SIAM J. SCI. COMPUT.
Vol. 21, No. 4, pp. 1632–1649

AN ENERGY-MINIMIZING INTERPOLATION FOR ROBUST MULTIGRID METHODS*

W. L. WAN†, TONY F. CHAN‡, AND BARRY SMITH§

Abstract. We propose a robust interpolation for multigrid based on the concepts of energy minimization and approximation. It can handle PDE coefficients of various types on structured or unstructured grids under one framework. The formulation is general; it can be applied to any dimension. We demonstrate numerically the effectiveness of the multigrid method in two dimensions by applying it to a discontinuous coefficient problem, an oscillatory coefficient problem, and an anisotropic problem. Empirically, the convergence rate is independent of the coefficients of the underlying PDE, in addition to being independent of the mesh size. The proposed method is primarily designed for second-order elliptic PDEs, with possible extensions to other classes of problems such as integral equations.

Key words. multigrid, interpolation, energy minimization, nonsmooth coefficient, elliptic differential equations

AMS subject classifications. 65F10, 65N22, 65N55

PII. S1064827598334277

1. Introduction.

Multigrid methods are widely used as efficient solvers for second-order elliptic partial differential equations (PDEs) because of their often optimal convergence behavior; that is, their convergence rate is independent of the mesh size. Optimal theory can be found, for example, in [3, 4, 5, 23, 32, 36, 43, 44]. However, the convergence rate may depend on the nature of the coefficients in the PDE. Typically, the convergence deteriorates as the coefficients become rougher. Specifically, if the coefficients are anisotropic [23], have large jumps [1, 6, 12, 13], or are highly oscillatory [19, 30, 39], standard multigrid methods will converge very slowly. Special techniques such as line Gauss–Seidel/block smoothing [6], semicoarsening [14, 15, 37], algebraic multigrid [2, 7, 33, 35, 38], frequency decomposition [16, 24, 39], and homogenization [19, 30] are used to handle some of these cases. In this paper, we study the design of multigrid methods from the energy minimization point of view, which gives powerful insight into the design of robust multigrid methods.

The success of multigrid hinges on the choice of the coarse grid points, the smoothing procedure, the interpolation operators, and the coarser grid discretization. In standard multigrid, full coarsening, Jacobi or Gauss–Seidel smoothing, and linear interpolation are usually used. Classical convergence theory shows that these simple

*Received by the editors February 18, 1998; accepted for publication (in revised form) December 2, 1998; published electronically March 21, 2000.
http://www.siam.org/journals/sisc/21-4/33427.html

†SCCM Program, Gates Building 2B, Stanford University, Stanford, CA 94305-9025 (wan@sccm.stanford.edu). Previous address: Department of Mathematics, UCLA, Los Angeles, CA 90095-1555. The author was partially supported by the grants listed under the second author, Lawrence Livermore National Laboratory under contract W07405-Eng-48, the Department of Energy under the Accelerated Strategic Computing Initiative (ASCI), and the Alfred P. Sloan Foundation as a doctoral dissertation fellow.

‡Department of Mathematics, UCLA, Los Angeles, CA 90095-1555 (chan@math.ucla.edu). The work of this author was partially supported by the NSF under contract ASC-9720257, NASA Ames under contract NAG2-1238, and Sandia National Laboratory under contract LG-4440.

§Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL 60439 (bsmith@mcs.anl.gov). The work of this author was supported by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Computational and Technology Research, U.S. Department of Energy, under contract W-31-109-Eng-38.

ingredients are enough to achieve optimal convergence for smooth coefficient problems. In general, however, these choices may lead to slow convergence. In one dimension, to remedy the situation, a more robust interpolation [23, 33, 42] can be used. It is obtained by solving local homogeneous PDEs, which are equivalent to minimizing the energy of the coarse grid basis functions.

The extension of this approach to higher dimensions is not obvious. Nonetheless, many attempts [1, 12, 23, 22, 28, 34, 42] have been made to set up similar local PDEs for defining a robust interpolation. In place of setting up PDEs, we consider an equivalent minimization formulation and derive a so-called energy-minimizing interpolation with special emphasis on its stability and approximation properties, which are essential for optimal convergence. In this paper, we propose an energy minimization with global constraint framework to construct robust interpolation which can handle elliptic PDEs with rough coefficients on structured or unstructured grids with one unifying approach. Moreover, it does not assume any underlying coarsening strategy, and hence can be used in combination with, for instance, semicoarsening or algebraic multigrid coarsening. This will be made more precise in section 3.

This energy-minimizing approach to determining appropriate interpolation operators has also been used for iterative substructuring [17]. We also note that general properties of interpolation operators for multigrid have been discussed by Vaněk, Mandel, and Brezina [40], in which they describe a smoothed aggregation method of achieving these properties. The connections between their and our approaches will be discussed in section 3.2.3. In their recent paper [31], Mandel, Brezina, and Vaněk interpret the smoothed aggregation approach as a projected steepest descent method for computing the interpolation operators formulated in section 3, and they also extend the formulation to systems of PDEs arising from linear elasticity problems.

In section 4, we give numerical examples mainly in two dimensions, including a discontinuous coefficient problem, an oscillatory coefficient problem, and an anisotropic problem. Empirically, the convergence rate is independent of the coefficients of the underlying PDE, in addition to being independent of the mesh size. In the one-dimensional case, this can be rigorously proved (see [41]), which has not been given in the literature before. Finally, we summarize our experience with several remarks in section 5.

We now set up notation to be used throughout the paper. Let $V = V^h$ and let $V_1 \subset V_2 \subset \cdots \subset V_J = V$ denote a sequence of nested subspaces of V defined by the span of nodal basis functions, $\{\phi_i^k\}_{i=1}^{n_k}, k = 1, \ldots, J$, at level k. The operator $A : V \to V$ is self-adjoint and induces the A-inner product, $(\cdot, \cdot)_A \equiv (A\cdot, \cdot)$. Also, we define $A_i : V_i \to V_i$ by $(A_i u_i, v_i) = (Au_i, v_i), u_i, v_i \in V_i$. Correspondingly, we have $R_i : V_i \to V_i$, which is an approximate inverse of A_i. Let $Q_i : V \to V_i$ and $P_i : V \to V_i$ be the projection operators with respect to the L^2 and the A inner product, respectively. In the following analysis, the generic constant C is independent of the mesh size h.

2. Stability and approximation property. Before we explain the formulation of the energy-minimizing interpolation, we first discuss our motivation from the classical results of multigrid and domain decomposition methods. Two key properties, stability and approximation, must be satisfied by the coarse subspaces and the smoothers [4, 23] in order to obtain optimal convergence results. These two terms occur frequently in the literature but often appear in slightly different forms. For example, in the subspace correction framework [44], these two properties are built into the estimate of a constant K_0, which in turn is used to prove optimal convergence,

together with another constant K_1. The definitions of K_0 and K_1 are as follows:

K_0: For any $v \in V$, there exists a decomposition $v = \sum_{i=1}^{J} v_i$ for $v_i \in V_i$ such that

$$\sum_{i=1}^{J} (R_i^{-1} v_i, v_i) \leq K_0 (Av, v), \tag{1}$$

where R_i is usually known as the smoother in the multigrid context.

K_1: For any $S \subset \{1, \ldots, J\} \times \{1, \ldots, J\}$ and $u_i, v_i \in V$ for $i = 1, \ldots, J$,

$$\sum_{(i,j) \in S} (T_i u_i, T_j u_j)_A \leq K_1 \left(\sum_{i=1}^{J} (T_i u_i, u_i)_A \right)^{\frac{1}{2}} \left(\sum_{j=1}^{J} (T_j v_j, v_j)_A \right)^{\frac{1}{2}}, \tag{2}$$

where $T_i = R_i A_i P_i$.

