ be a sequence of subsets of $\{1, ..., n\}$, given by

$$J_j \subset \{1, ..., n\} \quad j = 0, 1, ... \tag{4}$$

and $x(0) \in D_0$ be an initial vector. Then we call the sequence $x(j)_{\, j=0,1,...}$, given by

$$x_i(j+1) := \begin{cases} x_i(j) & , \text{if } \ i \notin J_j \\ G_i \left(\begin{pmatrix} x_1(s_1^1(j)) \\ \vdots \\ x_n(s_n^1(j)) \end{pmatrix}, ..., \begin{pmatrix} x_1(s_1^m(j)) \\ \vdots \\ x_n(s_n^m(j)) \end{pmatrix} \right) & , \text{if } \ i \in J_j \end{cases} \tag{5}$$

$$j = 0, 1, ...$$

an *asynchronous iteration with memory* and denote it by $(G, \mathcal{S}, \mathcal{J}, x(0))$.

In order to prove Theorems of convergence, we have to analyze the operator G and the sequences $\mathcal{S}$ and $\mathcal{J}$ on their properties. First we consider $\mathcal{S}$ and $\mathcal{J}$. The following terms are carried over from [3]:

Definition 1.2 Let $(G, \mathcal{S}, \mathcal{J}, x(0))$ be an asynchronous iteration with memory. Then

(i) $\mathcal{S}$ is called *admissible*, if for all $j \in \mathbb{N}_0$

$$s_j := \min\{ \, s \in \mathbb{N}_0 \, | \, s_i^l(j_0 + s) \geq j \ \ \forall \ j_0 \geq j \ , \ i = 1, ..., n \ , \ l = 1, ..., m \, \} \ . \tag{6}$$

exists.

(ii) $\mathcal{S}$ is said to be *regulated*, if

$$s := \max_{j,l,i} \ j - s_i^l(j) \tag{7}$$

exists.

(iii) $\mathcal{J}$ is called *admissible*, if for all $j \in \mathbb{N}_0$ there is a $t_j \in \mathbb{N}$, such that

$$J_j \cup J_{j+1} \cup ... \cup J_{j+t_j} = \{1, ..., n\} \tag{8}$$

holds.

(iv) $\mathcal{J}$ is *regulated*, if there is a number $t \in \mathbb{N}_0$, such that for all $j \in \mathbb{N}_0$

$$J_j \cup J_{j+1} \cup ... \cup J_{j+t} = \{1, ..., n\} \ . \tag{9}$$

The following criterion on the operator G was introduced in [1]:

Definition 1.3 Let G be an operator as in (1). Then G is said to be *pseudocontracting* on D_0, if there is a fixed point $\xi \in D_0$ of G and

(i) D_0 is closed,

(ii) for any diagonal matrix $P \in L(\mathbb{R}^n)$ with $P_{ii} \in \{0,1\}\ \forall\ i \in \{1,...,n\}$ it is

$$P\,x + (\,I\, -\, P\,)\,y \in D_0 \quad \forall\ x, y \in D_0,$$

(iii) $$|\,G(x^1,...,x^m) - \xi\,| \leq A\ \ \operatorname*{compmax}_j |\,x^j - \xi\,| \qquad \forall\ x^1,...,x^m \in D_0$$

for a matrix $A \in L(\mathbb{R}^n)$, such that there is a positive vector $v \in \mathbb{R}^n$, such that $|A|\,v \leq \omega\,v$ for an $\omega < 1$.

Here $|\cdot|$ denotes that matrix, whose entries are taken absolutely of the original one, and for a set of vectors $v^j\ j = 1,...,m$ from $\mathbb{R}^n$ we define

$$\operatorname*{compmax}_j\ v^j := \begin{pmatrix} \max\limits_j\ v_1^j \\ \vdots \\ \max\limits_j\ v_n^j \end{pmatrix} .$$

The conditions (i) and (ii), which we added to the criterion in [1], guarantee already that the iterations (5) are well-defined: They cannot leave D_0, since it is closed, and a combination of components of different vectors stays still within it.

For further investigations we need at last measures for the speed of convergence:

Definition 1.4 Let $x(j)\ {}_{j=0,1,...}$ be some sequence from $\mathbb{R}^n$, which converges to $\xi \in \mathbb{R}^n$ and $\|\cdot\|$ some vector norm, then

$$\mathcal{R}\ :=\ \limsup_{j\to\infty}\ \sqrt[j]{\|x(j) - \xi\|} \tag{10}$$

is called the *rate of convergence* of $x(j)\ {}_{j=0,1,...}$ and

$$\mathcal{O}\ :=\ \liminf_{j\to\infty}\ \sqrt[j]{-\log(\|x(j) - \xi\|)} \tag{11}$$

the *order of convergence* of $x(j)\ {}_{j=0,1,...}$. If $\mathcal{R} = 0$, then $x(j)\ {}_{j=0,1,...}$ is said to be *superlinearly convergent* .

