Journal of Computational Mathematics
Vol.31, No.1, 2013, 22–46.

http://www.global-sci.org/jcm
doi:10.4208/jcm.1203-m3960

OPTIMALITY OF LOCAL MULTILEVEL METHODS FOR ADAPTIVE NONCONFORMING P1 FINITE ELEMENT METHODS*

Xuejun Xu
LSEC, ICMSEC, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing, 100190, China
Email: xxj@lsec.cc.ac.cn
Huangxin Chen
School of Mathematical Sciences, Xiamen University, Xiamen 361005, China
LSEC, ICMSEC, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing 100190, China
Email: chx@xmu.edu.cn
R.H.W. Hoppe
Department of Mathematics, University of Houston, Houston, TX 77204-3008, USA
Institute for Mathematicd, University of Augsburg, D-86159, Augsburg, Germany
Email: rohop@math.uh.edu

Abstract

In this paper, a local multilevel product algorithm and its additive version are considered for linear systems arising from adaptive nonconforming P1 finite element approximations of second order elliptic boundary value problems. The abstract Schwarz theory is applied to analyze the multilevel methods with Jacobi or Gauss-Seidel smoothers performed on local nodes on coarse meshes and global nodes on the finest mesh. It is shown that the local multilevel methods are optimal, i.e., the convergence rate of the multilevel methods is independent of the mesh sizes and mesh levels. Numerical experiments are given to confirm the theoretical results.

Mathematics subject classification: 65F10, 65N30.
Key words: Local multilevel methods, Adaptive nonconforming P1 finite element methods, Convergence analysis, Optimality.

1. Introduction

Multigrid methods and other multilevel preconditioning methods for nonconforming finite elements have been studied by many researchers (cf. [4–7,16,22–25,27,30–32,36,38]). The BPX framework developed in [4] provides a unified convergence analysis for nonnested multigrid methods. Duan *et al.* [16] extended the result to general V-cycle nonnested multigrid methods, but only the case of full elliptic regularity was considered. Besides, Brenner [7] established a framework for the nonconforming V-cycle multigrid method under less restrictive regularity assumptions. All the above convergence results for nonconforming multigrid methods are based on the requirement of a sufficiently large number of smoothing steps at each level. For multilevel preconditioning methods, Oswald developed a hierarchical basis multilevel method [23] and a BPX-type multilevel preconditioner [24] for nonconforming finite elements. On the other

* Received July 29, 2011 / Revised version received March 5, 2012 / Accepted March 28, 2012 / Published online January 17, 2013 /

hand, Vassilevski and Wang [30] presented multilevel algorithms with only one smoothing step per level. These multilevel algorithms may be considered as successive subspace correction methods (SSC) (cf. [34] for details). They are completely different from standard nonconforming multigrid methods [5]. By using the well-known Schwarz framework, a uniform convergence result has been obtained. The idea that the conforming finite element spaces are contained in their nonconforming counterparts is essential in the analysis of the multilevel algorithms (cf. [30] for details). In this paper, we will use this idea to design optimal multilevel methods for adaptive nonconforming P1 element methods (ANFEM). We note that Hoppe and Wohlmuth [18] considered multilevel preconditioned conjugate gradient methods for nonconforming P1 finite element approximations with respect to adaptively generated hierarchies of nonuniform meshes based on residual type a posteriori error estimators.

Recent studies (cf., e.g., [2, 11–13, 19, 21, 28]) indicate optimal convergence properties of adaptive conforming and nonconforming finite element methods. Therefore, in order to achieve an optimal numerical solution, it is imperative to study efficient iterative algorithms for the solution of linear systems arising from adaptive finite element methods (AFEM). Since the number of degrees of freedom N per level may not grow exponentially with mesh levels, as Mitchell has pointed out in [20] for adaptive conforming finite element methods, the number of operations used for multigrid methods with smoothers performed on all nodes can be as bad as $O(N^2)$, and a similar situation may also occur in the nonconforming case.

For adaptive conforming finite element methods, the optimality of local multilevel methods for 2D and 3D $H^1(\Omega)$-elliptic problems has been studied in [17, 33, 35, 37]. The hierarchy of meshes used in the local multilevel methods can be obtained either by successive adaptive refinement of an initial coarse mesh or by successive coarsening of a fine mesh. Wu and Chen [33] applied the adaptively refined hierarchy of meshes generated by the newest vertex bisection and obtained uniform convergence for the multigrid V-cycle algorithm with local Gauss-Seidel smoother in 2D. The optimal multigrid methods developed by Xu, Chen and Nochetto [35] are based on the reconstruction of hierarchy of meshes. There are some assumptions of this strategy on the initial mesh and the fine mesh to guarantee that the compatible patches of meshes do exist (cf. [35]). We do not reconstruct a virtual refinement hierarchy of meshes in our algorithms, but use the hierarchy generated by the ANFEM. We also note that Dahmen and Kunoth [15] proved the optimality of BPX preconditioner for piecewise linear finite elements on the quasi-uniform meshes and the nonuniform meshes generated by red-green refinement. Brenner, Cui and Sung [8] proved the uniform convergence of W-cycle multigrid algorithm with sufficiently large number of smoothing steps for the symmetric interior penalty method on graded meshes in 2D. To our knowledge, so far there does not exist an optimal multilevel method for nonconforming finite element methods on locally refined meshes. Indeed, there are two difficulties in the theoretical analysis which need to be overcome. First, since the multilevel spaces are nonnested in this situation, we should consider how to design a stable decomposition of the finest nonconforming finite element space. The second difficulty is how to establish the strengthened Cauchy-Schwarz inequality on nonnested multilevel spaces. In this paper, we will construct a special prolongation operator from the coarse space to the finest space, and obtain the key global strengthened Cauchy-Schwarz inequality. Two multilevel methods, the product and additive version, are proposed. Applying the well-known Schwarz theory (cf. [29]), we show that local multilevel methods for adaptive nonconforming P1 finite element methods are optimal.

The remainder of this paper is organized as follows: In Section 2, we introduce some notations and briefly review nonconforming P1 finite element methods. Section 3 is concerned with the study of condition number estimates of linear systems arising from the ANFEM by applying the techniques presented by Bank and Scott in [1]. In Section 4, we analyze the stability property of a proposed prolongation operator, and address the multilevel methods featuring local Jacobi and local Gauss-Seidel smoothers. The convergence theory of the local multilevel methods is developed in Section 5 within the abstract framework of the Schwarz theory, and we also present the detailed analysis for the two types of local smoothers. In the final Section 6, we give some numerical experiments to confirm the theoretical results.

2. Notations and Preliminaries

Throughout this paper, we adopt standard notation from Lebesgue and Sobolev spaces theory (cf., e.g., [14]). In particular, we refer to $(\cdot,\cdot)$ as the inner product in $L^2(\Omega)$ and to $\|\cdot\|_{1,\Omega}$ as the norm in the Sobolev space $H^1(\Omega)$. Let C, with or without subscript, denote a generic positive constant which is independent of mesh sizes and mesh levels, but depends on the shape regularity of the meshes. These constants can take on different values in different occurrences.

Given a bounded Lipschitz polyhedron $\Omega \subset \mathbb{R}^d, d = 2, 3$, we consider the following second order elliptic boundary value problem

$$-\mathrm{div}(a(x)\nabla u) = f \quad \text{in } \Omega, \tag{2.1}$$

$$u = 0 \quad \text{on } \partial\Omega, \tag{2.2}$$

where the source function $f \in L^2(\Omega)$. The choice of a homogeneous Dirichlet boundary condition is made for ease of presentation only. Similar results are valid for other types of boundary conditions and equation (2.1) with a lower order term as well. We further assume that the coefficient function in (2.1) satisfies the following property: $a(x) \in W^{1,\infty}(\Omega)$ is a measurable function and there exist constants $\beta_1 \geq \beta_0 > 0$ such that

$$\beta_0 \leq a(x) \leq \beta_1 \quad \text{f.a.a. } x \in \Omega. \tag{2.3}$$

The weak formulation of (2.1) and (2.2) is to find $u \in H_0^1(\Omega)$ such that

$$a(u, v) = (f, v), \quad v \in H_0^1(\Omega), \tag{2.4}$$

where the bilinear form $a : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ is given by

$$a(u, v) = \big(a(x)\nabla u, \nabla v\big), \quad u, v \in H_0^1(\Omega). \tag{2.5}$$

Since the bilinear form (2.5) is bounded and V-elliptic, the existence and uniqueness of the solution of (2.4) follow from the Lax-Milgram theorem.

Let $\{\mathcal{T}_l, l = 0, 1, \cdots, L\}$ be a shape regular family of nested geometrically conforming simplicial triangulations of the computational domain Ω obtained by successive refinement of an intentionally chosen coarse mesh $\mathcal{T}_0$ using newest vertex bisection algorithm. Let $\#S$ denote the cardinality of any set S. For any $T \in \mathcal{T}_l$, h_T refers to the diameter of T. We refer to $\mathcal{N}_l$ as the set of interior vertices of $\mathcal{T}_l$. The set of faces on $\mathcal{T}_l$ is denoted by $\mathcal{F}_l$. Let $\mathcal{F}_l^0$ be the set of interior faces of $\mathcal{T}_l$ and $\mathcal{M}_l$ be the set of all the barycenters of $\mathcal{F}_l^0$. We denote by V_l the lowest order Crouzeix-Raviart nonconforming finite element space with respect to $\mathcal{T}_l$, i.e.,

$$V_l = \{v_l \in L^2(\Omega) \ : \ v_l|_T \in P_1(T),\ T \in \mathcal{T}_l\ ,\ \int_F [\![v_l]\!] ds = 0, F \in \mathcal{F}_l\},$$

where $F \in \mathcal{F}_l$, $[\![v_l]\!]$ refers to the jump of v_l across $F \in \mathcal{F}_l^0$ and is set to be $v_l|_F$ for $F \subset \partial\Omega$. Moreover, we define the conforming P1 finite element space on $\mathcal{T}_l$ by

$$V_l^c = \{v_l \in H_0^1(\Omega) \ : \ v_l|_T \in P_1(T), T \in \mathcal{T}_l\}.$$

The nonconforming finite element approximation of (2.4) is to find $u_l \in V_l$ such that

$$a_l(u_l, v_l) = (f, v_l), \quad v_l \in V_l, \tag{2.6}$$

where $a_l(\cdot, \cdot)$ stands for the mesh-dependent bilinear form

$$a_l(u_l, v_l) = \sum_{T \in \mathcal{T}_l} \big(a(x)\nabla u_l, \nabla v_l\big)_{0,T}. \tag{2.7}$$

Existence and uniqueness of the solution u_l again follow from the Lax-Milgram theorem. In the sequel, we refer to $\|\cdot\|_{A,l}$ as the mesh-dependent energy norm

$$\|v_l\|_{A,l}^2 = \sum_{T \in \mathcal{T}_l} a_l(v_l, v_l), \quad v_l \in V_l.$$

For brevity, when $l = L$ we will drop the subscript L from some of the above quantities, if no confusion is possible, e.g., we will write $a(\cdot, \cdot)$ instead of $a_L(\cdot, \cdot)$ and $\|\cdot\|_A$ instead of $\|\cdot\|_{A,L}$.

3. Condition Number Estimate

In this section, we consider the condition number estimate of the linear system arising from the discrete problem (2.6) based on a natural scaling of the basis functions on $\mathcal{T}_L$. The techniques presented in [1] and a prolongation operator from underlying nonconforming finite element space to finer conforming finite element space are used.

The computation of the solution u_L of (2.6) on $\mathcal{T}_L$ always requires to solve a matrix equation using particular basis functions for the finite element space V_L. Suppose that $\{\phi_i,\ i = 1, \cdots, N\}$ is a given basis for V_L, where N is the dimension of V_L, and define the matrix $\mathbf{A}$ and the vector $\mathbf{F}$ according to

$$(\mathbf{A})_{ij} := a(\phi_i, \phi_j) \quad \text{and} \quad (\mathbf{F})_i := (f, \phi_i), \quad i, j = 1, \cdots, N.$$

Then, equation (2.6) is equivalent to the linear algebraic system

$$\mathbf{AU} = \mathbf{F}, \tag{3.1}$$

where $u_L = \sum_{i=1}^N u_i\phi_i$ and $(\mathbf{U})_i = u_i$. We will specify conditions on V_L and the basis $\{\phi_i,\ i = 1, \cdots, N\}$ that will allow us to establish upper bounds for the condition number of $\mathbf{A}$.