THEOREM 2.1. *Let E_J be the iteration matrix given by the V-cycle multigrid, namely,*

$$u - u^{k+1} = E_J (u - u^k),$$

where u is the exact solution and u^k and u^{k+1} are two consecutive multigrid iterates. Then

$$E_J = (I - T_J)(I - T_{J-1}) \cdots (I - T_1)$$

and

$$\|E_J\|_A^2 \leq 1 - \frac{2 - \omega_1}{K_0 (1 + K_1)^2},$$

where $\omega_1 = \max_{1 \leq i \leq J} \rho(R_i A_i)$.

Proof. See [44]. □

By Theorem 2.1, the convergence rate can be improved by producing a smaller K_0 or K_1. In this paper, we propose an interpolation that will potentially decrease the size of the constant K_0 by reducing its dependence on the coefficients of the underlying elliptic PDE.

As shown in [44], the estimate of K_0 relies on two inequalities:

$$\|\tilde{Q}_1 v\|_A^2 + \sum_{k=2}^{J} \|(\tilde{Q}_k - \tilde{Q}_{k-1}) v\|_A^2 \leq C_0 \|v\|_A^2, \tag{3}$$

$$\|(\tilde{Q}_k - \tilde{Q}_{k-1}) v\| \leq C_1 h_k \|\tilde{Q}_k v\|_A \quad \forall k > 1, \tag{4}$$

where $\tilde{Q}_k : V \to V_k$ is any linear operator onto V_k.

Inequality (3) appears in the partition lemma, which is well known in the domain decomposition literature [18, 36]. In the multigrid context, however, this inequality typically is used only implicitly. Intuitively speaking, (3) says that given any $v \in V$, we must be able to decompose v into the subspaces such that the total energy of all the pieces v_i is bounded by a small constant factor of the original energy of v. Besides (3), we also require that functions on the coarser grids approximate those on the finer grids to at least first-order accuracy in h_k. This requirement is quantified by the inequality (4). If we have both (3) and (4), we can bound K_0 by a constant independent of the mesh size h.

LEMMA 2.2. *Let* $\omega_0 = \min_{2\le i\le J}(\rho(A_i)\lambda_{\min}(R_i))$. *Suppose* (3) *and* (4) *are satisfied. Then*

$$K_0 \le \frac{C}{\omega_0},$$

where C *is a constant independent of the mesh size.*

Proof. See [44]. ∎

To summarize, if the stability and approximation properties (3) and (4) are satisfied, mesh independent convergence follows. Thus, these two properties characterize a good coarse subspace. It is interesting to note that linear finite element subspaces are not compulsory for the V_k, though they are typically used or assumed in the classical analysis of multigrid methods. Moreover, the $\tilde{Q}_k$ in the approximation inequality (4) need not necessarily be the L^2 projections Q_k. Linear finite element and L^2 projections are simply two convenient and powerful tools for showing the stability and the approximation properties, but are not necessarily the only choice.

Mesh independent convergence, however, need not imply rapid convergence. The reason is that, in general, K_0 will depend on the PDE coefficients. The implicit dependence of the coefficients of the underlying PDE in the convergence rate may cause the multigrid method to converge very slowly, for example, when the coefficients are not smooth. In the following section, we construct coarse subspaces whose basis functions are, in general, different from piecewise linear finite elements but possess the stability and the approximation properties. In addition, the resulting multigrid algorithm is less sensitive to the coefficients than is the standard multigrid method. Furthermore, we show that these two concepts lead to an optimal convergence for a one-dimensional multigrid method, and we illustrate how they motivate a two-dimensional multigrid algorithm.

3. Energy-minimizing interpolation. In this section, we introduce the energy minimization approach to constructing the interpolation. The resulting formulation in the one-dimensional case is well known in the literature [23, 33, 42]. We explain the energy-minimizing interpolation in one dimension first and then in two dimensions.

3.1. One dimension. Consider the following model problem:

$$\begin{aligned} -\frac{d}{dx}a(x)\frac{d}{dx}u(x) &= f \qquad \text{in } (0,1), \\ u &= 0 \qquad \text{at } x=0 \text{ and } x=1, \end{aligned} \tag{5}$$

where $a(x)$ and $f(x)$ are integrable and $a(x)$ is uniformly positive.

Let $H^1(0,1)$ be the standard Sobolev space on the interval $[0,1]$ and $H_0^1(0,1)$ its subspace whose functions vanish at $x=0$ and $x=1$. Then the variational formulation of (5) is to find $u \in H_0^1(0,1)$ such that

$$a(u,v) = (f,v) \qquad \forall v \in H_0^1(0,1),$$

where

$$a(u,v) = \int_0^1 a(x)u'(x)v'(x)dx, \quad (f,v) = \int_0^1 f(x)v(x)dx.$$

Given a uniform grid with grid size $h = 1/n$, let $x_j^h = jh, j = 0,\dots,n$. Here, n is assumed to be a power of 2. Define the fine grid linear finite element space to be

$$V^h = \{v^h \in H_0^1(0,1) : v^h \text{ is linear on } [x_j^h, x_{j+1}^h], j = 0,\dots,n-1\}$$

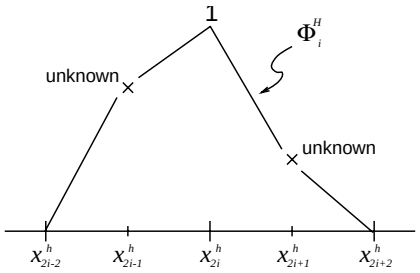

FIG. 1. *One-dimensional coarse grid basis function* ϕ_i^H *on its support* $[x_{2i-2}^h, x_{2i+2}^h]$.

and denote the set of nodal basis by $\{\phi_j^h\}_{j=1}^n$. The finite element approximation to the solution of (5) is the function $u^h \in V^h$, so that

$$a(u^h, v^h) = (f, v^h) \qquad \forall v^h \in V^h. \tag{6}$$

Let $u^h = \sum_{j=1}^n \mu_j \phi_j^h$ and $f = \sum_{j=1}^n \beta_j \phi_j^h$. Then (6) is equivalent to a linear system

$$\mathcal{A}^h \mu = \mathcal{M}^h b,$$

where $\mu = (\mu_1, \ldots, \mu_n)^T$, $b = (\beta_1, \ldots, \beta_n)^T$, $\mathcal{A}^h$ is the stiffness matrix, and $\mathcal{M}^h$ is the mass matrix. Define $\tilde{\mathcal{A}}^h$ to be the augmented stiffness matrix that also includes the equations of the boundary points. Thus, $\tilde{\mathcal{A}}^h$ is singular with the null space consisting of constant functions, and $\mathcal{A}^h$ is a submatrix of it.

Let $x_i^H = x_{2i}^h, i = 0, \ldots, n/2$, be the set of coarse grid points. Now we define a coarse subspace V^H for multigrid by defining the coarse grid nodal basis functions $\{\phi_i^H\}$. That is,

$$V^H = \text{span}\{\phi_i^H : i = 1, \ldots, m\}$$

and $m = n/2 - 1$. Since $\{\phi_i^H\}$ are nodal basis functions on the coarse grid, $\phi_i^H(x_{2i}^h) = 1$ and $\phi_i^H(x_{2i-2}^h) = \phi_i^H(x_{2i+2}^h) = 0$. We need only to define $\phi_i^H(x_{2i-1}^h)$ and $\phi_i^H(x_{2i+1}^h)$ (see Figure 1). For example, if we let them equal $1/2$, the basis functions $\{\phi_i^H\}$ are just linear finite elements, implying that the interpolation from the coarse grid to the fine grid is piecewise linear.

