BIT Numerical Mathematics
Accepted September 7, 2004, pp. 001–015
0006-3835/03/4301-0001 $16.00

ANALYSIS OF A CLASS OF PARALLEL MULTIGRID SMOOTHERS *

Dexuan Xie †

Department of Mathematical Sciences, University of Wisconsin, Milwaukee, Wisconsin, 53201-0413, USA. email: dxie@uwm.edu

Abstract.

This paper proposes and analyzes a class of multigrid smoothers called the parallel multiplicative (PM) smoother by subspace decomposition techniques. It shows that the well known additive and multiplicative smoothers and the JSOR smoother are special cases of the PM smoother, and their smoothing properties can be obtained directly from the PM analysis. Moreover, numerical results are presented in this paper to show that the JSOR smoother is more robust and effective than the damped Jacobi smoother on current MIMD parallel computers.

AMS subject classification (2000): 65N55, 65Y05.

Key words: Parallel multigrid methods, additive and multiplicative smoothers, JSOR.

1 Introduction.

The JSOR method is a family of linear stationary iterative methods defined by domain partitioning, and its convergence analysis leads to a general linear stationary iteration theory that includes the classic SOR and damped-Jacobi theories as two extreme cases [9]. One important application of the JSOR method is to define various parallel smoothers for parallel multigrid methods [2, 7, 8, 11]. Compared to the red-black (or multi-color) SOR smoother (one widely used parallel smoother), the JSOR smoother has great potential in solving unstructured finite element problems [1, 2], since it is defined on domain decomposition without using any reordering schemes. On unstructured finite element applications, it may be very difficult to construct a multi-coloring scheme since the number of required colors may increase dramatically. The JSOR smoother also can be implemented on current MIMD parallel computers as easily as the damped-Jacobi smoother (another widely used parallel smoother). Moreover, it can communicate data between processors as efficiently as the damped-Jacobi smoother. However, the JSOR smoother has not yet been analyzed mathematically.

Motivated by the study of JSOR smoothing properties, this paper proposes and analyzes a new smoother for the multigrid method by subspace decomposition techniques. The new smoother is called the parallel multiplicative (PM) smoother since various parallel versions of the sequential multiplicative smoother can be generated from it. In particular, this paper shows that the well known

*Received May 2004. Revised September 2004. Communicated by Per Lötstedt.

†This work was partially supported by the National Science Foundation through grant DMS-0241236.

additive and multiplicative smoothers [3, 4] are two extreme cases of the PM smoother, and their smoothing theorems given in [4] can be taken as the direct corollaries of the PM smoothing theorem presented in this paper. Furthermore, the JSOR smoother is shown to be a special PM smoother so that its smoothing properties can be obtained directly from the PM analysis. Finally, numerical results are presented in this paper, which show that the JSOR smoother is more robust and effective than the damped-Jacobi smoother. The numerical experiments were done on a MIMD parallel computer (the SGI Origin 2000 at the University of Wisconsin-Milwaukee) for an anisotropic model problem.

The remainder of this paper is organized as follows. Section 2 defines the PM smoother. Section 3 presents the PM analysis. Section 4 defines the JSOR smoother as a special PM smoother. Section 5 presents numerical results.

2 The PM Smoother.

Let M be a finite-dimensional space with inner product $(\cdot,\cdot)$, and A be a symmetric positive definite (SPD) operator from M to M. A smoother of the multigrid method is a linear operator R from M to M, with which a convergent iterative method for solving the linear system $Au = f$ is defined by

$$u^{(k+1)} = u^{(k)} + R(f - Au^{(k)}), \quad k = 0, 1, 2, \ldots, \tag{2.1}$$

where $u^{(0)}$ is an initial guess.

For two given positive integers p and t, the space M is partitioned into p subspace M_i, and each subspace M_i is divided into t smaller subspaces M_i^j such that

$$M = \sum_{i=1}^{p} M_i \quad \text{and} \quad M_i = \sum_{j=1}^{t} M_i^j \quad \text{for } i = 1, 2, \cdots, p, \tag{2.2}$$

where the two sums may or may not be direct, and t may have different values on different subspaces.

For $u \in M, u_i \in M_i$, and $u_i^j \in M_i^j$, the linear operators $Q_i, P_i, A_i, A_i^j, Q_i^j$, and P_i^j are defined as follows:

$$(Q_i u, v_i) = (u, v_i), \quad (AP_i u, v_i) = (Au, v_i), \quad \forall v_i \in M_i; \tag{2.3}$$

$$(A_i u_i, v_i) = (Au_i, v_i), \forall v_i \in M_i, \quad (A_i^j u_i^j, v_i^j) = (A_i u_i^j, v_i^j), \forall v_i^j \in M_i^j; \tag{2.4}$$

$$(Q_i^j u_i, v_i^j) = (u_i, v_i^j), \quad (A_i P_i^j u_i, v_i^j) = (A_i u_i, v_i^j), \quad \forall v_i^j \in M_i^j. \tag{2.5}$$

Obviously, A_i and A_i^j are SPD. From (2.3) to (2.5) it can follow that

$$A_i P_i = Q_i A \quad \text{and} \quad A_i^j P_i^j = Q_i^j A_i. \tag{2.6}$$

In terms of the above notation, the PM smoother R_{pm} is defined by

$$R_{pm} r = \sum_{i=1}^{p} v_i^{(t)}, \quad \forall r \in M, \tag{2.7}$$

where $v_i^{(t)}$ is the t-th iterate generated from the following iterative scheme

$$v_i^{(j)} = v_i^{(j-1)} + \omega(A_i^j)^{-1}Q_i^j(r_i - A_i v_i^{(j-1)}), \qquad j = 1, 2, \cdots, t. \tag{2.8}$$

Here $v_i^{(0)} = 0$, $r_i = Q_i r$, and ω is a relaxation parameter.

