Optimal-Order Nonnested Multigrid Methods for Solving Finite Element Equations I: On Quasi-Uniform Meshes

Author(s): Shangyou Zhang

Source: *Mathematics of Computation*, Vol. 55, No. 191 (Jul., 1990), pp. 23-36
Published by: American Mathematical Society
Stable URL: https://www.jstor.org/stable/2008790
Accessed: 22-10-2018 14:53 UTC

MATHEMATICS OF COMPUTATION
VOLUME 55, NUMBER 191
JULY 1990, PAGES 23-36

# OPTIMAL-ORDER NONNESTED MULTIGRID METHODS FOR SOLVING FINITE ELEMENT EQUATIONS I: ON QUASI-UNIFORM MESHES

SHANGYOU ZHANG

ABSTRACT. We prove that the multigrid method works with optimal computational order even when the multiple meshes are not nested. When a coarse mesh is not a submesh of the finer one, the coarse-level correction usually does not have the $a(\cdot, \cdot)$ projection property and does amplify the iterative error in some components. Nevertheless, the low-frequency components of the error can still be caught by the coarse-level correction. Since the (amplified) high-frequency errors will be damped out by the fine-level smoothing efficiently, the optimal work order of the standard multigrid method can still be maintained. However, unlike the case of nested meshes, a nonnested multigrid method with one smoothing does not converge in general, no matter whether it is a $V$-cycle or a $W$-cycle method.

It is shown numerically that the convergence rates of nonnested multigrid methods are not necessarily worse than those of nested ones. Since nonnested multigrid methods accept quite arbitrarily related meshes, we may then combine the efficiencies of adaptive refinements and of multigrid algorithms.

## 1. INTRODUCTION

Multigrid methods are very attractive because of their optimal order of computation. Many papers have been published studying multigrid methods. We refer to McCormick [10], Hackbusch et al. [6, 8, and 7] for references. Most of them deal with multigrid techniques in conforming finite element methods. Some are for nonconforming and mixed finite element methods (for example, [12, 16, 17, 13, 11, and 3]). In this paper, we generalize the multigrid method to cases of nonnested meshes and prove the convergence of the algorithm when solving finite element equations which arise from the Galerkin discretizations of second-order elliptic boundary value problems defined on polygonal domains.

In nonnested multigrid methods, a mesh could have quite arbitrary relations with its higher-level and lower-level mesh. Therefore, we require that two consecutive meshes be comparable, i.e., each triangle can be covered by finitely many triangles of its coarser level and of its finer level (cf. (1.6)), and that the number of unknowns of each level grow geometrically at a rate greater than 2. Under these conditions, we prove that the nonnested multigrid methods

Received June 30, 1988; revised May 8, 1989.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 65N30, 65F10.

are optimal-order algorithms. Our numerical results show that there is almost no difference in convergence rates between nonnested multigrid methods and nested ones if good coarser (finer) meshes have been used.

A motivation for the nonnested multigrid method is its application to 3-D finite element problems defined on tetrahedral meshes. In 2-D, we can subdivide a triangle into four congruent subtriangles by linking the midpoints of its edges. However we cannot usually subdivide a tetrahedron into eight identical subtetrahedra. (There exists only one, of unit size, which can be subdivided into eight identical ones; see [20]). For example, the subdivision of a regular tetrahedron gives four regular subtetrahedra and four subtetrahedra of another type. Degenerated tetrahedra could be generated in the sequence of refined meshes if inappropriate refinements are used. One way to resolve this problem would be to relax the condition of nested refinement, i.e., to use the nonnested multigrid methods. Another way of resolving the problem would be to choose some suitable methods to obtain a sequence of nested, quasi-uniform meshes. The latter approach was studied by the author in [20], and it was proved there that at most six different types of tetrahedra can be generated in the successive refinements of any tetrahedron if a special method is used. In this paper, we will treat 2-D nonnested multigrid methods only. The 3-D nonnested multigrid method will be studied in [15], where a different technique is used to prove convergence.

Besides the interest in 3-D problems, it is worthwhile knowing that the nestedness of meshes is not essential to multigrid methods. Therefore, better (adaptive) fine-level triangulations can be used. The trade-off might be the convergence rate, but not the optimal computational order of the algorithm. We confine ourselves to the case of quasi-uniform meshes in this paper and leave the case of non-quasi-uniform meshes to [23]. We note that our proof here might be used in other cases where the multiple finite element spaces are not naturally nested. For example, the author proved the optimal order of computation of the multigrid methods when applied to Hsieh-Clough-Tocher (macro), and several other $C^1$ finite element equations (cf. [21]). The author learned, after finishing this paper, that Bramble, Pasciak, and Xu in [2], too, proved convergence of the multigrid method on loosely coupled grids, using their framework. However, they considered symmetric multigrid methods, i.e., the number of presmoothing and that of postsmoothing are the same.

This paper is organized as follows. In the rest of §1, definitions and descriptions of nonnested multigrid methods will be given. In §2, some auxiliary results will be proved in preparation of the proofs for the convergence of nonnested multigrid methods and for optimal-order work estimates, which will be given in §3. In §4, we will give some numerical results.

