Journal of Computational Mathematics
Vol.30, No.3, 2012, 223–248.

http://www.global-sci.org/jcm
doi:10.4208/jcm.1109-m3401

# LOCAL MULTILEVEL METHODS FOR SECOND-ORDER ELLIPTIC PROBLEMS WITH HIGHLY DISCONTINUOUS COEFFICIENTS*

Huangxin Chen,
*School of Mathematical Sciences, Xiamen University, Xiamen, 361005, China*
*LSEC, ICMSEC, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, P.O. Box 2719, Beijing, 100190, China*
*Email: chx@xmu.edu.cn*
Xuejun Xu and Weiying Zheng
*LSEC, ICMSEC, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, P.O. Box 2719, Beijing 100190, China*
*Email: xxj@lsec.cc.ac.cn, zwy@lsec.cc.ac.cn*

**Abstract**

In this paper, local multiplicative and additive multilevel methods on adaptively refined meshes are considered for second-order elliptic problems with highly discontinuous coefficients. For the multilevel-preconditioned system, we study the distribution of its spectrum by using the abstract Schwarz theory. It is proved that, except for a few small eigenvalues, the spectrum of the preconditioned system is bounded quasi-uniformly with respect to the jumps of the coefficient and the mesh sizes. The convergence rate of multilevel-preconditioned conjugate gradient methods is shown to be quasi-optimal regarding the jumps and the meshes. Numerical experiments are presented to illustrate the theoretical findings.

*Mathematics subject classification:* 65F10, 65N30.
*Key words:* Local multilevel method, Adaptive finite element method, Preconditioned conjugate gradient method, Discontinuous coefficients.

## 1. Introduction

During the last two decades, adaptive finite element methods (AFEM) have been developed very rapidly and have become a popular and powerful tool in numerical solution of partial differential equations (PDEs). Quasi-optimal approximation results can be achieved by mesh adaptivity based on a posteriori error estimates (see, e.g., [6, 16, 32, 36]). In this paper, we also pursue asymptotically optimal methods for computing the solution of the discrete problem. By "optimal" we mean that the computation of the solution asymptotically only requires $O(N)$ operations where $N$ is the number of degrees of freedom (DOFs) on the underlying mesh. Multigrid or multilevel methods are among the most efficient and widely used methods for computing the approximate solution.

The uniform convergence of multigrid methods for conforming finite elements has been widely studied by many authors. We refer to [7–10, 12, 25, 33, 43] for a multigrid convergence theory on uniformly refined meshes. Since in AFEM the number of DOFs may not grow exponentially with the mesh levels, as Mitchell pointed out in [31], traditional multigrid methods,

* Received April 28, 2010 / Revised version received July 10, 2011 / Accepted September 1, 2011 / Published online May 7, 2012 /

which perform relaxations on all nodes, may require $O(N^2)$ operations for certain meshes. In order to overcome this issue, local multigrid methods adopt the idea of local smoothing, which restricts relaxations to new elements of each level. Local smoothing turns out to be very efficient on adaptively refined meshes (see, e.g., [26, 46, 48, 50] for elliptic problems with smooth coefficients). Motivated by the recent work of Xu and Zhu [49], we study local multiplicative and additive multilevel algorithms (LMMA and LMAA) for second-order elliptic problems with highly discontinuous coefficients. Different from the works of Chen, Holst, Xu and Zhu [18] for second-order elliptic problems with discontinuous coefficients and Hiptmair and Zheng [27] for Maxwell equations, our algorithm does not reconstruct a virtual refinement hierarchy of meshes. We assume that the meshes are generated by using AFEM based on a posteriori error estimates.

Given a bounded, polygonal or polyhedral domain $\Omega \subset R^d\,(d = 2, 3)$, we consider the following second-order elliptic problem:

$$-\mathrm{div}(\rho(\boldsymbol{x})\nabla u) = f \quad \text{in } \Omega, \tag{1.1}$$

$$u = 0 \quad \text{on } \partial\Omega, \tag{1.2}$$

where the source function $f \in L^2(\Omega)$. The coefficient $\rho$ is positive and piecewise constant and may have large jumps in $\Omega$. The homogeneous boundary condition in (1.2) is not essential to our theory and can be replaced with more general boundary conditions. Although problem (1.1)–(1.2) seems to be simple, it plays an important role in many practical applications: such as steady state heat conduction in composite materials, electromagnetism, and multiphase flow.

It is well known that the solution of problem (1.1)–(1.2) may have singularities near reentrant corners of the domain and jumps of the coefficient. The AFEM based on a posteriori error estimates is very efficient to capture local singularities of the solution. A considerable amount of work has been devoted to a posteriori error estimates for such problems. We refer to Bernardi and Verfürth [5], Petzoldt [35], and Chen and Dai [20] for residual-based error estimates, to Luce and Wohlmuth [29] for equilibrated error estimates, and to Cai and Zhang [14] for recovery-based error estimates. For adaptive nonconforming or mixed finite element methods, a posteriori error estimates have been studied by Ainsworth [1, 2] for equilibrated error estimates, by Chen, Xu, and Hoppe [19] for residual-based error estimates, and by Cai and Zhang [15] for recovery-based error estimates.

The purpose of this paper is to study local multilevel solvers for the adaptive finite element discretization of (1.1)–(1.2) and to prove the quasi-optimality of these solvers. It is known that the condition number of the discrete system of the problem (1.1)–(1.2) depends on the jumps of $\rho$ and on the mesh sizes. To reduce the condition number, multigrid methods and domain decomposition methods have been studied for quasi-uniform meshes (see, e.g., [17, 24, 30, 37, 40, 44]). In general, the convergence rate of local multilevel methods depends on the jump of the coefficient, the mesh sizes, or the mesh levels due to the lack of uniform stability estimates for the weighted $L^2$-projection (see, e.g., [11, 34, 42]). The convergence rate can be improved for some specific scenarios (see, e.g., [22, 23, 34, 45]). Recently, Xu and Zhu (see, e.g., [49, 51]) have proved quasi-uniform convergence of conjugate gradient methods preconditioned by multilevel methods and overlapping domain decomposition methods, respectively.

The objective of this paper is to extend the results of [49] to adaptively refined meshes which are generated by the "newest vertex bisection algorithm" [31, 46]. Using the abstract Schwarz theory, we prove that except for a few small eigenvalues, the *effective condition numbers*, i.e., the ratio of the maximum to the minimum of the remaining eigenvalues of the multilevel-preconditioned algebraic system, are bounded by $C|\log h_{\min}|^2$. Here the constant $C$ is independent of the jumps, the mesh sizes, and the mesh levels, and $h_{\min}$ is the minimum diameter of the triangles or tetrahedrons on the finest mesh. The main difficulty is how to obtain a stable multilevel decomposition of the finite element space on the finest mesh and how to prove the strengthened Cauchy-Schwarz inequality regarding this decomposition. We should point out that both local Jacobi smoother and local Gauss-Seidel smoother apply to the local multilevel methods.

The remainder of this paper is organized as follows. In Section 2, we introduce some notation, finite element spaces, and the preconditioned conjugate gradient method. In Section 3, we propose the local multiplicative and additive multilevel algorithms, i.e., local multigrid V-cycle and the local BPX preconditioner. In Section 4, we study the convergence of LMMA, the preconditioned conjugate gradient method by LMMA (LMMA-PCG), and the preconditioned conjugate gradient method by LMAA (LMAA-PCG). In Section 5, we study the multilevel decomposition of the finite element space on the finest mesh and prove the so-called strengthened Cauchy-Schwarz inequality. In Section 6, we present several numerical experiments to demonstrate our convergence theory.

## 2. Preliminaries

Throughout this paper, we denote by $(\cdot,\cdot)$ the standard inner product in $L^2(\Omega)$, by $\|\cdot\|_{1,\Omega}$ and $|\cdot|_{1,\Omega}$ the norm and semi-norm in $H^1(\Omega)$. Let $C$ with or without subscript stand for a generic positive constant which is independent of the jumps of $\rho$, the mesh sizes and the mesh levels, but depends on $\Omega$ and the shape regularity of the meshes. These constants can take on different values in different occurrences. We also introduce the weighted inner product and weighted norm in $L^2(\Omega)$:

$$(u,v)_\rho = (\rho u, v), \qquad \|v\|_{L^2_\rho(\Omega)} = (v,v)_\rho^{\frac{1}{2}} \qquad \forall\, u,v \in L^2(\Omega).$$

The weak formulation of (1.1) and (1.2) is: Find $u \in H_0^1(\Omega)$ such that

$$a(u,v) = (f,v) \qquad \forall v \in H_0^1(\Omega), \tag{2.1}$$

where $a : H_0^1(\Omega) \times H_0^1(\Omega) \mapsto \mathbb{R}^1$ is a bilinear form defined as follows

$$a(u,v) = (\rho(\boldsymbol{x})\nabla u, \nabla v) \qquad \forall u,v \in H_0^1(\Omega).$$

The existence and uniqueness of the solution $u$ follow from boundedness and coercivity of $a(\cdot,\cdot)$ by the Lax-Milgram lemma [21]. It is obvious that the weighted $H^1$-semi-norm coincides with the energy norm induced by $a(\cdot,\cdot)$, namely,

$$\|v\|_A := \sqrt{a(v,v)} = \|\nabla v\|_{L^2_\rho(\Omega)} \qquad \forall\, v \in H_0^1(\Omega).$$

Let $\mathcal{T}_h$ be a conforming triangulation of $\Omega$, that is, any two elements in $\mathcal{T}_h$ are either nonintersecting or intersecting with a common vertex or a common edge. Throughout the paper, we assume that any triangulation of $\Omega$ takes care of the discontinuity of $\rho$, namely, $\rho|_T$ is constant for any $T \in \mathcal{T}_h$. We define the linear Lagrangian finite element space on $\mathcal{T}_h$ by

$$V_h = \Big\{ v_h \in H_0^1(\Omega) :\ v_h|_T \in P_1(T),\ \forall\, T \in \mathcal{T}_h \Big\}.$$

The Galerkin approximation to (2.1) is: Find $u_h \in V_h$ such that

$$a(u_h, v_h) = (f, v_h) \qquad \forall v_h \in V_h. \tag{2.2}$$

Let the linear operator $A_h : V_h \mapsto V_h$ be defined by

$$(A_h w_h, v_h)_\rho = a(w_h, v_h) \qquad \forall\, w_h, v_h \in V_h.$$

Clearly $A_h$ is symmetric and positive definite (SPD) and (2.2) is equivalent to the following operator equation

$$A_h u_h = f_h, \tag{2.3}$$

where $f_h \in V_h$ satisfies $(f_h, v)_\rho = (f, v)$ for any $v \in V_h$.

Let $N_h$ be the dimension of $V_h$ and $\{\boldsymbol{x}_i^h,\, i = 1, \ldots, N_h\}$ be the set of interior vertices of $\mathcal{T}_h$. We denote by $\varphi_i^h \in V_h$ a natural scaling of nodal basis function (cf. [4]) belonging to $\boldsymbol{x}_i^h$, $1 \le i \le N_h$. Then the operator equation (2.3) is equivalent to the following algebraic system

$$\mathbf{A}_h \mathbf{U}_h = \mathbf{F}_h, \tag{2.4}$$

where the entries of the matrix $\mathbf{A}_h$ and the vectors $\mathbf{U}_h, \mathbf{F}_h$ are defined by

$$(\mathbf{A}_h)_{ij} := a(\varphi_i^h, \varphi_j^h), \quad (\mathbf{U}_h)_i := u_h(\boldsymbol{x}_i^h), \quad (\mathbf{F}_h)_i := (f, \varphi_i^h) \qquad \forall\, i, j = 1, \ldots, N_h.$$

Using the arguments in Bank and Scott [4], we know that the $\ell_2$-condition number $\kappa(\mathbf{A}_h)$ can be estimated as follows:

$$\begin{cases} \kappa(\mathbf{A}_h) \le C\mathcal{J}(\rho) N_h \left(1 + |\log(N_h h_{\min}^2)|\right) & \text{if } d = 2, \\ \kappa(\mathbf{A}_h) \le C\mathcal{J}(\rho) N_h^{2/3} & \text{if } d = 3, \end{cases} \qquad \mathcal{J}(\rho) = \frac{\max_{\boldsymbol{x}\in\Omega} \rho(\boldsymbol{x})}{\min_{\boldsymbol{x}\in\Omega} \rho(\boldsymbol{x})}.$$

The following lemma is to estimate the convergence rate of the PCG algorithm for the operator equation (2.3) (cf. e.g. [3, 49]).

**Lemma 2.1.** *Let $B_h$ be an SPD preconditioner of $A_h$ such that the spectrum of $B_h A_h$ satisfies*

$$0 < \lambda_1 \le \cdots \le \lambda_{m_0} \ll \lambda_{m_0+1} \le \cdots \le \lambda_{N_h}. \tag{2.5}$$

*Let $u_k$ be the $k$-th iterate of the PCG algorithm. Then*

$$\frac{\|u_h - u_k\|_A}{\|u_h - u_0\|_A} \le 2\,|\kappa(B_h A_h) - 1|^{m_0} \left( \frac{\sqrt{\lambda_{N_h}/\lambda_{m_0+1}} - 1}{\sqrt{\lambda_{N_h}/\lambda_{m_0+1}} + 1} \right)^{k-m_0} \qquad \forall\, k \ge m_0. \tag{2.6}$$

**Remark 2.1.** If the integer $m_0$ is very small, the convergence rate of the PCG algorithm will be dominated by $\kappa_{m_0+1}(B_h A_h) = \lambda_{N_h}/\lambda_{m_0+1}$ which is known as the "effective condition number". In the following we shall study the spectral distribution (2.5) of the preconditioned system, where the preconditioner $B_h$ will be defined by a local multilevel solver.