THEOREM 1. *Let Q_i, P_i^j, and A_i be the linear operators defined in (2.3) to (2.5). Then, the PM smoother R_{pm} has the following operator expression:*

$$R_{pm} = \sum_{i=1}^{p}(I - \mathcal{E}_i(\omega))A_i^{-1}Q_i, \tag{2.9}$$

where I denotes an identity operator, and

$$\mathcal{E}_i(\omega) = (I - \omega P_i^t)\cdots(I - \omega P_i^2)(I - \omega P_i^1). \tag{2.10}$$

PROOF. Set $v_i = A_i^{-1}r_i$. From (2.6) and (2.8) it follows the reduction formula:

$$\begin{aligned} v_i - v_i^{(j)} &= v_i - v_i^{(j-1)} - \omega(A_i^j)^{-1}Q_i^j(r_i - A_i v_i^{(j-1)}), \\ &= v_i - v_i^{(j-1)} - \omega(A_i^j)^{-1}Q_i^j A_i(v_i - v_i^{(j-1)}) \\ &= v_i - v_i^{(j-1)} - \omega P_i^j(v_i - v_i^{(j-1)}) = (I - \omega P_i^j)(v_i - v_i^{(j-1)}). \end{aligned}$$

This yields $v_i - v_i^{(t)} = \mathcal{E}_i(\omega)v_i$ with $\mathcal{E}_i(\omega)$ being defined in (2.10). Thus, for any $r \in M$,

$$\begin{aligned} R_{pm}r &= \sum_{i=1}^{p} v_i^{(t)} = \sum_{i=1}^{p} v_i - \sum_{i=1}^{p}(v_i - v_i^{(t)}) = \sum_{i=1}^{p} v_i - \sum_{i=1}^{p}\mathcal{E}_i(\omega)v_i \\ &= \sum_{i=1}^{p}(I - \mathcal{E}_i(\omega))v_i = \sum_{i=1}^{p}(I - \mathcal{E}_i(\omega))A_i^{-1}Q_i r, \end{aligned}$$

where $v_i = A_i^{-1}r_i$ and $r_i = Q_i r$ have been used. Since r is arbitrary, (2.9) follows directly from the above identity. □

Obviously, the additive smoother R_a defined in [4] and an extended multiplicative smoother R_m can be obtained from (2.9) as two extreme cases. In fact, when $t = 1$ and $p > 1$, the operator $P_i^1 = I$ and from (2.9) it follows

$$R_a = \omega \sum_{i=1}^{p} A_i^{-1}Q_i. \tag{2.11}$$

On the other hand, when $p = 1$ and $t > 1$, $M_1 = M$ so that $A_1 = A$ and $Q_1 = I$, and P_1^j can be simply denoted by P_j. Thus, R_m follows immediately from (2.9):

$$R_m = (I - E_t(\omega))A^{-1}, \tag{2.12}$$

where $E_t(\omega) = (I - \omega P_t)\cdots(I - \omega P_2)(I - \omega P_1)$. Moreover, R_m using $\omega = 1$ is the multiplicative smoother defined in [4]. Due to this reason, the smoother R_m defined in (2.12) is referred to as an extended multiplicative smoother.

3 The Smoothing Analysis.

In the analysis of multigrid methods, a smoother R is often assumed to satisfy the following two standard smoothing conditions.

(C.1) *There is a constant C_R independent of the grid level such that*

$$\frac{\|u\|^2}{\lambda} \leq C_R(\bar{R}u, u), \qquad \forall u \in M. \tag{3.1}$$

Here, $\bar{R}$ is either $(I-K^*K)A^{-1}$ or $(I-KK^*)A^{-1}$, $K=(I-RA)$, $K^* = I-R^TA$, R^T denotes the adjoint of R with respect to the inner product $(\cdot,\cdot)$, λ is the largest eigenvalue of A, and $\|u\|^2 = (u,u)$.

(C.2) *There is a constant θ independent of the grid level and less than 2 such that $T = RA$ satisfies*

$$(ATv, Tv) \leq \theta(ATv, v), \qquad \forall v \in M. \tag{3.2}$$

To verify (C.1) and (C.2), Bramble and Pasciak proposed the following two simple assumptions on subspace decomposition $M = \sum_{i=1}^{l} M_i$ (see [4]):

(A.1) *There exists a positive constant c_0 independent of any grid size such that for each $u \in M$, there is a decomposition $u = \sum_{i=1}^{l} u_i$ with $u_i \in M_i$ satisfying*

$$\sum_{i=1}^{l} \|u_i\|^2 \leq c_0 \|u\|^2. \tag{3.3}$$

(A.2) *There is a constant c_1 independent of any grid size such that*

$$\max_{1 \leq i \leq l} \sum_{j=1}^{l} \kappa_{ij} \leq c_1, \quad \textit{where} \quad \kappa_{ij} = \begin{cases} 0 & \textit{if } P_iP_j = 0, \\ 1 & \textit{otherwise} . \end{cases} \tag{3.4}$$

Since R_m is an extension of the multiplicative smoother defined in [4], it needs to be proved that R_m satisfies both (C.1) and (C.2).

THEOREM 2. *If the subspace decomposition $M = \sum_{i=1}^{t} M_i$ satisfies (A.1) and (A.2) with $l = t$, then for $\omega \in (0,2)$, (C.1) and (C.2) hold for the multiplicative smoother R_m defined in (2.12) with*

$$C_R = \frac{c_0(|1-\omega| + \omega c_1)^2}{\omega(2-\omega)}, \quad \theta = \frac{2\omega c_1}{2 + \omega(c_1 - 1)}. \tag{3.5}$$

PROOF. The multiplicative smoother R_m is first shown to satisfy (C.1).

Set $u = Av$ for $v \in M$. Since assumption (A.1) holds for the subspace decomposition $M = \sum_{i=1}^{t} M_i$, there exist $u_i \in M_i$ such that $u = \sum_{i=t}^{t} u_i$, and

$$\sum_{i=1}^{t} \|u_i\|^2 \leq c_0 \|u\|^2. \tag{3.6}$$

With (2.3), (3.6), and the Cauchy inequality,

$$\|u\|^2 = (u, u) = (u, \sum_{i=1}^{t} u_i) = \sum_{i=1}^{t}(u, u_i) = \sum_{i=1}^{t}(Q_i u, u_i)$$
$$\leq \sqrt{\sum_{i=1}^{t} \|u_i\|^2}\sqrt{\sum_{i=1}^{t} \|Q_i u\|^2} \leq c_0^{1/2}\|u\|\sqrt{\sum_{i=1}^{t} \|Q_i u\|^2},$$

which can be simplified as

$$\|u\|^2 \leq c_0 \sum_{i=1}^{t} \|Q_i u\|^2. \tag{3.7}$$

Let λ_i and λ be the largest eigenvalues of A_i and A, respectively. Clearly, $0 < \lambda_i \leq \lambda$, and $A_i = A_i^{1/2} A_i^{1/2}$ since A_i is SPD. Thus, by $u = Av$, (2.3), (2.4), and (2.6),

$$\|Q_i u\|^2 = \|Q_i A v\|^2 = \|A_i P_i v\|^2 = (A_i A_i^{1/2} P_i v, A_i^{1/2} P_i v)$$
$$\leq \lambda_i (A_i^{1/2} P_i v, A_i^{1/2} P_i v) = \lambda_i (A_i P_i v, P_i v)$$
$$\leq \lambda (A_i P_i v, P_i v) = \lambda (Av, P_i v) = \lambda (AP_i v, v).$$

