Analysis Of an Aggregation-Based Algebraic Multigrid Methods On Matrices Related to a 1D Model Problem

Meng-Huo Chen*

January 1, 2013

Abstract

We show that a pairwise aggregation-based algebraic 2-grid method, applied to the linear system $Ax = b$ arising from a 1D model problem for Poisson's equation with Dirichlet boundary conditions, reduces the A-norm of the error at each step by at least the factor $\sqrt{5/8}$. We then generalize this result to problems with the same eigenvectors but different eigenvalues from the model problem,and also to problems with different eigenvectors that are especially well-suited to the method. Finally, we discuss the reduction in the A-norm of the error when the 2-grid method is replaced by a multigrid V-cycle and indicate that conjugate gradient acceleration is required in order to improve the degraded performance of multigrid V-cycle.

1 Introduction

Recently aggregation-based algebraic multigrid methods with piecewise constant prolongation have received considerable attention. See, for instance, [3, 4]. Although these methods may require extra Krylov space iterations on coarse grids in order to perform well in a multigrid setting [2], their relative simplicity compared to other algebraic multigrid methods makes them attractive for analysis.

In [3] it was shown that a 2-grid pairwise aggregation-based method applied to Poisson-type problems with periodic boundary conditions gave rise to an iteration matrix with spectral radius (and hence *asymptotic* convergence rate) 0.5. In this paper we consider matrices related to the 1D model problem for Poisson's equation with Dirichlet boundary conditions, and instead of bounding the asymptotic convergence rate, we derive an estimate of the amount by which the A-norm of the error is reduced at each step. We show that for a 2-grid method the reduction in the A-norm of the error at each step is at least $\sqrt{5/8}$. Numerical computations indicate that the actual worst-case reduction factor is $\sqrt{2}/2$, and we indicate where the overestimate occurs in our proven result. An interesting point about methods with piecewise constant prolongation is that they may require more careful analysis than standard multigrid methods with, say, piecewise linear prolongation. Even for model problems, the coarse grid solve does not annihilate or nearly annihilate any eigencomponents of the error, and, inevitably, with piecewise constant prolongation the fine grid approximation obtained after the coarse grid solve has high frequency error that was not there previously.

2 Analysis of a 2-Grid Pairwise Aggregation Method

We first consider the linear system $Ax = b$ arising from Poisson's equation in one dimension with Dirichlet boundary conditions. Thus A is an $N \times N$ tridiagonal matrix with 2's on the main diagonal and -1's on the first sub and super-diagonal. (We ignore the scale factor $1/h^2$ that would arise in an

*University of Washington, Department of Applied Mathematics

actual finite difference scheme with spacing $h = 1/(N+1)$. Constant factors of this sort will not make a difference in the scheme that we analyze). Assume that N is even, and define the $N \times N/2$ prolongation matrix P by $P_{2l-1,l} = P_{2l,l} = 1$, $l = 1, ..., N/2$, with all other entries of P being 0. When P is applied to a vector v of length $N/2$, the result is a vector Pv of length N with each entry of v appearing twice. This is the piecewise constant prolongation operator. Define an $N/2 \times N/2$ matrix A_C by $A_C = P^T AP$. This is the usual Galerkin definition of the "coarse grid matrix" in a geometric multigrid method. It is easily seen that A_C is also a tridiagonal matrix with 2's on its main diagonal and -1's on the first sub and super-diagonal.

Starting with an approximate solution x_j , we consider the effect of a "coarse grid" solve followed by a damped Jacobi iteration with damping factor of 2. First the residual $r_j = b - Ax_j$ is computed. Then the residual after the coarse grid correction and post-smoothing is

$$r_{j+1} = (I - \frac{1}{4}A)(I - APA_C^{-1}P^T)r_j$$

Multiplying by A^{-1}, we obtain the equation for the error $e_{j+1} \equiv A^{-1}b - x_{j+1} \equiv x - x_{j+1}$:

