Asynchronous Multisplitting Methods with Flexible Communication for Pseudolinear P.D.E.

Didier El Baz *, Pierre Spiteri **, Jean Claude Miellou ***

* LAAS du CNRS, 7, avenue du Colonel Roche, 31077 Toulouse Cedex 4, France,
Fax: (33) 5 61 33 69 69, E-mail: elbaz@laas.fr
** ENSEEIHT-IRIT/LIMA, 2, rue Camichel, 31071 Toulouse Cedex 4, France,
*** LCS, Université de Franche Comté, 16, Route de Gray, 25030 Besançon Cedex, France.

Keywords: flexible communication, asynchronous iterations, parallel computing, multisplitting technics, boundary value problem.

AMS subject classification: 65 N 12 - 65 N 55 - 68 Q 10 - 68 Q 22

Abstract: We consider asynchronous iterations allowing exchanges between processors of partial updates which results from updating phases in progress, i.e. parallel schemes of computation for which every processor can have access to the current value of the components of the iteration vector updated by any other processor. This concept was introduced in [17] in order to speed up the convergence of parallel iterative algorithms in the context of M-functions. Application of asynchronous iterations with flexible communication to the solution of nonlinear boundary value problems via multisplitting methods is presented in this paper.

1. Introduction.

In this paper we consider asynchronous multisplitting methods with flexible communication. The originality of this class of parallel iterative methods is that updating phases can use data which are issued from computations in progress. This leads to a better coupling between communication and computation. Asynchronous iterations with flexible communication allows the exchange between processors of partial updates which results from updating phases in progress i.e. they correspond to parallel schemes of computation whereby every processor can have access to the current value of the components of the iteration vector updated by any other processor. In contrast with previous studies (see [3]-[6], [8], [9], [12], and [15]), the methods presented here permit one to consider block iterative schemes with delayed access to block subvectors, the components of which are relative to different iteration numbers or intermediate computations not explicitly labelled by an iteration number. Another important feature of this paper is the use of mappings which approximate the solution (see [16]). Asynchronous iterations with flexible communication were introduced in [17] in order to speed up the convergence of parallel iterative algorithms in the context of M-functions. This new model of parallel asynchronous iterations was applied with success to subdomain methods for the solution of partial differential equations (see [21]) and gradient type methods for network flow problems (see [10] and [11]). We note that computational results on shared memory and distributed memory multiprocessors are presented in detail in [10], [11] and [21]). Application of asynchronous iterations with flexible communication to the solution of nonlinear boundary value problems via multisplitting methods is presented in this paper. This includes in particular, discrete analogs of the multisubdomain Schwarz method (see [2]) and two-stage multisplitting methods(see

[2], [7], [13], [14], and [18]).

In Section 2 we present the main problem, give some background material on multisplitting and study the properties of the extended nonlinear problem. Section 3 deals with parallel asynchronous multisplitting methods with flexible communication.

2. Pseudolinear problem and formal multisplitting

We consider the solution of the following pseudolinear problem:

$$Ax^* + \varphi(x^*) = 0, \tag{2.1}$$

where $A \in L(R^n)$ is an M-matrix, $\varphi : R^n \to R^n$ is a continuous diagonal maximal monotone operator. Let the following regular splittings of matrix A :

$$A = M^l - N^l, \quad l = 1, ..., m, \tag{2.2}$$

where $(M^l)^{-1} \geq 0$ and $N^l \geq 0$. Let $F^l : R^n \to R^n, \ l = 1, ..., m$ be the fixed point mappings associated with problem (2.1) and defined by:

$$F^l(x) = y \text{ such that } M^l(y) = N^l(x) - \varphi(x). \tag{2.3}$$

A formal multisplitting associated to problem (2.1) is defined by the collection of fixed point problems (see [2]):

$$x^* - F^l(x^*) = 0, \ l = 1, ..., m. \tag{2.4}$$

Let $E = (R^n)^m$ and consider the following block-decomposition of E: $E = \prod_{l=1}^{m} E_l$, where $E_l = R^n$. Each subspace E_l is endowed with the natural (or component-wise) partial ordering associated with the cone $K_l = R^n_+$ of vectors with nonnegative components in R^n. Let $X \in E$. We have the following block-decomposition of X :

$$X = \{X_1, \ldots, X_l, \ldots, X_m\} \in \prod_{l=1}^{m} E_l,$$

Definition 1. The extended fixed point mapping $T : E \to E$ associated with the formal multisplitting is defined as follows (see [2]):