Hence, (3.7) can be estimated by

$$\|u\|^2 \leq c_0 \lambda \sum_{i=1}^{t} (AP_i v, v). \tag{3.8}$$

To estimate $\sum_{i=1}^{t}(AP_i v, v)$, set $E_i = (I - \omega P_i)(I - \omega P_{i-1}) \cdots (I - \omega P_1)$ for $i = 1, 2, \cdots, t$, and $E_0 = I$. Clearly, $E_i = (I - \omega P_i)E_{i-1}$, $E_{i-1} - E_i = \omega P_i E_{i-1}$, and

$$I = E_{i-1} + \sum_{j=1}^{i-1}(E_{j-1} - E_j) = E_{i-1} + \omega \sum_{j=1}^{i-1} P_j E_{j-1}. \tag{3.9}$$

Thus, the term $(AP_i v, v)$ in (3.8) can be written as

$$(AP_i v, v) = (AP_i v, (E_{i-1} + \omega \sum_{j=1}^{i-1} P_j E_{j-1})v) = (AP_i v, E_{i-1} v) \tag{3.10}$$
$$+ \omega \sum_{j=1}^{i-1}(AP_i v, P_j E_{j-1} v) = (1 - \omega)(AP_i v, P_i E_{i-1} v) + \omega \sum_{j=1}^{i}(AP_i v, P_j E_{j-1} v).$$

By the Cauchy inequality, (3.10) and the following inequality,

$$\left(\sum_{i,j=1}^{t} |(Au_i, v_j)| \right)^2 \leq n_0^2 \sum_{i=1}^{t}(Au_i, u_i) \sum_{j=1}^{t}(Av_j, v_j), \tag{3.11}$$

which is proved in [4], the sum $\sum_{i=1}^t (AP_i v, v)$ can be estimated as below:

$$\sum_{i=1}^{t}(AP_iv,v) = (1-\omega)\sum_{i=1}^{t}(AP_iv, P_iE_{i-1}v) + \omega\sum_{i=1}^{t}\sum_{j=1}^{i}(AP_iv, P_jE_{j-1}v)$$
$$\leq |1-\omega|\sum_{i=1}^{t}(AP_iv, P_iE_{i-1}v) + \omega\sum_{i=1}^{t}\sum_{j=1}^{t}|(AP_iv, P_jE_{j-1}v)|$$
$$\leq (|1-\omega| + \omega n_0)\sqrt{\sum_{i=1}^{t}(AP_iv,v)}\sqrt{\sum_{i=1}^{t}(AP_iE_{i-1}v, P_iE_{i-1}v)}.$$

Simplifying the above inequality yields

$$\sum_{i=1}^{t}(AP_iv,v) \leq (|1-\omega| + \omega c_1)^2\sum_{i=1}^{t}(AP_iE_{i-1}v, P_iE_{i-1}v), \tag{3.12}$$

where $n_0 \leq c_1$ has been used. Next, it is clear that

$$\begin{aligned}
&(AE_{i-1}v, E_{i-1}v) - (AE_iv, E_iv)\\
&= (AE_{i-1}v, E_{i-1}v) - (AE_iv, E_{i-1}v) + (AE_iv, E_{i-1}v) - (AE_iv, E_iv)\\
&= (A(E_{i-1} - E_i)v, E_{i-1}v) + (AE_iv, (E_{i-1} - E_i)v)\\
&= \omega(AP_iE_{i-1}v, E_{i-1}v) + \omega(AE_iv, P_iE_{i-1}v)\\
&= \omega(AP_iE_{i-1}v, (E_{i-1} + E_i)v) = \omega(AP_iE_{i-1}v, (2I - \omega P_i)E_{i-1}v)\\
&= \omega[2(AP_iE_{i-1}v, E_{i-1}v) - \omega(AP_iE_{i-1}v, P_iE_{i-1}v)]\\
&= \omega[2(AP_iE_{i-1}v, E_{i-1}v) - \omega(AP_iE_{i-1}v, E_{i-1}v)]\\
&= \omega(2-\omega)(AP_iE_{i-1}v, E_{i-1}v),
\end{aligned}$$

and $\sum_{i=1}^t [(AE_{i-1}v, E_{i-1}v) - (AE_iv, E_iv)] = (Av, v) - (AE_tv, E_tv)$. Thus, the sum $\sum_{i=1}^t (AP_iE_{i-1}v, P_iE_{i-1}v)$ in (3.12) can be expressed in the form

$$\sum_{i=1}^{t}(AP_iE_{i-1}v, E_{i-1}v) = \frac{1}{\omega(2-\omega)}[(Av,v) - (AE_tv, E_tv)]. \tag{3.13}$$

By $u = Av$ and $E_t = K_m$, where $K_m = I - R_mA$,

$$\begin{aligned}
&(Av,v) - (AE_tv, E_tv) = (u, A^{-1}u) - (AK_mv, K_mv)\\
&= (u, A^{-1}u) - (Av, K_m^*K_mv) = (u, A^{-1}u) - (u, K_m^*K_mA^{-1}u)\\
&= ((I - K_m^*K_m)A^{-1}u, u).
\end{aligned} \tag{3.14}$$

Substituting (3.13) and (3.14) into (3.12) yields

$$\sum_{i=1}^{t}(AP_iv,v) \leq \frac{(|1-\omega| + \omega c_1)^2}{(2-\omega)\omega}[(I - K_m^*K_m)A^{-1}u, u)]. \tag{3.15}$$

Applying (3.15) to (3.8) produces

$$\|u\|^2 \le c_0\lambda \frac{(|1-\omega|+\omega c_1)^2}{(2-\omega)\omega}((I-K_m^*K_m)A^{-1}u,u).$$

This proves that R_m satisfies (C.1) with C_R given in (3.5).

Next, the multiplicative smoother R_m is shown to satisfy (C.2).