The model problem to be solved is

$$\begin{aligned} -\partial x_1(a_1 \partial x_1 u) - \partial x_2(a_2 \partial x_2 u) + bu &= f \qquad \forall x \in \Omega, \\ u &= 0 \qquad \forall x \in \partial\Omega, \end{aligned} \tag{1.1}$$

where $\underline{a} < a_1(x), a_2(x) < \bar{a}$, $0 \le b(x) < \bar{b}$, for some positive constants $\underline{a}$, $\bar{a}$, $\bar{b}$, and

$$f \in H^{\alpha-1}(\Omega), \quad a_1, a_2, \in W^{1,\infty}(\Omega), \quad b \in L^{\infty}(\Omega) \tag{1.2}$$

for some $\alpha \in (0, 1]$. Here, $\Omega$ is a bounded polygonal domain. We use standard notations for Sobolev spaces and norms (cf. [1]). Finite element approximation problems for (1.1) can be stated as follows: Find $u_k \in V_k$ such that

$$a(u_k, v) = (f, v) \qquad \forall v \in V_k, \quad k = 1, 2, \dots, \tag{1.3}$$

where $a(u, v) = \int_\Omega (a_1 u_{x_1} v_{x_1} + a_2 u_{x_2} v_{x_2} + buv)\, dx$, $(f, v) = \int_\Omega fv\, dx$ and

$$V_k \overset{\text{def}}{=} \{v \in C_0(\Omega) \mid v|_K \text{ is linear } \forall K \in \mathcal{T}_k\} \subset H_0^1(\Omega).$$

Here, $\{\mathcal{T}_k, \ k = 1, 2, \dots\}$ is a family of triangulations on $\Omega$. We assume that the triangulations $\{\mathcal{T}_k\}$ are quasi-uniform:

$$h_K > \alpha_0 h_k, \tag{1.4}$$

$$\alpha_K > \alpha_0, \tag{1.5}$$

for any $K \in \mathcal{T}_k$, $k = 1, 2, \dots$, where $\alpha_0$ is a positive constant. Here, $h_K$, $\alpha_K$, and $h_k$ denote, respectively, the diameter of $K$, the smallest angle of $K$ and the maximum of all diameters of triangles in $\mathcal{T}_k$. $K$ is understood to be a closed triangle in this paper. We do not assume the nestedness of meshes, i.e., we can have $\mathcal{T}_{k-1} \not\subset \mathcal{T}_k$. But we assume that $\mathcal{T}_k$ can be finitely covered by $\mathcal{T}_{k-1}$ and by $\mathcal{T}_{k+1}$:

$$\begin{aligned} &\sup_{K \in T_k} \{\text{cardinality}(\{K' \in \mathcal{T}_{k-1} \mid K' \cap K \neq \varnothing\})\} \le \beta_0, \quad k = 2, 3, \dots, \\ &\sup_{K \in T_k} \{\text{cardinality}(\{K' \in \mathcal{T}_{k+1} \mid K' \cap K \neq \varnothing\})\} \le \beta_0, \quad k = 1, 2, \dots, \end{aligned} \tag{1.6}$$

and we also assume that

$$\alpha_1 N_k \le N_{k+1}, \quad \alpha_2^{-1} h_k \le h_{k+1} \le \alpha_2 h_k, \qquad k = 1, 2, \dots, \tag{1.7}$$

for some constants $\alpha_1 > 2$ and $\alpha_2 \ge 1$, where $N_k \overset{\text{def}}{=} \dim(V_k) \simeq h_k^{-2}$. For nested meshes, $\beta_0 = 4$ in (1.6), and $\alpha_1 \sim 4$, $\alpha_2 \sim 2$ in (1.7).

As usual, we define the energy norm by $|||u||| = \sqrt{a(u, u)} \quad \forall u \in H^1(\Omega)$, which is equivalent to the $H^1$ Sobolev norm in $H_0^1(\Omega)$. Further, in the finite-dimensional space $(V_k, (\cdot, \cdot))$, the bilinear form $a(\cdot, \cdot)$ defines a linear, symmetric and positive definite operator $A_k$: $a(v, w) = (v, A_k w) \quad \forall v, w \in V_k$. Then we can define the following discrete norms on $V_k$:

$$|||v|||_{s,k} = \sqrt{(v, A_k^s v)} \qquad \forall v \in V_k, \quad s \in \mathbf{R}^1. \tag{1.8}$$

We note that $|||v|||_{s,k}$ is, generally speaking, defined only for $v \in V_k$ if $s \neq 0, 1$, while

$$|||v|||_{1,k} = |||v|||, \qquad |||v|||_{0,k} = \|v\|_{L^2(\Omega)} \qquad \forall v \in H_0^1(\Omega).$$

By Lemma 1 of [1], we have

$$C\|v\|_{H^s(\Omega)} \le |||v|||_{s,k} \le C\|v\|_{H^s(\Omega)} \qquad \forall v \in V_k,\ s \in [0,1].$$

Here and later, $C$ denotes a generic constant. We use $\Lambda_k$ to denote the largest eigenvalue of $A_k$.

As the $\mathscr{T}_k$ are quasi-uniform, we have inverse inequalities (cf. [4]),

$$|||v|||_{s,k} \le Ch_k^{t-s}|||v|||_{t,k} \qquad \forall v \in V_k \ \ \& \ \ \forall k \ \text{ for } s > t, \tag{1.9}$$

and approximation properties (cf. [14] and the references in [1]): For any $u \in H^{\alpha+1}(\Omega) \cap H_0^1(\Omega)$, there exists an $\mathscr{I}_k u \in V_k$ such that

$$\|u - \mathscr{I}_k u\|_{L^2(\Omega)} + h_k\|u - \mathscr{I}_k u\|_{H^1(\Omega)} \le Ch_k^{\alpha}\|u\|_{H^{\alpha+1}(\Omega)}. \tag{1.10}$$

In particular, $I_k u$ can be taken as the $\mathscr{I}_k u$ if $u$ is continuous (cf. [14] and the proof of Proposition 2.2). Here, $I_k$ is the standard nodal value interpolator:

$$I_k u(x) = \sum_{n_i \in \mathscr{N}_k} u(n_i)\psi_{k,i}(x) \qquad \forall u \in C_0(\Omega), \tag{1.11}$$

where $\mathscr{N}_k$ is the set of nodal points corresponding to the triangulation $\mathscr{T}_k$ and $\psi_{k,i}(x)$ is the nodal basis (hat) function at the node $n_i$ of the triangulation $\mathscr{T}_k$. We assume the following elliptic regularity for (1.1) (cf. [1],[5])

$$\|u\|_{H^{\alpha+1}(\Omega)} \le C\|f\|_{L^{\alpha-1}(\Omega)}. \tag{1.12}$$

To finish this section, we define the multigrid schemes for solving (1.3). The multigrid method has two iterative processes (cf. [1]). The overall process involves solving (1.3) sequentially for $k = 1, 2, \ldots$ to get $\tilde{u}_k \simeq u_k$. To solve (1.3) on each level, $k$, we take the approximate solution of the $(k-1)$st level to be the initial guess and then we use the second, recursive process several times to get $\tilde{u}_k$. The second process involves solving more general problems: Find $u_k \in V_k$ such that

$$a(u_k, v) = \tilde{F}(v) \qquad \forall v \in V_k, \tag{1.13}$$

where $\tilde{F}$ is a linear functional on $V_k$ and can be represented in $(V_k, (\cdot,\cdot))$ as $\tilde{f}$: $\tilde{F}(v) = (\tilde{f}, v) \quad \forall v \in V_k$.

