

# Black Box Multigrid with coarsening by a factor of three<sup>‡</sup>

J. E. Dendy Jr and J. D. Moulton<sup>\*,†</sup>

*Applied Mathematics and Plasma Physics, MS B284, Los Alamos National Laboratory,  
Los Alamos, NM 87545, U.S.A.*

## SUMMARY

Black Box Multigrid (BoxMG) is a robust variational multigrid solver for diffusion equations on logically structured grids. BoxMG standardly uses coarsening by a factor of two. It handles cell-centered discretizations on logically rectangular grids by treating the cell-centers as the unknowns to be coarsened. Such a strategy does not preserve the cell structure. That is, coarse-grid cells are not the union of fine-grid cells. In some applications, such as local grid refinement, it is desirable that the cell structure be preserved. In this paper, we develop a method that employs coarsening by a factor of three. It is a natural generalization of standard BoxMG, using operator-induced interpolation (which approximately preserves the continuity of the normal flux), restriction as the transpose of interpolation, and Galerkin coarsening. In addition, we introduce a new colored block Gauss–Seidel scheme that is motivated by the form of the interpolation operator, dubbed ‘pattern’ relaxation. We present numerical results that demonstrate robustness of this method with respect to discontinuous diffusion coefficients, boundary conditions, and grid dimension. Published in 2010 by John Wiley & Sons, Ltd.

Received 27 May 2009; Revised 4 December 2009; Accepted 7 December 2009

KEY WORDS: multigrid; algebraic multigrid; variational coarsening

## 1. INTRODUCTION

The efficient and robust solution of the diffusion equation is important in a number of applications, including flow in porous media, conductivity and heat transfer in composite materials, and neutron transport. This equation may be written in the form

$$-\nabla \cdot (D\nabla U) + \sigma U = Q \quad \text{in } \Omega, \quad (1a)$$

$$D\nabla U \cdot \mathbf{n} + \gamma U = 0 \quad \text{on } \partial\Omega, \quad (1b)$$

\*Correspondence to: J. D. Moulton, Applied Mathematics and Plasma Physics, Theoretical Division, MS-B284, Los Alamos National Laboratory, NM 87545, U.S.A.

†E-mail: moulton@lanl.gov

‡This article is a U.S. Government work and is in the public domain in the U.S.A.

Contract grant/sponsor: U.S. Department of Energy; contract grant/number: DE-AC52-06NA25396

where  $\Omega$  is a bounded region in  $R^2$  with boundary  $\partial\Omega$ ,  $D$  is a symmetric positive-definite tensor,  $\sigma$  and  $\gamma$  are nonnegative. The diffusion coefficient,  $D$ , may represent diffusivity of neutrons, the absolute permeability of a porous medium, or the conductivity in a composite material. Thus, the diffusion coefficient may vary discontinuously by orders of magnitude, and may exhibit very fine-scale structure (i.e. media with layers, or inclusions). Similarly, the reaction or removal term,  $\sigma$ , and the source term  $Q$ , may vary discontinuously.

Early efforts to generalize classic geometric multigrid for (1) identified the treatment of highly discontinuous coefficients as a significant challenge. The Black Box Multigrid (BoxMG) method introduced in [1] was one of the first robust multigrid methods for these problems. This method extended ideas introduced in [2], from five-point to nine-point stencils, and hence, from diagonal to full tensor diffusion, and from orthogonal grids to reasonable quadrilateral meshes.

Two significant elements of this advance are critical to delivering a robust multigrid algorithm, namely the use of variational coarsening to construct the hierarchy of coarse-grid operators, and the development of operator-induced interpolation. Variational coarsening, employing the Galerkin coarse-grid operator [3] can be shown to be optimal in the sense that it minimizes the error in the range of interpolation. This has a number of important consequences. First, it implies that as long as an interpolation can be provided by the algorithm, the user needs to only provide a fine-grid problem, as all operators in the multigrid algorithm can be generated automatically. This observation motivated the name *Black Box*. Also, variational coarsening does not make assumptions about the form of the coarse-scale model. Specifically, in many diffusion applications the fine-scale structure of the coefficient is such that simply rediscrctizing on coarser levels with averaged coefficients will not satisfy the approximation property. For these reasons variational coarsening is used in virtually all variants of AMG (e.g. [4, 5]), including smoothed-aggregation (SA)-based AMG [6].

Operator-induced interpolation is an equally important part of this robust algorithm. In particular, it can be shown that for a bilinear FEM discretization of (1) and using bilinear interpolation, the Galerkin coarse-grid operator is equivalent to that derived by first taking the arithmetic average of the diffusion coefficient and then rediscrctizing at the coarse-scale. Thus, such a multigrid algorithm will have a convergence rate that is proportional to the jump in the diffusion coefficient. This is because the gradient of the solution is not continuous across discontinuities in the diffusion coefficient; instead, it is the normal component of the flux that is continuous. Operator-induced interpolation uses this fact to design a more robust interpolation, one that is consistent with the underlying problem and the homogenization of certain one-dimensional interface problems [7]. Later these ideas were extended to nonsymmetric problems, such as convection diffusion in [8] and dubbed matrix-dependent prolongations in [9]. Recent work has investigated the energy minimizing aspects of these interpolations [10], and the design and the adaptivity of interpolation for AMG [11] and SA-based AMG methods [12].

Continuing the development of solution algorithms for logically structured grids is important both for enhancing our understanding of fundamental issues, as well as for practical solvers for real applications. Logically structured body-fitted grids may be used to handle complex shapes, and irregular domains may be embedded in a rectangular domain. Furthermore, new hardware architectures, for example various accelerator configurations, are significantly more efficient with data on structured versus unstructured meshes. On these grids, standard coarsening by a factor of two is a very common method for creating a hierarchy of grids. Although this coarsening may seem most natural for vertex-based discretizations, the flexibility of BoxMG makes it possible to apply this same coarsening strategy directly to the dual-grid of cell-centers, as is shown in Figure 1

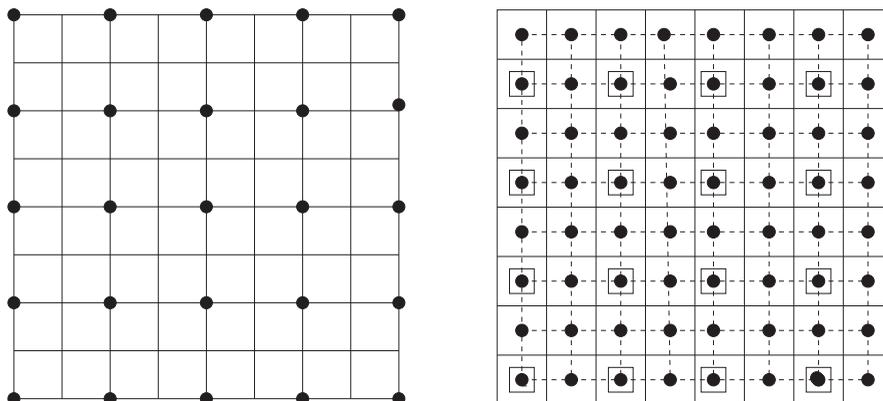


Figure 1. Standard coarsening of vertices (left) and the dual-grid of cell-centers (right).

(right). This observation was made in [1], and is applied frequently. However, in some cases, such as cell-based adaptive mesh refinement (AMR), retaining the cell structure throughout the grid hierarchy may be desirable.

Development of multigrid methods with cell-based coarsening initially focused on coarsening by a factor of two. The first author considered extending the vertex-based method in [2] to use cell-based coarsening for cell-centered difference schemes on uniform rectangular grids. In particular, a cell-based coarsening would replace each aggregate of four fine-scale cell-centered unknowns with a single coarse-grid cell-centered unknown. Although the coarse-grid would not be a subset of fine-grid unknowns, a generalization of bilinear interpolation is still readily obtained. However, in this case the Galerkin coarse-grid operator will have a 25-point stencil or an algebraic complexity of 5, and thus, is impractical. A number of remedies for this complexity problem have been proposed in the literature. First, a method that uses piecewise constant interpolation, with restriction the transpose of interpolation, and the Galerkin coarse-grid operator is considered in McCormick and Thomas (unpublished manuscript). Although this method diverges with V-cycles, it converges with W-cycles. They attribute this behavior to their violation of the well-known condition that the order of interpolation plus the order of restriction should be greater than the order of the operator (see, e.g. [3, 13, 14]). Similarly, in [15], a method is proposed that uses piecewise linear interpolation, with restriction the transpose of piecewise constant interpolation, and the Petrov–Galerkin coarse-grid operator. Although the Petrov–Galerkin operator is generally nonsymmetric, the coarse-grid operator in this method is symmetric. However, in certain cases it can be shown to be equivalent to discretizing on the coarse-grid with an arithmetically averaged diffusion coefficient. Thus, it is not surprising that even for W-cycles, the convergence results reported in [15] are not spectacular.