Let the computational domain $\Omega \subset \mathbb{R}^d$, $d \geq 2$. We assume that $\mathcal{T}_L$ contains at most $\alpha_1^{d/2} N$ elements, with α_1 denoting a fixed constant. The following estimate holds true (cf., e.g., [14]):

$$C\|v\|_{1,T}^2 \leq h_T^{d-2}\|v\|_{L^\infty(T)}^2 \leq C\|v\|_{L^{2d/(d-2)}(T)}^2, \quad T \in \mathcal{T}_L,\ v \in V_L,\ d \geq 3. \tag{3.2}$$

In the special case of two dimension, we supplement the following inequality to the latter one in (3.2),

$$\|v\|_{L^\infty(T)} \leq Ch_T^{-2/p}\|v\|_{L^p(T)}, \quad T \in \mathcal{T}_L,\ v \in V_L,\ 1 \leq p \leq \infty. \tag{3.3}$$

Under the assumptions on the domain Ω, there exists a continuous embedding $H^1(\Omega) \hookrightarrow L^p(\Omega)$. For $d \geq 3$, Sobolev's inequality

$$\|v\|_{L^{2d/(d-2)}(\Omega)} \leq C\|v\|_{1,\Omega}, \quad v \in H^1(\Omega). \tag{3.4}$$

holds true. In two dimensions, we have a more explicit estimate (cf., e.g., [1])

$$\|v\|_{L^p(\Omega)} \leq C\sqrt{p}\|v\|_{1,\Omega}, \quad v \in H^1(\Omega),\ p < \infty. \tag{3.5}$$

As far as the basis $\{\phi_i,\, i = 1, \cdots, N\}$ of V_L is concerned, we assume that it is a local basis:

$$\max_{1\leq i\leq N} \#\{T \in \mathcal{T}_L : \operatorname{supp}(\phi_i) \cap T \neq \varnothing\} \leq \alpha_2, \tag{3.6}$$

where α_2 is a fixed constant. Finally, we impose a more important assumption with regard to the scaling of the basis:

$$Ch_T^{d-2}\|v\|^2_{L^\infty(T)} \leq \sum_{\operatorname{supp}(\phi_i)\cap T\neq\varnothing} v_i^2 \leq Ch_T^{d-2}\|v\|^2_{L^\infty(T)}, \quad T \in \mathcal{T}_L, \tag{3.7}$$

where $v = \sum_{i=1}^N v_i\phi_i$ and $\{v_i\}_{i=1}^N$ is arbitrary. For instance, if $\{\varphi_i,\, i = 1, \cdots, N\}$ denotes the lowest order Crouzeix-Raviart nonconforming basis functions, we define a new scaled basis $\{\phi_i,\, i = 1, \cdots, N\}$ by $\phi_i := h_i^{(2-d)/2}\varphi_i$, where h_i is the diameter of the support of φ_i. Then, the new basis satisfies assumption (3.7). We also impose the same assumption (3.7) for the conforming finite element basis, when utilized in the sequel.

For the analysis of the condition number estimate, we propose a prolongation operator from V_L to $\widetilde{V}^c_{L+1}$, where $\widetilde{V}^c_{L+1}$ is the conforming finite element space based on $\widetilde{\mathcal{T}}_{L+1}$, and $\widetilde{\mathcal{T}}_{L+1}$ is a shape regular conforming mesh obtained from $\mathcal{T}_L$ by any bisection strategy. The prolongation operator $I_L^{L+1} : V_L \to \widetilde{V}^c_{L+1}$ is defined by $I_L^{L+1}v(x) = \beta_x$ if x is an interior vertex of $\widetilde{\mathcal{T}}_{L+1}$, where β_x is the average of v at x, i.e., $\beta_x = \sum_{x\in\overline{T}_s} v|_{T_s}/M$. Here M is the number of elements $T_s \in \mathcal{T}_L$ sharing x as a vertex. Moreover, $I_L^{L+1}v(x) = 0$, if the vertex x is located on the Dirichlet boundary.

In this section, $\widetilde{\mathcal{T}}_{L+1}$ is an auxiliary triangulation, only used in the analysis. In the case of two dimension, $\widetilde{\mathcal{T}}_{L+1}$ is obtained from $\mathcal{T}_L$ by subdividing each $T \in \mathcal{T}_L$ into 4 simplices by joining the midpoints of the edges. In the case $d \geq 3$, $\widetilde{\mathcal{T}}_{L+1}$ is always obtained from a refinement of $\mathcal{T}_L$ such that the set of nodes of degrees of V_L contains in $\mathcal{N}_{L+1}$. For instance, when $d = 3$, $\widetilde{\mathcal{T}}_{L+1}$ is obtained from $\mathcal{T}_L$ by subdividing each $T \in \mathcal{T}_L$ into 12 simplices by joining the vertices, barycenter of T and barycenters of faces of T. The stability analysis of the prolongation operator I_L^{L+1} has been derived when $\widetilde{\mathcal{T}}_{L+1}$ is obtained from $\mathcal{T}_L$ by the above bisection algorithm (cf. [30, Lemma 5.2]).

We now give bounds on the condition number of the matrix $\mathbf{A}$, where $\{\phi_i,\, i = 1, \cdots, N\}$ is the scaled basis for V_L satisfying the above assumptions. In the general case $d \geq 3$, we have the following result.

Theorem 3.1. *Suppose that the nonconforming finite element space V_L satisfies* (3.2) *and the basis $\{\phi_i,\, i = 1, \cdots, N\}$ satisfies* (3.6) *and* (3.7)*. Then, the ℓ_2-condition number $\kappa(\mathbf{A})$ of the matrix $\mathbf{A}$ is bounded by*

$$\kappa(\mathbf{A}) \leq CN^{2/d}. \tag{3.8}$$

Proof. We set $v = \sum_{i=1}^{N} v_i \phi_i$, then $a(v,v) = \mathbf{X}^t \mathbf{A} \mathbf{X}$, where $(\mathbf{X})_i = v_i$. By a similar technique as in the proof of Theorem 3.1 in [1], we have $a(v,v) \le C\mathbf{X}^t\mathbf{X}$. On the other hand, we apply the prolongation operator I_L^{L+1} to v, and set

$$I_L^{L+1} v = \sum_{x_i \in \mathcal{N}_{L+1}(\widetilde{\mathcal{T}}_{L+1})} I_L^{L+1} v(x_i) \widetilde{\psi}_{i,L+1},$$

where $\{\widetilde{\psi}_{i,L+1}\}$ is the conforming finite element basis of $\widetilde{V}_{L+1}^c$. By Hölder's inequality, Sobolev's inequality, and the stability of I_L^{L+1}, we derive a complementary inequality according to

$$\begin{aligned}
\mathbf{X}^t\mathbf{X} &\le \sum_{T\in\widetilde{\mathcal{T}}_{L+1}} \sum_{\operatorname{supp}(\widetilde{\psi}_{i,L+1})\cap T\neq\varnothing} I_L^{L+1} v^2(x_i) \le C \sum_{T\in\widetilde{\mathcal{T}}_{L+1}} h_T^{d-2} \|I_L^{L+1} v\|_{L^\infty(T)}^2 \\
&\le C \sum_{T\in\widetilde{\mathcal{T}}_{L+1}} \|I_L^{L+1} v\|_{L^{2d/(d-2)}(T)}^2 \le C N^{2/d} \|I_L^{L+1} v\|_{L^{2d/(d-2)}(\Omega)}^2 \\
&\le C N^{2/d} \|I_L^{L+1} v\|_{1,\Omega}^2 \le C N^{2/d} \|v\|_{1,L}^2 \le C N^{2/d} a(v,v).
\end{aligned}$$

Using the above estimates, we obtain $N^{-2/d}\mathbf{X}^t\mathbf{X} \le C\mathbf{X}^t\mathbf{A}\mathbf{X} \le C\mathbf{X}^t\mathbf{X}$, which implies

$$N^{-2/d} \le C\lambda_{\min}(\mathbf{A}) \quad \text{and} \quad \lambda_{\max}(\mathbf{A}) \le C.$$

Recalling $\kappa(\mathbf{A}) = \lambda_{\max}(\mathbf{A})/\lambda_{\min}(\mathbf{A})$, the above two estimates yield (3.8). □

In the special case $d = 2$, a similar result can be deduced as follows.

Theorem 3.2. *Suppose that the nonconforming finite element space V_L satisfies* (3.2) *and* (3.3)*, and that the basis $\{\phi_i,\ i = 1, \cdots, N\}$ satisfies* (3.6) *and* (3.7)*. Then, the ℓ_2-condition number $\kappa(\mathbf{A})$ of the matrix $\mathbf{A}$ is bounded by*

$$\kappa(\mathbf{A}) \le CN\big(1 + |\log(N h_{\min}^2)|\big), \tag{3.9}$$

where $h_{\min} = \min\{h_T :\ T \in \mathcal{T}_L\}$.

Proof. We set $v = \sum_{i=1}^{N} v_i\phi_i$, $(\mathbf{X})_i = v_i$ and $a(v,v) = \mathbf{X}^t\mathbf{A}\mathbf{X}$. As in the proof of the above theorem, it suffices to show that

$$CN^{-1}\big(1 + |\log(N h_{\min}^2)|\big)^{-1}\mathbf{X}^t\mathbf{X} \le \mathbf{X}^t\mathbf{A}\mathbf{X} \le C\mathbf{X}^t\mathbf{X}. \tag{3.10}$$

Actually, $a(v,v) \le C\mathbf{X}^t\mathbf{X}$ holds true as Theorem 4.1 in [1].

As far as the lower bound in (3.10) is concerned, as in the proof of Theorem 3.1 we have $(p > 2)$

$$\begin{aligned}
\mathbf{X}^t\mathbf{X} &\le \sum_{T\in\widetilde{\mathcal{T}}_{L+1}} \sum_{\operatorname{supp}(\widetilde{\psi}_{i,L+1})\cap T\neq\varnothing} I_L^{L+1} v^2(x_i) \le C \sum_{T\in\widetilde{\mathcal{T}}_{L+1}} \|I_L^{L+1} v\|_{L^\infty(T)}^2 \\
&\le C \sum_{T\in\widetilde{\mathcal{T}}_{L+1}} h_T^{-4/p} \|I_L^{L+1} v\|_{L^p(T)}^2 \le C\Big(\sum_{T\in\widetilde{\mathcal{T}}_{L+1}} h_T^{-4/(p-2)}\Big)^{(p-2)/p} \|I_L^{L+1} v\|_{L^p(T)}^2 \\
&\le C\Big(\sum_{T\in\widetilde{\mathcal{T}}_{L+1}} h_T^{-4/(p-2)}\Big)^{(p-2)/p} p\, \|I_L^{L+1} v\|_{1,\Omega}^2 \\
&\le C\Big(\sum_{T\in\widetilde{\mathcal{T}}_{L+1}} h_T^{-4/(p-2)}\Big)^{(p-2)/p} p\, a(v,v) \le CN\big(N h_{\min}^2\big)^{-2/p} p\, a(v,v).
\end{aligned}$$

The special choice $p = \max\{2, \left|\log(Nh_{\min}^2)\right|\}$ allows to conclude. $\square$

For a fixed triangulation, the conforming P1 finite element space is contained in the non-conforming P1 finite element space. Hence, the sharpness of the bounds in Theorem 3.2 can be verified by the same example as in [1].

4. Local Multilevel Methods

The above section clearly shows that for the solution of a large scale problem the convergence of standard iterations such as Gauss-Seidel or CG will become very slow. This motivates the construction of more efficient iterative algorithms for those algebraic systems resulting from adaptive nonconforming finite element approximations.

