

## HIGHER-DIMENSIONAL NONNESTED MULTIGRID METHODS

L. RIDGWAY SCOTT AND SHANGYOU ZHANG

**ABSTRACT.** Nonnested multigrid methods are shown to be optimal-order solvers for systems of finite element equations arising from elliptic boundary problems in any space dimension. Results are derived for Lagrange-type elements of arbitrary degree.

### 1. INTRODUCTION

Large linear systems of elliptic finite element equations can be solved in the optimal computational order, i.e., the costs are linear with respect to the number of unknowns in the systems, by multigrid methods (cf. [8, 9] and references therein). Various versions of multigrid methods have been developed and studied, for example, [2, 13, 4, 3, and 15]. In this paper, we will prove optimal computational order for *nonnested* multigrid methods for solving finite element equations arising from higher-dimensional elliptic problems. Our primary purpose is to provide a theoretical foundation for the general 3-dimensional nonnested multigrid code NMGTM (cf. [11]).

The principle of multigrid methods is to accelerate standard iterative techniques by doing corrections on coarser grids. The low-frequency iterative errors are corrected with less cost by solving a projected residual problem on a coarser mesh. To solve the coarse-level residual problem, we repeat the process recursively, iterating and then making corrections on an even coarser mesh. The multigrid method thus uses a sequence of grids which are usually obtained successively by subdivision. In one (resp. two) dimensions, each interval (resp. triangle) can be subdivided into 2 (resp. 4) congruent subintervals (resp. subtriangles), generating a family of nested meshes. Unfortunately, refining 3-dimensional tetrahedral meshes is much more complicated [16]. For example, when refining a regular tetrahedron by linking the midpoints of its edges, there are four subtetrahedra which have one edge of length  $\sqrt{2}$  times the length of the other edges, no matter how we link the midpoints. Further refinements of these subtetrahedra could potentially degenerate [16]. If we would like to refine the meshes nonuniformly to resolve perceived features in the solution, even greater difficulties would be encountered due to face and edge couplings between tetrahedra. More suitable meshes can potentially be obtained if we are not constrained to have nested meshes [11]. Thus, nonnested multiple grids would

---

Received June 21, 1990; revised October 1, 1990 and February 7, 1991.

1991 *Mathematics Subject Classification.* Primary 65N30, 65F10.

This work was supported in part by the National Science Foundation through award number DMS-8903548.

be a natural approach to avoid the difficulties caused by the nested-refinement constraint.

Although the projection property [2] of coarse-level corrections does not usually exist in nonnested multigrid methods, nonnested coarser meshes can capture low-frequency errors and the optimal order of the multigrid method can be kept. In this paper, we will show that nonnested multigrid methods have a constant rate of convergence and are of optimal order in solving finite element equations in any space dimension and for any degree of piecewise polynomial. Nonnested multigrid methods were proposed and studied in [15] and [3]. Two-dimensional nonnested multigrids were studied in [17, 18, and 19] for the cases where the meshes are quasi-uniform, nondegenerate (but non-quasi-uniform) and degenerate, respectively. In [11], it is demonstrated that the complexity of implementation of nonnested multigrid methods is essentially the same as for the nested case.

The organization of this paper is as follows. In §2, we give some necessary definitions and assumptions. In §3, some preliminary results are presented which concern interpolation operators, coarse-level projection operators and coarse-level correction operators. In §4, we apply the theory of [3] together with the results of §3 to prove constant-rate convergence of the  $W$ -cycle and the variable  $V$ -cycle nonnested multigrid methods (with sufficiently many smoothings). The theory of [3] is also used to show that more complex iterative schemes, such as conjugate gradients, can be used to improve convergence properties of nonnested multilevel methods. In addition, we observe that these convergence estimates can yield optimal-order work estimates.

## 2. DEFINITIONS AND ASSUMPTIONS

In this section, we will define the multilevel finite element problems and the nonnested multigrid schemes, and give our assumptions on elliptic regularity and on quasi-uniformity of the meshes. For simplicity, we consider a model problem,

$$(2.1) \quad \begin{aligned} -\nabla \cdot (a \nabla u) &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned}$$

where  $\Omega$  is a bounded, Lipschitz, polyhedral domain in  $\mathbb{R}^d$ ,  $a \in W^{1,\infty}(\Omega)$ ,  $a \geq C_0 > 0$  a.e. on  $\Omega$  and  $f \in L^2(\Omega)$ . We use  $W^{m,p}(\Omega)$  and  $H^\alpha(\Omega) = W^{\alpha,2}(\Omega)$  to denote usual Sobolev spaces (cf. [6, 7]). We assume that the variational solution of (2.1) is in  $H^{1+\alpha}(\Omega)$  for some  $\alpha > 0$  and that

$$(2.2) \quad \|u\|_{H^{1+\alpha}} \leq C \|f\|_{H^{-1+\alpha}}$$

for some  $C > 0$  independent of  $f$  (cf. [6]).