Remark, that these values exist in any case, if we accept $\mathcal{O} = \infty$, and that they are independent of the norm and the logarithm used. Before coming to our main results, we should mention the result of Baudet in [1]. Indeed, Baudet did not use exactly the same assumptions as in the definitions 1.2 and 1.3, but the following Theorem can be derived easily:

Theorem: 1.5 *Let G be a pseudocontracting operator on D_0 and $(G, \mathcal{S}, \mathcal{J}, x(0))$ be an asynchronous iteration with memory. Then $(G, \mathcal{S}, \mathcal{J}, x(0))$ converges to ξ, if $\mathcal{S}$ and $\mathcal{J}$ are admissible. If they are regulated, it converges at least linear with*

$$\mathcal{R} \leq \sqrt[s+t+1]{\omega}\,. \tag{12}$$

2 Newton-Schulz' type algorithms

Chazan and Miranker gave in [4] an example, which shows, that in the linear and regular case, the criterion, that G has to be pseudocontacting is also necessary for general convergence, if the criteria on $\mathcal{S}$ and $\mathcal{J}$ will not be sharpened. Assume for example, a regular linear system $Ax = b$ shall be solved by an asynchronous version of the Jacobi or Gauss-Seidel method, e.g. we split A into $A = B - C$, where B is the diagonal or lower triangle part of A, and observe the operator

$$G(x) = B^{-1}C\,x + B^{-1}b\,.$$

This operator is pseudocontracting on $\mathbb{R}^n$, if and only if $|\,B^{-1}\,C\,|$ is of spectral radius less than unit (compare [4]). But the only class of matrices, for which was shown, that this criterion holds in general are M-matrices and classes of matrices, which are related to them, like diagonal dominant matrices, H-matrices, etc. (see for example [2] for details). The class of M-matrices is admittedly important, but at least not the whole world of linear systems. For example, the likewise important class of positive definite matrices does not have this property, in general. Even such systems can be solved parallel by a synchronized multisplitting method, introduced in [6]. If a nonlinear system

$$F(x) = 0 \tag{13}$$

is smooth, we have to expect, that the same problems arise, e.g. there is only a hope, that corresponding asynchronous iterations are generally locally convergent, if the Jacobian of F at F's unique local root, is related to an M-matrix. A way out is perhaps the following: If in the splitting $A = B - C$ the *largest part* of A was given up to B, then $|\,C\,|$ is small and/or sparse, and we can expect, that the spectral radius of $|\,B^{-1}\,C\,|$ is small. Of course, in general this makes no sense in the linear case, because the invertation or LR-decompositioning of a matrix, which is not a lower or upper triangle matrix, is in general a synchronized task. Therefore, instead of inverting B by a synchronized method in order to get an iterative method, it generally seems to be more advantageous to invert A itself. In the nonlinear case, the problem lies different. If we have an initial vector $x(0)$ for solving (13), and there is no essentially different possibility, then to compute the next Newton step, we get

$$x(1) := G(x(0)) := x(0) - F'(x(0))^{-1} F(x(0)) . \tag{14}$$

Assume now, that we have typical conditions for an unique local root of F :

Definition 2.1 Let $F : D \subset \mathbb{R}^n \to \mathbb{R}^n$ be a function with $F(\xi) = 0$ for some $\xi \in D$. If there is an open neighborhood $D_1 \subset D$ of ξ, such that

$$\begin{array}{lll} \text{(i)} & F' & \text{exists and is continuous on } D_1 , \\ \text{(ii)} & \| F'(x) - F'(\xi) \| \leq L \| x - \xi \| & \text{for all } x \in D_1 \text{ and some } L > 0 , \\ \text{(iii)} & \det(F'(\xi)) \neq 0 , & \end{array} \tag{15}$$

then ξ is called a *regular root* of F.

Let G be the *iteration function* of an asynchronous iteration with $m = 1$. If G is differentiable in an open neighborhood of a fixed point ξ of G, then $\rho(| G'(\xi) |) < 1$ is that criterion, which guarantees that an asynchronous iteration method converges locally (s. [10]). $\rho(|G'(\xi)|)$ is even the *local* ω of definition 1.3, since $|G'(\xi)|$ is the limit of the matrices A_ε, which satisfy in $\varepsilon-$neighborhoods of ξ (iii) of definition 1.3. Now return to the case of equation (14). If we have already computed the first step of Newton's method and if $F'(x(0))$'s LR-decomposition or inverse is saved, one possibility to create an asynchronous iteration method is to avoid further evaluations and invertations of Jacobians and just to go on working with $F'(x(0))^{-1}$. Using the results of [10] and [7],§10.2.1., we have in the case of regular roots, that the iteration function

