MATHEMATICS OF COMPUTATION
VOLUME 64, NUMBER 209
JANUARY 1995, PAGES 23–49

OPTIMAL-ORDER NONNESTED MULTIGRID METHODS FOR SOLVING FINITE ELEMENT EQUATIONS III: ON DEGENERATE MESHES

SHANGYOU ZHANG

ABSTRACT. In this paper, we consider several model problems where finite element triangular meshes with arbitrarily small angles (high aspect ratios) are utilized to deal with anisotropy, interfaces, or singular perturbations. The constant-rate (independent of the number of unknowns, the smallest angle, the interface discontinuity, the singular-perturbation parameter, etc.) convergence of some special nonnested multigrid methods for solving the finite element systems on such degenerate meshes will be proved. Numerical data are provided to support the analysis in each case.

1. INTRODUCTION

To be a complete method for solving boundary value problems, not only should the method provide a discretization scheme to approximate the PDE, but also it should include a fast algorithm to solve the resulting discrete linear systems. The combination of the two determines the efficiency of the numerical method. Multigrid iterative methods provide optimal-order solvers for many finite element systems (cf. [18, 11] and the references therein). In a standard multigrid method, a nested family of triangular meshes is generated from an initial mesh by connecting midpoints of edges sequentially. But this sequence of nested grids may not be economical in many situations. Other refinement methods should be included in the multigrid method. Naturally, one would relax the condition of nestedness of grids. In our previous paper [25], a nonnested multigrid method is considered for boundary value problems with corner singularities, where nonnested meshes with high refinements at corners are employed.

Meshes with high aspect ratios (high refinement in some particular directions) are encountered often in practice, owing to the geometry of the domain, or to the roughness of the solution in some directions. We will study some model cases only in this paper. In §2, we will define a nonnested multigrid method to solve Poisson equations where finite element meshes have high aspect ratios. In §§3 and 4, a framework is presented to analyze nonnested multigrid methods on some degenerate meshes. Numerical data presented in §4 show that the convergence rate of nested multigrid methods deteriorates when the minimal-angle condition is violated, while the rate of the new nonnested multigrid method remains the same. In §5, two approaches will be studied to solve an

Received by the editor March 26, 1991 and, in revised form, May 3, 1993.
1991 *Mathematics Subject Classification.* Primary 65N30, 65F10.

anisotropic, singular perturbation problem, and the multigrid method is shown to be of constant convergence rate in both cases. In §6, the nested multigrid method is shown to have constant contractivity which is independent of interface jumpings and mesh sizes, when applied to solve for an interface problem in which the interface is covered completely by mesh lines. Also in §6 the nonnested multigrid is proved to have a constant rate of convergence, too, if a narrow region along the interface is covered by skinny triangles. Finally, in §7 we will show the constant contractivity of a nonnested multigrid method for the finite element equations arising from discretizing a convection-dominated convection-diffusion equation by either high-aspect-ratio meshes, or the streamline diffusion method. We should emphasize, however, that the domains in all the problems considered are squares, and that the meshes are "regular" in the sense that all elements are right triangles with one horizontal and one vertical edge. The analysis is not ready to be extended to general domains or to general nonconstant-coefficient problems. But the method should be applicable to general problems. There is one more remark. Some results in §§4–5 might also be obtained by doing Fourier Analysis as, for example, in [11], but not those in §§6–7.

To treat the meshes with small angles (of high aspect ratios) in the multigrid method, we modify two ingredients of the multigrid method while retaining its principles. The usual fine-level smoothing is replaced by a weighted smoothing where the weight at each node depends on those triangles carrying the node and on the local diffusion coefficients. Only those coarse meshes which are coarser in some directions are permitted in the coarse-level correction. The first modification is also necessary in the multigrid method for nonquasiuniform meshes (cf. [25] and the references therein). As in the standard multigrid method, the new fine-level smoothing reduces high-frequency components of iterative errors "uniformly" over the whole domain, while the new coarse-level correction corrects all low-frequency error components.

The approximation of finite elements with arbitrarily small angles has been studied in [3] and [13]. However, we do not use these estimates in this paper, nor do we use the elliptic regularity assumption. One could use approximation and regularity theories, but one would usually obtain a convergence rate depending on the minimal angle, or the singular perturbation parameter. The convergence of the multigrid method without elliptic regularity has been studied by Bramble et al. in [9]. Nonnested multigrid methods are considered in [8], too. The multigrid method for interface problems (finite element equations) can be found in [22] and [9]. Other related works will be mentioned in the subsequent sections.

2. Definitions

In this section, we define a nonnested multigrid method by considering finite element solutions of the following Poisson equation:

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega := (0,1)\times(0,1), \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{2.1}$$

The multilevel finite element spaces, $\{V_k \subset H_0^1(\Omega) | k = 1, 2, \dots\}$, consist of continuous, piecewise linear functions (cf. [10]) on meshes $\{\mathcal{T}_k,\ k = 1, 2, \dots\}$ as shown in Figure 1, where a coarse mesh is coarser only in the x_1-direction.

That is, refinements take place in the x_1-direction only. Further, each mesh $\mathcal{T}_k$ is uniform in the x_1-direction and the x_2-direction separately with mesh sizes h_k and H_k, respectively. Therefore, $h_k = h_{k-1}/2$ and $H_k = H_{k-1}$. The growth rate of the number of multilevel unknowns is about 2. We will consider some other meshes later. For convenience, we let the aspect ratio be

$$\rho_k := h_k/H_k \le 1. \tag{2.2}$$

Obviously, the multilevel finite element spaces are not nested: $V_{k-1} \not\subset V_k$. We remark that one can alternatively refine the mesh in each direction.

We define the bilinear form associated with the Laplace operator and a discrete inner product (scaled l^2 inner product) as follows:

$$a(u, v) := \int_\Omega \nabla u \cdot \nabla v \, dx \quad \forall u, v \in V_k, \tag{2.3}$$

$$(u, v)_k := \sum_{n \in \mathcal{N}_k} \frac{H_k^3}{h_k} u(n)v(n) \quad \forall u, v \in V_k, \tag{2.4}$$

where $\mathcal{N}_k$ is the set of vertices in the triangulation $\mathcal{T}_k$. We note that one would get different weighted fine-level smoothings by defining different discrete inner products $(\cdot, \cdot)_k$. This can be seen from (2.9) and (6.2). Let $A_k: V_k \to V_k$ be the symmetric, positive definite operator defined by $(A_k u, v)_k := a(u, v)$ $\forall v \in V_k$. As usual, we define a family of discrete norms by

$$|||v|||_{s,k}^2 := (A_k^s v, v)_k \quad \forall v \in V_k, \ 0 \le s \le 2. \tag{2.5}$$

For any $v \in V_k$,

$$|||v|||_{1,k} = \sqrt{a(v, v)} \quad \text{and} \quad |||v|||_{0,k} = \sqrt{(v, v)_k}.$$

Let $I_k: H_0^1(\Omega) \cap C(\Omega) \to V_k$ be the usual Lagrange nodal-value interpolation operator. Let the vector space $\tilde{V}_{k-1}$ be the image of I_k when restricted on V_{k-1}:

$$\tilde{V}_{k-1} := \text{Range}(I_k|_{V_{k-1}}). \tag{2.6}$$

We note that $I_k: V_{k-1} \to \tilde{V}_{k-1}$ is a one-to-one and onto mapping. (In general, I_k may be singular in the sense that $I_k v = 0$ for some $0 \ne v \in V_{k-1}$ if $\mathcal{T}_{k-1} \not\subset \mathcal{T}_k$; cf. [23].) We introduce two auxiliary operators for the two common coarse-level corrections in the nonnested multigrid method,

$$\begin{aligned} &Q_{k-1}: V_k \to V_{k-1}, \qquad a(Q_{k-1}e, v) = a(e, I_k v) \quad \forall v \in V_{k-1}, \\ &R_{k-1}: V_k \to \tilde{V}_{k-1}, \qquad a(R_{k-1}e, v) = a(e, v) \qquad \forall v \in \tilde{V}_{k-1}. \end{aligned} \tag{2.7}$$

Finite element approximations of (2.1) are defined as follows: Find $u_k \in V_k$ such that

$$a(u_k, v) = F(v) \quad \forall v \in V_k, \tag{2.8}$$

where $F(v) = \int_\Omega f(x)v(x)\,dx$. One kth-level multigrid iteration is defined (recursively) by the following two steps (cf. [5]) to produce a new iterative solution w_{m+1} (for u_k in (2.8) or for $\bar{q}$ in (2.11)) from a given initial guess w_0. First, m fine-level smoothings will be performed:

$$(w_l - w_{l-1}, v)_k = \lambda_k^{-1}(F(v) - a(w_{l-1}, v)) \quad \forall v \in V_k, \ 1 \le l \le m, \tag{2.9}$$

where λ_k is the maximal eigenvalue of the A_k in (2.5) and $F(v)$ is either defined in (2.8) or in (2.11) below. The second step is a coarse-level correction:

$$w_{m+1} = w_m + I_k q\,, \tag{2.10}$$

where $q \in V_{k-1}$ is the iterative solution, obtained by doing p (> 1) $(k-1)$st-level multigrid iterations with 0 initial guess for the residual problem: Find $\overline{q} \in V_{k-1}$ such that

$$a(\overline{q}\,, v) = F(I_k v) - a(w_m\,, I_k v) =: F_{\text{new}}(v) \quad \forall v \in V_{k-1}\,. \tag{2.11}$$

3. Preliminary analysis

In order to prove the constant rate of convergence for the nonnested multigrid method, we need a few lemmas concerning stability and approximability properties of the operators I_k and Q_k defined in §2. In the next two sections, we use explicit constants instead of the generic constant C in the estimation. Since we consider a special model problem, explicit constants may better exhibit the idea in estimation.