$$e_{j+1} = (I - \frac{1}{4}A)(I - PA_C^{-1}P^T A)e_j \tag{1}$$

The numerical computation exhibits that the spectral radius of the iteration matrix $(I - \frac{1}{4}A)(I - PA_C^{-1}P^T A)$ in (1) is 0.5, suggesting that asymptotically the 2-norm of the error will be reduced by a factor of 2 at each step. Because the iteration matrix is not symmetric, however, this does not imply that the 2-norm of the error is reduced by this factor at every step. For $N = 64$, for example, the 2-norm of the matrix in (1) was computed to be 20.36 and for $N = 128$ it was computed to be 40.74. This means that the 2-norm of the error may grow before eventually decreasing at the asymptotic rate determined by the spectral radius of the iteration matrix. If, instead, one computes the A-norm of the iteration matrix,

$$\begin{aligned} \|G\|_A \equiv \max_{\|v\|_A=1} \|Gv\|_A &= \max_{\|A^{1/2}v\|_A=1} \|A^{1/2}Gv\|_2 \\ &= \max_{\|w\|_2=1} \|A^{1/2}GA^{-1/2}w\|_2 \end{aligned} \tag{2}$$

one finds that it is always less than $\sqrt{2}/2$ and seems to approach $\sqrt{2}/2$ as $N \to \infty$. This is the best that one could hope to prove, and we will prove a slightly weaker result - that the A-norm of the error is reduced by at least the factor $\sqrt{5/8}$ at each step. Multiplying the matrix in (1) by $A^{1/2}$ on the left and by $A^{-1/2}$ on the right, we are interested in the quantity

$$\sigma \equiv \|(I - \frac{1}{4}A)(I - A^{1/2}PA_C^{-1}P^T A^{1/2})\|_2 \tag{3}$$

which is an upper bound on the factor by which the A-norm of the error is reduced at each step. From here on $\|\cdot\|$ without a subscript will always denote the 2-norm, while the A-norm will still be denoted as $\|\cdot\|_A$. From the approach in [1][pp. 183-187] and the discussion followed, and letting $C = PA_C^{-1}P^T$, we can write

$$\sigma \leq \max_{\substack{\|y\|=\|I-A^{1/2}CA^{1/2}\| \\ y \in \mathcal{R}(I-A^{1/2}CA^{1/2})}} \|(I - \frac{1}{4}A^{1/2}IA^{1/2})y\| \tag{4}$$

where $\mathcal{R}(\cdot)$ denotes the range. It is shown in [1][Theorem 12.1.1] that $\|I - A^{1/2}CA^{1/2}\| = 1$ and that $A^{1/2}$ times the range of $I - A^{1/2}CA^{1/2}$ is the null space of P^T. Thus

$$\sigma \leq \max_{\substack{\|y\|=1 \\ A^{1/2}y \in \mathcal{N}(P^T)}} \|(I - \frac{1}{4}A^{1/2}IA^{1/2})y\| = \max_{\substack{\|y\|=1 \\ A^{1/2}y \in \mathcal{N}(P^T)}} \|(I - \frac{1}{4}A)y\| \tag{5}$$

The eigenvalues λ_k and eigenvectors of A

$$\lambda_k = 2 - 2\cos\left(\frac{k\pi}{N+1}\right) \qquad k = 1, \ldots, N. \tag{6}$$

$$q_j^{(k)} = \sqrt{\frac{2}{N+1}}\sin\left(\frac{jk\pi}{N+1}\right) \qquad j, k = 1, \ldots, N. \tag{7}$$

where $q_j^{(k)}$ is the jth component of the kth eigenvector. Then we can write the eigendecomposition of A as $A = Q\Lambda Q^T$, where Λ is the diagonal matrix of eigenvalues in (6) and Q is the orthogonal matrix whose columns are the eigenvectors in (7). Furthermore, the eigenvalues of A have the following property:

$$\begin{aligned} \max_{i=1,\ldots,N/2}(1 - \frac{1}{4}\lambda_i)^2 &= (1 - \frac{1}{4}\lambda_1)^2 = \left(\frac{1}{2} + \frac{1}{2}\cos\frac{\pi}{N+1}\right)^2 < 1 \\ \max_{i=N/2+1,\ldots,N}(1 - \frac{1}{4}\lambda_i)^2 &= (1 - \frac{1}{4}\lambda_{\frac{N}{2}+1})^2 = \left(\frac{1}{2} + \frac{1}{2}\cos\frac{(N/2+1)\pi}{N+1}\right)^2 < \frac{1}{4} \end{aligned} \tag{8}$$