$$G(x) := x - F'(x(0))^{-1} F(x) ,$$

(it is exactly that function, which leads in a sequential algorithm to the so called *simplified Newton method*, see [7],§7.1.) is differentiable in an open neighborhood of ξ and that

$$G'(\xi) = I - F'(x(0))^{-1} F'(\xi) .$$

Thus, if $x(0)$ is close enough to a regular ξ, this asynchronous method is well-defined and convergent. But also the disadvantages of such a method are obvious: The area of attraction seems to be larger and we lose that preferred superlinear convergence of Newton's method. On the other hand $F'(x(0))^{-1}$ is also close to $F'(x(1))^{-1}$, and another way out could be to compute the entries of $F'(x)^{-1}$ in parallel and asynchronously by interpreting $F'(\xi)^{-1}$ itself as an unknown. By that, we will get in addition the inverse of F's Jacobian at ξ, which is perhaps wanted, and the area of convergence should be larger. At first, we will formulate our algorithm in a very general way.

Algorithm: Asynchronous Newton-Schulz'-type iteration methods

$$\text{Compute} \qquad (G = (G_1, G_2), \mathcal{S}, \mathcal{J}, (x(0), \mathrm{X}(0)) = X(0)),$$

$$\begin{aligned} \text{where} \quad & G_1 : (D_0 \times D_{F'})^m \subset (\mathbb{R}^n \times \mathbb{R}^{n^2})^m \to D_0 \subset \mathbb{R}^n, \\ & G_2 : (D_0 \times D_{F'})^m \subset (\mathbb{R}^n \times \mathbb{R}^{n^2})^m \to D_{F'} \subset \mathbb{R}^{n^2} \end{aligned}$$

and G_1 was given by

$$G_1(x^1, \mathrm{X}^1, ..., x^m, \mathrm{X}^m) := x^1 - \mathrm{X}^1 F(x^1). \tag{16}$$

In order to analyze these methods, we will identify, without taking care, the Banach spaces $(L(\mathbb{R}^n), \|\cdot\|)$ and $(\mathbb{R}^{n^2}, \|\cdot\|)$. But throughout the rest of the paper $\|\cdot\|$ will be a matrix norm, while $\|\cdot\|_\infty$ is the maximum norm on $\mathbb{R}^{n^2}$, respectively on $\mathbb{R}^n$. e is the vector in $\mathbb{R}^n$, whose entries are all equal to one, and $\mathcal{e}$ the corresponding one in $\mathbb{R}^{n^2}$.

3 Main results

Theorem: 3.1 *Let $F : D \subset \mathbb{R}^n \to \mathbb{R}^n$ be a function, and $\xi \in D$ be a regular root of it. $(G, \mathcal{S}, \mathcal{J}, X(0))$ be an asynchronous Newton-Schulz' type method. Then:*

(i) *If $G : (D_0 \times D_{F'})^m \subset D^m \times L(\mathbb{R}^n)^m \to D_0 \times D_{F'}$ is a pseudocontracting operator on $D_0 \times D_{F'}$, then $(G, \mathcal{S}, \mathcal{J}, X(0))$ converges to $(\xi, F'(\xi)^{-1})$, if $\mathcal{S}$ and $\mathcal{J}$ are admissible and $X(0) \in D_0 \times D_{F'}$.*

(ii) *Let $(x(j), \mathrm{X}(j))_{j=0,1,...}$ be the sequence $X(j)_{j=0,1,...}$ partitioned into sequences, which are from D_0, respectively from $D_{F'}$. Likewise, let $\mathcal{S}^1$ and $\mathcal{J}^1$ be that parts of $\mathcal{S}$ respectively $\mathcal{J}$, which belong to $x(j)_{j=0,1,...}$. If in addition to the assumptions of (i), $\mathcal{S}^1$ and $\mathcal{J}^1$ are regulated in the sense of definition 1.2 (ii) and (iv), then $x(j)_{j=0,1,...}$ converges superlinearly to ξ.*

(iii) *Let there be an open neighborhood $D_1 \times D_{F_1'}$ of $(\xi, F'(\xi)^{-1})$, such that G is well-defined and such that there are constants $K_0, K_1 > 0$, so that*

$$\begin{aligned} & \| G_2(x^0, \mathrm{X}^0, ..., x^m, \mathrm{X}^m) - F'(\xi)^{-1} \|_\infty \\ \leq \; & K_0 \max_{1 \leq l \leq m} \| x^l - \xi \|_\infty + K_1 \max_{1 \leq l \leq m} \| \mathrm{X}^l - F'(\xi) \|_\infty^2, \\ & \qquad \forall \; (x^1, \mathrm{X}^1, ..., x^m, \mathrm{X}^m) \in D_1^m \times D_{F_1'}^m , \end{aligned} \tag{17}$$

then G *is pseudocontracting on a closed set* $D_0 \times D_{F'}$, *which contains* $(\xi, F'(\xi)^{-1})$ *in its interior.*

(iv) *If, under the assumptions of (iii),* $\mathcal{S}$ *and* $\mathcal{T}$ *are regulated, then* $(G, \mathcal{S}, \mathcal{J}, X(0))$ *converges with*

$$\mathcal{O} \geq \sqrt[2\,(s+t+1)]{2} \tag{18}$$

to $(\xi, F'(\xi)^{-1})$, *if* $x(0) \in D_0$ *and* $\mathrm{X}(0) \in D_{F'}$.

