ELSEVIER

Journal of Computational and Applied Mathematics 123 (2000) 323–352

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.nl/locate/cam

Robust multigrid methods for nonsmooth coefficient elliptic linear systems

Tony F. Chan[a,*,1], W.L. Wan[b,2]

[a] *Department of Mathematics, University California at Los Angeles, Los Angeles, CA 90095-1555, USA*
[b] *SCCM Program, Gates Building 2B, Stanford University, Stanford, CA 94305-9025, USA*

Received 15 September 1999

Abstract

We survey the literature on robust multigrid methods which have been developed in recent years for solving second-order elliptic PDEs with *nonsmooth* coefficients. We highlight the key ideas of designing robust multigrid methods which are able to recover the usual multigrid efficiency for nonsmooth coefficient PDEs on structured or unstructured grids. In particular, we shall describe various approaches for constructing the interpolation and the smoothing operators, and the coarse grid points selections. © 2000 Elsevier Science B.V. All rights reserved.

1. Introduction

Multigrid methods are multilevel techniques for solving partial differential equations (PDEs) by eliminating errors in different parts of the spectrum on a sequence of coarse grids, or more generally, coarse subspaces. The basic principle is based on the interplay of smoothing and coarse grid correction which complement each other; the smooth errors not being reduced by smoothing are eliminated by coarse grid corrections. These techniques can generally be applied directly to PDEs but are of most interest when applied to the linear systems arising from their discretizations. Multigrid methods have been widely used in a broad variety of applications, from Poisson equations to full Navier–Stokes equations, from two-dimensional square domains to three-dimensional unstructured airfoil grids, etc. Multigrid has proved itself as a powerful and successful numerical technology for fast and efficient computations. In contrast with many other iterative methods such as classical relaxation methods, multigrid offers the capability of solving elliptic PDE problems with complexity

* Corresponding author.
E-mail addresses: chan@math.ucla.edu (T.F. Chan), wan@sccm.stanford.edu (W.L. Wan).

[1] Partially supported by the NSF under Contract ASC-9720257, NASA Ames under Contract NAG2-1238, and Sandia National Laboratory under Contract LG-4440.

[2] Partially supported by the Department of Energy under the Accelerated Strategic Computing Initiative (ASCI).

0377-0427/00/$ - see front matter © 2000 Elsevier Science B.V. All rights reserved.
PII: S0377-0427(00)00411-8

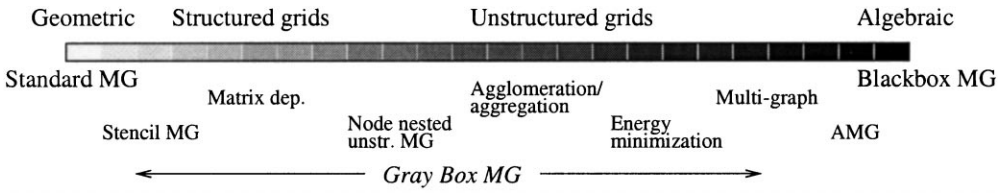

Fig. 1. A spectrum of multigrid methods.

and storage proportional to the number of unknowns, and its convergence rate is often independent of the problem size.

In this article, we survey robust multigrid methods in the literature which have been developed in recent years for solving second-order elliptic PDEs with *nonsmooth* coefficients. While multigrid converges rapidly for model problems such as the Poisson equation on a square, its convergence rate can be severely affected by PDEs with highly nonsmooth coefficients, or problems defined on complicated geometries, and unstructured grids. Unfortunately, these types of problems often arise in industrial applications, and hence traditional multigrid methods must be redesigned for them. The primary focus of this paper is in the design of robust multigrid methods which are able to retain the usual multigrid efficiency for smooth coefficient PDEs on structured grids. In particular, we shall describe various approaches for constructing the interpolation and the smoothing operators, and the coarse grid points selections.

General surveys of multigrid methods for solving different kinds of applications can be found in Brandt [20,21,23]. A survey on multilevel methods on unstructured grids can be found in Chan et al. [28]. Also, see [61,70] for a survey of parallel implementation of multigrid, which is not within the scope of this paper. Surveys on other aspects of multigrid methods can be found in [24,54,60,93]. We also note that the introductory note by Wagner [90] contains a lot of the details of the interpolation approaches discussed in this paper. Finally, we refer the readers to MGNet `http://www.mgnet.org` for a database of an extensive collection of multigrid papers in the literature.

The idea of multigrid was introduced and analyzed by Brakhage [15], and Fedorenko [47,48] in the 1960s, followed by Bachvalov [5]. Multigrid methods have not been paid much attention in the 1970s until the works of Astrachancer [2], Bank and Dupont [6], Brandt [19], Hackbusch [53], Nicolaides [74], and others showed that multigrid is indeed a very useful technique practically and theoretically. An enormous amount of progress has been achieved since then. Various multigrid methods have been developed, ranging from geometry specific to purely algebraic black box methods, and a spectrum of methods exist between the two extremes; see Fig. 1. We refer to this spectrum of methods as *gray box* methods: they require more information about the problem (e.g. grids, matrix graph, etc.) than a complete black box approach, but on the other hand, they can produce better robustness and performance.

Close to the geometric-dependent end of the spectrum where Cartesian grid was used, Alcouffe et al. [1] was one of the earliest papers to address the issue of nonsmooth coefficient PDE problems, and proposed robust interpolation methods for multigrid; see also [62]. Along this line were also the black box multigrid method by Dendy [38,39], and matrix-dependent approaches by de Zeeuw [102] and Reusken [76,77]. Other related approaches include frequency decomposition by Hackbusch [56], and filtering decomposition by Wittum [96,97]. The purely algebraic methods, on the other end of

the spectrum, were first proposed by Brandt et al. [25], and then popularized by Ruge and Stüben [79]. There is a recent resurgence of interest in AMG and other multigrid algorithms with focuses on parallel implementation and memory hierarchy aspects [26,36,37,43,44,64,75,87]. An introduction to AMG is recently given by Stüben [84]. See also the algebraic multilevel methods by Axelsson and Vassilevski [3,4], and an additive version of AMG by Grauschopf et al. [49]. The geometric unstructured multigrid methods were studied by Bank and Xu [9], Chan et al. [32], Guillard [52], Lallemand et al. [65], Morano et al. [72], and Xu [100]. The recent interest in energy minimization was studied by Brezina et al. [26], Brezina et al. [67,68], Chan et al. [33] with a local minimization perspective, and Wan et al. [92] with a global minimization perspective. Another recent interest is in the direction of bridging the gap between Gaussian elimination and multigrid; see [8,78]. Other multilevel methods include the hierarchical basis multigrid methods proposed by Yserentant [101], and Bank et al. [7], and the BPX method proposed by Bramble, Pasciak and Xu [17]. In Griebel [50], multilevel methods including multigrid and BPX were viewed as iterative methods on semidefinite systems. General multigrid references can be found in the books of Bramble [16], Briggs [27], Hackbusch [55], Smith et al. [80], and Wesseling [94]. Finally, we note that we are not able to survey the many more references in the literature here.

This paper is organized as follows: Section 1 begins with the basic principles of multigrid, and its classical convergence analysis. The design of robust multigrid will be discussed component by component. In Section 2, the construction of various sophisticated interpolation operators is described. Section 3 concerns the robustness and efficiency of smoothers. Algebraic and geometric coarsening strategies are covered in Section 4. Finally, Section 5 summarizes the current and future research on robust multigrid methods for elliptic linear systems.

In the rest of this section, we introduce the model problem and notation used in this paper, followed by the standard multigrid algorithm and the classical convergence analysis for smooth coefficient problems.

1.1. Elliptic PDEs

Elliptic PDE problems are among the most extensively investigated problems in applied mathematics. Their relation to many physical models is well known and the theoretical and numerical results obtained in this area are very useful in practice. The design of numerical methods for such model problems can often be adapted and applied to more complicated situations. Elliptic problems are also important in their own right, for instance, in the solution of the pressure equation arising from incompressible fluid problems, implicit time integration schemes, etc.

The model problem of primary interest is the following elliptic PDE which exhibits the fundamental properties and challenges that the elliptic problems above generally experience:

$$-\nabla \cdot a(x)\nabla u(x) = f(x) \quad x \in \Omega,$$
$$u = 0 \quad x \in \partial\Omega,$$

where $\Omega \subset \mathbb{R}^d$, $d=2,3$, is a polygonal or polyhedral domain, and $a(x)$, in general, is a $d \times d$ symmetric positive-definite matrix whose eigenvalues are bounded uniformly on $\bar{\Omega}$, and its coefficients can be oscillatory or discontinuous with large jumps across the interfaces. We note that Dirichlet boundary condition is used just for simplicity, and other boundary conditions are also permissible.

Many of the multigrid methods discussed in this paper apply to the discretization matrices given by finite element, finite difference or finite volume methods. For easy exposition, we set up notations based on finite element discretization. Let $H^1(\Omega)$ be the standard Sobolov space consisting of square integrable functions with square integrable derivatives of first order, and $H_0^1(\Omega)$ the subspace of $H^1(\Omega)$ whose functions vanish on $\partial\Omega$. Solving the PDE problem is equivalent to finding $u \in H_0^1(\Omega)$ such that

$$a(u,v) = (f,v) \quad \forall v \in H_0^1(\Omega), \tag{1}$$

where

$$a(u,v) = \int_\Omega a(x)\nabla u \cdot \nabla v \, \mathrm{d}x, \quad (f,v) = \int_\Omega fv \, \mathrm{d}x.$$

Suppose Ω is triangulated by quasi-uniform nonoverlapping simplices τ_i with size h, i.e., $\Omega = \bigcup_i \tau_i$. Define the finite element subspace by

$$V^h = \{v^h \in H_0^1(\Omega) \colon v^h|_{\tau_i} \in P_1(\tau_i), \ \forall i\},$$

where $P_1(\tau_i)$ is the set of linear functions on τ_i. The finite element approximation is the function $u^h \in V^h$ such that

$$a(u^h, v^h) = (f, v^h) \quad \forall v^h \in V^h. \tag{2}$$

Define a linear operator $A^h : V^h \to V^h$ by

$$(A^h u^h, v^h) = a(u^h, v^h) \quad u^h, v^h \in V^h.$$

Then (2) is equivalent to

$$A^h u^h = f^h, \tag{3}$$

where f^h is the projection of f on V^h. Let $\{\phi_j^h\}_{j=1}^n$ be the set of nodal basis of V^h. Write $u^h = \sum_{j=1}^n \mu_j^h \phi_j^h$, and $f^h = \sum_{j=1}^n b_j \phi_j^h$. Then (3) is equivalent to the linear system

$$\mathscr{A}^h \mu^h = b^h, \tag{4}$$

where $\mathscr{A}^h$ is the stiffness matrix, $\mu^h = (\mu_1^h, \ldots, \mu_n^h)^{\mathrm{T}}$, $b^h = \mathscr{M}^h(b_1^h, \ldots, b_n^h)^{\mathrm{T}}$, and $\mathscr{M}^h$ the mass matrix. It is well known that the condition number of $\mathscr{A}^h$ grows in the order of $\mathrm{O}(h^{-2})$, and hence classical iterative methods converge very slowly for large-scale problems. In the next sections, we describe a fast solution procedure – *multigrid* – for solving (4) whose convergence rate is often independent of the mesh size h.

Remark. We distinguish A^h, a linear operator of V^h, from $\mathscr{A}^h$, the corresponding stiffness matrix. The multigrid algorithms discussed in this paper are fast solution methods for solving the matrix equation (4).

1.2. Basic principles of multigrid

The idea of multigrid consists of two main components: *smoothing and coarse grid correction*. The smoothing process, usually carried out by a few iterations of a relaxation method, damps away the high frequency error components. The coarse grid correction process, carried out by a restriction,

a coarse grid solve, and an interpolation, eliminates the low-frequency error components. Hence, the key of multigrid is that smoothing and coarse grid correction complement each other. As a result, the combination of the two yields a significant error reduction, resulting in a fast solution procedure. Moreover, we gain efficiency since the coarse grid solves are less expensive than the fine grid one. The two-grid solution process is made more precise in the following.