Using eigendecomposition of A the equation (19) can rewritten as

$$\sigma \leq \max_{\substack{\|Q^Ty\|=1 \\ Q^Ty \in \Lambda^{-1/2}Q^T\mathcal{N}(P^T)}} \|(I - \frac{1}{4}\Lambda)Q^Ty\| \tag{9}$$

Further, using inequalities (8), the square of the right hand side of the above equation can be written as a sum of two half parts:

$$\|(I - \frac{1}{4}\Lambda)Q^Ty\|^2 = \sum_{i=1}^{\frac{N}{2}}(1 - \frac{1}{4}\lambda_i)^2(Q^Ty)^2 + \sum_{i=\frac{N}{2}+1}^{N}(1 - \frac{1}{4}\lambda_i)^2(Q^Ty)^2 \tag{10}$$

$$\leq \sum_{i=1}^{\frac{N}{2}}(Q^Ty)_i^2 + \frac{1}{4}\left(1 - \sum_{i=1}^{\frac{N}{2}}(Q^Ty)_i^2\right) \tag{11}$$

It hopes that the condition $Q^Ty \in \Lambda^{-1/2}Q^T\mathcal{N}(P^T)$, together with $\|Q^Ty\| = 1$, ensures that, for any N, the sum $\sum_{i=1}^{N/2}(Q^Ty)_i^2$ is bounded above by some number $r < 1$, then $\sigma^2 \leq \frac{3}{4}r + \frac{1}{4}$, which is strictly smaller than 1 and the bound is independent of N. In going from equality (10) to inequality (11), however, one may lose something and be unable to recover the *optimal* bound for σ^2. We will show that $r \leq \frac{1}{2}$, from which it follows that $\sigma \leq \sqrt{5/8}$.

Now, the null space of P^T is the span of the set of vectors whose only two nonzero components are a 1 in position $2l-1$ and a -1 in position $2l$, $l = 1, ..., N/2$. Since the matrix Q is symmetric, the space $\Lambda^{-1/2}Q^T \cdot \mathcal{N}(P^T)$ in (9) can be written as the span of the vectors

$$\Lambda^{-1/2}(q^{(1)} - q^{(2)}) \ , \ \Lambda^{-1/2}(q^{(3)} - q^{(4)}) \ , \ \ldots \ , \ \Lambda^{-1/2}(q^{(N-1)} - q^{(N)}) \tag{12}$$

Let K be the N by $N/2$ matrix whose columns are the vectors in (12). For any nonzero vector $Q^T y$ in the span of the columns of K,

$$Q^T y = Kv = \begin{pmatrix} K_1 v \\ K_2 v \end{pmatrix} \tag{13}$$

If one can show that $\|K_1 v\|^2 \leq \|K_2 v\|^2$ for any v, it will imply that $\sum_{i=1}^{N/2} (Q^T y)_i^2 \leq \frac{1}{2}$.

