Numerische Mathematik manuscript No.
(will be inserted by the editor)

Local Multilevel Preconditioners for Elliptic Equations with Jump Coefficients on Bisection Grids

Long Chen[1], **Michael Holst**[2], **Jinchao Xu**[3], **Yunrong Zhu**[2⋆]

[1] Department of Mathematics,
University of California at Irvine, CA 92697, USA
e-mail: chenlong@math.uci.edu
[2] Department of Mathematics,
University of California at San Diego, CA 92093, USA
e-mail: {mholst, zhu}@math.ucsd.edu
[3] Department of Mathematics,
Pennsylvania State University, University Park, PA 16802, USA
e-mail: xu@math.psu.edu

April 21, 2010

Summary The goal of this paper is to design optimal multilevel solvers for the finite element approximation of second order linear elliptic problems with piecewise constant coefficients on bisection grids. Local multigrid and BPX preconditioners are constructed based on local smoothing only at the newest vertices and their immediate neighbors. The analysis of eigenvalue distributions for these local multilevel preconditioned systems shows that there are only a fixed number of eigenvalues which are deteriorated by the large jump. The remaining eigenvalues are bounded uniformly with respect to the coefficients and the meshsize. Therefore, the resulting preconditioned conjugate gradient algorithm will converge with an asymptotic rate independent of the coefficients and logarithmically with respect to the meshsize. As a result, the overall computational complexity is nearly optimal.

Key words Local Multilevel Preconditioners, Multigrid, BPX, Discontinuous Coefficients, Adaptive Finite Element Methods, PCG, Effective Condition Number

Mathematics Subject Classification (2010): 65F08, 65F10, 65N30, 65N50, 65N55

⋆ Correspondence to: Yunrong Zhu (e-mail: zhu@math.ucsd.edu)

1 Introduction

We shall construct robust multilevel preconditioners for the finite element discretization of second order linear elliptic equations with strongly discontinuous coefficients. We extend corresponding results on uniform grids [57] to locally refined grids obtained by bisection methods.

Consider the following model problem :

$$-\nabla\cdot(a\nabla u)=f \text{ in } \Omega,\quad u=g_D \text{ on } \Gamma_D,\quad a\frac{\partial u}{\partial n}=g_N \text{ on } \Gamma_N \qquad (1.1)$$

where $\Omega\in\mathbb{R}^d$ is a polygon (for $d=2$) or polyhedron (for $d=3$) with Dirichlet boundary Γ_D and Neumann boundary Γ_N such that $\Gamma_D\cup\Gamma_N=\partial\Omega$. The diffusion coefficient $a=a(x)$ is piecewise constant. More precisely, the domain Ω is partitioned into M open disjoint polygonal or polyhedral regions Ω_i $(i=1,\cdots,M)$ and

$$a|_{\Omega_i}=a_i,\ \ i=1,\ldots,M$$

where each a_i is a positive constant. The regions Ω_i $(i=1,\cdots M)$ may possibly have complicated geometry but we assume that they are completely resolved by an initial triangulation $\mathcal{T}_0$. Our analysis can be carried through to more general cases when $a(x)$ varies moderately in each subdomain and to other types of boundary conditions in a straightforward way.

The problem (1.1) belongs to the class of interface problems or transmission problems, which are relevant to many applications such as groundwater flow [29], electromagnetics [27], semiconductor modeling [22,32], and fuelcells [50]. The coefficients in these applications may have large jumps across interfaces between regions with different material properties, i.e. $J(a):=\max_i a_i/\min_i a_i\gg 1$. Due to this $J(a)$ and the meshsize, the finite element discretization of (1.1) is usually very ill-conditioned, which leads to deterioration in the rate of convergence of multilevel and domain decomposition methods [3,26,47].

Only in some special circumstances, we are able to show the (nearly) uniform convergence of the multilevel and (overlapping) domain decomposition methods, see [12,48,49,23,37] for examples. For general cases, we usually need some special techniques to obtain robust iterative methods, see [16,41,25,1,45]. Recently, in [57,61] we analyzed the eigenvalue distributions of the standard multilevel and overlapping domain decomposition preconditioned systems, and showed that there are only a small fixed number of eigenvalues may deteriorate with respect to the discontinuous jump or meshsize, and that all the other eigenvalues are bounded below and above nearly uniformly with respect to the jump and meshsize. As a result, we proved that the convergence rate of the preconditioned conjugate

gradient methods is uniform with respect to the large jump, and depends logarithmically on meshsize. These results justified that the standard multilevel and domain decomposition preconditioners are efficient and robust for finite element discretization of (1.1) on *quasi-uniform grids*. In this paper, we extend our results to *locally refined grids*.

The discontinuity of diffusion coefficients causes a lack of regularity of the solution to (1.1), which, in turn, leads to deterioration in the rate of convergence for finite element approximations over quasi-uniform triangulations. Adaptive finite element methods through local mesh refinements can be applied to recover the optimal rate of convergence [15]. In order to achieve the optimal computational complexity in adaptive finite element methods, it is imperative to study fast algorithms for solving the linear system of equations arising from the finite element discretization. The distinct feature of applying multigrid methods on locally refined meshes is that the number of nodes of nested meshes obtained by local refinements may not grow exponentially, violating one of the key properties of multilevel methods on uniform meshes that leads to optimal $\mathcal{O}(N)$ complexity. Indeed, let N be the number of unknowns in the finest space, the complexity of smoothing can be as bad as $\mathcal{O}(N^2)$ [34]. This prevents direct application of algorithms and theories developed in [57] for quasi-uniform grids to locally refined grids.

To achieve optimal complexity $\mathcal{O}(N)$, the smoothing in each level must be restricted to the newly added unknowns and their neighbors; see, for example, [6,11,34]. Such methods are referred to as *local multilevel methods* in [6]. As an extreme case, one can preform the smoothing only on newly added nodes turning a coarse grid to a fine grid. The resulting method is known as the hierarchical basis method [59,8]. In two dimensions, the hierarchical basis methods are proven to be robust for jump coefficient problems on locally refined meshes (cf. [8]). In three dimensions, however, classic multilevel and domain decomposition methods, including the hierarchical basis multigrid methods, deteriorate rapidly due to the presence of discontinuity of coefficients. To obtain robust rates of convergence for multigrid methods, one has to use special coarse spaces [23,40] or assume that the distribution of diffusion coefficients satisfies the so called quasi-monotone condition [23]. Therefore the three dimensional case is much more difficult. There are other works [2,28] on optimal complexity of local multilevel methods in three dimensions, but the problems with discontinuous coefficients remain open.

We shall design and prove the efficiency and robustness of local multilevel preconditioners for the finite element discretization of problem (1.1) on bisection grids – one class of locally refined grids. In these preconditioners, we use a global smoothing in the finest mesh; and for each newly

added node, we perform smoothing only for three vertices - the new vertex and its two parents vertices (the vertices sharing the same edge with the new vertex). We analyze the eigenvalue distribution of the multilevel preconditioned matrix, and prove that there are only a fixed number of small eigenvalues deteriorated by the coefficient and mesh-size; the other eigenvalues are bounded nearly uniformly. Thus, the resulting preconditioned conjugate gradient algorithm converges uniformly with respect to the jump and logarithmically with respect to the mesh size of the discretization. We establish our results of this type in both two and three dimensions.

To attack the geometric structure of bisection grids, we use the decomposition of bisection grids developed in the recent work by Chen, Nochetto and Xu [19,56]. This approach enables us to introduce a natural decomposition of the finite element space into subspaces consisting only the newest vertices and their two parents vertices. In the analysis of these local multilevel preconditioners, one of the key ingredient is the *stable decomposition* (see Theorem 4.2). For the standard multilevel preconditioners on uniform mesh, in [57] we used the approximation and stability properties of the weighted L^2 projection ([12]) to construct a stable decomposition. This weighted L^2 projection is no longer applicable for the local multilevel preconditioners, since it is a global projection. In order to preserve the local natural of the highly graded meshes, we introduce a local interpolation operator, which we manage to prove similar approximation and stability properties (see Theorem 3.1 and 3.2) as the weighted L^2-projection. Our local quasi-interpolation operator and the corresponding analysis is much more delicate than that in [19,56] for the Poisson equation. We should remark that due to this space decomposition, we are able to remove the assumption, *nested local refinement*, which is used in most existing work on multilevel methods on local refinement grids [2,28].

The rest of the paper is organized as follows. In Section 2, we give some notation and recall some fundamental results as in [57]. In Section 4, we study bisection grids, and review some technical tools from [19,56]. Here we restrict ourself to a kind of special bisection scheme, namely the newest vertex bisection. Then in Section 4, we study some technical results of space decomposition, and present the optimal/stable decomposition and the strengthened Cauchy-Schwarz inequality on bisection grids. In Section 5, we analyze multilevel preconditioners, i.e., the BPX preconditioner and the multigrid V-cycle preconditioner, and prove convergence results for the preconditioned conjugate gradient algorithm. In Section 6, we present numerical experiments to justify our theoretical results.

We will use the following short notation, $x \lesssim y$ means $x \leq Cy$, $x \gtrsim y$ means $x \geq cy$ and $x \eqsim y$ means $cx \leq y \leq Cx$ where c and C are generic

positive constants independent of the variables appearing in the inequalities and any other parameters related to mesh, space and coefficients.

2 Preliminary

In this section, we introduce some notation, set up our problem, and review briefly some facts about the preconditioned conjugate gradient algorithm.

2.1 Notation and Problem

Given a set of positive constants $\{a_i\}_{i=1}^M$, we define the following weighted inner products on the space $H^1(\Omega)$

$$(u,v)_{0,a} = \sum_{i=1}^{M} a_i (u,v)_{L^2(\Omega_i)} \quad \text{and} \quad (u,v)_{1,a} = \sum_{i=1}^{M} a_i (\nabla u, \nabla v)_{L^2(\Omega_i)}$$

with the induced weighted L^2 norm $\|\cdot\|_{0,a}$, and the weighted H^1-seminorm $|\cdot|_{1,a}$, respectively. We denote by

$$\|u\|_{1,a} = \left(\|u\|_{0,a}^2 + |u|_{1,a}^2\right)^{\frac{1}{2}},$$

and the related inner product and the induced energy norm by

$$(u,v)_A = A(u,v) := (u,v)_{1,a}, \qquad \|u\|_A = \sqrt{A(u,u)}.$$

To impose the Dirichlet boundary condition in (1.1), we define

$$H^1_{g_D,\Gamma_D} = \{v \in H^1(\Omega) : v|_{\Gamma_D} = g_D \text{ in the trace sense}\},$$

and $H^1_D := H^1_{0,\Gamma_D}$. Given a shape regular triangulation $\mathcal{T}_h$, which could be highly graded, we define $\mathcal{V}_h$ as the standard piecewise linear and global continuous finite element space on $\mathcal{T}_h$. Given $f \in H^{-1}(\Omega)$ and $g_N \in H^{1/2}(\Gamma_N)$, the linear finite element approximation of (1.1) is the function $u \in \mathcal{V}_h \cap H^1_{g_D,\Gamma_D}$, such that

$$A(u,v) = \langle f, v\rangle + \int_{\Gamma_N} g_N v \quad \text{for all } v \in \mathcal{V}_h \cap H^1_D. \tag{2.1}$$

Given any $u_0 \in \mathcal{V}_h \cap H^1_{g_D,\Gamma_D}$, the problem (2.1) is equivalent to finding $u \in \mathcal{V}_h \cap H^1_D$ such that

$$A(u,v) = \langle f, v\rangle + \int_{\Gamma_N} g_N v - A(u_0,v) \quad \text{for all } v \in \mathcal{V}_{h,D}. \tag{2.2}$$

We thus consider the space $\mathcal{V}_{h,D} := \mathcal{V}_h \cap H^1_D$. The bilinear form $A(\cdot,\cdot)$ will then introduce a symmetric positive definite (with respect to standard L^2-inner product) operator, still denoted by A, from $\mathcal{V}_{h,D}$ to $\mathcal{V}_{h,D}$ as

$$(Au, v) = A(u, v).$$

Define $b \in \mathcal{V}_{h,D}$ as

$$(b, v) = \langle f, v\rangle + \int_{\Gamma_N} g_N v - A(u_0, v) \quad \text{for all } v \in \mathcal{V}_h \cap H^1_{0,\Gamma_D}.$$

We then get the following operator equation on $\mathcal{V}_{h,D}$

$$Au = b. \tag{2.3}$$

For simplicity, in the remainder of the paper, we should omit the subscript D in $\mathcal{V}_{h,D}$ without ambiguity.

We are interested in solving equation (2.3) by the preconditioned conjugate gradient methods with BPX and multigrid preconditioners.

2.2 Preconditioned Conjugate Gradient Method

Let B be a symmetric positive definite (SPD) operator. Applying it to both sides of (2.3), we get an equivalent equation

$$BAu = Bb. \tag{2.4}$$

We apply the conjugate gradient method to solve (2.4) and the resulting method is known as the *preconditioned conjugate gradient (PCG)* method, where B is called a *preconditioner*.

