TREATMENTS OF DISCONTINUITY AND BUBBLE FUNCTIONS IN THE MULTIGRID METHOD

SHANGYOU ZHANG AND ZHIMIN ZHANG

ABSTRACT. When multilevel finite element spaces are not nested, different intergrid transfer operators would lead to different multigrid algorithms. It is proposed in this paper that discontinuous functions be averaged to continuous functions and that the bubble functions be discarded in the multigrid transferring. Applications of the techniques to various problems are presented with convergence analysis. Numerical comparisons with other existing methods are provided.

1. INTRODUCTION

The multigrid method provides optimal-order algorithms for solving large linear systems of finite element and finite difference equations (cf. [3]). The multigrid theory is well established (cf. [12], [16], [3], [15] and [5]). However, in many situations, the multi-level discrete spaces are nonnested due to the nature of the underlying finite elements ([4], [7]–[8], [18], [25]) or due to the special structures of grids (cf. [6], [9] [20], [26], [27]). Some special treatments are then needed to extend the standard multigrid method.

Brenner [7] and Braess-Verfürth [4] have studied P_1 nonconforming multigrid methods. The *intergrid transfer* operator in [7] averages the midpoint values of coarse-level functions on the two sides of each edge. This operator is generalized in [4], where the averaging is weighted by the areas of two neighboring triangles, such that the nonconforming multigrid method can be covered by the standard multigrid theory. However, the analysis of [4] is not valid for variable coefficient equations. Brenner [8] and Verfürth [25] also have treated bubble functions in multigrid methods. They have two different averaging operators to transfer a coarse-level bubble (a cubic polynomial vanishing at three edges of a triangle) to four small bubbles on the four subtriangles. In this paper, numerical tests on [7], [4], [8], [25] will be presented along with some tests on several new algorithms.

Different from [7], [8], [4], [25], we average a discontinuous finite element function to a continuous one which is naturally in the higher-level nonconforming finite element (NCF) space. We also discard bubble functions in transferring them to higher levels. By using averaging functions and discarding bubbles we would still retain the smooth components of coarse-level solutions. It is shown that the new multigrid methods converge with constant rates and retain the optimal order of computation for several model problems. The direct estimate of strengthened Cauchy inequalities in [25] would limit the extension of the theory there, while the analysis here

is valid in higher-dimensional spaces and for higher degree elements with bubbles (cf. [2]).

The rest of the paper is organized as follows. In §2, the P_1 nonconforming multigrid method is defined. The constant rate of convergence and the optimal order of computation for the nonconforming multigrid method are proved in §3. Raviart-Thomas mixed triangular elements and their nonconforming finite element correspondence are introduced in §4. Section 5 is devoted to solving Raviart-Thomas finite element equations via nonconforming multigrid methods. A brief discussion on the multigrid method for the Wilson element is given in §6. A multigrid iteration for solving mini-element equations arising from discretizing stationary Stokes equations is defined in §7 and the convergence analysis is presented in §8. We present some numerical results in §9 with discussions on implementing the multigrid method. In some numerical computations, the cubic bubble functions are replaced by linear macro-bubbles to avoid higher order quadratures.

2. P1 NONCONFORMING ELEMENTS

For simplicity, we consider the following Poisson equation in a convex, bounded, polygonal domain in $\mathbf{R}^2$:

$$(2.1) \qquad -\Delta u = f \quad \text{in } \Omega, \qquad u = 0 \quad \text{on } \partial\Omega,$$

and its weak formulation: Find $u \in H_0^1(\Omega)$, such that

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

where $a(u, v) := \int_\Omega \nabla u \cdot \nabla v \, dx$, $(f, v) := \int_\Omega fv \, dx$ and $H_0^1(\Omega)$ is the space of Sobolev $W^{1,2}(\Omega)$ functions with zero trace on $\partial\Omega$. We note that the analysis here covers general self-adjoint elliptic problems.

Let $\{\mathcal{T}_k\}$ be a nested family of triangulations on Ω, where $\mathcal{T}_k$ is obtained by refining each triangle in $\mathcal{T}_{k-1}$ into four congruent subtriangles by connecting midpoints on edges. The mesh sizes of these quasiuniform grids have the relation $h_k = h_{k-1}/2$. The work here can be extended to the case of nonquasiuniform meshes as in [27]. Let V_k be the P_1 nonconforming space on grid $\mathcal{T}_k$, i.e., the space of all piecewise linear functions on $\mathcal{T}_k$ which are continuous at midpoints of all edges and vanish at midpoints of all boundary edges. The finite element problems arising from discretizing (2.1) read: Find $u_k \in V_k$, such that

$$(2.2) \qquad a_k(u_k, v) = (f, v) \qquad \forall v \in V_k,$$

where $a_k(u, v) := \sum_{K \in \mathcal{T}_k} \int_K \nabla u \cdot \nabla v \, dx$ for all $u, v \in V_k + H_0^1(\Omega)$. It is known from [22] that

$$(2.3) \qquad \|u - u_k\|_{L^2(\Omega)} + h_k |u - u_k|_{1,k} \leq C h_k^2 \|u\|_{H^2(\Omega)},$$

where u and u_k are solutions of (2.1) and (2.2) respectively, and $|v|_{1,k} := \sqrt{a_k(v, v)}$. Here and later, C stands for a generic constant independent of the multigrid index k.

Since $V_{k-1} \not\subset V_k$, we introduce auxiliary, P_1 conforming finite element spaces $\{\widetilde{V}_k\}$ (cf. [10]). We note that $\widetilde{V}_k = V_k \cap H_0^1(\Omega)$. It holds that ([10])

$$(2.4)$$
$$\inf_{v \in \widetilde{V}_k \subset V_k} \left\{ \|u - v\|_{L^2(\Omega)} + h_k |u - v|_{1,k} \right\} \leq C h_k^2 \|u\|_{H^2(\Omega)} \quad \forall u \in H^2(\Omega) \cap H_0^1(\Omega).$$

Noting that $\widetilde{V}_{k-1}$ is a subspace of both V_{k-1} and of V_k (in fact $\widetilde{V}_{k-1} = V_{k-1} \cap V_k$), we define an intergrid transfer operator $\widetilde{I}_k \ : \ V_{k-1} \to \widetilde{V}_{k-1} \subset V_k$: For any $v \in V_{k-1}$,

$$(2.5) \qquad (\widetilde{I}_k v)(a) = \frac{1}{j_a} \sum_{j=1}^{j_a} v|_{K_j}(a)$$

and $(\widetilde{I}_k v)(a) = 0$ if $a \in \partial\Omega$, where a is a vertex in $\mathcal{T}_{k-1}$ and $\{K_j\}_{j=1}^{j_a}$ are all the triangles in $\mathcal{T}_{k-1}$ which have a as a vertex. We will introduce a few more intergrid transfer operators in §9.

Let $A_k \ : \ V_k \to V_k$ be defined by $(A_k u, v) = a_k(u, v) \quad \forall u, v \in V_k$, and

$$(2.6) \qquad |v|_{s,k} := \sqrt{(A_k^s v, v)} \quad \forall v \in V_k, 0 \le s \le 2.$$

We note that $|v|_{0,k} = \|v\|_{L^2(\Omega)}$ and that $|v|_{1,k} = \sqrt{a_k(v, v)}$. Let Λ_k be the spectral radius of A_k. By a simple calculation (cf. [10]) on each triangle, we have the inverse estimate: $\Lambda_k \le C h_k^{-2}$. We now define a multigrid scheme for solving (2.2) (cf. [3]).

Definition 2.1 (one kth-level multigrid iteration).

(1) For $k = 1$, the original problem (2.2) or the residual problem (2.8a) defined below is solved exactly.

(2) For $k > 1$, w_{m+1} will be generated from w_0 by the following two steps. We do m smoothings:

$$(2.7) \quad (w_l - w_{l-1}, v) = \Lambda_k^{-1}((f, v) - a_k(w_{l-1}, v)) \qquad \forall v \in V_k, \quad l = 1, 2, ..., m,$$

and a coarse-level correction: $w_{m+1} = w_m + \widetilde{I}_k \epsilon$, where $\epsilon \in V_{k-1}$ is obtained by doing $p \ (> 1)$ (W-cycle iterations) $(k-1)$st-level multigrid iterations starting with 0 guess for the following residual problem: Find $\epsilon_a \in V_{k-1}$, such that

$$(2.8a) \qquad a_{k-1}(\epsilon_a, v) = (f, \widetilde{I}_k v) - a_k(w_m, \widetilde{I}_k v) =: (f_{\text{new}}, v) \quad \forall v \in V_{k-1}.$$

In nested multigrid methods, the prolongation operator $\widetilde{I}_k$ in Definition 2.1 is simply the identity operator. Indeed, noting $\widetilde{V}_k \subset V_k$, we can define such a coarse-level correction: Find $\widetilde{\epsilon} = \epsilon_b \in \widetilde{V}_{k-1}$ (or $\widetilde{\epsilon} = \epsilon_c \in \widetilde{V}_k$, resp.), such that

$$(2.8b,c) \qquad a(\widetilde{\epsilon}, v) = (f, v) - a_k(w_m, v) =: (f_{\text{new}}, v) \quad \forall v \in \widetilde{V} \subset V_k,$$

where in (2.8b) $\widetilde{V} = \widetilde{V}_{k-1}$ (in (2.8c) $\widetilde{V} = \widetilde{V}_k$, resp.) We note that the method with (2.8b) or (2.8c) is covered by a standard multigrid analysis (see the proof for Theorem 3.3 below or [3]). To use (2.8b) or (2.8c), one needs to implement both conforming and nonconforming elements. But for the purpose of comparison, we still include methods (2.8b,c) in our tests in §9.