We first compute inner products of the vectors in (12) and for this the Lagrange trigonometric identities

$$\sum_{i=1}^{p} \cos(i\theta) = -\frac{1}{2} + \frac{\sin((p+\frac{1}{2})\theta)}{2\sin(\frac{\theta}{2})} \tag{14}$$

$$\sum_{i=1}^{p} \sin(i\theta) = \frac{1}{2}\cot(\frac{\theta}{2}) - \frac{\cos((p+\frac{1}{2})\theta)}{2\sin(\frac{\theta}{2})} \tag{15}$$

will be needed. The result of the computations gives the following lemma:

Lemma 1. Let K be the N by $N/2$ matrix whose columns are the vectors in (12). Let K_1 consist of the first $N/2$ rows of K and let K_2 consist of the last $N/2$ rows of K. Then

$$K^T K = K_1^T K_1 + K_2^T K_2 = I - \frac{1}{N+1} ee^T$$

where e is the $(N/2)$-vector of 1's,

$$(K_2^T K_2)_{l,m} = \frac{(-1)^{l+m}}{2(N+1)} \csc\left(\frac{(l+m-\frac{1}{2})\pi}{N+1}\right), \qquad l \neq m$$

$$(K_2^T K_2)_{l,l} = \frac{1}{2} + \frac{1}{2(N+1)} \csc\left(\frac{(2l-\frac{1}{2})\pi}{N+1}\right)$$

$$(K_1^T K_1)_{l,m} = -\frac{1}{N+1} + \frac{(-1)^{l+m+1}}{2(N+1)} \csc\left(\frac{(l+m-\frac{1}{2})\pi}{N+1}\right), \qquad l \neq m$$

$$(K_1^T K_1)_{l,l} = \frac{1}{2} - \frac{1}{N+1} - \frac{1}{2(N+1)} \csc\left(\frac{(2l-\frac{1}{2})\pi}{N+1}\right)$$

The ratio of the two norms $\|K_1 v\|$, $\|K_2 v\|$ in equation (13)

$$\frac{v^T K_1^T K_1 v}{v^T K_2^T K_2 v}$$

is maximized by an eigenvector v corresponding to the largest eigenvalue of $(K_2^T K_2)^{-1}(K_1^T K_1)$. Thus we wish to show that the spectral radius $\rho((K_2^T K_2)^{-1}(K_1^T K_1))$ is less than 1. By the theory of P-regular splitting [6] and result in Lemma 1, we have the desired result as in Lemma 2:

Lemma 2. With K, K_1 and K_2 defined as in Lemma 1, we have

$$\rho((K_2^T K_2)^{-1}(K_1^T K_1)) < 1$$

With Lemma 1 and Lemma 2, we have now proved that any vector $Q^T y$ that has norm one and lies in the span of the vectors in (12) will be such that the sum of squares of its first $N/2$ components is less than $1/2$, and this proves the following theorem:

Theorem 3. The 2-grid pairwise aggregation method corresponding to the iteration matrix in (1) reduces the A-norm of the error at each step by at least the factor $\sqrt{5/8}$.

□

We have analyzed a method (1) that uses only post-smoothing, while the analysis in [3] used pre- and post-smoothing. In this case, equation (1) for the error is replaced by

$$e_{j+1} = (I - \frac{1}{4}A)(I - PA_c^{-1}P^T A)(I - \frac{1}{4}A)e_j \tag{16}$$

Again, one can study this problem numerically by simply computing the eigenvalues and different norms of the matrix in (16) for various values of N. Note, however, that the A-norm of this matrix is the 2-norm of the *symmetric* matrix

$$\begin{aligned} &A^{1/2}(I - \frac{1}{4}A)(I - PA_c^{-1}P^T A)(I - \frac{1}{4}A)A^{-1/2} \\ &= (I - \frac{1}{4}A)(I - A^{1/2}PA_c^{-1}P^T A^{1/2})(I - \frac{1}{4}A) \end{aligned} \tag{17}$$

This matrix has the same eigenvalues as the one in (16), so its 2-norm will be the same as the spectral radius of the matrix in (16). This spectral radius was again by numerical calculation to be less than 0.5 and approaching 0.5 as N increased, so in this case one can expect a reduction in the A-norm of the error by at least a factor of 2 at each step. Using the derivation and arguments used in obtaining Theorem 3, we can prove the following theorem:

Theorem 4. Let the error e_{j+1} at step $j+1$ satisfy (16). Then

$$\|e_{j+1}\|_A \leq \sqrt{\frac{17}{32}}\|e_j\|_A \tag{18}$$

Proof : The spectral radius and hence the 2-norm of the symmetric matrix in (16) is the same as the spectral radius of

$$(I - \frac{1}{4}A)^2(I - A^{1/2}PA_c^{-1}P^T A^{1/2})$$

By the same arguments used to obtain estimate (19) and (9), one can show that 2-norm, and hence the spectral radius, of this matrix is bounded above by

$$\max_{\substack{\|y\|=1 \\ A^{1/2}y \in \mathcal{N}(P^T)}} \|(I - \frac{1}{4}A)^2 y\| = \max_{\substack{\|Q^T y\|=1 \\ Q^T y \in \Lambda^{-1/2}Q^T\mathcal{N}(P^T)}} \|(I - \frac{1}{4}\Lambda)^2 Q^T y\| \tag{19}$$

Arguing as in (10), we can write

$$\begin{aligned} \|(I - \frac{1}{4}\Lambda)^2 Q^T y\|^2 &= \sum_{i=1}^{\frac{N}{2}}(1 - \frac{1}{4}\lambda_i)^2(Q^T y)^2 + \sum_{i=\frac{N}{2}+1}^{N}(1 - \frac{1}{4}\lambda_i)^2(Q^T y)^2 \\ &\leq \sum_{i=1}^{\frac{N}{2}}(Q^T y)_i^2 + \frac{1}{16}\left(1 - \sum_{i=1}^{\frac{N}{2}}(Q^T y)_i^2\right) \end{aligned} \tag{20}$$

Since it was shown that the conditions $Q^T y \in \Lambda^{-1/2}Q^T\mathcal{N}(P^T)$ and $\|Q^T y\| = 1$ imply that the first sum in (20) is bounded by $1/2$, it follows that the quantity in (20) is bounded by $1/16 + (15/16)\cdot(1/2) = 17/32$.

□

One might also consider only pre-smoothing, and the equation for the error becomes

$$e_{j+1} = (I - \frac{1}{4}A)(I - PA_c^{-1}P^TA)e_j \tag{21}$$

In this case we again found numerically that the spectral radius of the iterartion matrix was about 0.5, the A-norm of the matrix was about $\sqrt{2}/2$ for large N and the 2-norm of the matrix was only slightly larger than with both pre- and post-smoothing. The proven and computed results of this section are summarized in the table 1.

Table 1: Spectral radius and norms of the iteration matrix for a 2-grid aggregation method using damped Jacobi smoothing with damping factor 2.

Multigrid Iteration Matrix for damped Jacobi smoothing with factor 2			
	spectral radius	2-norm	A-norm
post-smoothing only	0.5	> 1	$\frac{\sqrt{2}}{2}$ (proved $\leq \sqrt{\frac{5}{8}}$)
pre-smoothing only	0.5	> 1	$\frac{\sqrt{2}}{2}$
pre- and post-smoothing	0.5	> 1	0.5 (proved $\leq \sqrt{\frac{17}{32}}$)

3 Generalization to Matrices with the Same Eigenvectors but Different Eigenvalues

The analysis in the previous section holds when the 2-grid pairwise aggregation procedure is applied to *any* symmetric positive definite linear system $\tilde{A}\tilde{x} = \tilde{b}$. Inequality (10) holds provided the eigenvalues $\tilde{\lambda}_i$ of $\tilde{A}$ satisfy

$$0 < \bar{\lambda}_i < 8, \quad i = 1, ..., \frac{N}{2}; \qquad \text{and} \qquad 2 < \bar{\lambda}_i < 6, \quad i = \frac{N}{2} + 1, ..., N \tag{22}$$

This type of eigenvalue distribution, in which the ratio of the largest to the $(N/2+1)$st eigenvalues is $O(1)$ independent of N is critical to the performance of any 2-grid-type method, as the coarse grid solve can, at most, reduce half of the eigencomponents and the relaxation scheme must effectively reduce the others.