This research suggests that the design requirements of a robust multigrid method with low operator complexity and without a nested hierarchy of grids are too demanding. However, the nested hierarchy may be restored if we consider cell-based coarsening by a factor of three, as is shown in Figure 3 (right). In terms of the BoxMG algorithm, the operator-induced interpolation may be extended to this case in such a way that the complexity of Galerkin coarse-grid operators is bounded by  $\frac{9}{5}$ , as in the coarsening-by-two case. In the case of cell-based AMR refinement by a factor of three, the results presented here are readily applicable. This is particularly relevant for the challenging problem of defining the discrete operator near the coarse–fine interfaces in

the presence of highly discontinuous coefficients. Adopting a coarsening-by-three strategy has the potential advantage of coarsening more quickly, ultimately reducing the number of levels, and thus, communication steps in a parallel implementation. It also brings the method closer to SA-AMG [6], suggesting the potential for a comparative analysis in the future. In addition, coarsening by even larger factors introduces connections to approximation methods such as the Multiscale Finite Element (MsFEM) method [16]. MsFEM creates coarse-scale discrete models of single-phase flow in porous media through a single coarsening step, typically by a factor of 10. A detailed comparison of BoxMG and MsFEM, however, is beyond the scope of this research. Here we focus on the development of the coarsening-by-three BoxMG algorithm, and leave a rigorous analysis of its convergence for future work.

In this paper, we develop an extension of the BoxMG algorithm using a coarsening factor of three. We introduce notation in Section 2, and then discuss the important features of the coarsening-by-two case, paying particular attention to the interpolation. We develop the extensions to this algorithm in Section 3, and discuss the additional features of the interpolation and smoothing that we use to achieve a robust algorithm. Numerical results are presented in Section 4 that demonstrate the mesh-independent convergence rates and robustness with respect to discontinuous coefficients.

## 2. BACKGROUND

### 2.1. Discretizations

We consider logically structured grids in two dimensions, and are interested in discretizations that generate, at most, a 9-point nearest neighbor stencil. The diffusion tensor and removal coefficient,  $D$  and  $\sigma$ , respectively, are assumed to be piecewise constant on the domain, and constant on each fine-grid cell. However, even in this restricted setting there are a variety of possible discretizations, and each impacts the properties of the resulting matrix. For example, Galerkin Finite Element methods (GFEM) can be applied either using bilinear basis functions on quadrilaterals, or linear basis functions on a triangulation of the quadrilateral grid. However, with strongly anisotropic diffusion, even if the anisotropy is aligned with the coordinate axes, the GFEM discretization does not lead to an M-matrix. Also, GFEM does not provide local mass conservation without post processing.

Thus, many discretizations are based on the first-order form of the diffusion equation, and explicitly enforce the continuity of the normal flux and local mass conservation. For example, using the Finite Volume Method (FVM) on an orthogonal grid with a diagonal diffusion tensor, it is possible to derive the *standard* 5-point difference formula for the diffusion equation. If the control volume is centered on a grid cell, then the resulting discretization has weights that are given by the harmonic average of the neighboring diffusion coefficients. In contrast, if the control volume is centered on a vertex (the dual mesh), then the weights of the discretization are given by the arithmetic average of the neighboring diffusion coefficients. Since these averages may differ by several orders of magnitude, the condition number of the matrix and the difficulty associated with coarsening across discontinuities in the two cases are very different. However, the choice may be dictated by the application. For example, the vertex-centered FVM is necessary for diffusion-synthetic acceleration of the neutron transport equation, as the cell-centered FVM is unstable in this case [17].

In the numerical examples (Section 4), we present results for both of these FVM discretizations, as well as for a GFEM discretization. Even more advanced discretizations may be required for general tensor diffusion on distorted quadrilateral meshes, such as Mixed Finite Element methods (e.g. [18]) and Mimetic Finite Difference methods (e.g. [19, 20]). But these do not lead to 9-point nearest neighbor stencils for the diffusion operator, and hence are beyond the scope of this solver.

## 2.2. Linear system and Multigrid

Let  $G^M$ , where  $M$  is to be determined, be a logically rectangular grid and let

$$A_M u_M = f_M \quad (2)$$

be the given linear system of equations on  $G^M$ . In BoxMG, a sequence of subgrids of  $G^M$  are derived along with auxiliary operators  $A_k$ ,  $k < M$ . The sequence is terminated when the cost of direct solution of

$$A_1 u_1 = f_1 \quad (3)$$

is trivial compared with the cost of direct solution of (2). One  $V(ID, IU)$ -cycle employs  $ID$  relaxation sweeps on (2), before forming the equation

$$A_{M-1} V_{M-1} = f^{M-1} = I_M^{M-1} (f_M - A_M v^M), \quad (4)$$

where  $v^M$  is the most recent iterate on grid  $G^M$ . To solve (4), recursion is employed, taking  $ID$  relaxation sweeps on grid  $G^k$  before visiting grid  $G^{k-1}$ ,  $M-1 \geq k \geq 2$ , and the equation

$$A_{k-1} V_{k-1} = f_{k-1} = I_k^{k-1} (f_k - A_k v_k).$$

$A_1 V_1 = f_1$  is solved directly and  $v_2 \leftarrow v_2 + I_1^2 V_1$  is performed. Then  $IU$  relaxation sweeps are performed on grid  $G^{k-1}$ , before performing  $v_k \leftarrow v_k + I_{k-1}^k V_{k-1}$ ,  $3 \leq k \leq M$ , and  $IU$  relaxation sweeps on (2).

## 2.3. BoxMG: coarsening by two

In [1] the sequence of grids,  $G^{k-1}$ ,  $k = M, \dots, 1$ , is obtained by standard coarsening, choosing every other  $x$ -point of every other logical  $y$ -grid-line, of  $G^k$ . A schematic of standard coarsening in the vertex-centered case is shown in Figure 1 (left), where the coarse grid,  $G^{M-1}$ , is illustrated by dots. An important feature of the BoxMG algorithm is that it handles grids of any dimension, as opposed to many geometric coarsening implementations that require grids of the form  $2^M + 1$ . This flexibility is shown schematically in Figure 1 (right) where standard coarsening is applied to the dual-grid of cell-centers (shown as dots). Here the coarse-grid, which is shown as squares, does not contain the right and the top boundary points; yet the performance of BoxMG is not degraded.

*Galerkin coarse-grid operator:* The coarse-grid operators  $A_{k-1}$ ,  $k = M, \dots, 2$ , are defined by

$$A_{k-1} = I_k^{k-1} A_k I_{k-1}^k, \quad (5)$$

where  $I_k^{k-1} = (I_{k-1}^k)^T$ , and where  $I_{k-1}^k$  is defined by operator-induced interpolation, described as follows.

*Operator-induced interpolation:* Perhaps the earliest example of operator-induced interpolation is [21]. The examples in that paper use the Laplace operator and exploit the discrete 5-point approximation and the discrete rotated 5-point approximation to recursively solve exactly difference equations to obtain approximations on finer grids. It is not clear how to employ such an approach for BoxMG. To define the operator-induced interpolation used in BoxMG when coarsening by two and to motivate its extension to coarsening by three, we first write the standard coarse/fine (C/F) splitting in block form,

$$A_h u = \begin{bmatrix} A_{ff} & A_{fc} \\ A_{fc}^T & A_{cc} \end{bmatrix} \begin{bmatrix} u_f \\ u_c \end{bmatrix} = \begin{bmatrix} f_f \\ f_c \end{bmatrix} = f, \quad (6)$$

where we have used the symmetry of the matrix. Here the intergrid transfer operators may be written as

$$I_{k-1}^k = \begin{bmatrix} I_{fc} \\ I_{cc} \end{bmatrix}, \quad I_k^{k-1} = (I_{k-1}^k)^T = [I_{fc}^T | I_{cc}], \quad (7)$$

with  $I_{cc}$  the identity. The *perfect* smoothing and intergrid transfer operators lead to a direct solver, but are obviously impractical. Nevertheless, if we define

$$I_{fc} = -A_{ff}^{-1} A_{fc}, \quad (8)$$

then the Galerkin operator is the Schur complement and we can construct an impractical but direct method (see, e.g. [5])