Lemma 3.1. *Let* $v \in V_k$. *If* $|||v|||_{0,k} \le C_0|||v|||_{1,k}$ *for some* $C_0 \ge 0$, *then* $|||v|||_{1,k} \le C_0|||v|||_{2,k}$.

Proof. Let $0 < \alpha_1 \le \alpha_2 \le \cdots \le \alpha_n$ be the eigenvalues of A_k: $A_k\theta_i = \alpha_i\theta_i$, where $|||\theta_i|||_{0,k} = 1$ and $n := \dim V_k$. We can expand v in the eigenvector basis: $v = \sum_{i=1}^n v_i\theta_i$. Consequently, $|||v|||_{s,k}^2 = \sum v_i^2\alpha_i^s$. For convenience, we let $\alpha_{n+1} = \infty$. We denote the index i_0 such that $C_0^{-2} < \alpha_{i_0}$ and $C_0^{-2} \ge \alpha_i$ for all $i < i_0$. By separating positive and negative terms, the lemma is proved as follows:

$$\begin{aligned} C_0^2|||v|||_{2,k}^2 - |||v|||_{1,k}^2 &= -\sum_{i=1}^{i_0-1} v_i^2(1 - \alpha_i C_0^2)\alpha_i + \sum_{i=i_0}^{n} v_i^2(\alpha_i C_0^2 - 1)\alpha_i \\ &\ge -\alpha_{i_0}\sum_{i=1}^{i_0-1} v_i^2(1 - \alpha_i C_0^2) + \alpha_{i_0}\sum_{i=i_0}^{n} v_i^2(\alpha_i C_0^2 - 1) \\ &= \alpha_{i_0}\left(\sum_{i=1}^{n} v_i^2(\alpha_i C_0^2 - 1)\right) = \alpha_{i_0}(C_0^2|||v|||_{1,k}^2 - |||v|||_{0,k}^2) \ge 0. \quad \square \end{aligned} \tag{3.1}$$

Lemma 3.2. *Let* $B = \text{tridiag}(-B_{i-1}\, B_i\, B_{i+1})$ *be a block tridiagonal matrix,*

$$B = \begin{pmatrix} B_0 + B_1 & -B_1 & & \\ -B_1 & B_1 + B_2 & -B_2 & \\ & \ddots & \ddots & \ddots \\ & & -B_{n-1} & B_{n-1} + B_n \end{pmatrix},$$

where $B_0, \ldots, B_n$ *are symmetric and nonnegative semidefinite. Then, so is* B.

Proof. Let $X = (X_1, X_2, \ldots, X_n)$ be an arbitrary vector in block form and $X_0 = X_{n+1} = 0$. We need to show $X^\top BX \ge 0$; this is done by the Cauchy-

Schwarz inequality:

$$X^\top BX = \sum_{i=1}^{n} X_i^\top(-B_{i-1}X_{i-1} + B_{i-1}X_i + B_iX_i - B_iX_{i+1})$$
$$\geq \sum_{i=1}^{n}\Bigg(- \frac{X_i^\top B_{i-1}X_i + X_{i-1}^\top B_{i-1}X_{i-1}}{2}$$
$$+X_i^\top B_{i-1}X_i + X_i^\top B_iX_i - \frac{X_i^\top B_iX_i + X_{i+1}^\top B_iX_{i+1}}{2}\Bigg)$$
$$= X_1^\top B_0X_1 + X_n^\top B_nX_n \geq 0. \quad \square$$

Lemma 3.3. *For all* $u \in V_k$, *there holds*

$$|||\mathrm{I}_k\mathrm{I}_{k-1}u|||_{2,k} \leq 2|||u|||_{2,k}. \tag{3.2}$$

Proof. We will do a direct estimate. We scale the usual nodal basis by $\sqrt{h_k/H_k^3}$ to be our basis $\{\varphi_n\}$ for V_k. The function φ_n is a hat function having nodal value $\sqrt{h_k/H_k^3}$ at the node n and zero values at the remaining nodes. Let $U = (u_1, u_2, \dots)$ be the coefficient vector of a function $u = \sum_n u_n\varphi_n$. With this special basis, we have, for all $u, v \in V_k$, that

$$(u, v)_k = U^\top V, \quad a(u, v) = U^\top A_kV, \quad \text{and} \quad |||u|||_{2,k} = U^\top A_kA_kU,$$

where A_k is the matrix form of the operator A_k. The matrix A_k is symmetric and positive definite, with the (i, j)th element being $a(\varphi_i, \varphi_j)$. By a simple calculation, one can find that the component of A_kU at the index corresponding to node i (see notations in Figure 1) is

$$(A_kU)_i = \rho_kH_k^{-2}(\rho_k^{-1}(2u_i - u_h - u_j) + \rho_k(2u_i - u_c - u_m)), \tag{3.3}$$

where $\rho_k = h_k/H_k$ as defined in (2.2). Summing over all interior nodes of $\mathcal{T}_k$ (the set is denoted by $\mathcal{N}_k$), we get

$$|||u|||_{2,k}^2 = \sum_{i\in\mathcal{N}_k} \rho_k^2H_k^{-4}(\rho_k^{-1}(2u_i - u_h - u_j) + \rho_k(2u_i - u_c - u_m))^2. \tag{3.4}$$

We now show that

$$|||u|||_{2,k}^2 \geq \sum_{i\in\mathcal{N}_k} \rho_k^2H_k^{-4}(\rho_k^{-2}(2u_i - u_h - u_j)^2 + \rho_k^2(2u_i - u_c - u_m)^2). \tag{3.5}$$

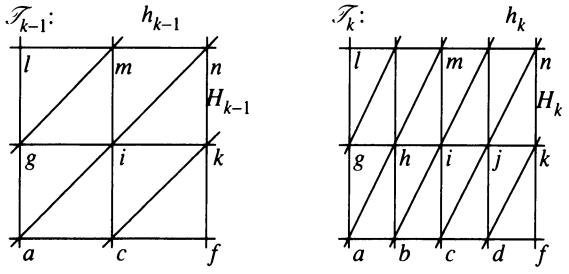

FIGURE 1. Two-level triangulations. The fine grid is finer only in one direction

We order the nodes in $\mathcal{N}_k$ along the x_1-direction first, then along the x_2-direction. Let $M = \text{tridiag}(-\rho_k^{-1}, 2\rho_k^{-1}, -\rho_k^{-1})$, and

$$M_1 = \begin{pmatrix} M & & & \\ & M & & \\ & & \ddots & \\ & & & M \end{pmatrix} \quad \text{and} \quad M_2 = \rho_k \begin{pmatrix} 2I & -I & & \\ -I & 2I & -I & \\ & \ddots & \ddots & \ddots \\ & & -I & 2I \end{pmatrix}, \tag{3.6}$$

where I is an identity matrix. It follows that $A_k = \rho_k H_k^{-2}(M_1 + M_2)$ and

$$\begin{aligned} |||u|||_{2,k}^2 &= U^\top A_k^2 U \\ &= \rho_k^2 H_k^{-4}(U^\top(M_1^2 + M_2^2)U + U^\top(M_1M_2 + M_2M_1)U). \end{aligned} \tag{3.7}$$

To show (3.5), it suffices to show that the symmetric matrix $M_1M_2 + M_2M_1$ is nonnegative semidefinite. Because $M_1M_2 + M_2M_1$ is a block tridiagonal matrix of the form B in Lemma 3.2 with all $B_i = \text{tridiag}(-1, 2, -1)$, inequality (3.5) follows.

Let $v = \mathrm{I}_k\mathrm{I}_{k-1}u$. We can find the coefficient vector V from U. For example, $v_h = (u_g + u_i)/2$ (see Figure 1). Similarly to (3.3), we have

$$\begin{aligned} (A_kV)_i &= \rho_k H_k^{-2}\left(\rho_k^{-1}\frac{2u_i - u_g - u_k}{2} + \rho_k(2u_i - u_c - u_m)\right), \\ (A_kV)_h &= \rho_k H_k^{-2}\left(\rho_k\frac{2u_g - u_a - u_l}{2} + \rho_k\frac{2u_i - u_c - u_m}{2}\right). \end{aligned}$$

Therefore,

$$\begin{aligned} |||\mathrm{I}_k\mathrm{I}_{k-1}u|||_{2,k}^2 &\le \sum_{i\in\mathcal{N}_{k-1}} \frac{\rho_k^2}{H_k^4}\left(\frac{3}{8}\rho_k^{-2}(2u_i - u_g - u_k)^2 + 3\rho_k^2(2u_i - u_c - u_m)^2\right) \\ &\quad + \sum_{h\in\mathcal{M}_{k-1}^1} \frac{\rho_k^2}{H_k^4}\left(\rho_k^2\frac{(2u_g - u_a - u_l)^2}{2} + \rho_k^2\frac{(2u_i - u_c - u_m)^2}{2}\right) \\ &= \sum_{i\in\mathcal{N}_{k-1}} \frac{\rho_k^2}{H_k^4}\left(\frac{3}{8}\rho_k^{-2}(2u_i - u_g - u_k)^2 + 4\rho_k^2(2u_i - u_c - u_m)^2\right), \end{aligned} \tag{3.8}$$

where $\mathcal{M}_{k-1}^1$ is the set of midpoints on the horizontal edges of $\mathcal{T}_{k-1}$ (see Figure 1). Here, in the first sum of (3.8), the inequality $(a+b)^2 \le \frac{3}{2}a^2 + 3b^2$ is applied. To bound $|||\mathrm{I}_k\mathrm{I}_{k-1}u|||_{2,k}$ by $|||u|||_{2,k}$, we insert the entries of U associated with the nodes in $\mathcal{M}_{k-1}^1$ into the sum (3.8) to get

$$(2u_i - u_g - u_k)^2 \le 8(2u_i - u_h - u_j)^2 + 4(2u_h - u_g - u_i)^2 + 4(2u_j - u_i - u_k)^2.$$