Proof : (i) follows directly from Theorem 1.5.

(ii): For abbreviation, we are using x_j^1 and X_j^1 for the values of x^1 and X^1, which are used in the $j+1-$st iteration step to update the components $i \in J_j^1$ of x. Because of (i) and since $\mathcal{J}$ and $\mathcal{S}$ are admissible, $\| \mathrm{X}_j^1 - F'(\xi)^{-1} \|_\infty \; {}_{j=0,1,...}$ and $\| x_j^1 - \xi \|_\infty \; {}_{j=0,1,...}$ converge to zero. Therefore, we can assume, that $D_0 \times D_{F'}$ is bounded and convex and that definition 2.1 holds on D_0 for $\| \cdot \|_\infty$. Therefore, by the mean value Theorem of [7],§3.2.12, we have

$$\| F(\xi) - F(x) + F'(\xi)(\xi - x) \|_\infty \leq \frac{L}{2} \| x - \xi \|_\infty^2 \quad \forall \; x \in D_0. \tag{19}$$

In addition, we can assume that for some $C_0 \in \mathbb{R}$

$$\| F(x) - F(\xi) \|_\infty \leq C_0 \| x - \xi \|_\infty \qquad \forall \, x \in D_0 \,. \tag{20}$$

Now, let sequences $\lambda_j \; {}_{j=0,1,...}$ and $\Lambda_j \; {}_{j=0,1,...}$ be defined by

$$\begin{aligned} \lambda_j &:= \sup_{j_0 \geq j} \| x_{j_0}^1 - \xi \|_\infty \\ \text{and} \qquad \Lambda_j &:= \sup_{j_0 \geq j} \| \mathrm{X}_{j_0}^1 - F'(\xi)^{-1} \|_\infty \end{aligned} \tag{21}$$

and for abbreviation denote $\| F'(\xi)^{-1} \|_\infty \frac{L}{2}$ by C_1, then, if $i \in J_j^1$:

$$\begin{aligned} &| \, x_i(j+1) - \xi_i \, | = | \, x_j^1 - \mathrm{X}_j^1 \, F(x_j^1) - \xi \, |_i \\ &\leq \;\; | \, \mathrm{X}_j^1 - F'(\xi)^{-1} \, |_i \, | \, F(x_j^1) - F(\xi) \, | \; + \; | \, x_j^1 - F'(\xi)^{-1} \, F(x_j^1) - \xi \, |_i \\ &\leq \;\; \Lambda_j \, e^t \, C_0 \, \| \, x_j^1 - \xi \, \|_\infty \, e \; + \; | \, F^{-1}(\xi) \, |_i \; | \, F(x_j^1) - F(\xi) + F'(\xi) \, (\xi - x) \, | \\ &\leq \;\; \Lambda_j \, n \, C_0 \, \lambda_j \; + \; C_1 \, n \, \lambda_j^2 \, . \end{aligned} \tag{22}$$

Now let

$$\omega_j := \Lambda_j \, n \, C_0 \; + \; C_1 \, n \, \lambda_j \quad \forall \; j \geq 0 \; , \tag{23}$$

and assume w.l.o.g., that $\omega_0 < 1$. Since $\mathcal{J}^1$ is regulated, there is some t^1, such that

$$J_j^1 \cup \, ... \, \cup J_{j+t^1}^1 \; = \; \{1, ..., n\} \qquad \forall \, j \geq 0 \; .$$

But this means

$$\| x(j_0+t^1+1)-\xi \|_\infty \;\le\; \omega_j\,\lambda_j \quad \forall\, j_0 \ge j \,, \forall\, j \ge 0\,, \tag{24}$$

and since $\mathcal{S}^1$ is regulated, there is a s^1, such that $s^1_l(j_0+t^1+1+s^1) \ge j+t^1+1 \;\forall\, j_0 \ge j\,,\, j \in \mathbb{N}_0$. Hence, replacing s^1+t^1 by p, we have

$$\lambda_{j+p+1} \;\le\; \omega_j\,\lambda_j \qquad \forall\, j \in \mathbb{N}_0\,. \tag{25}$$

Since for each fixed $j_0 \in \mathbb{N}_0$

$$\begin{aligned} \mathcal{R} &= \limsup_{j\to\infty} \sqrt[j]{\|x(j)-\xi\|_\infty} \\ &\le \limsup_{j\to\infty} \sqrt[j]{\omega_{j_0}^{\lfloor \frac{j-j_0}{p+1} \rfloor}\,\lambda_{j_0}} \\ &= \sqrt[p+1]{\omega_{j_0}}\;, \end{aligned} \tag{26}$$

(ii) is proved now, since $\omega_j\;{}_{j=0,1,\ldots}$ is a null sequence.