Since $T = R_mA$, $K_m = I - R_mA$, and $K_m = E_t$, T can be expressed as $T = I - E_t$. Thus, (3.9) follows $T = \omega\sum_{i=1}^t P_iE_{i-1}$ so that

$$(ATu,Tu) = \omega^2 \sum_{i,j=1}^{t} (AP_iE_{i-1}v, P_jE_{j-1}v).$$

Hence, by (2.3), (3.4), (3.11), and (3.13),

$$\begin{aligned}(ATu,Tu) &\le \omega^2 c_1 \sum_{i=1}^{t}(AP_iE_{i-1}v, P_iE_{i-1}v) = \frac{\omega c_1}{2-\omega}[(Au,u)-(AE_tu,E_tu)]\\ &= \frac{\omega c_1}{2-\omega}[(Au,u)-(A(I-T)u,(I-T)u)] = \frac{\omega c_1}{2-\omega}[2(ATu,u)-(ATu,Tu)].\end{aligned}$$

Solving the above inequality for (ATu,Tu) gives

$$(ATu,Tu) \le \frac{2\omega c_1}{2+\omega(c_1-1)}(ATu,u).$$

This completes the proof that (3.2) holds for R_m with θ given in (3.5). □

From Theorem 2 it immediately follows Theorem 3.2 in [4]. That is, the multiplicative smoother defined in [4] (i.e., R_m with $\omega = 1$) satisfies (C.1) and (C.2) with $C_R = c_0c_1^2$ and $\theta = 2c_1/(c_1+1)$.

To prove that the PM smoother R_{pm} satisfies both (C.1) and (C.2), two lemmas are introduced in the following.

LEMMA 1. *Let* $\bar{R} = (I-K^*K)A^{-1}$, *or* $\bar{R} = (I-KK^*)A^{-1}$, *where* $K^* = I - R^TA$, *and* $K = I - RA$. *If (3.2) holds with* $\theta \in (0,2)$, *then*

$$(\bar{R}u,u) \ge (2-\theta)(Ru,u), \qquad \forall u \in M. \tag{3.16}$$

PROOF. The inequality (3.16) is shown in the proof of Theorem 3.1 in [4]. It can also be proved as given below.

Set $\bar{R} = (I-K^*K)A^{-1}$. With the expressions of K^* and K, $\bar{R}$ can be rewritten as $\bar{R} = R^T + R - R^TAR$ such that

$$\begin{aligned}(\bar{R}u,u) &= ((R^T+R-R^TAR)u,u) = (R^Tu,u)+(Ru,u)-(R^TARu,u)\\ &= (u,Ru)+(Ru,u)-(ARu,Ru) = 2(Ru,u)-(ARu,Ru).\end{aligned} \tag{3.17}$$

By (3.2), $u = Av$, and $T = RA$,

$$\begin{aligned}(ARu,Ru) &= (ARAv,RAv) = (ATv,Tv) \le \theta(ATv,v)\\ &= \theta(Tv,Av) = \theta(RAv,Av) = \theta(Ru,u).\end{aligned}$$

Hence, $(\bar{R}u, u) \geq 2(Ru, u) - \theta(Ru, u) = (2-\theta)(Ru, u)$. This completes the proof, since the proof for $\bar{R} = (I - KK^*)A^{-1}$ can be done similarly. ◻

LEMMA 2. *Let R_{pm} and $\mathcal{E}_i$ be defined in (2.9) and (2.10), respectively. Then*

$$\mathcal{E}_i^* = (I - \omega P_i^1)(I - \omega P_i^2) \cdots (I - \omega P_i^t), \tag{3.18}$$

and

$$R_{pm}^T = \sum_{i=1}^{p} (I - \mathcal{E}_i^*) A_i^{-1} Q_i. \tag{3.19}$$

PROOF. From (2.5) it follows that

$$(A_i P_i^j u_i, v_i) = (A_i u_i, P_i^j v_i), \quad \forall u_i, v_i \in M_i,$$

and thus $(P_i^j)^* = P_i^j$. Hence, $(I - \omega P_i^j)^* = (I - \omega P_i^j)$, and (3.18) follows immediately.

Since A_i is SPD, it is clear that for $u, v \in M$, there exist u_i and v_i such that $Q_i u = A_i u_i$ and $Q_i v = A_i v_i$. Thus, with (2.3) and (3.18),

$$\begin{aligned} (\mathcal{E}_i A_i^{-1} Q_i u, Q_i v) &= (\mathcal{E}_i u_i, A_i v_i) = (A_i \mathcal{E}_i u_i, v_i) \\ &= (A_i u_i, \mathcal{E}_i^* v_i) = (Q_i u, \mathcal{E}_i^* A_i^{-1} Q_i v) = (u, \mathcal{E}_i^* A_i^{-1} Q_i v). \end{aligned}$$

Hence, with (2.9) and the above identity,

$$\begin{aligned} (R_{pm} u, v) &= (\sum_{i=1}^{p} (I - \mathcal{E}_i) A_i^{-1} Q_i u, v) = \sum_{i=1}^{p} ((I - \mathcal{E}_i) A_i^{-1} Q_i u, v) \\ &= \sum_{i=1}^{p} [(A_i^{-1} Q_i u, v) - (\mathcal{E}_i A_i^{-1} Q_i u, v)] = \sum_{i=1}^{p} [(A_i^{-1} Q_i u, Q_i v) - (\mathcal{E}_i A_i^{-1} Q_i u, Q_i v)] \\ &= \sum_{i=1}^{p} [(u, A_i^{-1} Q_i v) - (u, \mathcal{E}_i^* A_i^{-1} Q_i v)] = \sum_{i=1}^{p} (u, (I - \mathcal{E}_i^*) A_i^{-1} Q_i v) \\ &= (u, \sum_{i=1}^{p} (I - \mathcal{E}_i^*) A_i^{-1} Q_i v), \quad \forall u, v \in M. \end{aligned}$$

This proves expression (3.19). ◻

The PM smoothing theorem is now presented below.

THEOREM 3. *If the two subspace decompositions in (2.2) satisfy (A.1) and (A.2) with $l = p, t$, $c_0 = c_{0,a}, c_{0,m}$ and $c_1 = c_{1,a}, c_{1,m}$, respectively, then (C.1) and (C.2) hold for the PM smoother R_{pm} defined in (2.9) with*

$$C_R = \frac{c_{0,a} c_{0,m} (2c_{1,a} - \theta)(|1 - \omega| + \omega c_{1,m})^2}{c_{1,a} \omega (2 - \theta)(2 - \omega)}, \quad \theta = \frac{2\omega c_{1,a} c_{1,m}}{2 + \omega(c_{1,m} - 1)}, \tag{3.20}$$

PROOF. The proof is presented in two parts. In Part 1, the PM smoother R_{pm} is shown to satisfy (C.2), while in Part 2, R_{pm} is shown to satisfy (C.1).

Part 1: Prove that R_{pm} satisfies (C.2).