Let $\kappa(BA) = \lambda_{\max}(BA)/\lambda_{\min}(BA)$ be the (generalized) condition number of the preconditioned system BA. Starting from an arbitrary initial guess u_0, we have the following well known convergence rate estimate for the kth iteration u_k $(k \geq 1)$ in PCG (see e.g. [39])

$$\frac{\|u - u_k\|_A}{\|u - u_0\|_A} \leq 2\left(\frac{\sqrt{\kappa(BA)} - 1}{\sqrt{\kappa(BA)} + 1}\right)^k.$$

Therefore if the condition number $\kappa(BA)$ is uniformly bounded, then PCG algorithm converges uniformly. Here the uniformity means the independence of the size of the matrix A. Later on, when A is related to equation (1.1), we shall also discuss the uniformity of convergence with respect to the jump of diffusion coefficients.

If there are some isolated small or large eigenvalues, we can sharpen the above convergence rate estimate; see [5].

Theorem 2.1 *[5] Suppose that* $\sigma(BA) = \sigma_0(BA) \cup \sigma_1(BA)$ *such that there are* m *elements in* $\sigma_0(BA)$ *and* $\alpha \leq \lambda \leq \beta$ *for each* $\lambda \in \sigma_1(BA)$. *Then*

$$\frac{\|u - u_k\|_A}{\|u - u_0\|_A} \leq 2K \left(\frac{\sqrt{\beta/\alpha} - 1}{\sqrt{\beta/\alpha} + 1} \right)^{k-m}, \tag{2.5}$$

where

$$K = \max_{\lambda \in \sigma_1(BA)} \prod_{\mu \in \sigma_0(BA)} \left| 1 - \frac{\lambda}{\mu} \right|.$$

If there are only m small eigenvalues in $\sigma_0(BA)$, say

$$0 < \lambda_1 \leq \lambda_2 \cdots \leq \lambda_i \ll \lambda_{m+1} \leq \cdots \leq \lambda_n,$$

then

$$K = \prod_{i=1}^{m} \left| 1 - \frac{\lambda_n}{\lambda_i} \right| \leq \left(\frac{\lambda_n}{\lambda_1} - 1 \right)^m = (\kappa(BA) - 1)^m. \tag{2.6}$$

Therefore the convergence rate of PCG algorithm will be dominated by the factor $(\sqrt{\beta/\alpha} - 1)/(\sqrt{\beta/\alpha} + 1)$, i.e. by β/α where $\beta = \lambda_n(BA)$ and $\alpha = \lambda_{m+1}(BA)$. We define the "effective condition number" as follows.

Definition 1 *Let* $\mathcal{V}$ *be an* n*-dimensional Hilbert space and* $T : \mathcal{V} \to \mathcal{V}$ *be a symmetric and positive definite operator. For any integer* $m \in [1, n-1]$, *the* m*th* effective condition number *of* T *is defined by*

$$\kappa_m(T) = \frac{\lambda_n(T)}{\lambda_{m+1}(T)}$$

where $\lambda_{m+1}(T)$ *is the* $(m+1)$*-th minimal eigenvalue of* T.

As a corollary of Theorem 2.1, we have

$$\frac{\|u - u_k\|_A}{\|u - u_0\|_A} \leq 2(\kappa(BA) - 1)^m \left(\frac{\sqrt{\kappa_m(BA)} - 1}{\sqrt{\kappa_m(BA)} + 1} \right)^{k-m}. \tag{2.7}$$

From (2.7), given a tolerance ϵ, the number of iterations of the PCG method to reduce the relative error below the tolerence ϵ is [4,5]

$$m + \left\lceil \left(\log\left(\frac{2}{\epsilon}\right) + m |\log(\kappa(BA) - 1)| \right) / c_0 \right\rceil,$$

where $c_0 = \log\left((\sqrt{\kappa_m(BA)} + 1)/(\sqrt{\kappa_m(BA)} - 1) \right)$. Therefore if there exists an $m \geq 1$ such that the mth effective condition number is bounded uniformly, then the PCG algorithm will still converge almost uniformly, even though the standard condition number $\kappa(BA)$ may not be uniformly bounded.

To estimate the effective condition number $\lambda_{m+1}(A)$, we use a fundamental tool known as the *Courant "minimax" principle* (see e.g. [24]).

Theorem 2.2 *Let $\mathcal{V}$ be an n-dimensional Hilbert space with inner product $(\cdot,\cdot)_{\mathcal{V}}$ and $T : \mathcal{V} \to \mathcal{V}$ a symmetric positive operator on $\mathcal{V}$. Suppose $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ are the eigenvalues of T, then*

$$\lambda_{m+1}(T) = \max_{\dim(S)=m} \min_{0\neq v\in S^{\perp}} \frac{(Tv,v)_{\mathcal{V}}}{(v,v)_{\mathcal{V}}}$$

for $i = 1, 2, \cdots, n-1$. Especially, for any subspace $\mathcal{V}_0 \subset \mathcal{V}$ with $\dim(\mathcal{V}_0) = n - m$

$$\lambda_{m+1}(T) \geq \min_{0\neq v\in \mathcal{V}_0} \frac{(Tv,v)_{\mathcal{V}}}{(v,v)_{\mathcal{V}}}. \tag{2.8}$$

If both A and B are SPD operators, then BA is SPD in the inner product induced by B^{-1} and A. We shall apply Theorem 2.2 to $T = BA$ and $(u,v)_{\mathcal{V}} := (B^{-1}u,v)_{L^2}$. Therefore if we have an inequality of the type $(Av,v) \geq c(B^{-1}v,v)$ for all v in a suitable subspace $\mathcal{V}_0$ with $\dim(\mathcal{V}_0) = n - m$, we can get a lower bound of $\lambda_{m+1}(BA)$.

3 Local Quasi-interpolation

The theoretical justification of the robustness of multilevel preconditioners relies on establishing approximation and stability properties of certain interpolation operators. In [57,61], we used the weighted L^2-projection $Q_h^a : L^2(\Omega) \to \mathcal{V}_h$ defined by

$$(Q_h^a u, v_h)_{0,a} = (u, v_h)_{0,a} \quad \text{for all } v_h \in \mathcal{V}_h.$$

For the analysis of local multilevel preconditioners, we require the interpolation operator to preserve certain local structure. Therefore, the weighted L^2-projection, which is a global operator, is not appropriate. On the other hand, the standard nodal interpolation operator is local but not stable in the energy norm. Local quasi-interpolation, such as Scott-Zhang operators [42], are developed to achieve both locality and stability. The stability constant of the standard quasi-interpolation, however, depends on the jump of diffusion coefficients.

In this section, we construct a local and stable quasi-interpolation operator by gluing Scott-Zhang operators in each subdomains and interfaces. Our operator is stable uniformly with respect to the jump of coefficients and nearly uniform to the mesh size of the triangulation. We stress that this local quasi-interpolation operator is designed for the analysis and will not enter the algorithm.

3.1 Notation on Triangulations

Let us introduce some notation related to the domain and its triangulations. As we mentioned earlier, we assume that the polygonal or polyhedral subdomains Ω_i $(i = 1, \cdots, M)$ are open, disjoint to each other, and satisfy $\cup_{i=1}^{M} \overline{\Omega}_i = \overline{\Omega}$. We denote $\Gamma_{ij} = \partial\Omega_i \cap \partial\Omega_j$, or simply Γ if without ambiguity, as the *interface* between two subdomains Ω_i and Ω_j. The subdomains Ω_i $(i = 1, \cdots M)$ may possibly have complicated geometry but we assume that they are resolved by an initial *conforming* triangulation $\mathcal{T}_0$. Recall that a triangulation $\mathcal{T}$ is called *conforming* if the intersection of any two elements τ and τ' in $\mathcal{T}$ either consists of a common vertex, edge, face (when $d = 3$), or empty.

Let $\mathcal{N}$, $\mathcal{E}$ and $\mathcal{F}$ (when $d = 3$) denote the set of vertices, edges, and faces of $\mathcal{T}$ respectively. For each vertex $p \in \mathcal{N}$, we define local patch $\omega_p := \cup_{\tau \ni p} \tau$ and, for $\tau \in \mathcal{T}$, $\omega_\tau = \cup_{p \in \tau} \omega_p$. Similarly, on the $(d-1)$ dimensional interface Γ, o_p, o_e and o_f denote the intersection of corresponding local patches and the interface. The linear finite element space associated to $\mathcal{T}$ is denoted by $\mathcal{V}(\mathcal{T})$, or simply $\mathcal{V}$. More generally, for any subset $\S \subset \mathcal{T}$, $\mathcal{V}(\S)$ denote the finite element subspace restricted to the subset G. Similarly, we should denote $\mathcal{N}(G) \subset \mathcal{N}$, $\mathcal{E}(G) \subset \mathcal{E}$ and $\mathcal{F}(G) \subset \mathcal{F}$ as the set of vertices, edges, and faces in $\overline{G} \subset \overline{\Omega}$, respectively.

For each element $\tau \in \mathcal{T}$, we define $h_\tau = |\tau|^{1/d}$ and ρ_τ for the radius of its inscribed ball. In the whole paper, we assume that the triangulation is *shape regular* in the sense $h_\tau \eqsim \rho_\tau$. Let h denote the piecewise constant meshsize function with $h|_\tau = h_\tau$, and $h_{\min} := \min_{\tau \in \mathcal{T}} h_\tau$. We should also denote h_e by the length of an edge $e \in \mathcal{E}$ and h_f by the diameter of a face $f \in \mathcal{F}$. Moreover, we define h_p as the diameter of the local patch ω_p. By the shape regularity assumption, for all $e, f, \tau \subset \omega_p$, we have

$$h_p \eqsim h_e \eqsim h_f \eqsim h_\tau.$$

3.2 Technical Lemmas

Here for completeness, we quote some technical lemmas from [12], which will be used later for proving the approximation and stability of our local interpolation operator.

In two dimensions, it is well known that $H^1(\Omega)$ is not embedded into $L^\infty(\Omega)$. But for finite element functions, we can control the L^∞ norm by its H^1-norm with a factor $|\log h|^{1/2}$.

Lemma 3.1 ([12, Lemma 2.3]) *Let G be a bounded Lipschitz domain in $\mathbb{R}^2$ and $\mathcal{V}(G)$ be a finite element space based on a quasi-uniform triangulation of G with mesh size h. Then for all $v \in \mathcal{V}(G)$, it satisfies*

$$\|v\|_{L^\infty(G)} \lesssim |\log h|^{1/2} \|v\|_{H^1(G)}.$$

In three dimensions, the trace of a H^1-function on an edge is not well defined. But for a finite element function, its L^2-norm on an edge can be bounded by its H^1-norm with a factor $|\log h|^{1/2}$.

Lemma 3.2 ([12, Lemma 2.4]) *Let G be a polyhedral domain in $\mathbb{R}^3$ and $\mathcal{V}(G)$ be a finite element space based on a quasi-uniform triangulation of G with mesh size h. Then for all $v \in \mathcal{V}(G)$, and any edge E of G, there holds*

$$\|v\|_{L^2(E)} \lesssim |\log h|^{1/2} \|v\|_{H^1(G)}.$$

In the following application of Lemma (3.2), G is usually taken as a local patch which is quasi-uniform.

3.3 Stable Local Quasi-Interpolation

Given a conforming triangulation $\mathcal{T}$, the standard Scott-Zhang interpolation operator $\Pi : H^1(\Omega) \to \mathcal{V}(\mathcal{T})$ can be defined as follows. For any $p \in \mathcal{N}(\mathcal{T})$, choose an element $\tau \subset \omega_p$. The choice of τ is not unique. Let $\{\lambda_{\tau,i} : i = 1, \cdots, d+1\}$ be the barycentric coordinates of τ. One can define the L^2-dual basis $\{\theta_{\tau,i} : i = 1, \cdots, d+1\}$ of $\{\lambda_{\tau,i} : i = 1, \cdots, d+1\}$, namely, $\int_\tau \theta_{\tau,i}\lambda_{\tau,j} = \delta_{ij}$. We define a quasi-interpolation Π as

$$\Pi v = \sum_{p \in \mathcal{N}(\mathcal{T})} \left(\int_\tau \theta_p v \right) \phi_p,$$

where $\{\phi_p\}_{p \in \mathcal{N}(\mathcal{T})}$ is the set of nodal basis of $\mathcal{V}$. Note that if $v \in \mathcal{P}_1(\tau)$, then $\int_\tau \theta_p v = v(p)$ and thus Π will preserve linear polynomial in ω_τ.

The following properties of the quasi-interpolation Π can be found at [36,42].