(iii): Since (17), there are constants $C>0$ and $K>0$, such that G is well-defined on $(D_2 \times D_{F'_2})^m$, where

$$D_2 \times D_{F'_2} := \{(x,\mathrm{X}) \in \mathbb{R}^n \times \mathbb{R}^{n^2} \mid \|x-\xi\|_\infty \le C\,,\; \|\mathrm{X}-F'(\xi)^{-1}\|_\infty \le K\},$$

and such that (compare (22)):

$$\begin{pmatrix} \mathbb{R}^n \\ \mathbb{R}^{n^2} \end{pmatrix} \ni \quad \left| \begin{matrix} G_1(x^1,\mathrm{X}^1,\ldots,x^m,\mathrm{X}^m)-\xi \\ G_2(x^1,\mathrm{X}^1,\ldots,x^m,\mathrm{X}^m)-F'(\xi)^{-1} \end{matrix} \right| \;\le \tag{27}$$

$$\begin{pmatrix} (C_0\,n\,\|\mathrm{X}^1-F'(\xi)^{-1}\|_\infty + C_1\,\|x^1-\xi\|_\infty\;)\;e\;e^t & 0 \\ K_0\;e\;e^t & K_1\;\max_l \|\mathrm{X}^l-F'(\xi)^{-1}\|_\infty\;e\;e^t \end{pmatrix}$$

$$\cdot \begin{pmatrix} \operatorname{compmax}_{1\le l\le m}\;|x^l-\xi| \\ \operatorname{compmax}_{1\le l\le m}\;|\,\mathrm{X}^l-F'(\xi)^{-1}\,| \end{pmatrix}$$

$$\forall\,(x^1,\mathrm{X}^1,\ldots,x^m,\mathrm{X}^m) \in D_2 \times D_{F'_2}\;.$$

If $\alpha < \frac{1}{K_0\,n}$ is a positive number, there are positive numbers $d_{F'}$ and d_x, such that

$$d_{F'} \;<\; \min\left\{ \frac{1-K_0\,n\,\alpha}{K_1\,n^2}\,,\, \frac{1}{C_0\,n^2}\,,\, K \right\}$$

and

$$d_x \;<\; \min\left\{ \frac{1-C_0\,d_{F'}\,n^2}{C_1\;n}\,,\, d_{F'}\,\frac{1-K_1\,n^2\,d_{F'}}{K_0\,n}\,,\, C \right\}. \tag{28}$$

Now, G is well-defined on $(D_0 \times D_{F'})^m$, given by

$$\begin{aligned} D_0 &:= \{x \in \mathbb{R}^n \mid \|x-\xi\|_\infty \le d_x\} \\ D_{F'} &:= \{\mathtt{X} \in L(\mathbb{R}^n) \mid \|\mathtt{X} - F'(\xi)^{-1}\|_\infty \le d_{F'}\}. \end{aligned} \tag{29}$$

Moreover, G is pseudocontracting on the quader $D_0 \times D_{F'}$, since by (28) there is a $\omega < 1$, such that

$$\begin{aligned} &\begin{pmatrix} (C_0\, n\|\mathtt{X}^1 - F'(\xi)^{-1}\|_\infty + C_1\|x^1-\xi\|_\infty)\, e\, e^t & 0 \\ K_0\; e\, e^t & K_1 \max_l \|\mathtt{X}^l - F'(\xi)^{-1}\|_\infty\, e e^t \end{pmatrix} \begin{pmatrix} \alpha\, e \\ e \end{pmatrix} \\ &\quad \le \begin{pmatrix} (C_0\; n\; d_{F'} + C_1\; d_x\;)\; e\; e^t & 0 \\ K_0\; e\, e^t & K_1\; d_{F'}\; e\; e^t \end{pmatrix} \begin{pmatrix} \alpha\, e \\ e \end{pmatrix} \\ &\quad < \omega \begin{pmatrix} \alpha\, e \\ e \end{pmatrix} \qquad \forall\, (x^1, \mathtt{X}^1, ..., x^m, \mathtt{X}^m) \in (D_0 \times D_{F'})^m . \end{aligned}$$

Because of

$$C_0\; d_{F'}\; n^2 \;+\; C_1\; d_x\; n \;<\; 1$$

and

$$K_0\; n\; d_x \;+\; K_1\; d_{F'}\; n^2\; d_{F'} \;<\; d_{F'}, \tag{30}$$

we even have $\left\| G(X^1,...,X^m) - \binom{\xi}{F'(\xi)^{-1}} \right\|_\infty \le \max_l \left\| X^l - \binom{\xi}{F'(\xi)^{-1}} \right\|_\infty$ $\forall\, X^l \in D_0 \times D_{F'}$, $l = 1,...,m$, and therefore (iii).