From (3.5) and (3.8) we conclude that $|||\mathrm{I}_k\mathrm{I}_{k-1}u|||_{2,k}^2 \le 4|||u|||_{2,k}^2$. □

Lemma 3.4. *For all $v, w \in V_{k-1}$, there holds*

$$|a(v, w) - a(\mathrm{I}_k v, \mathrm{I}_k w)| \le \rho_k H_k |||\mathrm{I}_k v|||_{1,k} |||w|||_{2,k-1}. \tag{3.9}$$

Proof. We consider $a(v, w)$ and $a(\mathrm{I}_k v, \mathrm{I}_k w)$ on a rectangle $[acig]$ from the coarse-level triangulation (see Figure 1). Let $v_{\alpha\beta} = v(\alpha) - v(\beta)$, where $v(\alpha)$ is the nodal value of v at a vertex α. By the linearity of the functions involved,

we have

$$\int_{[acig]} \nabla v \nabla w \, dx = \frac{1}{4\rho_k}(v_{ca}w_{ca} + v_{ig}w_{ig}) + \rho_k(v_{ic}w_{ic} + v_{ga}w_{ga}),$$

$$\int_{[acig]} \nabla \mathrm{I}_k v \nabla \mathrm{I}_k w \, dx = \frac{1}{4\rho_k}(v_{ca}w_{ca} + v_{ig}w_{ig}) + \frac{\rho_k}{2}(v_{ic}w_{ic} + v_{ga}w_{ga}) + \rho_k \frac{v_{ic} + v_{ga}}{2} \frac{w_{ic} + w_{ga}}{2}.$$

Summing over all rectangles in $\mathcal{T}_{k-1}$, we obtain that

$$|||v|||_{1,k-1}^2 = \sum_{\text{all } [acig]} \frac{1}{4\rho_k}(v_{ca}^2 + v_{ig}^2) + \rho_k(v_{ic}^2 + v_{ga}^2), \tag{3.10}$$

$$|||\mathrm{I}_k v|||_{1,k}^2 = \sum_{\text{all } [acig]} \frac{1}{4\rho_k}(v_{ca}^2 + v_{ig}^2) + \rho_k \left(\frac{v_{ic}^2 + v_{ga}^2}{2} + \frac{(v_{ic} + v_{ga})^2}{4} \right), \tag{3.11}$$

and that

$$a(v, w) - a(\mathrm{I}_k v, \mathrm{I}_k w) = \sum_{\text{all } [acig]} \frac{\rho_k}{4}(v_{ic} - v_{ga})(w_{ic} - w_{ga}). \tag{3.12}$$

Combining the term $v_{ic}(w_{ic} - w_{ga})$ from the square $[acig]$ and the term $v_{ic}(w_{ic} - w_{kf})$ from the square $[cfki]$ in the sum of (3.12), we can rewrite it as

$$a(v, w) - a(\mathrm{I}_k v, \mathrm{I}_k w) = \sum_{i \in \mathcal{T}_{k-1}} \frac{\rho_k}{4} v_{ic}(2w_{ic} - w_{ga} - w_{kf}). \tag{3.13}$$

Noting the relation between nodal values and vector representations, for example, $w_i = w(i)\sqrt{H_k^3/(2h_k)}$, we can write down the corresponding form of (3.5) in V_{k-1} as

$$|||w|||_{2,k-1}^2 \geq \sum_{i \in \mathcal{T}_{k-1}} \rho_k^2 H_k^{-4}(\rho_k^{-2}(2w(i) - w(g) - w(k))^2 + 8\rho_k^2(2w(i) - w(c) - w(m))^2).$$

We then apply the Cauchy-Schwarz inequality and (3.10)–(3.13) to get

$$\begin{aligned} &|a(v, w) - a(\mathrm{I}_k v, \mathrm{I}_k w)| \\ &\leq \left\{ \sum_{i \in \mathcal{T}_{k-1}} \frac{\rho_k}{8} v_{ic}^2 \right\}^{1/2} \left\{ \sum_{i \in \mathcal{T}_{k-1}} \frac{\rho_k}{2}(2w_{ic} - w_{ga} - w_{kf})^2 \right\}^{1/2} \\ &\leq \frac{1}{2}|||\mathrm{I}_k v|||_{1,k} \left\{ \sum_{i \in \mathcal{T}_{k-1}} 2\rho_k(2w(i) - w(g) - w(k))^2 \right\}^{1/2} \\ &\leq \rho_k H_k |||\mathrm{I}_k v|||_{1,k} |||w|||_{2,k-1}. \quad \square \end{aligned} \tag{3.14}$$

Comparing (3.10) and (3.11), we can easily get the next corollary.

Corollary 3.5. *There holds*

$$\tfrac{1}{2}|||v|||_{1,k-1} \le |||I_k v|||_{1,k} \le |||v|||_{1,k-1} \quad \forall v \in V_{k-1}. \tag{3.15}$$

Corollary 3.6. *There holds*

$$|||Q_{k-1}w|||_{1,k-1} \le |||w|||_{1,k} \quad \forall w \in V_k. \tag{3.16}$$

Proof. By Corollary 3.5 and the definition of Q_{k-1} in (2.7), we get

$$\begin{aligned} |||Q_{k-1}w|||_{1,k-1} &= \sup_v a(Q_{k-1}w, v) = \sup_v a(w, I_k v) \\ &\le |||w|||_{1,k} \sup_v |||I_k v|||_{1,k} \le |||w|||_{1,k}, \end{aligned}$$

where the supremum is taken over all $v \in V_{k-1}$ with $|||v|||_{1,k-1} = 1$. □

Corollary 3.7. *For all $v, w \in V_{k-1}$, there holds*

$$|a(v, w) - a(I_k v, I_k w)| \le \frac{2\rho_k}{1+\sqrt{2}\rho_k}\frac{2\rho_k}{1+2\rho_k}|||I_k v|||_{1,k}|||w|||_{1,k-1}. \tag{3.17}$$

Proof. By the Cauchy-Schwarz inequality and (3.10)–(3.12), we can show (3.17) similarly as in (3.14). We omit the details. □

4. Constant rate of convergence

We will prove in this section constant convergence rate for the 2-level, and the W-cycle multilevel nonnested multigrid method defined in §2. We start with a few more lemmas.

Lemma 4.1. *For all $u \in V_k$, the following inequalities hold*:

$$|||u - I_k I_{k-1}u|||_{0,k} \le \frac{1}{\sqrt{2}}H_k|||u - I_k I_{k-1}u|||_{1,k}, \tag{4.1}$$

$$|||u - I_k I_{k-1}u|||_{1,k} \le \left(1 + \frac{\rho_k^2}{4}\right)|||u|||_{1,k}. \tag{4.2}$$

Proof. Let $v = u - I_k I_{k-1}u$. Considering $|||v|||^2_{1,k}$ on a rectangle $[bcih]$ from the mesh $\mathcal{T}_k$ (see Figure 1), we have

$$\begin{aligned} \int_{[bcih]} |\nabla v|^2\,dx &= \frac{1}{2\rho_k}((v(i)-v(h))^2 + (v(c)-v(b))^2) \\ &\quad + \frac{\rho_k}{2}((v(i)-v(c))^2 + (v(h)-v(b))^2). \end{aligned} \tag{4.3}$$

Summing (4.3) over all rectangles of $\mathcal{T}_k$, since v is linear on each triangle of $\mathcal{T}_k$ and vanishes at all nodes of $\mathcal{T}_{k-1}$, we can show (4.1) as follows:

$$\begin{aligned} |||v|||^2_{1,k} &= \sum_{h\in\mathcal{M}^1_{k-1}} \left\{\frac{2}{\rho_k}v(h)^2 + \frac{\rho_k}{2}(v(h)-v(b))^2\right\} \\ &\ge \sum_{h\in\mathcal{M}^1_{k-1}} \frac{2}{\rho_k}v(h)^2 = 2H_k^{-2}|||v|||^2_{0,k}, \end{aligned}$$

where $\mathcal{M}^1_{k-1}$ is the set of midpoints of the horizontal edges of $\mathcal{T}_{k-1}$. We remind the readers that $v(h) = u(h) - u(g)/2 - u(i)/2$ at a midpoint $h \in \mathcal{M}^1_{k-1}$ (see

(3.8) and Figure 1). Once again, summing (4.3) over all rectangles of $\mathcal{T}_k$, we get

$$\begin{aligned}
|||v|||_{1,k}^2 &= \sum_{[acig]} \frac{1}{\rho_k}(v(h)^2 + v(b)^2) + \rho_k(v(h) - v(b))^2 \\
&= \sum_{[acig]} \frac{1}{4\rho_k}((2u(h) - u(i) - u(g))^2 + (2u(b) - u(a) - u(c))^2) \\
&\qquad + \frac{\rho_k}{4}((2u(h) - u(i) - u(g)) - (2u(b) - u(a) - u(c)))^2 \\
&\le \sum_{[acig]} \frac{1}{2\rho_k}(u_{hi}^2 + u_{hg}^2 + u_{ba}^2 + u_{bc}^2) \\
&\qquad + \frac{\rho_k}{4}(u_{hi}^2 + u_{hg}^2 + u_{ba}^2 + u_{bc}^2 + 2u_{ag}^2 + 4u_{bh}^2 + 2u_{ci}^2) \\
&\le \left(1 + \frac{\rho_k^2}{2}\right) |||u|||_{1,k}^2,
\end{aligned}$$

where notations like $u_{ab} = u(a) - u(b)$ are used. □

Corollary 4.2. *The following inverse inequality holds*:

$$|||v|||_{1,k} \le 2\sqrt{2}H_k^{-1}|||v|||_{0,k} \quad \forall v \in V_k. \tag{4.4}$$