Lemma 3.3 *The interpolation operator Π satisfies the following properties:*

$$\textit{Stability:} \quad \|\Pi v\|_{L^2(\tau)} \lesssim \|v\|_{L^2(\omega_\tau)}, \quad \|\Pi v\|_{H^1(\tau)} \lesssim \|v\|_{H^1(\omega_\tau)}; \tag{3.1}$$

$$\textit{Locality:} \quad (\Pi v)|_\tau = v|_\tau \textit{ if } v \in \mathcal{V}(\omega_\tau); \tag{3.2}$$

$$\textit{Approximability:} \quad \|h^{-1}(v - \Pi v)\|_{L^2(\tau)} \lesssim \|v\|_{H^1(\omega_\tau)}. \tag{3.3}$$

We treat the interior of each subdomain and the interfaces separately and use a subscript to indicate different quasi-interpolations. For example $\Pi_i : L^2(\Omega_i) \to \mathcal{V}(\Omega_i)$ denotes a Scott-Zhang interpolation restricted to Ω_i and $\Pi_\Gamma : L^2(\Gamma) \to \mathcal{V}(\Gamma)$ on the interface $\Gamma \subset \partial\Omega_i$. On the interface, Π_Γ has similar properties as stated in Lemma 3.3:

$$\|\Pi_\Gamma v\|_{L^2(f)} \lesssim \|v\|_{L^2(o_f)}; \tag{3.4}$$

$$\|h^{-1}(v - \Pi_\Gamma v)\|_{L^2(f)} \lesssim \|v\|_{H^1(o_f)}. \tag{3.5}$$

We now construct a local interpolation operator $\mathcal{I}_h^a$ with desirable local approximation and stability properties in the weighted Sobolev norms. Given a $u \in H^1(\Omega)$, we define $\mathcal{I}_h^a u \in \mathcal{V}(\mathcal{T})$ such that for $p \in \mathcal{N}(\Omega_i)$

$$\mathcal{I}_h^a u(p) := \begin{cases} (\Pi_i u)(p), & \text{if } p \in \mathcal{N}(\Omega_i) \setminus \mathcal{N}(\partial\Omega_i), \\ (\Pi_\Gamma u)(p), & \text{if } p \in \mathcal{N}(\Gamma) \setminus \mathcal{N}(\partial\Gamma) \\ & \quad \text{for each interface } \Gamma \subset \partial\Omega_i, \\ 0, & \text{otherwise.} \end{cases} \tag{3.6}$$

For a vertex p, let us denote by σ_p the simplex chosen to define the nodal value at p. Then the interpolant $\mathcal{I}_h^a$ is uniquely determined by the mapping $p \to \sigma_p$. In (3.6), if p is in the interior of some subdomain Ω_i, then we choose a d-simplex $\sigma_p \subset \Omega_i$; if p is in the interior of the interface Γ, we choose $\sigma_p \subset \Gamma$ as a $(d-1)$-simplex on the interface. The choice of σ_p is not unique. However, in order to preserve the local structure of the adaptive grids, we need to choose σ_p carefully. This will be clear in Section 4 when we discuss the geometry of the bisection grids.

Now we are in the position to present the main result in this section:

Theorem 3.1 *Let $\Omega \subset \mathbb{R}^d$ with $d = 2$ or 3 and $\mathcal{T}_h$ be a triangulation of Ω with mesh size h. Then for all $u \in H^1(\Omega)$, we have*

$$\|h^{-1}(u - \mathcal{I}_h^a u)\|_{0,a,\Omega} \lesssim \left\| |\log h|^{1/2} u \right\|_{1,a,\Omega}.$$

Proof Here we present the proof for $d = 3$. The case for $d = 2$ can be proved similarly (indeed with simpler arguments by using the nodal interpolation and Lemma 3.1 cf. [8]). We define $w = \mathcal{I}_h^a u \in \mathcal{V}$ as (3.6), and denote $w_i := \Pi_i u$ for simplicity.

Note that $w(p) = w_i(p)$ for all interior nodes $p \in \mathcal{N}(\Omega_i) \setminus \mathcal{N}(\partial\Omega_i)$. By the equivalence of continuous and discrete L^2 norms, we obtain

$$\begin{aligned} &\|h^{-1}(w - w_i)\|^2_{L^2(\Omega_i)} \\ &\lesssim \sum_{p \in \mathcal{N}(\partial\Omega_i)} h_p (w - w_i)^2(p) \le \sum_{\Gamma \subset \partial\Omega_i} \sum_{p \in \mathcal{N}(\Gamma)} h_p (w - w_i)^2(p) \\ &\le \sum_{\Gamma \subset \partial\Omega_i} \left(\sum_{p \in \mathcal{N}(\Gamma)} h_p \left(\Pi_\Gamma u - w_i\right)^2(p) + \sum_{p \in \mathcal{N}(\partial\Gamma)} h_p w_i^2(p) \right) \\ &\lesssim \sum_{\Gamma \subset \partial\Omega_i} \left(\sum_{f \subset \mathcal{F}(\Gamma)} h_f^{-1} \|\Pi_\Gamma u - w_i\|^2_{L^2(f)} + \|w_i\|^2_{L^2(\partial\Gamma)} \right). \end{aligned}$$

We need to bound two terms appearing in the last expression. For the first term, we have

$$\begin{aligned} h_f^{-1}\|\Pi_\Gamma u - w_i\|_{L^2(f)}^2 &\lesssim h_f^{-1}\|u - w_i\|_{L^2(o_f)}^2 \\ &\lesssim h_f^{-2}\|u - w_i\|_{L^2(\omega_f)}^2 + \|u - w_i\|_{H^1(\omega_f)}^2 \\ &\lesssim \|u\|_{H^1(\omega_f)}^2, \end{aligned}$$

where recall that $o_f \subset \partial\Omega_i$ is the local patch associated with f on the interface, and $\omega_f \subset \Omega_i$. In the first step, we used the L^2 stability (3.4) of Π_Γ and the fact $\Pi_\Gamma w_i = w_i|_\Gamma$. In the second step, we used the trace theorem (cf. [12, Lemma 2.1]), and in the last inequality, we used Lemma 3.3 for the Scott-Zhang interpolation Π_i. Hence we have that

$$\sum_{\Gamma\subset\partial\Omega_i}\sum_{f\subset\mathcal{F}(\Gamma)} h_f^{-1}\|w_i - \Pi_\Gamma u\|_{L^2(f)}^2 \lesssim \|u\|_{H^1(\Omega_i)}^2.$$

For the second term, we bound it by using the discrete Sobolev inequality Lemma 3.2 on each local patch ω_p and the local H^1-stability (3.1) of Π_i to obtain

$$\sum_{\Gamma\subset\partial\Omega_i}\|w_i\|_{L^2(\partial\Gamma)}^2 \lesssim \left\||\log h|^{1/2} w_i\right\|_{H^1(\Omega_i)}^2 \lesssim \left\||\log h|^{1/2} u\right\|_{H^1(\Omega_i)}^2.$$

Consequently,

$$\|h^{-1}(w_i - w)\|_{L^2(\Omega_i)} \lesssim \left\||\log h|^{1/2} u\right\|_{H^1(\Omega_i)}.$$

Finally, by the triangle inequality and the approximation property (3.3) of Π_i we have

$$\begin{aligned} \|h^{-1}(u - \mathcal{I}_h^a u)\|_{L^2(\Omega_i)} &\le \|h^{-1}(u - w_i)\|_{L^2(\Omega_i)} + \|h^{-1}(w_i - \mathcal{I}_h^a u)\|_{L^2(\Omega_i)} \\ &\lesssim \|u\|_{H^1(\Omega_i)} + \left\||\log h|^{1/2} u\right\|_{H^1(\Omega_i)}. \end{aligned}$$

Multiplying by a suitable weight and summing up over all subdomains on both sides, we get the desired estimate.

□

In general, we can not replace $\|u\|_{1,a}$ by the energy norm $|u|_{1,a}$ in the above lemma; see [52] for a counter example. To be able to use $|u|_{1,a}$ in the estimate, we introduce a subspace $\widetilde{H}_D^1(\Omega)$ of $H_D^1(\Omega)$ as follows:

$$\widetilde{H}_D^1(\Omega) = \left\{ u \in H_D^1(\Omega) : \int_{\Omega_i} u \,\mathrm{d}x = 0 \quad \text{for all } i \in I \right\},$$

where I is the set of indices of all *floating* subdomains:

$$I = \{i : \text{ meas}(\partial\Omega_i \cap \Gamma_D) = 0\}.$$

Let $m_0 := \#I$ be the cardinality of I. We emphasize that m_0 is a constant, depending only on the distribution of the coefficients, and $m_0 \leq M$. In this subspace $\widetilde{H}^1_D(\Omega)$, the interpolation $\mathcal{I}^a_h$ has the following properties.

Theorem 3.2 *For any $v \in \widetilde{H}^1_D(\Omega)$, we have the approximation property of $\mathcal{I}^a_h$*

$$\left\|h^{-1}(v - \mathcal{I}^a_h v)\right\|_{0,a} \lesssim |\log h_{\min}|^{\frac{1}{2}} |v|_{1,a}, \tag{3.7}$$

and the stability of $\mathcal{I}^a_h$ in the energy norm

$$|\mathcal{I}^a_h v|_{1,a} \lesssim |\log h_{\min}|^{\frac{1}{2}} |v|_{1,a}. \tag{3.8}$$

Proof For $v \in \widetilde{H}^1_D(\Omega)$, it satisfies the Poincaré-Friedrichs inequality on each subdomain Ω_i. Therefore we get $\|v\|_{0,a} \lesssim |v|_{1,a}$. The inequality (3.7) then follows from Lemma 3.1.

To prove inequality (3.8), we use the inequality (3.7) and the local L^2 projection $Q_\tau : L^2(\tau) \to \mathcal{P}_0(\tau)$ defined by $Q_\tau u|_\tau = |\tau|^{-1} \int_\tau u \,\mathrm{d}x$. Then on each element $\tau \in \mathcal{T}_h$, we have

$$\begin{aligned}
|\mathcal{I}^a_h v|^2_{H^1(\tau)} &\lesssim |\mathcal{I}^a_h v - Q_\tau v|^2_{H^1(\tau)} \lesssim h_\tau^{-2} \|\mathcal{I}^a_h v - Q_\tau v\|^2_{L^2(\tau)} \\
&\lesssim h_\tau^{-2} \left(\|v - \mathcal{I}^a_h v\|^2_{L^2(\tau)} + \|v - Q_\tau v\|^2_{L^2(\tau)} \right) \\
&\lesssim h_\tau^{-2} \|v - \mathcal{I}^a_h v\|^2_{L^2(\tau)} + |v|^2_{H^1(\tau)}
\end{aligned}$$

where in the last inequality, we used the approximation properties of Q_τ. Multiplying by a suitable weight and summing up over all $\tau \in \mathcal{T}$ on both sides, we get

$$|\mathcal{I}^a_h v|^2_{1,a} \lesssim \left\|h^{-1}(v - \mathcal{I}^a_h v)\right\|^2_{0,a} + |v|^2_{1,a} \lesssim |\log h_{\min}| \, |v|^2_{1,a}$$

where in the last step, we used inequality (3.7).

□

Remark 3.1 When the coefficients satisfy the quasi-monotone assumption, the factor $|\log h_{\min}|$ can be removed by arguments on a modified local patch; see [23,38]. □

4 Bisection Grids and Space Decomposition

In this section, we give a short overview of the framework in the multilevel space decomposition on bisection grids in the recent work [19,56]. Most of the material in this section can be found there.

4.1 Bisection Methods

We recall briefly the bisection algorithm for the mesh refinements. Detailed discussions can be found in [10,17,34] and the references cited therein.

Given a conforming triangulation $\mathcal{T}$ of Ω, for each element $\tau \in \mathcal{T}$, we assign an edge of τ to be the *refinement edge* of τ, denoted by $e(\tau)$ or simply e without ambiguity. This procedure is called *labeling*. Given a set of elements marked for refinement, the refinement procedure consists two steps:

(1) bisect the marked element into two elements by connecting the middle point of the refinement edge to the vertices not contained in the refinement edge;
(2) assign refinement edges for two new elements.

Given a labeled initial grid $\mathcal{T}_0$ of Ω and a bisection method, we define

$$\mathbb{F}(\mathcal{T}_0) = \{\mathcal{T} : \mathcal{T} \text{ is refined from } \mathcal{T}_0 \text{ using the chosen bisection method }\},$$
$$\mathbb{T}(\mathcal{T}_0) = \{\mathcal{T} \in \mathbb{F}(\mathcal{T}_0) : \mathcal{T} \text{ is conforming}\}.$$

Namely $\mathbb{F}(\mathcal{T}_0)$ contains all triangulations obtained from $\mathcal{T}_0$ using the chosen bisection method. But a triangulation $\mathcal{T} \in \mathbb{F}(\mathcal{T}_0)$ could be non-conforming and thus we define $\mathbb{T}(\mathcal{T}_0)$ as a subset of $\mathbb{F}(\mathcal{T}_0)$ containing only conforming triangulations.