3. CONVERGENCE OF P1 NONCONFORMING MULTIGRID METHODS

The central task is to estimate the perturbation to the orthogonal-projection property caused by the intergrid transfer operator in the coarse-level correction (2.8a). The perturbation relates strongly to the stability of $\widetilde{I}_k$, which will be studied in the next lemma.

Lemma 3.1. *It holds that*

$$(3.1) \qquad |\widetilde{I}_k v|_{l,k} \le C|v|_{l,k-1} \quad \forall v \in V_{k-1} \quad for \ l = 0, 1.$$

Proof. The proof uses a straightforward triangle by triangle argument. Let $K \in \mathcal{T}_{k-1}$ have vertices $\{a_1, a_2, a_3\}$, $\{K_{ij}\}_{j=1}^{j_i} \subset \mathcal{T}_{k-1}$ having a_i as a vertex, and let $v_{ij} = v|_{K_{ij}}(a_i)$.

$$(3.2) \quad \|\widetilde{I}_k v\|_{L^2(K)}^2 \leq |K| \sum_{i=1}^{3} (\widetilde{I}_k v)^2(a_i) \leq C h_{k-1}^2 \sum_{i=1}^{3} \sum_{j=1}^{j_i} v_{ij}^2 j_i^{-2} \leq C \|v\|_{L^2(\bigcup_{i,j} K_{ij})}^2.$$

Summing (3.2) over all triangles in $\mathcal{T}_{k-1}$, (3.1) is proved for $l = 0$ where C depends only on the smallest angle (the same on all levels) in $\mathcal{T}_1$. Let w_{ij} be the nodal value of v at the midpoint of the edge shared by two triangles K_{ij} and $K_{i,j+1}$ (we let $K_{i,j_i+1} = K_{i1}$). Here we list $\{K_{ij}\}$ such that $w_{11} = w_{3j_3}$ and $w_{1j_1} = w_{2j_2}$. As $\widetilde{I}_k$ is an averaging operator, we get a telescoping sum in the estimate below.

$$\|\nabla \widetilde{I}_k v\|_{L^2(K)}^2 \leq C h_{k-1}^2 \left(\left(\widetilde{I}_k v(a_2) - \widetilde{I}_k v(a_1) \right)^2 + \left(\widetilde{I}_k v(a_3) - \widetilde{I}_k v(a_1) \right)^2 \right)$$

$$= C h_{k-1}^2 \left(\sum_{j=1}^{j_2-1} \frac{j}{j_2} (v_{2j} - w_{2j} + w_{2j} - v_{2,j+1}) + (v_{2j_2} - w_{2j_2}) \right.$$

$$\left. + w_{1j_1} - v_{1j_1}) - \sum_{j=1}^{j_1-1} \frac{j}{j_1} (v_{1j} - w_{1j} + w_{1j} - v_{1,j+1}) \right)^2$$

$$+ C h_{k-1}^2 \left(\sum_{j=1}^{j_3-1} \frac{j}{j_3} (v_{3j} - w_{3j} + w_{3j} - v_{3,j+1}) + (v_{3j_3} - w_{3j_3}) \right.$$

$$\left. + w_{11} - v_{11}) - \sum_{j=1}^{j_1-1} \frac{j_1 - j}{j_1} (v_{1j} - w_{1j} + w_{1j} - v_{1,j+1}) \right)^2$$

$$\leq C \|\nabla v\|_{L^2(\bigcup_{i,j} K_{ij})}^2.$$

Summing over all triangles, (3.1) is proved for $l = 1$. $\square$

To analyze the coarse-level corrections (2.8a–c), we introduce the following operators:

$$(3.3a) \qquad P_a : V_k \quad \rightarrow V_{k-1}, \qquad a_{k-1}(P_a v, w) = a_k(v, \widetilde{I}_k w) \quad \forall w \in V_{k-1},$$

$$(3.3b) \qquad P_b : D(P_b) \rightarrow \widetilde{V}_{k-1}, \qquad a(P_b v, w) = a_k(v, w) \qquad \forall w \in \widetilde{V}_{k-1},$$

where $D(P_b) = V_k + V_{k-1} + H_0^1(\Omega)$ in (3.3b). We note that P_b is an $a_k(\cdot, \cdot)$ orthogonal-projection operator. In particular, by (2.4) $|u - P_b u|_{1,k-1} \leq C h_{k-1} \|u\|_{H^2(\Omega)}$.

Lemma 3.2. *For any $v \in V_l$, $l = k - 1$ or $l = k$, it holds that*

$$|v - P_b v|_{0,l} \leq C h_l |v - P_b v|_{1,l}.$$

Proof. We use a standard duality argument. Let u be the solution of (2.1) with $f = v - P_b v$. Let u_l be the corresponding finite element solution of (2.2). Noting $P_b v \in \widetilde{V}_{k-1} \subset V_l$, the lemma is proved by the following estimate:

$$|v - P_b v|_{0,l}^2 = a_l(u_l, v - P_b v) = a_l(u_l - P_b u, , v - P_b v)$$

$$\leq |u_l - P_b u|_{1,l} \, |v - P_b v|_{1,l} \leq (|u_l - u|_{1,l} + |u - P_b u|_{1,l})|v - P_b v|_{1,l}$$

$$\leq C h_l \|u\|_{H^2(\Omega)} |v - P_b v|_{1,l} \leq C h_l |v - P_b v|_{0,l} |v - P_b v|_{1,l},$$

where (2.3)–(2.4) and $\|u\|_{H^2} \leq C\|f\|_{L^2} = C|v - \mathrm{P}_b v|_{0,l}$ (since Ω is convex, cf. [11]) are applied. $\qquad\square$

Theorem 3.3 (The two-level method). *Let $\epsilon = \epsilon_a$ in Definition 2.1. Given any $0 < \gamma < 1$, there exists an integer m independent of k such that*

$$(3.4) \qquad |u_k - w_{m+1}|_{1,k} \leq \gamma |u_k - w_0|_{1,k},$$

where u_k is defined in (2.2) and w_l in Definition 2.1.

Proof. Let the errors be denoted by $e_l = u_k - w_l$ for $0 \leq l \leq m + 1$. We need to show $|e_{m+1}|_{1,k} \leq \gamma |e_0|_{1,k}$. By Definition 2.1 and (3.3a–b), we have $\epsilon_a = \mathrm{P}_a e_m$ and consequently we get that

$$(3.5) \qquad |e_{m+1}|_{1,k} = |e_m - \widetilde{\mathrm{I}}_k \epsilon_a|_{1,k} = |e_m - \widetilde{\mathrm{I}}_k \mathrm{P}_a e_m|_{1,k}$$
$$\leq |e_m - \mathrm{P}_b e_m|_{1,k} + |\mathrm{P}_b e_m - \widetilde{\mathrm{I}}_k \mathrm{P}_a e_m|_{1,k}.$$

The first term in (3.5) can be estimated by Lemma 3.2 as follows, noting P_b is an $a_k(\cdot, \cdot)$ orthogonal-projection operator.

$$(3.6) \qquad |e_m - \mathrm{P}_b e_m|_{1,k}^2 \leq |e_m - \mathrm{P}_b e_m|_{0,k} |e_m|_{2,k} \leq Ch_k |e_m - \mathrm{P}_b e_m|_{1,k} |e_m|_{2,k}.$$

For the second term in (3.5), we derive by Lemma 3.1 that

$$(3.7) \qquad |\mathrm{P}_b e_m - \widetilde{\mathrm{I}}_k \mathrm{P}_a e_m|_{1,k} = |\widetilde{\mathrm{I}}_k (\mathrm{P}_b e_m - \mathrm{P}_a e_m)|_{1,k} \leq C|\mathrm{P}_b e_m - \mathrm{P}_a e_m|_{1,k-1}.$$

Noting that $\widetilde{\mathrm{I}}_k w = w$ for all $w \in \widetilde{V}_{k-1}$ and that

$$a(\mathrm{P}_b \mathrm{P}_a e_m, w) = a_{k-1}(\mathrm{P}_a e_m, w) = a_k(e_m, \widetilde{\mathrm{I}}_k w) = a_k(e_m, w) = a(\mathrm{P}_b e_m, w),$$

we have $\mathrm{P}_b \mathrm{P}_a e_m = \mathrm{P}_b e_m$. Therefore, we can apply Lemmas 3.1 and 3.2 to estimate the right-hand side of (3.7) as follows. For any $v \in V_{k-1}$,

$$(3.8) \qquad a_{k-1}(\mathrm{P}_b \mathrm{P}_a e_m - \mathrm{P}_a e_m, v) = a_{k-1}(\mathrm{P}_a e_m, \mathrm{P}_b v - v)$$
$$= a_k(e_m, \widetilde{\mathrm{I}}_k(\mathrm{P}_b v - v)) \leq |e_m|_{2,k} |\widetilde{\mathrm{I}}_k(\mathrm{P}_b v - v)|_{0,k}$$
$$\leq |e_m|_{2,k}(C|\mathrm{P}_b v - v|_{0,k-1}) \leq Ch_k |e_m|_{2,k} |\mathrm{P}_b v - v|_{1,k-1}$$
$$\leq Ch_k |e_m|_{2,k} |v|_{1,k-1}.$$

Therefore, the second term in (3.5) can be estimated by (3.7) and (3.8) as

$$(3.9) \qquad |\mathrm{P}_b e_m - \widetilde{\mathrm{I}}_k \mathrm{P}_a e_m|_{1,k} \leq C|\mathrm{P}_b \mathrm{P}_a e_m - \mathrm{P}_a e_m|_{1,k-1} \leq Ch_k |e_m|_{2,k}.$$

For the smoothing iteration (2.7) we have (cf. [3] for example)

$$(3.10) \qquad |e_m|_{2,k} \leq Cm^{-1/2} \Lambda_k^{-1/2} |e_0|_{1,k} \leq Cm^{-1/2} h_k^{-1} |e_0|_{1,k}.$$

Combining (3.5)–(3.6) and (3.9)–(3.10), (3.4) is proved by letting $m > (C/\gamma)^2$:

$$|e_{m+1}| \leq Cm^{-1/2} |e_0|_{1,k} \leq \gamma |e_0|_{1,k}. \quad \square$$

By a standard argument in [3] and Theorem 3.3, one can show the convergence for W-cycle iterations and the theorem of optimal-order computation for the multigrid method.