We begin with an initial guess μ^h and we smooth the error by applying a relaxation iteration

$$\mu_1^h = \mu^h + \mathscr{R}^h(b^h - \mathscr{A}^h\mu^h), \tag{5}$$

where $\mathscr{R}^h$ is the approximate inverse of $\mathscr{A}^h$ given by the relaxation method. Then, we improve μ_1^h by a coarse grid correction consisting of the following steps (Eqs. (6)–(8)): First, we restrict the residual to the coarse grid V^H:

$$r^H = \mathscr{I}_h^H(b^h - \mathscr{A}^h\mu_1^h), \tag{6}$$

where $\mathscr{I}_h^H$ is the restriction operator. Second, we solve the coarse grid error equation:

$$\mathscr{A}^H e^H = r^H, \tag{7}$$

where the coarse grid matrix $\mathscr{A}^H$ is formed by the Galerkin process: $\mathscr{A}^H = \mathscr{I}_h^H \mathscr{A}^h \mathscr{I}_H^h$. The Galerkin approach can be shown to be optimal for symmetric positive-definite elliptic problems. Here $\mathscr{I}_H^h = (\mathscr{I}_h^H)^{\mathrm{T}}$ is the interpolation operator. The coarse grid error is then interpolated back to the fine grid and the fine grid approximation is updated by

$$\mu_2^h = \mu_1^h + \mathscr{I}_h^H e^H. \tag{8}$$

Finally, we apply a post-smoothing at the end:

$$\mu_{\mathrm{new}}^h = \mu_2^h + \mathscr{R}^h(b^h - \mathscr{A}^h\mu_2^h). \tag{9}$$

Combining (5)–(9), the entire process can be summarized by the following result.

Lemma 1.1. *The iteration matrix M_{tg} of the two-grid method with ν_1 steps of pre-smoothing and ν_2 steps of post-smoothing is given by*

$$M_{\mathrm{tg}} = (I - \mathscr{R}_2^h\mathscr{A}^h)^{\nu_2}(I - \mathscr{I}_h^H(\mathscr{A}^H)^{-1}\mathscr{I}_h^H\mathscr{A}^h)(I - \mathscr{R}_1^h\mathscr{A}^h)^{\nu_1}, \tag{10}$$

where $\mathscr{R}_1^H$ and $\mathscr{R}_2^h$ denote the approximate inverses of $\mathscr{A}^h$ given by the pre- and post-smoother, respectively.

If we solve the coarse grid problem (7) recursively by the same two-grid method, we obtain a multigrid algorithm. We use the following notations for the multilevel methods in the next sections. Let $V_J = V^h$ be the fine grid space, and $V_1 \subset V_2 \subset \cdots \subset V_J$ be a sequence of nested coarse grid subspace of V^h. On each level k, $A^k : V_k \to V_k$ is the restriction of $A^J = A^h$ on V_k, and $R^k : V_k \to V_k$ is an approximate inverse of A^k given by the smoother. Their corresponding stiffness matrix is denoted by $\mathscr{A}^k$ and $\mathscr{R}^k$, respectively. Let $I_{k-1}^k : V_{k-1} \to V_k$ be the interpolation operator and its adjoint $I_k^{k-1} : V_k \to V_{k-1}$ the restriction operator. Their matrix representation is denoted by $\mathscr{I}_{k-1}^k$ and $\mathscr{I}_k^{k-1} = (\mathscr{I}_{k-1}^k)^{\mathrm{T}}$, respectively.

In addition, associated with each A^k, we define the A-inner product by $(\cdot,\cdot)_{A^k} \equiv (A^k\cdot,\cdot)$. Let $Q_k : V_j \to V_k$ and $P_k : V_J \to V_k$ be the projection operators with respect to the L^2 and A inner products, respectively.

1.3. Convergence theory

In view of (10), the classical convergence analysis involves the norm estimates of

$$\|I - \mathscr{R}^h \mathscr{A}^h\| \quad \text{and} \quad \|I - \mathscr{I}_h^H (\mathscr{A}^H)^{-1} \mathscr{I}_h^H \mathscr{A}^h\|.$$

See [16,22,55,71] for details. Here, we summarize the results in the literature based on the subspace correction framework developed by Xu [99]. The convergence of multigrid is governed by two constants K_0 and K_1 defined as

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

$$\sum_{i=1}^{J} (R_i^{-1} v_i, v_i) \leqslant K_0 (Av, v), \tag{11}$$

where R_i is the approximate inverse operator given by the smoother.

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

$$\sum_{(i,j)\in S} (T_i u_i, T_j u_j)_A \leqslant K_1 \left(\sum_{i=1}^{J} (T_i u_i, u_i)_A \right)^{1/2} \left(\sum_{j=1}^{J} (T_j v_j, v_j)_A \right)^{1/2}, \tag{12}$$

where $T_i = R_i A_i P_i$.

Theorem 1.2. *Let M_{mg} be the iteration matrix given by the V-cycle* multigrid. Then

$$\|M_{\mathrm{mg}}\|_A^2 \leqslant 1 - \frac{2 - \omega_1}{K_0 (1 + K_1)^2},$$

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

Proof. See [99]. □

By Theorem 1.2, the convergence rate can be improved by producing a smaller K_0 or K_1. By definition, it can be easily proved the following result.

Lemma 1.3.

$$K_1 \leqslant \omega_1 J.$$

Proof. See [99]. □

Thus, the estimate of K_0 is crucial. We analyze K_0 from the domain decomposition perspective. For second order scaler elliptic PDEs, K_0 depends on two inequalities:

$$\|Q_1 v\|_A^2 + \sum_{k=2}^{J} \|(Q_k - Q_{k-1}) v\|_A^2 \leqslant C_0 \|v\|_A^2, \tag{13}$$

$$\|(Q_k - Q_{k-1}) v\| \leqslant C_0 h_{k-1} \|Q_k v\|_A, \quad k > 1, \tag{14}$$

where $Q_k : V \to V_k$ is the L^2 projection. More precisely, we have the following estimate:

Lemma 1.4. *Suppose* (13) *and* (14) *are satisfied. Then*

$$K_0 \leqslant \frac{C_0}{\omega_0},$$

where $\omega_0 = \min_{1 \leqslant k \leqslant j} \lambda_{\min}(\mathscr{R}_k^{-1} \mathscr{A}_k)$.

Proof. See [99]. □

The *stability* inequality (13) is known as the partition lemma [66,69] which plays an essential role in the convergence analysis of domain decomposition methods. It requires that for any given $v \in V$, we must be able to decompose it into $v_k \in V_k$ such that the total energy of all the pieces v_k is bounded by a small constant factor of the original energy of v. In the multigrid context, it can be translated into the following: the coarse grid basis functions must have small energy. The *approximation* inequality (14) requires that the functions on the coarse grids approximate the fine grid functions to at least first-order accuracy. A sufficient condition is that the coarse subspace contains constant functions.

In conclusion, one major approach of improving robustness is to devise multigrid methods which lead to a small K_0. For instance, the constructions of the robust interpolation operators described in Section 2.5 are based on the stability and approximation inequalities.

1.4. Multigrid for nonsmooth coefficient PDEs

The success of multigrid hinges on the choice of coarse grids, and the smoothing, interpolation and coarse grid operators. In standard multigrid, full coarsening, damped Jacobi or Gauss–Seidel smoothing, and linear interpolation are often used. Classical convergence theory and practice shows that these simple choices are enough to achieve mesh independent convergence.

For PDE problems with nonsmooth coefficients, however, mesh-independent convergence does not necessarily result in fast convergence. The nonsmoothness of the PDE coefficients typically lead to a large constant C_0 in (13) and (14). Thus, multigrid converges slowly when the coefficients exhibit anisotropy [55], large jumps in discontinuity [1,19,38,39], or large oscillations [46,85]. Special techniques such as line Gauss–Seidel [19], semi-coarsening [40,41,81], algebraic multigrid [14,25,76,79,83], frequency decomposition [42,56,85], and homogenization [46], are used to handle some of these cases. In the next sections, we survey the state-of-the-art of each individual multigrid components and discuss how they bring insight into the design of robust multigrid methods.

2. Interpolation

Sophisticated designs of interpolation have been the key in developing robust multigrid methods. The many different methods can be generally divided into four categories ranging from geometric specific to purely algebraic. The structured grid approach takes advantages of the special PDE and algebraic structures associated with the Cartesian grids. The unstructured grid approach exploits the given grid information to derive interpolations. The algebraic multigrid approach, on the other

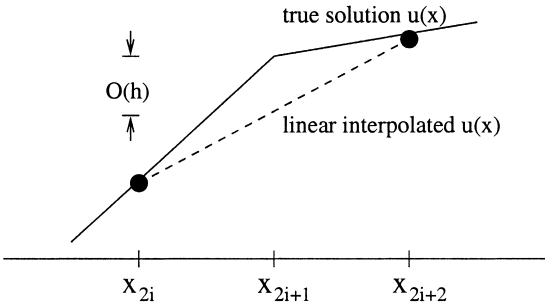

Fig. 2. Linear interpolation makes an O(h) error for a typical solution of the PDEs whose coefficient is piecewise constant.

hand, focuses on the algebraic aspect and derives interpolation from the residual equations. A recent approach constructs interpolation based on energy minimization which exploits the properties of the underlying PDEs while allowing general computational domains.

In the following, without loss of generality, we only discuss interpolation from coarse to fine, since the Galerkin process will automatically generate a multigrid method; see Section 1.2. Thus, superscripts h or H are used to denote quantities in the fine or coarse grid. Moreover, we sometimes describe the construction of coarse grid basis functions rather than the interpolation operators since they are essentially the same. In the finite element context, the coarse grid space V^H is often a subspace of V^h. Thus, if $\{\phi_j^h\}_{j=1}^n$ and $\{\phi_i^H\}_{i=1}^m$ are the nodal basis for V^h and V^H, respectively, then we have the following equality:

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

where $\mathcal{I}_h^H$ is the interpolation matrix. Hence, the set of coarse grid basis functions defines an interpolation, and vice versa. In particular, in the subsequent sections on the agglomeration unstructured grid approach and energy minimization approach, we shall describe the constructions of the coarse grid basis in place of interpolation. We note that the coarse grid subspaces need not be nested, for instance, in geometric unstructured grid multigrid methods. However, multigrid methods resulting from nested coarse subspaces are generally more robust, and hence we shall focus on this case in the next sections.

Before going on, we first discuss a well-known interpolation technique in one dimension, which is one of the earliest attempts to construct robust interpolation operator for nonsmooth coefficients. It turns out the basic design strategies in higher dimensions can be viewed as trying to extend this one-dimensional approach.

2.1. One dimension

For nonsmooth coefficient PDEs, linear interpolation is not able to accurately approximate the irregular shape of the numerical solutions during the multigrid process. For example, Fig. 2 shows a typical solution of the PDEs whose coefficient is piecewise constant. In the worst case, linear interpolation can make an O(h) error, which is much poorer than the usual O(h^2) error.

A robust interpolation can be constructed by solving local PDEs [55]. Given the values v_{2i} and v_{2i+2} at the coarse grid points x_{2i} and x_{2i+2}, respectively, the value v_{2i+1} is computed by solving a

homogeneous two-point boundary value problem:

$$-\frac{\mathrm{d}}{\mathrm{d}x}a(x)\frac{\mathrm{d}}{\mathrm{d}x}v(x)=0, \quad x\in(x_{2i},x_{2i+2}),$$
$$v(x_{2i})=v_{2i}, \quad v(x_{2i+2})=v_{2i+2}. \tag{15}$$

Suppose $a(x)$ is piecewise constant, for instance, $a(x)\equiv a^-$, $x_{2i}<x<x_{2i+1}$, and $a(x)\equiv a^+$, $x_{2i+1}<x<x_{2i+2}$. Then the finite element solution of (15) yields

$$v_{2i+1}=\frac{a^-}{a^-+a^+}v_{2i}+\frac{a^+}{a^-+a^+}v_{2i+2}. \tag{16}$$

The new interpolated solution is more accurate at the discontinuities. It is well known that the resulting multigrid is very robust and converges rapidly for nonsmooth coefficient $a(x)$.