Proof. Once again, summing (4.3) over all rectangles in $\mathcal{T}_k$, we conclude, by $\rho_k \le 1$, that

$$|||v|||_{1,k}^2 \le \sum_{i \in \mathcal{T}_k} (4\rho_k^{-1} + 4\rho_k)v(i)^2 \le 8H_k^{-2}|||v|||_{0,k}^2. \quad \square$$

Lemma 4.3. *There holds*

$$|||\mathrm{I}_k v|||_{0,k}^2 \le 2|||v|||_{0,k-1} \quad \forall v \in V_{k-1}. \tag{4.5}$$

Proof. By Definition (2.4), we have (see Figure 1)

$$\begin{aligned}
|||\mathrm{I}_k v|||_{0,k}^2 &= \sum_{i \in \mathcal{T}_{k-1}} \frac{H_k^3}{h_k} v(i)^2 + \sum_{h \in \mathcal{M}_{k-1}^1} \frac{H_k^3}{h_k} \left(\frac{v(i) + v(g)}{2}\right)^2 \\
&\le 2 \sum_{i \in \mathcal{T}_{k-1}} \frac{H_k^3}{h_k} v(i)^2 = 4|||v|||_{0,k-1}^2. \quad \square
\end{aligned}$$

We now prove a lemma showing an approximation property of the operator R_{k-1} defined in (2.7). We remark that R_{k-1} is an $a(\cdot, \cdot)$-orthogonal projection operator from the linear vector space V_k to its subspace $\tilde{V}_{k-1}$.

Lemma 4.4. *For all $v \in V_k$, there holds*

$$|||v - \mathrm{R}_{k-1}v|||_{1,k} \le \frac{3}{\sqrt{2}} H_k |||v|||_{2,k}. \tag{4.6}$$

Proof. Noticing $\mathrm{R}_{k-1}v \in V_k$, we can apply the Cauchy-Schwarz inequality to obtain that

$$|||v - \mathrm{R}_{k-1}v|||_{1,k}^2 = a(v - \mathrm{R}_{k-1}v, v) \le |||v - \mathrm{R}_{k-1}v|||_{0,k} |||v|||_{2,k}.$$

Let $w \in V_k$ with $|||w|||_{0,k} = 1$ and $u = A_k^{-1} w$, where the operator A_k is defined before (2.5). By (2.5), $|||u|||_{2,k} = 1$. Since $I_k I_{k-1} u \in \widetilde{V}_{k-1}$, we can insert $I_k I_{k-1} u$ in our estimation to get

$$\begin{aligned}
|||v - R_{k-1}v|||_{0,k} &= \sup_w (v - R_{k-1}v, w)_k = \sup_u a(v - R_{k-1}v, u) \\
&= \sup_u a(v - R_{k-1}v, u - I_k I_{k-1} u) \\
&\le |||v - R_{k-1}v|||_{1,k} \sup_u |||u - I_k I_{k-1} u|||_{1,k} \\
&\le \frac{H_k}{\sqrt{2}} |||v - R_{k-1}v|||_{1,k} \sup_u (|||u - I_k I_{k-1} u|||_{2,k}) \\
&\le \frac{3H_k}{\sqrt{2}} |||v - R_{k-1}v|||_{1,k},
\end{aligned}$$

where the Cauchy-Schwarz inequality, Lemma 4.1, Lemma 3.1, the triangle inequality, and Lemma 3.3 are applied. The assertion (4.6) is proved by combining the above two inequalities. □

Theorem 4.5 (Two-level methods). *Let $q = \overline{q}$ in (2.10). For any $0 < \gamma < 1$, there is an integer m independent of the level number k, such that*

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

where u_k, w_i, q, and $\overline{q}$ are defined in (2.8)–(2.11).

Proof. Let the iteration error be denoted by $e_l = u_k - w_l$, $0 \le l \le m+1$. Our goal is to show $|||e_{m+1}|||_{1,k} \le \gamma |||e_0|||_{1,k}$. By (2.9)–(2.11), (2.7), and the triangle inequality, it follows that

$$|||e_{m+1}|||_{1,k} \le |||e_m - R_{k-1}e_m|||_{1,k} + |||R_{k-1}e_m - I_k Q_{k-1} e_m|||_{1,k}. \tag{4.7}$$

The first term in (4.7) is analyzed in Lemma 4.4. To estimate the second term, by (2.7), the Cauchy-Schwarz inequality, and Lemma 4.3, we have, for any $v \in V_{k-1}$,

$$\begin{aligned}
a(R_{k-1}e_m - I_k Q_{k-1} e_m, I_k v) &= a(e_m, I_k v) - a(Q_{k-1}e_m, Q_{k-1} I_k v) \\
&= a(Q_{k-1}e_m, v - Q_{k-1} I_k v) = a(e_m, I_k(v - Q_{k-1} I_k v)) \\
&\le |||e_m|||_{2,k} |||I_k(v - Q_{k-1} I_k v)|||_{0,k} \\
&\le 2|||e_m|||_{2,k} |||v - Q_{k-1} I_k v|||_{0,k-1}.
\end{aligned} \tag{4.8}$$

Let $w \in V_{k-1}$ with $|||w|||_{0,k-1} = 1$ and $u = A_{k-1}^{-1} w$; then $|||u|||_{2,k-1} = 1$. By (2.7) and (3.9), it follows that

$$\begin{aligned}
(v - Q_{k-1} I_k v, w)_{k-1} &= a(v - Q_{k-1} I_k v, u) \\
&= a(v, u) - a(I_k v, I_k u) \le \rho_k H_k |||I_k v|||_{1,k}.
\end{aligned} \tag{4.9}$$

Taking the supremum over all such w in (4.9), we get $|||v - Q_{k-1} I_k v|||_{0,k-1} \le \rho_k H_k |||I_k v|||_{1,k}$. Noting $(R_{k-1}e_m - I_k Q_{k-1} e_m) \in \widetilde{V}_{k-1} = I_k V_{k-1}$, we have shown by (4.8)–(4.9) that

$$\begin{aligned}
|||R_{k-1}e_m - I_k Q_{k-1} e_m|||_{1,k} &= \sup_v \frac{a(R_{k-1}e_m - I_k Q_{k-1} e_m, I_k v)}{|||I_k v|||_{1,k}} \\
&\le 2\rho_k H_k |||e_m|||_{2,k}.
\end{aligned} \tag{4.10}$$

Using (4.6) and (4.10) in (4.7), we have

$$|||e_{m+1}|||_{1,k} \le (3/\sqrt{2} + 2\rho_k)H_k|||e_m|||_{2,k}.$$

By standard estimates (cf. [5] or the following (4.14)) and Corollary 4.2, we have the following smoothing property for (2.9):

$$|||e_m|||_{2,k}^2 \le (\lambda_k/2m)|||e_0|||_{1,k}^2 \le (4H_k^{-2}/m)|||e_0|||_{1,k}^2.$$

Thus, $|||e_{m+1}|||_{1,k} \le |||e_0|||_{1,k}(3\sqrt{2} + 4\rho_k)/\sqrt{m}$. Therefore, the theorem is proved by choosing $m \ge (3\sqrt{2}+4)^2/\gamma^2 \ge (3\sqrt{2}+4\rho_k)^2/\gamma^2$. □

We remark that the best possible coarse-level corrector $I_k q$ in $\tilde{V}_{k-1}$ for e_m with respect to the $|||\cdot|||_{1,k}$ norm (see (2.10)) is $R_{k-1}e_m$. The second term $|||R_{k-1}e_m - I_k Q_{k-1}e_m|||_{1,k}$ in (4.7) measures the perturbation to the projection in the nonnested multigrid scheme. When the aspect ratio is high, we may get a better estimate of this error by applying (3.17) instead of (3.9) in (4.7). In this way, it follows that

$$\begin{aligned} &|||R_{k-1}e_m - I_k Q_{k-1}e_m|||_{1,k} \\ &\quad \le \frac{2\rho_k}{1+\sqrt{2}\rho_k}\frac{2\rho_k}{1+2\rho_k}|||e_m|||_{1,k} =: C_1|||e_m|||_{1,k}, \end{aligned} \tag{4.11}$$

where $C_1 \sim \rho_k^2$.

In general, the number of smoothings, m, has to be sufficiently large in the nonnested multigrid method (see a numerical example in [24]), owing to the violation of the projection property in the coarse-level correction. But $m = 1$ would guarantee the convergence in the nested multigrid method (see [17]), even for nonsymmetric problems. For our model problem, we have by Corollaries 3.5 and 3.6 that

$$|||I_k Q_{k-1} v|||_{1,k} \le |||v|||_{1,k} \quad \forall v \in V_k. \tag{4.12}$$

This means that the iteration error would not be amplified in the nonnested coarse-level correction. We can apply (4.11)–(4.12) to prove constant-rate convergence for the W-cycle symmetric nonnested multigrid methods with one smoothing. In symmetric multigrid methods, m (post)smoothings are added after the coarse-level correction (2.10) (cf. [8]). That is, one cycle of the level-k multigrid iteration would generate w_{2m+1} from w_0 by (2.9)–(2.11) and doing m smoothings (2.9) on w_{m+1}.

Theorem 4.6 (Symmetric W-cycles with one smoothing). *Let m be a positive integer. Let the aspect ratio ρ_k be small enough such that the constant C_1 in* (4.11) *is less than* $36m/(36+m)^2$. *If* $|||\bar{q} - q|||_{1,k-1} \le \gamma^2|||\bar{q}|||_{1,k-1}$, *then*

$$|||u_k - w_{2m+1}|||_{1,k} \le \gamma|||u_k - w_0|||_{1,k},$$

where $\gamma = 36/(36+m)$, *and* $\bar{q}$, q, u_k, *and* w_i *are defined in* (2.9)–(2.11).