Since $\{\phi_i^H\}$ is a basis of V^H, which is a subspace of V^h, there exists a unique matrix $\mathcal{I}_h^H$ of size $n \times m$ such that

$$[\phi_1^H \cdots \phi_m^H] = [\phi_1^h \cdots \phi_n^h] \mathcal{I}_h^H.$$

The matrix $\mathcal{I}_h^H$ is usually known as the prolongation (or interpolation) matrix and its transpose $(\mathcal{I}_h^H)^T = \mathcal{I}_H^h$ as the restriction matrix in the multigrid context. Hence, the set of coarse grid basis functions defines an interpolation and vice versa. In the following, instead of deriving an interpolation method directly, we construct an energy-minimizing basis.

Formulation. As noted above, the interpolation is uniquely defined if the coarse grid basis functions $\{\phi_i^H\}$ are known. We can define $\phi_i^H(x)$ by solving the following local PDE problem in $[x_{i-1}^H, x_i^H] = [x_{2i-2}^h, x_{2i}^h]$:

$$\begin{aligned} &-\frac{d}{dx} a(x) \frac{d}{dx} \phi_i^H = 0 \qquad &&\text{in } [x_{2i-2}^h, x_{2i}^h], \\ &\phi_i^H(x_{2i-2}^h) = 0, \quad \phi_i^H(x_{2i}^h) = 1. \end{aligned} \tag{7}$$

We observe that the PDE formulation of the basis functions has a "physical" meaning attached to it. Specifically, it looks for basis functions that have small energy. It is best illustrated by the following result.

LEMMA 3.1. *An equivalent formulation of* (7) *is*

$$\begin{aligned} &\min \quad a(\phi_i^H, \phi_i^H) && \text{in } [x_{2i-2}^h, x_{2i}^h], \\ &\text{subject to } \phi_i^H(x_{2i-2}^h) = 0, \quad \phi_i^H(x_{2i}^h) = 1. \end{aligned} \tag{8}$$

Thus, the solution of the local PDE minimizes the energy of the coarse grid basis functions. This observation turns out to be very convenient for extending the idea to higher dimensions.

The solution of $\phi_i^H(x)$ on $[x_{2i-2}^h, x_{2i}^h]$ gives $\phi_i^H(x_{2i-1}^h)$. We can do the same for $\phi_i^H(x_{2i+1}^h)$ in $[x_{2i}^h, x_{2i+2}^h]$. The local PDE formulation calculates the "harmonic" function ϕ_i^H which minimizes the energy on its support. If $a(x) \equiv 1$, ϕ_i^H is a linear function and we recover the linear interpolation, that is, $\phi_i^H(x_{2i-1}^h) = \phi_i^H(x_{2i+1}^h) = 1/2$. In fact, in this case, ϕ_i^H is harmonic in the usual sense, and it has minimum energy. In general, instead of $1/2$, we have

$$\phi_i^H(x_{2i-1}^h) = -\frac{a(\phi_{2i-1}^h, \phi_{2i}^h)}{a(\phi_{2i-1}^h, \phi_{2i-1}^h)} = -\frac{\mathcal{A}_{2i-1,2i}^h}{\mathcal{A}_{2i-1,2i-1}^h}, \tag{9}$$

where $(\mathcal{A}_{ij}^h)$ is the stiffness matrix. Since our interpolation depends on the matrix $\mathcal{A}^h$, sometimes it is called a matrix-dependent interpolation in the algebraic multigrid context. The resulting interpolation was also described in [23, 33, 42] but from a different point of view. Our interpretation based on the energy-minimization principle provides a clue to developing similar interpolation operators in higher dimensions.

The approximation property (4) is closely related to preserving constant functions. In fact, the coarse space V^H constructed in this way automatically contains constant functions on the fine grid.

LEMMA 3.2.

$$\sum_{i=1}^{m} \phi_i^H(x) = 1.$$

Proof. Let $\psi^H(x) = \sum_{i=1}^m \phi_i^H(x)$. By (7), for $i = 1, \ldots, m$, ψ^H satisfies the following:

$$\begin{aligned} &-\frac{d}{dx} a(x) \frac{d}{dx} \psi^H = 0 && \text{in } [x_{2i-2}^h, x_{2i}^h], \\ &\psi^H(x_{2i-2}^h) = 1, \quad \psi^H(x_{2i}^h) = 1. \end{aligned}$$

By uniqueness, $\psi^H \equiv 1$ on $[x_{2i-1}^h, x_{2i}^h]$, and hence the result follows. ∎

Thus, the interpolation derived from the energy-minimizing coarse grid basis functions preserves constants.

THEOREM 3.3. *If the energy-minimizing interpolation operator derived from* (8) *is used together with damped Jacobi or Gauss–Seidel smoothings, the resulting multigrid convergence is independent of the mesh size and the PDE coefficient* $a(x)$.

The proof, which is given in detail in [41], uses the fact that $\{\phi_i^k\}_{k=1}^J$ obtained from recursively applying (8) to coarser levels contains an A-orthogonal hierarchical basis, and hence the damped Jacobi and Gauss–Seidel smoothings give an optimal

constant bound for K_0 and K_1 which are the essential elements for estimating multigrid convergence by Theorem 2.1.

Remarks. (1) If $a(x)$ is piecewise constant, this interpolation preserves the continuity of the flux, $a(x)\nabla u$, at the discontinuities [23]. (2) If red-black Gauss–Seidel is used as a smoother, the resulting multigrid method coincides with the cyclic reduction method in the numerical linear algebra context.

3.2. Higher dimensions. The construction of the energy-minimizing interpolation described in this section is valid for two and three dimensions. However, to facilitate understanding, we focus on the standard structured grid on the square domain Ω: $[0,1] \times [0,1]$ in two dimensions. The model problem is

$$\begin{aligned} -\nabla \cdot a(x,y)\nabla u(x,y) &= f(x,y) \quad && \text{in } \Omega, \\ u &= 0 && \text{on } \partial\Omega, \end{aligned} \tag{10}$$

with the same assumptions on $a(x,y)$ and $f(x,y)$ as before. Again, we use a finite element method to discretize (10).

3.2.1. Formulation. The extension to higher dimensions of the local PDE approach is difficult because there is no natural analog between one dimension and higher dimensions. For instance, in one dimension, the coarse grid points form the boundaries of the local subdomains so that well-posed PDEs can be easily defined. In higher dimensions, however, the boundaries consist of both coarse grid and noncoarse grid points, and hence local boundary value problems apparently do not exist. Nevertheless, several possibilities for setting up local PDEs are discussed in the literature, for instance, the stencil or the so-called black-box multigrid approach [1, 12, 13, 23, 22, 27, 28, 42, 45], the Schur complement approach [20, 29, 34], and the algebraic multigrid approach [2, 7, 33, 35, 40], each of which mimics the one-dimensional case in some way.