The local PDE approach has the property of preserving flux continuity. It can be proved [55] that the interpolated v given by (16) satisfies the jump condition

$$\lim_{x\to x_{2i+1}^-} a(x)v'(x)=\lim_{x\to x_{2i+1}^+} a(x)v'(x)$$

at x_{2i+1} which the exact solution does. In fact, the converse is also true; that is, if v satisfies the jump condition, then it solves the local PDE (15).

The interpolation can be interpreted by pure linear algebra. Ordering the noncoarse grid points μ_F and then the coarse grid points μ_C, we can write the permuted matrix, still denoted by $\mathscr{A}^h$, in a 2×2 block form

$$\begin{bmatrix}\mathscr{A}_{11} & \mathscr{A}_{12}\\ \mathscr{A}_{21} & \mathscr{A}_{22}\end{bmatrix}\begin{bmatrix}\mu_F\\ \mu_C\end{bmatrix}=\begin{bmatrix}b_F\\ b_C\end{bmatrix},$$

where $\mathscr{A}_{11}$ is a diagonal matrix. After eliminating μ_F, we obtain the Schur complement equation for μ_C:

$$\mathscr{S}\mu_C=b_C-\mathscr{A}_{21}\mathscr{A}_{11}^{-1}b_F, \tag{17}$$

where $\mathscr{S}=\mathscr{A}_{22}-\mathscr{A}_{21}\mathscr{A}_{11}^{-1}\mathscr{A}_{12}$. Define the interpolation and restriction matrices, respectively, by

$$\mathscr{I}_H^h=\begin{bmatrix}-\mathscr{A}_{11}^{-1}\mathscr{A}_{12}\\ I\end{bmatrix}, \quad \text{and} \quad \mathscr{I}_h^H=(\mathscr{I}_H^h)^{\mathrm{T}}=[-\mathscr{A}_{21}\mathscr{A}_{11}^{-1}, \quad I]. \tag{18}$$

Then $\mathscr{S}=\mathscr{I}_h^H\mathscr{A}^h\mathscr{I}_H^h$ is precisely the coarse grid matrix $\mathscr{A}^H$ in the multigrid context, and (17) is the usual coarse grid equation for μ_C with the right-hand side given by restriction: $\mathscr{I}_h^H[b_F,b_C]^{\mathrm{T}}$. The noncoarse grid values μ_F are obtained by backward substitution:

$$\mu_F=-\mathscr{A}_{11}^{-1}\mathscr{A}_{12}\mu_C+\mathscr{A}_{11}^{-1}b_C. \tag{19}$$

Using (19), we can write $[\mu_F,\mu_C]^{\mathrm{T}}$ as

$$\begin{bmatrix}\mu_F\\ \mu_C\end{bmatrix}=\mathscr{I}_H^h\mu_C+\begin{pmatrix}\mathscr{A}_{11}^{-1} & 0\\ 0 & 0\end{pmatrix}(b^h-\mathscr{A}^h\mathscr{I}_H^h\mu_C). \tag{20}$$

Thus, the backward substitution can be interpreted as applying a Jacobi relaxation smoothing on the noncoarse grid points to the interpolated solution $\mathscr{I}_H^h\mu_C$.

We note that the block Gaussian elimination can be also written as block LU form:

$$\mathscr{A}^h = \begin{bmatrix} \mathscr{A}_{11} & \mathscr{A}_{12} \\ \mathscr{A}_{21} & \mathscr{A}_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ \mathscr{A}_{21}\mathscr{A}_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} \mathscr{A}_{11} & 0 \\ 0 & \mathscr{A}^H \end{bmatrix} \begin{bmatrix} I & \mathscr{A}_{11}^{-1}\mathscr{A}_{12} \\ 0 & I \end{bmatrix},$$

where the interpolation $\mathscr{I}_H^h$ and the restriction $\mathscr{I}_h^H$ operators are related to the upper and lower triangular factors, respectively.

To summarize, the block Gaussian elimination process is a two-level multigrid with $\mathscr{I}_H^h$ and $\mathscr{I}_h^H$ defined as in (18), and post-smoothing step (20). Furthermore, the inversion of the Schur complement in (17) can be done by recursively applying the previous procedure to $\mathscr{S}$. The resulting algorithm is known as cyclic reduction, and the corresponding multigrid method is a backslash cycle [99]. Moreover, it can be easily verified that the particular matrix-dependent interpolation matrix defined in (18) is precisely the one obtained by solving local PDEs.

Lastly, one may interpret the local PDE solve as an energy minimization to be described in Section 2.5. These four interpretations of the same interpolation: local PDE solve, flux continuity, the Schur complement, and energy minimization constitute the basic design principles of constructing the robust interpolations discussed in the following sections.

Finally, we remark that convergence analysis of robust multigrid methods in general is very limited in the literature since the interpolation operator is usually complicatedly defined. However, in one dimension, we have the following result [91].

Theorem 2.1. *If the interpolation operator discussed above is used together with damped Jacobi or Gauss–Seidel smoothers, the resulting multigrid convergence is independent of the mesh size and the PDE coefficient* $a(x)$.

The proof uses the fact that the coarse grid basis functions from all the levels form an A-orthogonal hierarchical basis, and hence the damped Jacobi and Gauss–Seidel smoothers give an optimal constant bound for K_0 and K_1 which are the essential elements for estimating multigrid convergence by Theorem 1.2. Details can be found in [91].

2.2. Structured grid approaches

Structured grids, in particular, Cartesian grids, have been very popular in applications for their regularity in geometry and in the algebraic structure of the resulting discretization matrix. Thus efficient numerical methods can be easily derived and employed. This approach also includes nested finite element grids obtained by recursive refinement where the discontinuities of the PDE coefficients are aligned with all the coarse grids. As a result, according to the domain decomposition theory [18,45,98], it can be proved that the convergence rate is independent of the size of the jumps in the coefficient using even the linear interpolation. However, in general, the discontinuities do not align with some of the coarse grids. Then other approaches are needed; see Sections 2.3–2.5.

Assuming Cartesian grids, the structured grid multigrid methods can generally be categorized into two: the stencil and Schur complement approaches.

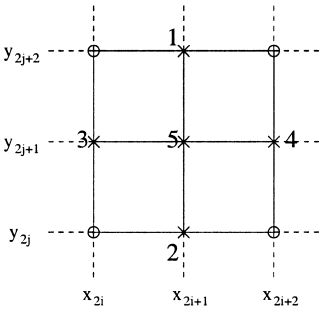

Fig. 3. A portion of a fine grid with coarse grid points denoted by circles, and noncoarse grid points by crosses.

2.2.1. Stencil

The one-dimensional local PDE technique cannot be applied directly to higher dimensions. Consider a portion of the fine grid with coarse grid points denoted by circles as shown in Fig. 3. In contrast to one dimension, the noncoarse grid points are not enclosed by only coarse grid points, and hence a local boundary value problem similar to (15) cannot be set up. The challenge of extending the one-dimensional technique to higher dimensions is to set up local PDEs appropriately.

Alcouffe et al. [1] used special harmonic averaging techniques to construct operator-induced interpolations, and in black box multigrid, Dendy [38,39] simplified the interpolation procedure by considering the stencil of the discrete operators. The key of the stencil approach is to first define interpolation on edges, i.e. noncoarse grid points lying on coarse grid lines (nodes 1–4 in Fig. 3), and then the interpolation at the interior (node 5) can be defined by a local PDE solve. In particular, black box multigrid defines the interpolation on the edges by preserving the continuity of flux across the interfaces. Thus, this multigrid method is efficient for PDE problems with discontinuous coefficients; the convergence rate is often independent of the mesh size and the size of the jumps.

The construction is as follows. Consider the nine-point stencil at the noncoarse grid points, for instance, node 1 (Fig. 3) lying on a horizontal edges:

$$\begin{bmatrix} \alpha_{NW}^{(1)} & \alpha_{N}^{(1)} & \alpha_{NE}^{(1)} \\ \alpha_{W}^{(1)} & \alpha_{C}^{(1)} & \alpha_{E}^{(1)} \\ \alpha_{SW}^{(1)} & \alpha_{S}^{(1)} & \alpha_{SE}^{(1)} \end{bmatrix}.$$

The first and third rows are lumped to the second row, thus producing a one-dimensional three-point stencil. The flux preserving interpolation (16) yields

$$\tilde{v}_{2i+1,2j+1}^{-h} = \frac{\alpha_{NW}^{(1)} + \alpha_{W}^{(1)} + \alpha_{SW}^{(1)}}{\alpha_{N}^{(1)} + \alpha_{C}^{(1)} + \alpha_{S}^{(1)}} v_{i,j+1}^{H} + \frac{\alpha_{NE}^{(1)} + \alpha_{E}^{(1)} + \alpha_{SE}^{(1)}}{\alpha_{N}^{(1)} + \alpha_{C}^{(1)} + \alpha_{S}^{(1)}} v_{i+1,j+1}^{H}. \tag{21}$$

The lumping is used to preserve the continuity of the flux on the average along the vertical segment through nodes 1 and 2. The interpolation on vertical coarse grid lines (nodes 3 and 4) are defined analogously. Finally, since all the values on the edges are known, the interpolated value at node 5

can be obtained by solving a local PDE problem as in (15) where the solution is given by

$$\tilde{v}_{2i+1,2j+1} = \frac{\alpha_{NW}^{(5)}}{\alpha_C^{(5)}} v_{i,j+1}^H + \frac{\alpha_N^{(5)}}{\alpha_C^{(5)}} \tilde{v}_{2i+1,2j+2}^h + \frac{\alpha_{NE}^{(5)}}{\alpha_C^{(5)}} v_{i+1,j+1}^H + \frac{\alpha_W^{(5)}}{\alpha_C^{(5)}} \tilde{v}_{2i,2j+1}^h + \frac{\alpha_E^{(5)}}{\alpha_C^{(5)}} \tilde{v}_{2i+2,2j+1}^h + \frac{\alpha_{SW}^{(5)}}{\alpha_C^{(5)}} v_{i,j}^H + \frac{\alpha_S^{(5)}}{\alpha_C^{(5)}} \tilde{v}_{2i+1,2j}^h + \frac{\alpha_{SE}^{(5)}}{\alpha_C^{(5)}} v_{i+1,j}^H. \tag{22}$$

Another stencil-based method is the matrix-dependent prolongation proposed by de Zeeuw [102]. It differs from the previous method in that the interpolated values on the edges are defined based on a decomposition of the nine-point stencil. Viewing the stencil as a 3×3 matrix, it can be written as a linear combination of nine basis matrices, or equivalently, stencils. De Zeeuw considered a particular set of basis stencils corresponding to the discretization of the first and second derivatives. The interpolation formula which depends on the coefficients of the linear combination is very technical, and we refer the interested readers to [90,102] for details. This approach coincides with the black box multigrid of Dendy [38] for solving the model equation (1), and can be directly applied to certain nonsymmetric problems such as convection diffusion equations.

2.2.2. Schur complement and lumping

In this approach, we exploit the special algebraic structure associated with the discretization matrix arising from Cartesian grids. As in one dimension, the five-point stencil matrix in two dimensions can be written in a 2×2 block form:

$$\mathscr{A}^h = \begin{bmatrix} \mathscr{A}_{11} & \mathscr{A}_{12} \\ \mathscr{A}_{21} & \mathscr{A}_{22} \end{bmatrix},$$

where $\mathscr{A}_{11}$ is diagonal if the red-black ordering is used. However, the coarse grid matrix $\mathscr{A}^H = \mathscr{S} = \mathscr{A}_{22} - \mathscr{A}_{21}\mathscr{A}_{11}^{-1}\mathscr{A}_{12}$ now corresponds to a nine-point stencil instead. Thus the algorithm cannot be repeated recursively.