To derive the operator-induced interpolation used in BoxMG, we divide the fine-points into two types: fine-points that are embedded in coarse-grid lines ( $\gamma$ ) and fine-points that are in the interior of a coarse-grid cell ( $\iota$ ). This division of points is shown in Figure 2. Then the original system of equations can be written in the form

$$A_h u = \left[ \begin{array}{cc|c} A_{\iota\iota} & A_{\iota\gamma} & A_{\iota c} \\ A_{\iota\gamma}^T & A_{\gamma\gamma} & A_{\gamma c} \\ \hline A_{\iota c}^T & A_{\gamma c}^T & A_{cc} \end{array} \right] \begin{bmatrix} u_\iota \\ u_\gamma \\ u_c \end{bmatrix} = \begin{bmatrix} f_\iota \\ f_\gamma \\ f_c \end{bmatrix} = f. \quad (9)$$

Next we consider the  $\gamma$  points, which we want to interpolate directly from the neighboring C points along the edge of a coarse-grid cell. Specifically, using the averaging approach discussed below we will remove the connections from  $\gamma$  points to  $\iota$  points and generate new weights for the remaining connections to C points. Thus, the approximate system that defines the interpolation may be written as

$$A_h u = \left[ \begin{array}{cc|c} A_{\iota\iota} & A_{\iota\gamma} & A_{\iota c} \\ 0 & \widehat{A}_{\gamma\gamma} & \widehat{A}_{\gamma c} \\ \hline A_{\iota c}^T & A_{\gamma c}^T & A_{cc} \end{array} \right] \begin{bmatrix} u_\iota \\ u_\gamma \\ u_c \end{bmatrix} = \begin{bmatrix} f_\iota \\ f_\gamma \\ f_c \end{bmatrix} = f. \quad (10)$$

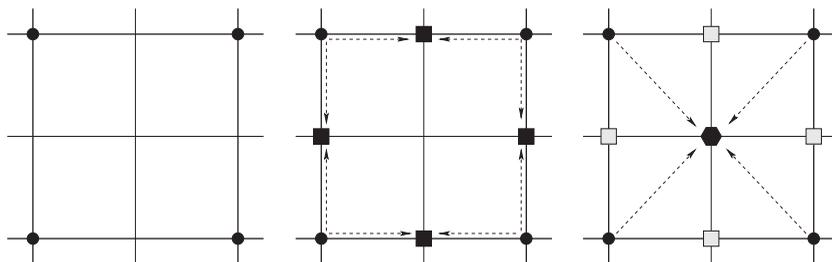


Figure 2. Schematic of interpolation design.

Now the block involving fine-grid unknowns (of both types) is block upper triangular, and moreover the diagonal blocks are themselves diagonal. Thus, the interpolation is readily defined by exactly inverting the homogeneous equation associated with this block,

$$\widehat{I}_{fc} = - \begin{bmatrix} A_{ii} & A_{i\gamma} \\ 0 & \widehat{A}_{\gamma\gamma} \end{bmatrix}^{-1} \begin{bmatrix} A_{ic} \\ \widehat{A}_{\gamma c} \end{bmatrix} = -\widehat{A}_{ff}^{-1} \widehat{A}_{fc}, \tag{11}$$

and by replacing  $I_{fc}$  in (7) by  $\widehat{I}_{fc}$ . Thus, the operator-induced interpolation in BoxMG is multipass interpolation [5] that takes advantage of the nearest neighbor stencil on a structured grid. Moreover, the only approximation in the interpolation is made along logical coarse-grid lines.

To develop the interpolation weights for the points embedded in the coarse-grid lines, we average the fine-grid stencil over the dimension transverse to each line. For example, consider writing the fine-grid stencil at a  $\gamma$  point on an  $x$ -line in compass-based notation,

$$\begin{bmatrix} -NW & -N & -NE \\ -W & O & -E \\ -SW & -S & -SE \end{bmatrix}_{2i+1,j},$$

where  $(2i+1, j)$  are fine-grid indices and where we have ignored symmetry to simplify the discussion. Now we assume that the error is constant in the transverse direction (i.e.  $y$ ) and varies along the coarse-grid line (i.e.  $x$ ). This leads to the one-dimensional stencil,

$$[-(W+SW+NW) \quad (O-S-N) \quad -(E+SE+NE)]_{2i+1,j}, \tag{12}$$

and defines the corresponding entries in the matrices,  $\widehat{A}_{\gamma\gamma}$  and  $\widehat{A}_{\gamma c}$ . If the fine-grid dimension in  $x$  is even, then (12) provides an extrapolation formula, and the Galerkin coarse-grid operator (5) still provides meaningful coefficients. Similarly, we may derive the approximations for the  $\gamma$  points embedded in  $y$ -lines.

Although, the definition of the interpolation seems complete, there are few points worth noting. First, this approximation has preserved the symmetry of the operator. Hence, defining the integrated stencil weights,

$$\overline{W} = W + SW + NW, \quad \overline{O} = O - S - N \quad \text{and} \quad \overline{E} = E + SE + NE,$$

and rewriting (12) as

$$[-\overline{W} \quad \overline{O} \quad -\overline{E}]_{2i+1,j}, \quad (13)$$

it is apparent that this is a one-dimensional diffusion equation that implicitly contains the discontinuous coefficients from the original problem. With respect to this one-dimensional discrete system, it is well known that the *perfect* interpolation is readily defined by  $[\widehat{A}_{\gamma\gamma}]^{-1}\widehat{A}_{\gamma c}$ .

Second, in the case of pure diffusion (i.e.  $\sigma=0$ ), this is a zero sum operator at interior points and on Neumann boundaries. Here we have  $\overline{O}=\overline{W}+\overline{E}$ , and the interpolation weights clearly preserve the constant or near null-space of the operator. In fact, we could substitute this in (13) to define the interpolation by,

$$[-\overline{W} \quad (\overline{W}+\overline{E}) \quad -\overline{E}]_{2i+1,j}. \quad (14)$$

Also, in the case  $\sigma=0$  this interpolation preserves the continuity of the normal flux [7], and does not make assumptions about the continuity of the gradient ( $x$  derivative). Moreover, in conjunction with Galerkin coarsening it effectively introduces the harmonic mean of the diffusion coefficient when coarsening across discontinuities, which corresponds with well-established results in homogenization [22].

Finally, we consider which of these definitions is the best when  $\sigma \neq 0$ , as this is effectively the case for boundary points where there is a Dirichlet or mixed boundary condition as well. We have found examples when one definition performs much better than the other, and the heuristics of which to choose have evolved over time. Letting  $\omega = \overline{W} + \overline{E}$ , the current implementation defines the diagonal weight as

$$[\text{diag}(\widehat{A}_{\gamma\gamma})]_{2i+1,j} = \begin{cases} \overline{O} & O > (1+\varepsilon)\omega, \\ \omega & O \leq (1+\varepsilon)\omega, \end{cases} \quad (15)$$

where  $\varepsilon = \min(|\overline{W}|/O, |\overline{E}|/O)$ . Note that if the points at the right boundary are extrapolated (i.e. if the fine-grid dimension  $x$  is even), then this definition uses  $\overline{O}$ , even if  $\sigma \equiv 0$ . Analogous behavior is defined for  $\gamma$  points on a  $y$ -line, and the interior points,  $i$ . Here the switch for interior points is necessary to ensure that a consistent approximation is made in  $A_{ii}$ . Furthermore, with extrapolation cases it was noted in [1] that convergence is helped by replacing  $u_k \leftarrow u_k + I_{k-1}^k u_{k-1}$  with

$$u_k \leftarrow u_k + I_{k-1}^k u_{k-1} + (f^k - A_k u_k)/O \quad (16)$$

at F points on the boundary. In effect, (16) adds on the result of a Jacobi relaxation (cf. [5]), for little cost, since the residual on  $G^k$  has already been computed and stored. Hence, BoxMG uses this Jacobi step at all F points.

*Smoothing:* With standard coarsening selected *a priori* the complementarity of coarse-grid corrections and smoothing hinges on the selection of appropriate smoothers. For problems with isotropic diffusion (i.e.  $D$  is a scalar) and uniform grids, red-black Gauss-Seidel has proved to be a robust and efficient smoother. For problems with anisotropic diffusion, we need a more powerful block smoother and use alternating red-black (zebra) line relaxation.