## 3. Local Multilevel Methods

Let $\{\mathcal{T}_l\}_{l=0}^L$ be a family of nested conforming triangulations of $\Omega$ such that $\mathcal{T}_0$ is a quasi-uniform initial mesh and $\mathcal{T}_l$ is a (local) refinement of $\mathcal{T}_{l-1}$, $l \ge 1$, using the "newest vertex bisection" algorithm. For any $0 \leq l \leq L$, we denote the linear Lagrangian finite element space on $\mathcal{T}_l$ by $V_l \subset H_0^1(\Omega)$ and define $A_l : V_l \to V_l$ by

$$(A_l v, w)_\rho = a(v, w) \qquad \forall v, w \in V_l.$$

Then the operator equation (2.3) on $\mathcal{T}_l$ can be written as: Find $u_l \in V_l$ such that

$$A_l u_l = f_l, \tag{3.1}$$

where $f_l \in V_l$ satisfies that $(f_l, v_l)_\rho = (f, v_l)$ for any $v_l \in V_l$. For $0 \leq l \leq L$, we also define the energy projection $P_l$: $H_0^1(\Omega) \mapsto V_l$ and the weighted $L^2$-projection $Q_l^\rho$: $L^2(\Omega) \mapsto V_l$ by

$$a(P_l v, w) = a(v, w) \qquad \forall\, v \in H_0^1(\Omega),\ w \in V_l, \tag{3.2}$$

$$(Q_l^\rho v, w)_\rho = (v, w)_\rho \qquad \forall\, v \in L^2(\Omega),\ w \in V_l. \tag{3.3}$$

For $1 \leq l \leq L$, denote by $\mathcal{N}_l$ the set of interior nodes of $\mathcal{T}_l$ and by $\widetilde{\mathcal{N}}_l$ the set of nodes on which local relaxations are carried out. We shall give the exact definition of $\widetilde{\mathcal{N}}_l$ in Section 5. For brevity, we set $\widetilde{\mathcal{N}}_l = \{\boldsymbol{x}_i^l,\ i = 1, \ldots, \widetilde{n}_l\}$ with $\widetilde{n}_l$ being the cardinality of $\widetilde{\mathcal{N}}_l$, and we refer to $\phi_i^l$ as the nodal basis function of $V_l$ belonging to the node $\boldsymbol{x}_i^l$. For notational ease we set $V_1^0 := V_0$ and $\widetilde{n}_0 := 1$. We define the energy projection and the weighted $L^2$-projection onto the one-dimensional space $V_i^l := \operatorname{span}\{\phi_i^l\}$ as follows:

$$P_i^l :\ H_0^1(\Omega) \mapsto V_i^l, \quad a(P_i^l v, \phi_i^l) = a(v, \phi_i^l) \qquad \forall\, v \in H_0^1(\Omega),$$

$$Q_i^{\rho,l} :\ L^2(\Omega) \mapsto V_i^l, \quad (Q_i^{\rho,l} v, \phi_i^l)_\rho = (v, \phi_i^l)_\rho \qquad \forall\, v \in L^2(\Omega).$$

Let $A_i^l : V_i^l \mapsto V_i^l$ be defined by

$$(A_i^l v, \phi_i^l)_\rho = a(v, \phi_i^l) \qquad \forall v \in V_i^l.$$

Then the well-known relationship holds:

$$Q_i^{\rho,l} A_l = A_i^l P_i^l.$$

Let $R_l^J : V_l \mapsto V_l$ and $R_l^G : V_l \mapsto V_l$ be the local smoothing operators which perform Jacobi and Gauss-Seidel relaxations at the nodes in $\widetilde{\mathcal{N}}_l$, $1 \leq l \leq L$. Moreover, we set $R_0^J = R_0^G = A_0^{-1}$ on the initial mesh $\mathcal{T}_0$. Then $R_l^J$ defines an additive smoother (cf. [8]):

$$R_l^J := \gamma \sum_{i=1}^{\widetilde{n}_l} (A_i^l)^{-1} Q_i^{\rho,l}, \qquad 1 \leq l \leq L, \tag{3.4}$$

with a scaling factor $\gamma > 0$, while $R_l^G$ defines a multiplicative smoother:

$$R_l^G := (I - E_l) A_l^{-1}, \quad E_l := (I - P_{\widetilde{n}_l}^l) \cdots (I - P_1^l), \quad 1 \leq l \leq L. \tag{3.5}$$

With $R_l^J$ and $R_l^G$ at hand, we construct the local multilevel algorithms for the adaptive finite element approximation to (2.1).

**Algorithm 3.1.** (*Local multilevel additive algorithm* (*LMAA*))
Given an initial guess $\hat{u}_0 \in V_L$, the $k$-th iterate of LMAA applied to (3.1) on $\mathcal{T}_L$ is defined by:

$$\hat{u}_k = \hat{u}_{k-1} + B_L^A (f_L - A_L \hat{u}_{k-1}), \qquad k \geq 1,$$

where $B_L^A = \sum_{l=0}^L R_l Q_l^\rho$ is an additive multilevel operator and the smoother $R_l$ can be either the local Jacobi smoother $R_l = R_l^J$ or the local Gauss-Seidel smoother $R_l = R_l^G$.

**Algorithm 3.2.** (*Symmetrical local multilevel additive algorithm* (*SLMAA*))
Given an initial guess $\hat{u}_0 \in V_L$, the $k$-th iterate of SLMAA applied to (3.1) on $\mathcal{T}_L$ is defined by:

$$\hat{u}_k = \hat{u}_{k-1} + \overline{B}_L^A (f_L - A_L \hat{u}_{k-1}), \qquad k \geq 1,$$

where $\overline{B}_L^A = \left(B_L^A + (B_L^A)^t\right)/2$ is the symmetrization of $B_L^A$.

**Algorithm 3.3.** (*Local multilevel multiplicative algorithm* (*LMMA*))
Given an initial guess $\hat{u}_0 \in V_L$, the $k$-th iterate of LMMA applied to (3.1) on $\mathcal{T}_L$ is defined by:

$$\hat{u}_k = \hat{u}_{k-1} + B_L^M (f_L - A_L \hat{u}_{k-1}), \qquad k \geq 1.$$

For any $g \in V_l$, the multiplicative multilevel operators $B_l^M$: $V_l \mapsto V_l$, $l \geq 0$ are recursively defined as follows: $B_0^M := A_0^{-1}$ and $B_l^M g = x_3$,

1. pre-smoothing: $x_1 = (R_l)^t g$;
2. correction: $x_2 = x_1 + B_{l-1}^M Q_{l-1}^\rho (g - A_l x_1)$;
3. post-smoothing: $x_3 = x_2 + R_l (g - A_l x_2)$,

where the smoother $R_l$ can be either the local Jacobi smoother $R_l = R_l^J$ or the local Gauss-Seidel smoother $R_l = R_l^G$.

## 4. The Abstract Schwarz Theory

In this section, we present an abstract Schwarz theory for the local multilevel methods. We shall adopt the abstract theory (cf. [41, 43]) to the LMMA, LMAA algorithms and the PCG algorithms for which LMMA and LMAA serve as preconditioners.

Let $M \geq 1$ be the smallest integer such that there exists a family of open polygonal or polyhedral subdomains $\{\Omega_i \subset \Omega : \ 1 \leq i \leq M\}$ satisfying

$$\bigcup_{i=1}^{M} \overline{\Omega}_i = \overline{\Omega}, \quad \Omega_i \cap \Omega_j = \emptyset \ \text{ if } \ i \neq j, \quad \text{and} \quad \rho_i := \rho|_{\Omega_i} = \text{Constant}.$$

We introduce the set of indices of subdomains which do not touch $\partial\Omega$:

$$\mathcal{I} = \{i : \ \partial\Omega_i \cap \partial\Omega = \emptyset, \ 1 \leq i \leq M\}. \tag{4.1}$$

As in [49], we define a subspace $\widetilde{V}_l \subset V_l$ by

$$\widetilde{V}_l = \Big\{ v \in V_l : \int_{\Omega_i} v(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = 0, \ i \in \mathcal{I} \Big\}. \tag{4.2}$$

Then using Poincáre's inequality and Friedrichs' inequality we have

$$\begin{aligned}\|v\|^2_{L^2_\rho(\Omega)} &= \sum_{i=1}^{M} \rho_i \|v\|^2_{L^2(\Omega_i)} = \sum_{i\in\mathcal{I}} \rho_i \|v\|^2_{L^2(\Omega_i)} + \sum_{i\in\{1,\dots,M\}\setminus\mathcal{I}} \rho_i \|v\|^2_{L^2(\Omega_i)} \\ &\le C\Big(\sum_{i\in\mathcal{I}} \rho_i\, |\nabla v|^2_{L^2(\Omega_i)} + \sum_{i\in\{1,\dots,M\}\setminus\mathcal{I}} \rho_i\, |\nabla v|^2_{L^2(\Omega_i)}\Big) \le C\, \|v\|^2_A, \quad \forall\, v\in \widetilde{V}_l,\end{aligned} \tag{4.3}$$

where the constant $C$ depends on $\Omega_1,\dots,\Omega_M$.

The abstract Schwarz theory depends greatly on two important properties of the finite element spaces $\{V_l\}_{l=0}^L$, that is, the existence of a stable multilevel decomposition of $V_L$ and the strengthened Cauchy-Schwarz inequality regarding the space decomposition. At this moment we simply state the two properties and postpone the proofs to the next section.

(A1) *Stability of the multilevel decomposition.* For any function $v\in V_L$, there exists a decomposition of $v$:

$$v = v_0 + \sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l} v_i^l, \qquad v_0\in V_0,\ v_i^l \in V_i^l, \tag{4.4}$$

and a positive constant $C_{\text{stab}}$ independent of $\mathcal{J}(\rho)$, $L$, and $h_{\min}$ such that

$$\|v_0\|^2_A + \sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l} \left\|v_i^l\right\|^2_A \le C_{\text{stab}} C_d^{h,\rho}\, \|v\|^2_A\,, \tag{4.5}$$

where $d$ is the dimension of $\Omega$ and

$$C_d^{h,\rho} := \begin{cases} \min\{|\log h_{\min}|^2, \mathcal{J}(\rho)\}, & \text{if } d=2,\\ \min\{h_{\min}^{-1}, \mathcal{J}(\rho)\}, & \text{if } d=3.\end{cases} \tag{4.6}$$

In particular, there also exists a positive constant $\widetilde{C}_{\text{stab}}$ independent of $\mathcal{J}(\rho)$, $L$, and $h_{\min}$ such that

$$\|v_0\|^2_A + \sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l} \left\|v_i^l\right\|^2_A \le \widetilde{C}_{\text{stab}} |\log h_{\min}|^2\, \|v\|^2_A \qquad \forall v\in \widetilde{V}_L. \tag{4.7}$$

(A2) *Strengthened Cauchy-Schwarz inequality.* For any functions

$$v_i^l, w_i^l \in V_i^l, \qquad 1\le i\le \widetilde{n}_l,\ 0\le l\le L,$$

there exists a constant $C_{\text{orth}}$ independent of $\mathcal{J}(\rho)$, $L$, and $h_{\min}$ such that

$$\sum_{l=0}^{L}\sum_{i=1}^{\widetilde{n}_l}\sum_{k=0}^{l-1}\sum_{j=1}^{\widetilde{n}_k} a(v_i^l, w_j^k) \le C_{\text{orth}}\Big(\sum_{l=0}^{L}\sum_{i=1}^{\widetilde{n}_l}\left\|v_i^l\right\|^2_A\Big)^{\frac12}\Big(\sum_{l=0}^{L}\sum_{i=1}^{\widetilde{n}_l}\left\|w_i^l\right\|^2_A\Big)^{\frac12}. \tag{4.8}$$

**Lemma 4.1.** *Let $T_l = R_l A_l P_l$ where $R_l = R_l^J$ or $R_l^G$, $0\le l\le L$. Then the following statements hold with a constant $C>0$ only depending on the domain and the shape regularity of the meshes:*

(E1) *Let* $T_A = \sum_{l=0}^{L} R_l A_l P_l$ *be the additive operator. Then*

$$\|v\|_A^2 \le C C_d^{h,\rho} a(T_A v, v) \qquad \forall v \in V_L,$$
$$\|v\|_A^2 \le C |\log h_{\min}|^2 a(T_A v, v) \qquad \forall v \in \widetilde{V}_L.$$

(E2) *For any* $v_l, w_k \in V_L$, $0 \le l, k \le L$, *we have*

$$\sum_{l=0}^{L} \sum_{k=0}^{l-1} a(T_l v_l, T_k w_k) \le C \Big( \sum_{l=0}^{L} a(T_l v_l, v_l) \Big)^{\frac{1}{2}} \Big( \sum_{k=0}^{L} a(T_k w_k, w_k) \Big)^{\frac{1}{2}}.$$

(E3) *There exists a constant* $0 < \omega_l < 2$ *independent of* $\mathcal{J}(\rho)$, $L$, $h_{\min}$ *such that*

$$\|T_l v\|_A^2 \le \omega_l a(T_l v, v) \qquad \forall v \in V_L,\ 0 \le l \le L.$$

*If* $R_l = R_l^J, 1 \le l \le L$, *the scaling factor should be so chosen such that* $\omega_l < 2$.