Given any triangulation $\mathcal{T}$, we define $\overline{\mathcal{T}}_0 = \mathcal{T}$, and the kth uniform refinement $\overline{\mathcal{T}}_k \quad (k \geq 1)$ being the triangulation obtained by bisecting all element in $\overline{\mathcal{T}}_{k-1}$ only once. Note that for a conforming initial triangulation $\mathcal{T}_0$ with arbitrary labeling, $\overline{\mathcal{T}}_k \in \mathbb{F}(\mathcal{T}_0)$ but not necessarily in the set $\mathbb{T}(\mathcal{T}_0)$ in general. Throughout this paper, we shall consider bisection methods which satisfy the following two assumptions:

(B1) Shape Regularity: $\mathbb{F}(\mathcal{T}_0)$ is shape regular.

(B2) Conformity of Uniform Refinement: $\overline{\mathcal{T}}_k(\mathcal{T}_0) \in \mathbb{T}(\mathcal{T}_0)$ for all $k \geq 0$.

In two dimensions, newest vertex bisection with compatible initial labeling [33] satisfies (B1) and (B2). In three and higher dimensions, the bisection method by Kossaczký [31] and Stevenson [44] will satisfy (B1) and (B2). We note that to satisfy assumption (B2), the initial triangulation is modified by further refinement of each element, which deteriorates the shape regularity. Although (B2) imposes a severe restriction on the initial labeling, it is crucial to control the number of elements added in the completion which is indispensable to establish the optimal complexity of adaptive finite element methods [35].

4.2 Compatible Bisections

For a vertex $p \in \mathcal{N}(\mathcal{T})$ or an edge $e \in \mathcal{E}(\mathcal{T})$, we define the *first ring* of p or e to be

$$\mathcal{R}_p = \{\tau \in \mathcal{T} \,|\, p \in \tau\}, \quad \mathcal{R}_e = \{\tau \in \mathcal{T} \,|\, e \subset \tau\},$$

and the local patch of p or e as $\omega_p = \cup_{\tau \in \mathcal{R}_p} \tau$, and $\omega_e = \cup_{\tau \in \mathcal{R}_e} \tau$. Note that ω_p and ω_e are subsets of Ω, while $\mathcal{R}_p$ and $\mathcal{R}_e$ are subsets of $\mathcal{T}$ which can be thought of as triangulations of ω_p and ω_e, respectively. The cardinality of a set S will be denoted by $\#S$.

Given a labeled triangulation $\mathcal{T}$, an edge $e \in \mathcal{E}(\mathcal{T})$ is called a *compatible edge* if e is the refinement edge of τ for all $\tau \in \mathcal{R}_e$. For a compatible edge, the ring $\mathcal{R}_e$ is called a *compatible ring*, and the patch ω_e is called a *compatible patch*. Let p be the midpoint of e and $\mathcal{R}_p$ be the ring of p in the refined triangulation. A *compatible bisection* is a mapping $b_e : \mathcal{R}_e \to \mathcal{R}_p$. We then define the addition

$$\mathcal{T} + b_e := \mathcal{T} \backslash \mathcal{R}_e \cup \mathcal{R}_p.$$

For a compatible bisection sequence $\mathcal{B} := (b_1, \cdots, b_k)$, the addition $\mathcal{T} + \mathcal{B}$ is defined as

$$\mathcal{T} + \mathcal{B} = ((\mathcal{T} + b_1) + b_2) + \cdots + b_k,$$

whenever the addition is well defined. Note that if $\mathcal{T}$ is conforming, then $\mathcal{T} + b_e$ is conforming for a compatible bisection b_e, whence compatible bisections preserve the conformity of triangulations.

We now present a decomposition of meshes in $\mathbb{T}(\mathcal{T}_0)$ using compatible bisections, which will be instrumental later. We only give a pictorial demonstration in Fig. 4.1 to illustrate the decomposition. For the proof, we refer to [56].

Theorem 4.1 (Decomposition of Bisection Grids) *Let $\mathcal{T}_0$ be a conforming triangulation. Suppose the bisection method satisfies assumptions (B2), i.e., for all $k \geq 0$ all uniform refinements $\overline{\mathcal{T}}_k$ of $\mathcal{T}_0$ are conforming. Then for any $\mathcal{T} \in \mathbb{T}(\mathcal{T}_0)$, there exists a compatible bisection sequence $\mathcal{B} = (b_1, b_2, \cdots, b_N)$ with $N = \#\mathcal{N}(\mathcal{T}) - \#\mathcal{N}(\mathcal{T}_0)$ such that*

$$\mathcal{T} = \mathcal{T}_0 + \mathcal{B}. \tag{4.1}$$

We point out that in practice it is not necessary to store $\mathcal{B}$ explicitly during the refinement procedure. Instead we can apply coarsening algorithms to find out the decomposition. We refer to Chen and Zhang [20] (see also Chen [18]) for a vertex-oriented coarsening algorithm and the application to multilevel preconditioners and multigrid methods.

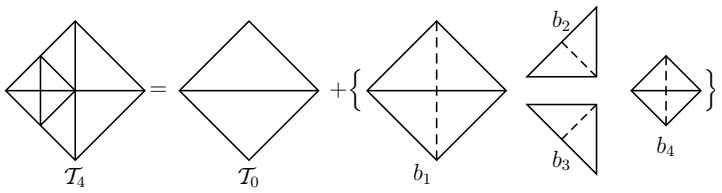


Fig. 4.1. A decomposition of a bisection grid.

For a compatible bisection $b_i \in \mathcal{B}$, we use the same subscript i to denote related quantities such as:

- e_i: the refinement edge;
- p_i: the midpoint of e_i;
- $\widetilde{\omega}_i = \omega_{p_i} \cup \omega_{p_{l_i}} \cup \omega_{p_{r_i}}$;
- $\mathcal{T}_i = \mathcal{T}_0 + (b_1, \cdots, b_i)$;
- ω_i: the patch of p_i i.e. ω_{p_i};
- p_{l_i}, p_{r_i}: two end points of e_i;
- h_i: the diameter of ω_i;
- $\mathcal{R}_i$: the first ring of p_i in $\mathcal{T}_i$.

4.3 Generation of Compatible Bisections

The generation of each element in the initial grid $\mathcal{T}_0$ is defined to be 0, and the generation of a child is 1 plus that of the father. The generation of an element $\tau \in \mathcal{T} \in \mathbb{F}(\mathcal{T}_0)$ is denoted by g_τ and coincides with the number of bisections needed to create τ from $\mathcal{T}_0$. For any vertex $p \in \mathcal{N}(\mathcal{T}_0)$, the generation of p is defined as the minimal integer k such that $p \in \mathcal{N}(\overline{\mathcal{T}}_k)$ and is denoted by g_p. In [56], we show that if $b_i \in \mathcal{B}$ is a compatible bisection, then all elements of $\mathcal{R}_i$ have the same generation g_i. Therefore we can introduce the concept of generation of compatible bisections. For a compatible bisection $b_i : \mathcal{R}_{e_i} \to \mathcal{R}_{p_i}$, we define $g_i = g(\tau), \tau \in \mathcal{R}_{p_i}$.

Throughout this paper we always assume $h(\tau) \eqsim 1$ for $\tau \in \mathcal{T}_0$. Then since a bisection of a simplex will reduce the volume by half, we have the following important relation between generation and mesh size

$$h_i \eqsim \gamma^{g_i}, \ \text{ with } \gamma = \left(\frac{1}{2}\right)^{1/d} \in (0,1).$$

In particular, we introduce a "level" (or generation) constant $L := \max_{\tau \in \mathcal{T}} g_\tau$. It is obvious that $L \eqsim \lceil |\log h_{\min}| \rceil$.

Different bisections with the same generation have disjoint local patches. Namely for two compatible bisections b_i and b_j with $g_j = g_i$, we then have $\omega_i \cap \omega_j = \varnothing$. A simple but important consequence is that, for all $u \in L^2(\Omega)$

and $k \geq 0$,

$$\sum_{g_i=k} \|u\|^2_{a,\widetilde{\omega}_i} \lesssim \|u\|^2_{a,\Omega}. \tag{4.2}$$

4.4 A Local Quasi-Interpolation

We define a sequences of quasi-interpolation operators recursively. Let $\mathcal{I}_0^a : \mathcal{V}(\mathcal{T}_N) \to \mathcal{V}_0$ be an arbitrary quasi-interpolation operator defined by (3.6). Assume $\mathcal{I}_{i-1}^a : \mathcal{V}(\mathcal{T}_N) \to \mathcal{V}(\mathcal{T}_{i-1})$ is defined. A compatible bisection b_i introduces a new vertex p_i from $\mathcal{T}_{i-1}$ to $\mathcal{T}_i = \mathcal{T}_{i-1} + b_i$. We construct $\mathcal{I}_i^a : \mathcal{V}(\mathcal{T}_N) \to \mathcal{V}(\mathcal{T}_i)$ as follows. For a boundary vertex $p \in \Gamma_D$, we simply define $(\mathcal{I}_i^a v)(p) = 0$ to reflect the vanishing boundary condition of v. To define the nodal value at the new vertex $p_i \notin \Gamma_D$, we choose σ_{p_i} as follows:

(i) if p_i is in the interior of some subdomain Ω_i, we choose σ_i as a d-simplex introduced by the bisection $\sigma_{p_i} \subset \omega_i$;
(ii) if p_i is in the interior of some interface Γ, we choose a $(d-1)$-simplex $\sigma_{p_i} \subset o_i$;
(iii) otherwise, we simply let $\sigma_{p_i} = \emptyset$ and define $(\mathcal{I}_i^a v)(p_i) = 0$.

For other vertices $p \in \mathcal{N}(\mathcal{T}_{i-1})$, let $\sigma_p \in \mathcal{T}_{i-1}$ be the simplex used to define $(\mathcal{I}_{i-1}^a v)(p)$, we update $(\mathcal{I}_i^a v)(p)$ according to the following cases:

(i) if $\sigma_p \subset \overline{\omega_p(\mathcal{T}_i)}$ we keep the nodal value, i.e., $(\mathcal{I}_i^a v)(p) = (\mathcal{I}_{i-1}^a v)(p)$;
(ii) otherwise we update σ_p as $\sigma_p \leftarrow \overline{\omega_p(\mathcal{T}_i)} \cap \sigma_p$ to define $(\mathcal{I}_i^a v)(p)$.

In either case, we ensure that the simplex $\sigma_p \subset \overline{\omega_p(\mathcal{T}_i)}$. In this way, we obtain a sequence of quasi-interpolation operators

$$\mathcal{I}_i^a : \mathcal{V}(\mathcal{T}_N) \to \mathcal{V}(\mathcal{T}_i), \quad i = 0 \cdots N.$$

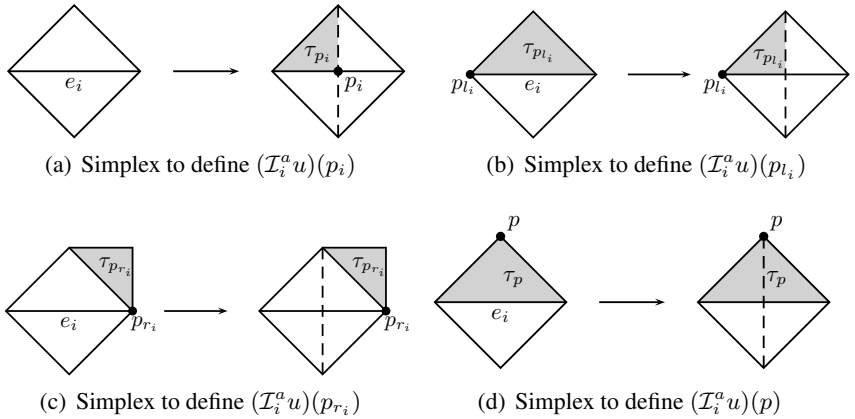
Note that in general $\mathcal{I}_N^a v \neq v$ since the simplex used to define nodal values of $\mathcal{I}_N^a v$ may not be in the finest mesh $\mathcal{T}_N$ but in $\mathcal{T}_{N-1}$. Figure 4.2 illustrates the choice of σ_p in different cases in 2D.