Set $T = R_{pm}A$ and $u = Tv$ for $v \in M$. By (2.6), (2.9) and the Cauchy inequality,

$$(Au, u) = (AR_{pm}Av, u) = \left(A\sum_{i=1}^{p}(I - \mathcal{E}_i)A_i^{-1}Q_iAv, u \right)$$
$$= \left(A\sum_{i=1}^{p}(I - \mathcal{E}_i)P_iv, u \right) \le \sqrt{\left(A\sum_{i=1}^{p}(I - \mathcal{E}_i)P_iv, \sum_{i=1}^{p}(I - \mathcal{E}_i)P_iv \right)}\sqrt{(Au, u)},$$

which can be simplified as

$$(ATv, Tv) \le \left(A\sum_{i=1}^{p}T_iv_i, \sum_{i=1}^{p}T_iv_i \right), \tag{3.21}$$

where $T_i = I - \mathcal{E}_i$ and $v_i = P_iv$. By (2.4) and (3.11), it is clear that

$$\left(A\sum_{i=1}^{p}T_iv_i, \sum_{i=1}^{p}T_iv_i \right) = \sum_{i,j=1}^{p}(AT_iv_i, T_jv_j) \le c_{1,a}\sum_{i=1}^{p}(A_iT_iv_i, T_iv_i). \tag{3.22}$$

Applying Theorem 2 to subspace M_i produces

$$(A_iT_iv_i, T_iv_i) \le \frac{2\omega c_{1,m}}{2 + \omega(c_{1,m} - 1)}(A_iT_iv_i, v_i), \quad i = 1, 2, \cdots, p. \tag{3.23}$$

By (2.3), (2.4), (2.6), $v_i = P_iv$ and $T_i = I - \mathcal{E}_i$, the term $(A_iT_iv_i, v_i)$ is rewritten as

$$(A_iT_iv_i, v_i) = (AT_iv_i, v_i) = (T_iv_i, Av_i) = (T_iv_i, AP_iv) = (T_iv_i, Av)$$
$$= (T_iP_iv, Av) = ((I - \mathcal{E}_i)A_i^{-1}A_iP_iv, Av) = ((I - \mathcal{E}_i)A_i^{-1}Q_iAv, Av).$$

Adding the above terms from $i = 1$ to $i = p$ gives

$$\sum_{i=1}^{p}(A_iT_iv_i, v_i) = \sum_{i=1}^{p}((I - \mathcal{E}_i)A_i^{-1}Q_iAv, Av) \tag{3.24}$$
$$= \left(\sum_{i=1}^{p}(I - \mathcal{E}_i)A_i^{-1}Q_iAv, Av \right) = (R_{pm}Av, Av) = (ATv, v).$$

Here $T = R_{pm}A$. Thus, a combination of (3.22), (3.23), and (3.24) with (3.21) yields

$$(ATv, Tv) \le c_{1,a}\sum_{i=1}^{p}(A_iT_iv_i, T_iv_i) \le c_{1,a}\frac{2\omega c_{1,m}}{2 + \omega(c_{1,m} - 1)}(ATv, v).$$

This completes the proof of Part 1.

Part 2: Prove that R_{pm} satisfies (C.1).

According to decomposition (2.2), for each $u \in M$, there exist $u_i \in M_i$ such that $u = \sum_{i=1}^{l} u_i$. Applying Theorem 2 to each $u_i \in M_i$ gives

$$\frac{\|u_i\|^2}{\lambda_i} \le c_{0,m} \frac{(|1-\omega| + \omega c_{1,m})^2}{(2-\omega)\omega} \left((I - \mathcal{E}_i^* \mathcal{E}_i) A_i^{-1} u_i, u_i\right), \tag{3.25}$$

where $i = 1, 2, \cdots, p$, λ_i is the largest eigenvalue of A_i, and $\mathcal{E}_i$ is given in (2.10).

Set $v_i = A_i^{-1} u_i$ and $u_i = Q_i u$. By (3.7), (3.25), and $\lambda_i \le \lambda$ for all $i = 1, 2, \cdots, p$, it is easy to obtain

$$\begin{aligned}
\|u\|^2 &\le c_{0,a} \sum_{i=1}^{p} \|Q_i u\|^2 \le c_{0,a} \sum_{i=1}^{p} c_{0,m} \frac{(|1-\omega| + \omega c_{1,m})^2}{(2-\omega)\omega} \lambda_i ((I - \mathcal{E}_i^* \mathcal{E}_i) A_i^{-1} u_i, u_i) \\
&\le c_{0,a} c_{0,m} \lambda \frac{(|1-\omega| + \omega c_{1,m})^2}{(2-\omega)\omega} \sum_{i=1}^{p} [(A_i^{-1} u_i, u_i) - (\mathcal{E}_i^* \mathcal{E}_i A_i^{-1} u_i, u_i)] \\
&= c_{0,a} c_{0,m} \lambda \frac{(|1-\omega| + \omega c_{1,m})^2}{(2-\omega)\omega} \sum_{i=1}^{p} [(A_i^{-1} u_i, u_i) - (\mathcal{E}_i^* \mathcal{E}_i v_i, A_i v_i)] \\
&= c_{0,a} c_{0,m} \lambda \frac{(|1-\omega| + \omega c_{1,m})^2}{(2-\omega)\omega} \sum_{i=1}^{p} [(A_i^{-1} u_i, u_i) - (A_i \mathcal{E}_i^* \mathcal{E}_i v_i, v_i)] \qquad (3.26) \\
&= c_{0,a} c_{0,m} \lambda \frac{(|1-\omega| + \omega c_{1,m})^2}{(2-\omega)\omega} \sum_{i=1}^{p} [(A_i^{-1} u_i, u_i) - (A_i \mathcal{E}_i v_i, \mathcal{E}_i v_i)].
\end{aligned}$$