(E4) *For any* $v_l, w_l \in V_L$, $0 \le l \le L$, *we have*

$$\sum_{l=0}^{L} a(T_l v_l, w_l) \le C \Big( \sum_{l=0}^{L} a(T_l v_l, v_l) \Big)^{\frac{1}{2}} \Big( \sum_{l=0}^{L} a(T_l w_l, w_l) \Big)^{\frac{1}{2}}.$$

*Proof.* The lemma can be proved upon using (A1)–(A2) and similar arguments as in [50]. We omit the details here. □

For Algorithm 3.3, we can easily derive a representation of the multigrid error propagation operator

$$I - B_L^M A_L = E_M E_M^*, \tag{4.9}$$

where $I$ is the identity operator on $V_L$, $E_M^*$ is the conjugate of the operator $E_M$, and

$$E_M := (I - T_L)(I - T_{L-1}) \cdots (I - T_0), \quad T_l = R_l A_l P_l, \quad 0 \le l \le L. \tag{4.10}$$

Using Lemma 4.1 and similar arguments as in [43], we obtain the following theorem.

**Theorem 4.1.** *Let* $B_L^M$ *be the multiplicative multilevel operator in Algorithm* 3.3 *and* $C_d^{h,\rho}$ *be the constant defined in* (4.6). *There exists a constant* $C > 0$ *only depending on the domain and the shape regularity of the meshes such that*

$$a((I - B_L^M A_L)v, v) \le \delta\, a(v, v) \qquad \forall v \in V_L, \tag{4.11}$$
$$a((I - B_L^M A_L)v, v) \le \widetilde{\delta}\, a(v, v) \qquad \forall v \in \widetilde{V}_L, \tag{4.12}$$

*where*

$$\delta := 1 - \frac{2-\omega}{C C_d^{h,\rho}}, \quad \widetilde{\delta} := 1 - \frac{2-\omega}{C|\log h_{\min}|^2}, \quad \omega := \max_{0 \le l \le L} \omega_l < 2.$$

Since $a((I - B_L^M A_L)v, v) = a(E_M^* v, E_M^* v) \ge 0$, we have $\lambda_{\max}(B_L^M A_L) \le 1$. From the estimate (4.11) the minimum eigenvalue of $B_L^M A_L$ reads

$$\lambda_{\min}(B_L^M A_L) = \inf_{v \in V_L, v \ne 0} \frac{a(B_L^M A_L v, v)}{\|v\|_A} \ge \frac{2-\omega}{C C_d^{h,\rho}}.$$

Denote by $m_0 = \#\mathcal{I}$ the cardinality of the index set $\mathcal{I}$ in (4.1). Obviously $m_0 \le M$ and $\dim(\widetilde{V}_L) = \dim(V_L) - m_0$ from (4.2). Then by (4.12) we have

$$\lambda_{m_0+1}(B_L^M A_L) \ge \inf_{v\in\widetilde{V}_L, v\neq 0} \frac{a(B_L^M A_L v, v)}{\|v\|_A} \ge \frac{2-\omega}{C|\log h_{\min}|^2}.$$

Since $\omega$ is independent of $\mathcal{J}(\rho)$, $L$, $h_{\min}$ by (E3) of Lemma 4.1, the $\ell_2$-condition number $\kappa(B_L^M A_L)$ and the *effective condition number* $\kappa_{m_0+1}(B_L^M A_L)$ can be bounded as follows:

$$\kappa(B_L^M A_L) \le C C_d^{h,\rho}, \qquad \kappa_{m_0+1}(B_L^M A_L) := \frac{\lambda_{\max}(B_L^M A_L)}{\lambda_{m_0+1}(B_L^M A_L)} \le C|\log h_{\min}|^2.$$

**Lemma 4.2.** *Let $B_L^A$ and $\overline{B}_L^A$ be the additive multilevel operators in Algorithm* 3.1 *and* 3.2 *respectively. Then the operators $T_A = \sum_{l=0}^L R_l A_l P_l = B_L^A A_L$, $R_l = R_l^J$ or $R_l^G$, and $\overline{T}_A = \frac{1}{2}(T_A + T_A^*) = \overline{B}_L^A A_L$ admit the following stability properties*

$$\|T_A v\|_A \le C\|v\|_A, \quad \|\overline{T}_A v\|_A \le C\|v\|_A \qquad \forall v \in V_L,$$

*where the constant $C > 0$ only depends on the domain and the shape regularity of the meshes.*

*Proof.* The lemma is a direct consequence of (E2) of Lemma 4.1. □

If $R_l^J$ is symmetric, then $T_A$ is symmetric with respect to $a(\cdot,\cdot)$. From Lemma 4.2 and (E1) of Lemma 4.1, we know that

$$\kappa(B_L^A A_L) \le C C_d^{h,\rho}, \qquad \kappa_{m_0+1}(B_L^A A_L) \le C|\log h_{\min}|^2.$$

If $T_A$ is nonsymmetric, we have the following estimates for Algorithm 3.2:

$$\kappa(\overline{B}_L^A A_L) \le C C_d^{h,\rho}, \qquad \kappa_{m_0+1}(\overline{B}_L^A A_L) \le C|\log h_{\min}|^2.$$

For convenience, we denote by LMAA-PCG, SLMAA-PCG, LMMA-PCG the PCG algorithms with Algorithm 3.1, 3.2, 3.3 as preconditioners respectively. Notice that Theorem 4.1 presents the convergence rate of Algorithm 3.3. To end this section, we conclude the convergence of the multilevel-preconditioned conjugate gradient methods, namely, LMAA-PCG, SLMAA-PCG, and LMMA-PCG.

**Theorem 4.2.** *Let $u_h$ be the finite element solution of* (2.2) *on $\mathcal{T}_L$ and $u_k$ be the $k$-th iterate of the LMMA-PCG algorithm, or the LMAA-PCG with local Jacobi smoothers, or the SLMAA-PCG algorithm. Then there exists a constant $C$ independent of $\mathcal{J}(\rho)$, $L$, $h_{\min}$ such that*

$$\frac{\|u_h - u_k\|_A}{\|u_h - u_0\|_A} \le 2\left(C_d^{h,\rho} - 1\right)^{m_0}\left(1 - \frac{2}{1 + C|\log h_{\min}|}\right)^{k-m_0}, \quad k \ge m_0,$$

*where $m_0 = \#\mathcal{I}$ is the cardinality of $\mathcal{I}$ in* (4.1) *and*

$$C_d^{h,\rho} := \begin{cases} \min\{|\log h_{\min}|^2, \mathcal{J}(\rho)\}, & \text{if } d = 2, \\ \min\{h_{\min}^{-1}, \mathcal{J}(\rho)\}, & \text{if } d = 3. \end{cases}$$

**Remark 4.1.** In Theorem 4.2, the integer $m_0$ only depends on $\Omega$ and the distribution of $\rho$. It may happen that $m_0 = 0$ for some instances. Thus for any $k > k_0$ with $k_0$ satisfying

$$2\left(C_d^{h,\rho} - 1\right)^{m_0}\left(1 - \frac{2}{1 + C|\log h_{\min}|}\right)^{k-m_0} \le 1,$$

the convergence rate of the PCG algorithms is

$$1 - \frac{2}{1 + C|\log h_{\min}|}.$$

**Remark 4.2.** If the coefficient $\rho$ is quasi-monotone, the convergence of multilevel methods can be proved independent of $\mathcal{J}(\rho), L, h_{\min}$ (see [47]). We do not elaborate on this issue in this paper.

## 5. Verification of the Two Properties (A1) and (A2)

This section is devoted to the verification of the two properties (A1) and (A2) of the finite element spaces. The key ingredient is to construct a local multilevel decomposition of $V_L$ regarding the adaptively refined meshes $\{\mathcal{T}_l\}_{l=0}^L$.

### 5.1. Quasi-interpolation operator

Local quasi-interpolation operators play an important role in the analysis of multilevel decomposition. In this section, we introduce an interpolation operator $\Pi_l$: $L^2(\Omega) \mapsto V_l$ which is a modification of the one studied by Hiptmair and Zheng in [28]. For any $T \in \mathcal{T}_l$, we define the dual basis function $\psi_i^T \in P_1(T)$ by the $L^2(T)$-duality to the barycentric coordinate functions $\lambda_i$, $i = 1, \dots, d+1$ on $T$ which satisfies

$$\int_T \psi_j^T(\boldsymbol{x})\lambda_i(\boldsymbol{x})d\boldsymbol{x} = \delta_{ij} \qquad \text{for } i, j = 1, \dots, d+1. \tag{5.1}$$

By computing the explicit representation of $\psi_j^T$ we have

$$C_0 \le |T|\,\|\psi_j^T\|^2_{L^2(T)} \le C_1 \quad \text{and} \quad C_0 \le \|\psi_j^T\|_{L^1(T)} \le C_1, \tag{5.2}$$

where $C_0$ and $C_1$ only depend on the shape regularity of $\mathcal{T}_l$, $0 \le l \le L$.

For $0 \le l \le L$, the local quasi-interpolation operators $\Pi_l : L^2(\Omega) \mapsto V_l$ are defined as follows:

$$\Pi_l v = \sum_{\boldsymbol{p} \in \mathcal{N}_l} \int_{T_{\boldsymbol{p}}^l} \psi_{\boldsymbol{p}}^{T_{\boldsymbol{p}}^l}(\boldsymbol{x}) v(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \cdot \phi_{\boldsymbol{p}}^l \qquad \forall v \in L^2(\Omega), \tag{5.3}$$

where $\phi_{\boldsymbol{p}}^l \in V_l$ is the nodal basis function belonging to $\boldsymbol{p}$, $T_{\boldsymbol{p}}^l \in \mathcal{T}_l$ satisfies $T_{\boldsymbol{p}}^l \subset \Omega_{\boldsymbol{p}}^l := \operatorname{supp}(\phi_{\boldsymbol{p}}^l)$, and $\psi_{\boldsymbol{p}}^{T_{\boldsymbol{p}}^l}$ is the dual basis function defined in (5.1) and belonging to $\boldsymbol{p} \in \mathcal{N}_l$. In view of (5.1), it is easy to see that

$$\Pi_l v = v \quad \forall\, v \in V_l. \tag{5.4}$$

It is clear that the definition of $\Pi_l$ depends on how to select $T_{\boldsymbol{p}}^l$ for each $\boldsymbol{p} \in \mathcal{N}_l$. We shall adapt the selection of $T_{\boldsymbol{p}}^l$ to our multilevel theory regarding the discontinuous coefficient $\rho$. Notice that $\rho$ is constant on any element of $\mathcal{T}_0$. For any $\boldsymbol{p} \in \mathcal{N}_0$, we select $T_{\boldsymbol{p}}^0 \in \mathcal{T}_0$ such that

$$T_{\boldsymbol{p}}^0 \subset \Omega_{\boldsymbol{p}}^0 \quad \text{and} \quad \rho|_{T_{\boldsymbol{p}}^0} = \max\{\rho|_T :\ T \subset \Omega_{\boldsymbol{p}}^0,\ T \in \mathcal{T}_0\}. \tag{5.5}$$

For $1 \le l \le L$ and $\boldsymbol{p} \in \mathcal{N}_l$, we select $T_{\boldsymbol{p}}^l$ successively according to the following policy:

1. For any vertex $\boldsymbol{p} \in \mathcal{N}_l \cap \mathcal{N}_{l-1}$, we choose a $T^l_{\boldsymbol{p}} \in \mathcal{T}_l$ such that $T^l_{\boldsymbol{p}} \subseteq T^{l-1}_{\boldsymbol{p}}$.
2. For any vertex $\boldsymbol{p} \in \mathcal{N}_l \setminus \mathcal{N}_{l-1}$, we choose $T^l_{\boldsymbol{p}} \in \mathcal{T}_l$ such that

$$T^l_{\boldsymbol{p}} \subset \Omega^l_{\boldsymbol{p}} \quad \text{and} \quad \rho|_{T^l_{\boldsymbol{p}}} = \max\{\rho|_T,\ T \subset \Omega^l_{\boldsymbol{p}},\ T \in \mathcal{T}_l\}.$$

**Lemma 5.1.** *There exists a constant $C > 0$ only depending on the domain and the shape regularity of the meshes such that*

$$\begin{aligned} &\|\Pi_0 v\|_A^2 \le C\widetilde{C}^h_d \|v\|_A^2 && \forall v \in V_L, \\ &\|\Pi_0 v\|_A^2 \le C \|v\|_A^2 && \forall v \in \widetilde{V}_L, \end{aligned}$$

*where $\widetilde{C}^h_d = |\log h_{\min}|$ if $d = 2$ and $\widetilde{C}^h_d = h^{-1}_{\min}$ if $d = 3$.*