In order to recover a five-point stencil structure, Reusken [76,77] applies a lumping strategy to the nine-point stencil coarse grid operator as follows. He replaced the nine-point stencil at a noncoarse grid point by a five-point stencil:

$$\begin{bmatrix} \alpha_{NW} & \alpha_N & \alpha_{NE} \\ \alpha_W & \alpha_C & \alpha_E \\ \alpha_{SW} & \alpha_S & \alpha_{SE} \end{bmatrix} \to \begin{bmatrix} 0 & \beta_N & 0 \\ \beta_W & \beta_C & \beta_E \\ 0 & \beta_S & 0 \end{bmatrix},$$

where

$$\beta_N = \alpha_N + \alpha_{NW} + \alpha_{NE}, \quad \beta_W = \alpha_W + \alpha_{NW} + \alpha_{SW},$$
$$\beta_C = \alpha_C - (\alpha_{NW} + \alpha_{NE} + \alpha_{SW} + \alpha_{SE}), \quad \beta_E = \alpha_E + \alpha_{NE} + \alpha_{SE},$$
$$\beta_E = \alpha_S + \alpha_{SW} + \alpha_{SE}.$$

The lumping procedure essentially substitutes the unknowns $v_{i-1,j+1}^H$, $v_{i+1,j+1}^H$, $v_{i-1,j-1}^H$, and $v_{i+1,j-1}^H$ by the unknowns $v_{i,j-1}^H$, $v_{i,j+1}^H$, $v_{i-1,j}^H$, $v_{i+1,j}^H$, and $v_{i,j}^H$ in the finite difference equation corresponding to the coarse grid point (x_i^H, y_j^H) based on a linear approximation; for instance, $v_{i-1,j+1}^H \approx -v_{i,j}^H + v_{i,j+1}^H + v_{i-1,j}^H$. In matrix form, the resulting discretization matrix becomes

$$\tilde{\mathscr{A}}^H = \begin{bmatrix} \tilde{\mathscr{A}}_{11} & \tilde{A}_{12} \\ \tilde{\mathscr{A}}_{21} & \tilde{A}_{22} \end{bmatrix},$$

where $\tilde{A}_{11}$ is now a diagonal matrix. Moreover, the interpolation and restriction operators given by (18) are local, and the entire procedure can be repeated recursively.

2.3. Unstructured grid approaches

Unstructured gridding, which has a high flexibility of capturing complex geometrical shapes and providing adaptive local refinements, are useful for solving problems involving rapidly changing solutions, irregular boundaries, and multiscale geometries. However, as a result, the computational grids do not have any particular nested grid hierarchical structure to be exploited. Thus, the structured grid multigrid methods must be redesigned to handle the irregularity without losing too much in terms of complexity and performance.

The two main difficulties of designing multigrid methods on unstructured grids are the extraction of a hierarchy of coarser grids from a given fine grid, and the definition of the interpolation operators between grids. In the following sections, we describe several approaches of solving the two problems with increasing mathematical structures and decreasing intuition.

2.3.1. Independent grids

The first approach is based on independently generated coarse grids and piecewise linear interpolation between the grids. Thus, one can use the grid generator which generates the unstructured fine grid to generate a sequence of coarser grids. Moreover, since the coarser grids consist of the usual finite elements, for instance, linear elements on triangles, linear interpolation and the coarse grid operator can be easily defined.

The advantage of this approach is convenience; the coarse grids can be generated by using the same grid generator which produced the original fine grid. The disadvantage is that the construction of the interpolation operator is very expensive since one has to identify which coarse triangles the noncoarse grid points are in. Another disadvantage is nonblack box nature of the coarse grid construction; the user is required to manually generate the grids.

2.3.2. Node nested grids

An alternative approach [29,32,52] is based on generating node-nested coarse grids, which are created by selecting subsets of a vertex set, retriangulating the subset, and using piecewise linear interpolation between the grids. This provides an automatic way of generating coarse grids and a simpler implementation ($O(n)$) of the interpolation. The main disadvantage is that critical geometrical details may be lost through the coarsening and retriangulation process, and hence special treatments are needed to preserve the important geometric features of the fine grid. Moreover, the coarse grid boundaries may not match that of the fine grid, and hence the boundary conditions must be incorporated properly, especially for Neumann boundary condition [30]. Another drawback is that in three dimensions, retetrahedralization can be problematic.

Remark. Both the independent grid and node nested grid approaches are not designed to be robust for PDE problems with nonsmooth coefficients since linear interpolation is used.

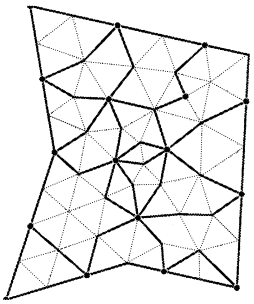

Fig. 4. Typical macroelements in a computational domain.

2.3.3. Agglomeration

To avoid the problem of losing geometrical details, a promising agglomeration technique [65,89] motivated by finite volume based methods is introduced. Instead of regenerating the coarse grids, neighboring fine grid elements are agglomerated together to form macroelements; see Fig. 4. For first order PDE problems, Venkatakrishnan and Mavriplis [89] used piecewise constant interpolation. More precisely, let τ_i be a macroelement and $\tau_i^H = \bigcup_{j\in N_i} \tau_j^h$, where N_i is the set of neighboring nodes. Then

$$v_j^h = c_i \equiv \text{ constant}, \quad j \in N_i.$$

However, their constant interpolation approach leads to slow convergence for second-order PDE problems since the basis is not stable. Within each macroelement, we need more robust weightings which mimic linear interpolation on structured grids.

In general, the coarse space V^H can be defined as the subspace spanned by a set of coarse grid basis functions $\{\phi_i^H\}$ constructed as follows. For each coarse grid point i, define

$$\phi_i^H = \sum_{j\in\tilde{N}_i} w_{ij}^h \phi_j^h + \phi_i^h, \tag{23}$$

where w_{ij}^h are appropriately chosen constants for robust interpolation, and

$$\tilde{N}_i = \{\, j\colon \mathscr{A}_{i,j}^h \neq 0, \text{ and line segment } [x_i, x_j] \text{ is an edge of a macroelement}\}.$$

Thus, the coarse grid basis functions are linear combinations of the fine grid basis, and V^H is a subspace of V^h; that is, we obtain a nested sequence of subspaces by recursive construction. Moreover, the interpolation weights are given by the coefficients w_{ij}^h. To summarize, the construction of V^H consists of two parts. We agglomerate the fine grid elements to form macroelements, and then we define robust coarse grid basis functions on the macroelements.

Smoothed aggregation: For second-order PDEs, the piecewise constant basis functions are not effective since they possess large energy norm due to the discontinuities. Vaněk et al. [88] proposed the smoothed aggregation approach which applies a relaxation method to *smooth* the piecewise constant basis, and hence reducing the energy norm. More precisely, similar to agglomeration, the computational nodes are aggregated into disjoint aggregates based on the AMG coarsening

technique (cf. Section 4.2). Trying to achieve the approximation property (14), one defines a tentative interpolation operator as the piecewise constant prolongator:

$$(\tilde{\mathscr{I}}_H^h)_{ij} = \begin{cases} 1 & \text{if } i \in \tau_j, \\ 0 & \text{otherwise,} \end{cases}$$

where τ_j is the jth aggregate. Since the induced piecewise constant basis functions exhibit large energy, the stability property (13) is violated. One may *smooth* the basis by applying a damped Jacobi smoother to $\tilde{\mathscr{I}}_H^h$ and obtain

$$\mathscr{I}_H^h = (I - \omega(\mathscr{D}^h)^{-1}\mathscr{A}^{\mathrm{F}})\tilde{\mathscr{I}}_H^h,$$

where $\mathscr{A}^{\mathrm{F}}$ is the *filtered matrix* of $\mathscr{A}^h$ defined as

$$\mathscr{A}_{ij}^{\mathrm{F}} = \begin{cases} \mathscr{A}_{ij}^h & \text{if } j \in N_i(\varepsilon),\ i \neq j, \\ \mathscr{A}_{ii}^h - \sum_{j=1, j\neq i}^{n} (\mathscr{A}_{ij}^h - \mathscr{A}_{ij}^{\mathrm{F}}) & i = j, \\ 0 & \text{otherwise} \end{cases}$$

and $N_i(\varepsilon) = \{j\colon |\mathscr{A}_{ij}^h| \geqslant \varepsilon\sqrt{\mathscr{A}_{ii}^h \mathscr{A}_{jj}^h}\}$. Basically, $\mathscr{A}^{\mathrm{F}}$ is obtained by lumping the small entries in $\mathscr{A}^h$ to the diagonal, thus controlling the number of nonzeros in the interpolation and coarse grid operators. Due to the smoothing effect of damped Jacobi, it smears the sharp edges of the coarse grid basis functions obtained from the piecewise constant prolongator, and hence the energies are reduced. Moreover, it can be proved that $\mathscr{I}_H^h$ preserves constant if $\tilde{\mathscr{I}}_H^h$ does.

2.3.4. Others

Other unstructured grid multigrid approaches have also been proposed. Bank and Xu [9] developed an effective coarsening and interpolation strategy using the geometrical coordinates of the fine grid. The basic idea is to treat the fine grid as if it came from a refinement procedure, and then recover the refinement structure through a symbolic Gaussian elimination. Another multigrid method based on incomplete Gaussian elimination was proposed by Reusken [78]. Hackbusch and Sauter [57] constructed a triangulation for the computational domain by adaptively refining a coarse triangulation of a rectangular domain covering the computational domain. Thus, a hierarchy of coarse grids is naturally embedded in the fine triangulation.

2.4. Algebraic multigrid approaches

The structured and unstructured grid approaches make use of the grid information either explicitly or implicitly and hence are geometry dependent. The algebraic multigrid (AMG) approach [79], on the other hand, exploits the algebraic information of the discretization matrix. This approach was first introduced by Brandt et al. [25] and later popularized by Ruge and Stüben [79]. Other related work have been studied by Huang [59], and Chang et al. [34] to extend AMG to matrices which are not symmetric M-matrices.

The success of AMG is that for symmetric positive-definite M-matrices, for instance, matrices arising from discretization of the Laplacian operator, AMG is able to identify algebraically the *smooth* errors obtained from standard relaxation methods such as Gauss–Seidel, and then construct

interpolation operators accordingly to eliminate such errors. In the following sections, we describe a definition of algebraic smooth errors and discuss how they motivate the construction of an interpolation operator.

2.4.1. Algebraic smoothness and strong connection

Let G^h be the iteration matrix of the relaxation smoother. In AMG, an error e^h is algebraically smooth if it is slow to converge with respect to G^h, i.e.,

$$\|G^h e^h\|_A \approx \|e^h\|_A.$$

For common relaxation smoothers, it can be argued [79] that an algebraically smooth error e^h is characterized by a small residual

$$r^h = \mathscr{A}^h e^h \approx 0,$$

in the sense that the residual norm is small compared to the error. Thus, we obtain a good approximation for e_i^h as a function of its neighboring values e_j^h by making $r_i^h = 0$:

$$r_i^h = \mathscr{A}_{ii}^h e_i^h + \sum_{j\in N_i} \mathscr{A}_{ij}^h e_j^h = 0, \tag{24}$$

where $N_i = \{j \neq i\colon \mathscr{A}_{ij}^h \neq 0\}$, the set of neighboring nodes of i. For symmetric M-matrices, the smooth error e^h often satisfies

$$\|e^h\|_A \ll \|e^h\|_D,$$

where $\|e^h\|_D^2 = (e^h, \mathscr{D}^h e^h)$, and $\mathscr{D}^h$ is the diagonal of $\mathscr{A}^h$. Note that $\|e^h\|_A$ essentially measures the norm of the residual. We have the following inequality:

$$\frac{1}{2}\sum_{i,j} -\mathscr{A}_{ij}^h (e_i^h - e_j^h)^2 + \sum_i \left(\sum_j \mathscr{A}_{ij}^h\right)(e_i^h)^2 \ll \sum_i \mathscr{A}_{ii}^h (e_i^h)^2. \tag{25}$$

If $\sum_{j\neq i} |\mathscr{A}_{ij}^h| \approx \mathscr{A}_{ii}^h$, for instance, $\mathscr{A}^h =$ Laplacian, then (25) can be written as

$$\sum_{j\neq i} \frac{\mathscr{A}_{ij}^h}{\mathscr{A}_{ii}^h} \frac{(e_i^h - e_j^h)^2}{(e_i^h)^2} \ll 1, \tag{26}$$

on the average for each i. Thus, if $|\mathscr{A}_{ij}^h/\mathscr{A}_{ii}^h|$ is relatively large, then e_i^h and e_j^h must be close, and hence e_j^h is not negligible compared to e_i^h. The nodes i and j are called strongly connected if $|\mathscr{A}_{ij}^h/\mathscr{A}_{ii}^h|$ is relatively large. This will be made more precise in (31). The strongly connectedness forms the basic notion for algebraic smoothing and interpolation.