4. The lowest order Raviart-Thomas elements

We consider first to solve (2.1) by the triangular Raviart-Thomas elements of the lowest order (cf. [19]):

$$RT_0^0(\mathcal{T}_k) := \left\{ \boldsymbol{\tau} \ : \ \boldsymbol{\tau}|_K = \begin{pmatrix} a + bx \\ c + by \end{pmatrix} \ \forall K \in \mathcal{T}_k, \text{ with normal}$$

components continuous on edges of elements in $\mathcal{T}_k\}$,

$$M_{-1}^0(\mathcal{T}_k) := \{v \ : \ v \text{ is piecewise constant on elements in } \mathcal{T}_k\}.$$

We omit the mixed formulation and the mixed element equation (cf. [1]). We will study the multigrid method for the equivalent (established in [1]) nonconforming finite element problems: Find $\phi_k \in U_k$ such that

$$(4.1) \qquad \sum_{K \in \mathcal{T}_k} \int_K \mathrm{P}_{RT_{-1}^0(\mathcal{T}_k)}^0 \nabla \phi_k \cdot \nabla v \, dx = \int \mathrm{P}_{M_{-1}^0(\mathcal{T}_k)}^0 f \cdot v \, dx \quad \forall v \in U_k,$$

where P_V^0 denotes the L^2 orthogonal-projection onto V and $RT_{-1}^0(\mathcal{I}_k) := \{\boldsymbol{\tau} \ : \ \boldsymbol{\tau}|_K = (a + bx, c + by) \ \forall K \in \mathcal{T}_k\}$. In (4.1) the multilevel nonconforming spaces are

$$U_k := V_k + B_k,$$

where V_k is the P_1 nonconforming space, and B_k is the space of piecewise cubic polynomials vanishing on all edges of $\mathcal{T}_k$. We remark that $v = v^l + v^b$ is unique where $v^l \in V_k$ and $v^b \in B_k$. Because $\mathrm{P}_{RT_{-1}^0(\mathcal{T}_k)}^0 \nabla u^l = \nabla u^l$ and because $\int_K \nabla u^l \cdot \nabla v^b \, dx = 0$, (4.1) can be decoupled as

$$\sum_{K \in \mathcal{T}_k} \int_K \nabla \phi_k^l \cdot \nabla v \, dx = \int_\Omega \mathrm{P}_{M_{-1}^0(\mathcal{T}_k)}^0 f v \, dx \quad \forall v \in V_k,$$

$$\sum_{K \in \mathcal{T}_k} \int_K \mathrm{P}_{RT_{-1}^0(\mathcal{T}_k)}^0 \nabla \phi_k^b \cdot \nabla v \, dx = \int_\Omega \mathrm{P}_{M_{-1}^0(\mathcal{T}_k)}^0 f v \, dx \quad \forall v \in B_k.$$

The first equation is exactly the P1 nonconforming equation (2.2), which can be solved by the multigrid method defined in §2. The second equation has a diagonal coefficient matrix. The equation is solved trivially. Because the two equations are uncoupled, and because we can recover the Raviart-Thomas solution from ϕ_k with a linear order of computation (see [1]), we get the following theorem by Theorem 3.3 (cf. [3]).

Theorem 4.1 (Optimal order MG in solving R-T equations). *The algorithm described in Definition 2.1 and above can produce Raviart-Thomas mixed element solutions up to the order of truncation error by order* $\dim(U_k)$ *(optimal order) computations.*

However, for a general-coefficient problem $\mathrm{div}(-\mathbf{A}(x)\nabla u) = f$, the above uncoupling of the linear and cubic elements in the linear system is not possible. We will consider next (cf. [1]) to solve

$$(4.2) \qquad \sum_{K \in \mathcal{T}_k} \int_K \mathrm{P}_{RT}^k (\mathbf{A} \nabla \phi_k) \cdot \nabla v \, dx = \int \mathrm{P}_{M_{-1}^0(\mathcal{T}_k)}^0 f v \, dx \quad \forall v \in U_k,$$

where $\mathbf{A}(x) \in W^{1,\infty}(\Omega)^{2\times 2}$ is a symmetric matrix, uniformly positive definite, and $\mathrm{P}_{RT}^k : L^2(\Omega) \to RT_{-1}^0(\mathcal{T}_k)$ is a weighted L^2 orthogonal-projector:

$$\int_\Omega \mathbf{A}^{-1}\left(\mathrm{P}_{RT}^k \boldsymbol{\sigma}\right) \cdot \boldsymbol{\tau}\, dx = \int_\Omega \mathbf{A}^{-1}\boldsymbol{\sigma} \cdot \boldsymbol{\tau}\, dx \quad \forall \boldsymbol{\tau} \in RT_{-1}^0(\mathcal{T}_k).$$

In addition to the discontinuity, bubble functions also cause nonnestedness ($B_{k-1} \not\subset B_k$). In transferring bubble functions, different from [8] and [25], we simply discard those bubble functions by defining $\mathrm{I}_k : U_{k-1} \to U_k$ as follows:

$$(4.3) \qquad \mathrm{I}_k u := \mathrm{I}_k(u^l + u^b) = \widetilde{\mathrm{I}}_k u^l + \mathrm{I}_k u^b = \widetilde{\mathrm{I}}_k u^l \quad \forall u = u^l + u^b \in U_{k-1},$$

where $\widetilde{\mathrm{I}}_k$ is defined in (2.5). We now define the bilinear forms for (4.2) and the discrete norms

$$a_k^b(u,v) := \sum_{K\in\mathcal{T}_k} \int_K \mathrm{P}_{RT}^k(\mathbf{A}\nabla u) \cdot \nabla v\, dx \quad \forall u, v \in U_k,$$

$$|u|_{s,k,b} := \sqrt{(\mathrm{A}_k^s u, u)} \quad \forall u \in U_k, \quad 0 \le s \le 2,$$

where $(\mathrm{A}_k u, v) := a_k^b(u,v)$. Let $\Lambda_{k,b}$ be the spectral radius of $\mathrm{A}_k u$.

Definition 4.2 (A multigrid method for (4.2)). The same as Definition 2.1 except (2.7) and (2.8a) are replaced by (4.4) and (4.5a) respectively.

$$(4.4) \qquad (w_l - w_{l-1}, v) = \Lambda_{k,b}^{-1}\left(\left(\mathrm{P}_{M_{-1}^0(\mathcal{T}_k)}^0 f, v\right) - a_k^b(w_{l-1}, v)\right) \qquad \forall v \in U_k,$$

$$(4.5a) \qquad a_{k-1}^b(\epsilon_a, v) = (Pf, \mathrm{I}_k v) - a_k^b(w_m, \mathrm{I}_k v) =: (f_{\mathrm{new}}, v) \qquad \forall v \in U_{k-1},$$

where in (4.5a) P is either $\mathrm{P}_{M_{-1}^0(\mathbf{T}_k)}^0$ if the k-th level is the highest level, or just the identity operator, i.e., depending on if the k-th level problem at the iteration is a finite element problem, or a coarse-level residual problem.

Noting $\widetilde{V}_{k-1} \subset \widetilde{V}_k \subset V_k \subset U_k$, the coarse-level problem could be defined by

$$(4.5b,c) \qquad a_k^b(\epsilon_b, v) = (Pf, v) - a_k^b(w_m, v) =: (f_{\mathrm{new}}, v) \quad \forall v \in \widetilde{V},$$

where $\widetilde{V}$ can be chosen either as $\widetilde{V}_{k-1}$ (4.5b) or $\widetilde{V}_k$ (4.5c). Because of the same reason mentioned after (2.8b,c), one has to program both conforming and nonconforming elements (two sets of data structures) if one uses (4.5b) or (4.5c) instead of (4.5a) in practice.

5. Multigrid convergence for Raviart-Thomas elements

The existence and uniqueness of a solution to (4.2) can be obtained from the formulation of Raviart-Thomas mixed method with Lagrange multipliers (cf [1]), or from the following lemma on the coercivity of $a_k^b(\cdot, \cdot)$.