We will present our local multilevel methods for adaptive nonconforming P1 finite element discretizations based on the Crouzeix-Raviart elements. As a prerequisite, for the shape regular family of nested conforming meshes $\{\mathcal{T}_l\}_{l=0}^L$ generated in the ANFEM using the newest vertex bisection algorithm, when $l < L$, we again use the prolongation operator $I_l^{l+1} : V_l \to V_{l+1}^c$ defined as in section 3. The prolongation operator I_l^{l+1} defines the values of $I_l^{l+1}v$ at the vertices of elements of level $l+1$, yielding a continuous piecewise linear function on $\mathcal{T}_{l+1}$. $I_l^{l+1}v$ is a function in V_{l+1}, and it naturally represents a function in the finest space V_L. Hence, the operator I_l given by

$$I_l v := I_l^{l+1} v, \quad v \in V_l,\ 0 \le l < L,$$

defines an intergrid operator from V_l to V_L. Now, proving stability of the "iterated" prolongation operator $I_l^L = \widetilde{I}_{L-1} \cdots \widetilde{I}_{l+1} I_l : V_l \to V_L^c \subset V_L$, where $\widetilde{I}_k : V_k^c \to V_{k+1}^c$ is the natural injection operator, is basically reduced to proving stability of the single two-level operator I_l while the influence of the multilevel operator $\widetilde{I}_{L-1} \cdots \widetilde{I}_{l+1}$ is basically taken care of by results for the nested conforming P1 case. This basic idea is the key in the design of our local multilevel methods.

Next we provide analysis of the stability estimate of the operator I_l.

Lemma 4.1. *Let $\mathcal{T}_{l+1}$ be obtained from $\mathcal{T}_l$ by the newest vertex algorithm. There exist positive constants $\widetilde{C}_I, C_I$ independent of mesh sizes and mesh levels such that*

$$(I_l v, I_l v) \le \widetilde{C}_I (v, v), \quad a(I_l v, I_l v) \le C_I a_l(v, v), \quad v \in V_l. \tag{4.1}$$

Proof. The first inequality in (4.1) is trivial for $I_l v$ being defined by local averaging. It suffices to derive the second one. For simplicity, we present the analysis in the 2D case. The three dimensional case can be derived similarly.

The origin of vertices of $T \in \mathcal{T}_{l+1}$ includes four cases depending whether the vertex of T is the midpoint of an edge or a vertex in $\mathcal{T}_l$. In particular, let m, n denote the number of vertices of $T \in \mathcal{T}_{l+1}$ representing midpoints or vertices in $\mathcal{T}_l$, respectively. Setting $S = \{(m,n) : m+n = 3,\ m, n = 0, 1, 2, 3\}$, we have $\#S = 4$. We only consider one of the possible cases: the vertices of $T \in \mathcal{T}_{l+1}$ are all vertices in $\mathcal{T}_l$, i.e., T is not refined in the transition from $\mathcal{T}_l$ to $\mathcal{T}_{l+1}$, e.g., $T_2 \in \mathcal{T}_{l+1}$ is also $K_2 \in \mathcal{T}_l$ in Fig. 4.1. A similar analysis can be carried out in all other cases.

Note that $a(I_l v, I_l v)|_{T_2}$ is equivalent to

$$\big(I_l v(x_1) - I_l v(x_2)\big)^2 + \big(I_l v(x_1) - I_l v(x_3)\big)^2. \tag{4.2}$$

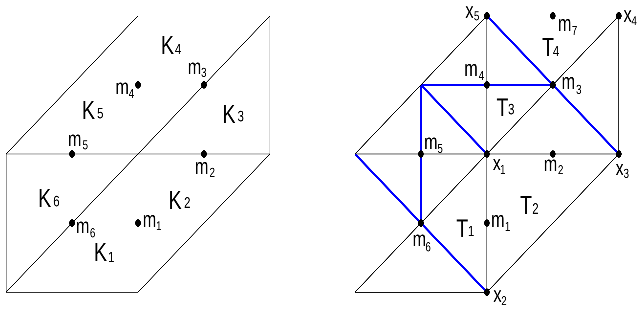

Fig. 4.1. The 2D case: the left figure illustrates a local grid of $\mathcal{T}_l$, and the right one displays its refinement as part of $\mathcal{T}_{l+1}$.

We recall that $I_l v(x_i)$ is the average of v at x_i over the triangles $K_i, i = 1, ..., M_{x_i}$, where M_{x_i} is the number of triangles containing x_i. Hence, the first term of (4.2) can be written as

$$\frac{1}{M_{x_1}} \sum_{i=1}^{M_{x_1}} \Big(v|_{K_i}(x_1) - v(m_1) \Big) + \frac{1}{M_{x_2}} \sum_{s=1}^{M_{x_2}} \Big(v(m_1) - v|_{K_s}(x_2) \Big). \tag{4.3}$$

A similar result can be obtained for the second term of (4.2). Since

$$v|_{K_i}(x_1) - v(m_1) = v|_{K_i}(x_1) - v(m_l) + \sum_{j=1}^{l-1} \Big(v(m_{j+1}) - v(m_j) \Big),$$

it suffices to find a constant C such that the first term of (4.3) can be bounded by

$$\sum_{i=1}^{M_{x_1}} \Big(v|_{K_i}(x_1) - v(m_1) \Big) \le C a_l(v,v)|_{\widetilde{K}}, \tag{4.4}$$

where $\widetilde{K} = \cup_{i=1}^{M_{x_1}} K_i$. The same analysis can be carried out for the second term of (4.3). Following (4.2-4.4), we get

$$a(I_l v, I_l v)|_{T_2} \le C a_l(v,v)|_{\widetilde{T}_2} \tag{4.5}$$

with some constant C, where $\widetilde{T}_2$ is a patch of triangles in $\mathcal{T}_l$ also containing the vertices of T_2.

For $T \in \mathcal{T}_{l+1}$, $\partial T \cap \partial\Omega \ne \varnothing$, let us assume $\partial T_4 \cap \partial\Omega \ne \varnothing$. Then, $a(I_l v, I_l v)|_{T_4}$ can be bounded by

$$C\Big(\big(I_l v(m_3) - I_l v(x_4)\big)^2 + \big(I_l v(m_3) - I_l v(x_5)\big)^2 \Big) = 2C\big(v(m_3) - v(m_7)\big)^2. \tag{4.6}$$

Combining (4.5), (4.6) and summing up all $T \in \mathcal{T}_{l+1}$ complete the proof. $\square$

With the above prolongation operator I_l, we define projections P_l, $P_l^0 : V_L \to V_l$ according to

$$a_l(P_l v, w) = a(v, I_l w), \quad (P_l^0 v, w) = (v, I_l w), \quad v \in V_L,\, w \in V_l.$$

For $0 \le l \le L$, we also define $A_l : V_l \to V_l$ by means of

$$(A_l v, w) = a_l(v, w), \quad w \in V_l.$$

For any node $\boldsymbol{p} \in \mathcal{N}_l$, we use the notation $\psi_l^{\boldsymbol{p}}$ to represent the associated nodal conforming basis function of V_l^c. Let $\widetilde{\mathcal{N}}_l$ be the set of new nodes and those old nodes where the support of the associated basis function has changed, i.e.,

$$\widetilde{\mathcal{N}}_l = \{\boldsymbol{p} \in \mathcal{N}_l : \boldsymbol{p} \in \mathcal{N}_l \setminus \mathcal{N}_{l-1} \text{ or } \boldsymbol{p} \in \mathcal{N}_{l-1} \text{ but } \psi_l^{\boldsymbol{p}} \neq \psi_{l-1}^{\boldsymbol{p}}\}.$$

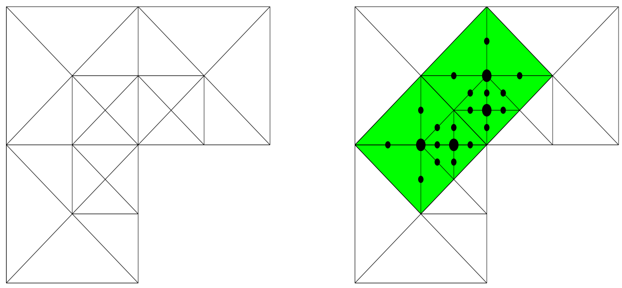

Fig. 4.2. The 2D case: coarse mesh (left), fine mesh (right) and illustration of $\widetilde{\mathcal{M}}_l$: the big nodes on the right refer to $\widetilde{\mathcal{N}}_l$, the small nodes refer to $\widetilde{\mathcal{M}}_l$, $l = 1, \cdots, L-1$.

For $1 \le l < L$, let $\widetilde{\mathcal{M}}_l$ represent the set of barycenters of faces on which local smoothers are performed (see Fig. 4.2 of the 2D case for an illustration):

$$\widetilde{\mathcal{M}}_l := \bigcup_{\boldsymbol{p} \in \widetilde{\mathcal{N}}_l} \mathcal{M}_l^0\big(\mathrm{supp}(\psi_l^{\boldsymbol{p}})\big), \tag{4.7}$$

where $\mathcal{M}_l^0\big(\mathrm{supp}(\psi_l^{\boldsymbol{p}})\big)$ represents the barycenters of interior faces $\mathcal{F}_l^0\big(\mathrm{supp}(\psi_l^{\boldsymbol{p}})\big)$. On the finest level L, we set $\widetilde{\mathcal{M}}_L = \mathcal{M}_L$.

For convenience, we set $\widetilde{\mathcal{M}}_l = \{\boldsymbol{b}_l^i, i = 1, \cdots, \widetilde{n}_l\}$, where $\widetilde{n}_l$ is the cardinality of $\widetilde{\mathcal{M}}_l$, and refer to φ_l^i as the lowest order Crouzeix-Raviart nonconforming finite element basis function of V_l. We also denote $V_0^1 := V_0$, $\widetilde{n}_0 = 1$. Then, for $i = 1, \cdots, \widetilde{n}_l$, let $P_l^i, Q_l^i : V_l \to V_l^i = \mathrm{span}\{\varphi_l^i\}$ be defined by

$$a_l(P_l^i v, \varphi_l^i) = a_l(v, \varphi_l^i), \quad (Q_l^i v, \varphi_l^i) = (v, \varphi_l^i), \quad v \in V_l\,,$$

and let $A_l^i : V_l^i \to V_l^i$ be defined by

$$(A_l^i v, \varphi_l^i) = a_l(v, \varphi_l^i), \quad v \in V_l^i.$$

It is easy to see that the following relationship holds true:

$$A_l^i P_l^i = Q_l^i A_l. \tag{4.8}$$

We assume that the local smoother $R_l : V_l \to V_l$ is nonnegative, symmetric or nonsymmetric with respect to the inner product $(\cdot, \cdot)$. It will be precisely defined and further studied in section 5. For $l = 1, \cdots, L-1$, R_l is only performed on local barycenters $\widetilde{\mathcal{M}}_l$. R_0 is solved directly, i.e., $R_0 = A_0^{-1}$. On the finest level, R_L is carried out on all interior barycenters $\mathcal{M}_L$. For simplicity, we set $A = A_L$ and denote by I_L and P_L the identity operator on the finest space V_L. We set

$$S_l := I_l R_l A_l P_l, \quad l = 0, 1, \cdots, L.$$

Now, we scale S_l as follows:

$$T_l := \mu_{L,l} S_l, \quad l = 0, 1, \cdots, L. \tag{4.9}$$

where $\mu_{L,l} > 0$ is a parameter, independent of mesh sizes and mesh levels, chosen to satisfy

$$a(T_l v, T_l v) \le \omega_l\, a(T_l v, v), \quad v \in V_L \ , \ w_l < 2.$$

We will also drop the subscript L from $\mu_{L,l}$ since no confusion is possible in the convergence analysis.

With the sequences of operators $\{T_l, l = 0, 1, \cdots, L\}$, we can now state the local multilevel methods for adaptive nonconforming P1 finite element methods as follows.

Algorithm 4.1. *Local multilevel product algorithm (LMPA) Given an arbitrarily chosen initial iterate $u^0 \in V_L$, we seek $u^n \in V_L$ as follows:*
1) *Let $v_0 = u^{n-1}$. For $l = 0, 1, \cdots, L$, compute v_{l+1} by*

$$v_{l+1} = v_l + T_l(u_L - v_l). \tag{4.10}$$

2) *Set $u^n = v_{L+1}$.*

Algorithm 4.2. *Local multilevel additive algorithm (LMAA)*
Let $T = \sum_{l=0}^{L} T_l$ and let u_L be the exact solution of (2.6) *on $\mathcal{T}_L$. Find $\widetilde{u}_L \in V_L$ such that*

$$T\widetilde{u}_L = \widetilde{f}, \tag{4.11}$$

where $\widetilde{f} = \sum_{l=0}^{L} T_l u_L$.

In view of the operator equation

$$A_l P_l = P_l^0 A,$$

the function $\widetilde{f}$ in (4.11) is formally defined by the exact finite element solution u_L which can be computed directly, and so does the iteration (4.10).

