Newton Additive and Multiplicative Schwarz Iterative Methods

Josep Arnal, Violeta Migallón,
José Penadés, and Daniel B. Szyld

Report 05-12-20
December 2005

This report is available in the World Wide Web at
`http://www.math.temple.edu/~szyld`

NEWTON ADDITIVE AND MULTIPLICATIVE SCHWARZ ITERATIVE METHODS*

JOSEP ARNAL†, VIOLETA MIGALLÓN†, JOSÉ PENADÉS†, AND DANIEL B. SZYLD‡

Abstract. Convergence properties are presented for Newton additive and multiplicative Schwarz iterative methods for the solution of nonlinear systems in several variables. These methods consist of approximate solutions of the linear Newton step using either additive or multiplicative Schwarz iterations, where overlap between subdomains can be used. Restricted versions of these methods are also considered. Numerical experiments on parallel computers are presented, indicating the effectiveness of these methods.

Key words. Nonlinear systems. Newton's method. Additive Schwarz. Multiplicative Schwarz. Iterative methods. Subspace correction. Parallel computing.

AMS subject classifications. 65H10, 65N55, 65Y05

1. Introduction and Preliminaries. We are interested in the parallel solution of the system of nonlinear equations

$$F(x) = 0, \tag{1.1}$$

where F is a map $F : \mathbb{R}^n \to \mathbb{R}^n$, and it is assumed that a solution $x^\star$ of (1.1) exists. A well-known method for solving the nonlinear system (1.1) is the classical Newton method; see, e.g., [15]. Given an initial vector $x^{(0)}$, this method produces the following sequence of vectors

$$x^{(\ell+1)} = x^{(\ell)} - \delta_x^{(\ell)}, \qquad \ell = 0, 1, \ldots, \tag{1.2}$$

where $\delta_x^{(\ell)}$ is the solution of the linear system

$$F'(x^{(\ell)})z = F(x^{(\ell)}), \tag{1.3}$$

with $F'(x)$ denoting the Jacobian of F at x.

If an iterative method is used to approximate the solution of (1.3), then one obtains a Newton-iterative method; see, e.g., [15], [16]. In this paper we analyze parallel Newton-iterative algorithms for the solution of the general nonlinear system (1.1) in which additive and multiplicative Schwarz iterations are used as secondary iterations to approximate the solution of the associated linear system (1.3) at each Newton step. These Schwarz methods are attractive because they are easily parallelizable, and because they allow for overlap, i.e., the same variable is updated by more than one processor; see, e.g., section 2, or the references [8], [17], [19]. In this paper, we prove the convergence of the Newton additive and multiplicative Schwarz iterative methods in two cases: when the Jacobian is symmetric positive definite, or a monotone matrix. In the monotone case, we also study restricted additive and multiplicative Schwarz iterative methods; see section 2, or the references [7], [11], [14], [18]. In all cases, we

*This version 20 December 2005

†Departamento de Ciencia de la Computación e Inteligencia Artificial, Universidad de Alicante, E-03080 Alicante, Spain (`{arnal, violeta, jpenades}@dccia.ua.es`). This research was supported by the Spanish Ministry of Science and Education under grant TIN2005-09037-C02-02.

‡Department of Mathematics, Temple University, Philadelphia, PA 19122, USA. `szyld@math.temple.edu`. Research supported in part by the U.S. Department of Energy under grant DE-FG02-05ER25672.

consider both exact and inexact local solvers. We also illustrate the effectiveness of the proposed methods using parallel computers with some numerical experiments.

Schwarz methods have been used essentially in two different forms for the solution of nonlinear systems. As we just mentioned, the approach considered in this paper, consists of the use of Newton's method, with the associated linear system solved by a Schwarz iterative method as a secondary (inner) iteration; see [4] for some comments on this approach for some special cases. Another approach consists in the use of Schwarz methods, directly on the nonlinear problem (1.1); see, e.g., [6], [9].

We briefly describe now the Newton-iterative methods. Consider for each x a splitting $F'(x) = M(x) - N(x)$, with $M(x)$ nonsingular. This splitting determines the iteration matrix $H(x) = M(x)^{-1}N(x) = I - M(x)^{-1}F'(x)$, and an iterative method for the solution of (1.3). At the ℓth Newton step, one computes m_ℓ iterations with such a splitting, and for simplicity we assume that the initial (inner) iterate is taken as the zero vector; cf. [16]. The Newton-iterative method starting with an initial vector $x^{(0)}$ is thus given by

$$x^{(\ell+1)} = x^{(\ell)} - (H(x^{(\ell)})^{m_\ell - 1} + H(x^{(\ell)})^{m_\ell - 2} + \cdots + I)M(x^{(\ell)})^{-1}F(x^{(\ell)}), \quad \ell = 0, 1, \ldots. \tag{1.4}$$

From $F'(x) = M(x) - N(x)$, we have that $F'(x)^{-1} = (I - H(x))^{-1}M(x)^{-1}$, and thus, we can express (1.4) as

$$x^{(\ell+1)} = x^{(\ell)} - A_{m_\ell}(x^{(\ell)})F(x^{(\ell)}) = G_{m_\ell}(x^{(\ell)}), \qquad \ell = 0, 1, \ldots, \tag{1.5}$$

where for each positive integer m

$$A_m(x) = (I - H(x)^m)F'(x)^{-1} \quad \text{and} \quad G_m(x) = x - A_m(x)F(x). \tag{1.6}$$

Schwarz iterations can be characterized by certain splittings [2], [10], [11], [14], and we review this in sections 2 and 4. Therefore, in this paper we analyze the convergence of the method (1.5) for the particular splittings which define Schwarz iterations; see sections 3 and 5. Numerical experiments are reported in section 6.

In the rest of this section, we present some auxiliary notation, review some convergence results of Newton-iterative methods, and some other preliminary results. We denote by $L(\mathbb{R}^n)$ the set of linear operators from $\mathbb{R}^n$ to $\mathbb{R}^n$. Let $\{x^{(\ell)}\}$ be a sequence in $\mathbb{R}^n$ convergent to $x^\star$. We define its convergence rate as

$$R_1(x^{(\ell)}) = \lim_{\ell \to \infty} \sup \|x^{(\ell)} - x^\star\|^{1/\ell}.$$

Let us consider the following usual conditions on F and on the splitting $F'(x) = M(x) - N(x)$. We suppose that there exists an $r_0 > 0$ such that

(i) F is differentiable on $S_0 \equiv \{x \in \mathbb{R}^n \ : \ \|x - x^\star\| < r_0\}$,
(ii) the Jacobian matrix at $x^\star$, $F'(x^\star)$, is nonsingular,
(iii) there exists an $L > 0$ such that for $x \in S_0$, $\|F'(x) - F'(x^\star)\| \leq L\|x - x^\star\|$,
(iv) $M(x)$ is continuous at $x^\star$,
(v) $M(x^\star)$ is nonsingular,
(vi) there exists λ, $0 < \lambda < 1$, such that $\|H(x^\star)\| < \lambda$,
(vii) there exists $L_1 < +\infty$ such that for $x \in S_0$, $\|M(x) - M(x^\star)\| \leq L_1\|x - x^\star\|$.

Under hypotheses (i)-(iii) it is well-known (see, e.g., [15]) that the iterative method (1.2) converges $Q-$quadratically to $x^\star$ for $x^{(0)}$ in a neighborhood of $x^\star$. Also, under

hypotheses (i)-(vi) the iterative method (1.5) converges for $x^{(0)}$ in a neighborhood of the solution $x^\star$ and moreover the following result holds.