(iv): Let again $x_j^l\ \forall\, j,l$ and $\mathtt{X}_j^l\ \forall\, j,l$ be that approximations for ξ, resp. $F'(\xi)^{-1}$, used in the $j+1$-th step of iteration to update the components $J_j^1 \subset \{1,...,n\}$ of x and the entries $J_j^2 \subset \{(1,1),...,(n,n)\}$ of $\mathtt{X}$. This means by (27):

$$\begin{aligned} |x_{i_0}(j+1) - \xi_{i_0}| &\le (C_0\, n^2\, \|\mathtt{X}_j^1 - F'(\xi)^{-1}\|_\infty \;+\; C_1\, n\, \|x_j^1 - \xi\|_\infty)\, \|x_j^1 - \xi\|_\infty \\ |\mathtt{X}(j+1)_{i_1\, i_2} - F'(\xi)^{-1}_{i_1\, i_2}| &\le K_0\, n \max_{1\le l\le m} \|x_j^l - \xi\|_\infty \;+\; K_1\, n^2 \max_{1\le l\le m} \|\mathtt{X}_j^l - F'(\xi)^{-1}\|_\infty^2 \\ &\qquad \forall\, i_0 \in J_j^1\,,\ (i_1,i_2) \in J_j^2 \quad \forall\, j \ge 0. \end{aligned} \tag{31}$$

After any $t_j + s_{j+s_j} + 1$ more steps of iteration, we therefore have

$$\begin{aligned} &\|x(j+t_j+s_{j+t_j}+1) - \xi\|_\infty \\ &\quad \le (C_0\, n^2 \max_{q\ge j} \|\mathtt{X}_q^1 - F'(\xi)^{-1}\|_\infty \;+\; C_1\, n \max_{q\ge j}\|x_q^1 - \xi\|_\infty) \max_{q\ge j} \|x_q^1 - \xi\|_\infty \end{aligned}$$

and likewise

$$\| \mathrm{X}(j + t_j + s_{j+t_j} + 1) - F'(\xi)^{-1} \|_\infty \tag{32}$$
$$\leq K_0 \, n \max_{\substack{q \geq j \\ 1 \leq l \leq m}} \|x_q^l - \xi\|_\infty + K_1 \, n^2 \max_{\substack{q \geq j \\ 1 \leq l \leq m}} \|\mathrm{X}_q^l - F'(\xi)^{-1}\|_\infty^2 \qquad \forall \, j \geq 0.$$

Again setting

$$\begin{aligned} \lambda_j \; &:= \; \sup_{j_0 \geq j} \|x(j_0) - \xi\|_\infty \,, \\ \text{and} \quad \Lambda_j \; &:= \; \sup_{j_0 \geq j} \|\mathrm{X}(j_0) - F'(\xi)^{-1}\|_\infty \qquad \forall \, j \geq 0, \end{aligned} \tag{33}$$

we get for the sequence p_j $_{j=0,1,\dots}$, defined recursively by

$$\begin{aligned} p_0 \; &:= \; 0 \\ p_{j+1} \; &:= \; p_j + t_{p_j} + s_{p_j + t_{p_j}} + 1 \qquad \forall \, j > 0 \,, \end{aligned} \tag{34}$$

that

$$\lambda_{p_{j+1}} \leq C_0 \, n^2 \, \Lambda_{p_j} \, \lambda_{p_j} + C_1 \, n \, \lambda_{p_j}^2 \qquad \forall \, j \geq 0 \,,$$

respectively

$$\Lambda_{p_{j+1}} \leq K_0 \, n \, \lambda_{p_j} + K_1 \, n^2 \, \Lambda_{p_j}^2 \qquad \forall \, j \geq 0 \,. \tag{35}$$

To verify our assertion, it suffices to show by induction, that

$$\lambda_{p_j} \leq \omega_0^{\sqrt{2}^{\,j}} \, d_x \qquad \forall \, j \geq 0$$

and

$$\Lambda_{p_j} \leq \omega_0^{\sqrt{2}^{\,j}} \, d_{F'} \qquad \forall \, j \geq 0 \tag{36}$$

for a $\omega_0 < 1$ holds, since this would mean

$$\begin{aligned} \mathcal{O} \quad &= \quad \liminf_{j \to \infty} \sqrt[j]{- \log \left(\| \, X(j) - \begin{pmatrix} \xi \\ F'(\xi)^{-1} \end{pmatrix} \|_\infty \right)} \\ &\geq \quad \liminf_{j \to \infty} \sqrt[p_j]{- \log(\omega_0^{\sqrt{2}^{j}})} \;\; \geq \;\; \sqrt[2(s+t+1)]{2} \,, \end{aligned} \tag{37}$$

because of $p_j \leq j \, (s + t + 1)$.