**Definition 1.1.** (The $k$th-level scheme I.)

(1) For $k = 1$, (1.3) or (1.13) is solved by any method:

$$a(\tilde{u}_1, v) = (f, v) \quad \text{or} \quad a(\tilde{u}_1, v) = \tilde{F}(v) \qquad \forall v \in V_1. \tag{1.14}$$

(2) For $k > 1$, a final guess $w_{m+1}$ will be generated from an initial guess $w_0$ as follows. $m$ steps of the so-called Jacobi-like smoothing iterations will be performed first:

$$(w_l - w_{l-1}, v) = \Lambda_k^{-1}(F(v) - a(w_{l-1}, v)) \qquad \forall v \in V_k, \quad l = 1, 2, \ldots, m. \tag{1.15}$$

To define $w_{m+1}$, we need to construct a coarse-level residual problem: Find $\bar{\varepsilon} \in V_{k-1}$, such that

$$a(\bar{\varepsilon}, v) = F(I_k v) - a(w_m, I_k v) \stackrel{\text{def}}{=} \tilde{F}(v) \qquad \forall v \in V_{k-1}, \tag{1.16}$$

where $I_k$ is defined in (1.11). Let $\varepsilon \in V_{k-1}$ be the approximation of $\bar{\varepsilon}$ obtained by applying $p$ iterations of the $(k-1)$st-level scheme to the residual equation (1.16) starting with initial guess zero. Finally, we set

$$w_{m+1} = w_m + I_k \varepsilon. \tag{1.17}$$

*Remark* 1.2. In generating the coarse-level residual problem, we have used $I_k v$ instead of $v$ in (1.16) to save computational work. Otherwise, we have to compute $a(\cdot, \cdot)$ and $(\cdot, \cdot)$ between functions of two consecutive levels:

$$a(\bar{\varepsilon}, v) = F(v) - a(w_m, v) \stackrel{\text{def}}{=} \tilde{F}(v) \qquad \forall v \in V_{k-1}. \tag{1.18}$$

Replacing (1.16) by (1.18), we can define another nonnested multigrid scheme:

**Definition 1.3.** (The $k$th-level scheme II.)
(1) If $k = 1$, use (1.14).
(2) If $k > 1$, use (1.15), (1.18) and (1.17).

*Remark* 1.4. In correcting the approximate solution $w_m$ by the solution of the coarser-level residual problem (1.16) or (1.18), we have used $I_k \varepsilon$ instead of $\varepsilon$ in (1.17) (here, $w_{m+1} = w_m + \varepsilon$ would be understood as a sum of two linear functionals of $V_k$). This is very important for implementation. Otherwise, we have to compute the representer $P_k \varepsilon$ of $\varepsilon$ in $V_k$ as a functional on $V_k$: Find $P_k \varepsilon \in V_k$ such that

$$a(P_k \varepsilon, v) = a(\varepsilon, v) \qquad \forall v \in V_k. \tag{1.19}$$

Here, $P_k : H_0^1(\Omega) \to V_k$ is the $a(\cdot, \cdot)$-projection operator. If we can invert (1.19) to get $P_k \varepsilon$ easily, we definitely should use:

$$w_{m+1} = w_m + P_k \varepsilon \tag{1.20}$$

instead of (1.17). Unfortunately, it is usually not practical to invert the problem (1.19) when $V_{k-1} \not\subset V_k$, except in a few cases (cf. [18]). In particular, if $\mathcal{T}_{k-1} \not\subset \mathcal{T}_k$, (1.19) is not feasible, since to solve it is the same as to solve (1.3). We note that when meshes are nested, (1.16) and (1.18) are identical, and so are (1.17) and (1.20).

## 2. Some auxiliary results

In this section, we will show some approximation properties of the nodal value interpolation operator $I_k$ when it transforms functions of $V_{k-1}$ to the finer-level space $V_k$. We then study the perturbation effects of the $I_k$ in the coarse-level residual problem (1.16).

**Lemma 2.1.** *Let* $\{ \mathcal{T}_k \}$ *satisfy* (1.5) *and* (1.6); *then*

$$|||w - I_k v||| \le \gamma_0 |||w - v||| \qquad \forall v \in V_{k-1} \quad \text{and} \quad \forall w \in V_k \tag{2.1}$$

*holds for some positive* $\gamma_0$ *independent of* $v$, $w$ *and* $k$, *where* $I_k$ *is the nodal interpolation operator defined in* (1.11).

*Proof.* Let $K$ be a triangle of $\mathcal{T}_k$ and $S_K$ be the union of all triangles in $\mathcal{T}_{k-1}$ which intersect $K$:

$$S_K = \bigcup \{K' \in \mathcal{T}_{k-1} \mid K' \cap K \neq \emptyset\}.$$

There can be no more than $\beta_0$ triangles in $S_K$, where $\beta_0$ is defined in (1.6). Noting that $w - I_k v$ is linear on $K$, we have

$$\begin{aligned} |w - I_k v|^2_{H^1(K)} &= C|(w - I_k v)(n_1) - (w - I_k v)(n_2)|^2 \\ &= C|(w - v)(n_1) - (w - v)(n_2)|^2 \end{aligned}$$

for some two vertices $n_1$ and $n_2$ of $K$. Let $K_1$ be the triangle, having $n_1$ and $n_2$ as its two vertices, on the other side of the line $\overline{n_1 n_2}$. Only when $\overline{n_1 n_2} \subset \partial\Omega$, may $K_1$ not exist. In this case, $(w - v)(n_1)$ and $(w - v)(n_2)$ are zero and the inequalities (2.2) below hold trivially.