4.5 Stable Space Decomposition

Let $\phi_{i,p} \in \mathcal{V}(\mathcal{T}_i)$ denote the nodal basis at node $p \in \mathcal{N}(\mathcal{T}_i)$. Motivated by the stable three-point wavelet construction by Stevenson [43], we define the subspaces $\mathcal{V}_0 = \mathcal{V}(\mathcal{T}_0)$, and

$$\mathcal{V}_i = \operatorname{span}\{\phi_{i,p_i}, \phi_{i,p_{l_i}}, \phi_{i,p_{r_i}}\}.$$

Let $\{\phi_p : p \in \Lambda\}$ be a basis of $\mathcal{V}(\mathcal{T}_N)$, where Λ is the index set of the basis functions, and let $\mathcal{V}_p$ be the 1-dimensional subspace spanned by the



(a) Simplex to define $(\mathcal{I}_i^a u)(p_i)$ (b) Simplex to define $(\mathcal{I}_i^a u)(p_{l_i})$

(c) Simplex to define $(\mathcal{I}_i^a u)(p_{r_i})$ (d) Simplex to define $(\mathcal{I}_i^a u)(p)$

Fig. 4.2. Update of nodal values $\mathcal{I}_i^a u$ to yield $\mathcal{I}_{i-1}^a u$: the element τ chosen to perform the averaging that gives $(\mathcal{I}_i^a u)(p)$ must belong to $\omega_p(\mathcal{T}_i)$. This implies $(\mathcal{I}_i^a - \mathcal{I}_{i-1}^a)u(p) \neq 0$ possibly for $p = p_i, p_{l_i}, p_{r_i}$ and $= 0$ otherwise.

nodal bases associated to p in the finest grid. We choose the following space decomposition:

$$\mathcal{V} := \sum_{p\in\Lambda} \mathcal{V}_p + \sum_{i=0}^{N} \mathcal{V}_i. \tag{4.3}$$

Recall that b_i only changes the local patches of two end points of the refinement edge e_i going from $\mathcal{T}_{i-1}$ to $\mathcal{T}_i$. By construction $(\mathcal{I}_i^a - \mathcal{I}_{i-1}^a)v(p) = 0$ for $p \in \mathcal{N}(\mathcal{T}_i), p \neq p_i, p_{l_i}$ or p_{r_i}, which implies $v_i := (\mathcal{I}_i^a - \mathcal{I}_{i-1}^a)v \in \mathcal{V}_i$. Although $\mathcal{I}_N^a v \neq v$ in general, the difference $v - \mathcal{I}_N^a v$ is of high frequency in the finest mesh. Let us write $v - \mathcal{I}_N^a v = \sum_{p\in\Lambda} v_p$ as the basis decomposition. We then obtain a decomposition

$$v = \sum_{p\in\Lambda} v_p + \sum_{i=0}^{N} v_i, \quad v_i \in \mathcal{V}_i, \tag{4.4}$$

where for convenience we define $\mathcal{I}_{-1}^a := 0$. Moreover, we introduce a subspace $\widetilde{\mathcal{V}} := \mathcal{V} \cap \widetilde{H}_D^1(\Omega)$. Then we have the following stable decomposition.

Theorem 4.2 (Stable Decomposition) *Given a triangulation $\mathcal{T}_N = \mathcal{T}_0 + \mathcal{B}$ in $\mathbb{T}(\mathcal{T}_0)$, let $L = \max_{\tau\in\mathcal{T}_N} g(\tau)$.*

(i) For any $v \in \mathcal{V}$, there exist $v_p \in \mathcal{V}_p$ $(p \in \Lambda)$ and $v_i \in \mathcal{V}_i$ $(i = 1, \cdots, N)$ such that $v = \sum_{p\in\Lambda} v_p + \sum_{i=0}^N v_i$ and

$$\sum_{p\in\Lambda} h_p^{-2}\|v_p\|_{0,a}^2 + \|v_0\|_{1,a}^2 + \sum_{i=1}^{N} h_i^{-2}\|v_i\|_{0,a}^2 \lesssim c_d(L)\|v\|_{1,a}^2, \tag{4.5}$$

where $c_d(L) = \begin{cases} 1, & d=1 \\ L^2, & d=2 \\ 2^L, & d=3 \end{cases}$.

(ii) For any $v \in \widetilde{\mathcal{V}}$, *there exist* $v_p \in \mathcal{V}_p$ $(p \in \Lambda)$ *and* $v_i \in \mathcal{V}_i$ $(i = 1, \cdots, N)$ *such that* $v = \sum_{p\in\Lambda} v_p + \sum_{i=0}^{N} v_i$ *and*

$$\sum_{p\in\Lambda} h_p^{-2}\|v_p\|_{0,a}^2 + \|v_0\|_{1,a}^2 + \sum_{i=1}^{N} h_i^{-2}\|v_i\|_{0,a}^2 \lesssim L^2\|v\|_{1,a}^2 \tag{4.6}$$

Proof The result of (i) is standard. We may use the standard nodal interpolation operator to define a decomposition using hierarchical basis (cf. [55]).

Now we prove (ii). Given a $v \in \widetilde{\mathcal{V}}$, we define $v_0 := \mathcal{I}_0^a v$ and $v_i := (\mathcal{I}_i^a - \mathcal{I}_{i-1}^a)v$. For $v - \mathcal{I}_N^a v = \sum_{p\in\Lambda} v_p$, by the approximability of the quasi-interpolation, cf. (3.7), we have

$$\sum_{p\in\Lambda} h_p^{-2}\|v_p\|_{0,a}^2 \lesssim \|h^{-1}(v - \mathcal{I}_N^a v)\|_{0,a}^2 \lesssim L\|v\|_{1,a}^2. \tag{4.7}$$

On the other hand, by Theorem 3.2 we obtain

$$\begin{aligned}
&\|\mathcal{I}_0^a v\|_{1,a}^2 + \sum_{i=1}^{N} h_i^{-2}\|(\mathcal{I}_i^a - \mathcal{I}_{i-1}^a)v\|_{0,a,\omega_i}^2 \\
&= \|\mathcal{I}_0^a v\|_{1,a}^2 + \sum_{l=1}^{L}\sum_{g_i=l} h_l^{-2}\|(\mathcal{I}_i^a - \mathcal{I}_{i-1}^a)v\|_{0,a,\omega_i}^2 \\
&\lesssim \left(\sum_{i=1}^{L} |\log h_{\min}|\right)\|v\|_{1,a}^2 \lesssim L^2\|v\|_{1,a}^2.
\end{aligned}$$

Then (4.6) follows by adding the above inequality to inequality (4.7).

□

Remark 4.1 The estimate (4.5) is not uniform for $d \geq 2$. For $d = 2$, $L \approx |\log h_{\min}|$ and the growth of $c_2(L)$ is acceptable. But for $d = 3$, the constant $c_3(L) = 2^L$ grows exponentially. This is the main reason that the hierarchical basis multilevel method deteriorates rapidly in 3D (cf. [60, 7]). For discontinuous coefficients problems, it seems unlikely to find a better decomposition with a better constants; see the counterexamples in [12, 37].

If the coefficients satisfy certain monotonicity, e.g. quasi-monotonicity (cf. [23,38]) in the local patches, one can show that the interpolation operator defined above is stable in the energy norm without deterioration. □

Remark 4.2 With a close look at the proof of (4.6), we may regroup the $v_i = (\mathcal{I}_i^a - \mathcal{I}_{i-1}^a)v$ into groups $\cup_{l=1}^{L'} G(l) = \{1, 2, \cdots, N\}$ such that for any $i, j \in G(l), \omega_j \cap \omega_i = \varnothing$ and therefore

$$\sum_{i=1}^{N} h_i^{-2} \|v_i\|_{0,a,\omega_i}^2 = \sum_{l=1}^{L'} \sum_{j \in G(l)} h_j^{-2} \|v_j\|_{0,a,\omega_i}^2 \leq L' |\log h_{\min}| \|v\|_{1,a}^2.$$

The constant L' could be much smaller than L; see Section 6 for numerical examples. □

4.6 Strengthened Cauchy-Schwarz inequality

An important tool in analysis of the multiplicative preconditioner is the following strengthened Cauchy-Schwarz inequality. A proof can be found in [19,56].

Lemma 4.1 (Strengthened Cauchy-Schwarz Inequality) *For any $u_i, v_i \in \mathcal{V}_i$, $i = 0, 1, \cdots, N$, we have*

$$\left| \sum_{i=0}^{N} \sum_{j=i+1}^{N} A(u_i, v_j) \right| \lesssim \left(\sum_{i=0}^{N} |u_i|_{1,a}^2 \right)^{1/2} \left(\sum_{i=0}^{N} h_i^{-2} \|v_i\|_{0,a}^2 \right)^{1/2}. \quad (4.8)$$

As a corollary of (4.8) and the inverse inequality, we have

$$\Big\| \sum_{i=0}^{N} u_i \Big\|_{1,a}^2 \lesssim \sum_{i=0}^{N} h_i^{-2} \|u_i\|_{0,a}^2 \quad (4.9)$$

5 Multilevel Preconditioners

In this section, we shall analysis the eigenvalue distribution of the BPX preconditioner and the multigrid V-cycle preconditioner on bisection grids, and prove the effective conditioner number is uniformly bounded.

5.1 BPX (Additive) Preconditioner

To simplify the notation, we include $\mathcal{V}_{N+1} = \mathcal{V}$ and rewrite our space decomposition as $\mathcal{V} = \sum_{i=0}^{N+1} \mathcal{V}_i$. Based on this space decomposition, we choose SPD smoothers $R_i : \mathcal{V}_i \to \mathcal{V}_i$ satisfying

$$(R_i^{-1} u_i, u_i)_{0,a} \eqsim h_i^{-2} (u_i, u_i)_{0,a}, \quad \forall u_i \in \mathcal{V}_i \ (i = 1, \cdots, N+1). \quad (5.1)$$

According to [57], both of the standard Jacobi and symmetric Gauss-Seidel smoother satisfy the above assumption. On the coarsest level, i.e. when $i = 0$, we choose the exact solver $R_0 = A_0^{-1}$. Let $Q_i^a : \mathcal{V} \to \mathcal{V}_i$ be the weighted L^2 projection. Then we can define the BPX-type preconditioner

$$B = \sum_{i=0}^{N+1} R_i Q_i^a. \tag{5.2}$$

It is well known [51,54,58] that the operator B defined by (5.2) is SPD, and

$$(B^{-1}v, v)_{0,a} = \inf_{\sum_{i=0}^{N+1} v_i = v} \sum_{i=0}^{N+1} (R_i^{-1} v_i, v_i)_{0,a}. \tag{5.3}$$

We have the following main result for BPX preconditioner.

Theorem 5.1 *Given a triangulation $\mathcal{T}_N = \mathcal{T}_0 + \mathcal{B}$ in $\mathbb{T}(\mathcal{T}_0)$, let $L = \max_{\tau \in \mathcal{T}_N} g(\tau)$. For the BPX preconditioner defined in* (5.2)*, we have*

$$\kappa(BA) \le C_1 c_d(L), \text{ and } \kappa_{m_0}(BA) \le C_0 L^2.$$

Consequently, we have the following convergence estimation of the BPX preconditioned conjugate gradient method:

$$\frac{\|u - u_k\|_A}{\|u - u_0\|_A} \le 2\left(C_1 c_d(L) - 1\right)^{m_0} \left(\frac{C_0 L - 1}{C_0 L + 1}\right)^{k - m_0}, \text{ for } k \ge m_0.$$

Proof First of all, let us estimate $\lambda_{\max}(BA)$. For any decomposition $v = \tilde{v} + \sum_{i=0}^{N} v_i, \tilde{v} \in \mathcal{V}, v_i \in \mathcal{V}_i$, we have

$$\begin{aligned} \|v\|_A^2 &\lesssim \|\tilde{v}\|_A^2 + \Big\| \sum_{i=0}^{N} v_i \Big\|_A^2 \\ &\le \|h^{-1}\tilde{v}\|_{0,a}^2 + \sum_{i=0}^{N} h_i^{-2} \|v_i\|_{0,a}^2 \\ &\le \sum_{i=0}^{N+1} (R_i^{-1} v_i, v_i)_{0,a}. \end{aligned}$$

In the second step, we used the inverse inequality and the inequality (4.9). In the third step, we used the assumption (5.1) of R_i . Taking infimum, we get

$$\|v\|_A^2 \lesssim \inf_{\sum_{i=0}^{N+1} v_i = v} \sum_{i=0}^{N+1} (R_i^{-1} v_i, v_i)_{0,a} = (B^{-1}v, v)_{0,a},$$

which implies that $\lambda_{\max}(BA) \lesssim 1$.

To estimate $\lambda_{\min}$, in view of (5.3) we choose the decomposition as in the stable decomposition Theorem 4.2 (see (4.5)) to conclude that

$$(B^{-1}v, v)_{0,a} \leq \sum_{i=0}^{N+1} (R_i^{-1}v_i, v_i)_{0,a} \lesssim c_d(L)(Av, v)_{0,a},$$

which implies that $\lambda_{\min}(BA) \gtrsim c_d(L)$. Therefore we have

$$\kappa(BA) \lesssim c_d(L).$$

On the other hand, if we apply (4.6) in the subspace $\widetilde{\mathcal{V}} \subset \mathcal{V}$, we obtain $\lambda_{m_0+1}(BA) \gtrsim L^2$ by the "min-max" Theorem 2.2. Hence we get an estimate of the effective condition number $\kappa_{m_0+1}(BA) \lesssim L^2$. The convergence rate estimate then follows by Theorem 2.1. This completes the proof.