## 3. BoxMG: COARSENING BY THREE

Not long after the publication of [1], discussions with Blair Swartz (private communication) led the first author to consider another possibility, but only with this paper has this idea been implemented. That possibility is to employ coarsening by a factor of three, which in the case of cell-centered unknowns is also nested coarsening by cells. This cell-based coarsening by three is shown schematically in Figure 3 (right), where the coarse-grid cells are shown by the dark lines and coarse-grid cell-center unknowns are shown as the dark dots. Naturally, coarsening by three can be applied to a vertex-centered discretization, as is shown schematically in Figure 3 (left). In addition, the proposed BoxMG algorithm does not place any restrictions on the dimension of the finest grid.

*Galerkin coarse-grid operator:* The coarse-grid operators  $A_{k-1}$ ,  $k = M, \dots, 2$ , are defined by (5), where  $I_k^{k-1} = (I_{k-1}^k)^T$  and where  $I_{k-1}^k$  is the operator-induced coarsening derived below. In the coarsening by two case, the computation of this triple matrix product (5) was coded explicitly, avoiding all unnecessary multiplication by zeros. However, this is not practical in this case, and instead we use a local blocking technique [23]. The entries in the coarse-grid operator are calculated at a grid point  $(i, j)$ , by applying  $I_k^{k-1} A_k I_{k-1}^k$  to an identity vector  $e_{k-1}$ , which is defined as being 1 at  $(i, j)$  and 0 at all other grid points. Since the grid is structured, the extent of the nonzeros is known *a priori*, and all calculations can be performed in small local arrays. This approach is reasonably efficient and readily parallelized.

*Operator-induced interpolation:* To develop the operator-induced interpolation for coarsening by three, we generalize the approach developed in the preceding section for coarsening by two. We consider the same splitting of points, first into coarse and fine-points, as shown in Figure 4 (left). This splitting does not change the form of the interpolation given in (7), and we still take  $I_{cc}$  to be the identity. Similarly, the use of the *perfect*  $I_{fc}$  is impractical and we seek a suitable approximation.

As before, we divide the fine-points into two types: fine-points that are embedded in coarse-grid lines ( $\gamma$ , shown as small boxes in Figure 4) and fine-points that are in the interior of a coarse-grid cell ( $\iota$ , shown as small hexagons in Figure 4). We then consider the averaging approach presented

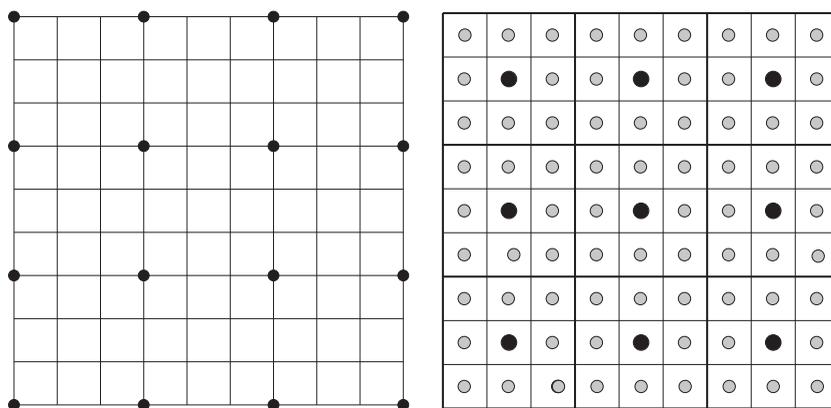


Figure 3. Standard coarsening of vertices (left) and the dual-grid of cell-centers (right).

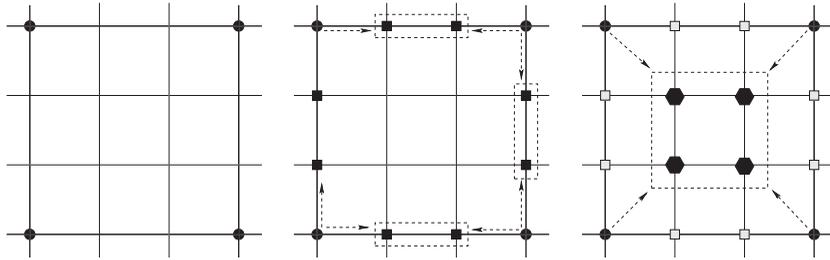


Figure 4. Schematic of interpolation for coarsening by three.

in the previous section to remove the connections from  $\gamma$  points to  $\iota$  points, and generate new weights for the remaining connections to C points. The block form of the approximate system remains the same as before (10), although we rewrite it here as

$$A_H u = \begin{bmatrix} A_{II} & A_{I\gamma} & A_{Ic} \\ 0 & \widehat{A}_{\gamma\gamma} & \widehat{A}_{\gamma c} \\ \hline A_{Ic}^T & A_{\gamma c}^T & A_{cc} \end{bmatrix} \begin{bmatrix} u_I \\ u_\gamma \\ u_c \end{bmatrix} = \begin{bmatrix} f_I \\ f_\gamma \\ f_c \end{bmatrix} = f, \tag{17}$$

because the dimensionality of the blocks has changed. In particular, in the coarsening-by-two case  $\widehat{A}_{\gamma\gamma}$  is a diagonal matrix, but here it is block diagonal, with each block a  $2 \times 2$  matrix. This is shown schematically in Figure 4 (middle) by the dashed lines that surround the two  $\gamma$  points on each line. Similarly, in the coarsening-by-two case,  $A_{II}$  is a diagonal matrix, but here it is block diagonal, with each block a  $4 \times 4$  matrix. This is shown schematically in Figure 4 (right) where the dashed line surrounds the four  $\iota$  points. Thus, it is still possible to define the multipass interpolation by exactly inverting the homogeneous equation associated with these blocks

$$\widehat{I}_{fc} = - \begin{bmatrix} A_{II} & A_{I\gamma} \\ 0 & \widehat{A}_{\gamma\gamma} \end{bmatrix}^{-1} \begin{bmatrix} A_{Ic} \\ \widehat{A}_{\gamma c} \end{bmatrix} = -\widehat{A}_{ff}^{-1} \widehat{A}_{fc} \tag{18}$$

and replacing  $I_{fc}$  with  $\widehat{I}_{fc}$  in (7).

Once again to develop the interpolation weights for the points embedded in the coarse-grid lines, we average the fine-grid stencils over the dimension transverse to each coarse-grid line. For example, consider a  $\gamma$  point on an  $x$ -line, and assume that the error is constant in the transverse direction (i.e.  $y$ ), and varies along the coarse-grid line (i.e.  $x$ ). This leads to the approximate one-dimensional stencils

$$\begin{bmatrix} -\overline{W}_{3i+1,j} & \overline{O}_{3i+1,j} & -\overline{E}_{3i+1,j} & 0 \\ 0 & -\overline{W}_{3i+2,j} & \overline{O}_{3i+2,j} & -\overline{E}_{3i+2,j} \end{bmatrix} \tag{19}$$

and defines the corresponding blocks in the matrices,  $\widehat{A}_{\gamma\gamma}$ , and  $\widehat{A}_{\gamma c}$ . Similarly, we may derive the approximations for the  $\gamma$  points embedded in  $y$ -lines. Although the cases for extrapolation are significantly more frequent here than in the coarsening-by-two case, the Galerkin coarse-grid operator still performs well. However, we did find it necessary to modify the smoothing operator, as discussed below.

At this point the definition of the interpolation is essentially complete, but a few observations and modifications motivated by the coarsening-by-two case (Section 2.3) are worth noting. First, the averaging preserves the symmetry of the operator, and hence leads to a one-dimensional diffusion operator that implicitly contains the discontinuous coefficients of the original problem. Moreover, for this one-dimensional case the *perfect* interpolation is readily defined by  $[\widehat{A}_{\gamma\gamma}]^{-1}\widehat{A}_{\gamma c}$ .

Second, in the case of pure diffusion (i.e.  $\sigma=0$ ), this is a zero sum operator at interior points and on Neumann boundaries. Here, as with the coarsening-by-two case, we have  $\overline{O} = \overline{W} + \overline{E}$ , and the interpolation weights clearly preserve the constant or near null-space of the operator. Thus, we could substitute this into (19) to define interpolation by

$$\begin{bmatrix} -\overline{W}_{3i+1,j} & (\overline{W} + \overline{E})_{3i+1,j} & -\overline{E}_{3i+1,j} & 0 \\ 0 & -\overline{W}_{3i+2,j} & (\overline{W} + \overline{E})_{3i+2,j} & -\overline{E}_{3i+2,j} \end{bmatrix}. \quad (20)$$

Hence, in this case it can be shown that this interpolation preserves the continuity of the normal flux for this one-dimensional diffusion problem. In addition, through the Galerkin coarsening it introduces the harmonic mean of the diffusion coefficient when coarsening across discontinuities.