THEOREM 1.1. [15] *Let us consider $F : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^n$ and let $x^\star \in D$ be such that $F(x^\star) = 0$. Assume that F satisfies:*

1. *$F(x)$ is G-differentiable on an open neighborhood $S_0 \subseteq D$ of $x^\star$,*
2. *$F'(x)$ is continuous at $x^\star$.*

Assume further that $F'(x) = M(x) - N(x)$, where $M, N : S_0 \longrightarrow L(\mathbb{R}^n)$ satisfy:

1. *$M(x)$ is continuous at S_0,*
2. *$M(x^\star)$ is nonsingular,*
3. *$\rho(H(x^\star)) < 1$ where $H(x) = M(x)^{-1}N(x)$.*

Then there exists an open neighborhood S of $x^\star$ such that, for each $x^{(0)} \in S$ and for each sequence of positive integers m_ℓ, the iteration given by (1.5) is well-defined and converges to $x^\star$. Moreover,

$$R_1(x^{(\ell)}) \leq \rho(H(x^\star))^{m'}, \quad m' = \lim_{\ell \to \infty} \inf m_\ell.$$

On the other hand, under hypotheses (i)-(vii), Sherman [16] proves the following result.

THEOREM 1.2. [16] *Assume that F satisfies* (i)-(iii) *and let $F'(x) = M(x) - N(x)$ be a splitting satisfying* (iv)-(vii). *Let $\{m_\ell\}_{\ell=0}^{\infty}$ be a sequence of positive integers and let*

$$m = \max \left[\{m_0\} \cup \left\{ m_\ell - \sum_{i=0}^{\ell-1} m_i \ : \ \ell = 1, 2, \ldots \right\} \right]. \tag{1.7}$$

Assume that $m < +\infty$, then there exist a neighborhood S of $x^\star$, and $c > 0$ with $\lambda < c < 1$ such that for $x^{(0)} \in S$ the sequence of iterates defined by (1.5) *satisfies*

$$\|x^{(\ell+1)} - x^\star\| \leq c^{m_\ell} \|x^{(\ell)} - x^\star\|. \tag{1.8}$$

We end this section by recalling a result that we use in the later sections. Given an iteration matrix, there is a unique splitting which defines it; see, e.g., [3].

LEMMA 1.3. *Let A and T be square matrices such that A and $I - T$ are nonsingular. Then, there exists a unique pair of matrices M, N, such that M is nonsingular, $T = M^{-1}N$ and $A = M - N$. The matrices are $M = A(I - T)^{-1}$ and $N = M - A = A((I - T)^{-1} - I)$.*

2. Additive Schwarz Methods. We now describe the additive Schwarz methods and give some auxiliary results. Consider a nonsingular linear system in $V = \mathbb{R}^n$

$$Az = b. \tag{2.1}$$

We consider p nonoverlapping subspaces $V_{k,0}$, $k = 1, \ldots, p$, which are spanned by columns of the identity I over $\mathbb{R}^n$ and which are then augmented to produce overlap. For a precise definition, let $S_n = \{1, \ldots, n\}$ and let

$$S_n = \bigcup_{k=1}^{p} S_{k,0}$$

be a partition of S_n into p disjoint, non-empty subsets. For each of these sets $S_{k,0}$ we consider a nested sequence of larger sets $S_{k,\delta}$ with

$$S_{k,0} \subseteq S_{k,1} \subseteq S_{k,2} \subseteq \cdots \subseteq S_n = \{1, \ldots, n\}, \tag{2.2}$$

so that we again have

$$S_n = \bigcup_{k=1}^{p} S_{k,\delta}$$

for all values of the integers $\delta \geq 0$, but for $\delta > 0$ the sets $S_{k,\delta}$ are not pairwise disjoint, i.e., there is *overlap*. A common way to obtain the sets $S_{k,\delta}$ is to add those indices to $S_{k,0}$ which correspond to nodes lying at distance δ or less from those nodes corresponding to $S_{k,0}$ in the (undirected) graph of A. This approach is particularly adequate in discretizations of partial differential equations where the indices correspond to the nodes of the discretization mesh; see, e.g., [7], [17].

Let $n_{k,\delta} = |S_{k,\delta}|$ denote the cardinality of the set $S_{k,\delta}$. For each nested sequence of the form (2.2) we can find a permutation π_k on $\{1, \ldots, n\}$ with the property that for all $\delta \geq 0$ we have $\pi_k(S_{k,\delta}) = \{1, \ldots, n_{k,\delta}\}$. We now build matrices $R_{k,\delta} \in \mathbb{R}^{n_{k,\delta} \times n}$ whose rows are precisely those rows j of the identity for which $j \in S_{k,\delta}$. Formally, such a matrix $R_{k,\delta}$ can be expressed as

$$R_{k,\delta} = [I_{k,\delta}|O]\,\pi_k \tag{2.3}$$

with $I_{k,\delta}$ the identity on $\mathbb{R}^{n_{k,\delta}}$. For each δ, we define the subspaces $V_{k,\delta} \subset V$ as the range of $R_{k,\delta}^T$. Thus, we have $\dim V_{k,\delta} = n_{k,\delta}$, $k = 1, \ldots, p$, and

$$\sum_{k=1}^{p} V_{k,\delta} = \left\{ x \in V : x = \sum_{k=1}^{p} v_k,\ v_k \in V_{k,\delta} \right\} = V.$$

In the sequel, for simplicity, and when no confusion may arise, we do not include δ in the subscripts whenever $\delta > 0$. We identify V_k with $\mathbb{R}^{n_k}$, an we have that $R_k : V \to V_k$ acts as a restriction operator and R_k^T is a prolongation operator from $\mathbb{R}^{n_k}$ to V.

In addition to the p subspaces just described (for a given overlap δ), some times a "coarse grid correction" is used; see, e.g., [19]. Formally, there is a set of nodes $S_0 \subset S_n$, with $n_0 = |S_0|$, a subspace $V_0 \subset V$, and a corresponding restriction operator R_0, which in this paper is assumed to be of the form (2.3).

Denote by $A_k = R_k A R_k^T$, the restriction of the operator A to the subspace V_k. In the cases considered here A_k is nonsingular. Then A_k is a symmetric permutation of an $n_k \times n_k$ principal submatrix of A. We also denote by

$$P_k = R_k^T A_k^{-1} R_k A = R_k^T (R_k A R_k^T)^{-1} R_k A.$$

Given an initial approximation $z^{(0)}$ to the linear system $Az = b$, the damped additive Schwarz (AS) iteration (see, e.g., [10], [12]), can be written as the iteration

$$z^{(s+1)} = z^{(s)} + \theta \sum_{k=1}^{p} R_k^T A_k^{-1} R_k (b - A z^{(s)}), \quad s = 0, 1, \ldots, \tag{2.4}$$

where $\theta > 0$ is the damping factor. The iteration matrix for the method (2.4) is given by

$$T_\theta = I - \theta \sum_{k=1}^{p} R_k^T A_k^{-1} R_k A, \tag{2.5}$$

and the error $e^{(s+1)} = z^{(s+1)} - z^\star$ satisfies $e^{(s+1)} = T_\theta e^{(s)}$, where $z^\star$ is the exact solution of (2.1). When we add a coarse grid correction, the sum in (2.5) goes from 0 to p.

Beginning with $z^{(0)}$, the restricted additive Schwarz (RAS) iteration consists of

$$z^{(s+1)} = z^{(s)} + \sum_{k=1}^{p} R_{k,0}^T A_k^{-1} R_k (b - A z^{(s)}), \quad s = 0, 1, \ldots, \tag{2.6}$$

and the corresponding iteration matrix is

$$T_{RAS} = I - \sum_{k=1}^{p} R_{k,0}^T A_k^{-1} R_k A. \tag{2.7}$$

It can be appreciated that the difference between AS and RAS is that in the latter case, $\theta = 1$, and that in one occurrence one uses the restriction $R_{k,0}$ to $V_{k,0} \subset V_k = V_{k,\delta}$. The advantage of RAS is that while the local problems are solved taking into account the overlap, the information collected (and usually send to the other processors) corresponds to the subspaces without the overlap $V_{k,0}$. Interprocessor communication is reduced resulting in less overall convergence time; see further [7] and our own numerical experiments reported in Table 6.1 (section 6). If a coarse grid correction is present, an additional term of the form $R_0^T A_0^{-1} R_0 A$ is added to the sum in (2.7).