Suppose matrix $\tilde{A}$ have the same eigenvectors as the matrix A of the previous section but different eigenvalues satisfying (22). The property that allowed us to prove $\sigma \leq \sqrt{5/8}$ was that the matrix $H = K_2^TK_2 - K_1^TK_1$ was positive definite, since this implied that after the coarse grid solve the error vector would have less than half of its weight in the direction of eigenvectors corresponding to the smaller eigenvalues $\lambda_1, ..., \lambda_{N/2}$. Suppose $\tilde{A} = QD\Lambda Q^T$, where Q and Λ are as before and D is a diagonal matrix with positive diagonal entries. Then the matrix K in Lemma 1 is replaced by a matrix $\tilde{K}$, whose columns are

$$(D\Lambda)^{-\frac{1}{2}}(q^{(1)} - q^{(2)}), ..., (D\Lambda)^{-\frac{1}{2}}(q^{(N-1)} - q^{(N)})$$

Let $\tilde{K}_1$ denote the first $N/2$ rows of $\tilde{K}$ and let $\tilde{K}_2$ denote the last $N/2$ rows of $\tilde{K}$. Also let D_1 denote the upper left $N/2$ by $N/2$ block of D and let D_2 denote the lower right $N/2$ by $N/2$ block of D. Then $\tilde{K}_1 = D_1^{-1/2}K_1$ and $\tilde{K}_2 = D_2^{-1/2}K_2$. Suppose the diagonal entries of D_1 are all greater than or equal

to 1 while the diagonal entries of D_2 are all less than or equal to 1. Assume also that the eigenvalues $\tilde{\lambda}_i = d_i \lambda_i$ satisfy 22. Then

$$\tilde{K_2}^T \tilde{K_2} - \tilde{K_1}^T \tilde{K_1} = K_2^T K_2 - K_1^T K_1 + K_2^T (D_2^{-1} - I) K_2 + K_1^T (I - D_1^{-1}) K_1$$

is positive definite because it is the sum of the positive definite matrix $K_2^T K_2 - K_1^T K_1$ and two positive semidefinite matrices $K_1^T (I - D_1^{-1}) K_1$ and $K_2^T (D_2^{-1} - I) K_2$. Therefore, Theorem 3 holds for $\tilde{A}$.

Figure 1 shows the magnitudes of entries in $\tilde{A}$, for $N = 64$. Note that $\tilde{A}$ is no longer tridiagonal (in fact, it is dense) and the off diagonal entries have different signs. The diagonal entries seem to be the largest, but the matrix is not diagonal dominant.

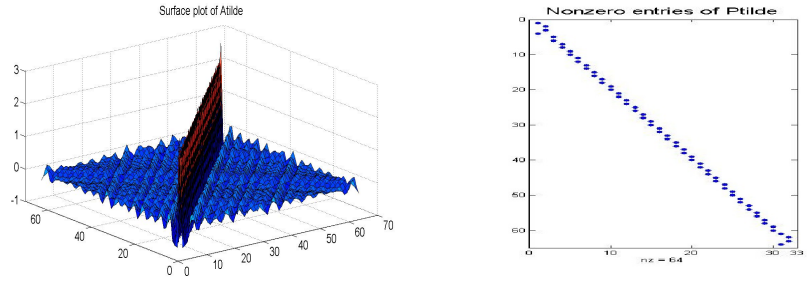

Figure 1: Surface plot of the entries of $\tilde{A}$ for $N = 64$, along with the nonzeros of the prolongation matrix $\tilde{P}$ chosen by an automatic aggregation procedure.

It seems interesting to see what a standard algorithm for choosing aggregates would do with such a matrix. Applying the aggregation algorithm given in [4] on matrix $\tilde{A}$, one sees the algorithm mostly aggregates successive unknowns with few exception near the end of the domain, as shown on the right hand side in the figure 1.

The spectral radius of iteration matrix (with post smoothing only) with $\tilde{P}$ is somewhat larger : 0.6743 instead of 0.4954 for successive pairwise aggregation. The A-norm is 0.8035 instead of 0.6277. Still, the automatic aggregation algorithm in [4] do a reasonable job on this dense matrix $\tilde{A}$.