Since $w - v$ is continuous and piecewise linear on the line $\overline{n_1 n_2}$, we can find a subinterval $[n_1', n_2'] \subset [n_1, n_2]$ such that

$$|(w - v)(n_1') - (w - v)(n_2')| \ge |(w - v)(n_1) - (w - v)(n_2)| / \beta_0,$$

where $n_1'$ and $n_2'$ are two intersection points between the line $\overline{n_1 n_2}$ and some two mesh lines of $\mathcal{T}_{k-1}$. Therefore, the line segment $\overline{n_1' n_2'}$ sits entirely inside some coarse-level triangle $K'$. By the minimal angle condition (1.5), it follows that we can draw a triangle $K_0$, having $n_1'$ and $n_2'$ as its two vertices, such that its area is greater than $C(\text{distance}(n_1', n_2'))^2$ and that

$$K_0 \subset K \cap K' \quad \text{or} \quad K_0 \subset K_1 \cap K'.$$

Noting that both $v$ and $w$ are linear on $K_0$, we obtain

$$\begin{aligned} |w - I_k v|^2_{H^1(K)} &\le C|(w - v)(n_1') - (w - v)(n_2')|^2 \\ &\le C|w - v|^2_{H^1(K_0)} \le C|w - v|^2_{H^1(S_K)}. \end{aligned} \tag{2.2}$$

Summing (2.2) over $\mathcal{T}_k$, we obtain by (1.6) that

$$\begin{aligned} |w - I_k v|^2_{H^1(\Omega)} &\le C \sum_{K \in \mathcal{T}_k} |w - v|^2_{H^1(S_K)} \\ &\le C\beta_0 \sum_{K' \in \mathcal{T}_{k-1}} |w - v|^2_{H^1(K')} = C|w - v|^2_{H^1(\Omega)}. \end{aligned}$$

Hence (2.1) follows from the fact that the norms $\|\cdot\|_{H^1(\Omega)}$, $|\cdot|_{H^1(\Omega)}$, and $|||\cdot|||$ are all equivalent in $H_0^1(\Omega)$. □

A different proof for Lemma 2.1 is originally given in [19], where an assumption, dominances of meshes, stronger than (1.6), is used. However, a stronger version of (2.2) is obtained:

$$|w - I_k v|^2_{H^1(K)} \le C|w - v|^2_{H^1(K)}. \tag{2.3}$$

**Proposition 2.2.** *Let* (1.4)–(1.7) *hold. Then the following estimates hold*:

$$\|v - I_k v\|_{L^2(\Omega)} \le Ch_k |||v||| \qquad \forall v \in V_{k-1}, \tag{2.4}$$

$$\|v - I_k v\|_{H^{1-\alpha}(\Omega)} \le Ch_k^{\alpha} |||v||| \qquad \forall v \in V_{k-1}. \tag{2.5}$$

*Proof.* The estimates (2.4) and (2.5) (only (2.5) is needed later) may be considered as being well known (cf. [2]), but we do still give a proof here since we cannot find any good reference. (2.4) could be proved directly (see [19]). But we use a result of the author in [22], which is generalized later by R. Scott and the author in [14]. We define a locally averaging Lagrange interpolation operator $\mathcal{I}_k$ (cf. (2.13) in [14] ), $\mathcal{I}_k : H_0^1(\Omega) \to V_k$. Here, if a node $n_i \in \mathcal{N}_k \cap \partial\Omega$, the averaging is taken on a boundary edge of $\mathcal{T}_{k-1}$; if a node $n_i \in \mathcal{N}_k \cap \Omega$, supposing $n_i \in K'$ for some $K' \in \mathcal{T}_{k-1}$, we choose, inside $K'$, the longest line segment among all containing $n_i$ for the averaging to obtain the nodal interpolating value at $n_i$. Since the averagings are taken on line segments of length $Ch_{k-1}$, by Theorem 4.1 in [14], we have

$$\begin{aligned} &\|u - \mathcal{I}_k u\|_{L^2(\Omega)} \le Ch_k \|u\|_{H^1(\Omega)}, \\ &\|u - \mathcal{I}_k u\|_{H^1(\Omega)} \le C\|u\|_{H^1(\Omega)}, \end{aligned} \qquad \forall u \in H^1(\Omega).$$

If we interpolate between $(I - \mathcal{I}_k)$ ($I$ is the identity) as a map from $H^1(\Omega)$ to $L^2(\Omega)$ and $(I - \mathcal{I}_k)$ as a map from $H^1(\Omega)$ to $H^1(\Omega)$, we obtain

$$\|(I - \mathcal{I}_k)u\|_{H^{1-\alpha}(\Omega)} \le Ch_k^{\alpha} \|u\|_{H^1(\Omega)} \qquad \forall u \in H^1(\Omega).$$

Noting that $I_k v \equiv \mathcal{I}_k v \ \ \forall v \in V_{k-1}$, the proposition is proved. □

To conclude this section, we prove a lemma concerning the difference between two coarse-level residual problems, (1.16) and (1.18).