Finite element discretizations for (2.1) read as follows. Find  $u_k \in V_k$  such that

$$(2.3) \quad a(u_k, v) = F(v) \quad \forall v \in V_k, \quad k = 1, 2, \dots,$$

where  $a(u, v) \stackrel{\text{def}}{=} \int_\Omega a \nabla u \cdot \nabla v \, dx$ ,  $F(v) \stackrel{\text{def}}{=} \int_\Omega f v \, dx$ , and

$$V_k \stackrel{\text{def}}{=} \{v \in C(\Omega) \mid v|_K \in \mathcal{P}_l^d \quad \forall K \in \mathcal{T}_k, \quad v|_{\partial\Omega} = 0\} \subset H_0^1(\Omega).$$

Here,  $\mathcal{P}_l^d$  denotes the space consisting of polynomials of degree  $l$  or less in  $d$  variables and  $\{\mathcal{T}_k, k = 1, 2, \dots\}$  is a family of quasi-uniform triangulations

on  $\Omega$  in the usual sense:  $\bar{\Omega} = \bigcup_{K \in \mathcal{T}_k} K$  and there are constants  $C_{QU} > 0$  and  $C_{ND} < \infty$  such that

$$(2.4) \quad C_{QU} h_k \leq h_K \leq C_{ND} \rho_K \quad \forall K \in \mathcal{T}_k, \quad k = 1, 2, \dots,$$

where  $h_K$  is the diameter of  $K$ ,  $h_k$  is the maximum of all  $h_K$  for  $K \in \mathcal{T}_k$  and  $\rho_K$  is the diameter of the maximal ball inscribed in  $K$ . In this paper,  $K$  is understood to be a closed,  $d$ -dimensional simplex. Without loss of generality, we assume  $l \geq \alpha$ . By (2.4), all Lagrange finite elements on each level,  $\{(K, P_K, \Sigma_K), K \in \mathcal{T}_k\}$ , are affine-equivalent to a single reference finite element  $(\hat{K}, \mathcal{P}_l^d, \hat{\Sigma}_K)$  (cf. [5] for notation). We do not assume  $\mathcal{T}_{k-1} \subset \mathcal{T}_k$ , the nestedness of meshes. The fact that *refinement* of the multiple meshes occurs is implied by the following three **assumptions**:

(A1) The numbers of simplices in the  $\mathcal{T}_k$ 's grow geometrically with a factor,  $\mathcal{F}$ , greater than 2.

(A2) Each simplex of  $\mathcal{T}_k$  intersects at most  $C_0$  simplices of  $\mathcal{T}_{k \pm 1}$  for some  $C_0$  independent of  $k$ .

(A3)  $C_1 h_k \leq h_{k-1} \leq C_2 h_k$  for some  $C_1 > 0$  and  $C_2 < \infty$  independent of  $k$ .

In standard, nested multigrid refinements, the factor in the first condition would be  $2^d$ . As a consequence of the quasi-uniformity condition (2.4), the three conditions (A1)–(A3) are interrelated; we have stated them separately for the purposes of clarity.

The bilinear form  $a(\cdot, \cdot)$  defines an inner product and a norm  $\|u\|^2 \stackrel{\text{def}}{=} a(u, u)$  in  $H_0^1(\Omega)$ . The  $\|\cdot\|$ -norm is equivalent to the  $H^1$ -norm. For computational reasons (cf. Bank and Dupont [2]), it is necessary to introduce an additional inner product that is equivalent to the  $L^2$  inner product. For example, let

$$b_k(v, w) := \sum_n v_n w_n (\phi_n, \phi_n),$$

where  $v_n$  denotes the value of  $v$  at the node  $n$ ,  $(\cdot, \cdot)$  is the usual  $L^2$  inner product on  $\Omega$  and  $\{\phi_n\}$  is the usual Lagrange basis for  $V_k$ .

Let  $A_k: V_k \rightarrow V_k$  be the symmetric positive definite operator defined by  $a(u, v) = b_k(A_k u, v) \quad \forall u, v \in V_k$ . We define a family of mesh-dependent norms:

$$\|u\|_{s,k}^2 = b_k(A_k^s u, u) \quad \forall u \in V_k, \quad 0 \leq s \leq 2, \quad k = 1, 2, \dots$$

We note that  $\|u\|_{0,k} = \|u\|_{L^2}$  and  $\|u\|_{1,k} = \|u\|$  for all  $u \in V_k$ . Also note that (2.3) may be written in the form

$$(2.3') \quad A_k u_k = f_k, \quad k = 1, 2, \dots,$$

where  $f_k \in V_k$  is determined by  $b_k(f_k, v) = F(v)$  for all  $v \in V_k$ .