*Proof.* For any $T \in \mathcal{T}_0$ with vertices $\boldsymbol{p}_i$, $1 \le i \le d+1$, we denote by $\phi_i = \phi^0_{\boldsymbol{p}_i}$, $T_i = T^0_{\boldsymbol{p}_i}$, $\psi_i = \psi^{T_i}_{\boldsymbol{p}_i}$ the nodal basis function, the selected element, and the dual basis function belonging to $\boldsymbol{p}_i$ respectively. From (3.3) we have

$$\|\Pi_0 v\|_A^2 \le \|\Pi_0(I - Q^\rho_0)v\|_A^2 + \|Q^\rho_0 v\|_A^2 .$$

By the definition of $\Pi_0$, direct calculations show that

$$\begin{aligned} \|\nabla\Pi_0(I - Q^\rho_0)v\|^2_{L^2_\rho(T)} &= \rho_T \|\nabla\Pi_0(I - Q^\rho_0)v\|^2_{L^2(T)} \\ &\le C\rho_T \sum_{i=1}^{d+1} \left| \int_{T_i} \psi_i(\boldsymbol{x})(I - Q^\rho_0)v(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \right|^2 \|\nabla\phi_i\|^2_{L^2(T)} \\ &\le C\rho_T h_T^{d-2}|T|^{-1} \sum_{i=1}^{d+1} \|(I - Q^\rho_0)v\|^2_{L^2(T_i)} \\ &\le C h_T^{-2} \|(I - Q^\rho_0)v\|^2_{L^2_\rho(D_T)} , \end{aligned}$$

where $D_T = \bigcup_{i=1}^{d+1} \Omega^0_{\boldsymbol{p}_i}$. Summing the above estimate over all elements in $\mathcal{T}_0$ leads to

$$\|\Pi_0(I - Q^\rho_0)v\|_A^2 \le C h_0^{-2} \|(I - Q^\rho_0)v\|^2_{L^2_\rho(\Omega)} ,$$

where $h_0$ is the mesh size of the initial mesh $\mathcal{T}_0$. By the argument in [11, Theorem 4.5], we obtain the following estimate for the weighted $L^2$-projection:

$$\begin{aligned} &\|Q^\rho_0 v\|_A^2 + h_0^{-2} \|(I - Q^\rho_0)v\|^2_{L^2_\rho(\Omega)} \le C\widetilde{C}^h_d \|v\|_A^2 && \forall\, v \in V_L, \\ &\|Q^\rho_0 v\|_A^2 + h_0^{-2} \|(I - Q^\rho_0)v\|^2_{L^2_\rho(\Omega)} \le C \|v\|_A^2 && \forall\, v \in \widetilde{V}_L. \end{aligned}$$

Combining the above estimates concludes the proof. □

### 5.2. Local multilevel decomposition

For any $v \in V_L$, (5.4) indicates the following multilevel decomposition of $v$ :

$$v = \sum_{l=0}^{L} v_l, \quad v_0 = \Pi_0 v, \quad v_l = (\Pi_l - \Pi_{l-1})v, \quad 1 \le l \le L. \tag{5.6}$$

From the definition of $\Pi_l$, it is clear that

$$v_l = (\Pi_l - \Pi_{l-1})v = \sum_{\boldsymbol{p}\in\widetilde{\mathcal{N}}_l} v^l_{\boldsymbol{p}}, \quad v^l_{\boldsymbol{p}} = v_l(\boldsymbol{p})\phi^l_{\boldsymbol{p}}, \quad 1 \le l \le L, \tag{5.7}$$

where $\widetilde{\mathcal{N}}_l$ is the set of smoothing nodes defined by

$$\widetilde{\mathcal{N}}_l := (\mathcal{N}_l \setminus \mathcal{N}_{l-1}) \bigcup \{\boldsymbol{p} \in \mathcal{N}_l \cap \mathcal{N}_{l-1} : \ \phi^l_{\boldsymbol{p}} \ne \phi^{l-1}_{\boldsymbol{p}} \ \text{ or } \ T^l_{\boldsymbol{p}} \ne T^{l-1}_{\boldsymbol{p}}\}.$$

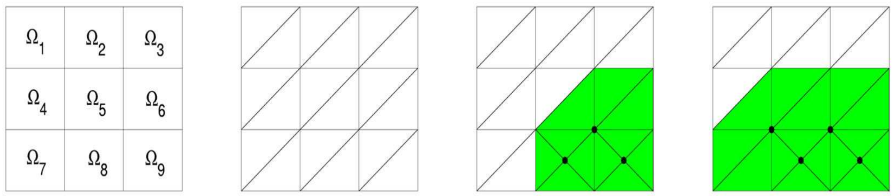


Fig. 5.1. The first figure shows the domain $\Omega$ and the distribution of $\rho$ such that $\rho_1 < \rho_2 < \cdots < \rho_8 < \rho_9$. The second figure shows the mesh $\mathcal{T}_{l-1}$. The third and fourth figures show the mesh $\mathcal{T}_l$. The black dots in the third figure show the nodes in $(\mathcal{N}_l \setminus \mathcal{N}_{l-1})$ and $\{\boldsymbol{p} \in \mathcal{N}_l \cap \mathcal{N}_{l-1} : \ \phi^l_{\boldsymbol{p}} \ne \phi^{l-1}_{\boldsymbol{p}}\}$. The black dots in the fourth figure show the nodes in $\widetilde{\mathcal{N}}_l$ to which local relaxations are restricted.

The local multilevel algorithms in [46,50] perform local relaxations on the nodes in $(\mathcal{N}_l \backslash \mathcal{N}_{l-1})$ and $\{\boldsymbol{p} \in \mathcal{N}_l \cap \mathcal{N}_{l-1} : \ \phi^l_{\boldsymbol{p}} \ne \phi^{l-1}_{\boldsymbol{p}}\}$. Our algorithms perform additional relaxations on the nodes in $\{\boldsymbol{p} \in \mathcal{N}_l \cap \mathcal{N}_{l-1} : \ T^l_{\boldsymbol{p}} \ne T^{l-1}_{\boldsymbol{p}}\}$ (see Figure 5.1 for the 2D case). Actually, these incremental relaxations do not have an impact on the optimality of the algorithms.

### 5.3. Stability estimate

The purpose of this section is to prove that the multilevel decomposition (5.6) satisfies the stability in (A1). The analysis relies on two assumptions on these meshes.

(H1) The shape regularity measures of the meshes $\mathcal{T}_0, \cdots, \mathcal{T}_L$ are uniformly bounded, that is, $\sigma(\mathcal{T}_l) \le C$ for all $0 \le l \le L$. Here $\sigma(\mathcal{T}_l)$ stands for the shape regularity measure of $\mathcal{T}_l$ and the constant $C$ is independent of the mesh sizes and the mesh levels.

(H2) There exists a constant integer $z > 0$ such that

$$\left[\ln(h_{T'} h_T^{-1})/\ln 2\right] \le z \qquad \forall T \in \mathcal{T}_l, \ \ 1 \le l \le L,$$

where $T' \in \mathcal{T}_{l-1}$ satisfying $T \subset T'$ and for any $\xi \ge 0$, $[\xi]$ stands for the largest integer less than or equal to $\xi$.

Assumption (H1) always holds for the popular bisection algorithms. Assumption (H2) implies that the adaptive refinement strategy should stop in finite bisections and is usually satisfied. We refer to [46] for a detailed proof of (H2) for the two-dimensional bisection algorithm.

Our theory depends on a close relationship between the adaptively refined meshes $\{\mathcal{T}_l\}_{l=0}^L$, and a sequence of quasi-uniformly refined meshes $\{\widehat{\mathcal{T}}_j\}_{j\ge 0}$. Here $\widehat{\mathcal{T}}_j$ is generated by connecting the edge midpoints of each element in $\widehat{\mathcal{T}}_{j-1}$ starting from $\widehat{\mathcal{T}}_0 = \mathcal{T}_0$. For $d = 2$, each triangle in $\widehat{\mathcal{T}}_{j-1}$ is subdivided into four congruent triangles by connecting the midpoints of the four edges.

For $d = 3$, each tetrahedron in $\widehat{\mathcal{T}}_{j-1}$ is subdivided into eight subtetrahedra by connecting the midpoints of the six edges.

For any $l \geq 0$ and $T \in \mathcal{T}_l$, there exists a $T_0 \in \mathcal{T}_0$ satisfying $T \subset T_0$. We define

$$n(T) = \left[\ln(h_{T_0} h_T^{-1})/\ln 2\right]. \tag{5.8}$$

It is easy to see that $n(T) = j$ for any $T \in \widehat{\mathcal{T}}_j$ and $j \geq 0$. The following lemma describes the relationship between $\{\mathcal{T}_l\}_{l=0}^L$ and $\{\widehat{\mathcal{T}}_j\}_{j\geq 0}$ which is used in our analysis.

**Lemma 5.2.** *For any $0 \leq l \leq L$ and $T \in \mathcal{T}_l$, there exists a $\widehat{T} \in \widehat{\mathcal{T}}_{n(T)}$ such that*

$$T \subset \widehat{T} \quad \text{and} \quad h_{\widehat{T}} \leq C h_T,$$

*where $C$ only depends on the shape regularity of the meshes.*

*Proof.* First we consider an arbitrary simplex $T$ and define an initial mesh of $T$ by $\mathcal{M}_0(T) = \{T\}$. Let $\widehat{\mathcal{M}}(T)$ be generated by a unform refinement of $\mathcal{M}_0(T)$, namely, by connecting the midpoints of the edges of $T$. Thus $\widehat{\mathcal{M}}(T)$ contains smaller elements:

$$\widehat{\mathcal{M}}(T) = \{\widehat{K}_1, \cdots, \widehat{K}_4\} \quad \text{for} \quad d = 2, \qquad \widehat{\mathcal{M}}(T) = \{\widehat{K}_1, \cdots, \widehat{K}_8\} \quad \text{for} \quad d = 3.$$

Clearly $h_{\widehat{K}} = 2^{-1} h_T$ for any $\widehat{K} \in \widehat{\mathcal{M}}(T)$ (see Figure 5.2 (right) for a 2D illustration).

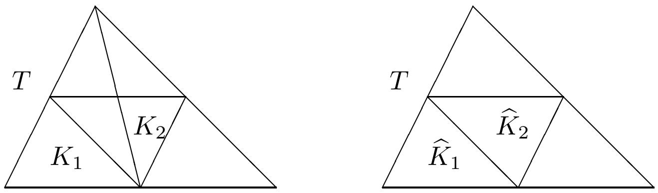


Fig. 5.2. Two elements satisfying $K_1 \subseteq \widehat{K}_1$, $K_2 \subseteq \widehat{K}_2$ in two dimension.

Furthermore, we generate a family of conforming meshes $\{\mathcal{M}_k(T)\}_{k=0}^I$ by successive bisections of $T$, where $\mathcal{M}_k(T)$ is a refinement of $\mathcal{M}_{k-1}(T)$. On the final mesh $\mathcal{M}_I(T)$, each triangular face of $T$ is subdivided as in the left picture of Figure 5.2. In this case, $\mathcal{M}_I(T)$ has 6 elements for $d = 2$ and 22 elements for $d = 3$:

$$\mathcal{M}_I(T) = \{K_1, \cdots, K_6\} \quad \text{for} \quad d = 2, \qquad \mathcal{M}_I(T) = \{K_1, \cdots, K_{22}\} \quad \text{for} \quad d = 3.$$

It is easy to see that for any $K \in \mathcal{M}_I(T)$, there exists a $\widehat{K} \in \widehat{\mathcal{M}}_1(T)$ such that

$$K \subset \widehat{K} \quad \text{and} \quad \left[\ln(h_T h_K^{-1})/\ln 2\right] = 1. \tag{5.9}$$

According to (5.9), for any $0 \leq l \leq L$ and $T \in \mathcal{T}_l$, there exist two sequences of elements $\{T_i\}_{i=0}^m$ and $\{\widehat{T}_i\}_{i=0}^m$ such that $T_0 = \widehat{T}_0 \in \mathcal{T}_0$ and

$$T_i \subset \widehat{T}_i \in \widehat{\mathcal{M}}(\widehat{T}_{i-1}), \quad \widehat{\mathcal{M}}(\widehat{T}_{i-1}) \subset \widehat{\mathcal{T}}_i, \quad 1 \leq i \leq m, \tag{5.10}$$

$$T_i \in \mathcal{M}_I(T_{i-1}), \quad \left[\ln\left(h_{T_0} h_{T_i}^{-1}\right)/\ln 2\right] = i, \quad 1 \leq i \leq m, \quad T \in \bigcup_{k=0}^{I-1} \mathcal{M}_k(T_m). \tag{5.11}$$

From (5.10)–(5.11) we conclude that

$$\begin{aligned} m = \left[\ln\left(h_{T_0}h_{T_m}^{-1}\right)/\ln 2\right] = \left[\ln\left(h_{T_0}h_{T}^{-1}\right)/\ln 2\right] = n(T), \\ T \subset T_m \subset \widehat{T}_m \in \widehat{\mathcal{T}}_m \quad \text{and} \quad h_{\widehat{T}_m} = 2^{-m}h_{T_0} \le Ch_{T_m} \le Ch_T. \end{aligned}$$

The proof is finished. □

**Lemma 5.3.** *Let* $v = \sum_{l=1}^{L}\sum_{\boldsymbol{p}\in\widetilde{\mathcal{N}}_l} v_{\boldsymbol{p}}^l$ *be the decomposition in* (5.6)–(5.7). *There exists a constant* $C > 0$ *only depending on the shape regularity of the meshes such that*

$$\sum_{l=1}^{L}\sum_{\boldsymbol{p}\in\widetilde{\mathcal{N}}_l} \left\|v_{\boldsymbol{p}}^l\right\|_A^2 \le CC_d^h \left\|v\right\|_A^2 \qquad \forall\, v \in V_L, \tag{5.12}$$

$$\sum_{l=1}^{L}\sum_{\boldsymbol{p}\in\widetilde{\mathcal{N}}_l} \left\|v_{\boldsymbol{p}}^l\right\|_A^2 \le C|\log h_{\min}|^2 \left\|v\right\|_A^2 \qquad \forall\, v \in \widetilde{V}_L, \tag{5.13}$$

*where* $C_d^h = |\log h_{\min}|^2$ *if* $d = 2$ *and* $C_d^h = h_{\min}^{-1}$ *if* $d = 3$.