**Lemma 2.3.** *Let* (1.4)–(1.7), (1.9), (1.10) *and* (1.12) *hold*; *then*

$$|||P_{k-1}w - Q_{k-1}w||| \le Ch_k^{\alpha} |||w|||_{\alpha+1,k} \qquad \forall w \in V_k, \tag{2.6}$$

*where* $P_{k-1}$ *is defined as in* (1.19) *and* $Q_{k-1} : V_k \to V_{k-1}$ *is defined by*

$$a(Q_{k-1}w, v) = a(w, I_k v) \qquad \forall v \in V_{k-1}. \tag{2.7}$$

*Proof.* We compare the definitions of $P_{k-1}$ and $Q_{k-1}$ to get

$$a(P_{k-1}w - Q_{k-1}w, v) = a(w, v - I_k v) \quad \forall v \in V_{k-1}.$$

Noting $P_{k-1}w$, $Q_{k-1}w \in V_{k-1}$, it follows by the Schwarz inequality that

$$
\begin{aligned}
|||P_{k-1}w - Q_{k-1}w||| &= \sup_{v\in V_{k-1},\ |||v|||=1} a(P_{k-1}w - Q_{k-1}, v) \\
&= \sup_v a(w, v - I_k v) = \sup_v a(w, P_k v - I_k v) \\
&\le \sup_v |||w|||_{\alpha+1,k} |||P_k v - I_k v|||_{1-\alpha,k} \\
&\le C \sup_v |||w|||_{\alpha+1,k} (||P_k v - v||_{H^{1-\alpha}(\Omega)} + ||v - I_k v||_{H^{1-\alpha}(\Omega)}),
\end{aligned} \tag{2.8}
$$

since $|||\cdot|||_{1-\alpha,k}$ and $\|\cdot\|_{H^{1-\alpha}(\Omega)}$ are equivalent in $V_k$. To estimate $||P_k v - v||_{H^{1-\alpha}(\Omega)}$, we can use a standard duality argument to get (cf. (3.14) in [1])

$$||v - P_k v||_{H^{1-\alpha}(\Omega)} \le C h_k^{\alpha} |||v - P_k v||| \le C h_k^{\alpha} |||v||| = C h_k^{\alpha}.$$

Using Proposition 2.2 in (2.8) yields (2.6). □

## 3. Convergence and work estimates of nonnested multigrid methods

In this section, the main theorems of the paper, the convergence and the optimal computational order of the nonnested mesh multigrid methods, will be proved. The method is based on the principle of mathematical induction. The proof follows the ideas in the proof of Theorem 1 of Bank and Dupont [1].

**Theorem 3.1.** *Let* (1.4)–(1.7), (1.9), (1.10) *and* (1.12) *hold and* $p > 1$ *be an integer. Then there exists a constant* $0 < \gamma < 1$ *and an integer* $m \ge 1$, *all independent of the level number* $k$, *such that, if*

$$|||\bar{\varepsilon} - \varepsilon||| \le \gamma^p |||\bar{\varepsilon}|||, \tag{3.1}$$

*then*

$$|||u_k - w_{m+1}||| \le \gamma |||u_k - w_0|||, \tag{3.2}$$

*where* $u_k$, $w_i$, $\varepsilon$, *and* $\bar{\varepsilon}$ *are defined in Definition* 1.1.

*Proof.* Let the iterative errors be denoted by $e_i = u_k - w_i$, $0 \le i \le m+1$. By expanding the initial error as a linear combination of the eigenfunctions of $A_k$, we get the well-known estimates for the Jacobi-like iteration (1.15) (cf. (3.8) and (3.13) in [1])

$$|||e_m||| \le |||e_0|||, \qquad |||e_m|||_{\alpha+1,k} \le C h_k^{-\alpha} m^{-\alpha/2} |||e_0|||. \tag{3.3}$$

Our aim is to estimate the final error $e_{m+1}$ after one cycle of $k$th-level iteration. By the definition (1.17), Lemma 2.1, the triangle inequality and the assumption (3.1), we get

$$
\begin{aligned}
|||e_{m+1}||| &= |||e_m - I_k \varepsilon||| \le \gamma_0 |||e_m - \varepsilon||| \le \gamma_0 |||e_m - \bar{\varepsilon}||| + \gamma_0 |||\bar{\varepsilon} - \varepsilon||| \\
&\le \gamma_0 (1 + \gamma^p) |||e_m - \bar{\varepsilon}||| + \gamma_0 \gamma^p |||e_m||| \\
&\le C |||e_m - P_{k-1} e_m||| + C |||P_{k-1} e_m - Q_{k-1} e_m||| + \gamma_0 \gamma^p |||e_m|||,
\end{aligned} \tag{3.4}
$$

where $Q_{k-1} e_m = \bar{\varepsilon}$ by (2.7) and (1.16).

For the first term on the right-hand side of (3.4), using the $a(\cdot, \cdot)$-projection property of $P_{k-1}$, we have

$$
\begin{aligned}
|||e_m - P_{k-1}e_m|||^2 &= a(e_m - P_{k-1}e_m, e_m) = a(e_m - P_kP_{k-1}e_m, e_m) \\
&\le |||e_m - P_kP_{k-1}e_m|||_{1-\alpha,k}\, |||e_m|||_{1+\alpha,k}.
\end{aligned} \tag{3.5}
$$

We note that due to nonnestedness the Schwarz inequality cannot be used directly to $a(e_m - P_{k-1}e_m, e_m)$. We then use a duality argument with (1.12) and (1.10): For any $\theta_0 \in H^{\alpha-1}(\Omega)$, let $\theta_2 \in H^{\alpha+1}(\Omega)$ solve