Obviously, there exists a unique solution $\widetilde{u}_L$ of (4.11) coinciding with u_L for (2.6) on $\mathcal{T}_L$. The conjugate-gradient method can be used to solve the new problem, if T is symmetric. We can also apply the conjugate gradient method to the symmetric version of LMAA (SLMAA) by solving

$$\overline{T}\widetilde{u}_L = \overline{f}$$

instead of (4.11), where $\overline{T} = \frac{1}{2}(T + T^*)$, $\overline{f} = \frac{1}{2}\sum_{l=0}^{L}(T_l + T_l^*)u_L$ and T^*, T_l^* denote the adjoint operator of T, T_l with respect to the inner product $a(\cdot,\cdot)$ respectively .

5. Convergence Theory

In this section, we provide an abstract theory concerned with the convergence of local multilevel methods for linear systems arising from adaptive nonconforming P1 finite element methods. We will use the well-known Schwarz theory developed in [29, 34, 39] to analyze the algorithms. This basic idea that the conforming P1 finite element spaces are contained in their nonconforming counterparts plays a key role in the decomposition of the nonconforming P1 finite element space V_L and the proof of a global strengthened Cauchy-Schwarz inequality.

5.1. Schwarz theory

Since the spaces $\{V_l\}_{l=0}^{L}$ are nonnested, we can not directly get the space decomposition of V_L by these nonconforming finite element spaces. Based on the basic idea $V_l^c \subset V_L, l = 0, 1, \cdots, L$, we can obtain a stable space decomposition of V_L as follows:

$$V_L = \sum_{l=0}^{L} V_l = V_0 + \sum_{l=1}^{L} \sum_{\boldsymbol{b}_l^i \in \widetilde{\mathcal{M}}_l} V_l^i. \tag{5.1}$$

In the next subsection we will illustrate how to construct the above space decomposition. In the following we show the two key properties of the space decomposition (5.1) in the Schwarz theory and delay the proof in the next subsection.

(S1) *Stability of space decomposition.* For any $v \in V_L$, there exists a decomposition

$$v = v_0 + \sum_{l=1}^{L} \sum_{i=1}^{\widetilde{n}_l} v_l^i, \qquad v_0 \in V_0,\ v_l^i \in V_l^i,$$

and a positive constant C_{stab}, independent of mesh sizes and mesh levels, such that

$$\|v_0\|_{A,0}^2 + \sum_{l=1}^{L} \sum_{i=1}^{\widetilde{n}_l} \|v_l^i\|_{A,l}^2 \le C_{\text{stab}} \|v\|_A^2 .$$

(S2) *Global strengthened Cauchy-Schwarz inequality.* For any functions

$$v_l^i, w_l^i \in V_l^i, \qquad 1 \le i \le \widetilde{n}_l,\ 0 \le l \le L,$$

there exists a positive constant C_{orth}, independent of mesh sizes and mesh levels, such that

$$\sum_{l=0}^{L} \sum_{i=1}^{\widetilde{n}_l} \sum_{k=0}^{l-1} \sum_{j=1}^{\widetilde{n}_k} a(I_l v_l^i, I_k w_k^j) \le C_{\text{orth}} \Big(\sum_{l=0}^{L} \sum_{i=1}^{\widetilde{n}_l} \|I_l v_l^i\|_A^2 \Big)^{\frac{1}{2}} \Big(\sum_{l=0}^{L} \sum_{i=1}^{\widetilde{n}_l} \|I_l w_l^i\|_A^2 \Big)^{\frac{1}{2}}. \tag{5.2}$$

Let $T_l = \mu_l I_l R_l A_l P_l$, $0 \le l \le L$, where $R_0 = A_0^{-1}$, for $1 \le l \le L-1$, R_l is local Jacobi or local Gauss-Seidel smoother, and R_L is a global Jacobi or Gauss-Seidel smoother. Based on the above two properties S1 and S2, we can further derive the following assumptions with a positive constant C independent of mesh sizes and mesh levels.

(A1) Let $T = \sum_{l=0}^{L} T_l$ be an additive operator, then

$$\|v\|_A^2 \le \frac{C}{\mu} a(Tv, v), \quad v \in V_L,$$

where $\mu = \min_{0 \le l \le L}\{\mu_l\}$.

(A2) There exists a constant $\omega_l \in (0, 2)$ which depends on μ_l but is independent of mesh sizes and mesh levels such that

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

(A3) For any $v_l, w_k \in V_L$, $0 \le l, k \le L$, we have

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

In the following subsections we will also apply the abstract theory to the algorithms LMPA and LMAA by verifying assumptions A1-A3 for the adaptive nonconforming P1 finite element method. There are two types of smoothers R_l, Jacobi and Gauss-Seidel iterations, which will be considered separately.

The abstract theory of local multilevel methods can be invoked due to the above statements. For the algorithm LMPA, the abstract theory provides an estimate for the energy norm of the error operator

$$E_M = (I - T_L)\cdots(I - T_1)(I - T_0),$$

where I is the identity operator in V_L. We have the following uniform convergence theorem.

Theorem 5.1. *Let the assumptions A1-A3 be satisfied. Then, for the algorithm LMPA, the energy norm of the error operator E_M can be bounded as follows (cf. [29, 34, 39])*

$$a(E_M v, E_M v) \le \delta\, a(v, v), \quad v \in V_L,$$

where $\delta = 1 - \frac{\mu(2-\omega)}{C}$, $\omega = \max\limits_{0\le l\le L}\{\omega_l\}$, C is a positive constant independent of mesh sizes and mesh levels.

For the additive local multilevel algorithm 4.2, the following theorem provides a spectral estimate for the operator $T = \sum_{l=0}^{L} T_l$ and its symmetric version $\overline{T}$.

Theorem 5.2. *Let the assumptions A1-A3 hold true. Then, there exists a positive constant C independent of mesh sizes and mesh levels such that (cf. [29, 34, 39])*

$$\frac{\mu}{C}\|v\|_A \le \|Tv\|_A \le C\|v\|_A, \quad \frac{\mu}{C}\|v\|_A \le \|\overline{T}v\|_A \le C\|v\|_A, \quad v \in V_L.$$

Theorem 5.2 implies that the ℓ_2-condition number of T and $\overline{T}$ can be bounded as follows:

$$\kappa(T) \le \frac{C}{\mu}, \quad \kappa(\overline{T}) \le \frac{C}{\mu}.$$

It should be pointed out that the convergence result for LMPA or for the preconditioned conjugate gradient method by LMAA depends on the parameter μ, which will be observed in our numerical experiments.

5.2. Stability of space decomposition and global strengthened Cauchy-Schwarz inequality

For any $v \in V_L$ we consider the decomposition

$$v = \sum_{l=0}^{L} v_l, \quad v_L = v - \widetilde{v}, \quad v_0 = \Pi_0\widetilde{v}, \quad v_l = (\Pi_l - \Pi_{l-1})\widetilde{v}, \quad l = 1, \cdots, L-1, \tag{5.3}$$

where $\widetilde{v} = \widetilde{\Pi}_{L-1}v$ and $\widetilde{\Pi}_{L-1}v$ represents a local regularization of v in V_{L-1}^c (c.f. [10]), e.g., by a Clément-type interpolation. The operator $\Pi_l : H_0^1(\Omega) \to V_l^c$ stands for the Scott-Zhang type quasi-interpolation operator [17, 26], and it is defined by

$$\Pi_l v = \sum_{\boldsymbol{p}\in\mathcal{N}_l} \psi_l^{\boldsymbol{p}} \int_{\sigma_{\boldsymbol{p}}} \theta_l^{\boldsymbol{p}} v, \quad v \in H_0^1(\Omega),$$

where $\sigma_{\boldsymbol{p}}$ is a d-simplex or $(d-1)$-simplex sharing the vertex $\boldsymbol{p}$, $\theta_l^{\boldsymbol{p}}$ is the $L^2(\sigma_{\boldsymbol{p}})$-dual basis function such that $\int_{\sigma_{\boldsymbol{p}}} \theta_l^{\boldsymbol{p}} w = w(\boldsymbol{p})$ for all $w \in P_1(\sigma_{\boldsymbol{p}})$. When $l \geq 1$, for any vertex $\boldsymbol{p} \in \mathcal{N}_l \cap \mathcal{N}_{l-1}$ satisfying $\psi_l^{\boldsymbol{p}} = \psi_{l-1}^{\boldsymbol{p}}$, we can choose the same $\sigma_{\boldsymbol{p}}$, thus

$$(\Pi_l v - \Pi_{l-1} v)(\boldsymbol{p}) = 0, \quad v \in H_0^1(\Omega),\ \boldsymbol{p} \in \mathcal{N}_l \setminus \widetilde{\mathcal{N}}_l.$$

By the space decomposition (5.3), for $1 \leq l \leq L-1$, we have

$$v_l \in \bigcup_{\boldsymbol{p}\in\widetilde{\mathcal{N}}_l} \operatorname{span}\{\psi_l^{\boldsymbol{p}}\} \subset \bigcup_{\boldsymbol{b}\in\widetilde{\mathcal{M}}_l} \operatorname{span}\{\varphi_l^{\boldsymbol{b}}\},$$

then $v_l = \sum_{\boldsymbol{b}\in\widetilde{\mathcal{M}}_l} v_l^{\boldsymbol{b}}, v_l^{\boldsymbol{b}} = v_l(\boldsymbol{b})\varphi_l^{\boldsymbol{b}}$, $\varphi_l^{\boldsymbol{b}} \in V_l$ is the lowest order Crouzeix-Raviart nonconforming finite element basis function corresponding to $\boldsymbol{b}$. Thus, for $1 \leq l \leq L-1$, the local smoothing nodes are those in (4.7). Since $v_L = v - \widetilde{v}$, the smoothing relaxation on the finest level L is done on all nodes of degrees of freedom $\mathcal{M}_L$. Although R_L is performed on $\mathcal{M}_L$, R_l is a local smoother on other levels $1 \leq l < L$. Thus the computational complexity of the local multilevel methods is also quasi-optimal. The next lemma shows the statement S1 of the stability estimate of the space decomposition (5.3).

Lemma 5.1. *Let $v \in V_L$, $v = v_0 + \sum_{l=1}^{L} \sum_{\boldsymbol{b}_l^i\in\widetilde{\mathcal{M}}_l} v_l^i$ be the space decomposition in* (5.3)*. There exists a positive constant C independent of mesh sizes and mesh levels such that*

$$\|v_0\|_{A,0}^2 + \sum_{l=1}^{L} \sum_{\boldsymbol{b}_l^i\in\widetilde{\mathcal{M}}_l} \|v_l^i\|_{A,l}^2 \leq C\,\|v\|_A^2\,. \tag{5.4}$$

Proof. For $1 \leq l \leq L-1$, $v_l = (\Pi_l - \Pi_{l-1})\widetilde{v} \in V_{L-1}^c \subset V_L$. By the stability estimate of the space decomposition of conforming P1 finite element space V_{L-1}^c (cf. [17, 33]) and the stability estimate of the local regularization operator $\widetilde{\Pi}_{l-1}$ [10], we deduce

$$\sum_{l=1}^{L-1} \sum_{i=1}^{\widetilde{n}_l} \|v_l^i\|_{A,l}^2 \leq C\|\widetilde{v}\|_A^2 = C\|\widetilde{\Pi}_{l-1}v\|_A^2 \leq C\|v\|_A^2.$$

On the coarsest level $l = 0$, we have

$$\|v_0\|_{A,0}^2 = a_0(\Pi_0\widetilde{v}, \Pi_0\widetilde{v}) \leq Ca(\widetilde{v}, \widetilde{v}) \leq C\|v\|_A^2.$$

On the finest level $l = L$, it follows from the local estimate of $\widetilde{\Pi}_{L-1}$ that

$$\sum_{i=1}^{\widetilde{n}_L} \|v_L^i\|_A^2 \leq C \sum_{i=1}^{\widetilde{n}_L} h_{L,i}^{-d} \|v - \widetilde{v}\|_{0,\Omega_L^i}^2 \leq C\|v\|_A^2,$$

where $\Omega_L^i = \operatorname{supp}(\varphi_L^i), h_{L,i} = \operatorname{diam}(\Omega_L^i)$. The stated result is obviously obtained by combining the above estimates. $\square$

For the local multilevel methods for adaptive conforming finite element methods, the associated global strengthened Cauchy-Schwarz inequality has been established in [13, 17, 35]. Due to the fact that the algorithms LMPA and LMAA are based on the basic property $V_l^c \subset V_L$, we can apply the techniques used for the conforming case (cf. [13, 17]) to prove the statement S2.

