LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

UCRL-PROC-231994

Distance-two interpolation for parallel algebraic multigrid

H. D. Sterck, R. Falgout, J. Nolting, U. M. Yang

June 20, 2007

SciDAC 2007
Boston, MA, United States
June 24, 2007 through June 28, 2007

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

Distance-two interpolation for parallel algebraic multigrid

H. De Sterck[1], R. D. Falgout[2], J. W. Nolting[3] and U. M. Yang[2]

[1] Waterloo University, Waterloo, Canada
[2] Lawrence Livermore National Laboratory, Box 808, Livermore, CA 94551
[3] University of Colorado at Boulder, Boulder, CO

E-mail: hdesterck@waterloo.ca, rfalgout@llnl.gov, Josh_Nolting@colorado.edu, umyang@llnl.gov

Abstract. In this paper we study the use of long distance interpolation methods with the low complexity coarsening algorithm PMIS. AMG performance and scalability is compared for classical as well as long distance interpolation methods on parallel computers. It is shown that the increased interpolation accuracy largely restores the scalability of AMG convergence factors for PMIS-coarsened grids, and in combination with complexity reducing methods, such as interpolation truncation, one obtains a class of parallel AMG methods that enjoy excellent scalability properties on large parallel computers.

1. Introduction

Algebraic multigrid (AMG) [1] is a very efficient parallel algorithm for solving sparse linear systems, $Au = f$, on unstructured grids. Unlike multigrid, it requires only the underlying matrix and nothing else. Additionally, AMG solvers and preconditioners are potentially scalable.

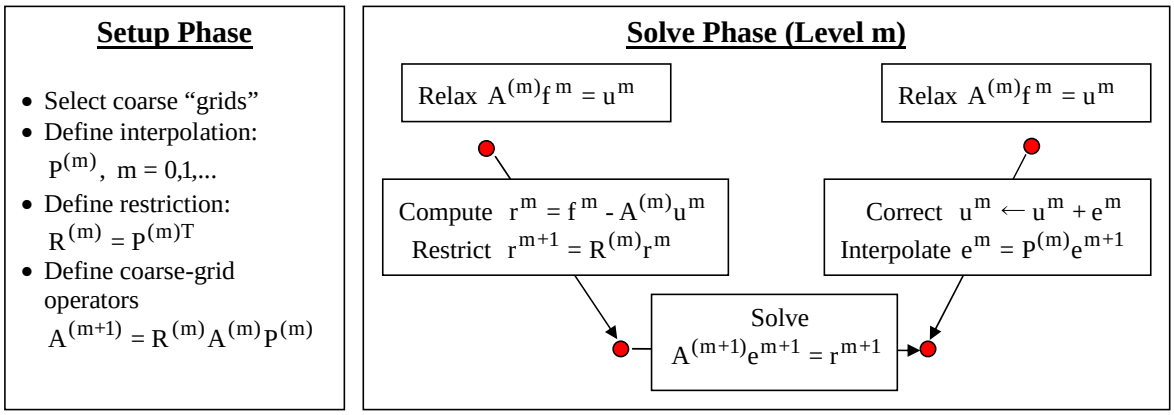
However, when applied to large three-dimensional problems, the "classical" AMG algorithm often generates unreasonably large complexities with regard to memory use as well as computational operations. In order to counter this complexity growth, the Parallel Modified Independent Set (PMIS) coarsening algorithm [2] was introduced. While the use of this coarsening algorithm in combination with a slight modification of Ruge and Stüben's classical interpolation scheme [5] leads to significantly lower complexities, it also leads to degraded convergence and reduced numerical scalability. For various test problems, such as isotropic and grid aligned anisotropic diffusion operators, it is possible to obtain fairly scalable results when AMG is used as a preconditioner in combination with Krylov methods. However, for more complicated problems, such as problems with rotated anisotropies or highly discontinuous material properties, this strategy is not sufficient. Therefore it was necessary to investigate improved interpolation operators.

In this paper we focus on our new extension of classical interpolation, called *extended+i interpolation*. Our investigation shows that this interpolation strategy significantly improves AMG convergence factors compared to classical interpolation. However, it also showed increased complexities. While the increase is not very significant for two-dimensional problems, it is of concern in the three-dimensional case. Therefore we also investigated complexity reducing strategies, such as the use of smaller sets of interpolation points and interpolation truncation.

The use of these strategies led to AMG methods with significantly improved overall scalability. For further details on this work as well as the use of additional long-distance interpolation operators see [3].