$$
a(\theta_2, v) = (\theta_0, v) \qquad \forall v \in H_0^1(\Omega);
$$

then

$$
\begin{aligned}
(\theta_0, e_m - P_kP_{k-1}e_m) &= a(\theta_2, e_m - P_kP_{k-1}e_m) \\
&= a(\theta_2, e_m - P_{k-1}e_m) + a(\theta_2, P_{k-1}e_m - P_kP_{k-1}e_m) \\
&= a(\theta_2 - \mathcal{I}_{k-1}\theta_2, e_m - P_{k-1}e_m) + a(\theta_2 - \mathcal{I}_k\theta_2, P_{k-1}e_m - P_kP_{k-1}e_m) \\
&\le |||\theta_2 - \mathcal{I}_{k-1}\theta_2|||\, |||e_m - P_{k-1}e_m||| \\
&\quad + |||\theta_2 - \mathcal{I}_k\theta_2|||\, |||P_{k-1}e_m - P_kP_{k-1}e_m||| \\
&\le Ch_{k-1}^{\alpha}|\theta_2|_{H^{\alpha+1}(\Omega)}|||e_m - P_{k-1}e_m||| \\
&\quad + Ch_k^{\alpha}|\theta_2|_{H^{\alpha+1}(\Omega)}|||P_{k-1}e_m - P_kP_{k-1}e_m||| \\
&\le Ch_k^{\alpha}|\theta_0|_{H^{\alpha-1}(\Omega)}(2|||e_m - P_{k-1}e_m||| + |||P_k(e_m - P_{k-1}e_m)|||) \\
&\le 3Ch_k^{\alpha}|\theta_0|_{H^{\alpha-1}(\Omega)}|||e_m - P_{k-1}e_m|||,
\end{aligned}
$$

where $h_k \sim h_{k-1}$ from (1.7) is used. Since $\theta_0$ is arbitrary, it follows, by Lemma 1 in [1], that

$$
\begin{aligned}
|||e_m - P_kP_{k-1}e_m|||_{1-\alpha,k} &\le C\|e_m - P_kP_{k-1}e_m\|_{H^{1-\alpha}(\Omega)} \\
&\le Ch_k^{\alpha}|||e_m - P_{k-1}e_m|||.
\end{aligned} \tag{3.6}
$$

By (3.6) and (3.3), we get from (3.5) that

$$
|||e_m - P_{k-1}e_m||| \le Cm^{-\alpha/2}|||e_0|||. \tag{3.7}
$$

The second term on the right-hand side of (3.4) has been estimated in Lemma 2.3. By (3.3), it follows that

$$
|||P_{k-1}e_m - Q_{k-1}e_m||| \le Ch_k^{\alpha}|||e_m|||_{1+\alpha,k} \le Cm^{-\alpha/2}|||e_0|||. \tag{3.8}
$$

By (3.7), (3.8), and (3.3), the estimate (3.4) becomes

$$
|||e_{m+1}||| \le (Cm^{-\alpha/2} + \gamma_0\gamma^p)|||e_0|||. \tag{3.9}
$$

To complete the proof, we can choose $0 < \gamma < 1$ small enough such that $\gamma^p\gamma_0 \le \gamma/2$, since we have $p > 1$. And then we can let $m$ be large enough such that $Cm^{-1/2} \le \gamma/2$. We can see that the choices of $\gamma$ and $m$ are independent of the level number $k$, since $C$ and $\gamma_0$ are independent of $k$. Using these bounds in (3.9), the assertion (3.2) follows. □

**Corollary 3.2.** *Theorem* 3.1 *holds when* "*Definition* 1.1" *is replaced by* "*Definition* 1.3".

*Proof.* The inequalities (3.4) in the proof of Theorem 3.1 become, in this case,

$$\begin{aligned} |||e_{m+1}||| &\le \gamma_0 |||e_m - P_{k-1}e_m||| + \gamma_0 |||\bar{\varepsilon} - \varepsilon||| \\ &\le C|||e_m - P_{k-1}e_m||| + \gamma_0 \gamma^p |||e_m|||, \end{aligned}$$

since $P_{k-1}e_m = \bar{\varepsilon}$ in the scheme II. Hence, by (3.7) and (3.3), the proof is completed. □

In practice, we would like to use some other simpler inner product $b_k(\cdot, \cdot)$ in the fine-level smoothing iteration (1.15) instead of the $L^2$-inner product there. In order to make Theorem 3.1 and the above corollary hold, we may choose $b_k(\cdot, \cdot)$ such that the norm induced by $b_k(\cdot, \cdot)$ is equivalent to the $L^2$-norm. As pointed out by Bank and Dupont in [1], there are many possible choices for $b_k(\cdot, \cdot)$. For example, we can let $b_k(\cdot, \cdot)$ be defined by the diagonal of the mass matrix if the standard nodal basis is used. Defining $\tilde{A}_k$ by $a(v, w) = b_k(\tilde{A}_k v, w) \quad \forall v, w \in V_k$ and defining $|||\cdot|||_{s,k,b}$ by $|||v|||^2_{s,k,b} = b_k(\tilde{A}_k^s v, w) \quad \forall v \in V_k$, we have the following equivalences of norms:

$$\begin{aligned} &\beta^{-1}|||v|||_{0,k} \le |||v|||_{0,k,b} \le \beta|||v|||_{0,k}, \\ &\beta^{-1}|||v|||_{1,k} \le |||v|||_{1,k,b} \le \beta|||v|||_{1,k} \qquad \forall v \in V_k \end{aligned}$$

for some positive constant $\beta$. Now we can replace (1.15) by the following smoothing iteration:

$$b_k(w_l - w_{l-1}, v) = \tilde{\Lambda}_k^{-1}(F(v) - a(w_{l-1}, v)) \qquad \forall v \in V_k, \quad l = 1, 2, \ldots, m, \tag{3.10}$$

where $\tilde{\Lambda}_k$ is the maximal eigenvalue of $\tilde{A}_k$. One iteration of (3.10) needs only $O(N_k)$ computations.

**Corollary 3.3.** *Theorem* 3.1 *and Corollary* 3.2 *hold if the fine-level smoothing* (1.15) *is replaced by* (3.10) *there.*

*Proof.* Repeating the proof of Theorem 3.1, we have the analogue of (3.5) as

$$|||e_m - P_{k-1}e_m|||^2 \le \beta|||e_m - P_k P_{k-1}e_m|||_{1-\alpha,k}|||e_m|||_{1+\alpha,k,b},$$

and the estimate of $|||e_m|||_{2,k}$ will be changed similarly for $|||e_m|||_{2,k,b}$, but with the right-hand constant $C$ depending on the $\beta$. The rest of the proof remains the same. □

**Theorem 3.4.** *Let the assumptions of Theorem* 3.1 *hold and let* $2 \le p < \alpha_1$ ($\alpha_1$ *is defined in* (1.7)). *There exists a constant* $r$ *and a constant* $\delta > 0$ *such that if*

$$|||u_1 - \hat{u}_1||| \le \delta C h_1$$

*holds for some constant* $C$, *then*

(1) $|||u_k - \hat{u}_k||| \leq \delta C h_k, \quad k \geq 2,$

(2) $|||u - \hat{u}_k||| \leq (1+\delta) C h_k, \quad k \geq 1,$

(3) *the cost of computing* $\hat{u}_k$ *is bounded by* $C_0 N_k$, *where* $C_0$ *is independent of the level number* $k$.