*Proof.* For any $1 \le l \le L$ and any vertex $\boldsymbol{p} \in \widetilde{\mathcal{N}}_l$, we choose an element $T' \in \mathcal{T}_{l-1}$ such that $\boldsymbol{p} \in \overline{T'}$ and define

$$\mathcal{T}_l(\boldsymbol{p}) = \{T \in \mathcal{T}_{l-1} :\ \overline{T'} \cap \overline{T} \ne \emptyset\} \quad \text{and} \quad n(l,\boldsymbol{p}) = \min\{n(T) : T \in \mathcal{T}_l(\boldsymbol{p})\},$$

where $n(T)$ is defined in (5.8). From Lemma 5.2, for any $T \in \mathcal{T}_l(\boldsymbol{p})$, there exists a $\widehat{T} \in \widehat{\mathcal{T}}_{n(l,\boldsymbol{p})}$ such that

$$T \subset \widehat{T} \quad \text{and} \quad h_T \ge Ch_{\widehat{T}} \ge C2^{-n(l,\boldsymbol{p})}h_0.$$

Let $\widehat{Q}_m^{\rho} : L^2(\Omega) \mapsto \widehat{V}_m$ be the weighted $L^2$-projection and $\widehat{Q}_m^{\rho} = \widehat{Q}_0^{\rho}$ if $m < 0$, where $\widehat{V}_m$ is the linear Lagrangian finite element space on $\widehat{\mathcal{T}}_m$. Clearly $\widehat{Q}_{n(l,\boldsymbol{p})}^{\rho}v$ is linear on each element of $\mathcal{T}_l(\boldsymbol{p})$. By the definition of $\Pi_l$, we have

$$\Pi_l\widehat{Q}_{n(l,\boldsymbol{p})}^{\rho}v(\boldsymbol{p}) = \widehat{Q}_{n(l,\boldsymbol{p})}^{\rho}v(\boldsymbol{p}) = \Pi_{l-1}\widehat{Q}_{n(l,\boldsymbol{p})}^{\rho}v(\boldsymbol{p}). \tag{5.14}$$

Notice that

$$\left\|v_{\boldsymbol{p}}^l\right\|_A^2 = |v_l(\boldsymbol{p})|^2 \left\|\phi_{\boldsymbol{p}}^l\right\|_A^2 \le C\rho_{T_{\boldsymbol{p}}^l}h_{T_{\boldsymbol{p}}^l}^{d-2}\left|v_l(\boldsymbol{p})\right|^2,$$

where $T_{\boldsymbol{p}}^l$ is the element in (5.3). Combining the above estimate and (5.14) yields

$$\begin{aligned} \sum_{l=1}^{L}\sum_{\boldsymbol{p}\in\widetilde{\mathcal{N}}_l} \left\|v_{\boldsymbol{p}}^l\right\|_A^2 &\le C\sum_{l=1}^{L}\sum_{\boldsymbol{p}\in\widetilde{\mathcal{N}}_l} \rho_{T_{\boldsymbol{p}}^l}h_{T_{\boldsymbol{p}}^l}^{d-2}|(\Pi_l - \Pi_{l-1})v(\boldsymbol{p})|^2 \\ &= C\sum_{l=1}^{L}\sum_{\boldsymbol{p}\in\widetilde{\mathcal{N}}_l} \rho_{T_{\boldsymbol{p}}^l}h_{T_{\boldsymbol{p}}^l}^{d-2}\left|(\Pi_l - \Pi_{l-1})\left(v - \widehat{Q}_{n(l,\boldsymbol{p})}^{\rho}v\right)(\boldsymbol{p})\right|^2. \end{aligned}$$

Set $w = v - \widehat{Q}_{n(l,\boldsymbol{p})}^{\rho}v$ for convenience. Then the definition of the quasi-interpolation operators (5.1)–(5.3) yields

$$|\Pi_l w(\boldsymbol{p})| \le \Big|\int_{T_{\boldsymbol{p}}^l}\psi_{\boldsymbol{p}}^{T_{\boldsymbol{p}}^l}(\boldsymbol{x})w(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\Big|, \qquad |\Pi_{l-1}w(\boldsymbol{p})| \le \sum_{\boldsymbol{q}\in S_{\boldsymbol{p}}}\Big|\int_{T_{\boldsymbol{q}}^{l-1}}\psi_{\boldsymbol{q}}^{T_{\boldsymbol{q}}^{l-1}}(\boldsymbol{x})w(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\Big|,$$

where $S_{\boldsymbol{p}} = \{\boldsymbol{q} \ : \ \boldsymbol{q} \in \widetilde{\mathcal{N}}_l \cap \mathcal{N}_{l-1}, \ \boldsymbol{p} \in \text{interior}(\Omega_{\boldsymbol{q}}^{l-1})\}$. Then using (H1) and (H2) we have

$$\begin{aligned}
&\rho_{T_{\boldsymbol{p}}^l} h_{T_{\boldsymbol{p}}^l}^{d-2} \left|(\Pi_l - \Pi_{l-1}) w(\boldsymbol{p})\right|^2 \\
\leq & C h_{T_{\boldsymbol{p}}^l}^{d-2} \left\{ \left|T_{\boldsymbol{p}}^l\right|^{-1} \|w\|_{L_\rho^2(T_{\boldsymbol{p}}^l)}^2 + \sum_{\boldsymbol{q}\in S_{\boldsymbol{p}}} \left|T_{\boldsymbol{q}}^{l-1}\right|^{-1} \|w\|_{L_\rho^2(T_{\boldsymbol{q}}^{l-1})}^2 \right\} \\
\leq & C h_{T_{\boldsymbol{p}}^l}^{-2} \|w\|_{L_\rho^2(D_{\boldsymbol{p}}^l)}^2 \leq C 2^{2n(l,\boldsymbol{p})} h_0^{-2} \|w\|_{L_\rho^2(D_{\boldsymbol{p}}^l)}^2 ,
\end{aligned}$$

where the constant $C$ depends on the integer $z$ in (H2) and $D_{\boldsymbol{p}}^l$ is the union of elements in $\mathcal{T}_l(\boldsymbol{p})$. For any fixed $m \geq 0$, the sub-domains in $\{D_{\boldsymbol{p}}^l : \ 1 \leq l \leq L, \ \boldsymbol{p} \in \widetilde{\mathcal{N}}_l, \ n(l,\boldsymbol{p}) = m\}$ are locally overlapping and their diameters are of the order $2^{-m} h_0$. Thus the union of these domains is also a subset of $\Omega$. It follows that

$$\begin{aligned}
&\sum_{l=1}^{L} \sum_{\boldsymbol{p}\in\widetilde{\mathcal{N}}_l} \left\|v_{\boldsymbol{p}}^l\right\|_A^2 \leq C \sum_{l=1}^{L} \sum_{\boldsymbol{p}\in\widetilde{\mathcal{N}}_l} 4^{n(l,\boldsymbol{p})} \left\|v - \widehat{Q}_{n(l,\boldsymbol{p})}^\rho v\right\|_{L_\rho^2(D_{\boldsymbol{p}}^l)}^2 \\
\leq & C \sum_{m=0}^{\widehat{L}} 4^m \sum_{l=1}^{L} \sum_{\substack{\boldsymbol{p}\in\widetilde{\mathcal{N}}_l, \\ n(l,\boldsymbol{p})=m}} \left\|v - \widehat{Q}_m^\rho v\right\|_{L_\rho^2(D_{\boldsymbol{p}}^l)}^2 \leq C \sum_{m=0}^{\widehat{L}} 4^m \left\|v - \widehat{Q}_m^\rho v\right\|_{L_\rho^2(\Omega)}^2 ,
\end{aligned}$$

where $\widehat{L} = \max\{n(l,\boldsymbol{p}) : \boldsymbol{p} \in \widetilde{\mathcal{N}}_l, \ 1 \leq l \leq L\}$, and we have $\widehat{L} \leq C \left|\log h_{\min}\right|$. Recall the estimates for the weighted $L^2$-projection on quasi-uniform meshes (cf. [11], Lemma 3.1-3.3 in [49]) :

$$\begin{aligned}
\sum_{m=0}^{\widehat{L}} 4^m \left\|v - \widehat{Q}_m^\rho v\right\|_{L_\rho^2(\Omega)}^2 &\leq C\, C_d^h \|v\|_A^2 \qquad \forall v \in V_L, \\
\sum_{m=0}^{\widehat{L}} 4^m \left\|v - \widehat{Q}_m^\rho v\right\|_{L_\rho^2(\Omega)}^2 &\leq C \left|\log h_{\min}\right|^2 \|v\|_A^2 \qquad \forall v \in \widetilde{V}_L.
\end{aligned}$$

This concludes the proof. □

In [50], it is proved that any $v \in V_L$ admits a multilevel decomposition

$$v = \widetilde{v}_0 + \sum_{l=1}^{L} \sum_{\boldsymbol{p}\in\widetilde{\mathcal{N}}_l} \widetilde{v}_{\boldsymbol{p}}^l, \quad \widetilde{v}_0 \in V_0, \quad \widetilde{v}_{\boldsymbol{p}}^l \in \text{span}\{\phi_{\boldsymbol{p}}^l\}$$

satisfying

$$\|\widetilde{v}_0\|_A^2 + \sum_{l=1}^{L} \sum_{\boldsymbol{p}\in\widetilde{\mathcal{N}}_l} \left\|\widetilde{v}_{\boldsymbol{p}}^l\right\|_A^2 \leq C \mathcal{J}(\rho) \|v\|_A^2 . \tag{5.15}$$

Clearly assumption (A1) follows from (5.15), Lemma 5.1, and Lemma 5.3.

### 5.4. Global strengthened Cauchy-Schwarz inequality

The strengthened Cauchy-Schwarz inequality has been established in [43] on quasi-uniform meshes. On adaptively refined meshes we need to establish a global strengthened Cauchy-Schwarz inequality. The following proof is different from [46] and [50] and does not elaborate on the meshes.

**Lemma 5.4.** *There exists a constant $C > 0$ only depending on the shape regularity of the meshes such that, for any functions*

$$v_i^l, w_i^l \in V_i^l, \qquad 1 \le i \le \widetilde{n}_l,\ 1 \le l \le L,$$

*the global strengthened Cauchy-Schwarz inequality holds*

$$\sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l}\sum_{k=1}^{l-1}\sum_{j=1}^{\widetilde{n}_k} a(v_i^l, w_j^k) \le C\Big(\sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l}\|v_i^l\|_A^2\Big)^{\frac{1}{2}}\Big(\sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l}\|w_i^l\|_A^2\Big)^{\frac{1}{2}}.$$

*Proof.* For convenience we introduce the generation $\mathcal{G}(T)$ of an element $T$ by the number of bisections for generating $T$ from one element in $\mathcal{T}_0$. It is reasonable to assume that

$$C_0\theta^m \le h_T \le C_1\theta^m, \quad m = \mathcal{G}(T), \quad \forall\, T \in \bigcup_{l=0}^{L}\mathcal{T}_l,$$

where $0 < \theta < 1$ is a constant that only depends on $\mathcal{T}_0$ and the shape regularity of the meshes. For the bisection algorithm that we are considering, $\theta \approx 2^{\frac{1}{1-2d}}$.

Then, we have

$$\begin{aligned} I_0 :=& \sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l}\sum_{k=1}^{l-1}\sum_{j=1}^{\widetilde{n}_k} a(v_i^l, w_j^k) \\ =& \sum_{l=1}^{L}\sum_{k=1}^{l-1}\sum_{m,n=0}^{\infty}\sum_{\substack{T\in\mathcal{T}_l\setminus\mathcal{T}_{l-1}\\ \mathcal{G}(T)=m}}\sum_{\substack{K\in\mathcal{T}_k\setminus\mathcal{T}_{k-1}\\ \mathcal{G}(K)=n}}\sum_{\substack{\boldsymbol{p}\in\mathcal{N}(T),\\ \boldsymbol{q}\in\mathcal{N}(K)}} a(\widetilde{v}_{\boldsymbol{p}}^l, \widetilde{w}_{\boldsymbol{q}}^k), \end{aligned} \tag{5.16}$$

where $\mathcal{N}(T)$ is the set of vertices of $T$ and

$$\widetilde{v}_{\boldsymbol{p}}^l = \begin{cases} v_{\boldsymbol{p}}^l/N_l(\boldsymbol{p}), & \text{if } \boldsymbol{p} \in \widetilde{\mathcal{N}}_l\ , \\ 0, & \text{otherwise,} \end{cases}$$

and $N_l(\boldsymbol{p})$ is the number of elements contained in $\mathcal{T}_l \setminus \mathcal{T}_{l-1}$ which share $\boldsymbol{p} \in \widetilde{\mathcal{N}}_l$. We note that $\widetilde{w}_{\boldsymbol{q}}^k$ is defined analogously. Suppose $m \le n$ and set

$$\widetilde{w}_n := \sum_{k=1}^{l-1}\sum_{\substack{K\in\mathcal{T}_k\setminus\mathcal{T}_{k-1}\\ \mathcal{G}(K)=n}}\sum_{\boldsymbol{q}\in\mathcal{N}(K)} \widetilde{w}_{\boldsymbol{q}}^k.$$

For any $T \in \mathcal{T}_l \setminus \mathcal{T}_{l-1}, \mathcal{G}(T) = m \le n, \boldsymbol{p} \in \mathcal{N}(T)$, we can derive that

$$a(\widetilde{v}_{\boldsymbol{p}}^l, \widetilde{w}_n) \le C\theta^{\frac{n-m}{2}}\|\nabla\widetilde{v}_{\boldsymbol{p}}^l\|_{L^2_\rho(\Omega_{\boldsymbol{p}}^l)}\|\nabla\widetilde{w}_n\|_{L^2_\rho(\Omega_{\boldsymbol{p}}^l)}. \tag{5.17}$$