Lemma 5.1 (Lemma 2.1 in [8]). *There exists a $C > 0$ such that*

$$|u|_{1,k,b} \le C|u|_{1,k} \qquad \forall u \in H_0^1(\Omega) \cup U_k,$$

$$|u|_{1,k,b} \ge C^{-1}|u|_{1,k} \qquad \forall u \in U_k.$$

Lemma 5.2. *It holds that*

$$|\mathrm{I}_k v|_{j,k,b} \le C|v|_{j,k-1,b} \quad \forall v \in U_k, \quad \text{for } j = 0, 1.$$

Proof. By a trianglewise calculation, we have (cf. Lemma 2.3 in [8]),

$$(5.1) \qquad C(|u^l|_{j,k} + |u^b|_{j,k}) \le |u^l + u^b|_{j,k} \quad \forall u = u^l + u^b \in U_k, \quad \text{for } j = 0, 1.$$

In fact $|u^l|_{1,k} + |u^b|_{1,k} = |u^l + u^b|_{1,k}$ (the case that $A(x)$ is the identity matrix). We now apply Lemma 5.1, Lemma 3.1 and the strengthened Cauchy inequalities (5.1) to get

$$|\mathbf{I}_k v|_{j,k,b} \le C|\mathbf{I}_k v|_{j,k} = C|\widetilde{\mathbf{I}}_k v^l|_{j,k} \le C|v^l|_{j,k-1} \le C|v|_{j,k-1} \le C|v|_{j,k-1,b}. \quad \square$$

Similar to (3.3a–b), we define the following operators:

$$(5.2a) \qquad \mathbf{P}_a^b : U_k \to U_{k-1}, \quad a_{k-1}^b(\mathbf{P}_a^b v, w) = a_k^b(v, \mathbf{I}_k w) \quad \forall w \in U_{k-1},$$

$$(5.2b) \qquad \mathbf{P}_b^b : U_k \to \widetilde{V}_{k-1}, \quad a_{k-1}^b(\mathbf{P}_b^b v, w) = a_k^b(v, w) \qquad \forall w \in \widetilde{V}_{k-1},$$

$$(5.2c) \qquad \mathbf{P}_j^b : U_j \to \widetilde{V}_{k-1}, \qquad a_j^b(\mathbf{P}_j^b v, w) = a_j^b(v, w) \qquad \forall w \in \widetilde{V}_{k-1},$$

where in (5.2c) $j = k - 1$ or $j = k$. We note that, unlike the $\mathbf{P}_b$ in (3.3b), $\mathbf{P}_b^b$, $\mathbf{P}_{k-1}^b$ and $\mathbf{P}_k^b$ are different operators caused by $\mathbf{P}_{RT}^k$.

Lemma 5.3. *It holds that, for $j = k - 1$ and $j = k$,*

$$(5.3) \qquad |v - \mathbf{P}_j^b v|_{0,j,b} \le Ch_j |v - \mathbf{P}_j^b v|_{1,j,b} \quad \forall v \in U_j.$$

Proof. The proof is similar to that for Lemma 3.2. But here we have variational crimes (first Strang lemma) caused by $\mathbf{P}_{RT}^k$ in the bilinear forms. Let $\phi_j = v - \mathbf{P}_j^b v$. Note that ϕ_j is $a_j^b(\cdot, \cdot)$-orthogonal in $\widetilde{V}_{k-1}$. We define three solutions using $(v - \mathbf{P}_j^b v)$ as the right-hand side function in the following three equations:

$$a_j^b(\tilde{u}_j, w) = (v - \mathbf{P}_j^b v, w) \qquad \forall w \in \widetilde{V}_{k-1},$$
$$a_j^b(u_j, w) = (v - \mathbf{P}_j^b v, w) \qquad \forall w \in U_j,$$
$$a^b(u, w) = (v - \mathbf{P}_j^b v, w) \qquad \forall w \in H_0^1(\Omega),$$

where $\tilde{u}_j \in \widetilde{V}_{k-1}$, $u_j \in U_j$, $u \in H^2 \cap H_0^1$, and $a^b(u, w) := \int_\Omega \mathbf{A} \nabla u \cdot \nabla v$. We note that both finite element solutions $\tilde{u}_j$ and u_j approximate u in the order of $O(h_j)$ in the norm $\|\mathbf{grad}_j(\cdot)\|_{L^2(\Omega)}$, where $\mathbf{grad}_j$ denotes the elementwise gradient operator (cf. for example, Proposition 2.1 in [8]). We now apply the standard duality analysis.

$$(5.4) \quad \begin{aligned} |v - \mathbf{P}_j^b v|_{0,j,b}^2 &= a_j^b(u_j, v - \mathbf{P}_j^b v) = a_j^b(u_j - \tilde{u}_j, v - \mathbf{P}_j^b v) \\ &\le |u_j - \tilde{u}_j|_{1,j,b} |v - \mathbf{P}_j^b v|_{1,j,b} \le C\|\mathbf{grad}_j(u_j - \tilde{u}_j)\|_{L^2(\Omega)} |v - \mathbf{P}_j^b v|_{1,j,b} \\ &\le C\left(\|\mathbf{grad}_j(u_j - u)\|_{L^2} + |u - \tilde{u}_j|_{H^1}\right) |v - \mathbf{P}_j^b v|_{1,j,b} \\ &\le Ch_j \|u\|_{H^2} |v - \mathbf{P}_j^b v|_{1,j,b} \le Ch_j |v - \mathbf{P}_j^b v|_{0,j,b} |v - \mathbf{P}_j^b v|_{1,j,b}. \quad \square \end{aligned}$$

Theorem 5.4 (The two-level method for R-T equations). *Let $\epsilon = \epsilon_a$ in Definition 4.2. Given any $0 < \gamma < 1$, there exists an integer m independent of the level number k such that*

$$|\phi_k - w_{m+1}|_{1,k,b} \le \gamma |\phi_k - w_0|_{1,k,b}.$$

Proof. By Definition 4.2 and the triangle inequality (see Theorem 3.3 for notations), we have

$$(5.5) \quad |e_{m+1}|_{1,k,b} \le |e_m - \mathbf{P}_k^b e_m|_{1,k,b} + |\mathbf{P}_k^b e_m - \mathbf{P}_b^b e_m|_{1,k,b} + |\mathbf{P}_b^b e_m - \mathbf{I}_k \mathbf{P}_a^b e_m|_{1,k,b}.$$

The first and the third term in (5.5) can be analyzed as in Theorem 3.3. To estimate the second term in (5.5), we define $f_k \in U_k$ and $u \in H_0^1(\Omega) \cap H^2(\Omega)$ by

$$(5.6) \qquad\qquad (f_k, v) := a_k(e_m, v) \qquad \forall v \in U_k$$

$$(5.7) \qquad\qquad -\nabla(\mathbf{A}(x)\nabla u) = f_k \qquad \text{weakly.}$$

We have $\|u\|_{H^2(\Omega)} \leq C|f_k|_{0,k,b} \leq C|e_m|_{2,k,b}$ since Ω is convex (cf. [11]). Comparing (5.6)–(5.7) and (5.2b,c), we realize that both $\mathrm{P}_k^b e_m$ and $\mathrm{P}_b^b e_m$ are finite element approximations (with variational crimes) of u in $\widetilde{V}_{k-1}$. Therefore, the second term in (5.5) can be estimated as in (5.4) to obtain

$$|\mathrm{P}_k^b e_m - \mathrm{P}_b^b e_m|_{1,k,b} \leq |\mathrm{P}_k^b e_m - u|_{1,k,b} + |u - \mathrm{P}_b^b e_m|_{1,k,b} \leq C h_{k-1}|e_m|_{2,k,b}.$$

The rest of the argument in the proof of Theorem 3.3 remains the same. $\qquad\square$

By Theorem 5.4, Theorem 4.1 holds for the general coefficient problem (4.2).

6. MULTIGRID METHODS FOR WILSON'S ELEMENT

In this section, we consider multigrid methods for 2D Wilson elements on general quadrilateral meshes. This falls into the framework for P_1 nonconforming elements. It is straightforward to generalize the analysis here to the case of 3-D Wilson bricks on rectangular brick meshes, or to the case of the modified Wilson elements [14], or to the case of solving linear elasticity problems (cf. [13]).

Let K be a convex quadrilateral with vertices $\{p_i = (x_1^i, x_2^i),\ 1 \leq i \leq 4\}$ and $\hat{K} = [-1, 1] \times [-1, 1]$ the reference square with vertices $\{\hat{p}_i\}$. There is a unique bilinear mapping such that

$$(6.1) \qquad\qquad F_K(\hat{K}) = K, \quad F_K(\hat{p}_i) = p_i, \quad 1 \leq i \leq 4.$$

On the reference square $\hat{K}$ the shape function has the following form:

$$(6.2)$$

$$\hat{u}(\hat{x}_1, \hat{x}_2) = \frac{1}{4}\Big\{(1 + \hat{x}_1)(1 + \hat{x}_2)u^1 + (1 - \hat{x}_1)(1 + \hat{x}_2)u^2 + (1 - \hat{x}_1)(1 - \hat{x}_2)u^3$$

$$+ (1 + \hat{x}_1)(1 - \hat{x}_2)u^4\Big\} + \frac{1}{8}\Big\{(1 - \hat{x}_1^2)u^5 + (1 - \hat{x}_2^2)u^6\Big\}.$$

On K the shape function is given by $u = \hat{u} \circ F_K^{-1}$. The last two terms in (6.2) can be called bubbles, which causes the discontinuity (and nonnestedness). A Wilson element function is determined by its values at the vertices p_i of K and the mean values of the second derivatives:

$$u^5 = \int_K J_K^{-1} \frac{\partial^2 \hat{u}}{\partial \hat{x}_1^2} \circ F_K^{-1}\, dx_1\, dx_2, \quad u^6 = \int_K J_K^{-1} \frac{\partial^2 \hat{u}}{\partial \hat{x}_2^2} \circ F_K^{-1}\, dx_1\, dx_2,$$

where J_K^{-1} is the determinant of the Jacobian for F_K^{-1}.

We consider solving (2.1) by the Wilson element. Let Ω be covered by a convex, quadrilateral grid $\mathcal{T}_1$ initially. The multilevel grids on Ω are defined by linking the midpoints on opposite edges of previous level quadrilaterals. The family of subdivisions $\{\mathcal{T}_k,\ \text{size}(\mathcal{T}_k) = h_k\}$ is regular and satisfies the following condition of Shi [20]:

Condition (B). There is a constant C depending only on the initial subdivision $\mathcal{T}_1$, such that $d_k \leq C h_k^2$ for level k.