*Here,* $\hat{u}_k$ *is obtained by doing* $r$ "*modified* $k$*th-level scheme* I" *defined by* (1.14), (3.10), (1.16) *and* (1.17) *with* $\hat{u}_{k-1}$ *as initial guess, and* $\hat{u}_1$ *is obtained by solving* (1.3) *directly.*

*Proof.* Noting the sparseness of the transfer matrix associated with $I_k$, the proof is the same as the one for Theorem 2 of [1]. □

**Corollary 3.5.** *Theorem* 3.4 *holds if the* "*modified* $k$*th-level scheme* I" *is replaced by the* "*modified* $k$*th-level scheme* II", *defined by* (1.14), (3.10), (1.18) *and* (1.17).

*Proof.* The difference in work estimate is only in evaluating inner products between the $k$th-level functions and nodal basis functions of $V_{k-1}$. Once $a(\cdot, \cdot)$ and the $L^2$ inner products between nodal basis functions of all two consecutive levels are computed, we need just order $N_k$ computations to evaluate those matrix vector products, since all matrices are sparse. □

## 4. Numerical experiments

The numerical problems to be dealt with are continuous piecewise linear finite element equations arising from the Poisson equation

$$\begin{cases} -\triangle u = f & \forall x \in \Omega = (0,1) \times (0,1), \\ u = 0 & \forall x \in \partial\Omega. \end{cases}$$

We have two sets of experiments on the two-level nonnested multigrid method defined in Definition 1.1. The fine triangulations are always uniform in the experiments, which have mesh sizes $h_f = 1/n_f$ for

$$n_f = 4, 6, 8, 10, 12, 14, 16, 18, \text{ and } 20,$$

where $n_f$ is the number of intervals in each direction. In the first set of experiments, the coarse meshes are obtained by shifting the standard nested coarse mesh ( (a): shifted by $1/n_f$ with one more grid point added in each direction, (b): no shifting, (c): shifted by $1/2n_f$ and (d): shifted by $1/n_f$ ). In the second set, the coarse meshes are again uniform, but with mesh sizes: $1/[n_f/2-1]$ (e), $1/[n_f/2]$ (f) (the case of the standard nested-meshes), $1/[n_f/2+1]$ (g) or $1/[n_f/2+2]$ (h). We computed the spectral radii of the two-level nonnested multigrid iterative operators and plotted them in Figure 1. For detailed descriptions and data, we refer the reader to [19]. In the left graph, (a1–h1) denote the contractive radii of the method of one smoothing with one exact coarse-level correction on meshes (a–h). We can see that the rates of (a1), (g1) and (h1)

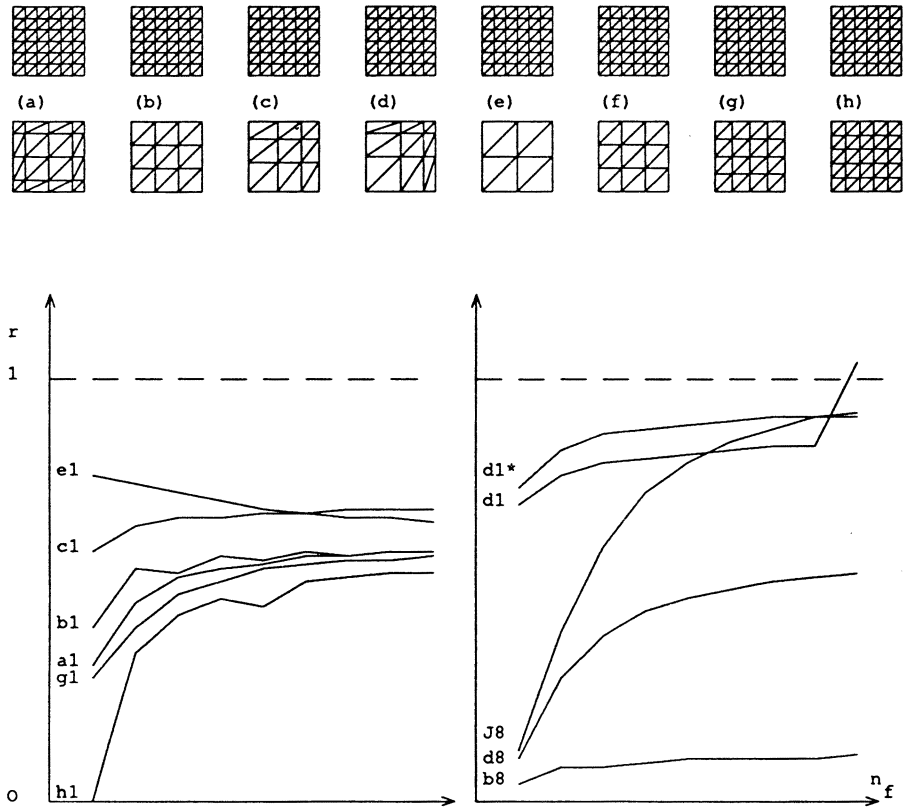


FIGURE 1

*Meshes and spectral radii of some two-level nonnested multigrid methods.*

are better than the rate in the case of nested meshes (b1). In fact, in (a), (g), and (h), there are more grid points on the coarse level. On the other hand, (d1) in the right graph shows that one smoothing nonnested multigrid methods do not converge in general (see [9] for the case of nested-meshes). However, as predicted by our theorem, (d8), in the right graph, with eight smoothings, converges.