Lemma 5.2. *Let V_L be decomposed as in* (5.1). *Then, the global strengthened Cauchy-Schwarz inequality* (5.2) *in statement S*2 *holds true.*

Proof. For convenience we introduce the generation of an element T, $\mathcal{G}(T)$, 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 and only depends on $\mathcal{T}_0$ and the shape regularity of the meshes. Let $\widetilde{\Omega}_l = \bigcup_{\boldsymbol{b}\in\widetilde{\mathcal{M}}_l} \mathrm{supp}\{\varphi_l^{\boldsymbol{b}}\}$, $\widetilde{\mathcal{T}}_l = \{T \in \mathcal{T}_l :\ T \subset \widetilde{\Omega}_l\}$. We have

$$\begin{aligned} I_0 :=& \sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l}\sum_{k=1}^{l-1}\sum_{j=1}^{\widetilde{n}_k} a(I_l v_l^i, I_k w_k^j) \\ =& \sum_{l=1}^{L}\sum_{k=1}^{l-1}\sum_{m,n=0}^{\infty} \sum_{\substack{T\in\widetilde{\mathcal{T}}_l\\ \mathcal{G}(T)=m}} \sum_{\substack{K\in\widetilde{\mathcal{T}}_k\\ \mathcal{G}(K)=n}} \sum_{\substack{\boldsymbol{b}\in\mathcal{M}(T)\\ \boldsymbol{a}\in\mathcal{M}(K)}} a(I_l\widetilde{v}_l^{\boldsymbol{b}}, I_k\widetilde{w}_k^{\boldsymbol{a}}), \end{aligned}$$

where $\mathcal{M}(T)$ denotes the set of barycenters of faces of T,

$$\widetilde{v}_l^{\boldsymbol{b}} = \begin{cases} v_l^{\boldsymbol{b}}/N_l(\boldsymbol{b}), & \text{if } \ \boldsymbol{b} \in \widetilde{\mathcal{M}}_l \ , \\ 0, & \text{otherwise,} \end{cases}$$

$N_l(\boldsymbol{b})$ is the number of elements contained in $\widetilde{\mathcal{T}}_l$ sharing the barycenter $\boldsymbol{b} \in \widetilde{\mathcal{M}}_l$, $N_l(\boldsymbol{b}) = 1$ or 2. $\widetilde{w}_k^{\boldsymbol{a}}$ can be defined similarly.

Suppose $m \le n$ and set

$$\widetilde{w}_n := \sum_{k=1}^{l-1} \sum_{\substack{K\in\widetilde{\mathcal{T}}_k\\ \mathcal{G}(K)=n}} \sum_{\boldsymbol{a}\in\mathcal{M}(K)} I_k\widetilde{w}_k^{\boldsymbol{a}}.$$

For any $T \in \widetilde{\mathcal{T}}_l, \mathcal{G}(T) = m \le n, \boldsymbol{b} \in \mathcal{M}(T), 1 \le l \le L-1$, we first prove the following estimate:

$$a(I_l\widetilde{v}_l^{\boldsymbol{b}}, \widetilde{w}_n) \le C\theta^{\frac{n-m}{2}} \|\nabla I_l\widetilde{v}_l^{\boldsymbol{b}}\|_{0,\widetilde{\Omega}_l^{\boldsymbol{b}}} \Big(\sum_{k=1}^{l-1} \sum_{\substack{K\in\widetilde{\mathcal{T}}_k\\ \mathcal{G}(K)=n}} \sum_{\boldsymbol{a}\in\mathcal{M}(K)} \|\nabla I_k\widetilde{w}_k^{\boldsymbol{a}}\|^2_{0,\widetilde{\Omega}_l^{\boldsymbol{b}}} \Big)^{\frac{1}{2}}, \tag{5.5}$$

where $\widetilde{\Omega}_l^{\boldsymbol{b}} = \mathrm{supp}(I_l\widetilde{v}_l^{\boldsymbol{b}})$. Let $\Omega_l^{\boldsymbol{b}} = \mathrm{supp}(\widetilde{v}_l^{\boldsymbol{b}}), \mathcal{T}_{l+1}(\widetilde{\Omega}_l^{\boldsymbol{b}}) = \{T' \in \mathcal{T}_{l+1} :\ T' \subset \widetilde{\Omega}_l^{\boldsymbol{b}}\}$. There exist positive constants t_0, s_0 depending only on the shape regularity of the meshes such that

$$\max_{\substack{T'\in\mathcal{T}_{l+1}\\ T'\subset\widetilde{\Omega}_l^{\boldsymbol{b}}}} \mathcal{G}(T') \le \min_{\substack{T'\in\mathcal{T}_{l+1}\\ T'\subset\widetilde{\Omega}_l^{\boldsymbol{b}}}} \mathcal{G}(T') + t_0, \qquad \min_{\substack{T'\in\mathcal{T}_{l+1}\\ T'\subset\widetilde{\Omega}_l^{\boldsymbol{b}}}} \mathcal{G}(T') \le \min_{\substack{T\in\mathcal{T}_l\\ T\subset\Omega_l^{\boldsymbol{b}}}} \mathcal{G}(T) + s_0.$$

If $n-m \le t_0+s_0$, the Cauchy-Schwarz inequality and local overlapping of subdomains $\{I_k\widetilde{w}_k^{\boldsymbol{a}} : \boldsymbol{a}\in\mathcal{M}_k\}$ on each level directly indicate (5.5). For the case $n-m > t_0+s_0$, we note that $I_l\widetilde{v}_l^{\boldsymbol{b}}$ is piecewise linear on $\mathcal{T}_{l+1}(\widetilde{\Omega}_l^{\boldsymbol{b}})$, $\widetilde{w}_n$ is piecewise linear in any $T'\in\mathcal{T}_{l+1}(\widetilde{\Omega}_l^{\boldsymbol{b}})$. We set

$$\xi_n := \sum_{k=1}^{l-1} \sum_{\substack{K\in\widetilde{\mathcal{T}}_k \\ \mathcal{G}(K)=n}} \sum_{\substack{\boldsymbol{a}\in\mathcal{M}(K) \\ \operatorname{supp}(I_k\widetilde{w}_k^{\boldsymbol{a}})\cap\partial T'\neq\emptyset}} I_k\widetilde{w}_k^{\boldsymbol{a}} \qquad \text{on } \partial T'.$$

Let $\{\widehat{\mathcal{T}}_j\}_{j\ge 0}$ be a sequence of quasi-uniformly refined meshes. Here $\widehat{\mathcal{T}}_j$ is generated by subdividing each element $T\in\widehat{\mathcal{T}}_{j-1}$ into 2^d simplices by connecting the edge midpoints starting from $\widehat{\mathcal{T}}_0=\mathcal{T}_0$. It is clear that

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

Since $I_l\widetilde{v}_l^{\boldsymbol{b}}$ is linear in T', using Green's formula we have

$$\begin{aligned}\int_{T'} \nabla I_l\widetilde{v}_l^{\boldsymbol{b}}\cdot\nabla\widetilde{w}_n &= \int_{\partial T'} \frac{\partial I_l\widetilde{v}_l^{\boldsymbol{b}}}{\partial\boldsymbol{n}}\widetilde{w}_n = \int_{\partial T'} \frac{\partial I_l\widetilde{v}_l^{\boldsymbol{b}}}{\partial\boldsymbol{n}}\xi_n = \int_{T'\cap\Gamma_{T'}} \nabla I_l\widetilde{v}_l^{\boldsymbol{b}}\cdot\nabla\xi_n \\ &\le \|\nabla I_l\widetilde{v}_l^{\boldsymbol{b}}\|_{L^2(\Gamma_{T'})}\|\nabla\xi_n\|_{L^2(\Gamma_{T'})} \le C\theta^{\frac{n-m}{2}}\|\nabla I_l\widetilde{v}_l^{\boldsymbol{b}}\|_{0,T'}\|\nabla\xi_n\|_{0,T'}.\end{aligned}$$

Summing up the above inequality over all $\mathcal{T}_{l+1}(\widetilde{\Omega}_l^{\boldsymbol{b}})$ yields (5.5). When $l=L$, we can also get the similar estimate as in (5.5). Applying (5.5) and the local overlapping of the supports of $I_l\widetilde{v}_l^{\boldsymbol{b}}$ and $I_k\widetilde{w}_k^{\boldsymbol{a}}$, we obtain

$$\begin{aligned} I_1 :&= \sum_{m=0}^{\infty}\sum_{n=m}^{\infty}\sum_{l=1}^{L} \sum_{\substack{T\in\widetilde{\mathcal{T}}_l \\ \mathcal{G}(T)=m}} \sum_{\boldsymbol{b}\in\mathcal{M}(T)} a\Big(I_l\widetilde{v}_l^{\boldsymbol{b}}, \sum_{k=1}^{l-1} \sum_{\substack{K\in\widetilde{\mathcal{T}}_k \\ \mathcal{G}(K)=n}} \sum_{\boldsymbol{a}\in\mathcal{M}(K)} I_k\widetilde{w}_k^{\boldsymbol{a}}\Big) \\ &\le C\sum_{m=0}^{\infty}\sum_{n=m}^{\infty}\theta^{\frac{n-m}{2}}\sum_{l=1}^{L} \sum_{\substack{T\in\widetilde{\mathcal{T}}_l \\ \mathcal{G}(T)=m}} \sum_{\boldsymbol{b}\in\mathcal{M}(T)} \|I_l\widetilde{v}_l^{\boldsymbol{b}}\|_A \Big(\sum_{k=1}^{l-1} \sum_{\substack{K\in\widetilde{\mathcal{T}}_k \\ \mathcal{G}(K)=n}} \sum_{\boldsymbol{a}\in\mathcal{M}(K)} \|I_k\widetilde{w}_k^{\boldsymbol{a}}\|_{A,\widetilde{\Omega}_l^{\boldsymbol{b}}}^2\Big)^{\frac12}. \end{aligned}$$

Since the matrix $\big(\theta^{|m-n|/2}\big)_{m,n=0}^{\infty}$ has the finite spectrum radius depending only on θ, we deduce

$$\begin{aligned} I_1 &\le C\sum_{m=0}^{\infty}\sum_{n=m}^{\infty}\theta^{\frac{n-m}{2}}\Big(\sum_{l=1}^{L} \sum_{\substack{T\in\widetilde{\mathcal{T}}_l \\ \mathcal{G}(T)=m}} \sum_{\boldsymbol{b}\in\mathcal{M}(T)} \|I_l\widetilde{v}_l^{\boldsymbol{b}}\|_A^2\Big)^{\frac12}\Big(\sum_{k=1}^{L} \sum_{\substack{K\in\widetilde{\mathcal{T}}_k \\ \mathcal{G}(K)=n}} \sum_{\boldsymbol{a}\in\mathcal{M}(K)} \|I_k\widetilde{w}_k^{\boldsymbol{a}}\|_A^2\Big)^{\frac12} \\ &\le C\Big(\sum_{m=0}^{\infty}\sum_{l=1}^{L} \sum_{\substack{T\in\widetilde{\mathcal{T}}_l \\ \mathcal{G}(T)=m}} \sum_{\boldsymbol{b}\in\mathcal{M}(T)} \|I_l\widetilde{v}_l^{\boldsymbol{b}}\|_A^2\Big)^{\frac12}\Big(\sum_{n=0}^{\infty}\sum_{k=1}^{L} \sum_{\substack{K\in\widetilde{\mathcal{T}}_k \\ \mathcal{G}(K)=n}} \sum_{\boldsymbol{a}\in\mathcal{M}(K)} \|I_k\widetilde{w}_k^{\boldsymbol{a}}\|_A^2\Big)^{\frac12} \\ &\le C\Big(\sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l}\big\|I_lv_l^i\big\|_A^2\Big)^{\frac12}\Big(\sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l}\big\|I_lw_l^i\big\|_A^2\Big)^{\frac12}. \end{aligned}$$

When $m>n$, the same arguments show that the remaining terms I_0-I_1 can also be bounded as the above estimate. Thus, we have

$$I_0 \le C\Big(\sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l}\big\|I_lv_l^i\big\|_A^2\Big)^{\frac12}\Big(\sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l}\big\|I_lw_l^i\big\|_A^2\Big)^{\frac12}. \tag{5.6}$$

Note that

$$\begin{aligned}
&\sum_{l=0}^{L}\sum_{i=1}^{\widetilde{n}_l}\sum_{k=0}^{l-1}\sum_{j=1}^{\widetilde{n}_k} a(I_l v_l^i, I_k w_k^j) \\
&= \sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l}\sum_{k=1}^{l-1}\sum_{j=1}^{\widetilde{n}_k} a(I_l v_l^i, I_k w_k^j) + \sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l} a(I_l v_l^i, I_0 w_0^1). \qquad (5.7)
\end{aligned}$$

It follows from (5.6) that

$$\begin{aligned}
\|\sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l} I_l v_l^i\|_A^2 &= 2\sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l}\sum_{k=1}^{l-1}\sum_{j=1}^{\widetilde{n}_k} a(I_l v_l^i, I_k v_k^j) + \sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l} \left\| I_l v_l^i \right\|_A^2 \\
&\le C \sum_{l=1}^{L}\sum_{i=1}^{\widetilde{n}_l} \left\| I_l v_l^i \right\|_A^2. \qquad (5.8)
\end{aligned}$$

Combining (5.7), (5.6), (5.8) and the Cauchy-Schwarz inequality yields (5.2) and completes the proof. □

In the next two subsections, we can apply the statements S1 and S2 to verify assumptions A1-A3 for the abstract convergence theory of algorithms LMPA and LMAA with local Jacobi and local Gauss-Seidel smoothers.