Very often in practice, instead of solving the local problems $A_k y_k = R_k(b - A z^{(s)})$ exactly, such linear systems are approximated. Let $\tilde{A}_k$ denote the approximation of A_k used, i.e., the inexact local solver is $\tilde{A}_k^{-1}$. By replacing A_k with $\tilde{A}_k$ in (2.4) and (2.5) (or in (2.6) and (2.7)) one obtains the damped (restricted) additive Schwarz iteration with inexact local solvers, and their iteration matrices are given by

$$\tilde{T}_\theta = I - \theta \sum_{k=1}^{p} R_k^T \tilde{A}_k^{-1} R_k A \quad \text{and} \quad \tilde{T}_{RAS} = I - \sum_{k=1}^{p} R_{k,0}^T \tilde{A}_k^{-1} R_k A.$$

We present now the splitting that defines the additive Schwarz method (2.4).

LEMMA 2.1. [12] *Let T_θ be the matrix given by* (2.5)*, there exist two matrices M_θ and N_θ such that $A = M_\theta - N_\theta$, M_θ is nonsingular and $T_\theta = M_\theta^{-1} N_\theta$. Moreover*

$$M_\theta^{-1} = \theta B = \theta \sum_{k=1}^{p} R_k^T A_k^{-1} R_k.$$

A similar result was shown in [11] for RAS: there exists a nonsingular matrix M_{RAS} such that $M_{RAS}^{-1} = \sum_{k=1}^{p} R_{k,0}^T A_k^{-1} R_k$, for which one can write $A = M_{RAS} - N_{RAS}$ and $T_{RAS} = M_{RAS}^{-1} N_{RAS}$. Similar splittings exist when a coarse grid correction is added [2], [11].

Note that by Lemma 1.3, these splittings are unique. In the next two subsections we recall some convergence results of the additive Schwarz method (2.4) and of the restricted additive Schwarz method (2.6). We also remark that when AS is used as a preconditioner, the preconditioner is M_1, i.e., with no damping, and the preconditioned matrix is $M_1^{-1} A = I - T_1$. Similarly, the preconditioned matrix with RAS is $M_{RAS}^{-1} A = I - T_{RAS}$.

2.1. Convergence for Symmetric Positive Definite Matrices. A matrix A is symmetric positive definite, denoted $A \succ O$, if it is symmetric and if for all vectors $x \neq 0$, $x^T Ax > 0$; $A \succ O$ implies $A^{-1} \succ O$. If $A \succ O$, an associated vector norm is defined as $\|v\|_A = (v^T Av)^{1/2}$. An operator norm is associated to the vector norm in the usual manner. A splitting $A = M - N$ is called P-regular if $M^T + N$ is positive definite.

LEMMA 2.2. [10] *Suppose that A is symmetric positive definite. Then $A = M - N$ is a P-regular splitting if and only if $\left\|M^{-1}N\right\|_A < 1$.*

It has been shown in [10] and [12] that if $A \succ O$ and $\theta < 1/p$, the iteration (2.4) converges, i.e., $\rho(T_\theta) < 1$. Furthermore, the following bound holds for some $\gamma > 0$,

$$\|T_\theta\|_A \leq \gamma < 1. \tag{2.8}$$

In the same references, the allowable range for the damping parameter θ is increased by considering coloring. The sets V_k can be colored using $q << p$ colors as follows, V_i and V_j have different color if $V_i \cap V_j \neq \{0\}$. It is proved that the expression (2.8) holds using the weaker condition $\theta < 1/q$. In summary, we have the following lemma.

LEMMA 2.3. *Assume that $A \succ O$, and let $A = M_\theta - N_\theta$ be the splitting given by Lemma 2.1. If $\theta < 1/q$, then this splitting is P-regular and thus $\|T_\theta\|_A < 1$.*

Next we study the case of inexact local solvers. When each inexact solver is symmetric positive definite, then it can be proved similarly to the exact case, as in [10] and [12], that $\tilde{B} = \sum_{k=1}^p R_k^T \tilde{A}_k^{-1} R_k$ is nonsingular.

LEMMA 2.4. *Suppose that $A \succ O$. If for a given $\mu > 0$, $A_k \preceq \mu \tilde{A}_k$, for $k = 1, \ldots, p$, then $A \preceq \mu q \tilde{M}$, where $\tilde{M}^{-1} = \sum_{k=1}^p R_k^T \tilde{A}_k^{-1} R_k$.*

From this lemma it follows that the unique splitting induced by $\tilde{T}_\theta$, $A = \tilde{M}_\theta - \tilde{N}_\theta$ (with $\tilde{M}_\theta^{-1} = \theta \tilde{B}$) is P-regular, assuming that $\theta < 1/(\mu q)$, and thus $\|\tilde{T}_\theta\|_A < 1$.

2.2. Convergence for Monotone Matrices. A nonsingular matrix A is called monotone if $A^{-1} \geq O$. A monotone matrix A is called a nonsingular M-matrix if it has nonpositive off-diagonal elements.

LEMMA 2.5. [10] *Let A be a nonsingular M-matrix. Let R_k be matrices of the form* (2.3). *Then $B = \sum_{k=1}^p R_k^T A_k^{-1} R_k$ is nonsingular. Moreover, $A_k^{-1} \leq R_k A^{-1} R_k^T$ for $k = 1, \ldots, p$.*

For a positive vector $\omega \in \mathbb{R}^n$ we define an associated vector norm as follows:

$$\|v\|_\omega = \max_{i=1,\ldots,n} |v_i|/\omega_i.$$

An operator norm is associated to the vector norm in the usual manner.

THEOREM 2.6. [10] *Let A be a nonsingular M-matrix. If $\theta < 1/q$, then the damped additive Schwarz iteration* (2.4) *converges to the solution of* (2.1), *and there exists a positive vector ω and $0 < \gamma < 1$ such that $\|T_\theta\|_\omega \leq \gamma$.*

THEOREM 2.7. [10] *Let A be a nonsingular M-matrix. Assume that $\tilde{A}_k$, $k = 1, \ldots, p$, are monotone matrices such that*

$$\tilde{A}_k^{-1}(\tilde{A}_k - A_k) \geq O, \ k = 1, \ldots, p.$$

Then, if $\theta \leq 1/q$, the damped additive Schwarz iteration with inexact local solvers converges to the solution of (2.1), *and there exists a positive vector ω and $0 < \gamma < 1$ such that $\|\tilde{T}_\theta\|_\omega \leq \gamma$.*

Similar results were shown in [11] in the case of RAS: If A is a nonsingular M-matrix, RAS converges to the solution of (2.1), and there exists a positive vector

ω and $0 < \gamma < 1$ such that $\|T_{RAS}\|_\omega \leq \gamma$. If the inexact local solvers satisfy the hypotheses of Theorem 2.7, we similarly have a positive vector ω and $0 < \gamma < 1$ such that $\|\tilde{T}_{RAS}\|_\omega \leq \gamma$. Similar results hold when a coarse grid correction is added; see [2], [11].

3. Newton Additive Schwarz Methods. In this section we consider the solution of the linear systems (1.3) by additive Schwarz methods of the form (2.4) and (2.6), and its variants, thus obtaining the Newton additive Schwarz methods. We define the corresponding iteration matrix as

$$T_\theta(x) = I - \theta \sum_{k=1}^{p} R_k^T F_k'(x)^{-1} R_k F'(x), \tag{3.1}$$

where $F_k'(x) = R_k F'(x) R_k^T$, $k = 1, \ldots, p$. Then, by Lemma 2.1 there exist two matrices $M_\theta(x)$ and $N_\theta(x)$ such that $F'(x) = M_\theta(x) - N_\theta(x)$, $M_\theta(x)$ is nonsingular, and $T_\theta(x) = M_\theta(x)^{-1} N_\theta(x)$. Moreover

$$M_\theta(x)^{-1} = \theta B(x) = \theta \sum_{k=1}^{p} R_k^T F_k'(x)^{-1} R_k. \tag{3.2}$$