2.4.2. Algebraic interpolation

Suppose $\mathscr{A}^h$ is a symmetric, weakly diagonally M-matrix. The derivation of the algebraic interpolation of Ruge and Stüben [79], again, stems from the idea of the one-dimensional interpolation, and has a strong connection with the stencil approach for Cartesian grids (Section 2.2.1). We start with the residual equation (24) corresponding to algebraic smooth errors where i is an index corresponding to a noncoarse grid point. Let C be the set of coarse grid points, and $C_i \subseteq N_i \cap C$ the set of coarse grid points in a neighborhood of i. Given the coarse grid values e_k^h, $k \in C_i$, we want to

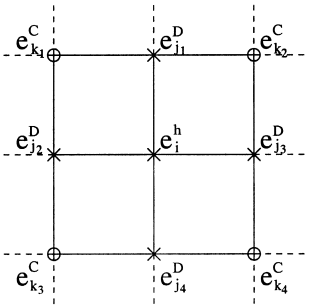

Fig. 5. A portion of a fine grid with coarse grid points denoted by circles, and noncoarse grid points by crosses. The coarse and noncoarse grid point connections with respect to e_i^h are denoted by the superscripts.

define a value e_i^h such that r_i^h is as small as possible. If C is selected such that $C_i = N_i$, then the choice of e_i^h

$$e_i^h = \sum_{k \in C_i} w_{ik}^h e_k^h, \quad w_{ik}^h = -\frac{\mathscr{A}_{ik}^h}{\mathscr{A}_{ii}} \tag{27}$$

leads to an ideal interpolation since $r_i^h = 0$. This is indeed equivalent to solving a local PDE with the ith node as interior noncoarse grid point and its neighbors as coarse grid points. However, such selection of C yields a dense coarse grid operator. Hence, in general, one has $C_i \subset N_i$ and $D_i \equiv N_i \setminus C_i \neq \emptyset$. For example, on a Cartesian grid with standard full coarsening (Fig. 5), the two set of variables in C_i and D_i are indicated by their superscripts.

Consider (24) again where r_i^h is to be made 0:

$$\mathscr{A}_{ii}^h e_i^h + \sum_{k \in C_i} \mathscr{A}_{ik}^h e_k^h + \sum_{j \in D_i} \mathscr{A}_{ij}^h e_j^h = 0. \tag{28}$$

The value e_i^h to be interpolated can be obtained by (28) provided e_k^h's and e_j^h's are known. Given only the coarse grid values e_k^h, the idea is to first interpolate the noncoarse grid values e_j^h, $j \in D_i$, by the e_k^h's $k \in C_i$. For $j \in D_i$, e_j^h is approximated by a weighted average:

$$e_j^h \approx \left(\sum_{k \in C_i} \mathscr{A}_{jk}^h e_k^h \right) \Big/ \left(\sum_{k \in C_i} \mathscr{A}_{jk}^h \right). \tag{29}$$

This local interpolation formula (29) is nothing but the one-dimensional local solve technique. Considering $e_{j_1}^D$ in Fig. 5, by formula (29), we have

$$e_{j_1}^D \approx \frac{\mathscr{A}_{j_1,k_1}^h e_{k_1}^C + \mathscr{A}_{j_1,k_2}^h e_{k_2}^C}{\mathscr{A}_{j_1,k_1}^h + \mathscr{A}_{j_1,k_2}^h}. \tag{30}$$

Comparing (30) with (21), and using the stencil terminology, we note that the one-dimensional interpolation (29) used by AMG is obtained by the second row of the stencil at the node $e_{j_1}^D$ whereas the one used by the black box multigrid in Section 2.2.1 is obtained by the average of the three rows of the stencil. Once e_j^h's are known, the interpolated value e_i^h is then given by (28), which is the same as the local PDE solve formula (22) used by Dendy's black box multigrid.

In general, the computation of (29) may still be too large. We want to interpolate e_i^h by only those e_j^h's which are significant. In view of the discussion after formula (26), the complexity can be reduced by the notion of strong connectedness. A point i is strongly connected to j if

$$-\mathscr{A}_{ij}^h \geqslant \theta \max_{l \neq i} \{-\mathscr{A}_{il}^h\} \tag{31}$$

with $0 < \theta \leqslant 1$ as an input parameter. One only considers strong connections in the construction. Specifically, denote by S_i the set of all strong connections of point i. Define $C_i = C \cap S_i$, and let $D_i^S = D_i \cap S_i$ and $D_i^W = D_i \setminus S_i$. For the weak connections ($j \in D_i^W$), e_j^h is simply replaced by e_i^h; i.e., lumping the weak entries to the diagonal. For the strong connections ($j \in D_i^S$), e_j^h is defined as in (29).

Other variants of the algebraic interpolation are discussed in [79]. See also the recent survey by Stüben [84].

2.5. Energy minimization approaches

The AMG approach is purely algebraic and potentially applies to more general problems than the other methods, but the underlying PDE and geometry information, if exist, may not be fully utilized. In this section, we discuss another approach based on energy minimization which bridges the gap between the two extremes: geometry dependent and purely algebraic. It exploits the properties of the underlying PDEs using primarily the algebraic information, thus allowing general computational domains. The essential idea is motivated by the classical multigrid convergence theory, in particular, the stability and approximation inequalities (13) and (14) described in Section 1.3. The key is to construct coarse grid basis which has minimal energy while preserving the zero energy modes. Vaněk et al. [88] identified altogether seven objectives the coarse grid basis should satisfy:

1. Compact support.
2. Coarse supports should follow strong couplings.
3. Reasonable geometry of supports.
4. Bounded intersections among the coarse supports.
5. Polynomial (null space) preserving property.
6. Small energy of coarse basis functions.
7. Uniform l^2 equivalence with L^2.

Based on these objectives, we try to construct coarse grid basis to achieve them, in particular, small energy, and null space preserving. For our model equation (1), the null space consists of constant functions.

2.5.1. Smoothed aggregation

In the smoothed aggregation method described in Section 2.3.3, one begins with the piecewise constant basis, which has high energy, and then smooths (or reduce the energy of) the basis by applying a relaxation method such as damped Jacobi to the interpolation operator. Thus, the energy of the basis is minimized locally. Moreover, it can be shown that the resulting basis preserves constants.

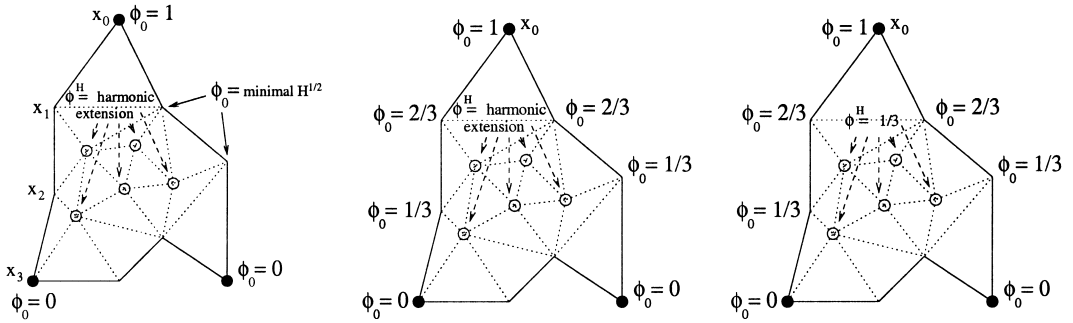

Fig. 6. Basis defined on macroelements. (Left) $H^{1/2}$ minimization on edges and harmonic extension inside. (Center) Graph distance weightings on edges and harmonic extension inside. (Right) Graph distance weightings on edges and inside.

We remark that the energy norm of the smoothed basis may not necessarily be minimal. One may further reduce the energy by applying more damped Jacobi steps which, however, increases the supports of the basis functions. Consequently, the coarse grid operator becomes denser for more smoothed interpolation operator.

2.5.2. Harmonic extension

A graph and agglomeration-based technique was proposed by Chan et al. [28,33]. The supports of the basis functions compose of macroelements formed by agglomeration. The basis functions on the edges are first constructed, and then they are extended by harmonic extension to the interiors. This procedure can be viewed as a local energy minimization. Moreover, if the basis preserves constant on the edges, it will also preserve constants in the interiors. There are several strategies to define the basis functions on the edges as well as in the interior with small energy. They are made more precise in the following.

$H^{1/2}$ norm minimization+harmonic extension: One defines the coarse grid basis functions on the edges as linear functions which are minimized in the $H^{1/2}$ norm-the interface analogue of the energy norm. More precisely, consider the coarse grid basis ϕ_i^H on a macroelement with coarse grid points denoted by the black dots; see Fig. 6. Suppose $\phi_i^H = \phi_0$, a linear function, on the edge formed by x_0, x_1, x_2 and x_3, i.e.,

$$\phi_0 = bx + c.$$

With two boundary conditions: $\phi_0(x_0)=1$, $\phi_0(x_3)=0$, and hence one degree of freedom, one requires that ϕ_0 minimizes the functional (discrete $H^{1/2}$ norm):

$$F(\phi_0) = \sum_{i=1}^{3} \sum_{j=i+1}^{3} \frac{h_i h_j}{h_{ij}^2} (\phi_0(x_i) - \phi_0(x_j))^2,$$

where h_i is the length of the edge (x_i, x_{i+1}) and $h_{ij} = |x_i - x_j|$. After incorporating the two boundary conditions, the one-dimensional minimization of $F(\phi_0)$ can be solved analytically. The same procedure is applied to the other edges, and the values at the interior points are obtained by harmonic extension.

Graph distance+harmonic extension: The $H^{1/2}$ norm minimization combined with the harmonic extension approach is robust but the entire procedure may be too complex. A simplified variant is to use a simpler boundary interpolation based on graph distance. Note that x_1 is distance 1 from x_0, x_2 distance 2 from x_0, etc. Define

$$\phi_i^H(x_j) = \frac{3-j}{3}$$

on the edge formed by x_0, x_1, x_2 and x_3, and piecewise linear on the edges. As in the previous approach, the values in the interior are given by the solution of a local PDE.

Pure graph distance: One may simplify the construction further by substituting the local PDE solve by a technique similar to graph distance. Suppose the macroelement has m number of coarse grid points on the boundary. Then the value of coarse grid basis function at each of the interior point is $1/m$. Thus, in our example, $\phi_i^H(x) = 1/3$, $x =$ interior points. Note that constants are preserved.

2.5.3. Energy-minimizing interpolation

The previous approaches construct basis which first satisfies the approximation property, followed by minimizing the energy locally. Another approach proposed by Wan et al. [92] is to prescribe the energy minimization and constant preserving explicitly into the formulation of the interpolation operator while fixing the size of the supports. As opposed to all the previous approaches, we determine the interpolation values on the edges and in the interior at the same time by a minimization formulation, and hence we do not actually identify edges nor interiors. Meanwhile, the constant preserving property is enforced by a constraint setup which globally couples all the individual basis functions.