We now define a symmetric, nonnested multigrid scheme (cf. [2]). The full multigrid method has two iterative processes. The overall process involves solving problems (2.3) sequentially for  $k = 1, 2, \dots$  to get  $\tilde{u}_k \sim u_k$ . To solve (2.3) on the  $k$ th level, we interpolate  $\tilde{u}_{k-1}$  as an initial guess and then apply one or more times the second, recursive process given in the following definition.

**Definition 2.1** (The  $k$ th level symmetric nonnested multigrid scheme).

- (1) For  $k = 1$ , (2.3) or (2.6) below is solved exactly.
- (2) For  $k > 1$ ,  $w_{2m_k+1}$  will be generated from the initial guess,  $w_0$ , as follows:

(2a)  $m_k$  presmoothings are performed to generate  $w_{m_k}$  :

$$(2.5) \quad b_k(w_l - w_{l-1}, v) = \lambda_k^{-1}(F(v) - a(w_{l-1}, v)) \quad \forall v \in V_k,$$

$$l = 1, 2, \dots, m_k,$$

where  $\rho(A_k)/\lambda_k \leq \omega$  for some fixed  $\omega$  satisfying  $0 < \omega < 2$  ( $\rho(A_k)$  is the spectral radius of  $A_k$ ) and  $F(\cdot)$  is either the  $F(\cdot)$  in (2.3) or the  $\tilde{F}(\cdot)$  in (2.6) below.

(2b)  $w_{m_k}$  is corrected by  $I_k q$  to generate  $w_{m_k+1}$  : Let  $\bar{q}$  solve the coarse-level residual problem,

$$(2.6) \quad a(\bar{q}, v) = F(I_k v) - a(w_{m_k}, I_k v) \stackrel{\text{def}}{=} \tilde{F}(v) \quad \forall v \in V_{k-1}.$$

Let  $q \in V_{k-1}$  be the approximation of  $\bar{q}$  obtained by applying  $p$  iterations of the  $(k - 1)$ st level scheme to (2.6) starting with initial guess zero. Then  $w_{m_k+1} = w_{m_k} + I_k q$ . Here,  $I_k: V_{k-1} \rightarrow V_k$  is the usual Lagrange interpolation operator.

(2c)  $m_k$  postsmoothings of the form (2.5) are performed to generate  $w_{2m_k+1}$  from  $w_{m_k+1}$ .

We let  $S_k$  denote the mapping given by (2.5), i.e.,  $w_l = S_k w_{l-1}$ . Thus  $w_m = S_k^m w_0$  and  $w_{2m+1} = S_k^m w_{m+1}$ . It is obvious that  $S_k$  satisfies condition (A.4) of [3]. If the meshes are nested,  $\mathcal{T}_{k-1} \subset \mathcal{T}_k$ , then the  $I_k$  above is just the identity operator and the scheme is a standard multigrid method.

### 3. PRELIMINARY RESULTS

In this section, we will present some lemmas concerning the stability and approximability of the operators  $I_k$  and two other operators  $P_k$  and  $Q_k$  defined below. We first extend the interpolation operator  $I_k: V_{k-1} \rightarrow V_k$  to a local-averaging operator, denoted also by  $I_k$ , as follows. For each Lagrange node  $n$  of  $\mathcal{T}_k$ , the corresponding nodal value is defined by averaging over a  $(d - 1)$ -dimensional set  $K_n$  fulfilling the conditions of the following lemma.

**Lemma 3.0.** *Let the assumptions (2.4) and (A1)–(A3) hold. Then for each Lagrange node  $n$  of  $\mathcal{T}_k$ , there is a set  $K_n$ , which is either a  $(d - 1)$ -dimensional simplex or the intersection of two  $d$ -dimensional simplices, fulfilling the following conditions:*

$$\begin{aligned} n &\in K_n, \\ K_n &\subset \partial\Omega \quad \text{if } n \in \partial\Omega, \\ K_n &\subset K \cap K' \quad \text{for some } K' \in \mathcal{T}_{k-1} \text{ and some } K \in \mathcal{T}_k, \\ \rho_{K_n} &\geq Ch_k \quad \text{for some } C > 0. \end{aligned}$$

*Proof of Lemma 3.0.* First assume that  $n \notin \partial\Omega$ . Let

$$\sigma_n = \bigcup_{n \in K \in \mathcal{T}_k} K.$$

Since  $n$  is a Lagrange node of  $\mathcal{T}_k$ , there is a ball,  $B_n$ , of radius  $Ch_k$  centered at  $n$  contained in  $\sigma_n$ . If  $n \in K' \in \mathcal{T}_{k-1}$ , then  $B_n \cap K'$  must itself contain a ball,  $B'$ , of radius  $C'h_k$ . Since  $B' = \bigcup_{n \in K \in \mathcal{T}_k} K \cap B'$  and there can be at most a fixed number of simplices satisfying  $n \in K \in \mathcal{T}_k$ , the one such that  $K \cap B'$  has the largest volume yields an appropriate set,  $K_n = K \cap K'$ .