2. AMG Interpolation

In AMG, the central idea is that "smooth error," e, that is not eliminated by relaxation must be removed by coarse-grid correction. This is done by solving the residual equation $Ae = r$ on a coarser grid, then interpolating the error back to the fine grid and using it to correct the fine-grid approximation. The coarse grid solution is achieved by recursion. AMG consists of the following two phases:



In the remainder of the paper we will use the following definitions and notations. One of the concepts used is *strength of connection*. A point j strongly influences a point i if $-a_{i,j} > \alpha \max_{k \neq i}(-a_{i,k})$, where $0 < \alpha \leq 1$. In the remainder of the paper, we will use $\alpha = 0.25$. We define the following sets: F is the set of all fine points, C contains all coarse points, $N_i = \{j | a_{ij} \neq 0\}$ is the set of neighbors of point i, S_i is the set of neighbors that strongly influence i, $F_i^s = F \cap S_i$ is the set of fine strong neighbor points of i, $C_i^s = C \cap S_i$ is the set of coarse strong neighbors of i, and $N_i^w = N_i \setminus (F_i^s \cap C_i^s)$ is the set of the remaining neighbors of i.

In classical AMG [5], the interpolation of the error at the F-point i takes the form

$$e_i = \sum_{j \in P_i} w_{ij} e_j, \tag{1}$$

where w_{ij} is an interpolation weight determining the contribution of the value e_j to e_i, and $P_i \subset C$ is the coarse interpolatory set of F-point i. In most classical approaches to AMG interpolation, P_i is a subset of the nearest neighbors of grid point i, i.e. $P_i \subset N_i$, and longer-range interpolation is not considered.

The assumption that algebraically smooth error has small residuals after relaxation implies that $A\,e \approx 0$, and can be rewritten as

$$a_{ii} e_i \approx - \sum_{j \in C_i^s} a_{ij} e_j - \sum_{j \in F_i^s} a_{ij} e_j - \sum_{j \in N_i^w} a_{ij} e_j. \tag{2}$$

From this expression, various interpolation formulae can be derived, such as the interpolation suggested by Ruge and Stüben in [5]. This interpolation (with a slight modification that was

proposed in [4] to avoid extremely large interpolation weights that can lead to divergence) is defined as follows:

$$w_{ij} = -\frac{1}{a_{ii} + \sum\limits_{k \in N_i^w} a_{ik}} \left(a_{ij} + \sum_{k \in F_i^s} \frac{a_{ik} \bar{a}_{kj}}{\sum\limits_{m \in P_i} \bar{a}_{km}} \right), \quad j \in P_i, \tag{3}$$

where

$$\bar{a}_{ij} = \begin{cases} 0 & \text{if } \operatorname{sign}(a_{ij}) = \operatorname{sign}(a_{ii}) \\ a_{ij} & \text{otherwise.} \end{cases}$$

We will refer to it as classical interpolation or 'class' in the remainder of the paper. Note that, when the two F-points i and k do not have a common C-point in C_i^s and C_k^s, such as is illustrated in Figure 1, the second denominator in (3) is vanishing. Therefore when using PMIS coarsening [2], we modified interpolation formula (3) such that if this case occurs, a_{ik} is added to the diagonal term (the term $a_{ii} + \sum_{k \in N_i^w} a_{ik}$ in Eq. (3)), i.e., the strongly influencing neighbor point k of i is treated like a weak connection of i. This modification can lead to inferior convergence. Another problem when using PMIS with classical interpolation is that it does not treat F-points like point p in Figure 1 correctly. Here we have an F-point that is not influenced by any points and has no coarse neighbors. This situation can occur for example if we have a fairly large strength threshold. For classical interpolation, the interpolated error in this point will vanish, and coarse grid correction will not be able to reduce the error in this point.

Both of these issues can be overcome by extending the classical interpolation formula so that the interpolatory set includes C-points that are distance two away from the F-point to be interpolated, i.e. applying the classical interpolation formula, but using the following interpolatory set $P_i = \hat{P}_i := C_i^s \cup \bigcup\limits_{j \in F_i^s} C_j^s$. In Figure 1, point i would now be interpolated using points l, m, n and o, instead of just l and m. This new interpolation formula, which we call *extended interpolation* deals efficiently with strong $F - F$ connections that do not share a common C-point, and in general leads to much better convergence. For numerical test results, see [3].

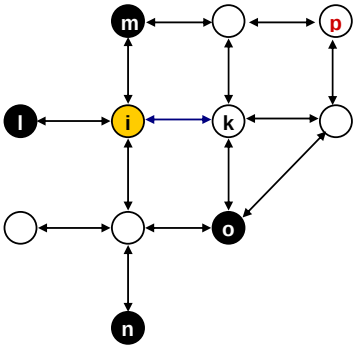


Figure 1. Example of a grid typical for PMIS; black points are C-points

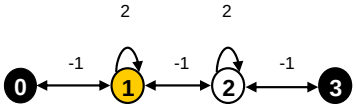


Figure 2. Finite difference 1D Laplace example.