With (2.3), $v_i = A_i^{-1} u_i$ and $u_i = Q_i u$,

$$\begin{aligned}
&(A_i^{-1} u_i, u_i) - (A_i \mathcal{E}_i v_i, \mathcal{E}_i v_i) = (A_i^{-1} u_i, Q_i u) - (A_i \mathcal{E}_i A_i^{-1} u_i, \mathcal{E}_i A_i^{-1} u_i) \\
&= (A_i^{-1} u_i, u) - (\mathcal{E}_i A_i^{-1} u_i, u) + (\mathcal{E}_i A_i^{-1} u_i, u) - (A_i \mathcal{E}_i A_i^{-1} u_i, \mathcal{E}_i A_i^{-1} u_i) \\
&= ((I - \mathcal{E}_i) A_i^{-1} u_i, u) + (\mathcal{E}_i A_i^{-1} u_i, (I - A_i \mathcal{E}_i A_i^{-1}) u_i) \\
&= ((I - \mathcal{E}_i) A_i^{-1} Q_i u, u) + (\mathcal{E}_i A_i^{-1} u_i, A_i (I - \mathcal{E}_i) A_i^{-1} u_i) \\
&= ((I - \mathcal{E}_i) A_i^{-1} Q_i u, u) + ((I - \mathcal{E}_i) A_i^{-1} u_i, A_i \mathcal{E}_i A_i^{-1} u_i - u_i + u_i) \\
&= ((I - \mathcal{E}_i) A_i^{-1} Q_i u, u) + ((I - \mathcal{E}_i) A_i^{-1} u_i, u_i) - ((I - \mathcal{E}_i) A_i^{-1} u_i, A_i (I - \mathcal{E}_i) A_i^{-1} u_i) \\
&= 2((I - \mathcal{E}_i) A_i^{-1} Q_i u, u) - ((I - \mathcal{E}_i) A_i^{-1} u_i, A_i (I - \mathcal{E}_i) A_i^{-1} u_i).
\end{aligned}$$

Summation over i gives

$$\begin{aligned}
&\sum_{i=1}^{p} [(A_i^{-1} u_i, u_i) - (A_i \mathcal{E}_i v_i, \mathcal{E}_i v_i)] \\
&= \sum_{i=1}^{p} 2((I - \mathcal{E}_i) A_i^{-1} Q_i u, u) - \sum_{i=1}^{p} ((I - \mathcal{E}_i) A_i^{-1} u_i, A_i (I - \mathcal{E}_i) A_i^{-1} u_i).
\end{aligned}$$

By the expression of R_{pm}, the first sum of the above identity can be written as

$$\sum_{i=1}^{p} 2((I - \mathcal{E}_i) A_i^{-1} Q_i u, u) = 2(\sum_{i=1}^{p} (I - \mathcal{E}_i) A_i^{-1} Q_i u, u) = 2(R_{pm} u, u).$$

As an application of (3.11), the second sum can be estimated as

$$\begin{aligned}&\sum_{i=1}^{p}((I-\mathcal{E}_i)A_i^{-1}u_i, A_i(I-\mathcal{E}_i)A_i^{-1}u_i)\\ &\geq c_{1,a}^{-1}(A\sum_{i=1}^{p}(I-\mathcal{E}_i)A_i^{-1}u_i, \sum_{i=1}^{p}(I-\mathcal{E}_i)A_i^{-1}u_i) = c_{1,a}^{-1}(AR_{pm}u, R_{pm}u).\end{aligned}$$

Thus, by Lemma 1, Theorem 2, and (3.17),

$$\begin{aligned}&\sum_{i=1}^{p}[(A_i^{-1}u_i, u_i) - (A_i\mathcal{E}_iv_i, \mathcal{E}_iv_i)] \leq 2(R_{pm}u, u) - c_{1,a}^{-1}(AR_{pm}u, R_{pm}u)\\ &\leq 2(1-c_{1,a}^{-1})(R_{pm}u, u) + c_{1,a}^{-1}[2(R_{pm}u, u) - (AR_{pm}u, R_{pm}u)]\\ &\leq \frac{2(1-c_{1,a}^{-1})}{2-\theta}(\bar{R}_{pm}u, u) + c_{1,a}^{-1}(\bar{R}_{pm}u, u) = \left[\frac{2(1-c_{1,a}^{-1})}{2-\theta} + c_{1,a}^{-1}\right](\bar{R}_{pm}u, u).\end{aligned}$$

Applying the above estimation to (3.26) gives

$$\|u\|^2 \leq c_{0,a}c_{0,m}\lambda\frac{(|1-\omega|+\omega c_{1,m})^2}{(2-\omega)\omega}\left[\frac{2(1-c_{1,a}^{-1})}{2-\theta} + c_{1,a}^{-1}\right](\bar{R}_{pm}u, u).$$

Hence, C_R is obtained as below:

$$C_R = c_{0,a}c_{0,m}\lambda\frac{(|1-\omega|+\omega c_{1,m})^2}{(2-\omega)\omega}\left[\frac{2(1-c_{1,a}^{-1})}{2-\theta} + c_{1,a}^{-1}\right],$$

which can be simplified as

$$C_R = \frac{c_{0,a}c_{0,m}(2c_{1,a}-\theta)(|1-\omega|+\omega c_{1,m})^2}{c_{1,a}\omega(2-\theta)(2-\omega)}.$$

This completes the proof that R_{pm} satisfies (C.1). Note that we only gave the proof for $\bar{R}_{pm} = (I-K^*K)A^{-1}$. The proof for $\bar{R}_{pm} = (I-KK^*)A^{-1}$ is similar. This completes the proof of Theorem 3. ☐

From Theorem 3, Theorem 2 and Theorem 3.1 in [4] follow immediately. In fact, when $p = 1$ and $t > 1$, R_{pm} is reduced to R_m with $c_{1,a} = c_{0,a} = 1$, $c_{0,m} = c_0$, and $c_{1,m} = c_1$. Thus, (3.20) is simplified as (3.5) so that Theorem 2 is followed from Theorem 3. Also, when $t = 1$ and $p > 1$, R_{pm} becomes R_a with $c_{1,m} = c_{0,m} = 1$, $c_{0,a} = c_0$, and $c_{1,a} = c_1$ such that (3.20) is simplified with $C_R = c_0c_1/[\theta(2-\theta)]$ and $\theta = \omega c_1$, which yields Theorem 3.1 in [4]. Note that Theorem 2 has been used in the proof of Theorem 3.

According to the multigrid theory [3, 4], a multigrid algorithm is convergent if it satisfies the "smoothing assumption" (C.1) and the following "regularity and approximation assumption":

$$|(\mathcal{A}_k(I-\mathcal{P}_{k-1})u, u)_k| \leq C_\alpha^2\left(\frac{\|\mathcal{A}_ku\|_k^2}{\lambda_k}\right)^\alpha(\mathcal{A}_ku, u)_k^{1-\alpha}, \quad \forall u \in \mathcal{M}_k, \tag{3.27}$$

where $k = 2, 3, \cdots, m$, $\alpha \in (0, 1]$, C_α is a positive constant independent of the coarse grid operator $\mathcal{A}_k$, and $\{\mathcal{M}_k\}_{k=1}^m$ denotes a sequence of finite-dimensional inner product spaces on which the multigrid algorithm is defined. Here the sequence $\{\mathcal{M}_k\}_{k=1}^m$ is assumed to satisfy that

$$\mathcal{M}_1 \subset \mathcal{M}_2 \subset \cdots \subset \mathcal{M}_m,$$

and the linear operators $\mathcal{P}_k$ and $\mathcal{A}_k$ are defined by

$$(\mathcal{A}\mathcal{P}_k u, v_k) = (\mathcal{A}u, v_k), \quad (\mathcal{A}_k u_k, v_k) = (\mathcal{A}u_k, v_k), \quad \forall v_k \in \mathcal{M}_k,$$

where $u \in \mathcal{M}_m$, $u_k \in \mathcal{M}_k$, and $\mathcal{A}$ denotes the SPD operator from $\mathcal{M}_m$ to $\mathcal{M}_m$.