4 Generalization to matrix with the different eigenvectors, same eigenvalues

As noted previously, when the eigenvalues of A are distributed as in (22), inequality (10) implies that the A-norm of the error is reduced at each step by a fixed amount independent of N, provided that any vecotr $Q^T y$ with norm 1 that lies in the space $\Lambda^{-1/2} Q^T \cdot \mathcal{N}(P^T)$, where the Q is the matrix of orthonormal eigenvectors of A, satisfies $\sum_{i=1}^{N/2} (Q^T y)_i^2$, for some $r < 1$ independent of N. An extreme case would be one in which pairs of successive column of Q^T had the same first $N/2$ components. In this case r would be 0 and the A-norm of the error would be reduced by at least a factor of 2 at each step. This corresponds to a problem in which the "coarse grid solve" completely annihilates error components in the direction of the first $N/2$ eigenvectors. If $s^{(1)}, ..., s^{(N/2)}$ are any $N/2$ orthogonal vectors of length $N/2$, each with norm $\sqrt{2}/2$, then

$$Q^T = \begin{pmatrix} s^{(1)} & s^{(1)} & \dots & s^{(N/2)} & s^{(N/2)} \\ s^{(1)} & -s^{(1)} & \dots & s^{(N/2)} & -s^{(N/2)} \end{pmatrix} \tag{23}$$

is such a matrix.

Numerical tests with matrices $\tilde{A}$ whose eigenvectors were the rows of Q^T in (23), where $s^{(1)}, ..., s^{(N/2)}$ were random orthogonal orthogonal $N/2$-vectors with norm $\sqrt{2}/2$, and whose eigenvalues were either those given in (6) or those randomly chosen in the previous section to satisfy (22), showed that the $\tilde{A}$-norm of the iteration matrix using post-smoothing only was less than 0.5 and close to this value for large N. Again the matrices were dense, with the largest entries on the diagonal, but not diagonally dominant.

Again, the automatic aggregation procedure described in [5] was tested on these problems. Interestingly, as shown in figure, the aggregates are quite different from the successive pairs used in the analysis, and the spectral radius and $\tilde{A}$ are very close to 1: 0.9982 for this particular problem.

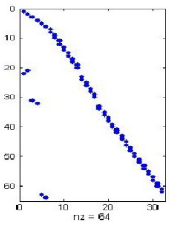

Figure 2: Nonzero entries of $\tilde{P}_2$

5 Multigrid

When the 2-grid method of the previous section is replaced by a multigrid method, the bound (4) on the reduction factor for the A-norm of the error is replaced by

$$\sigma \leq \max_{\substack{\|y\|=\|I-A^{1/2}\tilde{C}A^{1/2}\| \\ y\in\mathcal{R}(I-A^{1/2}\tilde{C}A^{1/2})}} \|(I-\frac{1}{4}A)y\| \tag{24}$$

where $\tilde{C} = P\tilde{A}_C^{-1}P^T$ and $\tilde{A}_C^{-1}$ is an approximation to A_C^{-1} resulting from work on coarse grids. As an example we consider a 3-grid V-cycle method with Jacobi pre and post-smoothing at each grid level with damping factor 2. We can derive the approximation $\tilde{A}_C^{-1}$ to A_C^{-1}:

$$\tilde{A}_C^{-1} = \frac{1}{4}I + \frac{1}{4}(I-\frac{1}{4}A_C) + (I-\frac{1}{4}A_C)P_C A_{CC}^{-1} P_C^T (I-\frac{1}{4}A_C) \tag{25}$$

where P_C denote prolongation from the coarsest grid to next finer level and $A_{CC} = P_C^T A_C P_C$. If more levels are used, then A_{CC}^{-1} in (25) can be replaced by a similar formula for $\tilde{A}_{CC}^{-1}$, etc., until the coarsest level is reached.