The idea is based on another interpretation of the one-dimensional interpolation now described. Consider the two-point boundary value problem (15) again. The equivalent variational formulation is given by

$$\begin{aligned} &\min \quad \|\phi_i^H\|_A \quad (x_{2i}^h, x_{2i+2}^h) \\ &\text{s.t.} \quad \phi_i^H(x_{2i}^h) = 1, \quad \phi_i^H(x_{2i+2}^H) = 0. \end{aligned} \tag{32}$$

Thus, the local PDE formulation in one dimension is precisely minimizing the energy of the coarse grid basis functions. Moreover, if constant functions are in the kernel of the differential operator, the minimal energy basis will automatically preserve constants [91].

The extension to higher dimensions, however, is not obvious. First, the basis $\{\phi_i^H\}$, each of which has minimum energy, does 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 (15) need to be modified. A clue is provided in the two-level setting. Let Π_H^h be the usual nodal value interpolant. By the Cauchy–Schwarz and Poincaré inequalities, we obtain a rough estimate

$$\|\Pi_H^h v^H\|_A = \left\| \sum_i v^H(x_i^H)\phi_i^H \right\|_A \leqslant \frac{C}{H} \left(\sum_i \|\phi_i^H\|_A^2 \right)^{1/2} \|v^H\|_A, \tag{33}$$

where C is a constant independent of h. Comparing (33) with the stability inequality (13), we see that the constant C_0 in (13) depends on the total energy of $\{\phi_i^H\}$. Thus, the 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.

Write the coarse grid basis function ϕ_i^H as in (23). We determine the coefficients w_{ij}^h by solving a constrained minimization problem:

$$\min \frac{1}{2} \sum_{i=1}^{m} \|\phi_i^H\|_A^2 \quad \text{s.t.} \quad \sum_{i=1}^{m} \phi_i^H(x) = 1 \quad \text{in } \bar{\Omega}. \tag{34}$$

Lemma 2.2. *An equivalent formulation of* (15) *and* (32) *is the global minimization*

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

Proof. See [92]. □

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

Remark. (1) The values of the basis functions are defined implicitly by the solution of (34) and are not known explicitly in general. However, for the Laplacian, we recover exactly the bilinear interpolation on Cartesian grids [91], which is known to lead to optimal multigrid convergence for Poisson equation. (2) Like algebraic multigrid, the construction is purely algebraic. In other words, geometry and in particular the grid information are not needed. However, if the additional knowledge of the geometry is useful, for instance, semi-coarsening on Cartesian grids for anisotropic problems, we can still apply the same formulation. In fact, the coarse grid points can be provided geometrically by semi-coarsening or interface preserving techniques (cf. Section 4.3), or algebraically by AMG coarsening. Moreover, the formulation of the interpolation remains valid even if the coarse grid points do not form an independent set. (3) Mandel et al. [67] generalized this approach to solve systems of elliptic PDEs arising from linear elasticity problems.

Solution of the minimization problem: A detailed solution procedure is described in [92], and we only discuss the main steps here. It can be shown that the minimization problem (34) can be written as a constrained quadratic minimization. Thus, Newton's method only takes one iteration to convergence, which, however, need to invert the Jacobian matrix. Since the solution of the minimization problem is used as an interpolation operator for multigrid only, we do not need an exact solution. An efficient approximation can be obtained by a preconditioned conjugate gradient method. Empirical evidence shows that in most cases, only a few iterations suffice, except for oscillatory PDE coefficients. In [67], Mandel et al. showed that the interpolation obtained from the first step of the steepest descent procedure in solving (34) yields the same result as the smoothed aggregation with a single smoothing step.

2.5.4. AMGe

The use of energy minimization in the formulation of the interpolation operator has shown to be powerful in the energy-minimizing interpolation approach. AMGe, algebraic multigrid based on element stiffness matrices, proposed by Brezina et al. [26], uses the local measures of algebraic smoothness derived from multigrid theory to construct the interpolation operator. The key observation is that the interpolation must be able to approximate an eigenvector with an error bound proportional

to the size of the associated eigenvalue. More precisely, the interpolation must be defined such that either of the following measures are bounded by a constant independent of h:

$$M_1(Q,e^h)=\frac{((I-Q)e^h,(I-Q)e^h)}{(Ae^h,e^h)},$$

$$M_2(Q,e^h)=\frac{(A(I-Q)e^h,(I-Q)e^h)}{(Ae^h,Ae^h)},$$

where Q is a projection onto the range of the interpolation matrix $\mathscr{I}_H^h$. Note that Q is related to $\mathscr{I}_H^h$ by

$$Q=[0 \quad \mathscr{I}_H^h],$$

if the unknowns corresponding to the noncoarse grid points are ordered before the coarse grid points. The boundedness of M_1 or M_2 requires Q to accurately interpolate the eigenvectors corresponding to small eigenvalues, but not necessarily as accurate as for the eigenvectors corresponding to large eigenvalues. In addition, the quantities M_1 or M_2, as opposed to the matrix entries used in standard AMG, give a potentially better measure of strong connectedness, especially for non-M-matrices.

In the previous approaches, the interpolation matrix is constructed by defining the coarse grid basis whose coefficients w_{ij} (cf. (23)) are the entries of the jth column of $\mathscr{I}_H^h$. Hence, the matrix $\mathscr{I}_H^h$ is constructed column by column whereas in AMGe, $\mathscr{I}_H^h$ is constructed row by row. Let q_i be the ith row of Q. Then, q_i is defined as the solution of the following min–max problem:

$$\min_{q_i} \max_{e^h \notin \mathrm{Null}(\mathscr{A}_i^h)} M_{i,p}(q_i,e^h) \tag{35}$$

for $p=1$ or 2. Here, $M_{i,p}(q_i,e^h)$ is a local measure derived from the corresponding global measure $M_p(Q,e^h)$ which is practically inaccessible, and $\mathscr{A}_i^h$ is the sum of local element stiffness matrices connected with i. It can be shown [26] that the solution of (35) is to fit the eigenvectors of $\mathscr{A}_i^h$ subject to the constraint that constants are preserved. Hence, it can be considered as another local energy minimization strategy.

Finally, we note that AMGe requires the knowledge of the element stiffness matrices which sometimes may not be conveniently available. Thus, this approach is less algebraic than the other energy-minimizing approaches.

3. Smoothing

Interpolation alone is not enough for fast convergence as the success of multigrid requires different components complement each other. The interpolation is effective only when the smoothers produce smooth errors either in the geometric sense, or in the algebraic sense (cf. Section 2.4). A classical example in the literature where smoothing plays an important role in improving the multigrid convergence is when solving PDEs with anisotropic coefficients, for instance,

$$\begin{aligned} &-\varepsilon u_{xx}-u_{yy}=f \quad \text{in } \Omega,\\ &u=0 \quad \text{on } \partial\Omega. \end{aligned} \tag{36}$$

Assuming standard coarsening, it can be shown by Fourier analysis [19,94] that point relaxation methods as smoothers are not effective for small ε since the errors are only smoothed in the y-direction,

and the errors in the x-direction can be highly oscillatory, leading to slow convergence of multigrid. In the next sections, we discuss the use of block\line relaxation methods, incomplete LU (ILU) factorization preconditioners, and sparse approximate inverses as smoothers for anisotropic as well as other types of elliptic PDE problems.

3.1. Block\line relaxation

A well-suited smoother can be derived from considering the limiting case when $\varepsilon \to 0$. The resulting PDE is a decoupled system of elliptic equations along the vertical lines, suggesting the use of block Gauss–Seidel as smoother where the blocks are associated with the lines in the y-direction, or equivalently, the direction of the anisotropy. It can be proved [19,55] that the two-grid method with this smoother converges independently of ε. In the case of variable coefficients where the anisotropy direction may change from y to x at different locations, one may alternate the direction of the line relaxation. Another option is to use the alternating direction implicit method [68].

The disadvantage of this approach is that the smoothing is most effective when the anisotropy is either in the x- or y-direction. Another problem is that it is essentially a Cartesian grid technique. Although similar idea can be adapted in unstructured grid computations [72], the determination of the lines or planes of anisotropy in three dimensions is complicated. Besides, inverting a line or a plane requires more work than point relaxations.

3.2. ILU

One needs a *direction free* and robust iterative method as smoother for solving anisotropic problems, and we shall discuss two possibilities in this and the next section. An incomplete LU factorization based on the stencil pattern was studied by Wesseling [94] and Wittum [95]. Given a five-point stencil matrix, for instance, one may use the stencil pattern for the incomplete L, U factors. Specifically, an incomplete LU factorization can be written as

$$\mathscr{A}^h = \mathscr{L}^h \mathscr{U}^h + \mathscr{E}^h,$$

where the incomplete lower triangular factor $\mathscr{L}^h$ has a three-point stencil structure corresponding to the lower triangular part of a five-point stencil matrix, and similarly for the upper triangular factor $\mathscr{U}^h$. Similar ideas can be applied to other stencil patterns. The resulting ILU smoother is effective for a wide range of directions of anisotropy. More precisely, suppose the model equation (36) is rotated by an angle θ. Hemker [58] showed that ILU is an effective smoother for $\pi/4 \leqslant \theta \leqslant \pi$. However, it may lead to divergence for $0 < \theta < \pi/4$. In the latter, the line relaxation smoother is still effective since the angle is small. One may combine both ideas and derive an incomplete line LU (ILLU) smoother [82] which uses block triangular factors for $\mathscr{L}^h$ and $\mathscr{U}^h$.

3.3. Sparse approximate inverse

The drawback of ILU smoothers is the lack of parallelism, since the (incomplete) LU factorization process is sequential in nature. Another class of direction free smoothers, which are inherently parallel, are sparse approximate inverses (SAI).

Most sparse approximate inverse (SAI) approaches seek a sparse matrix $\mathcal{M}^h$ so that the error of the residual is minimized in some measure. One of the earliest approaches was the Frobenius norm approach proposed by Benson [10] and Benson and Frederickson [13]:

$$\min_{\mathcal{M}^h} \|\mathcal{A}^h\mathcal{M}^h - I\|_F^2,$$

subject to some constraints on the number and position of the nonzero entries of $\mathcal{M}^h$. The minimization problem is equivalent to n independent least-squares problems:

$$\min_{m_j} \|\mathcal{A}^h m_j - e_j\|_2, \quad j = 1, \ldots, n, \tag{37}$$

where m_j and e_j are the jth column of $\mathcal{M}^h$ and I, respectively, and they can be solved in parallel. For efficient construction, the sparsity pattern may be selected as banded diagonal [63], for example, or determined adaptively by heuristic searching algorithms [35,51] which, however, may decrease parallelism.

Several SAI smoothers have been studied. Benson [11], and Benson and Banerjee [12] used a sparsity pattern based on graph neighbors. Recently, Huckle and Grote independently experimented a sparse approximate smoother based on SPAI [51] which adaptively search the nonzero pattern.

In the following, we describe the approach proposed by Tang and Wan [86]. Since the major cost of multigrid algorithms is smoothing, it is important to derive simple and yet effective sparsity patterns. In addition, the least-squares problems (37) must be solved efficiently. It turns out that a pre-defined pattern based on neighbors of the matrix graph is sufficient for effective smoothing [86]. Given a node j, define $L_k(j)$ as its k-level neighbor set in graph distance. For instance, $L_0(j)$ contains simply the set of stencil points in case of PDE problems. Furthermore, one modify the Frobenius norm approach (37) and introduce the (k,l)-level least-squares approximation:

$$\min_{m_j} \|\mathcal{A}^{k,l} m_j - e_j\|_2,$$

where $\mathcal{A}^{k,l} \equiv \mathcal{A}^h(L_k(j), L_l(j))$ is the (k,l)-level local submatrix of $\mathcal{A}^h$. The sparsity pattern is determined by the l-level neighbors, and the size of the least squares matrix is controlled by the selections of k and l. Hence, the two main issues of SAI smoothers are handled. Moreover, it can be proved that high frequency errors will be damped away efficiently for $k = 1$ and $l = 0$ [86].