If  $n \in \partial\Omega$ , define

$$\sigma_n = \bigcup_{n \in K \in \mathcal{T}_k} K \cap \partial\Omega.$$

Now there is a ball,  $B_n$ , of radius  $Ch_k$  centered at  $n$  such that  $B_n \cap \partial\Omega \subset \sigma_n$ . Again, if  $n \in K' \in \mathcal{T}_{k-1}$ , then  $B_n \cap K' \cap \partial\Omega \cap K$  must contain an appropriate simplex,  $K_n$ , for some  $n \in K \in \mathcal{T}_k$ .  $\square$

The averaging is done using a suitable [10] polynomial,  $\psi_n$ , on  $K_n$ . In particular,  $\psi_n \in \mathcal{P}_l$  is chosen such that

$$\int_{K_n} \psi_n(y)u(y) dy = v(n) \quad \forall v \in \mathcal{P}_l.$$

Let  $u$  be a function in  $W_p^l(\Omega)$  for some  $l > 1/p$ ;  $I_k u \in V_k$  is defined by

$$I_k u = \sum_n \int_{K_n} \psi_n(y)u(y) dy \phi_n,$$

where  $\{\phi_n\}$  denotes the usual Lagrange basis for  $V_k$ . Then,  $I_k$  preserves the homogeneous Dirichlet boundary condition and is identical to the Lagrange interpolation operator when restricted to  $V_{k-1} \cup V_k$ . We refer to [10] for more details regarding this averaging operator. It is shown in [10] that  $I_k$  has the simultaneous approximation property:

$$(3.1) \quad \|u - I_k u\|_{L^2} + h_k \|u - I_k u\|_{H^1} \leq Ch_k^\beta \|u\|_{H^{1+\beta}} \quad \forall u \in H^{1+\beta}(\Omega), \quad 0 \leq \beta \leq 1,$$

where, and in the sequel,  $C$  denotes a generic constant independent of  $k$ . We note that the way of selecting averaging simplices is not exactly the same as that used in [10], but the results and the analysis there remain the same here.

We define two linear operators,

$$(3.2) \quad \begin{aligned} P_{k-1}: H^1(\Omega) &\rightarrow V_{k-1}, & a(P_{k-1}u, v) &= a(u, v) \quad \forall v \in V_{k-1}, \\ Q_{k-1}: H^1(\Omega) &\rightarrow V_{k-1}, & a(Q_{k-1}u, v) &= a(u, I_k v) \quad \forall v \in V_{k-1}. \end{aligned}$$

$P_{k-1}$  is an  $a(\cdot, \cdot)$ -projection operator.  $Q_{k-1}$  is introduced for studying the coarse-level correction (2.6).

As a consequence of the quasi-uniformity condition (2.4), the following inverse inequality holds (cf. Theorem 3.2.6 in [5]):

$$(3.3) \quad \|u\| \leq Ch_k^{-1} \|u\|_{0,k},$$

where the constant depends on the polynomial degree  $l$ . By (3.1) and (3.3), we can show as in [2] that

$$(3.4) \quad \frac{1}{C} \|u\|_{H^\beta} \leq \|u\|_{\beta,k} \leq C \|u\|_{H^\beta} \quad \forall u \in V_k, \quad 0 \leq \beta \leq 1.$$

We note that the proof in [2] makes use of the fact that the Sobolev spaces can be viewed as real interpolation spaces,  $H^\beta = [L^2, H^1]_{\beta,2}$  (cf. [1]). This

fact may be deduced by generalizing Corollary 7.46 of [1] to Lipschitz domains using the Stein extension theorem [12] and an extension theorem with respect to the Sobolev norm used in [6]. The latter may be found in [14, Theorem 5.6].

By (3.1) and the definition of  $\|\cdot\|$ , we have

$$(3.5) \quad \|\mathbf{I}_k v\| \leq C_I \|v\| \quad \forall v \in H^1(\Omega)$$

for some  $C_I < \infty$ . Therefore,

$$(3.6) \quad \|\mathbf{Q}_{k-1} v\| \leq C_I \|v\| \quad \forall v \in H^1(\Omega),$$

since  $\|\mathbf{Q}_{k-1} v\|^2 = a(v, \mathbf{I}_k(\mathbf{Q}_{k-1} v)) \leq \|v\| \|\mathbf{I}_k(\mathbf{Q}_{k-1} v)\| \leq C_I \|v\| \|\mathbf{Q}_{k-1} v\|$ . It also follows by interpolating (3.1) that

$$(3.7) \quad \|v - \mathbf{I}_k v\|_{H^\beta} \leq C h_k^{1-\beta} \|v\| \quad \forall v \in V_{k-1}, \quad 0 \leq \beta \leq 1.$$

The following results use standard duality techniques. We give complete details since they are simplified by using the interpolant  $\mathbf{I}_k$ . We assume  $0 < \alpha \leq 1$ .