Proof. Let the errors be denoted by $e_l := u_k - w_l$, $0 \le l \le 2m+1$. Let B_k denote the multigrid reduction operator, i.e., $e_{2m+1} := B_k e_0$. Inductively, we can show by (4.12) that B_k is selfadjoint and nonnegative in $(a(\cdot,\cdot), V_k)$ (see (4.13) below) and that

$$B_k = S_k^m(I - I_k(I - B_{k-1}^2)Q_{k-1})S_k^m,$$

TABLE 1. Spectral radii of 2-level multigrid iterative operators for (2.1)

Nested, quasiuniform		Nested, degenerate		Nonnested, degenerate	
grid	ρ(CS)	grid	ρ(CS)	grid	ρ(CS)
10×10	0.7329	10×8	0.7842	10×8	0.6705
14×14	0.7416	14×8	0.8651	14×8	0.6063
18×18	0.7450	18×8	0.9101	18×8	0.5708
22×22	0.7467	22×8	0.9366	22×8	0.5499
26×26	0.7476	26×8	0.9531	26×8	0.5369

where $S_k := I - \lambda_k^{-1} A_k$ is the smoothing operator defined by (2.9). Since S_k is selfadjoint, we have, by the induction hypothesis $|||\overline{q} - q|||_{1,k-1} \le \gamma^2 |||\overline{q}|||_{1,k-1}$ and the stability of the coarse-level corrections (4.12),

(4.13)
$$\begin{aligned} a(B_k e_0, e_0) &= a((I - I_k Q_{k-1}) e_m, e_m) + a(B_{k-1}^2 Q_{k-1} e_m, Q_{k-1} e_m) \\ &\le a((I - I_k Q_{k-1}) e_m, e_m) + \gamma^2 a(Q_{k-1} e_m, Q_{k-1} e_m) \\ &= (1-\gamma^2) a((I - I_k Q_{k-1}) e_m, e_m) + \gamma^2 a(e_m, e_m) \\ &= (1-\gamma^2)(|||e_m - R_{k-1} e_m|||_{1,k}^2 + a((R_{k-1} - I_k Q_{k-1}) e_m, e_m)) \\ &\quad + \gamma^2 |||e_m|||_{1,k}^2 \\ &\le (1-\gamma^2) \frac{3^2}{2} H_k^2 |||e_m|||_{2,k}^2 + (\gamma^2 + (1-\gamma^2) C_1) |||e_m|||_{1,k}^2, \end{aligned}$$

where Lemma 4.2 in [25], (4.6), and (4.11) are used in the last step. By an eigenexpansion as in (3.1) (cf. [5] or [17], for example), it follows by the equality $|(1-x)^{2m}(1+2mx)|_{L^\infty(0,1)} = 1$ that

$$|||S_k^m v|||_{1,k}^2 + 2m\lambda_k^{-1} |||S_k^m v|||_{2,k}^2 \le |||v|||_{1,k}^2 \quad \forall v \in V_k. \tag{4.14}$$

Therefore, by Corollary 4.2,

$$\begin{aligned} a(B_k e_0, e_0) &\le (1-\gamma^2) \frac{3^2 8}{4m} (|||e_0|||_{1,k}^2 - |||e_m|||_{1,k}^2) + (\gamma^2 + C_1) |||e_m|||_{1,k}^2 \\ &\le \gamma (|||e_0|||_{1,k}^2 - |||e_m|||_{1,k}^2) + \gamma |||e_m|||_{1,k}^2 = \gamma |||e_0|||_{1,k}^2. \end{aligned}$$

The proof is completed by applying Lemma 4.2 in [25]. □

To conclude this section, we list the contraction numbers (the spectral radii of the multigrid iteration operators) of the two, nested and nonnested, multigrid methods with one smoothing ($m = 1$ in (2.9)) for the problem (2.8). For the standard multigrid method on uniform meshes, the iteration error is reduced by a factor about 0.75 after each iteration. However, when the meshes tend to degenerate, the factor deteriorates to 1 by the data in the second column of Table 1. The nonnested multigrid method defined via (2.9)–(2.11) provides a constant convergence rate (about 0.5) independent of mesh levels or aspect ratios of the meshes.

5. A PERTURBATION PROBLEM

In this section, we consider an anisotropic, singular perturbation problem. In order to achieve constant rates of convergence in the multigrid method, we

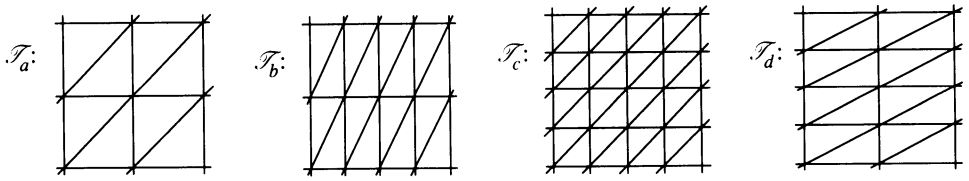

FIGURE 2. Nested meshes, $\mathcal{T}_a \subset \mathcal{T}_c$, and two intermediate meshes, $\mathcal{T}_b$ and $\mathcal{T}_d$

could either choose nested grids with high aspect ratios, or choose nonnested grids where the coarse grids are coarser only in the larger diffusion direction. We will discuss first the case of nested meshes. One purpose of presenting this problem is to indicate a proof of constant convergence rate for the standard multigrid methods on uniform grids without any elliptic regularity assumption. The general multigrid methods on quasiuniform meshes have been shown to yield constant convergence speed by Bramble et al. recently in [9] without the assumption of elliptic regularity.

We first consider piecewise linear finite elements on nested meshes (like $\mathcal{T}_a$ and $\mathcal{T}_c$ in Figure 2) with high aspect ratios to discretize the following problem:

$$\begin{aligned} -\partial_{11}u - \varepsilon\partial_{22}u &= f \quad \text{in } \Omega = (0,1)\times(0,1), \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{5.1}$$

where $0 < \varepsilon \leq 1$. The mesh sizes in the x_1- and x_2-directions are denoted by h_k and H_k, respectively (see Figure 1). We let the sequence of meshes be nested in a standard fashion, i.e., each coarse triangle is refined into four subtriangles by linking midpoints on the three edges. The aspect ratios remain constant on all levels, and we choose it to be

$$H_k/h_k = \sqrt{\varepsilon}, \qquad k = 1, 2, \ldots. \tag{5.2}$$

Unlike the meshes in §2, here we have $h_k = h_{k-1}/2$, $H_k = H_{k-1}/2$, and $H_k \leq h_k$ (just for the purpose of using the same figure). The bilinear form associated with (5.1) is $a(u, v) = \int_\Omega (\partial_1 u \partial_1 v + \varepsilon\partial_2 u \partial_2 v)\, dx$. We replace the discrete inner product (2.4) by

$$(u, v)_k = \sum_{n\in\mathcal{N}_k} H_k h_k u(n)v(n) \quad \forall u, v \in V_k. \tag{5.3}$$

We note that the norm induced by $(\cdot, \cdot)_k$ is now equivalent to the $L^2(\Omega)$ norm in V_k. The other definitions in §2 remain the same. In particular, the finite element problems and the multigrid method are defined as in (2.8)–(2.11). But the interpolation operator I_k is omitted in (2.9)–(2.11) since I_k is now the identity operator.

To analyze the multigrid scheme, we introduce two intermediate meshes ($\mathcal{T}_b$ and $\mathcal{T}_d$ in Figure 2) between two consecutive multilevel methods ($\mathcal{T}_a$ and $\mathcal{T}_c$ in Figure 2), both of which are coarser in one direction than the higher-level mesh and finer in the other direction than the lower-level mesh. We denote the Lagrange interpolation operator based on the grid $\mathcal{T}_b$ and $\mathcal{T}_d$ by I_b and I_d, respectively. We can see immediately that

$$\begin{aligned} &\mathrm{I}_k v = v = \mathrm{I}_k \mathrm{I}_b (\mathrm{I}_k \mathrm{I}_d v) \quad \forall v \in V_{k-1}, \\ &\mathrm{I}_{k-1} v = \mathrm{I}_{k-1} \mathrm{I}_b v \qquad\quad \forall v \in V_k. \end{aligned}$$

We first calculate the equality analogous to (4.3) (see Figure 1 for notation):

$$\int_{[bcih]} ((\partial_1 u)^2 + \varepsilon(\partial_2 u)^2)\,dx = \frac{H_k}{2h_k}((v(i)-v(h))^2 + (v(c)-v(b))^2) + \frac{H_k}{2h_k}((v(i)-v(c))^2 + (v(h)-v(b))^2), \tag{5.4}$$

where (5.2) is used. The idea of utilizing high-aspect ratio meshes can be seen clearly from (5.4). Once we choose the mesh ratio as in (5.2), ε will not appear any more in the analysis (see (5.4) and (5.6) below). The discrete diffusion coefficients in the x_1- and x_2-directions are then the same (see (5.4)). The same argument as used in Lemma 4.1 will show (4.1) again in the new case, by (5.4) and (5.3). We should notice the difference, however, that nonzero nodal values of $v - \mathrm{I}_k\mathrm{I}_{k-1}v = v - \mathrm{I}_{k-1}v$ now occur on all midpoints of coarse-level edges, while they occur only on horizontal edges in Lemma 4.1. Further, comparing (5.8) and (5.4), we can prove (4.4) of Corollary 4.2. It is straightforward to show Lemma 4.3 again. To prove Lemma 3.3 in the new setting, we apply the old Lemma 3.3 twice as follows (see Figure 2):

$$\begin{aligned} |||\mathrm{I}_k\mathrm{I}_{k-1}u|||_{2,k} &= |||\mathrm{I}_c\mathrm{I}_a u|||_{2,k} = |||\mathrm{I}_c\mathrm{I}_b(\mathrm{I}_c\mathrm{I}_d)u|||_{2,k} \\ &\le 2|||\mathrm{I}_c\mathrm{I}_d u|||_{2,k} \le 4|||u|||_{2,k}. \end{aligned} \tag{5.5}$$