It is clear that on each space $\mathcal{M}_k$ for $k = 2, \ldots, m$, the PM smoother R_{pm} can be defined and satisfies the smoothing assumption (C.1). Thus, if the regularity and approximation assumption (3.27) holds, the multigrid method using the PM smoother can be proved to be convergent by using the same arguments as the ones given in the proof of Theorem 1 in [5].

4 The JSOR Smoother as a Special PM Smoother.

Let e_i denote the i-th column of the identity matrix, and A be an $n \times n$ SPD matrix. Set $M = \text{Span}\{e_1, e_2, \ldots, e_n\}$ and $t = n/p$, where p is a positive integer such that t is a positive integer. The subspaces of (2.2) are constructed as below:

$$M_i = \text{span}\{e_{l_{i-1}+1}, e_{l_{i-1}+2}, \cdots, e_{l_i}\} \quad \text{and} \quad M_i^j = \text{span}\{e_{l_{i-1}+j}\}, \tag{4.1}$$

where $l_0 = 0, l_i = it$, $i = 1, 2, \ldots, p$, and $j = 1, 2, \ldots, t$.

Based on the subspaces $\{M_i\}_{i=1}^p$, A is partitioned into a $p \times p$ block matrix form, $A = (A_{ij})_{p \times p}$, and a vector r of M is partitioned into a block vector form, $r = (\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_p)^T$. Here $\mathbf{r}_i$ denotes the sub-vector of r on M_i. Thus, A_{ii} is the matrix form of the operator A_i, and $r_i = Q_i r = (0, 0, \cdots, 0, \mathbf{r}_i, 0, \cdots, 0)^T$. Furthermore, if $\mathbf{v}_i$ denotes the sub-vector of v on M_i, then the t-th iterate $v_i^{(t)}$ has the block form $v_i^{(t)} = (0, \cdots, 0, \mathbf{v}_i^{(t)}, 0, \cdots, 0)^T$. Here T denotes the transpose of a vector.

Let D_i, L_i and U_i be the main diagonal, strictly lower and strictly upper triangular matrices, respectively, satisfying $A_{ii} = D_i + L_i + U_i$. Then, $\mathbf{v}_i^{(t)}$ is expressed as $\mathbf{v}_i^{(t)} = \omega(D_i + \omega L_i)^{-1}\mathbf{r}_i$. Hence, the JSOR smoother R_{JSOR} follows from (2.7):

$$\begin{aligned} R_{JSOR} r = \sum_{i=1}^p v_i^{(t)} &= \omega((D_1 + \omega L_1)^{-1}\mathbf{r}_1, \cdots, (D_p + \omega L_p)^{-1}\mathbf{r}_p) \\ &= \omega(D + \omega B)^{-1} r, \end{aligned}$$

or equivalently, $R_{JSOR} = \omega(D + \omega B)^{-1}$, where D and B are two block diagonal matrices defined by $D = \text{diag}(D_1, D_2, \ldots, D_p)$ and $B = \text{diag}(L_1, L_2, \ldots, L_p)$. The abbreviated form $\text{diag}(\alpha_1, \alpha_2, \ldots, \alpha_m)$ denotes the diagonal matrix with the diagonal elements $\alpha_1, \alpha_2, \ldots, \alpha_m$. Clearly, D is the diagonal matrix of A.

The JSOR smoother is a family of smoothers, whose expression varies with p. With $p > 1$, the JSOR smoother can be implemented in parallel on p processors

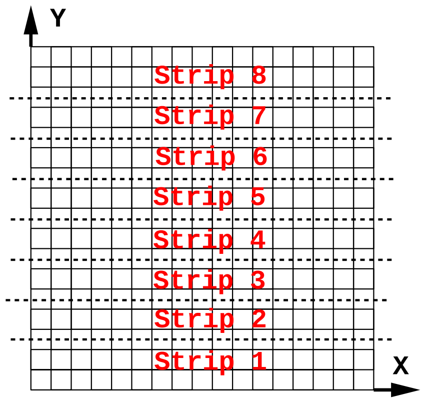

Figure 5.1: The 8-strips partition of the coarsest grid with $h = 1/16$. Here each strip contains two mesh lines. Each strip is assigned to one processor, along with its two neighboring mesh lines.

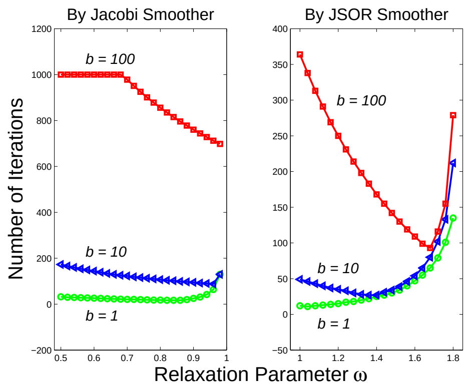

Figure 5.2: Comparison of the convergence of the parallel multigrid method using the JSOR smoother with that using the damped Jacobi smoother for solving the model problem (5.1) with $h = 1/256$ and $b = 1, 10, 100$.

if subspace M_i is assigned to processor i for $i = 1, 2, \ldots, p$. On a current MIMD parallel computer, each subspace M_i should be large enough (i.e., $p << n$) in order to overcome the overhead of interprocessor data communications. In the two extreme cases, $p = 1$ and $p = n$, the JSOR smoother is reduced to the familiar damped-Jacobi and SOR smoothers, respectively.

It is easy to show that the decompositions with subspaces given in (4.1) satisfy (A.1) and (A.2) with $c_{0,a} = c_{0,m} = c_{1,a} = c_{1,m} = 1$. In fact, for $r \in M$ and $r_i \in M_i$, it it clear that $\sum_{i=1}^{p} \|r_i\|^2 = \|r\|^2$ and the "interaction index" κ_{ij} defined in (3.4) is 1 if and only if $i = j$. Hence, from Theorem 3 it follows that the JSOR smoother satisfies both (C.1) and (C.2).