Our approach is based on the observation (8). The coarse grid basis functions $\{\phi_i^H\}$ should possess the least amount of energy while preserving constant functions. However, we can not use (8) directly in higher dimensions. First, the $\{\phi_i^H\}$, each of which has minimum energy, do not preserve constant functions. Second, the boundary of the support of each ϕ_i^H, in general, consists of both coarse and noncoarse grid points, and hence the boundary conditions of (8) must be modified. The stability inequality (3) in the two-level setting ($J = 2$) provides a hint on the quantity to be minimized. Let I^H be the usual nodal value interpolant. By the Cauchy–Schwarz and Poincaré inequalities, we have

$$\begin{aligned} \|I^H v\|_A &= \left\| \sum_i v(x_i^H)\phi_i^H \right\|_A \\ &\le \left(\sum_i \|\phi_i^H\|_A^2 \right)^{1/2} \left(\sum_i |v(x_i^H)|^2 \right)^{1/2} \\ &\le \left(\sum_i \|\phi_i^H\|_A^2 \right)^{1/2} \frac{C}{H} \|v\|_2 \\ &\le \left[\frac{C}{H} \left(\sum_i \|\phi_i^H\|_A^2 \right)^{1/2} \right] \|v\|_A, \end{aligned} \tag{11}$$

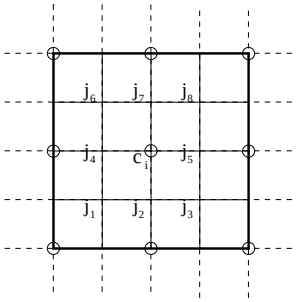

FIG. 2. *Two-dimensional coarse grid basis function ϕ_i^H on its support. ϕ_i^H is a linear combination of fine grid basis functions ϕ_j^h, $j = j_1, \ldots, j_8$, and c_i.*

where C is a constant independent of h. Comparing (11) with the stability inequality (3), we see that the constant C_0 in (3) depends on the total energy of $\{\phi_i^H\}$. Thus, our formulation is to minimize the sum of energies of $\{\phi_i^H\}$ so that the constant C_0 and hence the multigrid convergence will be improved.

The precise mathematical formulation is explained in the following. Given a set of m coarse grid points obtained by, for example, a maximal independent set or algebraic multigrid coarsening, denote their index set by $M = \{c_1, \ldots, c_m\}$. Write the coarse grid nodal basis function ϕ_i^H at node x_{c_i} as a linear combination of the fine grid ones,

$$\phi_i^H = \sum_{j \in N_{c_i} \setminus M} \varphi_j^i \phi_j^h + \phi_{c_i}^h, \tag{12}$$

where N_{c_i} is the set of neighboring nodes of c_i with respect to the mesh or matrix adjacency graph. Thus, ϕ_i^H is a local combination of $\phi_{c_i}^h$ and the fine grid basis functions whose corresponding node is adjacent to node x_{c_i} but not itself a coarse grid point. Figure 2 shows the support of ϕ_i^H in a two-dimensional regular grid. The indices j in the sum on the right-hand side of (12) correspond to $j_1, \ldots, j_8$. Since ϕ_i^H is a nodal basis function, the coefficient of $\phi_{c_i}^h$ is equal to 1. In view of the above calculation for estimating C_0, our formulation is to define the interpolation by solving a constrained minimization problem for $\{\varphi_j^i\}$,

$$\min \ \frac{1}{2} \sum_{i=1}^{m} \|\phi_i^H\|_A^2 \qquad \text{subject to} \qquad \sum_{i=1}^{m} \phi_i^H(x) = 1 \text{ in } \bar{\Omega}. \tag{13}$$

Notice that the minimization problem is solved up to and including the boundary of Ω. Usually, the grid points on the boundary with Dirichlet boundary condition are treated separately, and no coarse grid point is placed there. However, in our formulation, we compute all ϕ_i^H including the ones at the boundary, but only those not on the boundary with Dirichlet condition are used in the interpolation.

LEMMA 3.4. *An equivalent formulation of* (7) *and* (8) *is the global minimization*

$$\min \ \frac{1}{2} \sum_{i=1}^{m} \|\phi_i^H\|_A^2 \qquad \textit{subject to} \qquad \sum_{i=1}^{m} \phi_i^H(x) = 1 \ \textit{on} \ [0,1].$$

Thus, we see a way to naturally generalize the approach for generating a robust interpolation from one dimension to multiple dimensions.

Remarks. (1) The values of the basis functions are defined implicitly by the solution of (13) and are not known explicitly in general. However, for the Laplacian, we recover exactly the bilinear interpolation on tensor-product grids, which is known to lead to optimal multigrid convergence for Poisson equations.

LEMMA 3.5. *The solution of* (13) *gives the bilinear interpolation if* $a(x) \equiv 1$.

We also remark that if triangular grids are used, the linear interpolation is almost recovered; numerical experiments show that the interpolation values are close to 1/2.

(2) Like algebraic multigrid, the construction of the interpolation operator is purely algebraic. In other words, geometry and, in particular, the grid information are not needed. The formulation of the interpolation remains valid if the coarse grid points do not form an independent set. Independent sets are certainly beneficial to efficiency but are not necessary. In some situations, we may want to remove this requirement, for example, when semicoarsening is used.

(3) Since the construction is algebraic, we can recursively define a multilevel coarse grid basis functions $\{\phi_i^k\}$, $k = 1, \ldots, J$, satisfying (13).

(4) Finally, we remark that we may generalize the formulation further by putting in positive weights θ_i in front of $\|\phi_i^H\|_A^2$. Similarly, we have the following equivalence.

LEMMA 3.6. *Recall that* (7) *and* (8) *apply only for problems in one dimension. An equivalent formulation of* (7) *and* (8) *is the global weighted minimization*

$$\min \ \frac{1}{2}\sum_{i=1}^{m} \theta_i \|\phi_i^H\|_A^2 \qquad \textit{subject to} \qquad \sum_{i=1}^{m} \phi_i^H(x) = 1 \ \textit{on} \ [0,1]$$

for any sets of positive θ_i.

In our experience, special scalings in two or more dimensions, for instance, $\theta_i = 1/\tilde{\mathcal{A}}^h_{c_i,c_i}$, may improve the performance for problems such as discontinuous coefficient PDEs where the discontinuities do not align with any coarser grids. However, an optimal choice of θ_i has not yet been fully analyzed, and hence we shall not discuss this generalization further in the present paper.

3.2.2. Solution of the minimization problem. We describe a solution procedure for the minimization problem (13) below. For each i, write $\phi_i^H = \sum_{j=1}^n \varphi_j^i \phi_j^h$ and $\varphi^i = (\varphi_1^i, \ldots, \varphi_n^i)^T$. By (12), φ^i is a sparse vector. For example, in two dimensions, φ^i has at most nine nonzeros. For structured triangular grids, φ^i has at most seven nonzeros. Let $\Phi = [\varphi^1; \cdots; \varphi^m]$ be an $mn \times 1$ vector obtained by concatenating all the φ's. Note that $\|\phi_i^H\|_A^2 = \|\sum_{j=1}^n \varphi_j^i \phi_j^h\|_A^2 = (\varphi^i)^T \tilde{\mathcal{A}}^h \varphi^i$. (Recall that $\tilde{A}^h$ is the augmented stiffness matrix on the fine grid without incorporating any Dirichlet boundary condition.) Thus, (13) can be written as the following equivalent discrete linear constrained quadratic minimization problem:

$$\min \ \frac{1}{2}\Phi^T \mathcal{Q} \Phi \quad \text{subject to} \quad \mathcal{B}^T \Phi = \mathbf{1}. \tag{14}$$