**Lemma 3.1.** *Let the assumptions in §2 hold; then*

$$\|v - \mathbf{P}_{k-1} v\| \leq C h_k^\alpha \|v\|_{1+\alpha, k} \quad \forall v \in V_k.$$

We remark that this result does not follow simply by interpolation, since we do not necessarily know the result for any integer  $\alpha \geq 1$ .

*Proof of Lemma 3.1.* For any  $v \in V_k$ , we apply the Schwarz inequality to get

$$(3.8) \quad \begin{aligned} \|v - \mathbf{P}_{k-1} v\|^2 &= a(v, v - \mathbf{P}_{k-1} v) = a(v, v - \mathbf{P}_k \mathbf{P}_{k-1} v) \\ &\leq \|v\|_{1+\alpha, k} \|v - \mathbf{P}_k \mathbf{P}_{k-1} v\|_{1-\alpha, k}. \end{aligned}$$

We next use a duality argument. For any  $\tilde{u} \in H^{\alpha-1}(\Omega)$ , let  $u \in H_0^1(\Omega)$  solve

$$(3.9) \quad a(u, \phi) = \langle \tilde{u}, \phi \rangle \quad \forall \phi \in H_0^1(\Omega);$$

then, by (2.2),  $u \in H^{1+\alpha}(\Omega)$  and  $\|u\|_{H^{1+\alpha}} \leq C \|\tilde{u}\|_{H^{\alpha-1}}$ . Letting  $\phi = v - \mathbf{P}_k \mathbf{P}_{k-1} v$  in (3.9), we get, by (3.1),

$$(3.10) \quad \begin{aligned} \|v - \mathbf{P}_k \mathbf{P}_{k-1} v\|_{1-\alpha, k} &\leq C \sup_{\|\tilde{u}\|_{H^{\alpha-1}}=1} \langle \tilde{u}, v - \mathbf{P}_k \mathbf{P}_{k-1} v \rangle \\ &= C \sup_{\tilde{u}} a(u, v - \mathbf{P}_k \mathbf{P}_{k-1} v) \\ &= C \sup_{\tilde{u}} a(u, v - \mathbf{P}_{k-1} v) + a(u, \mathbf{P}_{k-1} v - \mathbf{P}_k \mathbf{P}_{k-1} v) \\ &= C \sup_{\tilde{u}} a(u - \mathbf{I}_{k-1} u, v - \mathbf{P}_{k-1} v) + a(u - \mathbf{I}_k u, \mathbf{P}_{k-1} v - \mathbf{P}_k \mathbf{P}_{k-1} v) \\ &\leq C \sup_{\tilde{u}} (\|u - \mathbf{I}_{k-1} u\| \|v - \mathbf{P}_{k-1} v\| + \|u - \mathbf{I}_k u\| \|\mathbf{P}_{k-1} v - \mathbf{P}_k \mathbf{P}_{k-1} v\|) \\ &\leq C \sup_{\tilde{u}} \|u\|_{H^{1+\alpha}} (h_{k-1}^\alpha \|v - \mathbf{P}_{k-1} v\| + h_k^\alpha \|\mathbf{P}_{k-1} v - \mathbf{P}_k \mathbf{P}_{k-1} v\|) \\ &\leq C \sup_{\tilde{u}} \|u\|_{H^{1+\alpha}} (h_{k-1}^\alpha \|v - \mathbf{P}_{k-1} v\| \\ &\quad + h_k^\alpha (\|v - \mathbf{P}_{k-1} v\| + \|v - \mathbf{P}_k \mathbf{P}_{k-1} v\|)) \\ &= C \sup_{\tilde{u}} \|u\|_{H^{1+\alpha}} (h_{k-1}^\alpha \|v - \mathbf{P}_{k-1} v\| \\ &\quad + h_k^\alpha (\|v - \mathbf{P}_{k-1} v\| + \|\mathbf{P}_k(v - \mathbf{P}_{k-1} v)\|)) \\ &\leq C h_k^\alpha \sup_{\tilde{u}} \|\tilde{u}\|_{H^{\alpha-1}} \|v - \mathbf{P}_{k-1} v\| = C h_k^\alpha \|v - \mathbf{P}_{k-1} v\|. \end{aligned}$$

Combining (3.8) and (3.10) proves the lemma.  $\square$

**Lemma 3.2.** *Let the assumptions in §2 hold; then*

$$\|P_{k-1}v - Q_{k-1}v\| \leq Ch_k^\alpha \|v\|_{1+\alpha, k} \quad \forall v \in V_k.$$