Such a ω_0 is given by

$$\omega_0 := \sqrt[\sqrt{2}]{\max \left\{ C_0 \, n^2 \, d_{F'} + C_1 \, d_x \, n \,, \; \frac{K_0 \, n \, d_x + K_1 \, n^2 \, d_{F'}^2}{d_{F'}} \right\}} \;. \tag{38}$$

Because of (30) we have, that at least $\omega_0 < 1$. If $j = 0$, the inequalities (36) are valid, because of $(x(0), \mathrm{X}(0)) \in D_0 \times D_{F'}$. If $j = 1$, we have by (35)

$$\lambda_{p_1} \leq (\, C_0 \, n^2 \, d_{F'} + C_1 \, n \, d_x \,) \, d_x \; \leq \; \omega_0^{\sqrt{2}} \, d_x$$

and

$$\Lambda_{p_1} \le K_0\, n\, d_x \,+\, K_1\, n^2\, d_{F'}^2 \,\le \omega_0^{\sqrt{2}}\, d_{F'}\,.$$

Now, let (36) be shown for all $1 \le q \le j$, then

$$\begin{aligned} \lambda_{p_{j+1}} &\le C_0\, n^2\, \Lambda_{p_j}\, \lambda_{p_j} \,+\, C_1\, n\, \lambda_{p_j}^2 \\ &\le (C_0\, n^2\, d_{F'} \,+\, C_1\, n\, d_x)\, \omega_0^{\sqrt{2}^{2j}}\, d_x \\ &< \omega_0^{\sqrt{2}^{j+1}}\, d_x \end{aligned}$$

and

$$\begin{aligned} \Lambda_{p_{j+1}} &\le K_0\, n\, \lambda_{p_j} \,+\, K_1\, n^2\, \Lambda_{p_j}^2 \\ &\le \frac{K_0\, n\, d_x \,+\, K_1\, n^2\, d_{F'}^2}{d_{F'}}\, \omega_0^{\sqrt{2}^j}\, d_{F'} \\ &\le \omega_0^{\sqrt{2}^{j+1}}\, d_{F'}\,. \end{aligned}$$

This completes the proof. □

We should mention, that we used very rough estimates, in order to prove the assertions. For a given problem, the constants should be much more moderate. In difference to the proof of (ii), we used in that one of (iv) at first the admissibility of the sequences $\mathcal{S}$ and $\mathcal{J}$. Therefore, if we compute, instead of (37), the rate of convergence of $\max\{\lambda_j, \Lambda_j\}_{j=0,1,\ldots}$, we can show superlinear convergence, even if all sequences are only admissible. For example, if we have $p_j = \lfloor \ln(j) \rfloor \ \forall\, j \in \mathbb{N}$, which is perhaps our intention, because some components or entries shall be computed more and more often, than other ones, we would get

$$\mathcal{R} \le \lim_{j\to\infty} \omega_0^{\sqrt{2}^{\frac{j}{\ln(j)}}} = 0\,.$$

The question now is how to get a second part of G, which satisfies (17). In [5] Helfrich combines Newton's method with that one of Schulz, which itself is sometimes called the *Newton method to invert matrices*. He showed semilocal and quadratical convergence for it, if the root of F is regular:

$$\begin{aligned} \mathrm{X}(j+1) &:= \mathrm{X}(j)\,(2\,I - F'(x(j))\,\mathrm{X}(j)) \\ x(j+1) &:= x(j) - \mathrm{X}(j+1)\; F(x(j)) \qquad \forall\, j = 0, 1, \ldots \end{aligned} \tag{39}$$

Now, exchanging $\mathrm{X}(j+1)$ by $\mathrm{X}(j)$ in the second equation, we already have a synchronized method, which converges by Theorem 3.1 with $\mathcal{O} \ge \sqrt{2}$ to $(\xi, F'(\xi)^{-1})$, since we would have $s = t = 0$ and the following Theorem, which allows at least much more and even asynchronous parallelism.

Theorem: 3.2 *Let $F : D \subset \mathbb{R}^n \to \mathbb{R}^n$ be a function, and $\xi \in D$ a regular root of it. Then*

$$G_2 : (\mathbb{R}^n \times L(\mathbb{R}^n))^{n^2+1} \to L(\mathbb{R}^n)$$

given by

$$G_2(x^1, \mathrm{X}^1, ..., x^{n^2+1}, \mathrm{X}^{n^2+1}) \ :=$$

$$\mathrm{X}^2 \left(2\,I - \begin{pmatrix} F'(x^2)_{11} & \cdots & F'(x^{n+1})_{1n} \\ \vdots & & \vdots \\ F'(x^{n^2-n+2})_{n1} & \cdots & F'(x^{n^2+1})_{nn} \end{pmatrix} \mathrm{X}^2 \right), \tag{40}$$