Here d_k denotes the maximal distance between the two midpoints of two diagonals of quadrilaterals on level k. $d_k q \equiv 0$ if $\mathcal{T}_1$ contains only parallelograms. The Wilson spaces are denoted by $W_k := \{u \mid u = 0$ at all boundary vertices and $u|_K = \hat{u} \circ F_K^{-1} \quad \forall K \in \mathcal{T}_k\}$, where F_K and $\hat{u}$ are defined in (6.1)–(6.2) respectively. The convergence of Wilson finite elements on quadrilateral grids is proved by Shi [21] under the Condition (B). The Wilson elements are nonconforming since they are not continuous on edges (but at vertices). We define $a_k(u, v)$ as in (2.2), but K stands for a quadrilateral now. As $W_{k-1} \not\subset W_k$, again, we discard bubble functions when doing the intergrid transfer. We introduce auxiliary (conforming) spaces $\widetilde{W}_k$, the bilinear elements on $\mathcal{T}_k$ (cf. [10]). Each function $u \in W_{k-1}$ can be uniquely written as $w + z$ where $w \in \widetilde{W}_{k-1}$ and z is a bubble (last two terms in (6.2)). Noting $\widetilde{W}_{k-1} \subset \widetilde{W}_k \subset W_k$, the intergrid transfer operator $\widetilde{I}_k$ is defined by

$$(6.3) \qquad \widetilde{I}_k \; : \; W_{k-1} \to W_k, \quad \widetilde{I}_k u = \widetilde{I}_k(w + z) = w \in \widetilde{W}_{k-1} \subset W_{k-1} \cap W_k.$$

The other definitions in §2 remain the same. We now prove Lemma 3.1 for Wilson elements.

Lemma 6.1. *It holds that*

$$|\widetilde{I}_k u|_{l,k} \leq C|u|_{l,k-1} \quad \forall u \in W_{k-1}, \quad l = 0, 1.$$

Proof. To prove the lemma, by the equivalence of norms, it suffices to prove the following two estimates on one quadrilateral $K \in \mathcal{T}_{k-1}$:

$$(6.4) \qquad \int_K \left| \nabla \widetilde{I}_k u \right|^2 dx \leq C \int_K |\nabla u|^2 \, dx,$$

$$(6.5) \qquad \int_K \left(\widetilde{I}_k u \right)^2 dx \leq C \int_K u^2 \, dx.$$

Let $u = w + z$ be decomposed as (6.3). Mapping K to the reference element, (6.4) is shown as follows:

$$C \int_K |\nabla w|^2 \, dx \leq \int_{\hat{K}} |\nabla \hat{w}|^2 \, d\hat{x} \leq \int_{\hat{K}} |\nabla \hat{w}|^2 \, d\hat{x} + \int_{\hat{K}} |\nabla \hat{z}|^2 \, d\hat{x}$$

$$= \int_{\hat{K}} |\nabla(\hat{w} + \hat{z})|^2 \, d\hat{x} \leq C \int_K |\nabla(w + z)|^2 \, dx.$$

We note that the Wilson bubbles are $(\nabla \cdot, \nabla \cdot)$-orthogonal to bilinear functions on squares. For the general coefficients problem, (6.4) can be easily proven by a strengthened Cauchy inequality (see (6.6) below), as we shall do next in proving (6.5). Since

$$(6.6) \qquad \left(\int_{\hat{K}} \hat{w} \hat{z} \, d\hat{x} \right)^2 \leq \gamma_0 \int_{\hat{K}} \hat{w}^2 \, d\hat{x} \int_{\hat{K}} \hat{z}^2 \, d\hat{x}$$

for some constant $0 \leq \gamma_0 < 1$, (6.5) follows the following argument:

$$\int_{\hat{K}} (\hat{w} + \hat{z})^2 \, d\hat{x} \geq (1 - \gamma_0) \int_{\hat{K}} \hat{w}^2 + \left(\left(\gamma_0 \int_{\hat{K}} \hat{w}^2 \right)^{1/2} - \left(\int_{\hat{K}} \hat{z}^2 \right)^{1/2} \right)^2$$

$$\geq (1 - \gamma_0) \int_{\hat{K}} \hat{w}^2. \quad \square$$

The remaining analysis in §3 stays the same. In particular we remark that the conforming space $\widetilde{W}_k$ has the same order approximation property as that of W_k, like the situation for the pair $(\widetilde{V}_k, V_k)$. We conclude that the multigrid iteration has a constant convergence rate for Wilson's elements, and is of optimal order (see [3] and Theorem 3.3 for the details).

7. MINI ELEMENTS FOR STATIONARY STOKES EQUATIONS

For brevity, we will restrict our analysis to the mini elements for 2-D Stokes equations only, although the work can be extended in several directions, for example, higher space-dimensional problems and higher-degree "mini elements" (cf. [2]). We consider the following 2D stationary Stokes equations on a convex polygonal domain:

(7.1)
$$\begin{cases} -\Delta\mathbf{u} + \mathbf{grad}\, p = \mathbf{f} & \text{in } \Omega, \\ \operatorname{div}\mathbf{u} = 0 & \text{in } \Omega, \\ \mathbf{u} = \mathbf{0} & \text{on } \partial\Omega. \end{cases}$$

We follow the notations of Verfürth in [24], [25]. The variational problem for (7.1) reads: Find $[\mathbf{u}, p] \in \mathcal{H} := \left(H_0^1(\Omega)\right)^2 \times \left(L^2(\Omega)/\mathbf{R}^1\right)$ such that

(7.2)
$$\mathcal{L}([\mathbf{u}, p]; [\mathbf{v}, q]) = (\mathbf{f}, \mathbf{v}) \quad \forall [\mathbf{v}, q] \in \mathcal{H},$$

where $(\cdot, \cdot)$ is the L^2-inner product, $\mathcal{L}([\mathbf{u}, p]; [\mathbf{v}, q]) := a(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) + b(\mathbf{u}, q)$, $a(\mathbf{u}, \mathbf{v}) := \int_\Omega \mathbf{grad}\mathbf{u} : \mathbf{grad}\mathbf{v}\, dx$, and $b(\mathbf{u}, p) := -\int_\Omega \operatorname{div}\mathbf{u}\, p\, dx$. Let $\{\mathcal{T}_k\}$ be as in §2. Define $\widetilde{V}_k$ and B_k as in §2 and §4, respectively. The multilevel mini (mixed) element spaces are (cf. [2])

$$\mathcal{H}_k := \left(\widetilde{V}_k \oplus B_k\right)^2 \times S_k^0,$$

where S_k^0 is the space of P_1 conforming functions ($\widetilde{V}_k$ functions without the boundary condition) on $\mathcal{T}_k$ with mean-value zero. The mini elements are stable for Stokes problems: there exists a constant C independent of k such that (cf. [2])

(7.3)
$$\sup_{\mathbf{v} \in (\widetilde{V}_k \oplus B_k)^2} \frac{b(\mathbf{v}, p)}{\|\mathbf{v}\|_{H^1(\Omega)^2}} \geq C\|p\|_{L^2} \quad \forall p \in S_k^0.$$

Again the multilevel spaces are nonnested, $\mathcal{H}_{k-1} \not\subset \mathcal{H}_k$, due to bubble functions. As in §4, we discard bubbles when doing the intergrid transfer:

(7.4)
$$\mathcal{I}_k : \mathcal{H}_{k-1} \to \widetilde{V}_{k-1}^2 \times S_{k-1}^0 =: \mathcal{H}_{k-1,l} \subset \mathcal{H}_k,$$
$$\mathcal{I}_k[\mathbf{v}, q]_{k-1} = \mathcal{I}_k[\mathbf{v}_{k-1,l} \oplus \mathbf{v}_{k-1,b}, q_{k-1}] = [\mathbf{v}_{k-1,l}, q_{k-1}].$$

In (7.4) and later, we use notations like

$$[\mathbf{v}, q]_k = [\mathbf{v}_k, q_k] \quad \text{and} \quad \mathcal{I}_k\mathbf{v}_{k-1} = \mathcal{I}_k(\mathbf{v}_{k-1,l} \oplus \mathbf{v}_{k-1,b}) = \mathbf{v}_{k-1,l}.$$

We define a mesh-dependent inner product as in [24]:

(7.5)
$$([\mathbf{u}, p]_k; [\mathbf{v}, q]_k)_{0,k} := (\mathbf{u}_k, \mathbf{v}_k) + h_k^2(p_k, q_k).$$

Let $\mathbf{L}_k : \mathcal{H}_k \to \mathcal{H}_k$ (symmetric but not positive definite) be defined by

$$(\mathbf{L}_k[\mathbf{u}, p]_k; [\mathbf{v}, q]_k)_{0,k} = \mathcal{L}([\mathbf{u}, p]_k; [\mathbf{v}, q]_k) \quad \forall [\mathbf{v}, q]_k \in \mathcal{H}_k.$$

We order its eigenvalues as $0 < |\lambda_1| \leq |\lambda_2| \leq \cdots \leq |\lambda_{N_k}|$. We define a family of mesh-dependent norms:

(7.6) $\||[\mathbf{u},p]_k\||^2_{s,k} := ((\mathbf{L}_k\mathbf{L}_k)^{s/2}[\mathbf{u},p]_k; [\mathbf{u},p]_k)_{0,k} \quad \forall [\mathbf{u},p]_k \in \mathcal{H}_k, \quad -2 \leq s \leq 2.$

We now define a multigrid scheme (cf. [24]) for solving the following discretized version of equation (7.2):

(7.7) $\mathcal{L}([\mathbf{u},p]_k; [\mathbf{v},q]) = (\mathbf{f},\mathbf{v}) \quad \forall [\mathbf{v},q] \in \mathcal{H}_k.$

Algorithm 7.1 (One k-th level MG iteration). One iteration will produce a solution $[\mathbf{u},p]^{m+1}_k \in \mathcal{H}_k$ from a given initial guess $[\mathbf{u},p]^0_k$ as follows.