□

From this convergence result, we can see that the convergence rate will deteriorate a little bit by $c_d(L)$ as L grows. But since m_0 is a fixed number, when k grows, the convergence rate will be controlled by the effective condition number, which is bounded uniformly with respect to the coefficient and logarithmically with respect to the meshsize. Notice that $L \eqsim |\log h_{\min}|$ and thus the asymptotic convergence rate of the PCG algorithm is $1 - \frac{1}{C|\log h_{\min}|}$ for $h < 1$.

Remark 5.1 The estimate $\kappa(BA) \leq C_1 c_d(L)$ is sharp in the sense that there exists an example on BPX preconditioner such that $\kappa(BA) \eqsim c_d(L)$; see [37]. □

Remark 5.2 Here we should emphasize that the convergence rate estimate in Theorem 5.1 holds for general substructures. In some special circumstance, for example "edge type" or "exceptional" in the terminology in [37], or "quasi-monotone" coefficient in [23], we can sharpen the convergence estimate in Theorem 5.1 by a modification of Theorem 4.2, see [37]. More precisely,

(i) For all the exceptional $\mathcal{T}_0$, which includes the quasi-monotone coefficients case, we have $\lambda_{\min}(BA) \geq C_0$ and thus $\kappa(BA) \leq C$ where C is a constant independent the coefficient a, meshsize h, and levels L. In this case, by the standard conjugate gradient theory, we get the convergence rate

$$\|u - u_k\|_A \leq 2\rho^k \|u - u_0\|_A$$

with $\rho = \frac{\sqrt{\kappa(BA)}-1}{\sqrt{\kappa(BA)}+1} < 1$. In this case, the convergence rate is bounded uniformly with respect to the coefficients and meshsize.

(ii) For $d = 2$ or $\mathcal{T}_0$ is of edge type in $d = 3$, we have $\lambda_{\min}(BA) \geq C_0^{-2}L^{-2}$. In this case, we obtain the convergence rate

$$\begin{aligned}\|u - u_k\|_A &\leq 2\left(\frac{\sqrt{\kappa(BA)}-1}{\sqrt{\kappa(BA)}+1}\right)^k \|u - u_0\|_A \\ &\leq 2\left(1 - \frac{1}{C_0 L}\right)^k \|u - u_0\|_A.\end{aligned}$$

For interested readers, we refer to [12,37,52,57]. So in fact, Theorem 5.1 provides an estimate of the convergence rate in the worst case. □

5.2 Multigrid (Multiplicative) Preconditioner

We shall use the following symmetric V-cycle multigrid as a preconditioner in PCG method and prove the efficiency of such a method. Let $A_i := A|_{\mathcal{V}_i}$. Then one step of the standard V-cycle multigrid $B : \mathcal{V} \to \mathcal{V}$ is recursively defined as follows:

Let $B_0 = A_0^{-1}$, for $i > 0$ and $g \in \mathcal{V}_i$, define $B_i g = w_3$.

(i) Presmoothing : $w_1 = R_i g$;
(ii) Correction: $w_2 = w_1 + B_{i-1}Q_{i-1}(g - A_i w_1)$;
(iii) Postsmoothing: $w_3 = w_2 + R_i^*(g - A_i w_2)$.

Set $B = B_{N+1}$.

For simplicity, we focus on the case of exact subspace solver, i.e., $R_i = A_i^{-1}$ for $i = 0, \cdots, N$ and for the finest level, R_{N+1} is chosen as Gauss-Seidal smoother, which can be also understood as the multiplicative method with exact local solvers applied to the nodal decomposition [53]. Let $P_p : \mathcal{V} \to \mathcal{V}_p$ and $P_i : \mathcal{V} \to \mathcal{V}_i$ be the orthogonal projection with respect to the inner product $(\cdot,\cdot)_a$. For our special choices of smoothers, we then have

$$\begin{aligned}I - R_{N+1}A &= \prod_{p\in\Lambda}(I - P_p), \\ I - B_N A &= \left(\prod_{i=0}^{N}(I - P_i)\right)^* \left(\prod_{i=0}^{N}(I - P_i)\right), \\ \|I - BA\|_A &= \left\|\prod_{i=0}^{N}(I - P_i)\prod_{p\in\Lambda}(I - P_p)\right\|_A^2\end{aligned}$$

For exact local solvers, we can apply the crucial X-Z identity [58] to conclude

$$\|I - BA\|_A = 1 - \frac{1}{1 + c_0}, \tag{5.4}$$

where

$$c_0 = \sup_{\|v\|_A=1} \inf_{v=\sum_{p\in\Lambda} v_p+\sum_{i=0}^N v_i} \left(\sum_{i=0}^N \Big\| P_i \sum_{j=i+1}^N v_j + P_i \sum_{p\in\Lambda} v_p \Big\|_A^2 + \sum_{p\in\Lambda} \big\| P_p \sum_{q>p} v_q \big\|_A^2 \right).$$

Theorem 5.2 *Given a triangulation $\mathcal{T}_N = \mathcal{T}_0 + \mathcal{B}$ in $\mathbb{T}(\mathcal{T}_0)$, let $L = \max_{\tau\in\mathcal{T}_N} g(\tau)$. For the multigrid V-cycle preconditioner B, we have*

$$\kappa(BA) \lesssim c_d(L), \quad \kappa_{m_0}(BA) \lesssim L^2.$$

Consequently, we have the following the convergence rate estimate of the BPX preconditioned conjugate gradient method:

$$\frac{\|u-u_k\|_A}{\|u-u_0\|_A} \le 2\left(C_1 c_d(L)-1\right)^{m_0} \left(\frac{C_0L-1}{C_0L+1}\right)^{k-m_0}, \ \textit{for } k \ge m_0.$$

Proof Since $I-BA$ is a non-expansion operator, we conclude $\lambda_{\max}(BA) \le 1$. Since $I-BA$ is SPD in the A-inner product and $\lambda_{\max}(BA) \le 1$, we have

$$\|I-BA\|_A = \max\{|1-\lambda_{\min}(BA)|, |1-\lambda_{\max}(BA)|\} = 1-\lambda_{\min}(BA).$$

To get an estimate on the minimum eigenvalue of BA, we only need to get a upper bound of the constant c_0 in (5.4).

To do so, for any $v \in \mathcal{V}$, we chose the decomposition in Theorem 4.2. That is,

$$v = \tilde{v} + \sum_{i=1}^N v_i, \text{ with } v_0 = \mathcal{I}_0^a v, \quad v_i = (\mathcal{I}_i^a - \mathcal{I}_{i-1}^a)v,$$

where $\tilde{v} = v - \mathcal{I}_N^a v = \sum_{p\in\Lambda} v_p$. Then by shape regularity of the triangulation, we have

$$c_0 \lesssim \sum_{i=0}^N \Big\| P_i \sum_{j=i+1}^N v_j \Big\|_A^2 + \sum_{i=0}^N \|P_i\tilde{v}\|_A^2 + \sum_{p\in\Lambda} \Big\| P_p \sum_{q>p} v_q \Big\|_A^2.$$

We estimate these three terms as follows. For the last term, by the finite overlapping of nodal bases, we have

$$\begin{aligned} &\sum_{p\in\Lambda} \big\| P_p \sum_{q>p} v_q \big\|_A^2 \lesssim \sum_{p\in\Lambda} \big\| \sum_{q>p} v_q \big\|_{A,\omega_p}^2 \lesssim \sum_{p\in\Lambda} \|v_p\|_{A,\omega_p}^2 \\ &\lesssim \sum_{p\in\Lambda} h_p^{-2} \|v_p\|_{0,a,\omega_p}^2 \lesssim \|h^{-1}(v-\mathcal{I}_N^a v)\|_{0,a}^2 \lesssim \|v\|_A^2. \end{aligned}$$

For the middle term, we regroup by generations and use (4.2) to get

$$\begin{aligned}\sum_{i=0}^{N}\left\|P_i\tilde{v}\right\|_A^2 &= \sum_{k=0}^{L}\sum_{l,g_l=k}\left\|P_l\tilde{v}\right\|_A^2\\ &\leq \sum_{k=0}^{L}\sum_{l,g_l=k}\|\tilde{v}\|_{A,\tilde{\omega}_l}^2\\ &\lesssim \sum_{k=0}^{L}\|\tilde{v}\|_A^2 = L\|\tilde{v}\|_A^2.\end{aligned}$$

For the first term, we define $u_i = P_i\left(\sum_{j=i+1}^{N} v_j\right)$ and $u_0 := P_0(v-v_0)$ and apply the strengthened Cauchy Schwarz inequality, cf. Lemma 4.1 to get

$$\begin{aligned}\sum_{i=0}^{N}\left\|P_i\sum_{j=i+1}^{N}v_j\right\|_A^2 &= \sum_{i=0}^{N}\sum_{j=i+1}^{N}A(u_i,v_j)\\ &\lesssim \|v-v_0\|_A^2 + \sum_{i=1}^{N}h_i^{-2}\|v_i\|_{0,a}^2\\ &\lesssim c_d(L)\|v\|_A^2.\end{aligned}$$

Here the constant $c_d(L)$ can be improved to L^2 if we consider the decomposition (4.6) of $v \in \widetilde{\mathcal{V}}$. Combined with the Mini-Max Theorem 2.2, yields $\lambda_{\min}(BA) \gtrsim c_d(L), \lambda_{m_0+1}(BA) \gtrsim L^{-2}$, and thus

$$\kappa(BA) \lesssim c_d(L), \kappa_{m_0+1}(BA) \lesssim L^2.$$

Finally, the convergence rate of the PCG method follows by Theorem 2.1.

□

Follow the same proof as Theorem 5.2, we can also obtain the following convergence result for the local multigrid V-cycle solver.

Corollary 5.1 *For the multigrid V-cycle algorithm defined above on bisection grids, we have*

$$\|E\|_A = \|I-BA\|_A = 1 - \frac{1}{1+c_0},$$

where $c_0 \lesssim c_d(L)$.

This corollary implies that multigrid alone is not robust, especially in 3D. In this case, the convergence rate of multigrid will be proportional to $1 - 2^{-L} \simeq 1 - h_{\min}^{-1}$, which deteriorates rapidly as the meshsize become small.

Remark 5.2 is also applicable here, i.e., all the above estimates are estimates for the worst case. For the special circumstances mentioned in Remark 5.2, the estimates can be improved in the same way.

6 Numerical Experiments

In this section, we present some numerical experiments to support the theoretical results in previous sections.

In the implementation of the adaptive loop, we use a modification of the error indicator presented in [38]. Some other a posteriori error indicators for jump coefficients problem (1.1) can be found in [9,21,46,14]. The adaptive algorithm using different error indicators will generate different grids. However, we should emphasize that the robustness of the local adaptive multilevel preconditioners is independent of how the grids are generated in the refinement procedure. Our theoretical results are applicable to all of these cases.

The implementation of the BPX preconditioner and the multigrid methods are standard, and can be found in, for example, [13,55]. The implementation of PCG algorithm can be found in [24,30,39]. All numerical examples are implemented by using iFEM [18]. We only present three-dimensional examples here and refer to [20] for two-dimensional ones. In the PCG algorithm, we use the stopping criterion

$$\frac{\|u^k - u^{k-1}\|_A}{\|u^k\|_A} \leq 10^{-10}.$$

In the implementation of the local multilevel preconditioners, we use an algorithm for coarsening bisection grids introduced by Chen and Zhang; see [20] for two dimensional case and [18] for three dimensional one. The coarsening algorithm will find all compatible bisections and regroup them, with possibly different generations, into groups $\cup_{l=1}^{L'} G(l) = \{1, 2, \cdots, N\}$ such that for any $i, j \in G(l), \omega_j \cap \omega_i = \varnothing$. Each coarsening step is corresponding to a *level* in the multilevel terminology, and the total number of levels is L'. There are two major benefits of using this coarsening algorithm.

(i) We do not need to store the complex bisection tree structure of the refinement procedure explicitly in the algorithm. Instead, we only need the grid information on the finest level and the coarsening subroutine will restore multilevel structure.

(ii) Our numerical evidence shows that the number of nodes will decrease around one half in one coarsening step. Therefore the constant L' is much smaller than the maximal generation $L \eqsim |\log h_{\min}|$; see Remark 4.2.

In what follows, we will use some shorthand notation for the different algorithms implemented.

- TPSMG stands for the V-cycle multigrid with *Three-Point Smoothing* (TPS), which only performs smoothing on new vertices and their two direct neighbors sharing the same edge.
- TPSMGCG is the PCG algorithm using the TPSMG as preconditioner.
- TPSBPXCG is the additive version of TPSMG preconditioner.

Among all these algorithms, the main focus of this paper is the behavior of TPSMGCG and TPSBPXCG. In the numerical experiments below, we also report some results for TPSMG for comparison.