5.3. Local Jacobi smoother

The local Jacobi smoother is defined as an additive smoother (cf. [3]):

$$R_l := \gamma \sum_{i=1}^{\widetilde{n}_l} (A_l^i)^{-1} Q_l^i, \quad 1 \le l \le L,$$

where $\gamma \in (0,1)$ is a suitably chosen positive scaling factor. Let $R_0 = A_0^{-1}$. Due to (4.8), we have

$$T_0 = \mu_0 I_0 P_0,\ T_l = \mu_l I_l R_l A_l P_l = \mu_l \gamma I_l \sum_{i=1}^{\widetilde{n}_l} P_l^i P_l, \quad l = 1, \cdots, L. \qquad (5.9)$$

Lemma 5.3. *Let $\{T_l\}_{l=0}^{L}$ be defined by* (5.9)*. Then, there exist positive constants C and $\mu = \min_{0\le l\le L}\{\mu_l\}$ such that the assumption A1 is satisfied.*

Proof. Due to the decomposition of any $v \in V_L$ in (5.3) and $I_l v_l = v_l$, $l = 0, 1, \cdots, L$, where v_l is defined by (5.3), there holds

$$a(v,v) = \sum_{l=0}^{L} a(v_l, v) = \sum_{l=0}^{L} a(I_l v_l, v) = \sum_{l=0}^{L} a_l(v_l, P_l v). \qquad (5.10)$$

For $1 \le l \le L$, $v_l = \sum_{\boldsymbol{b}\in\widetilde{\mathcal{M}}_l} v_l^{\boldsymbol{b}} := \sum_{i=1}^{\widetilde{n}_l} v_l^i$, we have

$$\begin{aligned}
a_l(v_l, P_l v) &= \sum_{i=1}^{\widetilde{n}_l} a_l(v_l^i, P_l v) = \sum_{i=1}^{\widetilde{n}_l} a_l(v_l^i, P_l^i P_l v) \\
&\le \sum_{i=1}^{\widetilde{n}_l} a_l^{\frac{1}{2}}(v_l^i, v_l^i) a_l^{\frac{1}{2}}(P_l^i P_l v, P_l^i P_l v)
\end{aligned}$$

$$\leq \Big(\sum_{i=1}^{\widetilde{n}_l} a_l(v_l^i, v_l^i) \Big)^{\frac{1}{2}} \Big(\sum_{i=1}^{\widetilde{n}_l} a(I_l P_l^i P_l v, v) \Big)^{\frac{1}{2}}. \tag{5.11}$$

By (5.10) and (5.11), we deduce

$$\begin{aligned} a(v,v) &= \sum_{l=0}^{L} a_l(v_l, P_l v) \\ &\leq \Big(a_0(v_0, v_0) + \sum_{l=1}^{L} \sum_{i=1}^{\widetilde{n}_l} a_l(v_l^i, v_l^i) \Big)^{\frac{1}{2}} \cdot \Big(a(I_0 P_0 v, v) + \sum_{l=1}^{L} \sum_{i=1}^{\widetilde{n}_l} a(I_l P_l^i P_l v, v) \Big)^{\frac{1}{2}}. \end{aligned}$$

Combining the above estimate and the stability estimate in (5.4) yields

$$a(v,v) \leq \widetilde{C} \Big(a(I_0 P_0 v, v) + \sum_{l=1}^{L} \sum_{i=1}^{\widetilde{n}_l} a(I_l P_l^i P_l v, v) \Big) \leq \frac{\widetilde{C}}{\gamma\mu} \sum_{l=0}^{L} a(T_l v, v) = \frac{\widetilde{C}}{\gamma\mu} a(Tv, v),$$

where the positive constant $\widetilde{C}$ is independent of mesh sizes and mesh levels. We thus obtain the stated result by setting $C = \widetilde{C}/\gamma$. $\square$

Lemma 5.4. *Let $\{T_l\}_{l=0}^{L}$ be defined by* (5.9). *Then we can choose suitable scaling factors γ and μ_l such that*

$$a(T_l v, T_l v) \leq \omega_l \, a(T_l v, v), \quad v \in V_L, \ \omega_l < 2.$$

Proof. If $l = 0$, the stability of I_0 implies

$$a(T_0 v, T_0 v) \leq \mu_0^2 C_I a_0(P_0 v, P_0 v) = \mu_0 C_I a(T_0 v, v), \quad v \in V_L.$$

Let $\omega_0 = \mu_0 C_I$ and choose

$$\mu_0 < 2/C_I, \tag{5.12}$$

then $\omega_l < 2$. When $l \geq 1$, we set $T_l^i = \{P_l^j : \mathrm{supp}(I_l P_l^i v) \cap \mathrm{supp}(I_l P_l^j v) \neq \varnothing, \ v \in V_l\}$ and

$$\gamma_{ij} = \begin{cases} 1, & \text{if } \mathrm{supp}(I_l P_l^i v) \cap \mathrm{supp}(I_l P_l^j v) \neq \varnothing \ , \\ 0, & \text{otherwise.} \end{cases}$$

The cardinality of T_l^i is bounded by a constant depending only on the shape regularity of the meshes. For any $v \in V_l$, the Cauchy-Schwarz inequality implies

$$\begin{aligned} \sum_{i,j=1}^{\widetilde{n}_l} |a(I_l P_l^i v, I_l P_l^j v)| &= \sum_{i,j=1}^{\widetilde{n}_l} \gamma_{ij} |a(I_l P_l^i v, I_l P_l^j v)| \\ &\leq \Big(\sum_{i,j=1}^{\widetilde{n}_l} \gamma_{ij} a(I_l P_l^i v, I_l P_l^i v) \Big)^{\frac{1}{2}} \Big(\sum_{i,j=1}^{\widetilde{n}_l} \gamma_{ij} a(I_l P_l^j v, I_l P_l^j v) \Big)^{\frac{1}{2}} \\ &\leq C_l \sum_{i=1}^{\widetilde{n}_l} a(I_l P_l^i v, I_l P_l^i v). \end{aligned} \tag{5.13}$$

Taking advantage of the definition of T_l and the stability of I_l, we have

$$a(T_l v, T_l v) = \mu_l^2 \gamma^2 a(\sum_{i=1}^{\widetilde{n}_l} I_l P_l^i P_l v, \sum_{i=1}^{\widetilde{n}_l} I_l P_l^i P_l v) \leq \mu_l^2 \gamma^2 \sum_{i,j=1}^{\widetilde{n}_l} |a(I_l P_l^i P_l v, I_l P_l^j P_l v)|$$

$$\leq \mu_l^2 \gamma^2 C_l \sum_{i=1}^{\widetilde{n}_l} a(I_l P_l^i P_l v, I_l P_l^i P_l v) \leq \mu_l^2 \gamma^2 C_I C_l \sum_{i=1}^{\widetilde{n}_l} a_l(P_l^i P_l v, P_l^i P_l v)$$

$$= \mu_l^2 \gamma^2 C_I C_l \sum_{i=1}^{\widetilde{n}_l} a(I_l P_l^i P_l v, v) = \mu_l \gamma C_I C_l a(T_l v, v).$$

The proof is completed by setting $\omega_l = \mu_l \gamma C_I C_l$ and choosing

$$0 < \gamma < 1, \quad 0 < \mu_l < \frac{2}{\gamma C_I C_l} \tag{5.14}$$

such that $\omega_l < 2$. We remark that due to the fact that I_L is the identity on the finest level we may choose $\mu_L = 1$ and $0 < \gamma < 1$ such that $\omega_L = \gamma C_L < 2$. □

Lemma 5.5. *Let $\{T_l\}_{l=0}^L$ be defined by* (5.9)*. There exists a positive constant C independent of mesh sizes and mesh levels such that the assumption A3 is satisfied.*

Proof. Let $\gamma_0 = 1, \gamma_l = \gamma, 1 \leq l \leq L$. The global strengthened Cauchy-Schwarz inequality (5.2) and the definition of T_l imply that

$$\sum_{l=0}^{L} \sum_{k=0}^{l-1} a(T_l v_l, T_k w_k)$$
$$\leq C\Big(\sum_{l=0}^{L} \sum_{i=1}^{\widetilde{n}_l} \gamma_l^2 \mu_l^2 \left\|I_l P_l^i P_l v_l\right\|_A^2\Big)^{\frac{1}{2}} \Big(\sum_{l=0}^{L} \sum_{i=1}^{\widetilde{n}_l} \gamma_l^2 \mu_l^2 \left\|I_l P_l^i P_l w_l\right\|_A^2\Big)^{\frac{1}{2}}.$$

Let $\xi_l = v_l$ or w_l, $0 \leq l \leq L$. By the stability of I_l, (5.12) and (5.14), we have

$$\gamma_l^2 \mu_l^2 \left\|I_l P_l^i P_l \xi_l\right\|_A^2 \leq C_I \gamma_l^2 \mu_l^2 a(P_l^i P_l \xi_l, P_l^i P_l \xi_l) \leq C \gamma_l \mu_l a(I_l P_l^i P_l \xi_l, \xi_l),$$

whence

$$\sum_{i=1}^{\widetilde{n}_l} \gamma_l^2 \mu_l^2 \left\|I_l P_l^i P_l \xi_l\right\|_A^2 \leq C a(T_l \xi_l, \xi_l).$$

Combining the above estimates completes the proof of the lemma. □

5.4. Local Gauss-Seidel smoother

In this subsection, we will verify assumptions A1-A3 for the multilevel methods with a local Gauss-Seidel smoother R_l which is defined by

$$R_l = (I - E_l) A_l^{-1}, \quad E_l = (I - P_l^{\widetilde{n}_l}) \cdots (I - P_l^1), \; l \geq 1.$$

Let $R_0 = A_0^{-1}$. We have

$$T_0 = \mu_0 I_0 P_0 \;, \; T_l = \mu_l I_l R_l A_l P_l = \mu_l I_l (I - E_l) P_l \;, \quad l = 1, \cdots, L. \tag{5.15}$$

Note that $I - E_l = \sum_{i=1}^{\widetilde{n}_l} P_l^i E_l^{i-1}$, where $E_l^0 = I, E_l^i = (I - P_l^i)\cdots(I - P_l^1)$. As Lemma 4.5 in [37], there holds

$$a_l(P_l v, P_l u) - a_l(E_l P_l v, E_l P_l u) = \sum_{i=1}^{\widetilde{n}_l} a_l(P_l^i E_l^{i-1} P_l v, E_l^{i-1} P_l u), \quad v, u \in V_L. \tag{5.16}$$