Finally, we consider which of these definitions is the best when  $\sigma \neq 0$ . As noted in Section 2.3, an effectively non-zero removal term appears at boundary points where there is a Dirichlet or mixed boundary condition as well as at points where  $\sigma \neq 0$ . Here, we propose to use the switch introduced in (15), for each point in the interpolation formula. Specifically, letting  $\omega = (\overline{W} + \overline{E})$ , we define

$$[\text{diag}(\widehat{A}_{\gamma\gamma})]_{3i+1,j} = \begin{cases} \overline{O}_{3i+1,j}, & O > [(1+\varepsilon)\omega]_{3i+1,j}, \\ \omega_{3i+1,j}, & O \leq [(1+\varepsilon)\omega]_{3i+1,j}, \end{cases} \quad (21)$$

where  $\varepsilon = \min(|\overline{W}|/O, |\overline{E}|/O)$ . Shifting this definition to the point  $(3i+2, j)$ , we complete the definition of the  $2 \times 2$  blocks of  $\widehat{A}_{\gamma\gamma}$  for  $\gamma$  points on an  $x$ -line. Analogous formulae are defined for  $\gamma$  points on a  $y$ -line, as well as the interior points,  $i$ . Here the switch for interior points is necessary to ensure that a consistent approximation is made in  $A_{ii}$ . Furthermore, in coarsening by three, as in coarsening by two, we embed the essentially free Jacobi step in the interpolation (16) as it improves convergence.

It is interesting to note that in the coarsening-by-three case the interpolation without the switch, based solely on (19) and its analogous formulae for the remaining fine-grid points, is considerably more robust than the corresponding formulae for the coarsening-by-two case. This robustness may be related to the robustness encountered in semicoarsening, [24, 25]. However, an example is presented in Section 4.4 that demonstrates the necessity of (21).

In addition, coarsening by a factor of  $n$ ,  $n > 3$ , could be considered; the diagonal blocks of  $A_{\gamma\gamma}$  would be  $(n-1) \times (n-1)$  matrices, and the diagonal blocks of  $A_{ii}$  would be  $(n-1)^2 \times (n-1)^2$  matrices. Although this is a costly endeavor, a two-level version of this coarsening procedure lies at the heart of the popular MsFEM [16]. But as the coarsening factor grows, it is not just the cost of defining the interpolation that raises concern, but the growing size of the one-dimensional problems and their impact on the approximation subspace. Thus, it is not surprising that the greatest efficiency is achieved with smaller coarsening factors. In fact, Brandt's earliest work [13] notes experimental evidence that coarsening by a factor of two is nearly optimal and should be used.

*Smoothing:* As with the coarsening-by-two case, standard coarsening places the burden of complementarity on the smoother. Once again we use red-black Gauss-Seidel for isotropic

diffusion, and alternating red–black lines for anisotropic diffusion. However, cell-based coarsening creates a new problem at the boundary, leaving two lines of points extrapolated at the right and the top boundaries, if the dimension of the fine-grid in that direction,  $x$  or  $y$ , respectively, is  $3m + 1$ . Although this does not cause any problem in the formation of the Galerkin coarse-grid operator, clearly the aggregate loss of accuracy in the correction can be significant. On these particular grids, we perform a relaxation sweep of these two extrapolated lines, before proceeding with the relaxation sweep of the entire grid.

We also consider an alternative smoothing for isotropic problems, which we call pattern relaxation. The idea is to perform relaxation sweeps compatible with the coarsening procedure, and hence, compatible with the structure of the interpolation operator. Specifically, since we consider at most a 9-point nearest neighbor stencil, the connectivity of the blocks used in the interpolation suggests a corresponding four-color block Gauss–Seidel relaxation scheme. In sweep one a point Gauss–Seidel sweep is performed on all the coarse-grid points ( $C$  points). In sweep two a block Gauss–Seidel sweep is performed on the four fine-grid points in the interior of each coarse-grid rectangle ( $i$  points). In the third (and fourth) sweep a block horizontal (vertical) Gauss–Seidel sweep is performed on the pairs of fine-points that lie on coarse-grid lines ( $\gamma$  points). If the  $2 \times 2$  and  $4 \times 4$  LU decompositions are pre-computed and stored, then the computational cost of pattern relaxation is very similar to point Gauss–Seidel relaxation. Experimentally (see Section 4.1), the convergence factor of  $V(1, 1)$  cycles with pattern relaxation falls in between the convergence factors of  $V(1, 1)$  and  $V(2, 2)$  cycles with colored point Gauss–Seidel relaxation. Note that pattern relaxation is potentially attractive for coarsening by a factor of large  $n$  on certain architectures.

#### 4. NUMERICAL EXAMPLES

The objective of the numerical examples is to demonstrate the robustness of the proposed BoxMG method with respect to discontinuous diffusion and removal coefficients,  $D$  and  $\sigma$ , respectively, different types of boundary conditions, and grid dimensions that are not optimal multiples of three. In all examples,  $V(1, 1)$  or  $V(2, 2)$  cycles are used, and a random initial guess for  $u_M$  is employed. Unless otherwise stated, the iteration is terminated when the relative  $\ell^2$  residual on  $G^M$  is less than  $10^{-6}$ . In the tables we report the average convergence factor,  $\rho_A$ , and the convergence factor of the last iteration,  $\rho_L$ . These convergence factors are defined as

$$\rho_A = \left( \frac{\|r_k^L\|}{\|r_k^0\|} \right)^{1/L}, \quad \rho_L = \left( \frac{\|r_k^L\|}{\|r_k^{L-1}\|} \right),$$

where  $L$  denotes the index of the last iteration.

##### 4.1. Poisson

In this first example, we consider Poisson’s equation

$$\Delta U = 0 \quad \text{in } \Omega = (0, 1) \times (0, 1), \quad (22a)$$

$$\frac{\partial U}{\partial \nu} = 0 \quad \text{on } \partial\Omega, \quad (22b)$$

because this simple elliptic problem provides an ideal setting to explore the robustness of the proposed algorithm with respect to grid dimensions that are not an optimal multiple of three. In addition, we consider standard 5-point discretizations that are either cell-centered or vertex-centered.

First, we consider red–black point Gauss–Seidel relaxation for both coarsening-by-two and coarsening-by-three BoxMG algorithms. In Table I, we highlight the performance of the existing coarsening-by-two BoxMG algorithm applied directly to the dual-grid of cell-centers. Although the grid is not optimal for coarsening by two excellent performance is still achieved, with  $\rho_L = 0.10$  for the largest grid considered. Next, in Table II, we summarize the performance of the proposed algorithm for cell-centered coarsening by a factor of three. In the first grouping, the grid dimension is of the form  $3m$ , and in the last it is  $3m + 2$ . Performance in these two cases is very similar, with  $V(1, 1)$  cycles giving  $\rho_L \approx 0.3$ , and with  $V(2, 2)$  cycles giving  $\rho_L \approx 0.09$ . As noted in the previous discussion of smoothing for coarsening by three, the  $3m + 1$  case is the most challenging because it leads to extrapolation of two lines of points on the right and the top boundaries. The second grouping in Table II is marked with an asterisk as this shows the degradation in performance that results if extra relaxation is not performed on these extrapolated points. In the third grouping, the extra relaxation on these extrapolated points is performed and the convergence is restored.

Next, we investigate the performance of the *pattern relaxation* that we introduced in the smoothing discussion of Section 3. The convergence rates of the coarsening-by-three BoxMG algorithm with pattern relaxation for a 5-point cell-centered discretization of Poisson are summarized in Table III. These results are much better than the red–black point Gauss–Seidel (Table II), with  $V(1, 1)$  cycles giving  $\rho_L \approx 0.1$  on the favorably dimensioned grids, and  $\rho_L \approx 0.2$  on the unfavorable  $3m + 1$  grids. This improvement in convergence rate is sufficient to justify the additional work and storage of the pattern relaxation. A similar improvement in convergence rates was observed for the isotropic examples that follow. Hence, in subsequent isotropic examples we present only the pattern relaxation results.

Coarsening by three can also be applied to vertex-centered discretizations. Since the results for (22) are about the same as in Table III, they are not presented here. In both the cell-centered and vertex-centered cases, convergence factors for  $V(1, 1)$  cycles are better for coarsening by two than by three. In [13] Brandt made the observation that coarsening by two is the most efficient coarsening factor in terms of convergence per work unit. However, the relaxation work on coarser grids for  $V(2, 2)$  cycles is less than the work on coarser grids for standard BoxMG with  $V(1, 1)$  cycles,  $\frac{1}{4}$  versus  $\frac{1}{3}$ . In the  $V(2, 2)$  case for coarsening by three, if a total of two relaxation sweeps, instead of four, were performed on the finest grid, then the work would be comparable to the work performed for a  $V(1, 1)$  cycle in the coarsening-by-two case; unfortunately,  $\rho_L$  for this variation is

Table I. Convergence of BoxMG for Poisson (22), using standard coarsening by two, red–black point Gauss–Seidel, and  $V(1, 1)$  cycles on the dual-grid of cell-centers.