5 Numerical Results.

Numerical experiments were conducted on the multigrid method for solving the linear system $A_h u_h = f_h$ arising from a 5-point finite difference approximation to the following anisotropic model problem:

$$\begin{cases} -(u_{xx} + bu_{yy}) = f, & \text{in } \Omega = (0,1)^2, \\ u = 0, & \text{on } \partial\Omega. \end{cases} \tag{5.1}$$

In the tests, $f(x, y) = 2\pi^2 \sin \pi x \sin \pi y$, and $b = 1, 10$, and 100. The multigrid algorithm was implemented in parallel on eight R12000 400 MHz processors of the SGI Origin 2000 at the University of Wisconsin-Milwaukee based on an 8-strips partition. To avoid idle processors on the coarse grid levels of the multigrid algorithm, the U-cycle approach was employed in the parallel implementation [11]. In these tests, the coarsest grid linear system was solved by the PSOR method (a parallel SOR method by mesh partitioning, which has the same convergence rate as the sequential SOR method) [10]. Both the PSOR method and

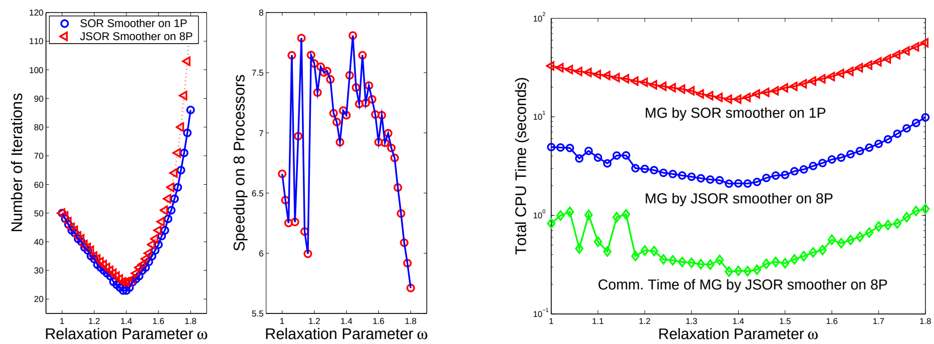

Figure 5.3: Comparison of the convergence and performance of the parallel multigrid method using the JSOR smoother with that of the sequential multigrid method using the SOR smoother for model problem (5.1) with $h = 1/1024$ and $b = 10$.

the JSOR smoother were defined on the 8-strips partitioning. Since the PSOR method requires at least two mesh lines per strip, the coarsest grid size was set as 1/16. The relaxation parameter ω was simply set to 1.7 for the PSOR method in all tests based on some numerical experiments. Other coarse grid sizes h_k were set to $1/2^k$ for $k = 4, 5, \cdots, m$ with the finest grid size $h_m = h$. The 8-strips partition of the coarsest grid is illustrated in Figure 5.1. Each strip of each grid level is assigned to one processor along with the two neighboring mesh lines. An initial guess was set to be zero for all the tests. The multigrid iterative sequence $\{u_h^{(j)}\}$ is accepted as converged when it satisfies the stopping rule $\|f_h - A_h u_h^{(j)}\|_2 / \|f_h\|_2 \leq 10^{-9}$, where $\|\cdot\|_2$ is a L_2 norm.

The numerical results were reported in Figures 5.2 to 5.3. From Figure 5.2 it can be seen that the multigrid method using the JSOR smoother has a much faster convergence rate (about 3 to 8 times faster) than using the damped Jacobi smoother for an anisotropic ratio b/a between 1 to 100. This shows that the JSOR smoother is more effective than the damped Jacobi smoother. Furthermore, Figure 5.3 shows that the convergence rate of the parallel multigrid method using the JSOR smoother can be close to that of the sequential multigrid method using the SOR smoother; the cost of interprocessor data communications for the multigrid method using the JSOR smoother is very small compared to the total CPU time cost. As displayed in the middle plot of Figure 5.3, the parallel multigrid method using the JSOR smoother had speedups ranging from 5.7 to 7.8 for the model problem (5.1) with $h = 1/1024$ on 8 processors of the SGI Origin 2000.

Note that the convergence rate of the multigrid method for solving the anisotropic problem can be further improved significantly by using semi-coarsening techniques [6]. Also, numerical results on the parallel performance of the parallel multigrid method using the JSOR smoother for solving various unstructured finite element applications (such as thin-body elasticity and eddy current approximations to Maxwell's equations) on up to 400 processors can be found in [2]. JSOR is referred to as Processor Block SOR in [2].

Acknowledgment.

The author would like to thank Professor L. Ridgway Scott for valuable discussions. He is also grateful to the anonymous referees for helpful comments.

REFERENCES

1. M. F. Adams, *Parallel multigrid solvers for 3d unstructured finite element problems in large deformation elasticity and plasticity,* International Journal for Numerical Methods in Engineering, 48 (2000), pp. 1241-1262.
2. M. F. Adams, M. Brezina, J. J. Hu, and R. S. Tuminaro, *Parallel multigrid smoothing: polynomial versus Gauss-Seidel,* J. Comp. Phys., 188 (2003), pp. 593-610.
3. J. H. Bramble, *Multigrid Methods,* Pitman Research Notes in Math. Series 294, Longman Scientific & Technical, 1993.
4. J. H. Bramble and J. E. Pasciak, *The analysis of smoothers for multigrid algorithms,* Math. Comp. 58 (1992), pp. 467-488.
5. J. H. Bramble and J. E. Pasciak, *New convergence estimate for multigrid algorithms,* Math. Comp. 49 (1987), pp. 311-329.
6. M. Prieto, R. Montero, D. Espadas, I. Llorente, and F. Tirado, *Parallel multigrid for anisotropic elliptic equations*, Journal of Parallel and Distributed Computing, Academic Press, 61 (2001), pp. 96-114.
7. B. F. Smith, P. Bjørstad, and W. D. Gropp, *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations*, Cambridge University Press, 1996.
8. U. Trottenberg, A. Schüller, and C. Oosterlee, *Multigrid*, Academic Press, 2000.
9. L. R. Scott and D. Xie, *Parallel linear stationary iterative methods*, IMA Volumes in Mathematics and Its Applications, Springer-Verlag, 120 (2000), pp. 31-55.
10. D. Xie and L. Adams, *New parallel SOR method by domain partitioning,* SIAM J. on Scientific Computing, 20 (1999), pp. 2261-2281.
11. D. Xie and L. R. Scott, *Parallel U-cycle multigrid method*, UH/MD 240, University of Houston, USA, (or Proceedings of the 8th Copper Mountain Conference on Multigrid Methods), 1997.