While the use of extended interpolation with PMIS significantly improves convergence compared to PMIS with classical interpolation, it does not always lead to the desired weights. Consider the case given in Figure 2. Here we have a one-dimensional Laplace problem generated

by finite differences. Points 1 and 2 are strongly connected F-points, and points 0 and 3 are coarse points. The interpolatory set for point 1 is {0, 3} when we include distance-two coarse neighbors. If we use extended interpolation to calculate $w_{1,0}$ and $w_{1,3}$, we obtain

$$w_{1,0} = 0.5, \qquad w_{1,3} = 0.5.$$

This is a better result than we would obtain for classical interpolation,

$$w_{1,0} = 1, \qquad w_{1,3} = 0,$$

but worse than the intuitively best interpolation weights,

$$w_{1,0} = 2/3, \qquad w_{1,3} = 1/3. \tag{4}$$

This can be remedied if we include not only connections a_{jk} from strong fine neighbors j of i to points k of the interpolatory set but also connections a_{ji} from j to point i itself. which leads to interpolation weights

$$w_{ij} = \frac{1}{\tilde{a}_{ii}} \left(a_{ij} + \sum_{k \in F_i^s} a_{ik} \frac{\bar{a}_{kj}}{\sum_{l \in \hat{P}_i \cup \{i\}} \bar{a}_{kl}} \right), \quad j \in \hat{P}_i, \tag{5}$$

where

$$\tilde{a}_{ii} = a_{ii} + \sum_{n \in N_i^w} a_{in} + \sum_{k \in F_i^s} a_{ik} \frac{\bar{a}_{ki}}{\sum_{l \in \hat{C}_i \cup \{i\}} \bar{a}_{kl}}. \tag{6}$$

We call this modification of extended interpolation *extended+i interpolation* and refer to it by 'ext+i'. If we apply it to the example in Figure 2 we obtain the weights in (4).

3. Numerical Results

We consider here a diffusion problem with highly discontinuous material properties on the unstructured domain illustrated in Figure 3. The results were obtained on the Linux cluster MCR at LLNL, using up to 1024 processors. Each processor owns about 90,000 rows of the matrix. AMG was used as a preconditioner for GMRES(10). For further details see [3]. PMIS is compared to the parallel coarsening scheme CLJP. The results in Figure 3 show that PMIS-ext+i (PMIS coarsening with ext+i interpolation) restores the convergence that was previously obtained with CLJP-class, however it also has larger operator complexities, C_{op}, than PMIS-class, and while these are significantly better than those of CLJP-class, they are larger than we would like them to be. (Here $C_{op} = \sum_{i=0}^{m} nnz(A^{(i)})/nnz(A)$, where $nnz(A)$ denotes the number of nonzeroes of A, m the number of multigrid levels and $A^{(0)} = A$.) Therefore, it was necessary to consider ways to reduce these complexities while (hopefully) retaining the improved convergence. One of the strategies to achieve this is interpolation truncation. There are essentially two ways we can truncate interpolation operators: by eliminating those weights whose absolute values are smaller than a chosen truncation factor, or by limiting the number of coefficients per row, i.e., choose only the k_{max} largest weights in absolute value. In both cases the new weights need to be re-scaled so that the total sums remain unchanged. Both approaches can lead to significant reductions in setup times and operator complexities, particularly for 3-dimensional problems, but if too much is truncated, the number of iterations rises significantly. In our experiment we used truncation by restricting the number of weights to 4. Applying this strategy to CLJP-class improved times significantly, but operator complexities were still too high. However we obtained excellent scalability for PMIS-ext+i(4), which achieved very good convergence and the overall best time with only slightly larger complexities than PMIS-class.

This work was performed under the auspices of U. S. Department of Energy by University of California Lawrence Livermore National Laboratory under contract W-7405-Eng-48.