We can apply the old version of Lemma 3.3 when interpolating functions between meshes $\mathscr{T}_c$ and $\mathscr{T}_b$ or between meshes $\mathscr{T}_c$ and $\mathscr{T}_d$, although the bilinear forms are different here. In fact, the new version of (3.5) is

$$|||u|||_{2,k}^2 \ge \sum_{i\in\mathscr{N}_k} H_k^{-4}((2u_i - u_h - u_j)^2 + (2u_i - u_c - u_m)^2). \tag{5.6}$$

Therefore, Lemma 3.3 holds when the $\mathrm{I}_k\mathrm{I}_{k-1}$ there is replaced by $\mathrm{I}_k\mathrm{I}_b$ or $\mathrm{I}_k\mathrm{I}_d$ (even the constant 2 in (3.2) is unchanged, since $\rho_k \le 1$ is not used in the proof). Lemma 3.4 holds trivially, since the left-hand side of (3.9) vanishes because of $\mathrm{I}_k v = v \ \ \forall v \in V_{k-1}$. Thus, $\mathrm{R}_{k-1} = \mathrm{I}_k\mathrm{Q}_{k-1}$. Therefore, we obtain that Theorem 4.5 holds and the (nested) multigrid method has a constant rate of convergence for the above problem:

Theorem 5.1 (Nested multilevel W-cycle methods for anisotropic problems). *For nested multigrids with the high aspect ratio defined in* (5.2), *the W-cycle multigrid method defined in* (2.9)–(2.11) *with $p = 2$ converges with a constant rate, γ, which is independent of the grid level k but depends only on the number of smoothings m, when it is applied to solve finite element equations arising from* (5.1).

In particular, letting $\varepsilon = 1$ in (5.1), we proved without any elliptic regularity assumption that the standard multigrid iteration has constant rate of convergence when solving finite element Poisson equations on uniform meshes. We could give the convergence rate quantitatively. Braess and Verfürth studied the rates of a special multigrid method on some uniform meshes (cf. [7, 21]) and did not use the elliptic regularity either.

We now consider solving (5.1) by finite elements on nonnested, but quasiuniform, meshes. Owing to the anisotropic diffusion, the variation of the iteration errors in the x_2-direction would be of low frequency if meshes are quasiuniform. We therefore use meshes like in Figure 1, where coarse meshes are only

coarser in the x_1-direction. Denoting mesh sizes in the x_1- and x_2-directions by h_k and H_k, respectively, we assume that $h_k \sim H_k$. The bilinear form $a(\cdot\,,\,\cdot)$ is again $\int_\Omega(\partial_1 u\partial_1 v + \varepsilon\partial_2 u\partial_2 v)\,dx$. But we do not use (5.3) as the discrete inner product $(\cdot\,,\,\cdot)_k$. The definition of $(\cdot\,,\,\cdot)_k$ in (2.4) and the other definitions in §2 remain the same. We note that the discrete norm $|||\cdot|||_{0,k}$ is here equivalent to the L^2 norm on V_k with the equivalence constants independent of k, since $h_k \sim H_k$.

It is straightforward to prove Lemma 3.3 and Lemma 3.4 in the new circumstance, noting that the only change is to insert ε in the estimation. For example, (3.5) becomes

$$|||u|||_{2,k}^2 \geq \sum_{i\in\mathcal{N}_k} \rho_k^2 H_k^{-4}(\rho_k^{-2}(2u_i - u_h - u_j)^2 + \varepsilon^2\rho_k^2(2u_i - u_c - u_m)^2),$$

where $\rho_k \sim 1$. In fact, (3.9) in Lemma 3.4 is improved to

$$|a(v\,,\,w) - a(\mathrm{I}_k v\,,\,\mathrm{I}_k w)| \leq \varepsilon^{1/2}\rho_k H_k |||\mathrm{I}_k v|||\,|||w|||_{2,k-1}. \tag{5.7}$$

We note that Lemma 3.3 holds no matter in which direction the previous levels are coarsened, and no matter how small ε is. However, Lemma 3.4 may depend on ε. For example, if we coarsen the meshes in the x_2-direction, (5.7) becomes

$$|a(v\,,\,w) - a(\mathrm{I}_k v\,,\,\mathrm{I}_k w)| \leq \varepsilon^{-1/2}\rho_k H_k |||\mathrm{I}_k v|||\,|||w|||_{2,k-1},$$

and the convergence rates of multigrid methods will depend on ε. The key point is to employ coarser meshes which are not coarser in the direction of small diffusion. The necessity of this special coarsening, in order to obtain a convergence rate independent of ε in the multigrid method, can be seen also from (4.1). If we used quasiuniform meshes or nonnested meshes where coarsening is in the x_2-direction, we would have an $\varepsilon^{-1/2}$ on the right of (4.1). Consequently, it appears that the standard nested multilevel, quasiuniform multigrid convergence rate would deteriorate as $\varepsilon \to 0$. This is shown by the data in the second column of Table 2. Corollary 4.2 and Lemma 4.3 can be shown here routinely. Therefore, Lemma 4.4 and Theorem 4.5 hold:

Theorem 5.2 (Two-level nonnested multigrid methods). *Let the multilevel meshes and multigrid methods be defined as in* §2 *and above, and let* $0 < \gamma < 1$. *There is an integer* m *independent of* k *or* $\varepsilon \in (0, 1)$ *such that*

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

where $\overline{q} = q$ *in* (2.10).

TABLE 2. Spectral radii of 2-level multigrid iterative operators for (5.1) where $\varepsilon = (8h_2)^2$. Here, h_2 is the mesh size in the x_2-direction

Nested, quasiuniform		Nested, degenerate		Nonnested, quasiuniform	
grid	ρ(CS)	grid	ρ(CS)	grid	ρ(CS)
10 × 10	0.7851	8 × 10	0.7312	10 × 10	0.6777
14 × 14	0.8658	8 × 14	0.7386	14 × 14	0.6152
18 × 18	0.9105	8 × 18	0.7414	18 × 18	0.5780
22 × 22	0.9368	8 × 22	0.7428	22 × 22	0.5555
26 × 26	0.9531	8 × 26	0.7435	26 × 26	0.5302

Finally, we list the convergence rates for two-level multigrid methods with one smoothing ($m = 1$ in (2.9)) in Table 2. From column 2 of Table 2, the standard nested multigrid methods on quasiuniform meshes clearly do not provide a convergence rate independent of ε. Either by employing high-aspect-ratio meshes, or specially coarsened quasiuniform meshes, the multigrid method can retain a constant convergence rate for all ε (column 4 and column 6 in Table 2).

6. An interface problem

In this section, we consider a simple model of interface problem. For the model problem, if mesh lines of finite elements can cover the interface line completely, the finite element method would yield the optimal order of approximation, $O(h)$, in the $H^1(\Omega)$-norm (see [2], Figure 3(B), and column 3 of Table 2). From Figure 3(B) the presence of the interface can hardly be seen from the nodal errors of the finite element solution. We will show, in this case, that the multigrid method on uniform, nested meshes can solve the finite element system in the optimal order of operations. In fact, it is shown that the rate of convergence of the multigrid method is independent of the jump at the interface. The standard multigrid theorem, for example, of Bank and Dupont [5] does not provide such an independence. Bramble et al. proved in [9] that the convergence rate for the multigrid method is independent of the jump for general interface problems, but the convergence rate depends on the level number. Some preliminary work on this subject can be found in [22]. Readers can find the multigrid method for solving finite difference interface equations in [1] and [11], where special treatment is needed in defining the finite difference scheme and intergrid transfer operators, while these problems are naturally resolved in the finite element method.

If the interface does not match any grid line completely (a practical situation), the finite element method can provide only order $O(h^{1/2})$ approximability in the $H^1(\Omega)$-norm (see Figure 3(C) and the data in column 5 of Table 3), since the PDE solution $u \in H^{3/2}(\Omega)$. One can see in Figure 3(C) that the error of a finite element solution is large near the interface. We then consider a model case where the grids (cf. Figure 3(E′) and 3(F′)), with high aspect ratios in a local region of interface, are aligned along the interface. With such grids, it is shown numerically that the convergence order can be brought back to $O(h)$ in the $H^1(\Omega)$-norm. Here, $h = N^{-1/2}$ is a pseudo mesh size, where N is the number of grid points. We will show that the two-level multigrid method converges at a speed independent of the number of unknowns and the interface jump, if the special coarsening method defined in §2 is used and a weighted finite-level smoothing is adopted.