Indeed, there exists a constant $t_0$ depending only on the shape regularity of the meshes such that

$$\max_{T'\in\mathcal{T}_l, T'\subset\Omega_{\boldsymbol{p}}^l}\mathcal{G}(T') \le \min_{T'\in\mathcal{T}_l, T'\subset\Omega_{\boldsymbol{p}}^l}\mathcal{G}(T') + t_0.$$

If $n - m \le t_0$, (5.17) holds true by the Cauchy-Schwarz inequality. For the case $n - m > t_0$, we note that $\widetilde{w}_n$ is piecewise linear in any $T' \in \mathcal{T}_l, T' \subset \Omega_{\boldsymbol{p}}^l$ and set

$$\widetilde{w}_n = \xi_n := \sum_{k=1}^{l-1}\sum_{\substack{K\in\mathcal{T}_k\setminus\mathcal{T}_{k-1}\\ \mathcal{G}(K)=n}}\sum_{\boldsymbol{q}\in\mathcal{N}(K)\cap\partial T'} \widetilde{w}_{\boldsymbol{q}}^k \qquad \text{on } \partial T'.$$

It is clear that

$$\operatorname{supp}(\xi_n) \cap T' \subset \Gamma_{T'} := \bigcup \{K \in \widehat{\mathcal{T}}_n : \ K \subset T' \ \text{ and } \ \partial K \cap \partial T' \neq \emptyset\}$$

is a narrow strip along the boundary of $T'$. Since $\widetilde{v}^l_{\boldsymbol{p}}$ is linear in $T'$, using Green's formula we have

$$\int_{T'} \rho \nabla \widetilde{v}^l_{\boldsymbol{p}} \cdot \nabla \widetilde{w}_n = \int_{\partial T'} \rho \frac{\partial \widetilde{v}^l_{\boldsymbol{p}}}{\partial \boldsymbol{n}} \widetilde{w}_n = \int_{\partial T'} \rho \frac{\partial \widetilde{v}^l_{\boldsymbol{p}}}{\partial \boldsymbol{n}} \xi_n = \int_{T' \cap \Gamma_{T'}} \rho \nabla \widetilde{v}^l_{\boldsymbol{p}} \cdot \nabla \xi_n$$
$$\leq |\rho_{T'}| \, \|\nabla \widetilde{v}^l_{\boldsymbol{p}}\|_{L^2(\Gamma_{T'})} \|\nabla \xi_n\|_{L^2(\Gamma_{T'})} \leq C \theta^{\frac{n-m}{2}} \left\|\nabla \widetilde{v}^l_{\boldsymbol{p}}\right\|_{L^2_\rho(T')} \|\nabla \xi_n\|_{L^2_\rho(T')} \, .$$

Summing over all $T' \subset \Omega^l_{\boldsymbol{p}}$ gives (5.17). Applying (5.17) and the local overlapping of the supports of $\widetilde{w}^k_{\boldsymbol{q}}$ and $\widetilde{v}^l_{\boldsymbol{p}}$, we have

$$I_1 := \sum_{m=0}^{\infty} \sum_{n=m}^{\infty} \sum_{l=1}^{L} \sum_{\substack{T \in \mathcal{T}_l \setminus \mathcal{T}_{l-1} \\ \mathcal{G}(T)=m}} \sum_{\boldsymbol{p} \in \mathcal{N}(T)} a(\widetilde{v}^l_{\boldsymbol{p}}, \sum_{k=1}^{l-1} \sum_{\substack{K \in \mathcal{T}_k \setminus \mathcal{T}_{k-1} \\ \mathcal{G}(K)=n}} \sum_{\boldsymbol{q} \in \mathcal{N}(K)} \widetilde{w}^k_{\boldsymbol{q}})$$
$$\leq C \sum_{m=0}^{\infty} \sum_{n=m}^{\infty} \theta^{\frac{n-m}{2}} \sum_{l=1}^{L} \sum_{\substack{T \in \mathcal{T}_l \setminus \mathcal{T}_{l-1} \\ \mathcal{G}(T)=m}} \sum_{\boldsymbol{p} \in \mathcal{N}(T)} \left\|\nabla \widetilde{v}^l_{\boldsymbol{p}}\right\|_{L^2_\rho(\Omega)} \cdot \Big( \sum_{k=1}^{l-1} \sum_{\substack{K \in \mathcal{T}_k \setminus \mathcal{T}_{k-1} \\ \mathcal{G}(K)=n}} \sum_{\boldsymbol{q} \in \mathcal{N}(K)} \left\|\nabla \widetilde{w}^k_{\boldsymbol{q}}\right\|^2_{L^2_\rho(\Omega^l_{\boldsymbol{p}})} \Big)^{\frac{1}{2}}.$$

It is known that the matrix $\left(\theta^{|m-n|/2}\right)_{m,n=0}^{\infty}$ has a finite radius of the spectrum depending only on $\theta$. Thus,

$$I_1 \leq C \sum_{m=0}^{\infty} \sum_{n=m}^{\infty} \theta^{\frac{n-m}{2}} \Big( \sum_{l=1}^{L} \sum_{\substack{T \in \mathcal{T}_l \setminus \mathcal{T}_{l-1} \\ \mathcal{G}(T)=m}} \sum_{\boldsymbol{p} \in \mathcal{N}(T)} \left\|\nabla \widetilde{v}^l_{\boldsymbol{p}}\right\|^2_{L^2_\rho(\Omega)} \Big)^{\frac{1}{2}} \cdot \Big( \sum_{k=1}^{L} \sum_{\substack{K \in \mathcal{T}_k \setminus \mathcal{T}_{k-1} \\ \mathcal{G}(K)=n}} \sum_{\boldsymbol{q} \in \mathcal{N}(K)} \left\|\nabla \widetilde{w}^k_{\boldsymbol{q}}\right\|^2_{L^2_\rho(\Omega)} \Big)^{\frac{1}{2}}$$
$$\leq C \Big( \sum_{m=0}^{\infty} \sum_{l=1}^{L} \sum_{\substack{T \in \mathcal{T}_l \setminus \mathcal{T}_{l-1} \\ \mathcal{G}(T)=m}} \sum_{\boldsymbol{p} \in \mathcal{N}(T)} \left\|\nabla \widetilde{v}^l_{\boldsymbol{p}}\right\|^2_{L^2_\rho(\Omega)} \Big)^{\frac{1}{2}} \cdot \Big( \sum_{n=0}^{\infty} \sum_{k=1}^{L} \sum_{\substack{K \in \mathcal{T}_k \setminus \mathcal{T}_{k-1} \\ \mathcal{G}(K)=n}} \sum_{\boldsymbol{q} \in \mathcal{N}(K)} \left\|\nabla \widetilde{w}^k_{\boldsymbol{q}}\right\|^2_{L^2_\rho(\Omega)} \Big)^{\frac{1}{2}}$$
$$\leq C \Big( \sum_{l=1}^{L} \sum_{i=1}^{\widetilde{n}_l} \left\|v^l_i\right\|^2_A \Big)^{\frac{1}{2}} \Big( \sum_{l=1}^{L} \sum_{i=1}^{\widetilde{n}_l} \left\|w^l_i\right\|^2_A \Big)^{\frac{1}{2}}. \tag{5.18}$$

If $m > n$, the same arguments show that the remaining terms $I_0 - I_1$ of the left hand side of (5.16) can also be bounded as follows

$$I_0 - I_1 \leq C \Big( \sum_{l=1}^{L} \sum_{i=1}^{\widetilde{n}_l} \left\|v^l_i\right\|^2_A \Big)^{\frac{1}{2}} \Big( \sum_{l=1}^{L} \sum_{i=1}^{\widetilde{n}_l} \left\|w^l_i\right\|^2_A \Big)^{\frac{1}{2}}. \tag{5.19}$$

Inserting (5.18) and (5.19) into (5.16) yields the stated result. This completes the proof. □

Now we come to the property (A2) in the previous section.

**Theorem 5.1.** *There exists a constant $C > 0$ only depending on the shape regularity of the meshes such that for any functions*

$$v^l_i, w^l_i \in V^l_i, \qquad 1 \leq i \leq \widetilde{n}_l, \ 0 \leq l \leq L,$$

*the global strengthened Cauchy-Schwarz inequality holds*

$$\sum_{l=0}^{L} \sum_{i=1}^{\widetilde{n}_l} \sum_{k=0}^{l-1} \sum_{j=1}^{\widetilde{n}_k} a(v^l_i, w^k_j) \leq C \Big( \sum_{l=0}^{L} \sum_{i=1}^{\widetilde{n}_l} \|v^l_i\|^2_A \Big)^{\frac{1}{2}} \Big( \sum_{l=0}^{L} \sum_{i=1}^{\widetilde{n}_l} \|w^l_i\|^2_A \Big)^{\frac{1}{2}}.$$

*Proof.* Note that

$$\sum_{l=0}^{L}\sum_{i=1}^{\widetilde{n}_l}\sum_{k=0}^{l-1}\sum_{j=1}^{\widetilde{n}_k} a(v_i^l, w_j^k) = \sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l}\sum_{k=1}^{l-1}\sum_{j=1}^{\widetilde{n}_k} a(v_i^l, w_j^k) + \sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l} a(v_i^l, w_1^0). \tag{5.20}$$

Since the supports of $\{v_i^l : 1, \ldots, \widetilde{n}_l\}$ are locally overlapped, an application of Lemma 5.4 shows that

$$\Big\| \sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l} v_i^l \Big\|_A^2 = 2\sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l}\sum_{k=1}^{l-1}\sum_{j=1}^{\widetilde{n}_k} a(v_i^l, v_j^k) + \sum_{l=1}^{L} \Big\| \sum_{i=1}^{\widetilde{n}_l} v_i^l \Big\|_A^2 \le C \sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l} \|v_i^l\|_A^2 .$$

We complete the proof by combining the above estimate, (5.20), Lemma 5.4 and the Cauchy-Schwarz inequality. □

## 6. Numerical Results

We present several numerical examples to demonstrate our convergence theory of multilevel methods. The implementation is based on the FFW toolbox [13] and the adaptive finite element package ALBERTA [38], [39].

Table 6.1: Example 6.1: Average error reduction factor and the number of iterations of PCG.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $R = 1.0$ | Level | | 6 | 7 | 8 | 9 | 10 | 11 |
| | DOFs | | 10153 | 22745 | 48440 | 101376 | 199012 | 408490 |
| | LMMA-PCG | $\alpha$ | 0.0907 | 0.0960 | 0.0860 | 0.0937 | 0.0849 | 0.0885 |
| | | **iter** | 6 | 6 | 6 | 6 | 6 | 6 |
| | SLMAA-PCG | $\alpha$ | 0.4743 | 0.4802 | 0.4744 | 0.4996 | 0.4893 | 0.5056 |
| | | **iter** | 19 | 19 | 18 | 20 | 19 | 20 |
| $R = 10^4$ | Level | | 7 | 9 | 10 | 12 | 14 | 16 |
| | DOFs | | 28811 | 69568 | 94270 | 128905 | 169872 | 220619 |
| | LMMA-PCG | $\alpha$ | 0.2656 | 0.3177 | 0.3349 | 0.3852 | 0.4311 | 0.4598 |
| | | **iter** | 12 | 13 | 14 | 16 | 17 | 19 |
| | SLMAA-PCG | $\alpha$ | 0.6324 | 0.7125 | 0.7311 | 0.7686 | 0.7928 | 0.8092 |
| | | **iter** | 33 | 46 | 47 | 55 | 61 | 67 |
| $R = 10^6$ | Level | | 7 | 9 | 10 | 12 | 14 | 15 |
| | DOFs | | 28745 | 73571 | 96955 | 137204 | 196927 | 224420 |
| | LMMA-PCG | $\alpha$ | 0.2356 | 0.2805 | 0.3006 | 0.3509 | 0.3854 | 0.4007 |
| | | **iter** | 13 | 14 | 15 | 15 | 18 | 19 |
| | SLMAA-PCG | $\alpha$ | 0.6339 | 0.6886 | 0.7033 | 0.7445 | 0.7665 | 0.7755 |
| | | **iter** | 40 | 48 | 49 | 54 | 63 | 66 |
| $R = 10^8$ | Level | | 7 | 9 | 10 | 12 | 14 | 15 |
| | DOFs | | 28744 | 73533 | 96913 | 139119 | 182107 | 208732 |
| | LMMA-PCG | $\alpha$ | 0.2140 | 0.2521 | 0.2748 | 0.3161 | 0.3515 | 0.3688 |
| | | **iter** | 14 | 15 | 16 | 17 | 18 | 19 |
| | SLMAA-PCG | $\alpha$ | 0.6163 | 0.6723 | 0.6831 | 0.7240 | 0.7494 | 0.7594 |
| | | **iter** | 43 | 51 | 53 | 59 | 66 | 69 |

In real computations, we have used the newest vertex bisection algorithm and the local error estimator defined in [20]. Given a finite element approximation $u_h$, for any $T \in \mathcal{T}_h$, the a posteriori error estimator is defined as

$$\eta_T^2 := h_T^2 \Lambda_T \|\rho_T^{-\frac{1}{2}} f\|_{L^2(T)}^2 + \frac{h_T}{2} \sum_{F \subset \partial T} \Lambda_F \|\rho_F^{-\frac{1}{2}} [\![\rho \nabla u_h]\!] \cdot \nu\|_{L^2(F)}^2, \tag{6.1}$$

where $F$ is a face of $T$ if $d = 3$, and $F$ is an edge of $T$ if $d = 2$, $[\![\rho \nabla u_h]\!]$ is the jump of $\rho \nabla u_h$ across $F$. The parameters $\Lambda_T, \Lambda_F, \rho_F$ in (6.1) are given by

$$\Lambda_T = \begin{cases} \max_{T' \in \Omega_T} \{\frac{\rho_T}{\rho_{T'}}\}, & \text{if } T \text{ has one singular node (cf. [20]),} \\ 1, & \text{otherwise,} \end{cases}$$

$\Lambda_F = \max_{T \in \Omega_F} \{\Lambda_T\}$, $\rho_F = \max_{T \subset \Omega_F} \{\rho_T\}$, where $\Omega_T = \{T' \in \mathcal{T}_h \ : \ \overline{T'} \cap \overline{T} \neq \emptyset\}$ and $\Omega_F = \{T \in \mathcal{T}_h \ : \ \partial T \cap F \neq \emptyset\}$. The global a posteriori error estimator on $\mathcal{T}_h$ is defined by

$$\eta_h := \Big( \sum_{T \in \mathcal{T}_h} \eta_T^2 \Big)^{\frac{1}{2}}.$$

Based on the above a posteriori error estimator and the AFEM algorithm in [16], we can mark and refine $\mathcal{T}_h$ adaptively.