Grid size	$\rho_A$	$\rho_L$
$8 \times 8$	0.070	0.112
$16 \times 16$	0.058	0.111
$32 \times 32$	0.062	0.120
$64 \times 64$	0.057	0.114
$128 \times 128$	0.054	0.106
$256 \times 256$	0.051	0.100

Table II. Convergence of BoxMG for Poisson (22), coarsening cell-centers by a factor of three, and using red–black point Gauss–Seidel relaxation.

Grid size	CF=3, V(1, 1)		CF=3, V(2, 2)	
	$\rho_A$	$\rho_L$	$\rho_A$	$\rho_L$
9 × 9	0.199	0.259	0.045	0.063
27 × 27	0.208	0.288	0.051	0.090
81 × 81	0.226	0.296	0.055	0.094
243 × 243	0.225	0.299	0.055	0.094
10 × 10*	0.383	0.426	0.131	0.216
28 × 28*	0.305	0.425	0.100	0.208
82 × 82*	0.307	0.432	0.115	0.219
244 × 244*	0.296	0.437	0.096	0.219
10 × 10	0.177	0.284	0.051	0.110
28 × 28	0.227	0.298	0.056	0.101
82 × 82	0.229	0.306	0.058	0.105
244 × 244	0.227	0.302	0.057	0.099
11 × 11	0.160	0.224	0.031	0.056
29 × 29	0.198	0.289	0.045	0.088
83 × 83	0.211	0.296	0.050	0.090
245 × 245	0.213	0.298	0.050	0.091

Note the ‘\*’ marks the case of a  $(3m+1) \times (3m+1)$  grid without extra relaxation on extrapolated points at the right and the top boundaries.

Table III. Convergence of BoxMG for Poisson (22), coarsening cell-centers by a factor of three, and using *pattern relaxation*.

Grid size	CF=3, V(1, 1)		CF=3, V(2, 2)	
	$\rho_A$	$\rho_L$	$\rho_A$	$\rho_L$
9 × 9	0.083	0.110	0.007	0.014
27 × 27	0.063	0.102	0.014	0.042
81 × 81	0.070	0.103	0.009	0.029
243 × 243	0.068	0.102	0.009	0.029
10 × 10	0.151	0.244	0.031	0.085
28 × 28	0.135	0.238	0.008	0.075
82 × 82	0.121	0.231	0.021	0.069
244 × 244	0.119	0.239	0.022	0.091
11 × 11	0.070	0.101	0.009	0.025
29 × 29	0.062	0.099	0.009	0.025
83 × 83	0.063	0.101	0.009	0.025
245 × 245	0.063	0.100	0.009	0.024

about the same as  $\rho_L$  for V(1,1), leading to the conclusion that the extra relaxation sweeps on the finest grid are necessary for convergence factors comparable to the V(1, 1), CF=2 case. For this example, the extra work pays off. A simple counting argument shows that V(2, 2) with CF=3 and

pattern relaxation costs about 1.7 times as much  $V(1, 1)$  with  $CF=2$  and color point-Gauss–Seidel. Thus, if the ratio of the convergence factors of the former to the convergence factors of the latter is less than 1.7, the former is competitive with the latter. But for the non-constant coefficient examples below, the results are not as good, and the former is not quite competitive with the latter.

#### 4.2. A thin layer

In the next example a thin layer is considered, to explore robustness with respect to a discontinuous diffusion coefficient. This example is the most challenging of the family of layer problems studied by Khalil and Wesseling in [15]. In fact, it was originally suggested by Brandt as it is a simple example that forces coarsening across discontinuities. Wesseling's approach required W-cycles to achieve a convergence rate that depended weakly on  $h$ . The problem is

$$\nabla \cdot (D \nabla U) = 0 \quad \text{in } \Omega = (0, 1) \times (0, 1), \quad (23a)$$

$$\frac{\partial U}{\partial \nu} = 0 \quad \text{on } \partial \Omega, \quad (23b)$$

where  $D$  has the values shown in Figure 5. The results for standard BoxMG and the results for BoxMG with coarsening by a factor of three are summarized in Table IV.

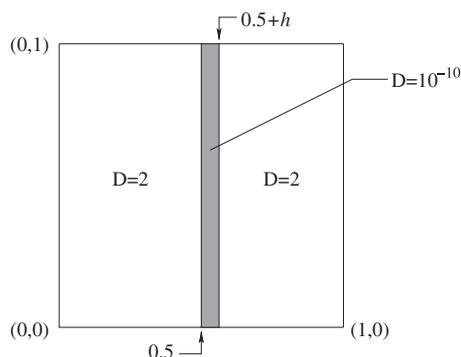


Figure 5. The thin layer suggested by Brandt forces coarsening across the discontinuities.

Table IV. Convergence of BoxMG for the thin layer (23), coarsening factors of two and three.

Grid size	CF=2, $V(1, 1)$		CF=3, $V(1, 1)$		CF=3, $V(2, 2)$	
	$\rho_A$	$\rho_L$	$\rho_A$	$\rho_L$	$\rho_A$	$\rho_L$
$8 \times 8$	0.113	0.173	0.169	0.250	0.028	0.062
$16 \times 16$	0.072	0.126	0.153	0.241	0.030	0.082
$32 \times 32$	0.073	0.125	0.161	0.246	0.040	0.089
$64 \times 64$	0.061	0.117	0.131	0.234	0.025	0.072
$128 \times 128$	0.056	0.110	0.123	0.236	0.023	0.075
$256 \times 256$	0.053	0.106	0.124	0.267	0.023	0.090

### 4.3. The Checkerboard problem

The next example is for a checkerboard problem:

$$\nabla \cdot (D \nabla U) = 0 \quad \text{in } \Omega = (0, 1) \times (0, 1), \quad (24a)$$

$$\frac{\partial U}{\partial \nu} = 0 \quad \text{on } \partial \Omega, \quad (24b)$$

where  $D$  has the values shown in Figure 6. The checkerboard problem is one of the standard test cases for the neutron transport and the neutron diffusion equations, and as such, was considered for vertex-centered differencing in [2]. The parameter  $a$  shifts the location of the cross (singularity), and hence, controls at what resolution it will be coarsened away. For the case  $a=0$ , Table V shows the results for BoxMG with standard coarsening by a factor of two on the dual mesh, and for cell-centered coarsening by a factor of three. Results are similar for the case that  $a=h$ , and are shown in Table VI.

### 4.4. Mixed boundary conditions

In this section, we consider two simple examples with vacuum boundary conditions that illustrate the necessity of the weighting switch developed in Section 3. In fact, we have three possible definitions of the interpolation to consider. The first defines the weights by the direct averaging of

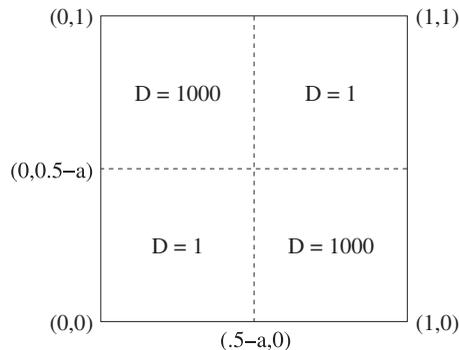


Figure 6. The classic checkerboard pattern.

Table V. BoxMG for (24), coarsening factors of two and three, cell-centered,  $a=0$ .

Grid size	CF=2, $V(1, 1)$		CF=3, $V(1, 1)$		CF=3, $V(2, 2)$	
	$\rho_A$	$\rho_L$	$\rho_A$	$\rho_L$	$\rho_A$	$\rho_L$
$8 \times 8$	0.057	0.093	0.131	0.181	0.019	0.026
$16 \times 16$	0.073	0.121	0.168	0.239	0.043	0.100
$32 \times 32$	0.075	0.127	0.136	0.231	0.027	0.081
$64 \times 64$	0.058	0.114	0.139	0.237	0.045	0.167
$128 \times 128$	0.056	0.110	0.124	0.231	0.023	0.076
$256 \times 256$	0.055	0.114	0.123	0.245	0.021	0.075

Table VI. BoxMG for (24), coarsening factors of two and three, cell-centers,  $a=h$ .