Unfortunately, as noted elsewhere [4, 2], performance of the aggregation method degrades as the number of grids is increased. For a fixed number of grids, the spectral radius and A-norm of the iteration matrix still appear to be bounded below 1 by an amount of that is independent of N, but that amount decreases with the number of grid levels used. Table 2 shows computational observation of these quantities for large N using 2, 3, 4, or 5 grid levels. The spectral radius of the iteration matrix appears to be equal to $1 - 1/2^{p-1}$, where p is the number of grid levels.

6 Conjugate gradient acceleration

The degradation of performance of aggregation multigrid V-cycle can be improved by using conjugate gradient (CG) acceleration where multigrid method is used as a preconditioner. The convergence improved significantly, but the number of CG iteration increases as the grid size (and number of grid level) increases. A numerical result is shown in table 3.

Table 2: Spectral radius and A-norms of the iteration matrix for a multigrid aggregation method using damped Jacobi smoothing with damping factor 2.

Multigrid Iteration Matrix for damped Jacobi smoothing with factor 2		
	spectral radius	A-norm
2-grid	$0.5 = 1/2$	0.7071
3-grid	$0.75 = 3/4$	0.7906
4-grid	$0.875 = 7/8$	0.8839
5-grid	0.9375=15/16	0.9395

Table 3: CG acceleration with multigrid V-cycle preconditioner

CG acceleration with multigrid V-cycle preconditioner		
N	No. of iteration	$\kappa = \frac{\lambda_{max}}{\lambda_{min}}$
64 (2-grid)	7	1.99
128 (3-grid)	10	3.97
256 (4-grid)	15	7.92
512 (5-grid)	21	15.83
1024 (6-grid)	30	30.71
2048 (7-grid)	42	61.37

On the other hand, if multigrid W-cycle is used as a preconditioner of conjugate gradient method, a preliminary numerical result shown in table 4 exhibits that the number of CG iterations hardly grows for N being large. The operation count grows as $O(N \log N)$, which is much mild than the V-cycle CG method.

Table 4: CG acceleration with multigrid W-cycle preconditioner

CG acceleration with multigrid V-cycle preconditioner	
N	No. of iteration
64 (2-grid)	6
128 (3-grid)	8
256 (4-grid)	9
512 (5-grid)	10
1024 (6-grid)	10
2048 (7-grid)	10
4096 (8-grid)	11

7 Conclusion

In this paper, we have derived the bounds for the reduction number for two grid method with pairwise aggregation method and damped Jacobi smoothing with damping factor 2. The analysis can be extended to a dense matrix $\tilde{A}$ with same eigenvectors as A and eigenvalues distribution satisfying (22) and a dense matrix $\tilde{A}_2$ with the same eigenvalues as A and eigenvectors satisfying an extreme case in (23).

The performance of aggregation-based multigrid V-cycle method appear to be degraded as the number of grid levels increases. Conjugate gradient acceleration improves the performance of aggregation-based multigrid, but the number of cycles increases significantly as the size of the problem increases. A better remedy is to use W-cycle multigrid as a preconditioner for conjugate gradient iteration.

References

[1] A. Greenbaum, *Iterative Methods for Solving Linear Systems, SIAM*, Philadelphia, 1997.

[2] Y. Notay and P. S. Vassilevski , *Recursive Krylov-based multigrid cycles*, Numer. Linear Algebra Appl. 2006; **0**: 1-20

[3] A. C. Muresan and Y. Notay , *Analysis of aggregation-based multigrid*, SIAM. J. Sci. Comput., 30(2008), pp 1082 - 1103.

[4] A. Napov and Y. Notay, *Algebraic analysis of aggregation-based multigrid*, Numer. Linear Algebra Appl., 18 (2011), pp 539 - 564

[5] Y. Notay, *An aggregation based algebraic multigrid method*, Electronic Trans. Numer. Anal., 37 (2010), pp 123 - 146

[6] M. Benzi, *Splitting of symmetric matrices and a question of Ortega*, Lin. Alg. Appl., 429 (2008), pp. 2340-2343