$$T(X) = Y, \text{ such that } Y_l = F^l(Z_l) \text{ with } Z_l = \sum_{k=1}^{m} W_{lk} X_k, l = 1, ..., m, \tag{2.5}$$

where W_{lk} are nonnegative diagonal weighting matrices satisfying for all $l \in \{1, \ldots, m\}$

$$\sum_{k=1}^{m} W_{lk} = I,$$

I being is the identity matrix in $L(E_l)$. Let the following block-decomposition of the mapping T :

$$T(X) = \{T_1(X), \ldots, T_l(X), \ldots, T_m(X)\} \in \prod_{l=1}^{m} E_l.$$

The extended fixed point mapping T is associated with the following extended nonlinear problem:

$$a(X^*) = 0, \tag{2.6}$$

where the mapping $a : E \to E$ is given by:

$$a(X) = A^e X + \varphi^e(X), \tag{2.7}$$

the function $\varphi^e : E \to E$ being the extended monotone perturbation operator and for all $l \in \{1, \ldots, m\}$:

$$A_l^e X = M^l X_l - N^l \sum_{k=1}^{m} W_{lk} X_k. \tag{2.8}$$

In the sequel $a_l(X_1, \ldots, X_{l-1}, Y_l, X_{l+1}, \ldots, X_m)$ will also be denoted by $a_l(Y_l\,; X)$.

Proposition 1. Let the above assumptions hold. The mapping a is a continuous surjective M-function.

Proof: We recall that an M-function is off-diagonally antitone and inverse isotone (see [19] and [20]). It follows from (2.8) and the definition of matrices W_{lk} that the off-diagonal entries of the matrix A^e are nonpositive since A is an M-matrix. Let $U = \{u, ..., u\}$ be the vector of E with block-components equal to the eigenvector u associated with the greatest eigenvalue of the Jacobi matrix $D^{-1}(D - A)$, where D denotes the diagonal part of matrix A and consider the Jacobi matrix J derived from the extended system (2.6). For all $l \in \{1, \ldots, m\}$, we have:

$$J_l u = (D^e)_l^{-1} D_l^e u - (D^e)_l^{-1} M^l u + (D^e)_l^{-1} N^l \sum_{k=1}^{m} W_{lk} u =$$

$$(D^e)_l^{-1}(D_l^e - M^l + N^l)u = D^{-1}(D - A)u = \rho(D^{-1}(D - A))u.$$

It follows from [15] that the Jacobi matrix J is contracting relative to a weighted maximum norm and the result follows from Proposition 2.4.17 in [19]. The continuity and surjectivity of a follows from the continuity and maximal monotonicity of φ.

It follows from Proposition 1 that

$$\begin{cases} \text{For all } l \in \{1, \ldots, m\} \text{ and all } X \in E, \text{ the mapping:} \quad Y_l \to a_l(Y_l, X), \\ \text{is a continuous surjective } M\text{-function of } E_l \text{ onto } E_l, \end{cases} \tag{2.9}$$

(see [20, theorem 3.5]). Moreover it follows from Proposition 1 that

$$\text{(2.10)} \qquad \begin{cases} \text{for all } l \in \{1, \ldots, m\} \text{ and } X \in E \text{ the problem:} \quad a_l(Y_l\,;\,X) = 0, \\ \text{has a unique solution } Y_l. \end{cases}$$

It follows from the above assumptions that T is isotone on E (see [16]).

Definition 2. A vector $Y \in R_+^{nm}$ is an a-supersolution (an a-subsolution) if $a(Y) \geq 0 \quad (a(Y) \leq 0)$.

We note that the above multisplitting corresponds to the discrete analog of the multisubdomain Schwarz method.

We also note that the above results are still valid in the case where φ is a multivalued diagonal maximal monotone operator.

3. Asynchronous multisplitting methods with flexible communication

In Section 2 we have presented a collection of fixed point mappings associated with the exact solution of the problem (2.1). In this Section we consider fixed point mappings associated with approximate solutions of the problem (2.1). Such mappings were introduced in [16] and are used in this paper for the solution of problem (2.1) via parallel asynchronous iterative methods with flexible communication.