Grid size	CF=2, $V(1, 1)$		CF=3, $V(1, 1)$		CF=3, $V(2, 2)$	
	$\rho_A$	$\rho_L$	$\rho_A$	$\rho_L$	$\rho_A$	$\rho_L$
$8 \times 8$	0.038	0.075	0.083	0.095	0.014	0.024
$16 \times 16$	0.045	0.072	0.162	0.231	0.043	0.082
$32 \times 32$	0.059	0.101	0.081	0.140	0.016	0.056
$64 \times 64$	0.062	0.105	0.153	0.267	0.056	0.229
$128 \times 128$	0.069	0.109	0.072	0.114	0.015	0.046
$256 \times 256$	0.070	0.112	0.125	0.247	0.023	0.092

the operator, without considering whether the interpolation preserves a constant. We will refer to this as the *averaging* interpolation and it is given by (13), for CF=2, and by (19) for CF=3. The second uses a similar averaging of the operator, but explicitly assumes a zero-row-sum operator to enforce interpolation of a constant. This is dubbed the *constant-preserving* interpolation and is given by (14) for CF=2, and by (20) for CF=3. Finally, we can use a switch based on the operator to determine which of these interpolations, the *averaging* or the *constant-preserving*, should be used. This is simply referred to as the *switching* interpolation, and is defined by (15) for CF=2, and by (21) for CF=3.

To develop our examples, we consider standard 5-point cell-centered discretizations of

$$-\nabla \cdot (D\nabla U) = 0 \quad \text{in } \Omega = (0, X_R) \times (0, Y_T), \quad (25a)$$

$$D\nabla U \cdot \mathbf{n} = 0 \quad \text{on } \{0\} \times (0, Y_T) \cup (0, X_R) \times \{0\} \cup \{X_R\} \times (0, Y_T), \quad (25b)$$

$$D\nabla U \cdot \mathbf{n} + \frac{1}{2}U = 0 \quad \text{on } (0, X_R) \times \{Y_T\}, \quad (25c)$$

with  $D$  a diagonal tensor and the associated matrix an M-matrix. In this case the discretization of the vacuum boundary condition leads to a diagonally dominant stencil along the top boundary, akin to a non-zero removal term. The strength of the diagonal dominance depends on the mesh spacing,  $h$ , and the  $y$ -component of the diffusion tensor,  $D_{yy}$ . Now, to demonstrate the need for the switched interpolation we consider two cases that explore this dependence. First we consider isotropic diffusion on a large domain,

$$D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad X_R = Y_T = 128. \quad (26)$$

Here,  $h \gg D_{yy} = 1$ , making the stencil at the top boundary strongly diagonally dominant and making the use of constant-preserving interpolation suspect. In fact, the results in Table VII show that for both coarsening by two and coarsening by three, use of constant-preserving interpolation leads to slow convergence. In contrast, Table VIII shows that using the averaging interpolation leads to good results for this case. In this case, there is no difference between the convergence rates of the averaging and switching interpolation; hence we do not explicitly show the latter.

Table VII. Convergence of BoxMG for the large-domain problem (26), with *constant-preserving* interpolation.

Grid size	CF=2, $V(1, 1)$		CF=3, $V(1, 1)$	
	$\rho_A$	$\rho_L$	$\rho_A$	$\rho_L$
$8 \times 8$	0.305	0.374	0.072	0.101
$16 \times 16$	0.417	0.532	0.444	0.555
$32 \times 32$	0.483	0.620	0.063	0.099
$64 \times 64$	0.500	0.721	0.433	0.575
$128 \times 128$	0.495	0.698	0.307	0.588
$256 \times 256$	0.460	0.714	0.506	0.745

Table VIII. Convergence of BoxMG for the large-domain problem (26), with *averaging* interpolation.

Grid size	CF=2, $V(1, 1)$		CF=3, $V(1, 1)$	
	$\rho_A$	$\rho_L$	$\rho_A$	$\rho_L$
$8 \times 8$	0.037	0.055	0.116	0.148
$16 \times 16$	0.072	0.124	0.153	0.239
$32 \times 32$	0.062	0.129	0.133	0.203
$64 \times 64$	0.060	0.117	0.128	0.237
$128 \times 128$	0.058	0.114	0.096	0.206
$256 \times 256$	0.056	0.111	0.118	0.245

Second, we show an example for which averaging interpolation does not always give good results, and hence, the switching interpolation is necessary. Specifically, we consider a constant-coefficient anisotropic problem

$$D = \begin{bmatrix} 1 & 0 \\ 0 & 100 \end{bmatrix}, \quad X_R = Y_T = 1. \quad (27)$$

and employ  $y$ -line relaxation. Here, the problematic grid dimensions are somewhat different for coarsening by two and coarsening by three. Specifically, for coarsening by two the inadequacy of averaging interpolation is most apparent on grids that are of dimension  $2^M + 1$ , and is shown in Table IX. On these grids the line of cell-centers next to the top boundary, which has the diagonally dominant stencils, appears on all coarser levels. However, as the dimension of the fine mesh increases sufficiently,  $h \ll D_{yy}$ , the boundary stencils become essentially zero row sum. At this critical resolution,  $129 \times 129$  in Table IX, the averaging interpolation is essentially constant preserving, and good convergence rates are restored. In contrast, the good convergence results for the switching interpolation (see Table IX) are largely independent of the grid dimension.

Coarsening by three honors the cell structure and is most intuitively viewed as an aggregation strategy. On each successive level, the aggregates grow, and hence the cell-center migrates away from the boundary and the strength of the diagonal dominance is weakened. For this reason, averaging interpolation with CF=3 is generally more robust for this problem. However, we have found that grids of dimension  $3m+2$  may be problematic, as is shown in Table X. In particular, the averaging interpolation performs poorly for the  $8 \times 8$  and  $32 \times 32$  grids. Once again, as the

Table IX. Convergence of BoxMG for the anisotropic problem (27), with CF=2 and y-line relaxation.

Grid size	V(1, 1) and averaging		V(1, 1) and switching	
	$\rho_A$	$\rho_L$	$\rho_A$	$\rho_L$
$9 \times 9$	0.564	0.913	0.0001	0.0005
$17 \times 17$	0.553	0.954	0.003	0.014
$33 \times 33$	0.531	0.981	0.004	0.034
$65 \times 65$	0.502	0.994	0.005	0.045
$129 \times 129$	0.007	0.198	0.004	0.042
$257 \times 257$	0.006	0.084	0.005	0.045

Table X. Convergence of BoxMG for the anisotropic problem (27), with CF=3 and y-line relaxation.

Grid size	V(1, 1) and averaging		V(1, 1) and switching	
	$\rho_A$	$\rho_L$	$\rho_A$	$\rho_L$
$8 \times 8$	0.062	0.357	0.010	0.143
$17 \times 17$	0.009	0.176	0.005	0.054
$32 \times 32$	0.504	0.986	0.005	0.048
$65 \times 65$	0.006	0.069	0.005	0.060
$128 \times 128$	0.005	0.051	0.005	0.050
$257 \times 257$	0.005	0.074	0.005	0.046

dimension of the fine-mesh increases, the averaging interpolation approaches constant-preserving interpolation, and good convergence rates are restored. The convergence rates for the switching interpolation are also presented in Table X, and clearly show the robustness of this algorithm.

Thus, switching interpolation appears to be the most robust choice. There are more complicated examples for non-zero removal and vacuum boundary conditions, and with discontinuous  $D$ , but it is interesting that the above difficulties can be illustrated with such simple examples.

#### 4.5. Non-zero removal

The final example is Experiment No. 4 from [26], which develops an SA method. This anisotropic problem with a non-zero removal term may be written as

$$-\nabla \cdot (D\nabla u) + qu = f(x, y) \quad \text{in } \Omega = (0, 1) \times (0, 1), \quad (28a)$$

$$u = 0 \quad \forall (x, y) \in \partial\Omega, \quad (28b)$$

where the diffusion tensor is defined in quadrants of the domain by

$$D = \begin{bmatrix} a & 0 \\ 0 & 1/a \end{bmatrix}, \quad a(x, y) = \begin{cases} 10^{-2}, & (x, y) \in (0, 0.5) \times (0, 0.5), \\ 1, & (x, y) \in (0, 0.5) \times (0.5, 1), \\ 10^2, & (x, y) \in (0, 0.5) \times (0, 1). \end{cases} \quad (29)$$

Table XI. BoxMG for the Vanek problem.