*Proof.* We first show  $\|\phi - P_k\phi\|_{H^{1-\alpha}} \leq Ch_k^\alpha \|\phi\| \quad \forall \phi \in V_{k-1}$ . Replacing the  $\phi$  in (3.9) by  $\phi - P_k\phi$ , we get

$$\begin{aligned} \|\phi - P_k\phi\|_{H^{1-\alpha}} &= \sup_{\|\tilde{u}\|_{H^{\alpha-1}}=1} (\tilde{u}, \phi - P_k\phi) = \sup_{\tilde{u}} a(u, \phi - P_k\phi) \\ (3.11) \quad &\leq \sup_{\tilde{u}} \|u - I_k u\| \|\phi - P_k\phi\| \leq Ch_k^\alpha \sup_{\tilde{u}} \|u\|_{H^{1+\alpha}} \|\phi - P_k\phi\| \\ &\leq Ch_k^\alpha \|\phi - P_k\phi\| \leq Ch_k^\alpha \|\phi\|. \end{aligned}$$

For any  $v \in V_k$ , (3.2) implies that

$$\begin{aligned} \|P_{k-1}v - Q_{k-1}v\| &= \sup_{\phi \in V_{k-1}, \|\phi\|=1} a(P_{k-1}v - Q_{k-1}v, \phi) \\ &= \sup_{\phi} a(v, \phi - I_k\phi) = \sup_{\phi} a(v, P_k\phi - I_k\phi) \\ &\leq \sup_{\phi} \|v\|_{1+\alpha, k} \|P_k\phi - I_k\phi\|_{1-\alpha, k} \\ &\leq C \sup_{\phi} \|v\|_{1+\alpha, k} (\|\phi - P_k\phi\|_{H^{1-\alpha}} + \|\phi - I_k\phi\|_{H^{1-\alpha}}) \\ &\leq C \sup_{\phi} h_k^\alpha \|v\|_{1+\alpha, k} \|\phi\| \leq Ch_k^\alpha \|v\|_{1+\alpha, k}, \end{aligned}$$

where (3.4), (3.11) and (3.7) are used.  $\square$

#### 4. CONVERGENCE THEORY

In this section, we apply the theory of [3] to establish the constant rate of convergence of nonnested multigrid methods and show that their work estimates can be optimal order. The key assumption (A.3) in [3] is of the form

$$(4.1) \quad |a((I - I_k Q_{k-1})v, v)|^{1/2} \leq Ch_k^\beta \|v\|_{2, k}^\beta \|v\|_{1, k}^{1-\beta} \quad \forall v \in V_k.$$

This estimate can be established for the nonnested multigrid method of Definition 2.1 as follows. From (3.2),

$$\begin{aligned} a((I - I_k Q_{k-1})v, v) &= a(v, v) - a((I_k Q_{k-1})v, v) \\ &= a(v, v) - a(Q_{k-1}v, Q_{k-1}v) \\ &= a(v - Q_{k-1}v, v + Q_{k-1}v). \end{aligned}$$

From (3.6), we thus have

$$\begin{aligned} |a((I - I_k Q_{k-1})v, v)| &\leq \|v - Q_{k-1}v\| \|v + Q_{k-1}v\| \\ &\leq (1 + C_I) \|v - Q_{k-1}v\| \|v\|. \end{aligned}$$

Combining Lemmas 3.1 and 3.2, we have

$$\|v - Q_{k-1}v\| \leq Ch_k^\alpha \|v\|_{1+\alpha, k}.$$

Expanding using eigenfunctions [2] and applying Hölder's inequality shows that

$$\|v\|_{1+\alpha, k} \leq \|v\|_{1, k}^{1-\alpha} \|v\|_{2, k}^\alpha \quad \forall v \in V_k.$$

Therefore we have shown that

$$|a((I - I_k Q_{k-1})v, v)| \leq Ch_k^\alpha \|v\|_{1,k}^{2-\alpha} \|v\|_{2,k}^\alpha.$$

Thus we have proved (4.1) with  $\beta = \alpha/2$ . The following is a consequence of Theorem 7 in [3].

**Theorem 4.1** (*W-cycle scheme*). *Suppose that (2.2), (2.4) and (A1-3) hold. Let  $p \geq 2$  in Definition 2.1. For any  $0 < \gamma < 1$ , there is an integer  $m$  independent of the level number  $k$ , such that*

$$(4.2) \quad \|u_k - w_{2m+1}\| \leq \gamma \|u_k - w_0\|,$$

where  $u_k$  and  $w_i$  are defined in (2.3) and Definition 2.1, respectively.