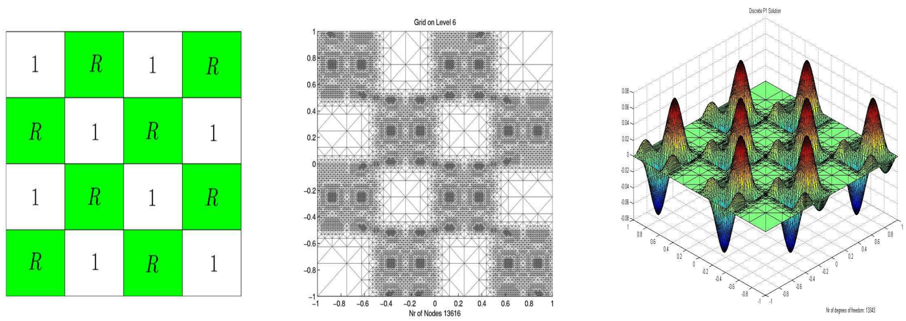


Fig. 6.1. The distribution of $\rho$ (left). A locally refined mesh of $\Omega$ (middle). The surface plot of the discrete solution (right).

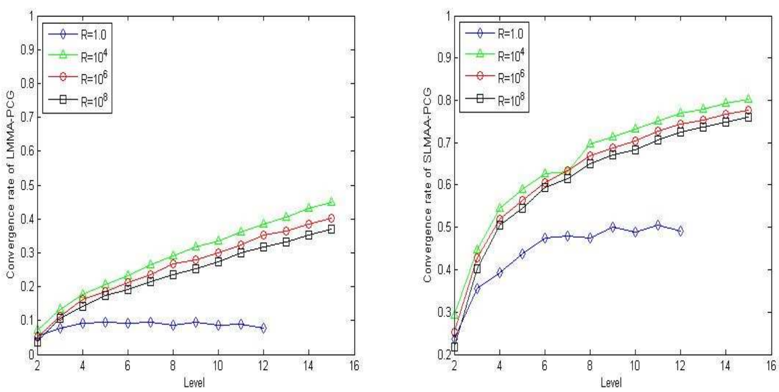


Fig. 6.2. Average error reduction factor of LMMA-PCG (left) and SLMAA-PCG (right).

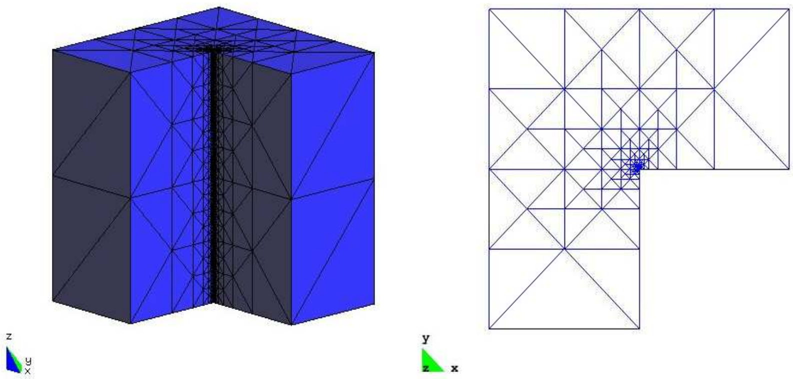

Fig. 6.3. A locally refined mesh with $1,537,132$ elements for the case of $\epsilon = 10^{-6}$.

In the following experiments, Algorithm LMMA and LMAA are mainly used as preconditioners for the conjugate gradient method. Let the discrete problem on $\mathcal{T}_L$ be

$$\mathbf{A}_L \mathbf{U}_L = \mathbf{F}_L.$$

We set the initial guess $\mathbf{U}_L^0$ by the solution of the previous level, i.e., $\mathbf{U}_L^0 = \mathbf{I}_{L-1}\mathbf{U}_{L-1}$, where $\mathbf{I}_{L-1} : \mathbb{R}^{N_{L-1}} \mapsto \mathbb{R}^{N_L}$ is the transfer matrix. Let $\mathbf{r}^k = \mathbf{F}_L - \mathbf{A}_L \mathbf{U}_L^k$ be the residual of the equation at the $k$-th iteration. The PCG algorithm stops when

$$\|\mathbf{r}^k\|/\|\mathbf{r}^0\| \leq 10^{-6}, \tag{6.2}$$

where $\|\mathbf{v}\|$ is the $l^2$-norm of the vector $\mathbf{v}$. We define the average error reduction factor of the PCG algorithm by

$$\alpha = (\sqrt{e_k}/\sqrt{e_0})^{1/\mathbf{iter}},$$

where **iter** is the number of iterations required to achieve (6.2) and

$$e_0 = (\mathbf{r}^0)^t \mathcal{B}_L \mathbf{r}^0, \quad e_k = (\mathbf{r}^k)^t \mathcal{B}_L \mathbf{r}^k, \quad k \geq 1.$$

Here $\mathcal{B}_L$ can be any of the local multilevel algorithms in Algorithm 3.1–3.3. We shall use local Gauss-Seidel smoothers in Algorithm 3.1–3.3 for all the examples.

**Example 6.1.** We consider (1.1)–(1.2) in two dimensions with

$$f = 2\pi^2 \sin(4\pi x_1)\cos(4\pi x_2), \qquad \Omega = (-1,1)\times(-1,1).$$

The coefficient $\rho$ is piecewise constant and has a checkerboard distribution on $\Omega$, where $R$ is a positive constant (see Figure 6.1).

In Fig. 6.1, the left picture shows the distribution of the coefficient $\rho$ which takes value 1 in the white regions and value $R$ in the shadow regions. The middle picture shows a locally refined mesh at the 6-th adaptive iteration for $R = 10^6$, and the right picture shows a surface plot of the associated discrete solution. We find that the mesh is refined considerably in the regions where the solution is rapidly varying.

In Fig. 6.2 and Table 6.1, the reduction factors and the number of iterations of algorithms LMMA-PCG and SLMAA-PCG are shown for different coefficients $R = 10^i, i = 0, 4, 6, 8$. When

Table 6.2: Example 6.2: Average error reduction factor and the number of iterations of LMMA and LMMA-PCG.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $\epsilon = 10^{-4}$ | Level | | 8 | 9 | 10 | 11 | 12 | 13 |
| | $N_{el}$ | | 48572 | 96612 | 193596 | 385880 | 770316 | 1537432 |
| | LMMA | $\alpha$ | 0.7555 | 0.7847 | 0.8089 | 0.8289 | 0.8457 | 0.8603 |
| | | **iter** | 30 | 33 | 37 | 40 | 44 | 48 |
| | LMMA -PCG | $\alpha$ | 0.2963 | 0.3255 | 0.3472 | 0.3758 | 0.4044 | 0.4200 |
| | | **iter** | 12 | 13 | 15 | 16 | 17 | 17 |
| $\epsilon = 10^{-6}$ | Level | | 8 | 9 | 10 | 11 | 12 | 13 |
| | $N_{el}$ | | 48572 | 96612 | 193596 | 385880 | 770316 | 1537132 |
| | LMMA | $\alpha$ | 0.7556 | 0.7848 | 0.8089 | 0.8290 | 0.8458 | 0.8600 |
| | | **iter** | 30 | 33 | 37 | 40 | 44 | 48 |
| | LMMA -PCG | $\alpha$ | 0.2964 | 0.3256 | 0.3472 | 0.3759 | 0.4045 | 0.4271 |
| | | **iter** | 12 | 13 | 15 | 16 | 17 | 19 |
| $\epsilon = 10^{-8}$ | Level | | 8 | 9 | 10 | 11 | 12 | 13 |
| | $N_{el}$ | | 48572 | 96612 | 193596 | 385880 | 770316 | 1537132 |
| | LMMA | $\alpha$ | 0.7556 | 0.7848 | 0.8089 | 0.8290 | 0.8458 | 0.8600 |
| | | **iter** | 30 | 33 | 37 | 40 | 44 | 48 |
| | LMMA -PCG | $\alpha$ | 0.2964 | 0.3256 | 0.3472 | 0.3759 | 0.4045 | 0.4271 |
| | | **iter** | 12 | 13 | 15 | 16 | 17 | 19 |

$R = 1$, both algorithms show uniform convergence with respect to mesh sizes and mesh levels. When $R = 10^i, i = 4, 6, 8$, the convergence rates of LMMA-PCG and SLMAA-PCG increase slightly with respect to the number of mesh levels. However we can see that the convergence rates for these three cases are almost the same regardless the jumps of $\rho$. The convergence rates agree well with our theoretical results, i.e. $1 - \frac{2}{C|\log h_{\min}|+1}$. From Table 6.1, we also note that the multiplicative algorithm LMMA-PCG performs much better than the additive algorithm SLMAA-PCG.

**Example 6.2.** We consider (1.1) with an inhomogeneous boundary condition. Here $\Omega$ is the "L-shaped" domain

$$\Omega = (-1, 1)^3 \ \setminus (0, 1) \times (-1, 0) \times (-1, 1).$$

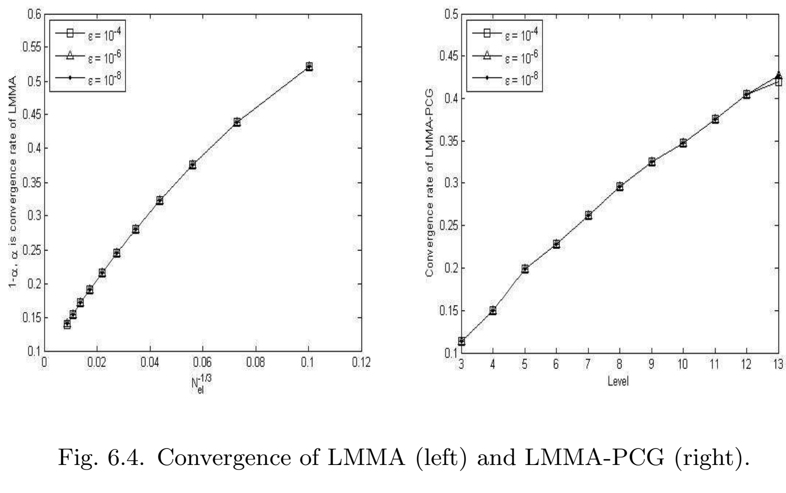


Fig. 6.4. Convergence of LMMA (left) and LMMA-PCG (right).

The coefficient function is defined by

$$\rho(\boldsymbol{x}) = \begin{cases} \epsilon, & \text{if } \boldsymbol{x} \in (0,1)\times(0,1)\times(-1,1)\bigcup(-1,0)\times(-1,0)\times(-1,1), \\ 1, & \text{elsewhere.} \end{cases}$$

The Dirichlet boundary condition and the right-hand side $f$ are chosen such that the exact solution is $u = r^{2/3}\sin(\frac{2}{3}\theta)$ in the cylindrical coordinates $(r,\theta,z)$.

Fig. 6.4 and Table 6.2 show that the convergence rate $\alpha$ of LMMA is uniform with respect the choices of $\epsilon$ or jumps of the coefficient. We also observe that $1-\alpha \propto N_{el}^{-1/3}$ where $N_{el}$ is the number of elements of the underlying mesh. We also note that the LMMA-PCG converges much faster than the LMMA. Figure 6.3 shows a locally refined mesh with $1,537,132$ elements for $\epsilon = 10^{-6}$ featuring pronounced local refinements near the reentrant corner.

**Example 6.3.** We consider (1.1) defined on a domain with an inner screen:

$$\Omega := (-1,1)^3 \setminus \Gamma, \quad \Gamma = \{(0,y,z):\ y,z \in [-1/3,\ 1/3]\}.$$

We choose the right-hand side according to $f = 1.0$ and consider the Dirichlet boundary condition by $u|_\Gamma = 0$, $u|_{\partial\Omega\setminus\Gamma} = 1.0$. The coefficient is defined as follows (cf. Figure 6.5):

$$\rho(\boldsymbol{x}) = \begin{cases} \epsilon, & \text{in } \bigcup_{i=1}^4 \Omega_i, \\ 1, & elsewhere, \end{cases}$$

where

$$\Omega_1 = (-1/3,\ -2/3)\times(0,\ 1/3)\times(0,\ 1/3),\ \Omega_2 = (-1/3,\ -2/3)\times(-1/3,\ 0)\times(-1/3,\ 0),$$
$$\Omega_3 = (1/3,\ 2/3)\times(0,\ 1/3)\times(0,\ 1/3),\ \Omega_4 = (1/3,\ 2/3)\times(-1/3,\ 0)\times(-1/3,\ 0).$$

Our computations show that the LMMA needs more than one thousand iterations to achieve (6.2) for $\epsilon \le 10^{-4}$. Thus the LMMA is unfavorable for this example and we only show the numerical results from the LMMA-PCG.