is well-defined in some open neighborhood $(D_1 \times D_{F_1'})^{n^2+1}$ of $(\xi, F'(\xi)^{-1})$, and there are constants $K_0, K_1 \geq 0$, such that

$$\| \, G_2(x^1, \mathrm{X}^1, ..., x^{n^2}, \mathrm{X}^{n^2}) - F'(\xi)^{-1} \, \|_\infty \tag{41}$$

$$\leq \quad K_0 \max_{0 \leq l \leq n^2+1} \| \, x^l - \xi \, \|_\infty + K_1 \, \| \, \mathrm{X}^2 - F'(\xi)^{-1} \, \|_\infty^2$$

$$\forall \ (x^1, X^1, ..., x^{n^2+1}, \mathrm{X}^{n^2+1}) \in (D_1 \times D_{F_1'})^{n^2+1}.$$

Proof : G_2 is well-defined, since ξ is regular. For abbreviation, we denote the approximation for $F'(\xi)$ in (40) by F_x'. Let $\| \cdot \|$ be a matrix norm on $L(\mathbb{R}^n)$ and $c, C \geq 0$ be constants, such that

$$c \, \|A\|_\infty \leq \|A\| \leq C \, \|A\|_\infty \qquad\qquad \forall \ A \in L(\mathbb{R}^n).$$

Then we have

$$\|F_x' - F'(\xi)\| \leq C \, L \ \max_l \ \|x^l - \xi\|_\infty \ ,$$

if definition 2.1(ii) validates for $\| \cdot \|_\infty$. W.l.o.g. let $D_{F_1'}$ be bounded, say

$$\mathrm{X} \in \ D_{F_1'} \ \Rightarrow \ \|\mathrm{X}\| \leq K \qquad \text{for some} \ \ K > 0 \ .$$

Noweven

$$\|\mathrm{X}^2 \, (2\,I - F_x' \, \mathrm{X}^2 \,) \ - \ F'(\xi)^{-1}\|_\infty$$

$$\leq \ \frac{1}{c} \, \|F'(\xi)^{-1}\| \, \|F'(\xi) \ \mathrm{X}^2 - F'(\xi) \ \mathrm{X}^2 \ F'(\xi) \ \mathrm{X}^2 + F'(\xi) \ \mathrm{X}^2 - I \, \|$$

$$+ \ \frac{1}{c} \|\mathrm{X}^2\|^2 \ \|F_x' - F'(\xi)\| \tag{42}$$

$$\leq \ \frac{1}{c} \, \|F'(\xi)^{-1}\| \, \|F'(\xi)\mathrm{X}^2 - I\|^2 \ + \ K^2 \, \frac{C}{c} \, L \ \max_l \ \|x^l - \xi\|_\infty$$

$$\leq \ \frac{C^2}{c} \, \| \, F'(\xi)^{-1} \, \| \, \|F'(\xi)\|^2 \ \ \|\mathrm{X}^2 - F'(\xi)^{-1}\|_\infty^2 \ + \ K^2 \, \frac{C}{c} \, L \ \max_l \ \|x^l - \xi\|_\infty \ . \ \square$$

Remark, that it is possible to compute some *area of attraction* of such a Newton-Schulz type method explicitly by (42) and (30). Remark also, that we have given implicitly an asynchronous method for inverting matrices, which converges superlinearly, if moderate conditions hold. Observe further, that all partial derivatives of F in (40) can be evaluated asynchronously at different values of x, e.g. in practice, on different processors. This could be useful for many problems, where the evaluation of the partial derivatives is a hard task.

Remark: The results of this paper are part of [8].

References

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

[2] A. Berman and R.J. Plemmons. *Nonnegative matrices in the mathematical sciences* Academic Press. 1979.

[3] R. Bru, L. Elsner, and M. Neumann. *Models of parallel chaotic iteration methods* Linear Algebra and Its Applications, 103: 175-192. 1988.

[4] D. Chazan and W. Miranker. *Chaotic relaxation* . Linear Algebra and Its Applications, 2: 199-222. 1969.

[5] H.-P. Helfrich. *Ein modifiziertes Newton'sches Verfahren* . ISNM, 12: 61-70. 1967.

[6] D.P. O'Leary and R.E. White. *Multi-splittings of matrices and parallel solution of linear systems* . SIAM J. of Alg. Disc. Meth., 6(4): 630-640. 1985.

[7] J.M. Ortega and W.C. Rheinboldt. *Iterative solution of nonlinear equations in several variables* . Academic Press. 1970.

[8] M. Pott. *Asynchrone Iterationsverfahren für lineare und nichtlineare Gleichungssysteme.* Dissertation, Universität Bielefeld, 1994.

[9] G. Schulz. *Iterative Berechnung der reziproken Matrix* . Z. angew. Math. Mech., 13(1): 57-59. 1933.

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