Similarly, using the restricted additive Schwarz method for the solution of (1.3), we get the iteration matrix

$$T_{RAS}(x) = I - \sum_{k=1}^{p} R_{k,0}^T F_k'(x)^{-1} R_k F'(x), \tag{3.3}$$

and the matrix defining the induced splitting is

$$M_{RAS}(x)^{-1} = \sum_{k=1}^{p} R_{k,0}^T F_k'(x)^{-1} R_k.$$

With this notation, as in (1.4), the Newton additive Schwarz iteration as well as the restricted case can be expressed as

$$\begin{aligned} x^{(\ell+1)} &= x^{(\ell)} - (T(x^{(\ell)})^{m_\ell - 1} + T(x^{(\ell)})^{m_\ell - 2} + \cdots + I) M(x^{(\ell)})^{-1} F(x^{(\ell)}), \\ &\ell = 0, 1, \ldots, \end{aligned} \tag{3.4}$$

where the iteration matrices are $T(x) = T_\theta(x)$ or $T(x) = T_{RAS}(x)$, with $M(x) = M_\theta(x)$ or $M(x) = M_{RAS}(x)$, respectively. As in (1.5) and (1.6), we rewrite (3.4) as

$$x^{(\ell+1)} = x^{(\ell)} - A_{m_\ell}(x^{(\ell)}) F(x^{(\ell)}) = G_{m_\ell}(x^{(\ell)}), \quad \ell = 0, 1, \ldots, \tag{3.5}$$

where for each positive integer m

$$A_m(x) = (I - T(x)^m) F'(x)^{-1} \quad \text{and} \quad G_m(x) = x - A_m(x) F(x). \tag{3.6}$$

A coarse grid correction can be considered as well by adding a term of the form $R_0^T F_0'(x)^{-1} R_0 F'(x)$ in the corresponding sum in (3.1) or (3.3).

If the local linear problems (1.3) are solved approximately using the additive Schwarz (or restricted additive Schwarz) iteration with inexact local solvers, the structure of the iterations and the operators (3.1)–(3.6) is maintained, with the only difference that the local solver $F_k'(x)$ is replaced by an inexact local solver $\tilde{F}_k'(x)$. Thus,

the iteration matrices in this case are

$$\tilde{T}_\theta(x) = I - \theta \sum_{k=1}^{p} R_k^T \tilde{F}_k'(x)^{-1} R_k F'(x) \quad \text{and} \quad \tilde{T}_{RAS}(x) = I - \sum_{k=1}^{p} R_{k,0}^T \tilde{F}_k'(x)^{-1} R_k F'(x).$$

We proceed now to discuss the convergence of the Newton additive Schwarz methods. We begin with a well-known result.

LEMMA 3.1. [15] *Let $G : D \subset \mathbb{R}^m \to L(\mathbb{R}^n)$ be continuous at $x^\star \in D$, and $G(x^\star)$ be nonsingular. Then, there exist $\delta > 0$, $\beta > 0$, such that for $x \in D \cap \{x : \|x - x^\star\| \leq \delta\}$, $G(x)$ is nonsingular, $G(x)^{-1}$ is continuous at $x^\star$, and $\|G(x)^{-1}\| \leq \beta$.*

LEMMA 3.2. *Let $G : D \subset \mathbb{R}^m \to L(\mathbb{R}^n)$ be Lipschitz continuous at $x^\star \in D$, and $G(x^\star)$ be nonsingular. Then $G(x)^{-1}$ is Lipschitz continuous at $x^\star$.*

Proof. Since G is Lipschitz continuous at $x^\star \in D$ there exists ϵ such that $\|G(x) - G(x^\star)\| \leq \epsilon \|x - x^\star\|$ in a neighborhood of $x^\star$. Then, using Lemma 3.1 we obtain

$$\begin{aligned} \left\|G(x)^{-1} - G(x^\star)^{-1}\right\| &= \left\|G(x)^{-1} G(x^\star) G(x^\star)^{-1} - G(x)^{-1} G(x) G(x^\star)^{-1}\right\| \\ &\leq \left\|G(x)^{-1}\right\| \|G(x^\star) - G(x)\| \left\|G(x^\star)^{-1}\right\| \\ &\leq \epsilon\beta \left\|G(x^\star)^{-1}\right\| \|x - x^\star\|. \end{aligned}$$

∎

Our convergence results are valid in two distinct cases, when the Jacobian $F'(x)$ is symmetric positive definite, or when it is a monotone matrix. These two cases correspond to the theory described in sections 2.1 and 2.2.

THEOREM 3.3. *Suppose that F satisfies* (i)–(iii) *and $F(x^\star) = 0$. Let $F'(x)$ be symmetric positive definite (or a nonsingular M-matrix) in a neighborhood of $x^\star$ and suppose that $\theta < 1/q$. Let $\{m_\ell\}_{\ell=0}^{\infty}$ be a sequence of positive integers. Suppose that $m < +\infty$, where m is as in* (1.7)*, then there exist a neighborhood S of $x^\star$ and $0 < c < 1$, such that for $x^{(0)} \in S$, the sequence of iterates defined by the Newton additive Schwarz method* (3.5) *using $T = T_\theta$ converges to $x^\star$ and satisfies* (1.8)*.*

Proof. By Lemma 2.1 there exist two matrices $M_\theta(x)$ and $N_\theta(x)$ such that $F'(x) = M_\theta(x) - N_\theta(x)$, $M_\theta(x)$ is given in (3.2), it is nonsingular, and $T_\theta(x) = M_\theta(x)^{-1} N_\theta(x)$. If we prove that this splitting satisfies conditions (iv)–(vii), the result will follow from Theorem 1.2. Since $F'(x)$ is continuous and nonsingular at $x^\star$, from (3.2) it follows that $M_\theta(x)^{-1}$ and $M_\theta(x)$ are continuous at $x^\star$. Then condition (iv) is satisfied. Condition (v) is satisfied by Lemma 2.1 and condition (vi) by Lemma 2.3 (or by Lemma 2.6 in the M-matrix case). In order to prove condition (vii), we have to demonstrate that there exists $L_1 < +\infty$ such that for x in a neighborhood of $x^\star$, $\|M_\theta(x) - M_\theta(x^\star)\| \leq L_1 \|x - x^\star\|$. By Lemma 3.2, it is sufficient to prove that $M_\theta(x)^{-1}$ is Lipschitz continuous at $x^\star$. Then, from (3.2), it is sufficient to prove that $F_k'(x)^{-1}$ is Lipschitz continuous. Since $F_k'(x) = R_k F'(x) R_k^T$, the Lipschitz continuity of $F_k'(x)$ at $x^\star$ is obtained easily from the Lipschitz continuity of $F'(x)$ given by (iii). ∎

The following result establishes the convergence of Newton additive Schwarz with inexact solvers.

THEOREM 3.4. *Suppose that F satisfies* (i)–(iii) *and $F(x^\star) = 0$. Let $F'(x)$ be symmetric positive definite (or a nonsingular M-matrix) in a neighborhood of $x^\star$. Consider, for each x in this neighborhood, p inexact local solvers $\tilde{F}_k'(x)$, $k = 1, \ldots, p$, symmetric positive definite (or monotone matrices), approximations of $F_k'(x)$, $k = 1, \ldots, p$, such that there exists $\mu > 0$, with $F_k'(x) \preceq \mu \tilde{F}_k'(x)$, $k = 1, \ldots, p$, and suppose*

that $\theta < 1/\mu q$ (or such that $\tilde{F}'_k(x)^{-1}(\tilde{F}'_k(x) - F'_k(x)) \geq O$ for $k = 1, \ldots, p$, and suppose that $\theta < 1/q$). Let $\{m_\ell\}_{\ell=0}^{\infty}$ be a sequence of positive integers. Assume that the inexact local solvers $\tilde{F}'_k(x)$, $k = 1, \ldots, p$, are continuous and nonsingular at $x^\star$.