Throughout the paper we adopt the following notation for order intervals. Let $X_l, Y_l \in E_l$, be such that $X_l \leq Y_l$; then an order interval in E_l is denoted by:

$$\langle X_l, Y_l \rangle_l = \{X_l' \in E_l \mid X_l \leq X_l' \leq Y_l\}.$$

Similarly, let $X, Y \in E$, be such that $X \leq Y$; then an order interval in E is denoted by:

$$\langle X, Y \rangle = \{X' \in E \mid X \leq X' \leq Y\}.$$

Definition 3. Let a be an M-function and T defined by (2.5). Then the mapping T^a with components T_l^a is an a-supermapping associated with T if for all $l \in \{1, \ldots, m\}$ and Y element of the domain: $\{Y \in E \mid a_l(Y) \geq 0\}$, we have $T_l^a(Y) \leq Y_l$, $a_l(T_l^a(Y), Y) \geq 0$, and $T_l^a(Y) \neq Y_l$ if $T_l(Y) \neq Y_l$.

Remark 1. We can analogously define a-submappings associated with T by reversing the inequalities.

The class of asynchronous iterations presented in this paper allows flexible communication between processors. We introduce first the concepts of steering and sequence of delays.

Definition 4. A steering of components of the iteration vector $\mathcal{S}$, is a sequence $\{s(p)\}$ where $p \in N$, the set of natural integers, such that

$$\text{(3.1)} \qquad \text{for all } p \in N, \quad s(p) \in \{1, \ldots, m\}.$$

Definition 5. A sequence of maximal delays in the access to the updated components of the iteration vector, is denoted by $\{r(p)\}$ (or $\mathcal{R}$) and defined by

$$r(p) = \{r_1(p), \ldots, r_l(p), \ldots, r_m(p)\} \in N^m, \text{ for all } p \in N,$$

where for all $l \in \{1, \ldots, m\}$ and $p \in N$, $0 \leq r_l(p) \leq p$ and $r_l(p) = 0$ if $l = s(p)$.

We note that the latest condition has been introduced in order to simplify the presentation of the algorithms. This assumption is more restrictive than the one used in the classical model, however it is natural since it implies that each processor has access to its own working space without delays. We introduce the functions $\rho_l : N \to N, l \in \{1, \ldots, m\}$, defined by $\rho_l(p) = p - r_l(p)$ and satisfying for all $l \in \{1, \ldots, m\}$ and $p \in N$:

$$0 \leq \rho_l(p) \leq p, \tag{3.2}$$

$$\rho_l(p) = p, \text{ if } l = s(p). \tag{3.3}$$

In the sequel, we will denote by $\rho(p)$ the mapping from N^m into itself, with components $\rho_l(p), l \in \{1, \ldots, m\}$. We introduce now the sequence $\{K_l^p\}, p \in N$ defined by

$$K_l^p = \{k \in N \mid s(k) = l, 0 \leq k < p\}, l \in \{1, \ldots, m\}, p \in N. \tag{3.4}$$

The elements of the set K_l^p correspond to the iteration numbers between 0 and p at which the l-th block-component of the iteration vector is updated. We note that

(3.5) $\left\{K_l^p\right\}$ is a denumerable sequence of finite elements of the set of parts of N.

Moreover $\left\{K_l^p\right\}$ is nested:

$$K_l^p \subseteq K_l^{p+1}. \tag{3.6}$$

Definition 6. Asynchronous multisplitting iterations with flexible communication associated with the a-supermapping T^a are sequences $\{Y^p\}$ defined recursively as follows. For all $p \in N$,

$$\begin{cases} Y_l^{p+1} = T_l^a(\tilde{Y}^p), & \text{if } l = s(p), \\ Y_l^{p+1} = Y_l^p, & \text{if } l \neq s(p), \end{cases} \tag{3.7}$$

where $\tilde{Y}^p$ is implicitly defined by

$$\tilde{Y}^0 = Y^0, \text{ an } a\text{-supersolution} \tag{3.8}$$

and

$$\tilde{Y}^p \in \langle Y^p, Y^{\rho(p)} \wedge \tilde{Y}^q \rangle, \text{ if } p \geq 1, \tag{3.9}$$

the block-components of the vector $Y^{\rho(p)} \in E$ are the subvectors $Y_l^{\rho_l(p)}, l \in \{1, \ldots, m\}$, and $q = \max_{k \in K_{s(p)}^p} k$.