We usually do not have a constant rate of convergence for the  $V$ -cycle scheme unless more stringent conditions hold. As suggested by Bramble et al. in [3], one can vary the number of smoothings,  $m_k$ , at different levels,  $k$ , to improve stability. The resulting "variable  $V$ -cycle" nonnested multigrid scheme has the attractive feature that its computational cost is comparable to the standard  $V$ -cycle scheme. Suppose that there are constants  $1 < \kappa_0 \leq \kappa_1 < \mathcal{F}$  (see (A1)) such that

$$(4.3) \quad \kappa_0 m_k \leq m_{k-1} \leq \kappa_1 m_k.$$

The following is a consequence of Theorem 6 in [3].

**Theorem 4.2** (*variable V-cycle scheme*). *Suppose that (2.2), (2.4), (A1-3) and (4.3) hold. Let  $p = 1$  in Definition 2.1. For any  $0 < \gamma < 1$ , (4.2) holds provided that the number of smoothings on the finest mesh,  $m_k$ , is sufficiently large.*

Standard convergence proofs and optimal-order work estimates [2] (e.g., as in Theorem 4.4 to follow) simply use the fact that the multigrid solution operator can yield an operator,  $\mathcal{G}_k$ , whose contraction number may be made sufficiently small independent of  $k$ . This can be achieved either by taking  $m_k$  sufficiently large or by iterating the  $k$ th level iteration some number,  $r$ , of times. (In the latter case, the contraction number is bounded by  $\gamma^r$ .) The proof of Theorem 7 in [3] shows that  $\gamma \leq Cm^{-\alpha/2}$  for the  $W$ -cycle. Thus, it would appear to be more efficient to pick  $r > 1$ . For example,  $r = 2$  with  $m$  smoothing steps involves essentially the same work as  $r = 1$  with  $2m$  smoothing steps. The former strategy would give an error reduction proportional to  $m^{-\alpha}$  whereas the latter would reduce by  $(2m)^{-\alpha/2}$ . (However, the resulting work estimates are of optimal order for all  $r \geq 1$ .)

Although optimal-order work estimates cannot be improved in terms of asymptotic order, there is clearly the possibility of using something more sophisticated than simple iteration to accelerate convergence, as observed in [3] and elsewhere. For example, one could use the conjugate gradient iteration [3]. Moreover, such iterative methods can yield optimal-order solvers under less stringent conditions. The following is also a consequence of Theorem 6 in [3].

**Theorem 4.3** (*variable V-cycle/conjugate gradient scheme*). *Suppose that (2.2), (2.4), (A1)-(A3) and (4.3) hold. Let  $p = 1$  in Definition 2.1. Let the operator induced by the  $k$ th level iteration given by Definition 2.1 be denoted by  $\mathcal{B}_k$ , that is (see (2.3')),  $\mathcal{B}_k f_k := w_{2m_k+1}$ , where  $w_0 = 0$ . Then  $r$  steps of the*



*preconditioned conjugate gradient method for solving  $A_k u_k = f_k$ , using  $\mathcal{B}_k$  as a preconditioner, has a contraction rate  $\gamma(r) < 1$  that tends to zero as  $r$  increases, independently of  $k$  for any  $m_k \geq 1$ .*

Nonnested multigrid methods must be implemented appropriately in order to obtain optimal-order work estimates. In the nested case, it is clear how one can compute the interpolants used in (2.6) and elsewhere in an optimal amount of work. Thus, we make the following assumption regarding the implementation of the method in the general case.

(A4) There is a constant  $C$  independent of the level,  $k$ , such that the interpolant  $I_k v$  can be computed for any  $v \in V_{k-1}$  in an amount of work bounded by  $C \dim V_k$ .

In [11], we show how this can be done in a simple way with some restrictions on the relationship between successive meshes.

With regard to convergence, note that

$$(4.4) \quad \|u - u_k\| \leq C_u h_k^\theta$$

for some  $\alpha \leq \theta \leq l$  (recall that  $\alpha$  is the degree of elliptic regularity in (2.2) and  $l$  is the degree of the piecewise polynomial approximation defined subsequent to (2.3)). In particular, if  $u \in H^{1+\theta}$ , then (4.4) holds with  $C_u = C \|u\|_{H^{1+\theta}}$ . This gives us a quantitative goal for the accuracy of an iterative solution method.

Theorems 4.1–4.3 above give bounds for operators,  $\mathcal{O}_k$ , that take an initial guess in  $V_k$  of the solution of (2.3) and then decrease the error by an amount,  $\gamma$ , that can be made arbitrarily small by increasing either (or both) of the parameters  $m$  and  $r$ . By the estimates in [2], it follows by induction that the nonnested multigrid schemes are optimal-order solvers for the finite element equations (2.3).