(1) Then there exists an open neighborhood S of $x^\star$ such that for each $x^{(0)} \in S$ and for each sequence of positive integers m_ℓ, the sequence of iterates defined by the Newton additive Schwarz method with inexact local solvers converges to $x^\star$. Moreover,

$$R_1(x^{(\ell)}) \leq \rho(\tilde{T}_\theta(x^\star))^{m'}, \quad m' = \lim_{\ell \to \infty} \inf m_\ell.$$

(2) Suppose that $\tilde{F}'_k(x)$, $k = 1, \ldots, p$, are Lipschitz continuous at $x^\star$. If $m < +\infty$, where m is as in (1.7)*, then there exist a neighborhood S of $x^\star$, and $0 < c < 1$ such that for $x^{(0)} \in S$ the sequence of iterates defined by the Newton additive Schwarz method with inexact local solvers converges to $x^\star$ and satisfies* (1.8)*.*

Proof. Using Lemma 2.4, the unique splitting induced by $\tilde{T}_\theta(x^\star)$, $F'(x^\star) = \tilde{M}_\theta(x^\star) - \tilde{N}_\theta(x^\star)$, with

$$\tilde{M}_\theta(x^\star)^{-1} = \theta \tilde{B}(x^\star) = \theta \sum_{k=1}^{p} R_k^T \tilde{F}'_k(x^\star)^{-1} R_k,$$

is P-regular, since $\theta < 1/(\mu q)$, and then $\|\tilde{T}_\theta(x^\star)\|_{F'(x^\star)} < 1$. In the monotone case, by Theorem 2.7, there exists a vector $\omega > 0$ such that $\|\tilde{T}_\theta(x^\star)\|_\omega < 1$. From the hypotheses on the inexact local solvers, we deduce that conditions (iv) and (v) are satisfied for the matrix $\tilde{M}_\theta(x)$. In the case (1), using Theorem 1.1, the result is proved. In the case (2) we only need to prove that $\tilde{M}_\theta(x)$ is Lipschitz continuous at $x^\star$ and then the result follows from Theorem 1.2. The proof of the Lipschitz continuity follows from the Lipschitz continuity of $\tilde{F}'_k(x)$, $k = 1, \ldots, p$, in a way similar to that described in Theorem 3.3. ∎

For nonsingular M-matrices, we obtain similar convergence results for RAS for exact and inexact local solvers, using essentially the same proofs as for Theorems 3.3 and 3.4. We summarize them in the following two theorems.

THEOREM 3.5. *Suppose that F satisfies* (i)–(iii) *and $F(x^\star) = 0$. Let $F'(x)$ be a nonsingular M-matrix in a neighborhood of $x^\star$. Let $\{m_\ell\}_{\ell=0}^{\infty}$ be a sequence of positive integers. Suppose that $m < +\infty$, where m is as in* (1.7)*. Then there exist a neighborhood S of $x^\star$ and $0 < c < 1$, such that for $x^{(0)} \in S$, the sequence of iterates defined by the Newton restricted additive Schwarz method* (3.5) *using $T = T_{RAS}$ converges to $x^\star$ and satisfies* (1.8)*.*

THEOREM 3.6. *Suppose that F satisfies* (i)–(iii) *and $F(x^\star) = 0$. Let $F'(x)$ be a nonsingular M-matrix in a neighborhood of $x^\star$. Consider, for each x in this neighborhood of $x^\star$, p inexact local solvers $\tilde{F}'_k(x)$, $k = 1, \ldots, p$, monotone matrices which are approximations of $F'_k(x)$, $k = 1, \ldots, p$, such that $\tilde{F}'_k(x)^{-1}(\tilde{F}'_k(x) - F'_k(x)) \geq O$ for $k = 1, \ldots, p$. Let $\{m_\ell\}_{\ell=0}^{\infty}$ be a sequence of positive integers. Suppose that the inexact local solvers $\tilde{F}'_k(x)$, $k = 1, \ldots, p$, are continuous and nonsingular at $x^\star$.*

(1) Then there exists an open neighborhood S of $x^\star$ such that for each $x^{(0)} \in S$ and for each sequence of positive integers m_ℓ, the sequence of iterates defined by the Newton restricted additive Schwarz method with inexact local solvers converges to $x^\star$. Moreover,

$$R_1(x^{(\ell)}) \leq \rho(\tilde{T}_{RAS}(x^\star))^{m'}, \quad m' = \lim_{\ell \to \infty} \inf m_\ell.$$

(2) Suppose that $\tilde{F}_k'(x)$, $k = 1, \ldots, p$, are Lipschitz continuous at $x^\star$. If $m < +\infty$, where m is as in (1.7)*, then there exist a neighborhood S of $x^\star$, and $0 < c < 1$ such that for $x^{(0)} \in S$ the sequence of iterates defined by the Newton restricted additive Schwarz method with inexact local solvers converges to $x^\star$ and satisfies* (1.8)*.*

As a review of the proofs of the results in this section indicates, the addition of a coarse grid correction can be included in each case, and the results continue to hold. We omit the details.

4. Multiplicative Schwarz Methods. The multiplicative Schwarz method to solve the linear system $Az = b$ can be expressed as

$$z^{(s+1)} = Tz^{(s)} + c, \quad s = 0, 1, \ldots, \tag{4.1}$$

with the iteration matrix

$$T = T_{MS} = (I - P_p)(I - P_{p-1}) \cdots (I - P_1) = \prod_{k=p}^{1} (I - P_k),$$

and a certain vector c. The matrices P_k are given by

$$P_k = R_k^T A_k^{-1} R_k A = R_k^T (R_k A R_k^T)^{-1} R_k A.$$

This iteration corresponds to successive subspace corrections on the subspaces $V_1, \ldots, V_p$. When multiplicative Schwarz is used as a preconditioner, the preconditioned matrix is $I - T_{MS}$. We mention that on sequential configurations, multiplicative Schwarz is faster than additive Schwarz; see, e.g., [13]. For parallel implementations additive Schwarz is usually faster, since it is easily parallelizable.

There is also a restricted multiplicative Schwarz method, which can be expressed as (4.1) with iteration matrix

$$T = T_{RMS} = (I - Q_p)(I - Q_{p-1}) \cdots (I - Q_1) = \prod_{k=p}^{1} (I - Q_k),$$

and $Q_k = R_{k,0}^T A_k^{-1} R_k A$.

In the case of inexact local solvers, the iteration matrices are given by

$$\tilde{T} = \tilde{T}_{MS} = \prod_{k=p}^{1} (I - R_k^T \tilde{A}_k^{-1} R_k A) \text{ and } \tilde{T} = \tilde{T}_{RMS} = \prod_{k=p}^{1} (I - R_{k,0}^T \tilde{A}_k^{-1} R_k A). \tag{4.2}$$

We review here convergence results for the multiplicative Schwarz method (4.1) when the coefficient matrix is either symmetric positive definite, or a nonsingular M-matrix. In the latter case, we also consider the restricted version.

THEOREM 4.1. [2] *Suppose that A is symmetric positive definite. Then the multiplicative Schwarz method* (4.1) *converges to the solution of $Az = b$ for any choice of the initial guess $z^{(0)}$. In fact, $\rho(T_{MS}) \leq \|T_{MS}\|_A < 1$. Furthermore, there exists a unique splitting $A = M_{MS} - N_{NS}$ such that $T_{MS} = M_{MS}^{-1} N_{MS}$, and this splitting is P-regular.*

THEOREM 4.2. [2], [14] *Let A be a nonsingular M-matrix. Then the multiplicative Schwarz iteration* (4.1) *with either $T = T_{MS}$ or $T = T_{RMS}$ converges to the solution*

of $Az = b$ for any choice of the initial guess $z^{(0)}$. In fact, for any $\omega = A^{-1}e > 0$ with $e > 0$, we have $\rho(T) \leq \|T\|_\omega < 1$. Furthermore, there exists a unique splitting $A = M - N$ such that $T = M^{-1}N$, and this splitting is nonnegative, i.e., $T \geq O$.