The symbol $\mathbf{1}$ denotes a vector of all 1's. The $mn \times mn$ SPD matrix $\mathcal{Q}$ is block diagonal with each block equal to $\tilde{\mathcal{A}}_i^h$ which is defined as

$$(\tilde{\mathcal{A}}_i^h)_{kl} = \begin{cases} \tilde{\mathcal{A}}_{kl}^h & \text{if } \varphi_k^i \neq 0 \text{ and } \varphi_l^i \neq 0, \\ \delta_{kl} & \text{otherwise.} \end{cases}$$

The $n \times mn$ rectangular matrix $\mathcal{B}^T = [\mathcal{J}_1^T \cdots \mathcal{J}_m^T]$, where $\mathcal{J}_i = \mathcal{J}_i^T$ is a matrix corresponding to the restriction operator that maps v to v_i such that $(v)_k = (v_i)_k$ on

$\mathrm{supp}(\phi_i^H)$ and $(v_i)_k = 0$ otherwise. More precisely,

$$(\mathcal{J}_i)_{kl} = \begin{cases} 1 & \text{if } k = l \text{ and } \varphi_k^i \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

It is clear that $\mathcal{J}_i^T \varphi^i = \varphi^i$, and hence $\mathcal{B}^T \Phi = \sum_{i=1}^m \mathcal{J}_i^T \varphi^i = \sum_{i=1}^m \varphi^i = \mathbf{1}$. We solve the discrete linearly constrained minimization problem (14) by the Lagrange multiplier formulation, which is equivalent to

$$\begin{bmatrix} \mathcal{Q} & \mathcal{B} \\ \mathcal{B}^T & 0 \end{bmatrix} \begin{bmatrix} \Phi \\ \Lambda \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{1} \end{bmatrix}, \tag{15}$$

where Λ is an $n \times 1$ vector of Lagrange multipliers. If Λ is known, Φ can be computed by solving

$$\mathcal{Q}\Phi = -\mathcal{B}\Lambda. \tag{16}$$

Since $\mathcal{Q}$ is block diagonal and inverting each block corresponds to solving a matrix of at most 9×9 in size, it is trivial to compute Φ once Λ is known. Thus, the entire minimization procedure is reduced to solving for the Lagrange multipliers Λ via

$$(\mathcal{B}^T \mathcal{Q}^{-1} \mathcal{B})\Lambda = -\mathbf{1}. \tag{17}$$

Note that $\mathcal{B}$ and $\mathcal{Q}^{-1}$ are sparse matrices. We can solve the linear system by the conjugate gradient (CG) method.

The solution of (17) could be costly. Depending on the conditional number of $\mathcal{Q}^{-1}$, the CG iteration may converge slowly. We shall discuss how to speed up the process. First, we need not compute $(\mathcal{B}^T \mathcal{Q}^{-1} \mathcal{B})^{-1}\mathbf{1}$ exactly because we are merely computing the interpolation to be used in the multigrid method. In fact, the numerical results in section 4 indicate that Λ is usually accurate enough when the relative residual of (17) is less than 10^{-2}.

Besides, we have a readily obtainable initial guess for Λ. Consider (16). Multiplying both sides by $\mathcal{B}^T$, we have

$$\Lambda = -(\mathcal{B}^T \mathcal{B})^{-1} \mathcal{B}^T \mathcal{Q} \Phi.$$

By a direct computation, one verifies that $\mathcal{B}^T \mathcal{B}$ is a diagonal matrix. Hence, this gives an easy way to compute an initial guess for Λ from Φ. Since the interpolation weights are between 0 and 1, the solution Φ usually is not very far from the linear interpolation. It may be advantageous to use the linear interpolation as an initial guess for Φ, which in turn provides an initial guess for Λ.

It is interesting to note that $\tilde{\mathcal{A}}^h$ is a free and natural preconditioner for $\mathcal{B}\mathcal{Q}^{-1}\mathcal{B}$. By the definition of $\mathcal{B}$ and $\mathcal{Q}$, rewrite the product $\mathcal{B}^T \mathcal{Q}^{-1} \mathcal{B}$ as a sum of matrices:

$$\mathcal{B}^T \mathcal{Q}^{-1} \mathcal{B} = \sum_{i=1}^m \mathcal{J}_i^T (\tilde{\mathcal{A}}_i^h)^{-1} \mathcal{J}_i = \sum_{i=1}^m \mathcal{R}_i^T (\mathcal{R}_i \tilde{\mathcal{A}}^h \mathcal{R}_i^T)^{-1} \mathcal{R}_i,$$

where $\mathcal{R}_i$ is the submatrix of the nonzero rows of $\mathcal{J}_i$ and it is sometimes known as the restriction matrix in the domain decomposition context. Clearly, $\mathcal{B}^T \mathcal{Q}^{-1} \mathcal{B}$ is an overlapping additive Schwarz preconditioner of $\tilde{\mathcal{A}}^h$. Unfortunately, $\tilde{\mathcal{A}}^h$ is singular in our case. A simple remedy is to use $\tilde{\mathcal{A}}^h + \eta \mathcal{I}$ instead as the preconditioner.

Because of the potentially high cost of computing Φ, the energy-minimizing interpolation is aimed at problems for which linear interpolation does not work well. Quite often, we may need to solve the same system many times, for instance, in time-dependent problems. The rapid convergence of each multigrid solve compensates for the expensive setup cost.

Remark. In [31], Mandel, Brezina, and Vaněk propose the use of steepest descent for solving (14) in place of the Newton's method on the Lagrangian described above.

3.2.3. Connections to other approaches. As noted above, the entire procedure of constructing the interpolation is algebraic; therefore it can be considered as a type of algebraic multigrid. In fact, it is related to the one derived by Vaněk, Mandel, and Brezina [40]. In their approach, groups of fine grid nodes are agglomerated to form larger elements, or macroelements. In each agglomerated region (which can be thought of a subdomain in the domain decomposition context), a value of 1 is assigned to each node as an initial guess of the coarse grid basis. Because of the high energy of the piecewise constant basis functions, they are *smoothed* by a few steps of Jacobi iteration. Our energy-minimizing coarse grid basis can also be thought of as being formed by agglomerating nearby fine grid nodes, but the agglomeration only occurs at the neighboring nodes of the coarse grid nodes. Also there are overlaps among agglomerated regions, while there are none in the approach of Vaněk, Mandel, and Brezina. Moreover, the support of their basis functions will increase when the Jacobi "smoothing" steps are applied to the basis functions. In our approach, the supports are fixed and the energy is minimized by solving the minimization problem (13). In the recent paper of Mandel, Brezina, and Vaněk [31], they show that the first step of the steepest descent procedure in [31] yields the same result as the smoothed aggregation with a single smoothing step.

Because of the agglomeration view of the construction, our approach is also related to the one derived by Chan et al. [10, 11]. They explicitly form the macroelements by agglomeration using standard graph theoretical techniques. Then they describe ways of defining the coarse grid basis functions. One way is the following: The noncoarse grid points on the edge of a macroelement are assigned a value using the graph distance, and those noncoarse grid points in the interior are obtained by solving a local homogeneous PDE. Our approach does not prescribe a value on the edges of the macroelements first and then solve for the interior points. Rather, we take all the unknowns together and solve for all the values simultaneously by solving the minimization problem.

4. Numerical results. In this section, we present results of numerical experiments mainly in two dimensions to verify that the multigrid algorithm resulting from the energy-minimizing interpolation has optimal convergence behavior and is robust with respect to the coefficients of the PDEs. Meanwhile, we compare the results with algebraic multigrid (AMG) [21], whose implementation is based on [35], to show that the energy-minimizing interpolation obtained from an abstract formulation leads to a multigrid convergence rate comparable to AMG. The problems are chosen to illustrate that the energy-minimizing multigrid is applicable to various kinds of coefficients such as discontinuous, oscillatory, and anisotropic.