6.1 Example 1: Subdomains Touching Dirichlet Boundary

As the first example, we consider an example from [38]. In particular, we consider the domain $\Omega = (-1,1)^3$ with the subdomains $\Omega_1 = (0,1)^3$ and $\Omega_2 = \Omega \setminus \bar{\Omega}_1$. We set the coefficients $a_1 = 1$ and $a_2 = \epsilon$. We choose $f = 0$ and impose Dirichlet conditions

$$u_{\{-1\}\times[-1,1]\times[-1,1]} = 0, \quad u_{\{1\}\times[-1,1]\times[-1,1]} = 1,$$

and homogenous Neumann boundary conditions on the remaining boundary. Note that both Ω_1 and Ω_2 have a nontrivial portion of the Dirichlet boundary.

For this problem, singularities occur along three interior edges of Ω_1. Figure 6.1 shows an adaptive mesh and corresponding finite element approximation after several local refinements. To view the mesh around the singularity, we cut the part $\{(x,y,z) : x > 0, y > 0\}$ of the domain Ω. As we can see from Figure 6.1, the adaptive algorithm captures the singularity quite well.

To test the robustness of the preconditioners with respect to the jump of coefficients, we vary $\epsilon = 10^{-4}, 10^{-2}, 10^{2}, 10^{4}$. Figure 6.2 shows the condition number of the TPSBPXCG and TPSMGCG with respect the degree of freedoms (DOFs) and the varies coefficients. As we can see from the figure, the condition numbers are robust with respect the choice of ϵ. Also, the range of the condition numbers are quite robust with respect to the DOFs ($\kappa(BA) \in (1,23)$ for TPSBPXCG and $\kappa(BA) \in (1,6.5)$ for TPSMGCG).

Figure 6.3 shows the eigenvalue distributions for the TPSMGCG and TPSBPXCG preconditioned systems. As we can see from the figure, there

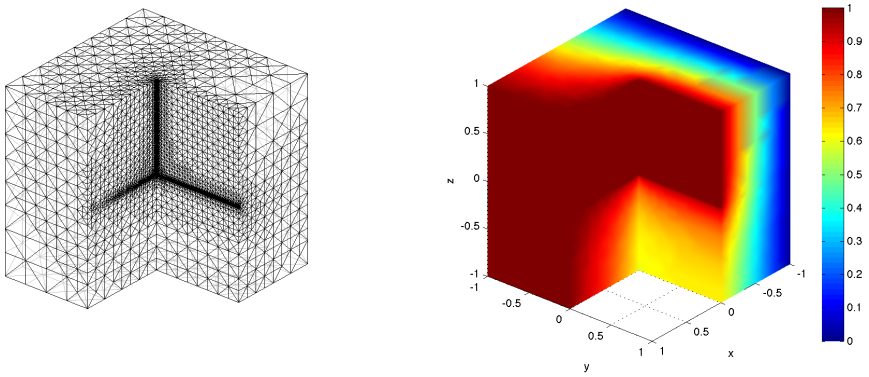
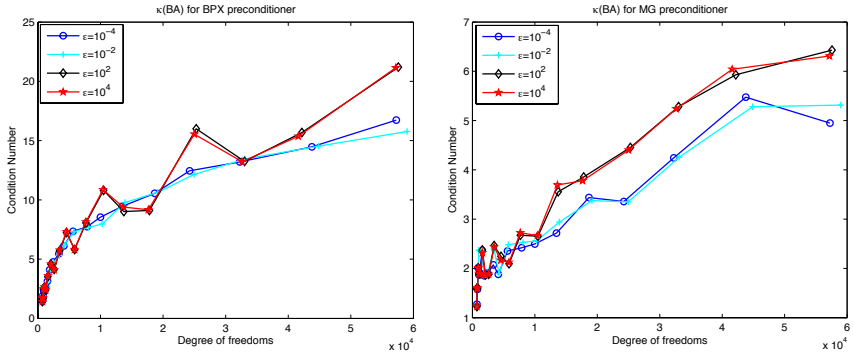


Fig. 6.1. An adaptive mesh and finite element solution for Example 1 with $\epsilon = 10^{-4}$ and 32236 vertices.



(a) TPSBPXCG (b) TPSMGCG

Fig. 6.2. Example 1: The Estimated Condition Numbers $\kappa(BA)$

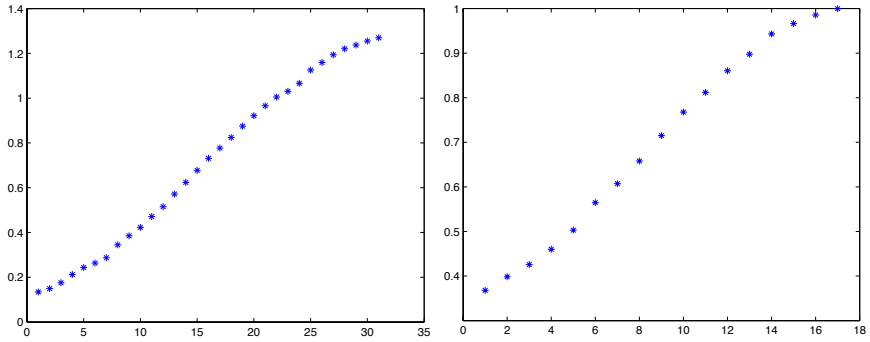
is no obvious "bad" eigenvalue. This agrees with the theoretical results, because both Ω_1 and Ω_2 has nontrivial portion of Dirichlet boundary, i.e., the number of floating domains $m_0 \equiv 0$.

6.2 Example 2: Floating Subdomains with Cross Point

For the second example, inspired by [37,52,57], we consider solving the model equation (1.1) in the cubic domain $\Omega = (-1,1)^3$. Let the coefficient $a(x)$ be the constants $a_1 = a_2 = 1$ and $a_3 = \epsilon$ on the three regions Ω_1, Ω_2 and Ω_3 respectively (see Figure 6.4), where

$$\Omega_1 = (-0.5,0)^3, \Omega_2 = (0,0.5)^3 \text{ and } \Omega_3 = \Omega \setminus (\overline{\Omega}_1 \cup \overline{\Omega}_2).$$

We choose $f = 1$ and impose the same boundary conditions as in Example



(a) Eigenvalues for TPSBPXCG (b) Eigenvalues for TPSMGCG

Fig. 6.3. Example 1: Eigenvalues of BA when $\epsilon = 10^{-4}$ with 13459 vertices

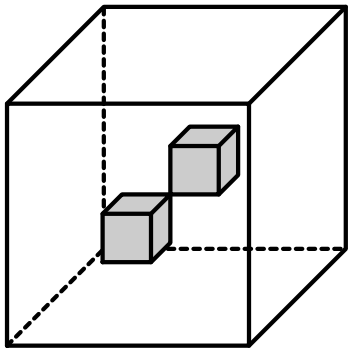


Fig. 6.4. *Example 2: the coefficients* $a_1 = a_2 = 1$ *in the gray domains* Ω_1 *and* Ω_2*, and* $a_3 = \epsilon$ *in the rest of the domain.*

1: Dirichlet conditions

$$u_{\{-1\}\times[-1,1]\times[-1,1]} = 0, \quad u_{\{1\}\times[-1,1]\times[-1,1]} = 1,$$

and homogenous Neumann boundary conditions on the remaining boundary.

For this problem, singularities occur along edges of Ω_1 and Ω_2. Figure 6.5 shows an adaptive mesh and the corresponding finite element approximation after several iterations of the adaptive algorithm. To view the mesh around the singularity, we only show half of the domain Ω.

For comparison, we also present the number of iterations for TPSMG algorithms (reducing the relative error in energy norm to the tolerance 10^{-10}) in Tables 6.1. This table shows that the TPSMG algorithm itself will deteriorate quickly when ϵ is small. On the other hand, if ϵ is large, the standard MG algorithm will converge uniformly. This is because the coefficient in

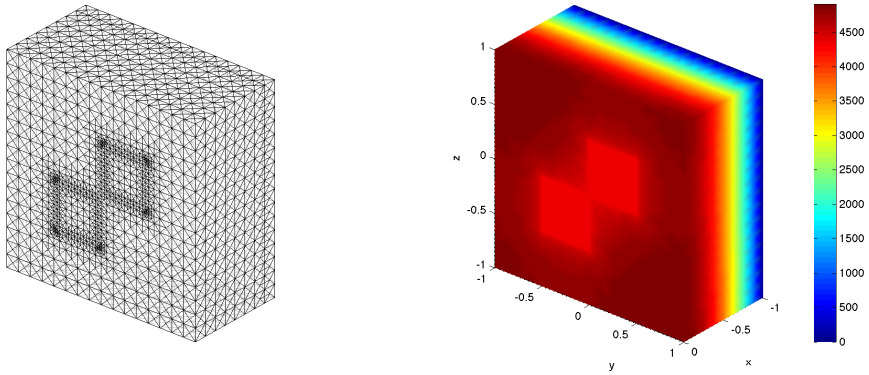


Fig. 6.5. An adaptive mesh and finite element approximation for Example 2 with $\epsilon = 10^{-4}$ and 36466 vertices.

Ω_3, which contains the Dirichlet boundary, is dominant. In this case, we could use the standard multigrid analysis (as in [53]) to show the robustness of the preconditioners.

$\epsilon = 10^{-4}$	
4913	41
5505	62
6617	89
8666	99
10585	98
12411	125
16353	154
21248	182
27755	197
36466	178
43271	238
51163	283
72349	395
89146	424
104747	413

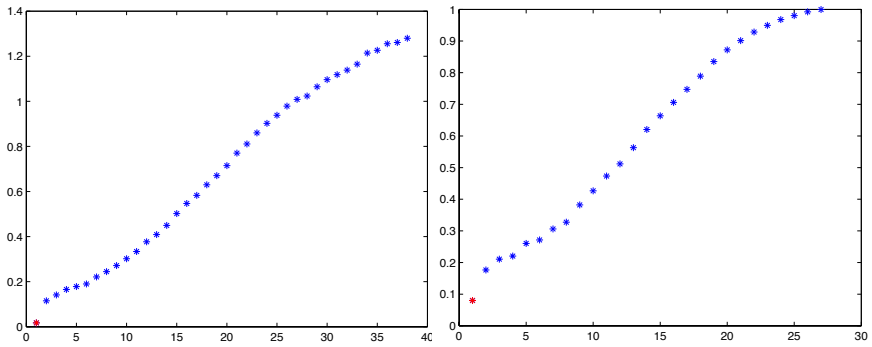
$\epsilon = 10^{-2}$	
4913	46
5550	51
6743	61
8907	65
10729	66
13281	86
17146	90
23139	90
28613	160
37338	175
43610	149
52715	154
72967	238
89320	165
113131	294

$\epsilon = 10^{2}$	
4913	16
5279	37
5867	43
6522	48
7562	68
9493	61
11858	49
15257	68
20649	61
27946	49
36735	52
48890	58
68297	71
89872	55
119109	61

$\epsilon = 10^{4}$	
4913	16
5269	37
5863	42
6493	45
7531	68
9419	59
11721	46
14941	69
20065	59
27199	47
35601	59
47743	55
66989	71
88079	57
116739	56

Table 6.1. Number of iterations of TPSMG for Example 2.

Figure 6.6 shows the eigenvalue distributions for the TPSMGCG and TPSBPXCG preconditioned systems. As we can see from the figure, there is one small eigenvalue for both preconditioned systems. This agrees with the theoretical results, the number of small eigenvalues is bounded by the floating subdomains $m_0 \equiv 2$.



(a) Eigenvalues for TPSBPXCG (b) Eigenvalues for TPSMGCG

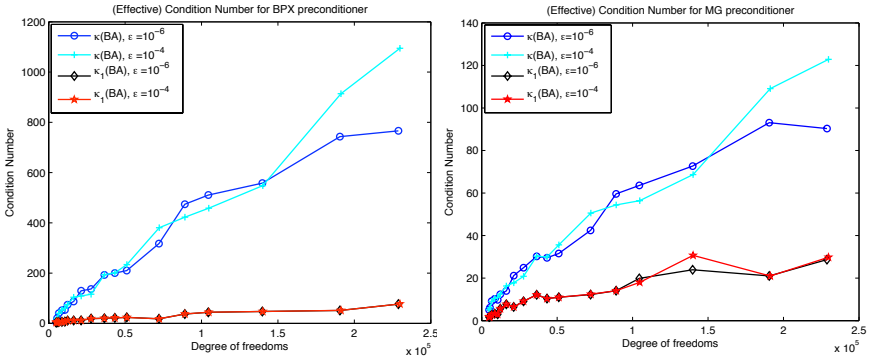
Fig. 6.6. Example 2: Eigenvalues of BA when $\epsilon = 10^{-4}$ with 12411 vertices

Figure 6.7 shows the condition number and effective condition number of TPSBPXCG and TPSMGCG preconditioned systems. From Figure 6.7, we observed that when ϵ is small, the condition number deteriorates ($\kappa(BA) \in [3, 1100]$ for TPSBPXCG, and $\kappa(BA) \in [3, 125]$ for TPSMGCG) as we can see from the figure). Comparing with the first example, the condition number is more sensitive to the coefficient in this case.