(1) *Smoothing.* This is done on the normal equation. For $l = 1, 2, \ldots, m$, $[\mathbf{u},p]^l_k$ is defined by

$$([\mathbf{w},r]^l_k; [\mathbf{v},q])_{0,k} = (\mathbf{f},\mathbf{v}) - \mathcal{L}([\mathbf{u},p]^{l-1}_k; [\mathbf{v},q]) \qquad \forall [\mathbf{v},q] \in \mathcal{H}_k,$$

$$([\mathbf{u},p]^l_k - [\mathbf{u},p]^{l-1}_k; [\mathbf{v},q])_{0,k} = |\lambda_{N_k}|^{-2}\mathcal{L}([\mathbf{w},r]^l_k; [\mathbf{v},q]) \quad \forall [\mathbf{v},q] \in \mathcal{H}_k.$$

(2) *Correction.* Let $[\mathbf{u},p]_{k-1} \in \mathcal{H}_{k-1}$ solve

(7.9) $\mathcal{L}([\mathbf{u},p]_{k-1}; [\mathbf{v},q]) = (\mathbf{f},\mathcal{I}_k\mathbf{v}) - \mathcal{L}([\mathbf{u},p]^m_k; \mathcal{I}_k[\mathbf{v},q]) \quad \forall [\mathbf{v},q] \in \mathcal{H}_{k-1}.$

The operator $\mathcal{I}_k$ in (7.9) and (7.10) below is defined in (7.4). If $k = 2$, let $[\mathbf{u},p]^{\sim}_1 = [\mathbf{u},p]_1$. If $k > 2$, let $[\mathbf{u},p]^{\sim}_{k-1}$ be the approximation of $[\mathbf{u},p]_{k-1} \in \mathcal{H}_{k-1}$ obtained by doing μ $(k-1)$-st level multigrid iterations with starting guess zero for some $\mu \geq 2$. Then the new iterative solution after one k-th level iteration is defined by

(7.10) $[\mathbf{u},p]^{m+1}_k := [\mathbf{u},p]^m_k + \mathcal{I}_k[\mathbf{u},p]^{\sim}_{k-1}.$

8. CONVERGENCE ANALYSIS OF MULTIGRID MINI ELEMENTS

Verfürth has set up in [24] a framework for analyzing the multigrid mixed-element method defined in Algorithm 7.1. Our main work is to prove the stability of the intergrid transfer operator which throws away coarse-level bubbles. But this has been proved in Lemma 5.2. Here we only need to change a few minor points of the analysis of Verfürth [24] because of the nonnestedness. Let $[\mathbf{u}^*_k, p^*_k] \in \mathcal{H}_k$ denote the solution of (7.7) and

(8.1) $[\mathbf{e}^l_k, \epsilon^l_k] := [\mathbf{u}^*_k - \mathbf{u}^l_k, p^*_k - p^l_k]$

be the error of the l-th iterate defined in Algorithm 7.1, $0 \leq l \leq m + 1$. The following fine-level smoothing property is shown in [24]:

(8.2) $\||[\mathbf{e}^m_k, \epsilon^m_k]\||_2 \leq Ch_k^{-2}m^{-1/2}\||[\mathbf{e}^0_k, \epsilon^0_k]\||_0.$

The next lemma is used for analyzing the coarse-level correction (7.9)–(7.10).

Lemma 8.1. *Let $[\mathbf{u}_k, p_k] \in \mathcal{H}_k$ be L^2 orthogonal to $\mathcal{H}_{k-1,l}$, i.e.*

$$(\mathbf{u}_k, \mathbf{v}) + h_k^2(p_k, q) = 0 \quad \forall [\mathbf{v},q] \in \mathcal{H}_{k-1,l},$$

where $\mathcal{H}_{k-1,l}$ is defined in (7.4). Then it holds that

$$\||[\mathbf{u}_k, p_k]\||_{-2} \leq Ch_k^2\||[\mathbf{u}_k, p_k]\||_0.$$

Proof. This lemma is almost identical to Lemma 4.2 in [24] except the space $\mathcal{H}_{k-1,l}$ here is replaced by $\mathcal{H}_{k-1}$. But the proof in [24] does not need the inf-sup condition (7.3) on the subspace. It only requires the approximation property. We note that the space $\mathcal{H}_{k-1,l}$ has the same order of approximation as that for $\mathcal{H}_{k-1}$ and that $\mathcal{H}_{k-1,l} \subset \mathcal{H}_k$. Therefore the proof in [24] remains the same.   □

Theorem 8.2 (Convergence of the two-level mini-element multigrid method). *For any $0 < \gamma < 1$ there is a number m independent of k, such that*

$$\||[\mathbf{e}_k^{m+1}, \epsilon_k^{m+1}]\||_0 \leq \gamma \||[\mathbf{e}_k^0, \epsilon_k^0]\||_0,$$

where $[\mathbf{e}_k^{m+1}, \epsilon_k^{m+1}]$ are defined in (8.1).

Proof. We will apply the framework of Verfürth [24]. Let $[\mathbf{w}_{k-1}, r_{k-1}]$ be the L^2 orthogonal projection of the iterative error $[\mathbf{e}_k^m, \epsilon_k^m]$ in $\mathcal{H}_{k-1,l}$. Let $[\mathbf{u}_{k-1}^*, p_{k-1}^*]$ be the exact solution of the residual problem (7.9). Since $\mathcal{H}_{k-1,l} \subset \mathcal{H}_{k-1}$, it follows by a duality argument that

$$\||[\mathbf{e}_k^m, \epsilon_k^m] - [\mathbf{u}_{k-1}^*, p_{k-1}^*]\||_0 \leq Ch_k \inf_{[\mathbf{v},q] \in \mathcal{H}_{k-1}} \{\|\mathbf{e}_k^m - \mathbf{v}\|_1 + \|\epsilon_k^m - q\|_0\}$$

$$\leq Ch_k\{\|\mathbf{e}_k^m - \mathbf{w}_{k-1}\|_1 + \|\epsilon_k^m - r_{k-1}\|_0\}$$

$$\leq C\||[\mathbf{e}_k^m, \epsilon_k^m] - [\mathbf{w}_{k-1}, r_{k-1}]\||_0,$$

where in the last step we used an inverse inequality on P1 elements. Because of the nonnestedness of spaces, the $(m+1)$-st iterative error is actually

$$\||[\mathbf{e}_k^{m+1}, \epsilon_k^{m+1}]\||_0 = \||[\mathbf{e}_k^m, \epsilon_k^m] - \mathcal{I}_k[\mathbf{u}_{k-1}^*, p_{k-1}^*]\||_0$$

$$\leq \||[\mathbf{e}_k^m, \epsilon_k^m] - \mathcal{I}_k[\mathbf{w}_{k-1}, r_{k-1}]\||_0 + \||\mathcal{I}_k[\mathbf{w}_{k-1}, r_{k-1}] - \mathcal{I}_k[\mathbf{u}_{k-1}^*, p_{k-1}^*]\||_0$$

$$= \||[\mathbf{e}_k^m, \epsilon_k^m] - [\mathbf{w}_{k-1}, r_{k-1}]\||_0 + \||\mathcal{I}_k([\mathbf{w}_{k-1}, r_{k-1}] - [\mathbf{u}_{k-1}^*, p_{k-1}^*])\||_0$$

$$\leq \||[\mathbf{e}_k^m, \epsilon_k^m] - [\mathbf{w}_{k-1}, r_{k-1}]\||_0 + C\||[\mathbf{w}_{k-1}, r_{k-1}] - [\mathbf{u}_{k-1}^*, p_{k-1}^*]\||_0,$$

where Lemma 5.2 is applied in the last step. Combining the above two estimates we obtain

$$\||[\mathbf{e}_k^{m+1}, \epsilon_k^{m+1}]\||_0 \leq C\||[\mathbf{e}_k^m, \epsilon_k^m] - [\mathbf{w}_{k-1}, r_{k-1}]\||_0.$$

Applying (8.2) and Lemma 8.1, we can get now (cf. [24]) that

$$\||[\mathbf{e}_k^m, \epsilon_k^m] - [\mathbf{w}_{k-1}, r_{k-1}]\||_0^2 \leq C\||[\mathbf{e}_k^m, \epsilon_k^m]\||_2\||[\mathbf{e}_k^m, \epsilon_k^m] - [\mathbf{w}_{k-1}, r_{k-1}]\||_{-2}$$

$$\leq \frac{C}{\sqrt{m}}\||[\mathbf{e}_k^0, \epsilon_k^0]\||_0\||[\mathbf{e}_k^m, \epsilon_k^m] - [\mathbf{w}_{k-1}, r_{k-1}]\||_0.$$

Therefore, the proof is completed by choosing m large enough as

$$\||[\mathbf{e}_k^{m+1}, \epsilon_k^{m+1}]\||_0 \leq \frac{C}{\sqrt{m}}\||[\mathbf{e}_k^0, \epsilon_k^0]\||_0. \quad \square$$

9. NUMERICAL TESTS

In our numerical tests, we solve (2.1) where

$$(9.1) \qquad f(x,y) = 2y(1-y) + 2x(1-x) \quad \text{and} \quad \Omega = (0,1) \times (0,1).$$