THEOREM 4.3. [2] *Suppose that A is symmetric positive definite. Then the multiplicative Schwarz iteration with iteration matrix $\tilde{T}_{MS}$ in (4.2) and with inexact local solvers $\tilde{A}_k$ such that $A_k = \tilde{A}_k - (\tilde{A}_k - A_k)$ are P-regular splittings, converges to the solution of $Az = b$ for any choice of the initial guess $z^{(0)}$. In fact, $\rho(\tilde{T}_{MS}) \leq \|\tilde{T}_{MS}\|_A < 1$. Furthermore, there exists a unique splitting $A = \tilde{M}_{MS} - \tilde{N}_{NS}$ such that $\tilde{T}_{MS} = \tilde{M}_{MS}^{-1}\tilde{N}_{MS}$, and this splitting is P-regular.*

THEOREM 4.4. [2], [14] *Let A be a nonsingular M-matrix. Then the multiplicative Schwarz iteration with either iteration matrix $\tilde{T}$ in (4.2) and with inexact local solvers $\tilde{A}_k$ such that $\tilde{A}_k \geq A_k$, converges to the solution of $Az = b$ for any choice of the initial guess $z^{(0)}$. In fact, for any $\omega = A^{-1}e > 0$ with $e > 0$, we have $\rho(\tilde{T}) \leq \|\tilde{T}\|_\omega < 1$. Furthermore, there exists a unique splitting $A = \tilde{M} - \tilde{N}$ such that $\tilde{T} = \tilde{M}^{-1}\tilde{N}$, and this splitting is nonnegative.*

5. Newton Multiplicative Schwarz Methods. In this section we analyze the iterative Newton methods obtained solving the linear systems (1.3) using the multiplicative Schwarz method (4.1). We call this combined method Newton multiplicative Schwarz. We start by giving algebraic representations of it. Under the hypotheses of Theorems 4.1 or 4.2, there exist two matrices $M_{MS}(x)$ and $N_{MS}(x)$ such that $F'(x) = M_{MS}(x) - N_{MS}(x)$, $M_{MS}(x)$ is nonsingular and $T_{MS}(x) = M_{MS}(x)^{-1}N_{MS}(x)$. By Lemma 1.3, this splitting satisfies

$$M_{MS}(x)^{-1} = (I - T_{MS}(x))F'(x)^{-1} = (I - \prod_{k=p}^{1}(I - P_k(x)))F'(x)^{-1}, \tag{5.1}$$

where

$$P_k(x) = R_k^T F_k'(x)^{-1} R_k F'(x) = R_k^T (R_k F'(x) R_k^T)^{-1} R_k F'(x). \tag{5.2}$$

Similarly, there exists a splitting $F'(x) = M_{RMS}(x) - N_{RMS}(x)$, with $T_{RMS}(x) = M_{RMS}(x)^{-1}N_{RMS}(x)$ and we have

$$M_{RMS}(x)^{-1} = (I - T_{RMS}(x))F'(x)^{-1} = (I - \prod_{k=p}^{1}(I - Q_k(x)))F'(x)^{-1},$$

with

$$Q_k(x) = R_{k,0}^T F_k'(x)^{-1} R_k F'(x) = R_{k,0}^T (R_k F'(x) R_k^T)^{-1} R_k F'(x).$$

With this notation, the Newton Multiplicative Schwarz iteration can be represented as follows

$$\begin{aligned} x^{(\ell+1)} &= x^{(\ell)} - (T(x^{(\ell)})^{m_\ell - 1} + T(x^{(\ell)})^{m_\ell - 2} + \cdots + I) M(x^{(\ell)})^{-1} F(x^{(\ell)}), \\ &\ell = 0, 1, \ldots, \end{aligned} \tag{5.3}$$

with iteration matrices $T(x) = T_{MS}(x)$ or $T(x) = T_{RMS}(x)$, and with $M(x) = M_{MS}(x)$ or $M(x) = M_{RMS}(x)$, respectively. The iteration (5.3) can also be written as (3.5)-(3.6) for these iteration matrices.

If the local linear problems (1.3) are solved approximately using the multiplicative Schwarz (or restricted multiplicative Schwarz) iteration with inexact local solvers, the structure of the iterations and the operators just defined is maintained, with the only difference that the local solver $F_k'(x)$ is replaced by an inexact local solver $\tilde{F}_k'(x)$. Thus, the iteration matrices in this case are

$$\tilde{T}_{MS}(x) = \prod_{k=p}^{1} (I - R_k^T \tilde{F}_k'(x)^{-1} R_k F'(x)) \quad \text{and} \tag{5.4}$$

$$\tilde{T}_{RMS}(x) = \prod_{k=p}^{1} (I - R_{k,0}^T \tilde{F}_k'(x)^{-1} R_k F'(x)). \tag{5.5}$$

Here we prove the convergence of Newton multiplicative Schwarz methods with exact or inexact local solvers, when the Jacobian is either positive definite or a nonsingular M-matrix. In the latter case, we also consider the restricted method. To that end, we first prove an auxiliary result on Lipschitz continuous maps.

LEMMA 5.1. *Let G, $H : D \subset \mathbb{R}^m \to L(\mathbb{R}^n)$ be Lipschitz continuous at $x^\star \in D$, then $G(x) + H(x)$ and $G(x)H(x)$ are Lipschitz continuous at $x^\star$.*

Proof. For $G(x) + H(x)$ the proof is trivial. For the case $G(x)H(x)$, since G and H are Lipschitz continuous at $x^\star$ there exist constants L_1, L_2, and K, such that in a neighborhood of $x^\star$

$$\begin{aligned} \|G(x) - G(x^\star)\| &\leq L_1 \|x - x^\star\|, \\ \|H(x) - H(x^\star)\| &\leq L_2 \|x - x^\star\|, \\ \|H(x)\| &\leq K. \end{aligned}$$

Using these expressions we obtain

$$\begin{aligned} &\|G(x)H(x) - G(x^\star)H(x^\star)\| \\ &\quad = \|G(x)H(x) - G(x^\star)H(x) + G(x^\star)H(x) - G(x^\star)H(x^\star)\| \\ &\quad \leq \|G(x)H(x) - G(x^\star)H(x)\| + \|G(x^\star)H(x) - G(x^\star)H(x^\star)\| \\ &\quad \leq \|G(x) - G(x^\star)\|\|H(x)\| + \|G(x^\star)\|\|H(x) - H(x^\star)\| \\ &\quad \leq L_1 \|x - x^\star\| K + \|G(x^\star)\| L_2 \|x - x^\star\| \\ &\quad \leq (L_1 K + \|G(x^\star)\| L_2) \|x - x^\star\|. \end{aligned}$$

∎

THEOREM 5.2. *Suppose that F satisfies* (i)-(iii) *and $F(x^\star) = 0$. Let $F'(x)$ be a symmetric positive definite matrix (or a nonsingular M-matrix) in a neighborhood of $x^\star$. Let $\{m_\ell\}_{\ell=0}^{\infty}$ be a sequence of positive integers. Suppose that $m < +\infty$, where m is as in* (1.7)*, then there exist a neighborhood S of $x^\star$, and $0 < c < 1$ such that for $x^{(0)} \in S$ the sequence of iterates defined by the Newton multiplicative Schwarz method converges to $x^\star$ and satisfies $\|x^{(\ell+1)} - x^\star\| \leq c^{m_\ell} \|x^{(\ell)} - x^\star\|$.*

Proof. By Theorem 4.1 (or by Theorem 4.2) there exist two matrices $M_{MS}(x)$ and $N_{MS}(x)$, the first of which is nonsingular, such that $F'(x) = M_{MS}(x) - N_{MS}(x)$, and $T_{MS}(x) = M_{MS}(x)^{-1} N_{MS}(x)$. Moreover, by Lemma 1.3 this splitting satisfies (5.1)-(5.2). Since $F'(x)$ is continuous and nonsingular at $x^\star$, from (5.1) it follows that $M_{MS}(x)^{-1}$ and $M_{MS}(x)$ are continuous at $x^\star$. Then condition (iv) is satisfied. Conditions (v) and (vi) are guaranteed by Theorem 4.1 (or by Theorem 4.2). In order to prove condition (vii) we have to demonstrate that there exists $L_1 < +\infty$ such that, for x in a neighborhood of $x^\star$, $\|M_{MS}(x) - M_{MS}(x^\star)\| \leq L_1 \|x - x^\star\|$. By

Lemma 3.2 it suffices to prove that $M_{MS}(x)^{-1}$ is Lipschitz continuous at $x^\star$. Using the expression (5.1) and Lemma 5.1, the result follows from the Lipschitz continuity of $F'_k(x)^{-1}$ proved in Theorem 3.3. □