Let $\Omega_1 = (0, 1/2)\times(0, 1)$ and $\Omega_2 = (1/2, 1)\times(0, 1)$ be the two subdomains of $\Omega = (0, 1) \times (0, 1)$ cut by a vertical line $\Gamma = \{(x_1, x_2) \in \Omega | x_1 = 1/2\}$. Let $a(x_1, x_2)$ be a discontinuous function which has two positive, constant values on the two subdomains. We consider the following interface problem:

$$
\begin{aligned}
-a\Delta u &= f \quad \text{in } \Omega_1 \cup \Omega_2, \\
u &= 0 \quad \text{on } \partial\Omega, \\
[u] = [a\, du/d\mathbf{n}] &= 0 \quad \text{across } \Gamma,
\end{aligned}
\tag{6.1}
$$

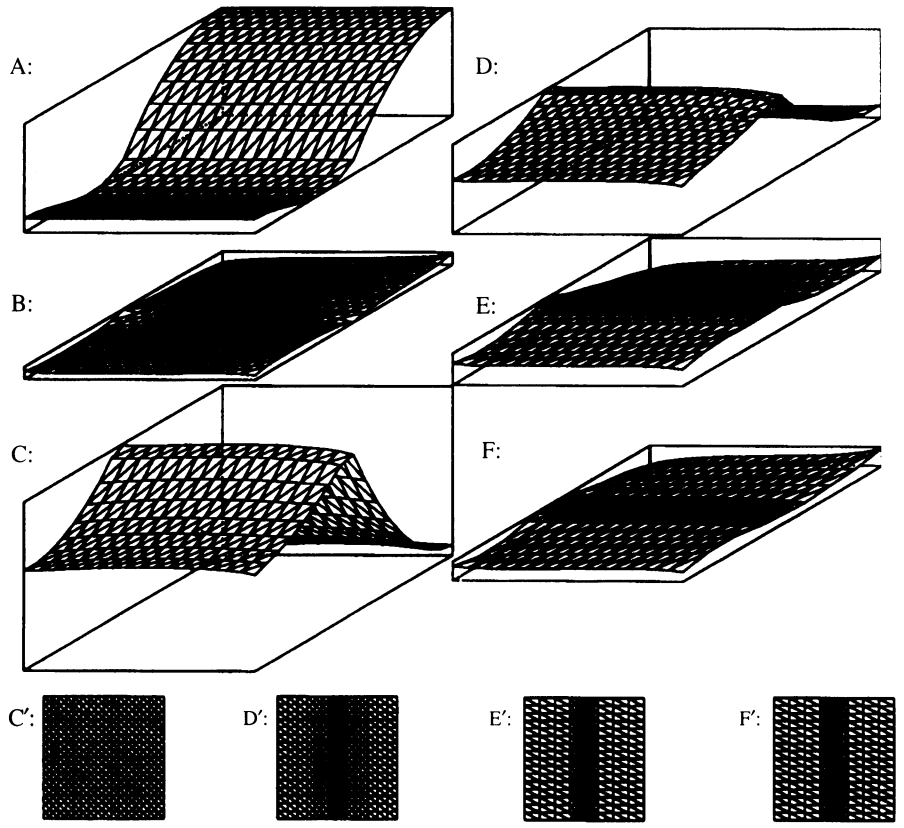

FIGURE 3. (A) Finite element solution of (6.1) on mesh (C');
(B) Nodal error on mesh (C') when interface matches a grid line;
(C) Error on (C'), the interface does not match any grid line;
(D) Error on (D') (no match). The mesh is mapped by $x^{1.2}$;
(E) Error on (E') (no match). The mesh is mapped by $x^{1.6}$;
(F) Error on (F') (no match). The mesh is mapped by x^2.

where $[v]$ denotes the jump in a quantity v across the interface Γ, and $\mathbf{n}$ is the normal to the interface. Without loss of generality, we can scale (6.1) such that

$$a(x_1, x_2) = a > 0 \quad \text{in } \Omega_1 \qquad \text{and} \qquad a(x_1, x_2) = 1 \quad \text{in } \Omega_2.$$

Since there are no cross interfaces and the interface intersects the boundary at a right angle, the solution u of (6.1) is an H^2 function in Ω_1 and Ω_2 separately, if $f \in L^2(\Omega)$ (cf. [15] and [2]). Therefore, if the two subdomains are triangulated separately, then the optimal convergence order would be obtained for the finite element methods in solving (6.1). We will show that the multigrid methods have constant convergence rate (independent of the jump of $a(x)$ across Γ) for solving such finite element equations.

We let $\{\mathcal{T}_k\}$ be nested, uniform meshes like $\mathcal{T}_c$ in Figure 2. To cover the interface by grid lines, the mesh size of $\mathcal{T}_k$ is set to $h_k = 1/(i_0 2^k)$ for some positive integer i_0. Let V_k be the space of continuous, piecewise linear functions defined on $\mathcal{T}_k$. The bilinear form associated with (6.1) is $a(u, v) =$

$\int_\Omega a(x)\nabla u\cdot\nabla v\,dx$. The discrete bilinear form (2.4) is replaced by

$$(u,v)_k=\sum_{n\in\mathcal{N}_k}h_k^2\tilde{a}(n)u(n)v(n)\quad\forall u,v\in V_k,\tag{6.2}$$

where

$$\tilde{a}(n)=\begin{cases}a & \text{if } n\in\Omega_1,\\ (a+1)/2 & \text{if } n\in\Gamma,\\ 1 & \text{if } n\in\Omega_2.\end{cases}$$

It is necessary to have $a(x)$ involved in $(\cdot,\cdot)_k$ in order to have the convergence rate independent of $a(x)$. This can be seen from the smoothing iteration (2.9). The discrete norms $|||\cdot|||_{s,k}$ and the multigrid algorithm are defined as in §2. We note $\mathrm{I}_k v=v\ \ \forall v\in V_{k-1}$, since we have $V_{k-1}\subset V_k$ here.

Similarly to the nested case considered in §5, we introduce two auxiliary grids $\mathcal{T}_b$ and $\mathcal{T}_d$ (see Figure 2) to link $\mathcal{T}_{k-1}$ and $\mathcal{T}_k$. However, the perturbation on the diffusion coefficient this time occurs only in a part of the domain, which requires special consideration. The key issue is to prove (3.5), which is implied by $U^\top(M_1M_2+M_2M_1)U\geq 0$ (see the proof of (3.5)). We will keep M_2 in the same form as in (3.6) by scaling the basis appropriately as follows. Let $\{\varphi_n\}\subset V_k$ be the nodal basis with nodal value $(\tilde{a}(n)h_k^2)^{-1/2}$, where the nodes $\mathcal{N}_k$ are listed along the x_1-direction first, and then along the other direction, as in §5. Let $u=\sum u_n\varphi_n\in V_k$ and U be its coefficient vector. We have by (6.2)

$$(u,v)_k=V^\top U,\qquad (a(u,\varphi_n))=h_k^{-2}A_3^{-1/2}(A_1+A_2)A_3^{-1/2}U.$$

Here $A_1=\mathrm{diag}(D_1)$, a block diagonal matrix, $A_2=\mathrm{tridiag}(-D_2,2D_2,-D_2)$, a block tridiagonal matrix, and $A_3=\mathrm{diag}(D_2)$, a block diagonal matrix, where D_1 and D_2 are respectively

(6.3)
$$\begin{pmatrix}2a & -a & & & & & \\ \ddots & \ddots & \ddots & & & & \\ & -a & 2a & -a & & & \\ & & -a & a+1 & -1 & & \\ & & & -1 & 2 & -1 & \\ & & & & \ddots & \ddots & \ddots\end{pmatrix},\qquad \begin{pmatrix}a & & & & & & \\ & \ddots & & & & & \\ & & a & & & & \\ & & & \frac{a+1}{2} & & & \\ & & & & 1 & & \\ & & & & & \ddots & \end{pmatrix}.$$

Denote the semipositive definite matrix $A_3^{-1/2}A_1A_3^{-1/2}$ by $\widetilde{M}_1$; then, if $w=\mathrm{A}_k u$, its coefficient vector is $W=h_k^{-2}(\widetilde{M}_1+M_2)U$. The matrix M_2 is block tridiagonal of the form $(-I,2I,-I)$ (see (3.6)). Here the operator A_k is again defined by $(\mathrm{A}_k u,v)_k=a(u,v)$. We thus obtain the corresponding form of (3.7) as

$$\begin{aligned}|||u|||_{2,k}=(\mathrm{A}_k u,\mathrm{A}_k u)_k=W^\top W&=\frac{1}{h_k^4}U^\top(\widetilde{M}_1+M_2)(\widetilde{M}_1+M_2)U\\ &\geq\frac{1}{h_k^4}U^\top(\widetilde{M}_1^2+M_2^2)U,\end{aligned}\tag{6.4}$$

by Lemma 3.2. The rest of the proof in Lemma 3.3 remains the same. We

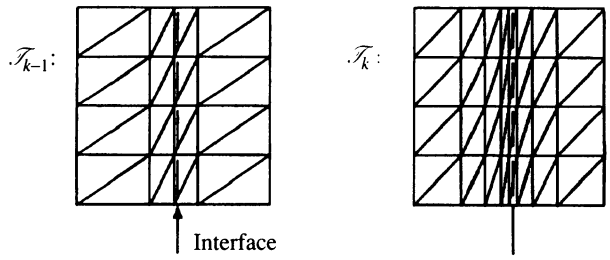

FIGURE 4. Two level grids for (6.1) where the jump is not on a grid line

certainly need to repeat the arguments used in the first case of §5. Since the diffusion coefficient $a(x)$ is added to the discrete bilinear forms $(\cdot\,,\,\cdot)_k$, Lemma 4.1 and Corollary 4.2 follow immediately.

Theorem 6.1 (Multilevel W-cycles for the interface problem). *Let the multigrid method be defined as above and in* §2. *Let* $p = 2$ *in the definition of* q *in* (2.10). *For any* $0 < \gamma < 1$, *there is an integer* m *independent of the level number* k *and the* $a(x)$ *in* (6.1) *such that*

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

where u_k, w_i, *and* q *are defined in* (2.8)–(2.11).

We next consider the case where the interface is between grid lines and the grids are of high aspect ratios near the interface (see Figure 4). This would be somewhat practical because, for example, the interfaces may move with time in many application problems. As we have seen in Figure 3(C), the error needs to be reduced near Γ. One way to increase the approximability of finite elements is to use augmented finite elements (cf. [4] and [19]), which will not be discussed here. One could also use small triangles near the interface line, but this is not economical. A mesh finer in the cross-interface direction would increase the approximability of finite element solutions in this particular case. Let us check the data in Table 3 (next page). Here the exact solution is shown in Figure 3(A). Column 7 and column 9 in Table 3 indicate that the degenerate-mesh finite element solutions have the optimal order (in terms of the number of unknowns) of convergence. In column 3 of Table 3, we have uniform grids where the interface matches the center grid line. The nodal errors of the three cases can be compared in Figure 3(E), Figure 3(F), and Figure 3(B).