We note that $\tilde{Y}^p \in \langle Y^p, Y^{\rho(p)} \rangle$ if $K_l^p = \emptyset$, where $\emptyset$ denotes the empty set.

Remark 2. Asynchronous iterations defined by (3.7)-(3.9) are general iterative methods whereby iterations are carried out in parallel by m processors in arbitrary order and possibly without any synchronization. Flexible communication between processors is the main originality of this class of algorithms. Indeed we note that the block-components $\tilde{Y}_j^p$ of the iteration vector used at the time of the updating of Y_l^{p+1} are taken anywhere in the interval $\langle Y_j^p, Y_j^{\rho(p)} \wedge \tilde{Y}_j^q \rangle_j$, where $\tilde{Y}_j^q$ denotes the value used at the time of the latest updating of the block-component Y_l and $Y_j^{\rho(p)}$ is related to the non-deterministic behavior of the iterative scheme. The introduction of order intervals permits us to take into account data coming from computations which are in progress and which are not explicitly labelled by an iteration number. So, there is a better coupling between communication and computation and we may expect a faster convergence. Moreover each communication of a block subvector is not necessarily associated with a given iteration number. The exchanged values of the components of each block subvector can be relative to different iteration numbers.

Proposition 2. Let a be a continuous surjective M-function, T the fixed point mapping defined by (2.5), T^a an a-supermapping associated with T, $Y^0 \in E$ an a-supersolution, $\mathcal{S}$, and $\mathcal{R}$ a steering and a sequence of delays, respectively. Then, the asynchronous multisplitting iteration $\{Y^p\}$ given by (3.7)-(3.9) is well defined and satisfies:

$$Y^* \leq \cdots \leq Y^{p+1} \leq Y^p \leq \cdots \leq Y^0,$$

and $\lim_{p\to\infty} Y^p = Y^*$, where Y^* is an a-supersolution of problem (2.6).

Proof: The result follows from Proposition 1 and Proposition 6 in [17].

We now introduce two classes of a-supermappings for which we have convergence properties
Definition 7. The mapping T^a with components T_l^a defined on the domain $\{Y \in E \mid a_l(Y) \geq 0\}$ is an a-supermapping of the first kind associated with T if T^a is an a-supermapping and there exists an a-supermapping T'^a with components $T_l'^a$ defined on the domain $\{Y \in E \mid a_l(Y) \geq 0,\}$ such that for all $l \in \{1, \ldots, m\}$ and Y element of the domain $\{Y \in E \mid a_l(Y) \geq 0\}$, we have

$$T_l'^a(Y) \in \langle T_l^a(Y), Y_l \rangle_l. \tag{3.10}$$

and

$$Y^p \downarrow Y^*, p \to \infty \text{ implies } T_l'^a(Y^p) \downarrow T_l'^a(Y^*), p \to \infty \text{ for all } l \in \{1, \ldots, m\}, \tag{3.11}$$

where the notation $Y^p \downarrow Y^*, p \to \infty$ means that

$$Y^* \leq \cdots \leq Y^{p+1} \leq Y^p \leq \cdots \leq Y^0 \text{ and } \lim_{p\to\infty} Y^p = Y^*.$$

Remark 3. The relation (3.11) can be interpreted as a property of continuity of T'^a related to the partial ordering.

Proposition 3. Let the assumptions of Proposition 2 hold. Let T^a be an a-supermapping of the first kind associated with T and assume that the steering $\mathcal{S}$ satisfies:

$$\text{for all } l \in \{1, \ldots, m\}, \text{ the set } \{p \in N \mid s(p) = l\} \text{ is infinite and} \tag{3.12}$$

$$\text{(3.13)} \qquad \text{for all } l \in \{1, \ldots, m\}, \quad \lim_{p \to \infty} \rho_l(p) = +\infty.$$

Then, the asynchronous multisplitting iteration $\{Y^p\}$ defined by (3.7)-(3.9) satisfies $Y^p \downarrow X^*$, where X^* is the unique solution of problem (2.6).

Proof: The result follows from Proposition 1 and Proposition 7 in [17].

Definition 8. The mapping T^a with components T_l^a is an a-supermapping of the second kind associated with T if T^a is an a-supermapping and for all $l \in \{1, \ldots, m\}$ and Y element of the domain: $\{Y \in E \mid a_l(Y) \geq 0\}$, there exists $\delta > 0$ such that

$$\text{(3.14)} \qquad \|Y_l - T_l^a(Y)\|_l \geq \delta \|Y_l - T_l(Y)\|_l,$$

where $\|.\|_l$ is a norm defined in E_l.