THEOREM 5.3. *Suppose that F satisfies* (i)-(iii) *and $F(x^\star) = 0$. Let $F'(x)$ be symmetric positive definite (or a nonsingular M-matrix) in a neighborhood of $x^\star$. Consider, for each x in this neighborhood, p inexact local solvers $\tilde{F}'_k(x)$, $k = 1, \ldots, p$, symmetric positive definite (or monotone matrices), approximations of $F'_k(x)$, $k = 1, \ldots, p$, such that $F'_k(x) = \tilde{F}'_k(x) - (\tilde{F}'_k(x) - F'_k(x))$ are P-regular splittings for $k = 1, \ldots, p$ (or such that $\tilde{F}'_k(x) - F'_k(x) \geq O$ for $k = 1, \ldots, p$). Let $\{m_\ell\}_{\ell=0}^{\infty}$ be a sequence of positive integers. Assume that the inexact local solvers $\tilde{F}'_k(x)$, $k = 1, \ldots, p$, are continuous and nonsingular at $x^\star$.*

(1) Then there exists an open neighborhood S of $x^\star$ such that for each $x^{(0)} \in S$ and for each sequence of positive integers m_ℓ, the sequence of iterates defined by the Newton multiplicative Schwarz method with inexact local solvers converges to $x^\star$. Moreover,

$$R_1(x^{(\ell)}) \leq \rho(\tilde{T}_{MS}(x^\star))^{m'}, \quad m' = \lim_{\ell \to \infty} \inf m_\ell.$$

(2) Suppose that $\tilde{F}'_k(x)$, $k = 1, \ldots, p$, are Lipschitz continuous at $x^\star$. If $m < +\infty$, where m is as in (1.7)*, then there exist a neighborhood S of $x^\star$, and c with $\lambda < c < 1$ such that for $x^{(0)} \in S$, the sequence of iterates defined by the Newton multiplicative Schwarz method with inexact local solvers converges to $x^\star$ and satisfies* (1.8)*.*

Proof. From the hypotheses on the inexact local solvers, we deduce that conditions (iv) and (v) are satisfied for the matrix $\tilde{M}_{MS}(x)$. For the case of $F'(x)$ symmetric positive definite, using Theorem 4.3, the unique splitting induced by $\tilde{T}_{MS}(x^\star)$, $F'(x^\star) = \tilde{M}_{MS}(x^\star) - \tilde{N}_{MS}(x^\star)$ with $\tilde{T}_{MS}(x)$ as in (5.4), is P-regular and then $\|\tilde{T}_{MS}(x^\star)\|_{F'(x^\star)} < 1$. For the case of $F'(x)$ a nonsingular M-matrix, we use Theorem 4.4, and thus we have an induced nonnegative splitting and there exists a positive vector ω such that $\|\tilde{T}_{MS}(x^\star)\|_\omega < 1$. In the case (1), using Theorem 1.1 the result is proved. In the case (2) we only need to prove that $\tilde{M}_{MS}(x)$ is Lipschitz continuous at $x^\star$ and then the result follows from Theorem 1.2. The proof of the Lipschitz continuity follows from the Lipschitz continuity of $\tilde{F}'_k(x)$, $k = 1, \ldots, p$, in a way similar to that described in Theorem 3.3. □

When $F'(x)$ is a nonsingular M-matrix, convergence results similar to those in Theorems 5.2 and 5.3 can be shown for the restricted multiplicative Schwarz method using essentially the same proofs. We omit the details.

6. Numerical Experiments. Our goal in this section is to illustrate the performance of a parallel implementation of the Newton restricted additive Schwarz method. As mentioned earlier, multiplicative Schwarz is not as effective in a parallel environment. Experience has shown that RAS has better parallel performance than AS; see, e.g., [5], [7], [18], and the experiments illustrating this at the end of the section. Therefore, we mostly confine our numerical experiments to the restricted additive Schwarz case.

We take advantage of readily available software, namely the PETSc library of routines [1], where RAS is the default Schwarz method. Since coarse gird corrections are not available in PETSc, we do not include them in the experiments. In this manner, the reader can both reproduce our results, and appreciate that the methods

proposed can be implemented without extensive new programming. For these reasons we also use the PETSc default convergence tests for the linear and nonlinear solvers, and for each of them, we use the default values for the parameters. In all the experiments reported here, the methods stopped when $\|F(x^{(\ell)})\|_2 \leq 10^{-8}\|F(x^{(0)})\|_2$.

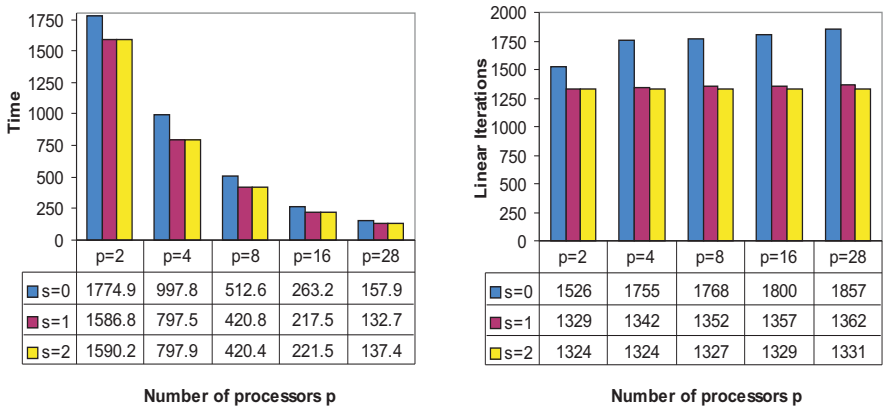
As our illustrative example, we consider a nonlinear radiative transport partial differential equation in three dimensions. The model problem is

$$-Div(\alpha \mathcal{F}^{\beta}(\nabla \mathcal{F})) = 0,$$

where we have chosen $\beta = 2.5$ and $\alpha = 1$ in our experiments. The domain is the unit cube and we consider Dirichlet boundary conditions on two opposite faces $\mathcal{F} = 1$, and Neumann boundary conditions $\partial \mathcal{F}/\partial n = 0$ on the other four faces. A finite volume approximation with the usual seven-point stencil is used to discretize the boundary value problem, obtaining a nonlinear system of equations of the form (1.1). The Jacobians in this case are nonsymmetric matrices, and we use the GMRES method for the solution of the local problems.

Our experiments are performed on a cluster of 28 nodes with two Intel Xeon processors (2.4 GHz, 1 GB DDR RAM, 512 KB L2 cache) per node connected via a Myrinet network (2.0 Gigabit/sec.). In our experiments we have used only one processor per node. The initial vector used was $x^{(0)} = (1, \ldots, 1)^T$. All times are reported in seconds.

We have obtained results for systems of different sizes and for different levels of the overlap ($s = 0, 1, 2, 3, 4, 5, \ldots$). Using an overlap of $s = 0$ results in an additive Schwarz variant that is equivalent to the block Jacobi preconditioner. In order to focus our discussion, we present here results obtained with nonlinear systems of size $n = 250047$, $n = 493039$, $n = 970299$, and $n = 1953125$, corresponding to discretizations of the cube with 65, 81, 101, and 127 points in each direction, respectively.