In all the numerical examples, the computational domain is $\Omega = [0,1] \times [0,1]$ with homogeneous Dirichlet boundary condition. In the multigrid procedure a V-cycle is used with two pre- and two postpointwise Gauss–Seidel smoothings. The iteration was terminated when the relative residual ℓ_2-norm was less than 10^{-6}. The number of

TABLE 1
Number of V-cycles using bilinear and energy-minimizing interpolations when $a(x) = 1 + xe^y$.

	MGBL				EMMG(10^{-1})				EMMG(10^{-12})			
h	4	5	6	7	4	5	6	7	4	5	6	7
1/16	5	-	-	-	5	-	-	-	5	-	-	-
1/32	5	5	-	-	5	5	-	-	5	5	-	-
1/64	5	5	5	-	5	5	5	-	5	5	5	-
1/128	5	5	5	5	5	5	5	5	5	5	5	5

multigrid levels is such that the coarsest grid is a single point, or as otherwise stated. All the coarse grid operators are obtained by the standard Galerkin process.

In section 3.2.1, we mentioned that it is not necessary to compute the Lagrange multipliers to machine precision; they are solved only approximately in all the examples below. Moreover, we used piecewise linear or bilinear interpolation as our initial guess for the minimization problem. Besides, as discussed in section 3.2.2, the augmented stiffness matrix $\tilde{\mathcal{A}}^h$, or more precisely, $\tilde{\mathcal{A}}^h + \eta\mathcal{I}$, is a free preconditioner for solving the Lagrange multiplier equation (17). In all the numerical examples, this preconditioner is used with η chosen as 10^{-3}.

Example 1. In this example we verify numerically that the convergence rate does not depend on the mesh size and the number of levels. Here we consider the following PDE with a smooth coefficient:

$$-\nabla \cdot (1 + xe^y)\nabla u = 1.$$

Table 1 shows the number of multigrid iterations to converge. We denote the multigrid method with bilinear interpolation by MGBL and our energy-minimizing multigrid method by EMMG(ϵ), where ϵ specifies the stopping criterion for CG applied to the Lagrange multiplier equation (17). More precisely, the CG iteration is stopped when the relative residual norm is less than ϵ. We see that when the optimization problem is effectively solved ($\epsilon = 10^{-12}$), the convergence rate is independent of the mesh size h and the number of levels. In fact, we observe that same convergence rate can be achieved even if the optimization problem is solved approximately ($\epsilon = 10^{-1}$). Thus, we may reduce the cost by applying significantly fewer number of CG iterations as shown in Table 2, which gives the number of conjugate gradient iterations at each multigrid level to solve (17).

We remark that this example is used to illustrate the optimal convergence of EMMG(ϵ) and the effect of varying ϵ only. It is not cost-effective to use energy-minimizing interpolation when bilinear interpolation works well.

Example 2. We compare the multigrid method using bilinear interpolation with that using energy-minimizing interpolation by solving the following discontinuous coefficient problem [1, modified Example I]:

$$-\nabla \cdot a(x, y)\nabla u = 1,$$

where

$$a(x, y) = \begin{cases} a^+ & 0.5 - h \le x \le 0.5 + h \quad \text{and} \quad 0.5 - h \le y \le 0.5 + h, \\ a^- & \text{otherwise.} \end{cases}$$

TABLE 2
Number of CG iterations at each multigrid level with varying ϵ when $a(x) = 1 + xe^y$.

h	level	EMMG(10^{-1})	EMMG(10^{-12})
1/16	4	1	53
	2	1	22
1/32	5	1	98
	3	1	33
	2	1	22
1/64	6	1	180
	4	1	53
	2	1	22
1/128	7	1	309
	5	1	98
	3	1	33
	2	1	22

TABLE 3
Number of V-cycles using bilinear, energy-minimizing, and AMG interpolations for the discontinuous coefficient problem. The jump $a^+ = 10, 10^2, 10^4$. $*$ More than 100 V-cycles required for convergence.

	MGBL			EMMG(10^{-3})			AMG		
h	10	10^2	10^4	10	10^2	10^4	10	10^2	10^4
1/16	12	41	$*$	5	5	5	5	6	7
1/32	11	33	$*$	5	5	5	6	7	7
1/64	11	31	$*$	5	5	5	6	9	10
1/128	10	30	$*$	6	6	6	7	9	10

Notice that the discontinuities do not align with any coarser grid. We fix $a^- = 1$ and vary a^+ from 10 to 10^4. The convergence results are given in Table 3. Here $*$ denotes convergence beyond 100 multigrid iterations. Consistent with the classical theory, the convergence rate of the standard multigrid does not depend on the mesh size h. However, the convergence rate deteriorates substantially as the jump of the discontinuity increases. On the other hand, the convergence of the energy-minimizing multigrid and AMG do not depend either on the mesh size or the size of the jump.

Table 4 shows the average number of CG iterations on the fine grid. It is computed as follows. One CG iteration on the first coarse grid is counted as 1/4 of a CG iteration on the fine grid and so on. From the column under "Ex2" of Table 4, we see that by applying only a few CG iterations to construct the energy-minimizing interpolation, the convergence of multigrid is improved significantly over that using linear interpolation. This result demonstrates that the extra cost of solving the minimization problem is justified by the much faster convergence of the multigrid method.

Example 3. We solve another PDE to demonstrate the robustness of the energy-minimizing multigrid method. The coefficient is oscillatory, and the equation is [25],

TABLE 4

Average number of CG iterations on the fine grid for the discontinuous coefficient problem (Example 2), oscillatory coefficient problem (Example 3), and anisotropic problem (Example 4).

	Ex2: EMMG(10^{-3})		Ex3: EMMG(10^{-2})		Ex4: EMMG(10^{-2})
h	$a^+ = 10$	$a^+ = 10^4$	$\eta = 0.1$	$\eta = 0.01$	
1/16	8.69	10.69	36.56	1.31	8.50
1/32	7.63	10.88	51.83	70.34	9.25
1/64	6.66	8.66	96.64	117.83	8.69
1/128	6.66	8.66	144.66	280.15	8.81

TABLE 5

Number of V-cycles using bilinear, energy-minimizing, and AMG interpolations for the oscillatory coefficient problem. $\eta = 0.1, 0.01$. $*$ More than 100 V-cycles required for convergence.

	MGBL		EMMG(10^{-2})		AMG	
h	0.1	0.01	0.1	0.01	0.1	0.01
1/16	$*$	4	7	5	6	4
1/32	51	$*$	7	14	8	6
1/64	65	58	7	7	10	9
1/128	66	$*$	7	10	10	16

[26, Example 7.4]:

$$-\nabla \cdot \frac{1}{(2 + P\sin(x/\eta))(2 + P\sin(y/\eta))}\nabla u = 1.$$

We chose $P = 1.99$ and η=0.1 and 0.01. The difficulty of this problem is the small scale generated by small η. As the mesh size increases with coarser grids, eventually the coarser grids do not resolve the oscillation of the coefficient, thus slowing down the standard multigrid convergence. The results are shown in Tables 5 and 4 (column under "Ex3"). Both EMMG and AMG converge rapidly. However, the minimization problem is more difficult to solve this time since the coefficient is very rough.