The instability of the coarse-level correction in the nonnested multigrid method (d1) is caused by the perturbation of the $I_k$ in (1.16) and (1.17). By our previous analysis, we have

$$\begin{aligned}
|||e_{m+1}||| &= |||e_m - I_k\bar{e}||| \le \gamma_0|||e_m - P_{k-1}e_m||| + \gamma_0|||P_{k-1}e_m - Q_{k-1}e_m||| \\
&\le \gamma_0|||e_m||| + \gamma_0 \sup_{v\in V_{k-1}} \frac{a(e_m, v - I_k v)}{|||v|||} \\
&\le \gamma_0|||e_m||| \left[1 + (1+\gamma_0)\sup_{v} \frac{|||P_k v - v|||}{|||v|||}\right] \\
&\le (2\gamma_0 + \gamma_0^2)|||e_m||| = C(\gamma_0)|||e_m|||,
\end{aligned}$$

where $C(\gamma_0)$ could be larger than one. But, when the meshes are nested, we have the $a(\cdot, \cdot)$ projection property for the coarse-level correction:

$$|||e_{m+1}||| = |||e_m - P_{k-1}e_m||| \leq |||e_m|||.$$

This instability could be reduced somewhat by modifying (1.17) to $w_{m+1} = w_m + \rho I_k \varepsilon$ with small $\rho$. The (d1*) in the right graph is the radius for $\rho = 0.25$. However, small $\rho$ would reduce the rate of convergence of the iteration. In the right graph, we also plot the radii for eight smoothings (J8), and for the case of nested-meshes with eight smoothings (b8).

Although the rates for (a), (g) and (h) are better than those in the case of nested meshes (b) and (f), we do not suggest using larger coarse-level spaces. There are two reasons for this. One is that fine coarse-meshes could cost more computations. The other is that unnecessarily fine coarse-meshes could cause $I_k$ to be very degenerate ($I_k v = 0$ for some $0 \neq v \in V_{k-1}$), which might make some effort of computing (1.16) worthless and might also increase the instability (the $\beta_0$ in (1.6) increases).

Finally, we note that in implementation of multigrid methods, no matter whether the meshes are nested or not, the codes are the "same", since the $I_k$ is needed to transform the internal representations of coarse-level functions to the internal representations of them as fine-level functions in the case of nested meshes, too.

## Acknowledgments

The author is indebted to the referee who suggested the method for proving Lemma 2.1, which considerably simplifies the original proof. The author would like to thank Professor Ridgway Scott, the author's thesis advisor, for his advice and help.

## Bibliography

1. R. Bank and T. Dupont, *An optimal order process for solving finite element equations*, Math. Comp. **36** (1981), 35–51.
2. J. H. Bramble, J. E. Pasciak, and J. Xu, *The Analysis of multigrid algorithms with nonnested spaces or non-inherited quadratic forms* (preprint).
3. S. C. Brenner, *An optimal-order multigrid method for $P_1$ nonconforming finite elements*, Math. Comp. **52** (1989), 1–15.
4. P. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland, Amsterdam, New York, Oxford, 1978.
5. P. Grisvard, *Elliptic problems in nonsmooth domains*, Pitman, 1985.
6. W. Hackbusch and U. Trottenberg, *Multigrid methods* II, Lecture Notes in Math., vol. 1228, Springer-Verlag, Berlin and New York, 1986.
7. ———, *Multigrid methods*, Lecture Notes in Math., vol. 960, Springer-Verlag, Berlin and New York, 1982.
8. W. Hackbusch, *Multigrid methods and applications*, Springer, 1985.
9. J. Mandel, *Multigrid convergence for nonsymmetric, indefinite variational problems and one smoothing step*, Appl. Math. Comput. **19** (1986), 201–216.

10. S. F. McCormick, *Multigrid Methods,* Frontiers in Applied Mathematics, SIAM, Philadel phia, PA, 1987.
11. P. Peisker, *A multilevel algorithm for the biharmonic problem*, Numer. Math. **46** (1985) 623–634.
12. P. Peisker and D. Braess, *A conjugate gradient method and a multigrid algorithm for Mor ley's finite element approximation of the biharmonic equation*, Numer. Math. **50** (1987) 567–586.
13. J. Pitkäranta and T. Saarinen, *A multigrid version of a simple finite element method for th Stokes problem*, Math. Comp. **45** (1985), 1–14.
14. L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp. **54** (1990), (to appear).
15. ———, *A nonnested multigrid method for three dimensional boundary value problems: An introduction to NMGTM code*, in preparation.
16. R. Verfürth, *A multilevel algorithm for mixed problems*, SIAM J. Numer. Anal. **21** (1984) 264–271.
17. ———, *A combined conjugate gradient-multigrid algorithm for the numerical solution of the Stokes problem*, IMA J. Numer. Anal. **4** (1984), 441–455.
18. ———, *Multilevel algorithms for mixed problems.* II. *Treatment of the mini-element*, SIAM J. Numer. Anal. **25** (1988), 285–293.
19. S. Zhang, *Multi-level iterative techniques*, Ph. D. thesis, Pennsylvania State University, 1988.
20. ———, *Successive subdivisions of tetrahedra and multigrid methods on tetrahedral meshes*, submitted, also in [19].
21. ———, *Multigrid methods for solving biharmonic, $C^1$ finite element equations*, Numer. Math. **56** (1989), 613–624.
22. ———, *An optimal order interpolation operator from $H_0^1$ to piecewise linear functions $S_0^1$ on polygons*, §6.3 in [19].
23. ———, *Optimal-order nonnested multigrid methods for solving finite element equations* II: *Non-quasiuniform meshes*, Math. Comp. **55** (1990), (to appear).

DEPARTMENT OF MATHEMATICS, PURDUE UNIVERSITY, WEST LAFAYETTE, INDIANA 47907