Fig. 6.6 displays four sections of a locally refined mesh with $1,154,472$ elements for $\epsilon = 10^{-6}$, three of which are at $x = 2/3, 0, -2/3$ and the other one is at $y = 0$. We observe that the mesh is locally refined near the boundary of the "screen" and the sub-domains $\Omega_1,\ldots,\Omega_4$. Table 6.3 shows the convergence results of the LMMA-PCG. Although LMMA shows an unpleasant

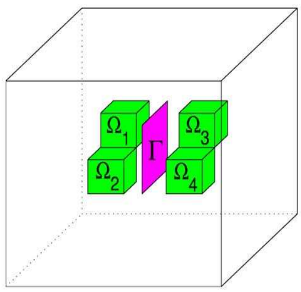


Fig. 6.5. The domain $\Omega$, subdomains $\Omega_1,\ldots,\Omega_4$, and the inner screen $\Gamma$.

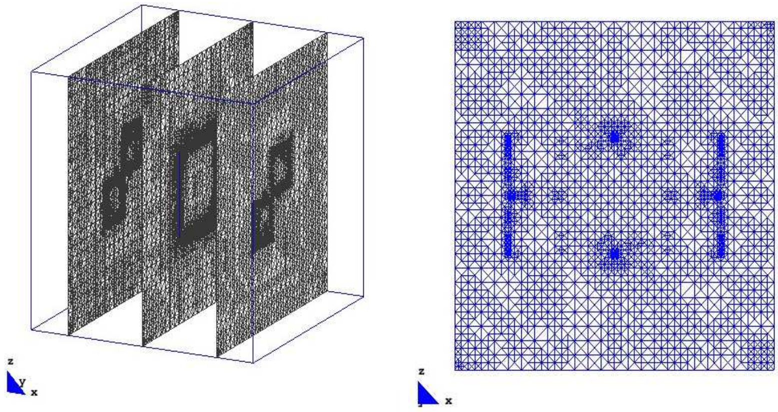

Fig. 6.6. A locally refined mesh with $1,154,472$ elements for $\epsilon = 10^{-6}$. Three sections at $x = 2/3, 0, -2/3$ (left). The section at $y = 0$ (right).

Table 6.3: Example 6.3: Average reduction factor and the number of iterations of LMMA-PCG.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $\epsilon = 10^{-2}$ | Level | | 4 | 6 | 8 | 10 | 12 | 13 |
| | $N_{el}$ | | 16944 | 45056 | 121944 | 266196 | 984020 | 1350936 |
| | LMMA | $\alpha$ | 0.2029 | 0.2396 | 0.2623 | 0.2571 | 0.2830 | 0.2967 |
| | -PCG | **iter** | 9 | 10 | 11 | 11 | 11 | 12 |
| $\epsilon = 10^{-4}$ | Level | | 4 | 6 | 8 | 10 | 12 | 13 |
| | $N_{el}$ | | 19460 | 38104 | 118468 | 321120 | 853716 | 1158312 |
| | LMMA | $\alpha$ | 0.3860 | 0.3800 | 0.4001 | 0.4361 | 0.4420 | 0.4649 |
| | -PCG | **iter** | 15 | 17 | 19 | 21 | 21 | 21 |
| $\epsilon = 10^{-6}$ | Level | | 4 | 6 | 8 | 10 | 12 | 13 |
| | $N_{el}$ | | 19532 | 38176 | 118932 | 320388 | 852004 | 1154472 |
| | LMMA | $\alpha$ | 0.4507 | 0.4380 | 0.4785 | 0.4969 | 0.4059 | 0.4313 |
| | -PCG | **iter** | 21 | 24 | 26 | 27 | 19 | 21 |
| $\epsilon = 10^{-8}$ | Level | | 4 | 6 | 8 | 10 | 12 | 13 |
| | $N_{el}$ | | 19532 | 38176 | 118932 | 320388 | 852004 | 1154472 |
| | LMMA | $\alpha$ | 0.5107 | 0.4837 | 0.5301 | 0.5685 | 0.4470 | 0.4565 |
| | -PCG | **iter** | 26 | 31 | 33 | 33 | 25 | 21 |

convergence behavior for $\epsilon \leq 10^{-4}$, it proves to be an efficient and robust preconditioner for the conjugate gradient method. This again justifies our theoretical analysis.

**Remark 6.1.** After the submission of this paper, we found another work on the same topic by Chen et al. [18] which appeared on the internet in June 2010. The two works are fully independent. The local multilevel method in [18] is based on the mesh hierarchy obtained by some coarsening strategy for bisection grids, while our method is based on adaptively refined meshes using a posteriori error estimates. This also results in different proofs for the uniform convergence of the multilevel method.

**Acknowledgements.** The work of the second author was supported by the National Basic Research Program under the Grant 2011CB30971 and National Science Foundation of China (11171335). The work of the third author was supported in part by China NSF under the grants 11031006 and 11171334, by the Funds for Creative Research Groups of China (Grant No. 11021101), and by the National Magnetic Confinement Fusion Science Program (Grant No.2011GB105003).

## References

[1] M. Ainsworth, Robust a posteriori error estimation for nonconforming finite element approximation, *SIAM J. Numer. Anal.,* **42** (2005), 320-2341.

[2] M. Ainsworth, A posteriori error estimation for lowest order Raviart-Thomas mixed finite elements, *SIAM J. Sci. Comput.,* **30** (2007), 189-204.

[3] O. Axelsson, Iteration number for the conjugate gradient method, *Math. Comput. Simulat.,* **61** (2003), 421-435.

[4] R.E. Bank and L.R. Scott, On the conditioning of finite element equations with highly refined meshes, *SIAM J. Numer. Anal.,* **26** (1989), 1383-1394.

[5] C. Bernardi and R. Verfürth, Adaptive finite element methods for elliptic equations with non-smooth coefficients, *Numer. Math.,* **85** (2000), 579-608.

[6] P. Binev, W. Dahmen, and R. DeVore, Adaptive finite element methods with convergence rates, *Numer. Math.,* **97** (2004), 219-268.

[7] D. Braess and W. Hackbusch, A new convergence proof for the multigrid method including the V-cycle, *SIAM J. Numer. Anal.,* **36** (1983), 967-975.

[8] J.H. Bramble, Multigrid Methods, Pitman, Boston, 1993.

[9] J.H. Bramble and J.E. Pasicak, New estimates for multigrid algorithms including the V-cycle, *Math. Comput.,* **60** (1993), 447-471.

[10] J.H. Bramble, J.E. Pasicak, J. Wang, and J. Xu, Convergence estimates for product iterative methods with applications to domain decomposition, *Math. Comput.,* **57** (1991), 23-45.

[11] J.H. Bramble and J. Xu, Some estimates for a weighted $L^2$ projection, *Math. Comput.,* **56** (1991), 463-476.

[12] S.C. Brenner, Convergence of the multigrid V-cycle algorithms for the second order boundary value problems without full elliptic regularity, *Math. Comput.,* **71** (2002), 507-525.

[13] A. Byfut, J. Gedicke, D. Günther, J. Reininghaus, S. Wiedemann et al., FFW Documentation, Humboldt University of Berlin, Germany, 2007.

[14] Z. Cai and S. Zhang, Recovery-based error estimator for interface problems: conforming linear elements, *SIAM J. Numer. Anal.,* **47** (2009), 2132-2156.

[15] Z. Cai and S. Zhang, Recovery-based error estimators for interface problems: mixed and nonconforming finite elements, *SIAM J. Numer. Anal.,* **48** (2010), 30-52.

[16] J.M. Cascon, C. Kreuzer, R.H. Nochetto, and K.G. Siebert, Quasi-optimal convergence rate for an adaptive finite element method, *SIAM J. Numer. Anal.,* **46** (2008), 2524-2550.

[17] T. Chan and W. Wan, Robust multigrid methods for nonsmooth coefficient elliptic linear systems, *J. Comput. Appl. Math.,* **123** (2000), 323-352.

[18] L. Chen, M. Holst, J. Xu, and Y. Zhu, Local multilevel preconditioners for elliptic equations with jump coefficients on bisection grids, Arxiv preprint, June, 2010, `http://arxiv.org/abs/1006.3277`.

[19] H. Chen, X. Xu, and R.H.W. Hoppe, Convergence and optimality of adaptive nonconforming finite element methods for nonsymmetric and indefinite problems, *Numer. Math.,* **116** (2010), 383-419.

[20] Z. Chen and S. Dai, On the efficiency of adaptive finite element methods for elliptic problems with discontinuous coeficients, *SIAM J. Sci. Comput.,* **24** (2001), 443-462.

[21] P.G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam, 1978.

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

[23] M. Dryja, M.V. Sarkis, and O.B. Widlund, Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions, *Numer. Math.,* **72** (1996), 313-348.

[24] M. Dryja and O.B. Widlund, Schwarz methods of neumann-neumann type for three-dimensional elliptic finite element problems, *Commun. Pur. Appl. Math.,* **48** (1995), 121-155.

[25] W. Hackbusch, Elliptic Differential Equations Theory and Numerical Treatment, Springer, Berlin-Heidelberg-New York, 2003.

[26] R. Hiptmair, H. Wu, and W. Zheng, Uniform convergence for adaptive multigrid methods in $H^1(\Omega)$ and $\mathbf{H}(\mathbf{curl},\Omega)$, preprint, 2010.

[27] R. Hiptmair and W. Zheng, Local multigrid in H(curl), *J. Comput. Math.*, **27** (2009), 573-603.

[28] R. Hiptmair and W. Zheng, Local multigrid in H(curl), LSEC Research Report 2007–04, `http://www.cc.ac.cn/2007research_ report.html`.

[29] R. Luce and B.I. Wohlmuth, A local a posteriori error estimator based on equilibrated fluxes, *SIAM J. Numer. Anal.,* **42** (2004), 1394-1414.

[30] J. Mandel and M. Brezina, Balancing domain decomposition for problems with large jumps in coefficients, *Math. Comput.,* **65** (1996), 1387-1401.

[31] W.F. Mitchell, Optimal multilevel iterative methods for adaptive grids, *SIAM J. Sci. Comput.,* **13** (1992), 146-167.

[32] P. Morin, R.H. Nochetto, and K.G. Siebert, Convergence of adaptive finite element methods, *SIAM Rev.,* **44** (2002), 631-658.

[33] P. Oswald, Multilevel Finite Element Approximation: Theory and Applications, Stuttgart: Teubner Verlag, 1994.

[34] P. Oswald, On the robustness of the BPX-preconditioner with respect to jumps in the coefficients, *Math. Comput.,* **68** (1999), 633-650.

[35] M. Petzoldt, A posteriori error estimators for elliptic equations with discontinuous coefficients, *Adv. Comput. Math.,* **16** (2002), 47-75.

[36] R. Stevenson, Optimality of a standard adaptive finite element method, *Found. Comput. Math.*, **7** (2007), 245-269.

[37] R. Scheichl and E. Vainikko, Additive Schwarz with aggregation-based coarsening for elliptic problems with highly variable coefficients, *Computing,* **80** (2007), 319-343.

[38] A. Schmidt and K. Siebert, ALBERTA-An adaptive hierarchical finite element toolbox, Website, ALBERTA is available online from `http://www.alberta-fem.de`.

[39] A. Schmidt and K. Siebert, Design of Adaptive Finite Element Software: The Finite Element Toolbox ALBERTA, Lecture Notes in Computational Science and Engineering, Springer, Heidelberg, 2005.

[40] B.F. Smith, A domain decomposition algorithm for elliptic problems in three dimensions, *Numer. Math.,* **60** (1991), 219-234.

[41] A. Toselli and O. Widlund, Domain Decomposition MethodsAlgorithms and Theory, Springer, Berlin, 2005.

[42] J. Xu, Counterexamples concerning a weighted $L^2$ projection, *Math. Comput.,* **57** (1991), 563-568.

[43] J. Xu, Iterative methods by space decomposition and subspace correction, *SIAM Rev.,* **34** (1992), 581-613.

[44] J. Xu and J. Zou, Some nonoverlapping domain decomposition methods, *SIAM Rev.,* **40** (1998), 857-914.

[45] J. Wang, New convergence estimates for multilevel algorithms for finite-element approximations, *J. Comput. Appl. Math.,* **50** (1994), 593-604.

[46] H.J. Wu and Z.M. Chen, Uniform convergence of multigrid V-cycle on adaptively refined finite element meshes for second order elliptic problems, *Science in China,* **39** (2006), 1405-1429.

[47] H.J. Wu and W. Zheng, Uniform convergence of multigrid V-cycle on adaptively refined finite element meshes for elliptic problems with discontinuous coefficients, submitted, 2010.

[48] 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, Springer, 2009, 599-659.

[49] J. Xu and Y. Zhu, Uniformly convergent multigrid methods for elliptic problems with strongly discontinuous coefficients, *Math. Mod. Meth. Appl. S.,* **18** (2008), 77-105.

[50] X. Xu, H. Chen, and R.H.W. Hoppe, Optimality of local multilevel methods on adaptively refined meshes for elliptic boundary value problems, *J. Numer. Math.,* **18** (2010), 59-90.

[51] Y. Zhu, Domain decomposition preconditioners for elliptic equations with jump coefficients, *Numer. Linear Algebra,* **15** (2008), 271-289.