We remark that the nonuniform number of V-cycles to convergence for the case $\eta = 0.01$ may be because the mesh size h is not small enough to resolve the coefficient $a(x, y)$ for the first couple of values of h.

Example 4. We show by an anisotropic problem that the constrained energy minimization formulation will generate the correct coarse grid basis functions, provided that the set of coarse grid points is selected appropriately. The model equation is

$$10^{-4}u_{xx} + u_{yy} = 1.$$

For this kind of problem, semicoarsening [14, 15, 37] is often used, and the interpolation is only done in the y direction but not in the x direction. Suppose the semicoarsening is used to select the coarse grid points and the constrained energy minimization is applied to construct the interpolation operator. Since the construction of the energy-minimizing interpolation is purely algebraic, it may interpolate in both directions depending on the adjacency graph and the mesh topology. Numerical results show that the constrained energy minimization formulation will automatically determine

TABLE 6
Number of V-cycles using bilinear, energy-minimizing, and AMG interpolations for the anisotropic coefficient problem. $*$ More than 100 V-cycles required for convergence.

h	MGBL	EMMG(10^{-2})	AMG
1/16	$*$	5	15
1/32	$*$	6	17
1/64	$*$	6	20
1/128	$*$	6	22

TABLE 7
Number of V-cycles using linear and energy-minimizing interpolations for the matrix defined in Example 5.

h	MGBL	EMMG
1/32	$*$	5
1/64	$*$	5
1/128	$*$	5

the interpolation weighting to be about 1/2 in the y direction and negligibly small in the other directions. If the small elements in the interpolation operator are truncated, it will recover precisely the standard multigrid with semicoarsening. Table 6 shows the fast convergence of EMMG for the anisotropic problem when semicoarsening is used.

The average number of CG iterations for solving the minimization problem as shown in Table 4 (column under "Ex4") is calculated differently from the others. Since semicoarsening is used, one CG iteration on the coarse grid is counted as 1/2 of a CG iteration on the fine grid.

Example 5. We show that the energy minimization principle is not restricted to PDEs by considering a tridiagonal matrix $\mathcal{A}^h$ whose diagonal entries are 2 and off-diagonal entries are 1. This matrix has the same spectrum as the one-dimensional Laplacian operator and hence it is very ill-conditioned. It may be considered as the discretization matrix of a one-dimensional *banded convolution operator* arising from image processing [8]. It is known that standard multigrid converges slowly for this kind of operator.

For this problem, we obtained ϕ_i^H from solving the local PDEs (7), not from the minimization problem (13), since constant functions are not in the kernel of $\mathcal{A}^h$. The convergence results of the multigrid methods using linear and energy-minimizing interpolations are shown in Table 7. The $*$ in the first column indicates that the standard multigrid takes more than 100 V-cycles to converge. The poor convergence comes from the effect of smoothing and the way the interpolation is done. The eigenfunctions of the operator $\mathcal{A}^h$ corresponding to small energy are oscillatory, whereas those corresponding to large energy are relatively smooth. As a result of standard relaxation smoothings, the errors become more oscillatory. Figure 3 shows the effect of 4 and 8 iterations of Gauss–Seidel smoothing applied to a smooth initial error. Such a phenomenon was also discussed in [9]. Hence, if we use linear interpolation, it will not be able to approximate the oscillatory error on the coarser subspaces. This causes the failure of the standard multigrid method.

On the other hand, the multigrid method using energy-minimizing interpolation works fine and shows no deterioration, because the energy minimization captures the

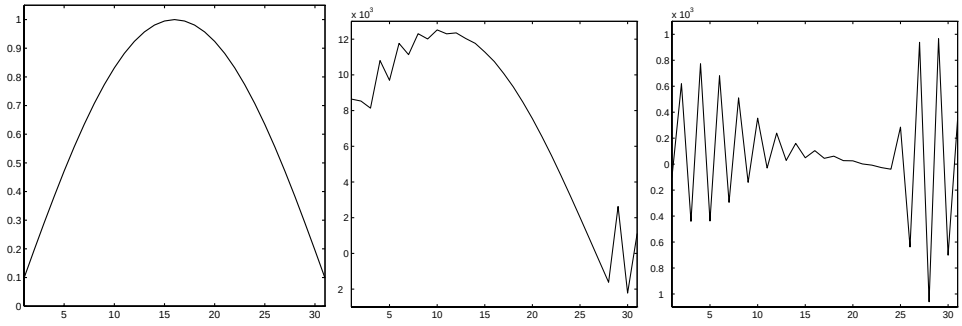

FIG. 3. *Left to right: errors after* 0, 4, *and* 8 *Gauss–Seidel iterations when* $\mathcal{A}^h$ *is as defined in Example* 5.

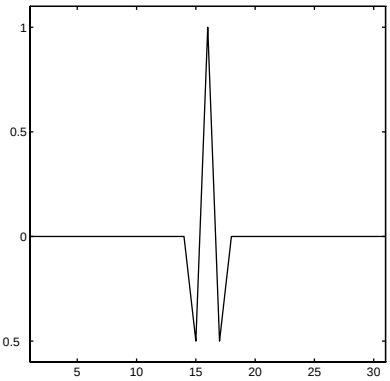

FIG. 4. *A coarse grid basis function obtained by the energy minimization when* $\mathcal{A}^h$ *is as defined in Example* 5.

property of this type of operators and produces oscillatory coarse grid basis functions (see Figure 4). This consistency enables a good approximation on the coarser subspaces, and hence the multigrid convergence is much better. The AMG algorithm, however, breaks down for this problem since the matrix $\mathcal{A}^h$ is not an M-matrix which is essential for the definition of strong coupling for this implementation of AMG.

Remark. The coarse grid basis functions obtained by solving the local PDEs do not preserve constants, an approach that is natural because the operator $\mathcal{A}$ does not annihilate constant functions. If we were to extend our minimization formulation to this case in higher dimensions, we would have to modify the constraint in (13).

5. Concluding remarks. Through the analytical and numerical results we have demonstrated that energy-minimizing and constant preserving are two key properties of the coarse grid interpolation required to obtain a robust multigrid method. An obvious drawback to the construction of the robust interpolation is the expensive solution of the minimization problem. An inexact preconditioned conjugate gradient method with the linear interpolation as initial guess is proposed to overcome this problem. More efficient methods for solving the minimization problem need to be derived and studied.

Finally, because of the algebraic nature of the construction of the interpolation, our method is also applicable to complicated geometries, for instance, unstructured

grids, but these cases are not discussed in the present paper.

Acknowledgments. We would like to acknowledge the permission of German National Research Center for the use of their AMG codes to generate the numerical results in section 4.

REFERENCES

[1] R. E. Alcouffe, A. Brandt, J. E. Dendy, Jr., and J. W. Painter, *The multi-grid method for the diffusion equation with strongly discontinuous coefficients*, SIAM J. Sci. Statist. Comput., 2 (1981), pp. 430–454.

[2] D. Braess, *Towards algebraic multigrid for elliptic problems of second order*, Computing, 55 (1995), pp. 379–393.

[3] J. Bramble, *Multigrid Methods*, Longman Scientific and Technical, Harlow, UK, 1993.

[4] J. Bramble, J. Pasciak, J. Wang, and J. Xu, *Convergence estimates for multigrid algorithms without regularity assumptions*, Math. Comp., 57 (1991), pp. 23–45.

[5] J. Bramble, J. Pasciak, J. Wang, and J. Xu, *Convergence estimates for product iterative methods with applications to domain decomposition and multigrid*, Math. Comp., 57 (1991), pp. 1–21.