More importantly, SAI smoothers have the flexibility of using larger values of k and l to improve the smoothing quality for difficult PDE problems. The potential higher computational cost can be reduced by dropping strategies. For anisotropic coefficient PDEs, the matrix $\mathcal{A}^h$ and its inverse typically have many small entries. Thus, one may drop the neighbors with weak connections in $\mathcal{A}^{k,l}$ before computing the approximate inverse. This is essentially the same idea as line relaxation which only applies to structured grids. One may further reduce the cost by discarding small entries in $\mathcal{M}^h$. It has been shown empirically [86] that the resulting complexity is only twice as expensive as point Gauss–Seidel for the anisotropic problem (36). In addition, since the determination of the lines or planes of anisotropy is done algebraically and automatically, SAI smoothing is applicable to both structured and unstructured grid computations in higher dimensions.

4. Coarsening

The design of interpolation and smoothing operators has been the main focus in improving multigrid performance. Coarsening, the selection of coarse grid points, can be as important as interpolation and smoothing, for instance, AMG coarsening [79] and semi-coarsening [40,41,81] are both critical components in their respective algorithms. The former selects coarse grid points algebraically according to strong connections and is robust for discontinuous and anisotropic coefficient PDEs. The latter selects coarse grid points geometrically according to the direction of strong coupling for anisotropic coefficient PDEs. Another approach, interface preserving coarsening [91], selects coarse grid points geometrically according to the shape of interfaces.

4.1. Semi-coarsening

For anisotropic coefficient problems, special smoothing techniques are discussed in Section 3 to improve multigrid efficiency. Coarsening has also shown to be another approach to recover fast multigrid convergence. The failure of standard multigrid is that the errors in the direction of weak anisotropy are not smoothed. Thus, they cannot be solved on the coarse grid. In the case of structured grids, one can apply standard coarsening to the direction of strong coupling only, i.e., the y-direction for the model problem (36), and select all the grid points as coarse grid points in the other directions, resulting in alternating $y =$ constant lines of coarse grid points. The drawback, however, is that the overall cost of multigrid will increase.

A related coarsening technique is the use of multiple semi-coarsened grids proposed by Mulder [73]. For nonlinear PDE problems where the direction of anisotropy changes from time to time, Mulder performed the coarse grid correction on two semi-coarsened grids in both x- and y-direction on each level of grid. The complexity of the resulting algorithm turns out still to be proportional to the number of unknowns. Frequency decomposition multigrid, proposed by Hackbusch [56], is another method using several grid corrections. Three additional fully coarsened grids are formed by shifting the standard coarse grid by one grid point in the x- and/or y-direction. Moreover, special prolongation and restriction operators are used to include also the high frequencies on the additional coarse grids so that the union of the ranges of the prolongation operators is the same as the fine grid function space. The filtering decomposition by Wittum [96,97] is another coarse grid correction method. Instead of including all the high frequencies, the coarse grid operator is required to have the same effect as the fine grid operator on a selected set of vectors, for instance, discrete sine functions with different frequencies. This principle is similar to the probing method proposed by Chan and Mathew [31] in domain decomposition.

4.2. AMG coarsening

In algebraic multigrid [79], the selection of coarse grid points ties strongly with the algebraic multigrid interpolation. Divide the fine grid points into the set of coarse (C) and noncoarse (F) grid points. In the ideal case where C is chosen such that for each noncoarse grid point $i \in F$, its neighbors are all coarse gird points, i.e., $N_i = C_i$ (Section 2.4), the algebraic interpolation defined in (27) is exact; it is just Gaussian elimination as described in one dimension. Otherwise, the interpolation needs to approximate the values at the noncoarse grid point connections as given in

(29). On the other hand, the notion of strong connectedness is introduced to maintain sparsity by ignoring weak connections. Moreover, the approximation in (29) is more accurate if many strong noncoarse grid connection neighbors of j are actually in C_i. Summing up, there are two criteria for algebraic coarsening:

1. For each noncoarse grid point i, each strong connection $j \in S_i$ should either be a coarse grid point ($j \in C_i$), or should be strongly connected to at least one point in C_i.
2. The set of coarse grid point should form a maximal independent set with respect to strong connections.

The two criteria, in general, are conflicting with each other. Usually, the second criterion is used to select a potential small set of coarse grid points. Then noncoarse grid points may be added to satisfy the first criterion.

4.3. Other approach

Algebraic coarsening selects coarse grid points based on matrix entries. A recent approach, the interface preserving coarsening, can be considered as its geometric counterpart which is specifically designed for discontinuous coefficient PDEs. For this class of problems, multigrid is typically improved by a sophisticated interpolation such as those described in Section 2 which captures the discontinuous behavior of the derivatives of the solution along the interfaces. This is particularly important since the interface may not necessarily align with the coarse grids as usually demanded by theory [18,45,98]. However, linear interpolation can be just fine if the interface aligns with *all* coarse grids. The main idea of interface preserving coarsening [91] is thus to select coarse grid points which resolve the shape of the interface on all coarse grids.

5. Conclusion

Significant advances have been made in robust multigrid methods for elliptic linear systems in recent years. The variety of techniques developed have been steadily maturing, but will still have not quite reached the holy grail for multigrid methods that is algebraic and easily parallelizable, with complexity proportional to the number of unknowns, and with rate of convergence independent of the mesh size, the nature of the PDE coefficients, and the computational grids. The gray box philosophy may ultimately lead to a more flexible approach to developing multigrid algorithms which can make optimal use of any available information. Our discussion has been confined to the algorithmic developments. Parallelization and theoretical issues are nonetheless essential for the practical and intelligent use of multigrid in large scale numerical simulations.

There is still plenty of room for improvements in every aspects of robust multigrid to come in the years ahead. For instance, it is likely that the energy minimization principle will continue to be useful in constructing robust interpolation operators as new techniques are developed. On the other hand, the interplay between coarse grid basis and interpolation provides another perspective for constructing robust interpolation, for instance, through the use of special finite element basis recently developed for discontinuous and oscillatory coefficient PDEs.

References

[1] R.E. Alcouffe, A. Brandt, J.E. Dendy Jr., J.W. Painter, The multi-grid method for the diffusion equation with strongly discontinuous coefficients, SIAM J. Sci. Statist. Comput. 2 (4) (1981) 430–454.

[2] G.P. Astrachancer, An iterative method of solving elliptic net problems, USSR Comput. Math. Math. Phys. 11 (1971) 171–182.

[3] O. Axelsson, P. Vassilevski, Algebraic multilevel preconditioning methods, Part I, Numer. Math. 56 (1989) 157–177.

[4] O. Axelsson, P. Vassilevski, Algebraic multilevel preconditioning methods, Part II, SIAM J. Numer. Anal. 27 (1990) 1569–1590.

[5] N.S. Bachvalov, On the convergence of a relaxation method with natural constraints on the elliptic operator, USSR Comput. Math. Phys. 6 (1966) 101–135.

[6] R.E. Bank, T. Dupont, An optimal order process for solving elliptic finite element equations, Math. Comp. 36 (1981) 35–51.

[7] R.E. Bank, T.F. Dupont, H. Yserentant, The hierarchical basis multigrid method, Numer. Math. 52 (1988) 427–458.

[8] R.E. Bank, R.K. Smith, The incomplete factorization multigraph algorithm, SIAM J. Sci. Comput. 20 (1999) 1349–1364.

[9] R.E. Bank, J. Xu, An algorithm for coarsening unstructured meshes, Numer. Math. 73 (1996) 1–36.

[10] M.W. Benson, Iterative solution of large scale linear systems, Master's Thesis, Lakehead University, Thunder Bay, Ontario, 1973.

[11] M.W. Benson, Frequency domain behavior of a set of parallel multigrid smoothing operators, Internat. J. Comput. Math. 36 (1990) 77–88.

[12] M.W. Benson, R.N. Banerjee, An approximate inverse based multigrid approach to the biharmonic problem, Internat. J. Comput. Math. 40 (1991) 201–210.

[13] M.W. Benson, P.O. Frederickson, Iterative solution of large sparse linear systems arising in certain multidimensional approximation problems, Utilitas Math. 22 (1982) 127–140.

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

[15] H. Brakhage, Über die numerishe behandlung von integralgleichungen nach der quadraturformelmethode, Numer. Math. 2 (1960) 183–196.

[16] J. Bramble, Multigrid Methods, Longman Scientific & Technical, Essex, UK, 1993.

[17] J. Bramble, J. Pasciak, J. Xu, Parallel multilevel preconditioners, Math. Comp. 55 (1990) 1–22.

[18] J. Bramble, J. Xu, Some estimates for a weighted l^2 projection, Math. Comp. 56 (1991) 463–476.

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

[20] A. Brandt, Guide to multigrid developments, in: W. Hackbusch, U. Trottenberg (Eds.), Multigrid Methods, Lectures Notes in Mathematics, Vol. 960, Springer, Berlin, 1982, pp. 220–312.

[21] A. Brandt, Multigrid techniques: 1984 guide with applications to fluid dynamics, GMD-Studien No. 85, Gesellschaft für Mathematik und Datenverarbeitung, St. Augustin, 1984.

[22] A. Brandt, Algebraic multigrid theory: the symmetric case, Appl. Math. Comput. 19 (1986) 23–56.

[23] A. Brandt, The gauss center research in multiscale scientific computation, Electron. Trans. Numer. Anal. 6 (1997) 1–34.

[24] A. Brandt, General highly accurate algebraic coarsening schemes, Proceedings of the Nineth Copper Mountain Conference on Multigrid Methods, Copper Mountain, CO, 1999.

[25] A. Brandt, S. McCormick, J.W. Ruge, Algebraic multigrid (AMG) for automatic multigrid solution with application to geodetic computations, Technical Report, Inst. for Computational Studies, Fort Collins, CO, 1982.

[26] M. Brezina, A.J. Cleary, R.D. Falgout, V.E. Henson, J.E. Jones, T.A. Manteuffel, S.F. McCormick, J.W. Ruge, Algebraic multigrid based on element interpolation (AMGe), Technical Report UCRL-JC-131752, Lawrence Livermore National Laboratory, 1998.

[27] W. Briggs, A Multigrid Tutorial, SIAM, Philadelphia, PA, 1987.

[28] T.F. Chan, S. Go, L. Zikatanov, Lecture Notes on Multilevel Methods for Elliptic Problems on Unstructured Grids, Lecture Course 28th Computational Fluid Dynamics, March 3–7, 1997, von Karman Institute for Fluid Dynamics, Belgium, 1997.