$q$	CF=2, $V(1, 1)$		CF=3, $V(1, 1)$		CF=3, $V(2, 2)$	
	$\rho_A$	$\rho_L$	$\rho_A$	$\rho_L$	$\rho_A$	$\rho_L$
0.1	0.020	0.365	0.145	0.595	0.055	0.452
1.0	0.020	0.365	0.145	0.595	0.055	0.452
10	0.020	0.365	0.145	0.595	0.055	0.452

The discretization in [26] uses  $P1$  elements on a ‘regular grid’, with 160 000 nodes. Similarly, in this study we use  $P1$  elements on a triangular grid obtained by subdividing a rectangular grid with southwest–northeast diagonals, where the rectangular grid has mesh size  $\Delta x = \Delta y = \frac{1}{400}$ . The values we consider for the removal term,  $q=0.1$ ,  $q=1$ , and  $q=100$ , are those used in [26]. For regions with anisotropic coefficients, their method employs aggregation of elements that results in semi-coarsening of shape functions. Thus, the number of unknowns on the first coarse-grid for (28) should be at least  $(2)(\frac{3}{4}) + (\frac{1}{4}) = \frac{7}{4}$  times the number of unknowns on the finest grid. Since there is bound to be some overlap between the semi-coarsened anisotropic subregions and the standard-coarsened isotropic subregion, we will assume the number to be closer to 2 than to  $\frac{7}{4}$ . Because of the semi-coarsening aggregation strategy, Vanek *et al.* [26] is able to employ a pointwise pre-smoothing strategy of a sweep of forward point Gauss–Seidel followed by a backward sweep of point SOR with  $\omega=1.85$ . Similarly, the post-smoothing strategy consists of a backward sweep of point SOR with  $\omega=1.85$  followed by a forward sweep of point Gauss–Seidel. Thus for comparison purposes, we employ  $V(2,2)$  cycles with alternating red–black line relaxations. Neither the initial guess nor  $f$  is specified in [26]; hence, we take  $f \equiv 0$  and a random initial guess for  $U^M$ . We terminate the iteration when the relative  $\ell^2$  residual on  $G^M$  is less than  $10^{-5}$ , as in [26]. We also show results for a  $V(1,1)$  cycle. The results are shown in Table XI.

For  $V(2, 2)$ , the average convergence factor is better than the  $\rho_A=0.1$  reported in [26]. Since the method in that reference is an SA method, it is obviously able to handle more general problems than the method reported in this paper. For this particular example, a more careful comparison would require evaluating convergence factors versus actual work. More generality for the method in this paper could be attained by extending it to locally refined grids.

## 5. CONCLUSIONS

We have generalized BoxMG to allow for coarsening by a factor of three and have presented some numerical examples to compare it with the classic BoxMG. In the development of this generalization, we have included a new discussion on the motivation for the switching mechanism used in the interpolation operators, and we have introduced ‘pattern’ relaxation. This new colored block Gauss–Seidel scheme is motivated by the form of the interpolation operator and was demonstrated to be more effective than colored point Gauss–Seidel relaxation. The numerical examples demonstrate the robustness of the new algorithm with respect to discontinuous coefficients, and flexibility with respect to grid dimensions. However, when coarsening by a factor of three,  $V(2, 2)$  cycles are required to obtain a convergence rate that is better than  $V(1, 1)$  cycles when coarsening

by a factor of two, and for all except constant-coefficient problems, the extra work is not justified. Thus, if one does not require coarsening by cells, then there is no real justification for coarsening by a factor of three. But if coarsening by cells is desired, then BoxMG with coarsening by a factor of three has been shown to be an effective algorithm.

## ACKNOWLEDGEMENTS

This work was funded by the Department of Energy at Los Alamos National Laboratory under contracts DE-AC52-06NA25396 and the DOE Office of Science Advanced Computing Research (ASCR) program in Applied Mathematical Sciences. The authors thank the referees for their constructive comments that improved this paper.

## REFERENCES

1. Dendy JE. Black box multigrid. *Journal of Computational Physics* 1982; **48**:366–386.
2. Alcouffe RE, Brandt A, Dendy JE Jr, Painter JW. The multi-grid method for the diffusion equation with strongly discontinuous coefficients. *SIAM Journal on Scientific Computing* 1981; **2**(4):430–454.
3. Brandt A. *Multigrid Techniques: 1984 Guide with Applications to Fluid Dynamics*. The Weizmann Institute of Applied Science: Rehovot, Israel, 1984.
4. Brandt A, McCormick SF, Ruge JW. Algebraic multigrid (AMG) for sparse matrix equations. In *Sparsity and its Applications*, Evans DJ (ed.). Cambridge University Press: Cambridge, 1984.
5. Trottenberg U, Oosterlee C, Schüller A. Introduction to the AMG methodology for scalar applications, including many results and performance measurements. In *Multigrid*, Stüben K (ed.). Academic Press: New York, 2001; 413–532.
6. Vanek P, Brezina M, Mandel J. Convergence of algebraic multigrid based on smoothed aggregation. *Numerische Mathematik* 2001; **88**(3):559–579.
7. Moulton JD, Dendy JE, Hyman JM. The black box multigrid numerical homogenization algorithm. *Journal of Computational Physics* 1998; **141**:1–29.
8. Dendy JE Jr. Black box multigrid for nonsymmetric problems. *Applied Mathematics and Computation* 1983; **13**:261–283.
9. de Zeeuw PM. Matrix-dependent prolongations and restrictions in a black box multigrid solver. *Journal of Computational Physics* 1990; **33**(1):1–27.
10. Mandel J, Brezina M, Vanek P. Energy optimization of algebraic multigrid bases. *Computing* 1999; **62**(3):205–228.
11. Brezina M, Falgout R, MacLachlan S, Manteuffel T, McCormick S, Ruge J. Adaptive algebraic multigrid. *SIAM Journal on Scientific Computing* 2006; **27**(4):1261–1286.
12. Brezina M, Falgout R, MacLachlan S, Manteuffel T, McCormick S, Ruge J. Adaptive smoothed aggregation ( $\alpha$ SA) multigrid. *SIAM Review* 2005; **47**(2):317–346.
13. Brandt A. Multi-level adaptive solutions to boundary-value problems. *Mathematics of Computation* 1977; **31**:333–390.
14. Hackbusch W. *Multi-grid Methods and Applications*. Springer Series in Computational Mathematics, vol. 4. Springer: Berlin, Heidelberg, 1985.
15. Khalil M, Wesseling P. Vertex-centered and cell-centered multigrid for interface problems. *Journal of Computational Physics* 1992; **98**(1):1–10.
16. Hou T, Wu X, Cai Z. Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients. *Mathematics of Computation* 1999; **68**:913–943.
17. Alcouffe RE. Diffusion synthetic acceleration for the diamond-difference discrete-ordinates equations. *Nuclear Science and Engineering* 1977; **64**(2):344–355.
18. Brezzi F, Douglas J Jr, Fortin M, Marini LD. Efficient rectangular mixed finite elements in two and three space variables. *Mathematical Modelling and Numerical Analysis* 1987; **21**(4):581–604.
19. Kuznetsov Y, Lipnikov K, Shashkov M. The mimetic finite difference method on polygonal meshes for diffusion-type problems. *Computational Geosciences* 2004; **8**(4):301–324.
20. Brezzi F, Lipnikov K, Simoncini V. A family of mimetic finite difference methods on polygonal and polyhedral meshes. *Mathematical Models and Methods in Applied Sciences (M3AS)* 2005; **15**(10):1533–1551.

21. Hyman JM. Mesh refinement and local inversion of elliptic differential equations. *Journal of Computational Physics* 1977; **23**:124–134.
22. Bensoussan A, Lions JL, Papanicolaou G. *Asymptotic Analysis for Periodic Structures, Studies in Mathematics and its Applications*, vol. 5. North–Holland: New York, 1978.
23. Dendy JE, Moulton JD. Alternative RAP. *Abstracts for the Eighth Copper Mountain Conference on Iterative Methods*, Copper Mountain, CO, 28 March–2 April 2004.
24. Schaffer S. A semicoarsening multigrid method for elliptic partial differential equations with highly discontinuous and anisotropic coefficients. *SIAM Journal on Scientific Computing* 1998; **20**(1):228–242.
25. Dendy JE, Ida MP, Rutledge JM. A semicoarsening multigrid algorithm for SIMD machines. *SIAM Journal on Scientific Computing* 1992; **13**(6):1460–1469.
26. Vanek P, Mandel J, Brezina M. Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems. *Computing* 1996; **56**(3):179–196.