Lemma 5.6. *Let $\{T_l\}_{l=0}^L$ be defined by* (5.15)*. Then there exists $\omega_l < 2$ such that*

$$a(T_l v, T_l v) \leq \omega_l \, a(T_l v, v), \quad v \in V_L.$$

Proof. When $l = 0$, the estimate has been obtained in Lemma 5.4. If $l \geq 1$, due to the definition of T_l we have

$$\begin{aligned} a(T_l v, T_l v) &= \mu_l^2 a(I_l(I - E_l)P_l v, I_l(I - E_l)P_l v) \\ &= \mu_l^2 \sum_{i,j=1}^{\widetilde{n}_l} a(I_l P_l^i E_l^{i-1} P_l v, I_l P_l^j E_l^{j-1} P_l v). \end{aligned}$$

Using (5.16), the similar techniques as in (5.13) and the stability of I_l, we obtain

$$\begin{aligned} a(T_l v, T_l v) &\leq \mu_l^2 C_l \sum_{i=1}^{\widetilde{n}_l} a(I_l P_l^i E_l^{i-1} P_l v, I_l P_l^i E_l^{i-1} P_l v) \\ &\leq \mu_l^2 C_I C_l \sum_{i=1}^{\widetilde{n}_l} a_l(P_l^i E_l^{i-1} P_l v, P_l^i E_l^{i-1} P_l v) \\ &= \mu_l^2 C_I C_l \Big(a_l(P_l v, P_l v) - a_l(E_l P_l v, E_l P_l v) \Big) \\ &= \mu_l^2 C_I C_l \Big(a_l(P_l v, P_l v) - a_l((I - (I - E_l))P_l v, (I - (I - E_l))P_l v) \Big) \\ &= \mu_l^2 C_I C_l \Big(2a_l((I - E_l)P_l v, P_l v) - a_l((I - E_l)P_l v, (I - E_l)P_l v) \Big) \\ &\leq \mu_l^2 C_I C_l \Big(2a_l((I - E_l)P_l v, P_l v) - \frac{1}{C_I} a(I_l(I - E_l)P_l v, I_l(I - E_l)P_l v) \Big) \\ &= 2\mu_l C_I C_l a(T_l v, v) - C_l a(T_l v, T_l v), \end{aligned} \tag{5.17}$$

whence

$$a(T_l v, T_l v) \leq \frac{2\mu_l C_I C_l}{1 + C_l} a(T_l v, v).$$

Setting $\omega_l = \frac{2\mu_l C_I C_l}{1+C_l}$, and choosing $\mu_l < \frac{1+C_l}{2C_I C_l}$ such that $\omega_l < 2$, the lemma is proved. We remark that we can choose $\mu_L = 1$, since I_L is the identity. $\square$

Lemma 5.7. *Let $\{T_l\}_{l=0}^L$ be defined by* (5.15)*. Then the assumption A*1 *is satisfied.*

Proof. In view of Lemma 5.3, for any $v \in V_L$, there holds

$$a(v, v) \leq C\Big(a(I_0 P_0 v, v) + \sum_{l=1}^{L} \sum_{i=1}^{\widetilde{n}_l} a_l(P_l^i P_l v, P_l^i P_l v) \Big). \tag{5.18}$$

By the identity $I - E_l^{i-1} = \sum_{j=1}^{i-1} P_l^j E_l^{j-1}$, we deduce

$$\sum_{i=1}^{\widetilde{n}_l} a_l(P_l^i P_l v, P_l^i P_l v) = \sum_{i=1}^{\widetilde{n}_l} a_l(P_l^i P_l v, P_l^i E_l^{i-1} P_l v) + \sum_{i=1}^{\widetilde{n}_l} \sum_{j=1}^{i-1} a_l(P_l^i P_l v, P_l^i P_l^j E_l^{j-1} P_l v)$$
$$\leq \Big(\sum_{i=1}^{\widetilde{n}_l} a_l(P_l^i P_l v, P_l^i P_l v) \Big)^{\frac{1}{2}} \Big(\sum_{i=1}^{\widetilde{n}_l} a_l(P_l^i E_l^{i-1} P_l v, E_l^{i-1} P_l v) \Big)^{\frac{1}{2}}$$
$$+ \sum_{i,j=1}^{\widetilde{n}_l} \big| a_l(P_l^i P_l v, P_l^j E_l^{j-1} P_l v) \big|.$$

Similar to (5.13), the local overlapping of subdomains $\{\widetilde{\Omega}_l^{\boldsymbol{b}} : \ \boldsymbol{b} \in \widetilde{\mathcal{M}}_l\}$ implies

$$\sum_{i,j=1}^{\widetilde{n}_l} \big| a_l(P_l^i P_l v, P_l^j E_l^{j-1} P_l v) \big|$$
$$\leq C \Big(\sum_{i=1}^{\widetilde{n}_l} a_l(P_l^i P_l v, P_l^i P_l v) \Big)^{\frac{1}{2}} \Big(\sum_{i=1}^{\widetilde{n}_l} a_l(P_l^i E_l^{i-1} P_l v, E_l^{i-1} P_l v) \Big)^{\frac{1}{2}}.$$

Furthermore, using the estimate in (5.17), we have

$$\sum_{i=1}^{\widetilde{n}_l} a_l(P_l^i P_l v, P_l^i P_l v) \leq C \sum_{i=1}^{\widetilde{n}_l} a_l(P_l^i E_l^{i-1} P_l v, E_l^{i-1} P_l v) \leq \frac{C}{\mu_l} a(T_l v, v). \tag{5.19}$$

The assertion follows from (5.18) and (5.19) that

$$a(v, v) \leq C \sum_{l=0}^{L} \frac{1}{\mu_l} a(T_l v, v) \leq \frac{C}{\mu} a(T v, v),$$

where $\mu = \min_{0 \leq l \leq L} \{\mu_l\}$. □

Lemma 5.8. *Let $\{T_l\}_{l=0}^L$ be defined by* (5.15)*. Then the assumption A*3 *is satisfied.*

Proof. By the global strengthened Cauchy-Schwarz inequality (5.2) and the definition of T_l, we have

$$\sum_{l=0}^{L} \sum_{k=0}^{l-1} a(T_l v_l, T_k w_k)$$
$$\leq C \Big(\sum_{l=0}^{L} \sum_{i=1}^{\widetilde{n}_l} \mu_l^2 \left\| I_l P_l^i E_l^{i-1} P_l v_l \right\|_A^2 \Big)^{\frac{1}{2}} \Big(\sum_{l=0}^{L} \sum_{i=1}^{\widetilde{n}_l} \mu_l^2 \left\| I_l P_l^i E_l^{i-1} P_l w_l \right\|_A^2 \Big)^{\frac{1}{2}}.$$

For $0 \leq l \leq L$, let $\xi_l = v_l$ or w_l. Using the estimate in (5.17), the stability of I_l and the choice of μ_l in Lemma 5.6, we deduce

$$\sum_{i=1}^{\widetilde{n}_l} \mu_l^2 \left\| I_l P_l^i E_l^{i-1} P_l \xi_l \right\|_A^2 \leq C a(T_l \xi_l, \xi_l).$$

In particular, we have $\mu_0^2 \left\| I_0 P_0 \xi_0 \right\|_A \leq C a(T_0 \xi_0, \xi_0)$. Combining the above estimates allows to conclude. □

6. Numerical Results

In this section, we present two 2D examples to illustrate the optimality of algorithm 4.1 and algorithm 4.2. The implementation is based on the FFW toolbox [9]. We mention here that the local multilevel methods and the corresponding convergence estimates can be extended to the nonhomogeneous Dirichlet boundary condition. The local error estimators and the MARK strategy for the selection of elements and edges for refinement have been realized as in the algorithm ANFEM in [13]. In the following examples, both LMPA and LMAA are considered as preconditioners for the conjugate gradient method, i.e., a symmetric version of LMPA (SLMPA) has been used in the computations. Likewise, a symmetric version of LMAA (SLMAA) is employed when the smoother is nonsymmetric, otherwise, LMAA is directly applied. The algorithms LMPA and LMAA require $O(N\log N)$ and $O(N)$ operations respectively, where N is the number of degrees of freedom (DOFs) (cf. [30]).

The estimate (4.1) in Lemma 4.1 indicates that the prolongation operator I_l from V_l to V_L would increase the energy by a constant C_I at worst, which is essential in the convergence analysis of the local multilevel methods. We can weaken the influence by a well chosen scaling number $\mu_{L,l}$ in (4.9). As seen from Theorem 5.1 and Theorem 5.2, the uniform convergence rate of LMPA or the preconditioned conjugate gradient method by LMAA will deteriorate for decreasing scaling number $\mu = \min_{0\le l\le L}\{\mu_{L,l}\}$. This property can be observed in the following Example 6.1. We always choose $\mu_{L,L} = 1$ in the computations.

At the lth level, the discrete problem is $\mathbf{A}_l\mathbf{u}_l = \mathbf{F}_l$, where $\mathbf{A}_l$ is the stiffness matrix. For the preconditioned conjugate gradient method, the iteration stops when it satisfies

$$\|\mathbf{r}_l^0 - \mathbf{A}_l\mathbf{r}_l^n\|_0 \le \epsilon\|\mathbf{r}_l^0\|_0, \quad \epsilon = 10^{-6},$$

where $\{\mathbf{r}_l^k : k = 1, 2, \cdots\}$ stands for the set of iterative solutions of the residual equation $\mathbf{A}_l\mathbf{x} = \mathbf{r}_l^0$.

Let $\mathbf{u}_l^0 = \widetilde{\mathbf{I}}_l\mathbf{I}_{l-1}\mathbf{u}_{l-1}$, where $\mathbf{I}_{l-1}$ and $\widetilde{\mathbf{I}}_l$ are matrix representation of the prolongation operators $I_{l-1} : V_{l-1} \to V_l^c$ and $\widetilde{I}_l : V_l^c \to V_l$ respectively, and $\mathbf{r}_l^n = \mathbf{F}_l - \mathbf{A}_l\mathbf{u}_l^n$, $n = 0, 1, 2, \cdots$, be the residual at nth iteration. Set

$$\epsilon_0 = (\mathbf{r}_l^0)^t\mathbf{B}_l\mathbf{r}_l^0, \qquad \epsilon_n = (\mathbf{r}_l^n)^t\mathbf{B}_l\mathbf{r}_l^n,$$

where $\mathbf{B}_l$ is defined by the local multilevel iteration. The number of iteration steps required to

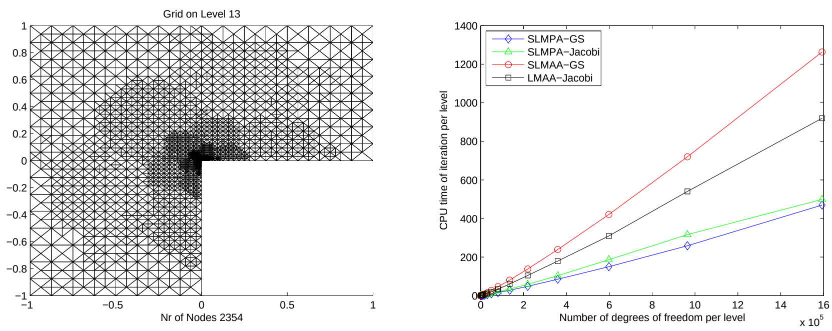

Fig. 6.1. Example 6.1: Locally refined mesh (left) at the 13th refinement level and CPU time (right) for SLMPA-GS, SLMPA-Jacobi, SLMAA-GS and LMAA-Jacobi.

Table 6.1: Example 6.1: Number of iterations and average reduction factor ρ on each level for the respective algorithms with scaling number $\mu_{L,l} = 0.8$ and $\mu_{L,L} = 1$, $0 \le l \le L-1, L \ge 1$. For the conjugate gradient method without preconditioning, only the number of iterations is given.

Level	DOFs	CG	SLMPA-GS		SLMPA-Jacobi		SLMAA-GS		LMAA-Jacobi	
		iter	iter	ρ	iter	ρ	iter	ρ	iter	ρ
13	6831	206	9	0.2203	12	0.3184	34	0.6732	46	0.7475
15	18121	310	10	0.2395	12	0.3179	35	0.6807	48	0.7567
17	49825	458	10	0.2430	12	0.3141	36	0.6853	49	0.7614
19	135060	700	10	0.2391	12	0.3079	35	0.6847	49	0.7624
20	219441	858	10	0.2405	12	0.3052	35	0.6838	50	0.7640
22	598091	1331	10	0.2353	12	0.2988	35	0.6845	49	0.7629
23	964580	1491	10	0.2356	12	0.2970	35	0.6848	50	0.7645
24	1592958	1715	10	0.2315	11	0.2873	35	0.6840	49	0.7631

Table 6.2: Example 6.1: Average reduction factors ρ (SLPMA-GS) for different scaling numbers.