**Theorem 4.4** (optimal computational order for full multigrid). *Suppose that the various assumptions of Theorems 4.1, 4.2 or 4.3, respectively, hold and that  $r$  and  $m$  are chosen so that the operators,  $\mathcal{O}_k$ , decrease the error by an amount  $\gamma \leq 1/(1 + 2C_2^\theta)$ , where  $C_2$  is the constant in assumption (A3) and  $\theta$  appears in (4.4). For the  $W$ -cycle, suppose that  $2 \leq p < \mathcal{F}$  (cf. (A1)). Let  $\tilde{u}_0 = 0$  and define  $\tilde{u}_k \in V_k$  for  $k \geq 1$  by applying  $\mathcal{O}_k$  with initial guess  $I_k \tilde{u}_{k-1}$  and  $F$  as given in (2.3). If assumption (A4) holds, then the total amount of computational work and storage required to determine  $\tilde{u}_k$  is bounded by  $C \dim V_k$ . Moreover, if (4.4) holds, then*

$$\|u_k - \tilde{u}_k\| \leq C_u h_k^\theta.$$

The proof of this theorem is similar to one in [2]. We have

$$\begin{aligned} \varepsilon_k := \|u_k - \tilde{u}_k\| &\leq \gamma \|u_k - \tilde{u}_{k-1}\| \\ &\leq \gamma (\|u_k - u\| + \|u - u_{k-1}\| + \|u_{k-1} - \tilde{u}_{k-1}\|) \\ &\leq \gamma (C_u h_k^\theta + C_u h_{k-1}^\theta + \varepsilon_{k-1}) \\ &\leq \gamma (C_u (1 + C_2^\theta) h_k^\theta + \varepsilon_{k-1}) \leq C_u h_k^\theta \end{aligned}$$

by induction, because  $\gamma \leq 1/(1 + 2C_2^\theta)$ . The error estimate then follows from the triangle inequality. The work estimates are identical to those in [2].

For the conjugate gradient/variable  $V$ -cycle scheme, it is always sufficient to pick  $m_k = 1$ . This is a considerable simplification in that the number of

unspecified parameters is reduced. Since  $p = 1$  in this case as well, the only parameter to be chosen is the number,  $r$ , of conjugate gradient iterations.

#### ACKNOWLEDGMENT

We thank Jim Bramble for his suggestions regarding the applications of the theory in [3] and the presentation of results.

#### BIBLIOGRAPHY

1. R. A. Adams *Sobolev spaces*, Academic Press, New York, 1975.
2. R. Bank and T. Dupont, *An optimal order process for solving finite element equations*, Math. Comp. **36** (1981), pp. 35–51.
3. J. H. Bramble, J. E. Pasciak, and J. Xu, *The analysis of multigrid algorithms with nonnested spaces or noninherited quadratic forms*, Math. Comp. **56** (1991), 1–34.
4. S. C. Brenner, *An optimal-order multigrid method for  $P_1$  nonconforming finite elements*, Math. Comp. **52** (1989), pp. 1–15.
5. P. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland, Amsterdam, New York, Oxford, 1978.
6. M. Dauge, *Elliptic boundary value problems on corner domains*, Lecture Notes in Math., vol. 1341, Springer, Berlin and New York, 1988.
7. P. Grisvard, *Elliptic problems in nonsmooth domains*, Pitman, New York and London, 1985.
8. W. Hackbusch, *Multigrid methods and applications*, Springer, Berlin and New York, 1985.
9. S. F. McCormick, ed., *Multigrid Methods*, Frontiers in Applied Mathematics, SIAM, Philadelphia, PA, 1987.
10. L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp. **54** (1990), 483–493.
11. L. R. Scott and S. Zhang, *A nonnested multigrid method for three dimensional boundary value problems: An introduction to the NMGTM code*, in preparation.
12. Elias M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Univ. Press, Princeton, NJ, 1970.
13. R. Verfürth, *A multilevel algorithm for mixed problems*, SIAM J. Numer. Anal. **21** (1984), pp. 264–271.
14. J. Wloka, *Partial differential equations*, Cambridge Univ. Press, London and New York, 1987.
15. S. Zhang, *Multi-level iterative techniques*, Ph.D. thesis, Pennsylvania State University, 1988.
16. —, *Successive subdivisions of tetrahedra and multigrid methods on tetrahedral meshes*, (submitted), also in [15].
17. —, *Optimal-order nonnested multigrid methods for solving finite element equations I: On quasiuniform meshes*, Math. Comp. **55** (1990), 23–36.
18. —, *Optimal-order nonnested multigrid methods for solving finite element equations II: On non-quasi-uniform meshes*, Math. Comp. **55** (1990), 439–450.
19. —, *Optimal-order nonnested multigrid methods for solving finite element equations III: On degenerate meshes*, (submitted).

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF HOUSTON, HOUSTON, TEXAS 77204-3476

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF DELAWARE, NEWARK, DELAWARE 19716