Proposition 4. Let the assumptions of Proposition 2 hold and assume that T^a is an a-supermapping of the second kind associated with T. Then, the asynchronous multisplitting iteration $\{Y^p\}$ defined by (3.7)-(3.9) satisfies $Y^p \downarrow X^*$, where X^* is the unique solution of problem (2.6).

Proof: The result follows from Proposition 1 and Proposition 8 in [17].

Parallel asynchronous multisplitting methods with flexible communication can be applied for example to the solution of the following problem.

Find v such that

$$\text{(5.6)} \qquad \begin{cases} -\Delta v + e^{cv} = f & \text{everywhere in } \Omega, \\ v \mid \Omega = 0, \end{cases}$$

where Ω is an open domain of R^2 (or R^3), c is positive real number, and we consider discretization using five-point difference equations for the Laplace operator. Inner iteration can be for example a conjugate gradient method used with an LU incomplete preconditionning.

Bibliography

[1] Bahi, J. and Miellou, J.-C. Contractive mappings with maximum norms: comparisons of constants of contraction and application to asynchronous iterations. Parallel Computing 19, 511-523 (1993).

[2] Bahi, J. ; Miellou, J.-C. and Rhofir K. Asynchronous multisplitting methods for nonlinear fixed point problems LCS Report (1997), to appear in Numerical Algorithms.

[3] Baudet, G.M. Asynchronous iterative methods for multiprocessors. J.A.C.M. 25, 226-244 (1978).

[4] Bertsekas, D.P. and El Baz, D. Distributed asynchronous relaxation methods for convex network flow problems. SIAM J. on Control and Optimization 25, 74-85 (1987).

[5] Bertsekas, D.P. and Tsitsiklis, J. Parallel and Distributed computation, Numerical Methods. Englewood cliffs: Prentice Hall 1989.

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

[7] Bru, R. ; Migallon, V. ; Penadés J. ; Szyld D. , Parallel synchronous and asynchronous two-stage multisplitting methods, ETNA, 3, 24-38, (1995).

[8] El Baz, D. M-functions and parallel asynchronous algorithms. SIAM J. Numer. Anal. 27, 136-140 (1990).

[9] El Baz, D. Distributed asynchronous gradient algorithms for convex network flow problems. Computational Optimization and Applications, 5, 187-205 (1996).

[10] El Baz, D. ; Spiteri, P. ; Miellou, J.-C. ; and Gazen D. Asynchronous iterative methods with flexible communication for nonlinear network flow problems Journal of Parallel and Distributed Computing, 38, 1-15 (1996).

[11] El Baz, D. ; Gazen, D. ; Miellou, J.-C. and Spiteri, P. Mise en oeuvre de méthodes itératives asynchrones avec communication flexible. Application à la résolution d'une classe de problèmes d'optimisation, Calculateurs Parallèles, 8, 393-410, (1996).

[12] El Tarazi, M.N. Some convergence results for asynchronous algorithms. Num. Math. 39, 325-340 (1982).

[13] Frommer, A. ; Szyld, D. Asynchronous two-stage iterative methods, Num. Math. 69, 141-153 (1994).

[14] Jones, M. ; Szyld, D. Two-stage multisplitting methods with overlapping blocks, Numer. Linear Algebra Appl., 3 113-124 (1996).

[15] Miellou, J.-C. Algorithmes de relaxation chaotique à retards. R.A.I.R.O. R1, 55-82 (1975).

[16] Miellou, J.-C. Asynchronous iterations and order intervals, Parallel Algorithms (M. Cosnard et al. eds) Amsterdam: North-Holland, 85-96 (1986).

[17] Miellou, J.-C. ; El Baz, D. ; Spiteri P. A new class of asynchronous iterative algorithms with order intervals, LCS report 1994 to appear in Mathematics of Computation.

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

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

[20] Rheinboldt, W.C. On M-functions and their application to nonlinear Gauss-Seidel iterations and to network flows, J. Math. Anal. and Appl. 32, 274-307 (1970).

[21] Spiteri, P. ; Miellou, J.-C. and El Baz D. Asynchronous alternating Schwarz method for the solution of nonlinear partial differential equations. IRIT report 95-17-R, LCS 195.10, LAAS Report 95309 (1995).