[6] A. Brandt, *Multi-level adaptive solutions to boundary-value problems*, Math. Comp., 31 (1977), pp. 333–390.

[7] A. Brandt, S. McCormick, and J. W. Ruge, *Algebraic multigrid (AMG) for automatic multigrid solution with application to geodetic computations*, in Sparsity and Its Applications, D. J. Evans, ed., Cambridge University Press, Cambridge, UK, 1984.

[8] R. Chan, T. Chan, M. Ng, W. Tang, and C. Wong, *Preconditioned iterative methods for high-resolution image reconstruction with multisensors*, in Proceedings SPIE 3461, Advanced Signal Processing Algorithms, Architectures and Implementation VI, 1998.

[9] R. H. Chan, T. F. Chan, and W. L. Wan, *Multigrid for Differential-Convolution Problems Arising from Image Processing*, Technical Report CAM 97-20, Department of Mathematics, UCLA, Los Angeles, CA, 1997.

[10] T. F. Chan, S. Go, and L. Zikatanov, *Lecture notes on multilevel methods for elliptic problems on unstructured grids*, in Lecture Course 28th Computational Fluid Dynamics, von Karman Institute for Fluid Dynamics, Belgium, 1997.

[11] T. F. Chan, J. Xu, and L. Zikatanov, *An Agglomeration Multigrid Method for Unstructured Grids*, Technical Report CAM 98-8, Department of Mathematics, UCLA, Los Angeles, CA, 1998.

[12] J. E. Dendy, Jr., *Black box multigrid*, J. Comput. Phys., 48 (1982), pp. 366–386.

[13] J. E. Dendy, Jr., *Black box multigrid for nonsymmetric problems*, Appl. Math. Comput., 13 (1983), pp. 261–283.

[14] J. E. Dendy, Jr., M. P. Ida, and J. M. Rutledge, *A semi-coarsening multigrid algorithm for SIMD machines*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 1460–1469.

[15] J. E. Dendy, Jr., S. F. McCormick, J. W. Ruge, T. F. Russell, and S. Schaffer, *Multigrid methods for three-dimensional petroleum reservoir simulation*, in 10th Symposium on Reservoir Simulation, Houston, TX, Society of Petroleum Engineers (SPE), 1989, pp. 19–25.

[16] J. E. Dendy, Jr., and C. C. Tazartes, *Grandchild of the frequency decomposition multigrid method*, SIAM J. Sci. Comput., 16 (1994), pp. 307–319.

[17] M. Dryja, B. Smith, and O. Widlund, *Schwarz analysis of iterative substructuring algorithms for problems in three dimensions*, SIAM J. Numer. Anal., 31 (1994), pp. 1662–1694.

[18] M. Dryja and O. B. Widlund, *Towards a unified theory of domain decomposition algorithms for elliptic problems*, in Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, T. F. Chan, R. Glowinski, J. Periaux, and O. B. Widlund, eds., SIAM, Philadelphia, PA, 1990, pp. 3–21.

[19] B. Engquist and E. Luo, *Convergence of a multigrid method for elliptic equations with highly oscillatory coefficients*, SIAM J. Numer. Anal., 34 (1997), pp. 2254–2273.

[20] J. Fuhrmann, *A modular algebraic multilevel method*, in Virtual Proceedings of the Ninth International GAMM-Workshop on Parallel Multigrid Methods, 1996.

[21] GMD software, http://mgnet.ccs.uky.edu/mgnet/Codes/gmd.

[22] W. Hackbusch, *Ein iteratives Verfahren zur schnellen Auflosung elliptischer Randwert probleme*, Technical Report 76-12, Universität zu Köln, Köln, Germany, 1976.

[23] W. Hackbusch, *Multi-grid Methods and Applications*, Springer-Verlag, Berlin, 1985.

[24] W. Hackbusch, *The frequency decomposition multigrid method, part* I: *Application to anisotropic equations*, Numer. Math., 56 (1989), pp. 229–245.
[25] T. Y. Hou and X. H. Wu, *A multiscale finite element method for elliptic problems in composite materials and porous media*, J. Comput. Phys., 134 (1997), pp. 169–189.
[26] T. Y. Hou, X. H. Wu, and Z. Cai, *Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients*, Math. Comp., 68 (1996), pp. 913–943.
[27] R. Kettler, *Analysis and comparison of relaxation schemes in robust multigrid and preconditioned conjugate gradient methods*, in Multigrid Methods, W. Hackbusch and U. Trottenberg, eds., Springer-Verlag, New York, 1981, pp. 502–534.
[28] R. Kettler and J.A. Meijerink, *A Multigrid Method and a Combined Multigrid-Conjugate Gradient Method for Elliptic Problems with Strongly Discontinuous Coefficients in General Domains*, Shell Publ. 604, KSEPL, Rijswijk, The Netherlands, 1981.
[29] S. Knapek, *Matrix-dependent multigrid-homogenization for diffusion problems*, SIAM J. Sci. Comput., 20 (1998), pp. 515–533.
[30] E. Luo, *Multigrid Method for Elliptic Equation with Oscillatory Coefficients*, Ph.D. thesis, Department of Mathematics, UCLA, Los Angeles, 1993.
[31] J. Mandel, M. Brezina, and P. Vaněk, *Energy Optimization of Algebraic Multigrid Bases*, Technical Report UCD/CCM 125, Center for Computational Mathematics, University of Colorado, Denver, CO, 1998.
[32] S. McCormick, *Multigrid Methods*, SIAM, Philadelphia, PA, 1988.
[33] A. Reusken, *Multigrid with matrix-dependent transfer operators for a singular perturbation problem*, Computing, 50 (1993), pp. 199–211.
[34] A. Reusken, *Multigrid with matrix-dependent transfer operators for convection-diffusion problems*, in Multigrid Methods, vol. IV, Internat. Ser. Numer. Math. 116, Birkhäuser-Verlag, Basel, 1994, pp. 269–280.
[35] J. W. Ruge and K. Stüben, *Algebraic multigrid*, in Multigrid Methods, S. McCormick, ed., SIAM, Philadelphia, PA, 1987, pp. 73–130.
[36] B. Smith, P. Bjorstad, and W. Gropp, *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations*, Cambridge University Press, Cambridge, UK, 1996.
[37] R. A. Smith and A. Weiser, *Semicoarsening multigrid on a hypercube*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 1314–1329.
[38] K. Stüben, *Algebraic multigrid: Experience and comparisons*, Appl. Math. Comput., 13 (1983), pp. 419–451.
[39] S. Ta'asan, *Multigrid Methods for Highly Oscillatory Problems*, Ph.D. thesis, Weizmann Institute of Science, Rehovat, Israel, 1984.
[40] P. Vaněk, J. Mandel, and M. Brezina, *Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems*, Computing, 56 (1996), pp. 179–196.
[41] W. L. Wan, *Scalable and Multilevel Iterative Methods*, Ph.D. thesis, Department of Mathematics, UCLA, Los Angeles, CA, 1998.
[42] P. Wesseling, *An Introduction to Multigrid Methods*, Wiley, Chichester, UK, 1992.
[43] J. Xu, *Theory of Multilevel Methods*, Ph.D. thesis, Cornell University, Ithaca, NY, 1989.
[44] J. Xu, *Iterative methods by space decomposition and subspace correction*, SIAM Rev., 34 (1992), pp. 581–613.
[45] P. M. De Zeeuw, *Matrix-dependent prolongations and restrictions in a blackbox multigrid solver*, J. Comput. Appl. Math., 33 (1990), pp. 1–27.