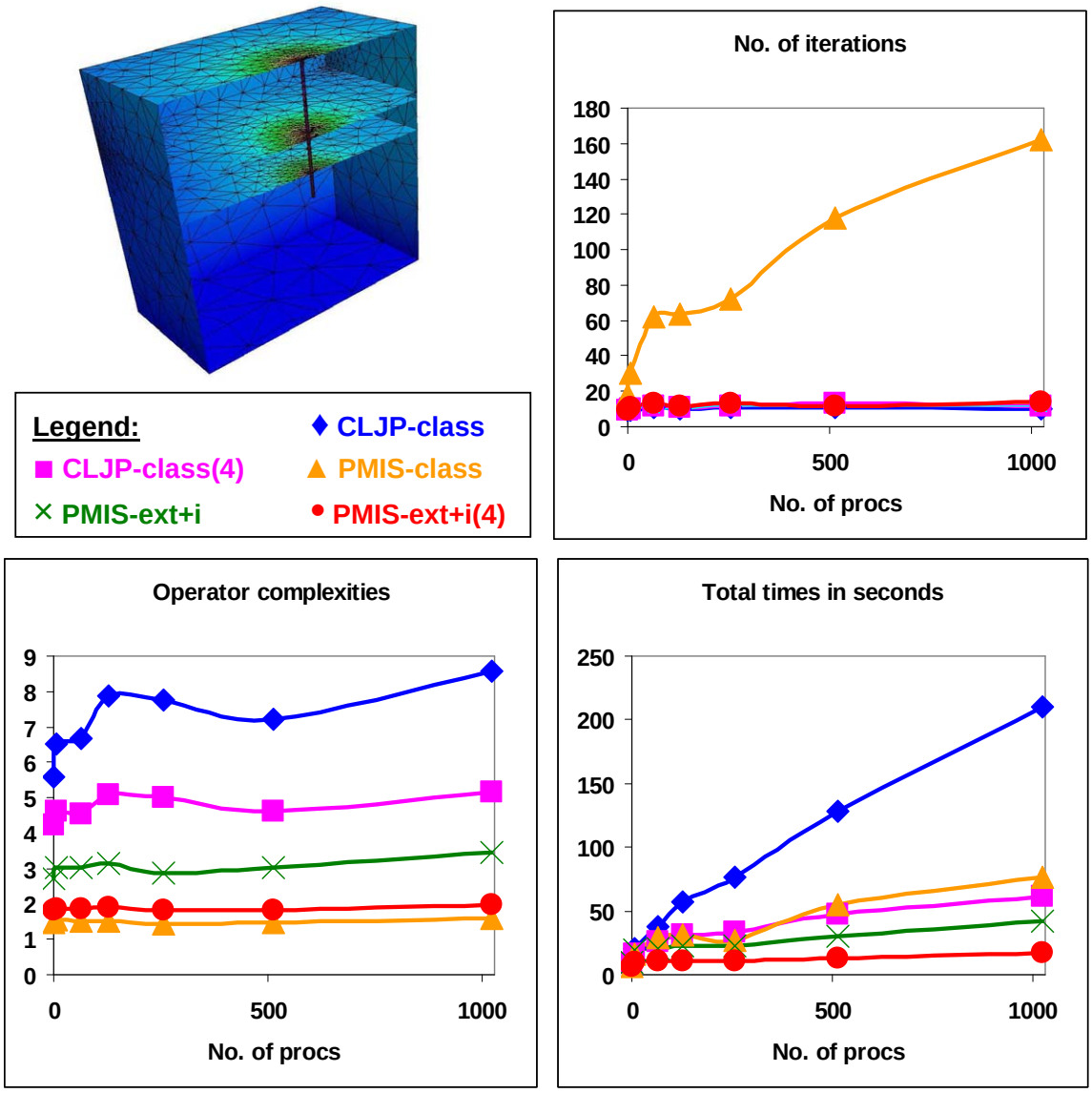


Figure 3. Results for a diffusion problem with highly discontinuous material properties on the domain illustrated in the upper left corner.

References

[1] A. Brandt, S. F. McCormick, and J. W. Ruge. Algebraic multigrid (AMG) for sparse matrix equations, in *Sparsity and Its Applications*, D. J. Evans, ed., Cambridge University Press, Cambridge, 1984.

[2] H. De Sterck, U. M. Yang, and J. J. Heys. Reducing Complexity in Parallel Algebraic Multigrid Preconditioners, *SIAM Journal on Matrix Analysis and Applications* 27 (2006) 1019–1039.

[3] H. De Sterck, R. D. Falgout, J. Nolting, and U. M .Yang, Distance-two interpolation for parallel algebraic multigrid, submitted to *Numerical Linear Algebra with Application*. Also available as LLNL Technical Report UCRL-JRNL-230844, 2007.

[4] V. E. Henson and U. M. Yang, BoomerAMG: a parallel algebraic multigrid solver and preconditioner, *Applied Numerical Mathematics* 41 (2002) 155–177.

[5] J. W. Ruge and K. Stüben, Algebraic multigrid (AMG), in : S. F. McCormick, ed., *Multigrid Methods, vol. 3 of Frontiers in Applied Mathematics* (SIAM, Philadelphia, 1987) 73–130.