Ω is uniformly triangulated as in Figure 1. The energy norm squares of the conforming finite element (CFE) and the nonconforming finite element (NCFE) solutions are listed in Table 1. We note that for CFE, $\|u - u_k\|_{H^1}^2 = \|u\|_{H^1}^2 - \|u_k\|_{H^1}^2$. The relative nodal errors on level 5 are about 0.5% for CFE and 0.15% for NCFE, but $\dim(\widetilde{V}_k) \sim \dim(V_k)/3$. This is interesting. The solution in V_k is better than that in $\widetilde{V}_k$, but the cost for NCFE is not necessarily more than that of CFE. As a matter of fact, in our code, the NCFE consumes a little less computing time, partially because a node is shared by 6 triangles in CFE but only 2 in NCFE. We remark that the stiffness matrices are not generated in our code, and that the matrix-vector product is evaluated triangle by triangle.
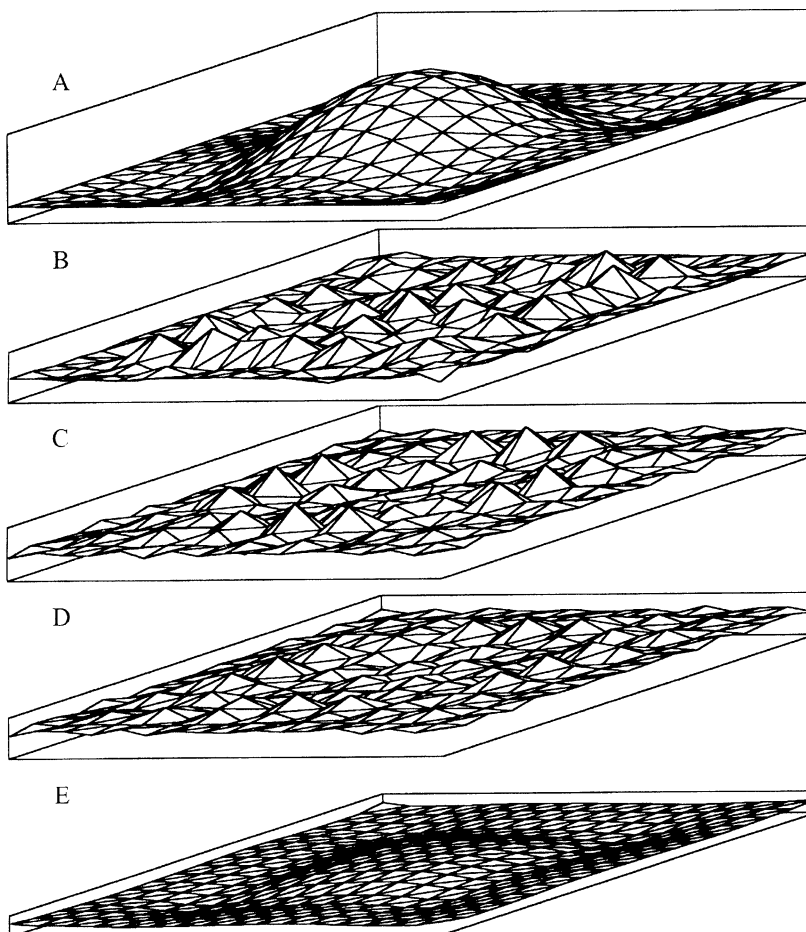
FIGURE 1. Nodal errors before (A) and after doing a coarse-level correction by I_{k-1}^k (B), or by $\widetilde{I}_k$ (C), or by $\widehat{I}_k$ (D), or by P1 CFE nested correction in $\widetilde{V}_k$ (E).

TABLE 1. Comparison of conforming and nonconforming finite elements.

level	number of unknowns		$\|u_k\|_{1,k}^2$	
	CFE	NCFE	CFE	NCFE
4	81	208	0.020 762 645 0	0.022 354 189 9
5	289	800	0.021 848 924 6	0.022 255 785 9
6	1,089	3,136	0.022 128 362 3	0.022 230 649 5
7	4,225	12,416	0.022 198 723 6	0.022 224 331 3
8	16,641	49,408	0.022 216 345 5	0.022 222 749 6
∞	∞	∞	0.022 222 222 2 $\ldots$	0.022 222 222 2 $\ldots$

We then test several "prolongation" (intergrid transfer) operators for NCFE. In Figure 1, the nodal errors before and after doing a coarse-level correction are plotted. The operator of Brenner [7] and Braess-Verfürth [4] (I_{k-1}^k) is used in

Figure 1B, while $\widetilde{I}_k$ (defined in (2.5)) is used in Figure 1C. It seems the difference between the two is small. In Figure 1D, another averaging operator $\widehat{I}_k$ is introduced:

$$\widehat{I}_k : V_{k-1} \to \widetilde{V}_k \subset V_k; \quad \widehat{I}_k v(n) = \begin{cases} \widetilde{I}_k v(n) & \text{if } n \text{ is a vertex in } \mathcal{T}_{k-1}, \\ v(n) & \text{if } n \text{ is a midpoint in } \mathcal{T}_{k-1}. \end{cases}$$

We remark that the operator $\widehat{I}_k$ is also used by Vassilevski and Wang in [23] recently. The graph shows that $\widehat{I}_k$ is better than both I_{k-1}^k and $\widetilde{I}_k$. If we use P1 CFE as the coarse-level subspaces (a nested multigrid method now), then both higher- and lower-frequency components of the error are reduced as shown in Figure 1E.

The operation counts for various "prolongation" operators are listed below (note that $\dim(V_k) \approx 12 \dim(\widetilde{V}_{k-1})$):

Operators	$\widetilde{I}_k$	I_{k-1}^k	$\widehat{I}_k$	$\overline{I}_k$
Operations (times $\dim V_k$)	6	7	4	3

Here we introduce a simpler prolongation operator $\overline{I}_k$: Let c be the midpoint of an edge $\overline{ab}$ in $\mathcal{T}_{k-1}$. $\overline{I}_k$ simply takes the value $v(c)$ as the values of $\overline{I}_k v$ at the midpoints of $\overline{ac}$ and $\overline{cb}$. It seems that this operator transfers NCFE roughly. But the convergence rate of the multigrid method defined by this operator is not worse than that with $\widetilde{I}_k$ (see Table 2).

In Table 2, we list the number of V-cycles needed for a multigrid method with $p = 1$ and $m = 8$ in Definition 2.1. By Figure 1D–E, the multigrid method seems to have a better convergence rate using coarse-level correction (2.8c) than using (2.8a) with $\widetilde{I}_k$ replaced by $\widehat{I}_k$. But it turns out that the two rates are about the same by the data in Table 2. In the last column of Table 2, we list the data for CFE. By Table 2, one may conclude that the NCFE multigrid with various perturbations in (2.8a) is almost as good as the CFE multigrid.

TABLE 2. Numbers of V-cycle iterations.

level	interpolator					CFE
	$\widetilde{I}_k$	I_{k-1}^k	$\widehat{I}_k$	I (2.8c)	$\overline{I}_k$	
4	8	6	6	6	7	3
5	9	7	7	7	11	6
6	9	8	7	7	10	7
7	9	8	7	7	9	7
8	9	8	7	7	9	6

In Table 3, the spectral radii are presented for solving (4.2). The smoothing and correction operators are defined respectively by $S := (I - \Lambda_k^{-1} A_k)$ and $C := (I - I_k A_{k-1}^{-1} I_k^\top A_k)$ (cf. (2.7–2.8)). Here, I_k is either defined by (4.3) or the I_{k-1}^k defined in [7] and [4]. In matrix form, the A_k for (4.2) is a 2-block diagonal matrix. The matrix in the first block is the A_k for P1 NCFE (2.2). The second block is diagonal, generated by bubble functions. We scale bubbles such that each has a nodal value $\sqrt{60}/3$ at its barycenter to get basis functions. Therefore the spectrum of the multigrid operator associated with this lower block is very close to zero, no matter that the bubbles are thrown away (our method) or projected to fine-level bubbles ([8]).

TABLE 3. Spectral radii of 2-level multigrid operators for (4.2).

grid (# unknowns)	interpolator I_{k-1}^k			interpolator $\widetilde{I}_k$		
	$\rho(S^4C)$	$\rho(SC)$	$\rho(C)$	$\rho(S^4C)$	$\rho(SC)$	$\rho(C)$
4×4 (40)	0.5422	0.8516	1.0000	0.7145	0.9194	1.0000
8×8 (176)	0.5708	1.0745	1.4144	0.7681	0.9361	1.0000
12×12 (408)	0.6295	1.1957	1.5865	0.7774	0.9390	1.0000
16×16 (736)	0.6548	1.2445	1.6554	0.7804	0.9399	1.0000
20×20(1160)	0.6672	1.2684	1.6888	0.7818	0.9403	1.0000

As we know, the orthogonal-projection would be lost in (2.8a) when the multilevel spaces are nonnested. The number of fine-level smoothings has to be sufficiently large in this case. This is verified in column 4 of Table 3. The V-cycle using Brenner's method with one step of smoothing diverges according to the data in column 3 of Table 3. Surprisingly, one smoothing is enough to ensure the convergence of nonnested multigrids when our new intergrid transfer operator $\widetilde{I}_k$ is in use. Columns 2 and 5 of Table 3 show that Brenner's method is better (consistent with Table 2). But Columns 3 and 6 of Table 3 imply that our method is better. It tells us that $\widetilde{I}_k$ is more stable than I_{k-1}^k. In fact, it can be seen in the analysis of Oswald [17] that $|v|_1 < |I_{k-1}^k v|_1 \leq C|v|_1$.