[29] T.F. Chan, S. Go, J. Zou, Multilevel domain decomposition and multigrid methods for unstructured meshes: algorithms and theory, Technical Report CAM 95-24, Department of Mathematics, UCLA, May 1995.
[30] T.F. Chan, S. Go, J. Zou, Boundary treatments for multilevel methods on unstructured meshes, Technical Report CAM 96-30, Department of Mathematics, UCLA, September 1996.
[31] T.F. Chan, T.P. Mathew, The interface probing technique in domain decomposition, SIAM J. Matrix Anal. Appl. 13 (1992) 212–238.
[32] T.F. Chan, B. Smith, J. Zou, Multigrid and domain decomposition methods for unstructured meshes, in: I.T. Dimov, Bl. Sendov, P. Vassilevski (Eds.), Proceedings of the Third International Conference on Advances in Numerical Methods and Applications, Sofia, Bulgaria, World Scientific, Singapore, 1994, pp. 53–62.
[33] T.F. Chan, J. Xu, L. Zikatanov, An agglomeration multigrid method for unstructured grids, Technical Report CAM 98-8, Department of Mathematics, UCLA, February 1998.
[34] Q. Chang, Y.S. Wong, Z. Li, New interpolation formulas of using geometric assumptions in the algebraic multigrid method, Appl. Math. Comput. 50 (1992) 223–254.
[35] E. Chow, Y. Saad, Approximate inverse preconditioners via sparse–sparse iterations, SIAM J. Sci. Comput. 19 (1998) 995–1023.
[36] A.J. Cleary, R.D. Falgout, V.E. Henson, J.E. Jones, Coarse-grid selection for parallel algebraic multigrid, Technical Report UCRL-JC-130893, Lawrence Livermore National Laboratory, 1998.
[37] A.J. Cleary, R.D. Falgout, V.E. Henson, J.E. Jones, T.A. Manteuffel, S.F. McCormick, G.N. Miranda, J.W. Ruge, Robustness and scalability of algebraic multigrid, Technical Report UCRL-JC-130718, Lawrence Livermore National Laboratory, 1998.
[38] J.E. Dendy Jr., Black box multigrid, J. Comput. Phys. 48 (1982) 366–386.
[39] J.E. Dendy Jr., Black box multigrid for nonsymmetric problems, Appl. Math. Comput. 13 (1983) 261–283.
[40] J.E. Dendy Jr., M.P. Ida, J.M. Rutledge, A semi-coarsening multigrid algorithm for SIMD machines, SIAM J. Sci. Statist. Comput. 13 (1992) 1460–1469.
[41] J.E. Dendy Jr., S.F. McCormick, J.W. Ruge, T.F. Russell, S. Schaffer, Multigrid methods for three-dimensional petroleum reservoir simulation, 10th Symposium on Reservoir Simulation, Houston, Society of Petroleum Engineers (SPE), February 6–8 1989, pp. 19–25.
[42] J.E. Dendy Jr., C.C. Tazartes, Grandchild of the frequency decomposition multigrid method, SIAM J. Sci. Comput. 16 (2) (1994) 307–319.
[43] C.C. Douglas, Caching in with multigrid algorithms: problems in two dimensions, Parallel Algorithms Appl. 9 (1996) 195–204.
[44] C.C. Douglas, U. Rüde, J. Hu, M. Bittencourt, A guide to designing cache aware multigrid algorithms, submitted for publication.
[45] M. Dryja, M. Sarkis, O. Widlund, Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions, Numer. Math. 72 (1996) 313–348.
[46] B. Engquist, E. Luo, Convergence of a multigrid method for elliptic equations with highly oscillatory coefficients, SIAM J. Numer. Anal. 34 (1997) 2254–2273.
[47] R.P. Fedorenko, A relaxation method for solving elliptic difference equations, USSR Comput. Math. Phys. 1 (1961) 1092–1096.
[48] R.P. Fedorenko, The speed of convergence of one iterative process, USSR Comput. Math. Phys. 4 (1964) 227–235.
[49] T. Grauschopf, M. Griebel, H. Regler, Additive multilevel-preconditioners based on bilinear interpolation, matrix-dependent geometric coarsening and algebraic multigrid coarsening for second order elliptic pdes, Appl. Numer. Math. 23 (1997) 63–96.
[50] M. Griebel, Multilevel algorithms considered as iterative methods on semidefinite systems, SIAM J. Sci. Comput. 15 (1994) 547–565.
[51] M. Grote, T. Huckle, Parallel preconditioning with sparse approximate inverses, SIAM J. Sci. Comput. 18 (1997) 838–853.
[52] H. Guillard, Node-nested multigrid method with Delaunay coarsening, Technical Report 1898, INRIA, Sophia Antipolis, France, 1993.
[53] W. Hackbusch, A fast iterative method solving Poisson's equation in a general region, in: R. Bulirsch, R.D. Grigorieff, J. Schröder (Eds.), Proceedings, Oberwolfach, July 1976, Lecture Notes in Mathematics, Vol. 631, Springer, Berlin, 1978, pp. 51–62.

[54] W. Hackbusch, Survey of convergence proofs for multigrid iterations, in: J. Frehse, D. Pallaschke, U. Trottenberg (Eds.), Special Topics of Applied Mathematics, North-Holland, Amsterdam, 1980, pp. 151–164.
[55] W. Hackbusch, Multi-grid Methods and Applications, Springer, Berlin, 1985.
[56] W. Hackbusch, The frequency decomposition multigrid method, Part I: application to anisotropic equations, Numer. Math. 56 (1989) 229–245.
[57] W. Hackbusch, S. Sauter, Adaptive composite finite elements for the solutions of pdes containing non-uniformly distributed micro-scales, Mat. Model. 8 (1996) 31–43.
[58] P.W. Hemker, Multigrid methods for problems with a small parameter, in: D.F. Griffiths (Ed.), Proceedings, Dundee, June–July 1983, Lecture Notes in Mathematics, Vol. 1066, Springer, Berlin, 1983, pp. 106–121.
[59] W.Z. Huang, Convergence of algebraic multigrid methods for symmetric positive definite matrices with weak diagonal dominance, Appl. Math. Comput. 46 (1991) 145–164.
[60] D.C. Jespersen, Recent developments in multigrid methods for the steady Euler equations, Lecture Notes for Lecture Series on Computational Fluid Dynamics, von Karman Institute for Fluid Dynamics, Rhode-St.-Genese, Belgium, 1984.
[61] J. Jones, S. McCormick, Parallel multigrid methods, in: D. Keyes, A. Sameh, V. Venkatakrishnan (Eds.), Parallel Numerical Algorithms, NASA/LaRC Interdisciplinary Series in Science and Engineering, Kluwer, Dordrecht, 1997.
[62] R. Kettler, 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, 1981.
[63] L.Yu. Kolotilina, A.Yu. Yeremin, Factorized sparse approximate inverse preconditionings I, theory, SIAM J. Matrix Anal. Appl. 14 (1993) 45–58.
[64] A. Krechel, K. Stüben, Operator dependent interpolation in algebraic multigrid, in: Lecture Notes in Computational Science and Engineering, Vol. 3. Proceedings of the Fifth European Multigrid Conference, Stuttgart, October 1–4, 1996, Springer, Berlin, 1998.
[65] M.H. Lallemand, H. Steve, A. Dervieux, Unstructured multigridding by volume agglomeration: current status, Comput. Fluids 21 (1992) 397–433.
[66] P.L. Lions, On the Schwarz alternating method, I, in: R. Glowinski, G.H. Golub, G.A. Meurant, J. Periaux (Eds.), First International Conference on Domain Decomposition Methods for Partial Differential Equations, SIAM, Philadelphia, PA, 1988.
[67] J. Mandel, M. Brezina, P. Vaně k, Energy optimization of algebraic multigrid bases, Technical Report UCD/CCM 125, Center for Computational Mathematics, University of Colorado, Denver, 1998.
[68] G.I. Marchuk, Splitting and alternating direction methods, in: P.G. Ciarlet, J.L. Lions (Eds.), Handbook of Numerical Analysis, Vol. 1, North-Holland, Amsterdam, 1990.
[69] A.M. Matsokin, S.V. Nepomnyaschikh, A Schwarz alternating method in a subspace, Soviet Math. 29 (10) (1985) 78–84.
[70] O.A. McBryan, P.O. Frederickson, J. Linden, A. Schuller, K. Solchenbach, K. Stüben, C.A. Thole, U. Trottenberg, Multigrid methods on parallel computers a survey of recent developments, Impact Comput. Sci. Eng. 3 (1991) 1–75.
[71] S. McCormick, Multigrid methods for variational problems: general theory for the V-cycle, SIAM J. Numer. Anal. 22 (1985) 634–643.
[72] E. Morano, D.J. Mavriplis, V. Venkatakrishnan, Coarsening strategies for unstructured multigrid techniques with application to anisotropic problems, SIAM J. Sci. Comput. 20 (1999) 393–415.
[73] W. Mulder, A high-resolution Euler solver, Technical Report CAM 89-08, Department of Mathematics, UCLA, April 1989.
[74] R.A. Nicolaides, On multiple grid and related techniques for solving discrete elliptic systems, J. Comput. Phys. 19 (1975) 418–431.
[75] C.W. Oosterlee, The covergence of parallel multiblock multigrid, Appl. Numer. Math. 19 (1995) 115–128.
[76] A. Reusken, Multigrid with matrix-dependent transfer operators for a singular perturbation problem, Computing 50 (1993) 199–211.
[77] A. Reusken, Multigrid with matrix-dependent transfer operators for convection–diffusion problems, in: Multigrid Methods, Vol. IV, Lectures in Mathematics, Springer, Berlin, 1993.
[78] A. Reusken, A multigrid method based on incomplete Gaussian elimination, J. Numer. Linear Algebra Appl. 3 (1996) 369–390.

[79] J.W. Ruge, K. Stüben, Algebraic multigrid, in: S. McCormick (Ed.), Multigrid Methods, SIAM, Philadelphia, PA, 1987, pp. 73–130.
[80] B. Smith, P. Bjørstad, W. Gropp, Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations, Cambridge University Press, Cambridge, 1996.
[81] R.A. Smith, A. Weiser, Semicoarsening multigrid on a hypercube, SIAM J. Sci. Statist. Comput. 13 (1992) 1314–1329.
[82] P. Sonneveld, P. Wesseling, P.M. de Zeeuw, Multigrid and conjugate gradient methods as convergence acceleration techniques, in: H. Holstein, D.J. Paddon (Eds.), Short Course on Multigrid Methods, Proceedings, Bristol, September 1983, Oxford Press, Oxford, 1985.
[83] K. Stüben, Algebraic multigrid: experience and comparisons, Appl. Math. Comput. 13 (1983) 419–451.
[84] K. Stüben, Algebraic multigrid (AMG): an introduction with applications, Technical Report 53, GMD, March 1999, in: U. Trottenbery, C.W. Oosterlee, A. Schüller (Eds.), Multigrid (appendix), Academic Press, New York, 1999, to appear.
[85] S. Ta'asan, Multigrid methods for highly oscillatory problems, Ph.D. Thesis, Weizmann Institute of Science, Rehovat, Israel, 1984.
[86] W.P. Tang, W.L. Wan, Sparse approximate inverse smoother for multi-grid, SIAM J. Matrix Anal. Appl. (1999), to appear.
[87] U. Trottenberg, C.W. Oosterlee, A. Schüller, Basic, Parallel and Adaptive Multigrid Methods, Academic Press, New York, 1998, to appear.
[88] P. Vaně k, J. Mandel, M. Brezina, Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems, Computing 56 (1996) 179–196.
[89] V. Venkatakrishnan, D.J. Mavriplis, Agglomeration multigrid for the three-dimensional Euler equations, Technical Report 94-5, ICASE, NASA Langley, USA, 1994.
[90] C. Wagner, Introduction to algebraic multigrid, Course Notes of an Algebraic Multigrid Course, University of Heidelberg, 1999.
[91] W.L. Wan, Scalable and multilevel iterative methods, Ph.D. Thesis, Department of Mathematics, UCLA, Los Angeles, June 1998.
[92] W.L. Wan, T.F. Chan, B. Smith, An energy-minimizing interpolation for robust multigrid, SIAM J. Sci. Comput. (1998), to appear.
[93] P. Wesseling, A survey of fourier smoothing analysis results, in: W. Hackbusch, U. Trottenberg (Eds.), Multigrid Methods III, International Series of Numerical Mathematics, Vol. 98, Birkhäuser, Basel, 1991, pp. 105–127.
[94] P. Wesseling, An Introduction to Multigrid Methods, Wiley, Chichester, 1992.
[95] G. Wittum, On the robustness of ilu smoothing, SIAM J. Sci. Statist. Comput. 10 (1989) 699–717.
[96] G. Wittum, An ILU-based smoothing correction scheme, in: W. Hackbusch (Ed.), Parallel Algorithms for PDEs, Proceedings, Kiel, January, 1990, Notes on Numerical Fluid Mechanics, Vol. 31, Vieweg, Braunschweig, 1991.
[97] G. Wittum, Filternde zerlegungen: schnelle löser für große gleichungssysteme (Filtering decomposition: fast solvers for large systems of equations), Teubner Skripten zur Numerik, Teubner, Stuttgart, 1992.
[98] J. Xu, Theory of multilevel methods, Ph.D. Thesis, Cornell University, Ithaca, New York, 1989.
[99] J. Xu, Iterative methods by space decomposition and subspace correction, SIAM Rev. 34 (1992) 581–613.
[100] J. Xu, The auxiliary space method and optimal multigrid preconditioning techniques for unstructured grid, Computing 56 (1996) 215–235.
[101] H. Yserentant, On the multilevel splitting of finite element spaces, Numer. Math. 49 (1986) 379–412.
[102] P.M. De Zeeuw, Matrix-dependent prolongations and restrictions in a blackbox multigrid solver, J. Comput. Appl. Math. 33 (1990) 1–27.