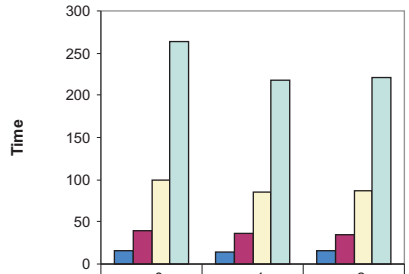


(a) Execution Time (b) Number of Linear Iterations

FIG. 6.1. *Newton Restricted Additive Schwarz, $n = 1953125$.*

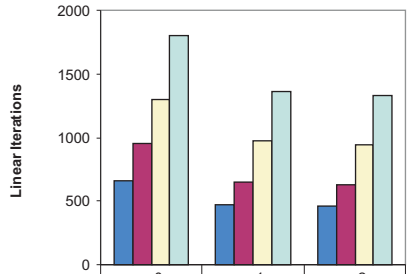
In Figure 6.1 we present the results for the nonlinear system of size $n = 1953125$. This figure illustrates, for different number of processors, the influence of the overlap ($s = 0, 1, 2$) on the execution time and on the number of linear iterations needed for convergence. In our examples, the number of nonlinear iterations needed for

convergence is the same, namely 7 iterations, for all choices of s. It can be observed that the number of linear iterations decreases as the overlap level s increases. Of course, the larger the overlap s, the larger are the local problems, and thus, their solution is expected to take longer. As a consequence, the overall computational time starts to decrease as the overlap level increases up to some "optimal" value of s after which the time increases. The optimal value of the overlap s depends on the system size and the number of processors. We also note that the number of linear iterations remains pretty constant when doubling the number of processors, and that the execution time is essentially halved. We have observed this phenomenon on other numerical experiments not reported here as well.



	s=0	s=1	s=2
n=250047	16.1	14.9	15.7
n=493039	40.2	36.4	34.3
n=970299	98.8	85.2	86.7
n=1953125	263.2	217.5	221.5

(a) Execution Time



	s=0	s=1	s=2
n=250047	660	469	456
n=493039	955	651	625
n=970299	1301	971	944
n=1953125	1800	1357	1329

(b) Number of Linear Iterations

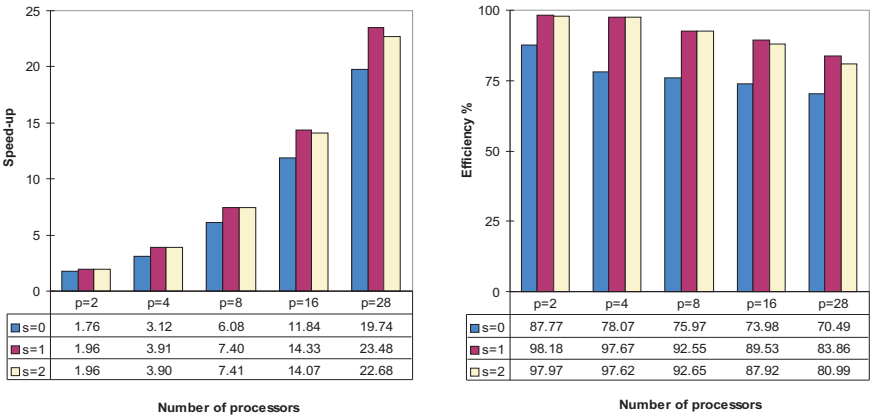
FIG. 6.2. *Newton Restricted Additive Schwarz,* $p = 16$.

We collect the computing times of the experiments with the four meshes in Figure 6.2 in which $p = 16$ processors are used. We would like to mention that the growth of number of linear iterations with the growth of the number of variables (i.e., with the decrease of the discretization parameter) is relatively slow, namely by only $30\% - 50\%$ when the number of variables increases by a factor of two. It can also be observed that doubling the number of variables may increase the total parallel computation time by a factor of about two and a half.

In Figure 6.3 we explicitly compare the Newton restricted additive Schwarz method with the well-known sequential Newton GMRES method; this is the sequential method for which we have obtained better results. If we calculate the speed-up setting the mentioned sequential method as the reference algorithm, and consider the overlap, efficiencies of about $81\% - 98\%$ are achieved depending on the number of processors used. Note that using an overlap of $s = 0$, that is, with block Jacobi preconditioner, we achieve efficiencies of only about $70\% - 88\%$, when the number of processors varies from 28 to 2.

We end this section with an example which illustrates the fact that RAS is faster than AS in a parallel environment for these nonlinear problems; see Table 6.1.

7. Conclusions. We have presented theoretical results showing the convergence of Newton Schwarz iterative methods. These are methods for the solution of nonlinear systems, where the linear step of the Newton method is solved using Schwarz itera-



(a) Speed-up (b) Efficiency

FIG. 6.3. *Newton Restricted Additive Schwarz,* $n = 1953125$.

Processors	AS time (secs.)	RAS time (secs.)
8	432.4	160.5
16	231.5	85.2

TABLE 6.1
Execution times for AS and RAS, $n = 970299$*, overlap* $s = 1$.

tions. Our convergence theory encompasses restricted additive Schwarz (RAS), which is a particularly effective method in parallel implementations. We present numerical experiments using RAS, implemented with the freely available software PETSc [1]. We observe that Newton-RAS is almost optimal in the sense that for the same differential equation, reducing the mesh size so that the number of variables increases by a factor of two, only increases the total number of linear iterations by about 20%, and the computational effort by a factor of about two and a half.

REFERENCES

[1] B. Balay, K. Buschelman, W. D. Gropp, D. Kaushik, M. Knepley, L. Curfman, B. F. Smith, and H. Zhang. PETSc Users Manual. Technical Report ANL-95/11 - Revision 2.1.5, Argonne National Laboratory, 2002.

[2] Michele Benzi, Andreas Frommer, Reinhard Nabben, and Daniel B. Szyld. Algebraic theory of multiplicative Schwarz methods. *Numerische Mathematik*, 89:605–639, 2001.

[3] Michele Benzi and Daniel B. Szyld. Existence and Uniqueness of Splittings for Stationary Iterative Methods with Applications to Alternating Methods. *Numerische Mathematik*, 76:309–321, 1997.

[4] Xiao-Chuan Cai and Maksymilian Dryja. Domain Decomposition Methods for Monotone Nonlinear Elliptic Problems. in *Contemporary Mathematics*, v. 180, David E. Keyes and Jinchao Xu, eds. pp. 21–27, AMS, Providence, R.I., 1994.

[5] Xiao-Chuan Cai, Charbel Farhat and Marcus Sarkis. A minimum overlap restricted additive Schwarz preconditioner and applications to 3D flow simulations. *Contemporary Mathematics*, 218:479–485, 1998.

[6] Xiao-Chuan Cai and David E. Keyes, Nonlinearly Preconditioned Inexact Newton Algorithms. *SIAM Journal on Scientific Computing*, 24:183–200, 2002.

[7] Xiao-Chuan Cai and Marcus Sarkis. A restricted additive Schwarz preconditioner for general sparse linear systems. *SIAM Journal on Scientific Computing*, 21:792–797, 1999.

[8] Tony Chan and Tarek Mathew. Domain decomposition algorithms, *Acta Numerica*, 3:61–144, 1994.

[9] Maksymilian Dryja and Wolfgang Hackbusch, On the nonlinear domain decomposition method, *BIT*, 37:296–311, 1997.

[10] Andreas Frommer and Daniel B. Szyld, Weighted max norms, splittings, and overlapping additive Schwarz iterations. *Numerische Mathematik*, 83:259–278, 1999.

[11] Andreas Frommer and Daniel B. Szyld. An algebraic convergence theory for restricted additive Schwarz methods using weighted max norms. *SIAM Journal on Numerical Analysis*, 39:463–479, 2001.

[12] Wolfgang Hackbusch, *Iterative Solution of Large Sparse Systems of Equations*, Springer Verlag, New York-Berlin-Heidelberg, 1994.

[13] Reinhard Nabben. Comparisons between additive and multiplicative Schwarz iterations in domain decomposition methods. *Numerische Mathematik*, 95:145–162, 2003.

[14] Reinhard Nabben and Daniel B. Szyld. Convergence theory of restricted multiplicative Schwarz methods. *SIAM Journal on Numerical Analysis*, 40:2318–2336, 2003.

[15] James M. Ortega and Warner C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, San Diego-New York-London, 1970. Reprinted by SIAM, Philadelphia, 2000.

[16] Andrew H. Sherman, On Newton-iterative methods for the solution of systems of nonlinear equations. *SIAM Journal on Numerical Analysis*, 15:755–771, 1978.

[17] Barry F. Smith, Petter E. Bjørstad, and William D. Gropp. *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations*. Cambridge University Press, Cambridge - New York - Melbourne, 1996.

[18] Xue-Cheng Tai. A space decomposition method for parabolic problems. *Numerical Methods for Partial Differential Equations*, 14:27–46, 1998.

[19] Andrea Toselli and Olof B. Widlund. *Domain Decomposition Methods - Algorithms and Theory*, volume 34 of *Springer Series in Computational Mathematics*. Springer, Berlin, Heidelberg, 2005.