Level	$\mu_{L,0} = \cdots = \mu_{L,L-1} = \alpha$, $\mu_{L,L} = 1$					
	$\alpha = 1.8$	$\alpha = 1.5$	$\alpha = 1$	$\alpha = 0.5$	$\alpha = 0.2$	$\alpha = 0.1$
13	0.2448	0.2340	0.2196	0.2737	0.4100	0.5125
15	0.2479	0.2410	0.2393	0.2738	0.4234	0.5292
17	0.2508	0.2444	0.2426	0.2722	0.4194	0.5274
19	0.2484	0.2408	0.2387	0.2668	0.4148	0.5225
20	0.2482	0.2419	0.2400	0.2567	0.4088	0.5215
22	0.2666	0.2368	0.2347	0.2488	0.4023	0.5182
23	0.2666	0.2371	0.2351	0.2451	0.3979	0.5088
24	0.2678	0.2331	0.2310	0.2423	0.3949	0.5058

achieve the desired accuracy is denoted by **iter**. We further denote by $\rho = (\sqrt{\epsilon_n}/\sqrt{\epsilon_0})^{1/\textbf{iter}}$ the average reduction factor.

Example 6.1. On the L-shaped domain $\Omega = [-1,1] \times [-1,1] \backslash (0,1] \times [-1,0)$, we consider the following elliptic boundary value problem

$$\begin{aligned} -\Delta(0.5u) + u &= f(x,y) \quad \text{in } \Omega, \\ u &= g(x,y) \quad \text{on } \partial\Omega, \end{aligned}$$

where f and g are chosen such that $u(r,\theta) = r^{\frac{2}{3}} \sin(\frac{2}{3}\theta)$ is the exact solution (in polar coordinates).

For ease of notation, we refer to SLMPA-GS, SLMPA-Jacobi, SLMAA-GS and LMAA-Jacobi as the preconditioned conjugate gradient method by SLMPA, SLMAA LMAA with local Gauss-Seidel smoother or local Jacobi smoother, respectively. For the Jacobi iteration, the scaling factor is chosen according to $\gamma = 0.8$.

At first, we choose $\mu_{L,l} = 0.8$ $(0 \le l < L)$ to illustrate the optimality of our algorithms. The left one of Fig. 6.1 displays the locally refined mesh at the 13th refinement level. As seen from Table 6.1, the number of iterative steps of the conjugate gradient method without preconditioning (CG) increases quickly with the mesh levels. However, for the algorithms SLMPA-GS, SLMPA-Jacobi, SLMAA-GS and LMAA-Jacobi we observe that the number of iteration steps and the average reduction factors are all bounded independently of mesh sizes and mesh levels. These results and the right one of Fig. 6.1, displaying the CPU times (in seconds) for the respective algorithms, demonstrate the optimality of the algorithms and thus confirm the theoretical analysis.

Next, we choose different scaling numbers to illustrate how they influence the convergence behavior of the local multilevel methods. We only list the results for SLMPA-GS. A similar behavior can be observed for the other algorithms. We choose $\mu_{L,0} = \cdots = \mu_{L,L-1} = \alpha$ and $\mu_{L,L} = 1$, and thus $\mu = \min\{\alpha, 1\}$. Table 6.2 shows that for a fixed α, SLMPA-GS converges almost uniformly. The last four numbers of each row in Table 6.2 show that for a fixed level the average reduction factor of SLMPA-GS deteriorates for decreasing μ. We also note that for $\mu = 1$ the convergence rate of SLMPA-GS deteriorates only with respect to ω_l (the spectral bound of T_l), which increases linearly with $\mu_{L,l}$. If $\alpha \geq 1$, then $\mu = \min\{\alpha, 1\} = 1$, and the convergence rate will also deteriorate as α increases. This is also observed for the first numbers of each row in Table 6.2. From Table 6.2 we can also observe that although $\mu_{L,l}$ will influence the convergence of local multilevel methods, the algorithms are not very sensitive with the the choice of scaling factor.

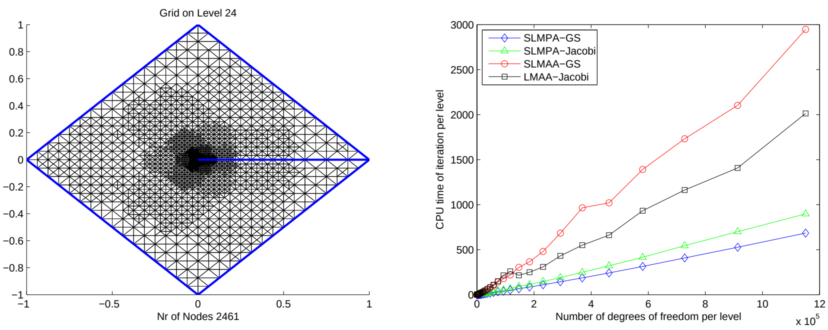

Fig. 6.2. Example 6.2: Locally refined mesh (left) at the 24th refinement level and CPU time (right) for SLMPA-GS, SLMPA-Jacobi, SLMAA-GS and LMAA-Jacobi.

Table 6.3: Example 6.2: Number of iterations and average reduction factors ρ on each level for the respective algorithms. For the conjugate gradient method without preconditioning, only the number of iterations is given.

Level	DOFs	CG	SLMPA-GS		SLMPA-Jacobi		SLMAA-GS		LMAA-Jacobi	
		iter	iter	ρ	iter	ρ	iter	ρ	iter	ρ
28	18206	287	11	0.2610	14	0.3756	51	0.7720	65	0.8154
32	46105	417	10	0.2403	14	0.3615	52	0.7745	65	0.8161
34	73571	523	11	0.2628	14	0.3773	55	0.7854	70	0.8271
36	116866	634	10	0.2511	13	0.3309	52	0.7768	63	0.8105
40	292148	942	10	0.2513	14	0.3708	55	0.7880	70	0.8286
42	462599	1168	10	0.2395	12	0.3181	52	0.7765	64	0.8141
44	727564	1404	10	0.2337	13	0.3511	53	0.7808	68	0.8243
46	1150917	1536	10	0.2435	14	0.3615	54	0.7852	70	0.8275

Example 6.2. We consider Poisson's equation

$$-\Delta u = 1 \qquad \text{in}\, \Omega,$$

with Dirichlet boundary conditions on a domain with a crack, namely $\Omega = \{(x, y) : |x| + |y| \leq 1\}\backslash\{(x, y) : 0 \leq x \leq 1, y = 0\}$. The exact solution is $r^{\frac{1}{2}} \sin(\theta/2) - \frac{1}{4}r^2$ (in polar coordinates).

In this example, we choose $\mu_{L,L} = 1$, $\mu_{L,l} = 1$ and $\mu_{L,l} = 0.8$ ($0 \leq l < L, L \geq 1$), respectively, for the local multilevel methods with local Gauss-Seidel smoother and local Jacobi smoother.

The left one of Fig. 6.2 displays the locally refined mesh at the 24th refinement level. The numbers in Table 6.3 and the CPU times (in seconds) displayed in the right one of Fig. 6.2 show a similar behavior as in the previous example and thus also support the theoretical findings.

Acknowledgements. The work of the first author was supported by the National Basic Research Program under the Grant 2011CB30971 and National Science Foundation of China (11171335). The work of the second author was supported by the National Natural Science Foundation of China (Grant No. 11201394) and the Fundamental Research Funds for the Central Universities (Grant No. 2012121003).

References

[1] 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.

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

[3] J.H. Bramble, Multigrid Methods, Pitman, 1993.

[4] J.H. Bramble, J.E. Pasciak, and J. Xu, The Analysis of Multigrid algorithms with nonnested spaces or non-inherited quadratic forms, *Math. Comp.*, **56** (1991), 1–34.

[5] S.C. Brenner, An optimal multigrid method for P1 nonconforming finite elements, *Math. Comp.*, **52** (1989), 1–16.

[6] S.C. Brenner, Convergence of nonconforming multigrid methods without full elliptic regularity, *Math. Comp.*, **68** (1999), 25–53.

[7] S.C. Brenner, Convergence of nonconforming V-cycle and F-cyle multigrid algorithms for second order elliptic boundary value problems, *Math. Comp.*, **73** (2004), 1041–1066.

[8] S.C. Brenner, J. Cui, and L. Sung, Multigrid Methods for the Symmetric Interior Penalty Method on Graded Meshes, *Numer. Linear Algebra Appl.*, **16** (2009), 481–501.

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

[10] C. Carstensen, Quasi-interpolation and a posteriori error analysis in finite element methods, *RAIRO Model. Math. Anal. Numer.*, **33** (1999), 1187–1202.

[11] C. Carstensen and R.H.W. Hoppe, Convergence analysis of an adaptive nonconforming finite element method, *Numer. Math.*, **103** (2006), 251–266.

[12] 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.

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

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

[15] W. Dahmen and A. Kunoth, Multilevel preconditioning, *Numer. Math.*, **63** (1992), 315–344.

[16] H. Duan, S. Gao, R. Tan, and S. Zhang, A generalized BPX framework covering the V-cycle nonnested multigrid methods, *Math. Comp.*, **76** (2007), 137–152.

[17] R. Hiptmair, H. Wu, and W. Zheng, On uniform convergence theory of local multigrid methods in $H^1(\Omega)$ and $\boldsymbol{H}(\mathbf{curl}, \Omega)$, preprint, 2010.

[18] R.H.W. Hoppe and B. Wohlmuth, Adaptive multilevel iterative techniques for nonconforming finite element discretizations, *East-West J. Numer. Math.*, **3** (1995), 179–197.

[19] K. Mekchay and R.H. Nochetto, Convergence of adaptive finite element methods for general second order linear elliptic PDEs, *SIAM J. Numer. Anal.*, **43** (2005), 1803-1827.

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

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

[22] P. Oswald, Intergrid transfer operators and multilevel preconditioners for nonconforming discretizations, *Appl. Numer. Math.*, **23** (1996), 139–158.

[23] P. Oswald, On a hierarchical basis multilevel method with nonconforming P1 elements, *Numer. Math.*, **62** (1992), 189–212.

[24] P. Oswald, On a BPX-preconditioner for P1 elements, *Computing*, **51** (1993), 125–133.

[25] P. Oswald, Optimality of multilevel preconditioning for nonconforming P1 finite elements, *Numer. Math.*, **111** (2008), 267–291.

[26] R. Scott and S. Zhang, Finite element interpolation of nonsmooth functions satisfying boundary conditions, *Math. Comp.*, **54** (1990), 483–493.

[27] Z. Shi and X. Xu, On a note for nonconforming multigrid for nonsymmetric and indefinite problems, *Chinese J. Numer. Math. Appl.*, **22** (2000), 102–108.

[28] R. Stevenson, Optimality of a standard adaptive finite element method, *Foundations of Computional Mathematics*, **2** (2007), 245–269.

[29] A. Toselli and O. Widlund, Domain Decomposition Methods – Algorithms and Theory, Springer-Verlag, Berlin, 2005.

[30] P.S. Vassilevski and J. Wang, An application of the abstract multilevel theory to nonconforming finite element methods, *SIAM J. Numer. Anal.*, **32** (1995), 235–248.

[31] S.P. Wong, Preconditioning nonconforming finite element methods for treating Dirichlet boundary conditions, II, *Numer. Math.*, **62** (1992), 413–437.

[32] S.P. Wong, Preconditioning P1 nonconforming finite elements: condition numbers and singular value distributions, *Numer. Math.*, **65** (1993), 253–271.

[33] 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.

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

[35] 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.

[36] X. Xu and J. Chen, Multigrid for the Mortar element method for P1 nonconforming element, *Numer. Math.*, **88** (2001), 381–398.

[37] 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.

[38] X. Xu, L. Li, and W. Chen, A multigrid method for the Mortar-type Morley element approximation of a plate bending problem, *SIAM J. Numer. Anal.*, **39** (2002), 1712–1731.

[39] H. Yserentant, Old and new convergence proofs for multigrid methods, *Acta Numerica*, **2** (1993), 285–326.