We now present the convergence rate of a multigrid method for solving (9.1) via Wilson's elements on uniform square grids. As we pointed out in the proof for Lemma 6.1, Wilson's "bubbles" are $(\nabla \cdot, \nabla \cdot)$-orthogonal to bilinear functions. The stiffness matrix has 2 diagonal blocks where the block for bubbles is diagonal. The rate of convergence of multigrid using Wilson elements is the same as that for bilinear elements. We list here the spectral radii of 2-level multigrid methods with one step of presmoothing:

grid	6×6	10×10	14×14	18×18	22×22	26×26	30×30
unknowns	97	281	561	937	1409	1977	2641
spectral radius	0.4302	0.4751	0.4874	0.4924	0.4949	0.4963	0.4973

Next we consider the multigrid method in §7 for mini elements. We replace the cubic bubbles in the standard mini element by macro linear-bubbles. This is to avoid using high-order numerical quadratures. Let O be the barycenter of a triangle $T = \triangle ABC \in \mathcal{T}_k$. The spaces of macro linear-bubble functions are defined by

$$\overline{B}_k := \text{span} \left\{ \phi \mid \phi \text{ is linear on } \triangle ABO, \triangle ACO \text{ and } \triangle BCO, \right.$$
$$\left. \phi \text{ is continuous on } T \text{ and } \phi|_{\partial T} = 0 \quad \forall T \in \mathcal{T}_k \right\}.$$

Following [2], we can show that the stability condition (7.3) holds for the modified mini element:

$$\mathcal{H}_k := \left(\widetilde{V}_k \oplus \overline{B}_k \right)^2 \times S_k^0.$$

The multigrid analysis remains valid. We construct all matrices in the multigrid method (uniform grids on the unit square domain) and use MATLAB to find the spectral radii of iterative operators.

Columns 3 and 4 of Table 4 consist of the data of the multigrid method when the bubbles are discarded or averaged by the Verfürth method [25], respectively.

TABLE 4. Spectral radii of 2–level multigrid mini-elements.

grid	smoothings	MG iterations	
(# unknowns)	$\rho(S^8)$	$\rho(C_Z S^8)$	$\rho(C_V S^8)$
$4 \times 4(106)$	0.99981616	0.8867	0.8868
$6 \times 6(242)$	0.99995724	0.9033	0.9169
$8 \times 8(434)$	0.99999154	0.9211	0.9201
$10 \times 10(682)$	0.99999688	0.9219	0.9214
$12 \times 12(986)$	0.99999870	0.9219	0.9215

TABLE 5. Maximal errors at vertices and barycenters.

grid	$L^\infty(\Omega)$	l^∞-error at barycenters(c), nodes(n)		
(# unknowns)	$|u_{k,b}|$	$|u - u_k|(c)$	$|u - u_{k,l}|(c)$	$|u - u_{k,l}|(n)$
$2 \times 2(\ 26)$	0.1070	0.8865	0.9144	0.0000
$4 \times 4(\ 106)$	0.1027	0.5607	0.6111	0.2990
$8 \times 8(\ 434)$	0.0308	0.1535	0.1566	0.0564
$16 \times 16(1762)$	0.0590	0.0533	0.0518	0.0373
$32 \times 32(7106)$	0.0028	0.0101	0.0098	0.0039

Both methods have constant rates of convergence. We remark that the original algorithm of [25] contains a minor error. It is necessary to include the intergrid transfer operator into the coarse-level residual equation (7.9) (cf. [4] [6], [7], [26] and [27]). The idea of discarding bubble functions is based on the fact that bubble functions only provide the stability but not the approximation property. We next numerically check it. The test problem is the Stokes equations (7.1) on the unit square with $\mathbf{f} = -\Delta \mathbf{curl} g$ where $g = 100x^2(1-x)^2 y^2(1-y)^2$. The exact solution is $\mathbf{curl} g$. In column 2 of Table 5, the maximal norms of bubble functions are listed. The maximal nodal errors of the mini–element solutions are shown in columns 3–5. We can see that the convergence of the mini element is of the same order as that of the linear part of the mini element, both $O(h_k^2)$ in l^∞-norm.

To conclude, we would remark again that the way to treat the discontinuity and bubbles in the multigrid method is not unique. Each method may have its advantages and disadvantages, depending on the specific implementation.

DEDICATION

We dedicate this paper to Professor James H. Bramble. An early version of this paper [28] was dedicated to Professor Bramble at the workshop on Mathematics of Computation in Partial Differential Equations held at Cornell University, January 25–27, 1991, in honor of his 60th birthday.

REFERENCES

1. D. N. Arnold and F. Brezzi, *Mixed and nonconforming finite element methods: implementation, postprocessing and error estimates,* RAIRO Modél. Math. Anal. Numér. **19** (1985), 7 – 32. MR **87g:**65126

2. D. N. Arnold, F. Brezzi and M. Fortin, *A stable finite element for the Stokes equations,* Calcolo **21** (1984), pp. 337 – 344. MR **86m:**65136

3. R. Bank and T. Dupont, *An optimal order process for solving finite element equations,* Math. Comp. **36** (1981), 35 – 51. MR **82b:**65113

4. D. Braess and R. Verfürth, *Multigrid methods for nonconforming finite element methods,* SIAM J. Numer. Anal. **27** (1990), 979 – 986. MR **91j:**65164

5. J. H. Bramble, Multigrid methods, Longman, Essex, 1993. MR **95b**:65002

6. 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. MR **91h**:65159

7. S. C. Brenner, *An optimal-order multigrid method for P1 nconforming finite elements,* Math. Comp. **52** (1989), 1 – 15. MR **89f**:65119

8. _____, *A multigrid algorithm for the lowest order Raviart-Thomas mixed triangular finite element method,* SIAM J. Numer. Anal. **29** (1992). MR **93j**:65175

9. _____, *An optimal order nonconforming multigrid method for the biharmonic equation,* SIAM J. Numer. Anal. **26** (1989), 1124 – 1138. MR **90i**:65189

10. P. G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam, New York, Oxford, 1978. MR **58**:25001

11. P. Grisvard, "Elliptic Problems in Nonsmooth Domains", Pitman Pub. Inc., 1985. MR **86m**:35044

12. W. Hackbusch, *On the convergence of a multi-grid iteration applied to finite element equations*, Report 77-8, Universität zu Köln, July 1977.

13. P. Lesaint, *On the convergence of Wilson's nonconforming element for solving the elastic problem*, Comput. Methods Appl. Mech. Engrg. **7** (1976), pp. 1 – 16. MR **56**:13717

14. P. Lesaint and M. Zlámal, *Convergence of the nonconforming Wilson element for arbitrary quadrilateral meshes*, Numer. Math. **36** (1980), pp. 33 – 52. MR **81m**:65171

15. S. F. McCormick (ed), "Multigrid Methods", Frontiers in Applied Mathematics, SIAM, Philadelphia, 1987. MR **89m**:65004

16. R. A. Nicolaides, *On the l^2 convergence of an algorithm for solving finite element equations,* Math. Comp. **31** (1977), 892 – 906. MR **58**:8239

17. P. Oswald, *On a hierarchical basis multilevel method with nonconforming P1 elements*, Numer. Math. **62** (1992), 189 – 212. MR **93b**:65059

18. P. Peisker and D. Braess, *A conjugate gradient method and a multigrid algorithm for Morley's finite element approximation of the biharmonic equation,* Numer. Math. **50** (1987), 567 – 586. MR **88e**:65147

19. P.-A. Raviart and J. M. Thomas, *A mixed finite element method for second order elliptic problems,* in "Mathematical Aspects of the Finite Element Method", Lecture Notes in Mathematics, 606, Springer-Verlag, Berlin, 1977. MR **58**:3547

20. R. Scott and S. Zhang, *Higher dimensional nonnested multigrid methods*, Math. Comp. **58** (1992), pp. 457–466. MR **92g**:65133

21. Z. Shi, *A convergence condition for the quadrilateral Wilson element,* Numer. Math. **44** (1984), pp. 349 – 361. MR **86d**:65151

22. G. Strang and G. Fix, "An Analysis of the Finite Element Method", Prentice-Hall, Englewood Cliffs, N. J., 1973. MR **56**:1747

23. 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. MR **96a**:65173

24. R. Verfürth, *A multilevel algorithm for mixed problems,* SIAM J. Numer. Anal. **21** (1984), 264 – 271. MR **85f**:65112

25. R. Verfürth, *Multilevel algorithms for mixed problems. II. Treatment of the mini element,* SIAM J. Numer. Anal. **25** (1988), 285 – 293. MR **89a**:65175

26. J. Xu, *Theory of multilevel methods,* Report AM–48 (1989), Department of Mathematics, Pennsylvania State University.

27. S. Zhang, *Optimal order nonnested multigrid methods for solving finite element equations II: on non-quasiuniform meshes*, Math. Comp. **55** (1990), 439 – 450. MR **91g**:65269

28. S. Zhang and Z. Zhang, *Treatment of discontinuity and bubble functions in multigrid methods I. P1 nonconforming elements and Raviart-Thomas elements, II. mini elements and Wilson elements*, Technical Report 153 and 154, March, 1991, Center for Applied Mathematics, Purdue University; Also Technical Report 1991-5 and 1991-7, Center for the Mathematics of Waves, University of Delaware.

DEPARTMENT OF MATHEMATICAL SCIENCES, UNIVERSITY OF DELAWARE, NEWARK, DELAWARE 19716
E-mail address: szhang@math.udel.edu

DEPARTMENT OF MATHEMATICS, TEXAS TECH UNIVERSITY, LUBBOCK, TEXAS 79409
E-mail address: zhang@ttmath.ttu.edu