On the other hand, if we get rid of the first small eigenvalue, the effective condition number $\kappa_1(BA)$ (the black-diamond and red-star lines) of TPSBPXCG and TPSMGCG preconditioned systems are almost identical for different ϵ. This indicates that the effective condition numbers are uniform with respect to the jumps. Moreover, as we can see from Figure 6.7, $\kappa_1(BA)$ are mildly increasing with respect to the DOFs ($\kappa_1(BA) \in [1, 80]$ for TPSBPXCG, and $\kappa_1(BA) \in [1, 30]$ for TPSMGCG). These results agree with our theoretical expectations from Section 5.

7 Conclusion

In this paper, we designed local multilevel preconditioners based on the decomposition of the finite element space into 3-point subspaces for the highly graded mesh obtained from adaptive bisection algorithms. To analyze the behavior of the local multilevel preconditioners, we introduced a local interpolation operator and proved some approximation and stability properties of it. Based on these properties, we showed the decomposition of the finite element space is stable , which is a key ingredient in the multilevel analysis. This enabled us to analyze the eigenvalue distributions of the preconditioned systems. In particularly, we showed that there are only a small fixed number of eigenvalues are deteriorated by the coefficients and



(a) TPSBPXCG (b) TPSMGCG

Fig. 6.7. Example 2: $\kappa(BA)$ and $\kappa_1(BA)$ for the cases $\epsilon = 10^{-6}, 10^{-4}$ w.r.t the DOFs.

meshsize, and the other eigenvalues are uniformly bounded with respect to the coefficients and logarithmically depends on the meshsize. As a result, we proved the asymptotic convergence rate of the PCG algorithm is uniform with respect to the coefficient and nearly uniform with respect to the meshsize. Moreover, the overall computation complexity of these multilevel preconditioner are nearly optimal. Numerical experiments justified our theoretical results.

Acknowledgement

The first author is supported in part by NSF Grant DMS-0811272, and in part by NIH Grant P50GM76516 and R01GM75309. This work is also partially supported by the Beijing International Center for Mathematical Research. The second and fourth authors were supported in part by NSF Awards 0715146 and 0915220, and DTRA Award HDTRA-09-1-0036. The third author was supported in part by NSF DMS-0609727, NSFC-10528102 and Alexander von Humboldt Research Award for Senior US Scientists.

References

1. B. Aksoylu, I. Graham, H. Klie, and R. Scheichl. Towards a rigorously justified algebraic preconditioner for high-contrast diffusion problems. *Computing and Visualization in Science*, 11(4):319–331, 2008.
2. B. Aksoylu and M. Holst. Optimality of multilevel preconditioners for local mesh refinement in three dimensions. *SIAM Journal on Numerical Analysis*, 44(3):1005–1025, 2006.
3. R. E. Alcouffe, A. Brandt, J. E. Dendy, and J. W. Painter. The multi–grid methods for the diffusion equation with strongly discontinuous coefficients. *SIAM Journal on Scientific and Statistical Computing*, 2:430–454, 1981.

4. O. Axelsson. *Iterative solution methods*. Cambridge University Press, Cambridge, 1994.
5. O. Axelsson. Iteration number for the conjugate gradient method. *Mathematics and Computers in Simulation*, 61(3-6):421–435, 2003. MODELLING 2001 (Pilsen).
6. D. Bai and A. Brandt. Local mesh refinement multilevel techniques. *SIAM Journal on Scientific and Statistical Computing*, 8:109–134, 1987.
7. R. E. Bank. Hierarchical bases and the finite element method. *Acta Numerica*, 5:1–43, 1996.
8. R. E. Bank, T. Dupont, and H. Yserentant. The hierarchical basis multigrid method. *Numerische Mathematik*, 52:427–458, 1988.
9. C. Bernardi and R. Verfürth. Adaptive finite element methods for elliptic equations with non-smooth coefficients. *Numerische Mathematik*, 85(4):579–608, 2000.
10. P. Binev, W. Dahmen, and R. DeVore. Adaptive finite element methods with convergence rates. *Numerische Mathematik*, 97(2):219–268, 2004.
11. J. H. Bramble, J. E. Pasciak, and J. Xu. Parallel multilevel preconditioners. *Mathematics of Computation*, 55(191):1–22, 1990.
12. J. H. Bramble and J. Xu. Some estimates for a weighted L^2 projection. *Mathematics of Computation*, 56:463–476, 1991.
13. W. L. Briggs, V. E. Henson, and S. F. McCormick. *A Multigrid Tutorial*. SIAM Books, Philadelphia, 2000. Second edition.
14. Z. Cai and S. Zhang. Recovery-based error estimator for interface problems: Conforming linear elements. *SIAM Journal on Numerical Analysis*, 47(3):2132–2156, 2009.
15. J. M. Cascon, C. Kreuzer, R. H. Nochetto, and K. G. Siebert. Quasi-optimal convergence rate for an adaptive finite element method. *SIAM Journal on Numerical Analysis*, 46(5):2524–2550, 2008.
16. T. F. Chan and W. L. Wan. Robust multigrid methods for nonsmooth coefficient elliptic linear systems. *Journal of Computational and Applied Mathematics*, 123(1-2):323–352, 2000.
17. L. Chen. Short implementation of bisection in MATLAB. In P. Jorgensen, X. Shen, C.-W. Shu, and N. Yan, editors, *Recent Advances in Computational Sciences – Selected Papers from the International Workship on Computational Sciences and Its Education*, pages 318 –332. World Scientific Pub Co Inc, 2007.
18. L. Chen. iFEM: an integrate finite element methods package in MATLAB. Technical report, University of California at Irvine, 2009.
19. L. Chen, R. H. Nochetto, and J. Xu. Multilevel methods on graded bisection grids I: H^1 system. *Preprint*, 2007.
20. L. Chen and C.-S. Zhang. A coarsening algorithm and multilevel methods on adaptive grids by newest vertex bisection. *To appear in J. Comp. Math.*, 2009.
21. Z. Chen and S. Dai. On the efficiency of adaptive finite element methods for elliptic problems with discontinuous coefficients. *SIAM Journal on Scientific Computing*, 24(2):443–462, 2002.
22. R. K. Coomer and I. G. Graham. Massively parallel methods for semiconductor device modelling. *Computing*, 56(1):1–27, 1996.
23. M. Dryja, M. V. Sarkis, and O. B. Widlund. Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions. *Numerische Mathematik*, 72(3):313–348, 1996.
24. G. H. Golub and C. F. Van Loan. *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, third edition, 1996.
25. I. Graham, P. Lechner, and R. Scheichl. Domain decomposition for multiscale pdes. *Numerische Mathematik*, 106(4):589–626, June 2007.

26. I. G. Graham and M. J. Hagger. Unstructured additive schwarz-conjugate gradient method for elliptic problems with highly discontinuous coefficients. *SIAM Journal on Scientific Computing*, 20:2041–2066, 1999.
27. B. Heise and M. Kuhn. Parallel solvers for linear and nonlinear exterior magnetic field problems based upon coupled FE/BE formulations. *Computing*, 56(3):237–258, 1996. International GAMM-Workshop on Multi-level Methods (Meisdorf, 1994).
28. R. Hiptmair and W. Zheng. Local Multigrid in H (curl). *Journal of Computational Mathematics*, 27(5):573–603, 2009.
29. C. E. Kees, C. T. Miller, E. W. Jenkins, and C. T. Kelley. Versatile two-level Schwarz preconditioners for multiphase flow. *Comput. Geosci.*, 7(2):91–114, 2003.
30. C. T. Kelley. *Iterative methods for linear and nonlinear equations*, volume 16 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1995. With separately available software.
31. I. Kossaczký. A recursive approach to local mesh refinement in two and three dimensions. *Journal of Computational and Applied Mathematics*, 55:275–288, 1994.
32. J. Meza and R. Tuminaro. A Multigrid Preconditioner for the Semiconductor Equations. *SIAM Journal on Scientific Computing*, 17:118–132, 1996.
33. W. F. Mitchell. A comparison of adaptive refinement techniques for elliptic problems. *ACM Transactions on Mathematical Software (TOMS) archive*, 15(4):326 – 347, 1989.
34. W. F. Mitchell. Optimal multilevel iterative methods for adaptive grids. *SIAM Journal on Scientific and Statistical Computing*, 13:146–167, 1992.
35. R. Nochetto, K. Siebert, and A. Veeser. Theory of adaptive finite element methods: An introduction. In R. DeVore and A. Kunoth, editors, *Multiscale, Nonlinear and Adaptive Approximation*, pages 409–542. Springer, 2009. Dedicated to Wolfgang Dahmen on the Occasion of His 60th Birthday.
36. P. Oswald. *Multilevel Finite Element Approximation, Theory and Applications*. Teubner Skripten zur Numerik. Teubner Verlag, Stuttgart, 1994.
37. P. Oswald. On the robustness of the BPX-preconditioner with respect to jumps in the coefficients. *Mathematics of Computation*, 68:633–650, 1999.
38. M. Petzoldt. A posteriori error estimators for elliptic equations with discontinuous coefficients. *Advances in Computational Mathematics*, 16(1):47–75, 2002.
39. Y. Saad. *Iterative methods for sparse linear systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, second edition, 2003.
40. M. Sarkis. Nonstandard coarse spaces and schwarz methods for elliptic problems with discontinuous coefficients using non-conforming elements. *Numerische Mathematik*, 77(3):383–406, 1997.
41. R. Scheichl and E. Vainikko. Additive schwarz with aggregation-based coarsening for elliptic problems with highly variable coefficients. *Computing*, 80(4):319–343, Sept. 2007.
42. R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Mathematics of Computation*, 54:483–493, 1990.
43. R. Stevenson. Stable three-point wavelet bases on general meshes. *Numerische Mathematik*, 80(1):131–158, 1998.
44. R. Stevenson. The completion of locally refined simplicial partitions created by bisection. *Mathemathics of Computation*, 77:227–241, 2008.
45. P. S. Vassilevski. *Multilevel block factorization preconditioners*. Springer, New York, 2008.
46. M. Vohralık. Guaranteed and fully robust a posteriori error estimates for conforming discretizations of diffusion problems with discontinuous coefficients. Technical Report Preprint R08009, Laboratoire Jacques-Louis Lions, 2008.

47. C. Vuik, A. Segal, and J. A. Meijerink. An efficient preconditioned cg method for the solution of a class of layered problems with extreme contrasts in the coefficients. *Journal of Computational Physics*, 152(1):385–403, June 1999.
48. J. Wang. New convergence estimates for multilevel algorithms for finite-element approximations. *Journal of Computational and Applied Mathematics*, 50:593–604, 1994.
49. J. Wang and R. Xie. Domain decomposition for elliptic problems with large jumps in coefficients. In *the Proceedings of Conference on Scientific and Engineering Computing*, pages 74–86. National Defense Industry Press, 1994.
50. Z. Wang, C. Wang, and K. Chen. Two-phase flow and transport in the air cathode of proton exchange membrane fuel cells. *J.Power Sources*, 94:40–50, 2001.
51. O. B. Widlund. Some Schwarz methods for symmetric and nonsymmetric elliptic problems. In D. E. Keyes, T. F. Chan, G. A. Meurant, J. S. Scroggs, and R. G. Voigt, editors, *Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 19–36, Philadelphia, 1992. SIAM.
52. J. Xu. Counter examples concerning a weighted L^2 projection. *Mathematics of Computation*, 57:563–568, 1991.
53. J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Review*, 34:581–613, 1992.
54. J. Xu. A new class of iterative methods for nonselfadjoint or indefinite problems. *SIAM Journal on Numerical Analysis*, 29:303–319, 1992.
55. J. Xu. An introduction to multigrid convergence theory. In R. Chan, T. Chan, and G. Golub, editors, *Iterative Methods in Scientific Computing*. Springer-Verlag, 1997.
56. J. Xu, L. Chen, and R. Nochetto. Optimal multilevel methods for H (grad), H (curl), and H (div) systems on graded and unstructured grids. In *Multiscale, Nonlinear and Adaptive Approximation*, pages 599–659. Springer, 2009.
57. J. Xu and Y. Zhu. Uniform convergent multigrid methods for elliptic problems with strongly discontinuous coefficients. *Mathematical Models and Methods in Applied Science*, 18(1):77 –105, 2008.
58. J. Xu and L. Zikatanov. The method of alternating projections and the method of subspace corrections in Hilbert space. *Journal of The American Mathematical Society*, 15:573–597, 2002.
59. H. Yserentant. Two preconditioners based on the multi-level splitting of finite element spaces. *Numerische Mathematik*, 58:163–184, 1990.
60. H. Yserentant. Old and new convergence proofs for multigrid methods. *Acta Numerica*, pages 285–326, 1993.
61. Y. Zhu. Domain decomposition preconditioners for elliptic equations with jump coefficients. *Numerical Linear Algebra with Applications*, 15(2-3):271–289, 2008.