As the minimal angle approaches zero, convergence of the standard multigrid method would be arbitrarily slow (see column 4 in Table 1). The essential factor for the deterioration is that coarse-level corrections can no longer capture the low-frequency components of iteration errors which oscillate in the direction of larger grid width. In the model case, a simple but effective modification would be to coarsen the meshes only in the x_1-direction (see Figure 4). The analysis of this multigrid method is almost the same as the previous one in this section.

Let V_k be the space of piecewise linear, continuous functions on $\mathcal{T}_k$ (see Figure 4). We list the nodes of $\mathcal{T}_k$, $\{n_{i,j}\}$, along the x_1-direction first, then

TABLE 3. Spectral radii of 2-level multigrid iterative operators for (5.1) where $\varepsilon = (8h_2)^2$. Here, h_2 is the mesh size in the x_2-direction. The x_1-coordinates of grid points are mapped by the mapping $(x_1 - 0.5)^\alpha$ toward the interface Γ. The β in $|||u - u_k|||_{1,k}^2 = N_k^{-\beta/2}$ is estimated by the errors on two consecutive levels

	Matched, $\alpha = 1$		Unmatched, $\alpha = 1$, 1.5 and 2					
grid	$\|\|\|u_k\|\|\|_{1,k}^2$	β	$\|\|\|u_k\|\|\|_{1,k}^2$	β	$\|\|\|u_k\|\|\|_{1,k}^2$	β	$\|\|\|u_k\|\|\|_{1,k}^2$	β
1×1	0.09375	-	0.07291	-	0.07291	-	0.2791	-
2×2	0.12500	0.55	0.12500	-	0.12500	-	0.12500	-
4×4	0.17187	1.43	0.17122	1.4	0.17315	1.5	0.15808	0.8
8×8	0.19238	1.91	0.19073	1.6	0.19019	1.5	0.18582	1.6
16×16	0.19805	1.97	0.19653	1.4	0.19708	1.8	0.19560	1.7
32×32	0.19951	1.99	0.19851	1.2	0.19917	1.8	0.19884	1.9
64×64	0.19987	1.99	0.19930	1.2	0.19976	1.8	0.19969	1.9
128×128	0.19997	1.99	0.19966	1.1	0.19993	1.9	0.19992	2.1

those along the other direction. Let

(6.5)
$$\tilde{A} = \begin{pmatrix} \ddots & & & & \\ & a & & & \\ & & \tilde{a} & & \\ & & & 1 & \\ & & & & \ddots \end{pmatrix},$$

$$D_1 = \begin{pmatrix} \alpha_1 + \alpha_2 & -\alpha_2 & & & \\ -\alpha_2 & \alpha_2 + \alpha_3 & -\alpha_3 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \alpha_{n-1} + \alpha_n & \end{pmatrix}, \quad D_2 = \begin{pmatrix} \beta_1 & & & \\ & \beta_2 & & \\ & & \ddots & \\ & & & \beta_{n-1} \end{pmatrix},$$

where

$$\alpha_i = \frac{H_k}{h_{h,i}}, \qquad \beta_i = \frac{h_{k,i} + h_{k,i+1}}{2H_k}, \qquad h_{k,i} = \operatorname{dist}(n_{i,j}, n_{i-1,j}),$$

and H_k is the uniform grid width in the x_2-direction. The $\tilde{a}$ in (6.5) depends on the location of the interface between two nearest vertical grid lines (see (6.2) and Figure 4). Let the coefficient vector of $u = \sum u_{i,j}\varphi_{i,j}$ be U. Let $(A_1 + A_2)U := (a(u, \varphi_{i,j}))$, where $\varphi_{i,j}$ denotes the standard nodal basis function at $n_{i,j}$. Then $A_1 = \operatorname{diag}(\tilde{A}D_1)$, a block diagonal matrix, and $A_2 = \operatorname{tridiag}(-\tilde{A}D_2, 2\tilde{A}D_2, -\tilde{A}D_2)$, a block tridiagonal matrix.

The bilinear form associated with (6.1) is again $a(u, v) = \int_\Omega a\nabla u \cdot \nabla v \, dx$. But the discrete inner product (6.2) is modified to, according to (6.5) and reflecting the β_i,

$$(u, v)_k = \sum_{n_{i,j} \in \mathcal{N}_k} \tilde{a}(n_{i,j}) \frac{H_k^2}{\beta_i} u(n_{i,j}) v(n_{i,j}) \quad \forall u, v \in V_k, \tag{6.6}$$

where $\tilde{a}(n_{i,j})$ is a diagonal element of $\tilde{A}$ defined in (6.5). We assume

$$h_{k,i} \le H_k \quad \text{and} \quad \underline{c} < \frac{h_{k,i}}{h_{k,i+1}} < \bar{c} \quad \forall n_{i,j} \in \mathcal{N}_k . \tag{6.7}$$

Here we do not allow a sudden change of the x_1-width. As in (6.4), if we scale the nodal basis by $(\tilde{a}(n_{i,j})H_k^2/\beta_i)^{-1/2}$, we again get by Lemma 3.2 the corresponding form of (3.6),

$$|||u|||_{2,k} = \frac{1}{H_k^4} U^\top (\tilde{\tilde{M}}_1 + M_2)(\tilde{\tilde{M}}_1 + M_2)U \ge \frac{1}{H_k^4} U^\top (\tilde{\tilde{M}}_1^2 + M_2^2)U ,$$

where $M_2 = \text{tridiag}(-I, 2I, -I)$ as before, and $\tilde{\tilde{M}}_1$ is a positive semidefinite, block tridiagonal matrix. It is essential in this analysis to get $M_2 = \text{tridiag}(-I, 2I, -I)$ by a careful scaling, in order to show (3.5). Therefore, by the slow variance in the triangle size (6.7), we can repeat the proof in §§3–4 and obtain the following theorem.

Theorem 6.2 (2-level method for meshes with skinny triangles near the interface). *Let the multigrid method be defined as above and in* §2. *Let* $\bar{q} = q$ *in* (2.10). *For any* $0 < \gamma < 1$, *there is an integer* m *independent of the level number* k *and the* $a(x)$ *in* (6.1) *such that*

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

where u_k, w_i, q, *and* $\bar{q}$ *are defined in* (2.8)–(2.11).

One may include the refinement method for nested meshes here. Near the interface, where skinny triangles occur, one can use those coarser triangles which are coarser in one direction. Away from the interface, one can use the nested grids (each coarse triangle consists of four subtriangles). Between the two regions, one can use the technique of the nonnested multigrid method in [24] to get a smooth transition.

7. A CONVECTION-DOMINATED CONVECTION-DIFFUSION PROBLEM

We will study two finite element methods for solving a model convection-dominated convection-diffusion problem. To dominate the convection term numerically, we choose the finite elements in such a way that the mesh size in the direction of convection is smaller than ε, the diffusion coefficient. For economical reasons, the mesh size in the crosswind direction is not that small. This results in high-aspect-ratio meshes. Such finite elements avoid the wiggles of the Galerkin method with quasiuniform finite elements (see Figure 5 on next page). Three finite element solutions for the perturbation problem (7.1) with $f = 1$ are shown in Figure 5. Clearly, the solution on the high-aspect-ratio grid is much better. Since the compensation of the streamline diffusion method to the right-hand side does not change the source vector (see (7.12)) in this test case, the streamline diffusion method just adds an artificial diffusion in the streamline direction. For references of the streamline diffusion method, readers can consult [14] and the references therein. In this section, we will show that a nonnested multigrid iteration has a constant convergence rate which does not depend on the perturbation parameter, nor on the grid size and the aspect ratio. We will show briefly that the special nonnested multigrid method would also

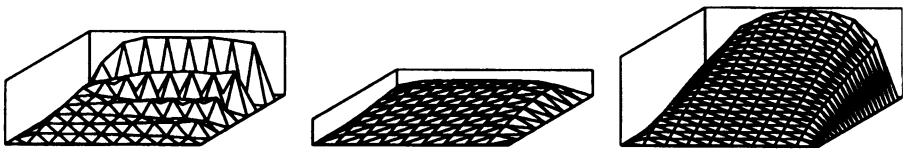

FIGURE 5. Finite element solutions for (7.1). Left, the Galerkin method on a uniform mesh; Center, the streamline diffusion method on a uniform mesh; Right, the Galerkin method on a high-aspect-ratio mesh

converge at a constant rate when applied with the streamline diffusion method. A purpose of this section is to show that the coarse-level correction can catch slowly varying components of the iteration errors even in very nonsymmetric, ill-conditioned linear systems.

We will consider a steady-state convection-dominated convection-diffusion problem:

$$\begin{aligned} -\varepsilon\Delta u + \partial_1 u &= f \quad \text{in } \Omega = (0,1)\times(0,1), \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{7.1}$$

where ε is a positive constant. We define finite element spaces $\{V_k\}$ as in §2, where triangular meshes have grid widths H_k and h_k in the x_2- and x_1-directions, respectively. We require that

$$h_k = \delta_k \varepsilon \quad \text{and} \quad h_k = \rho_k H_k \quad \text{for some } \delta_k \le 1 \text{ and } \rho_k \le 1, \tag{7.2}$$

and that the coarse-level meshes are only coarser in the x_1-direction, as shown in Figure 1, i.e., $h_k = h_{k-1}/2$, $H_k = H_{k-1}$. Similarly to the principle of the streamline diffusion method, choosing a small grid width in the convection direction could control the nonsymmetric term $\partial_1 u$ to result in an M-matrix for the algebraic system. Owing to the structure of the crosswind boundary layer, one may also want $H_k \le \sqrt{\varepsilon}$. If one also requires $H_k \le \varepsilon$, one could prove convergence of the multigrid method by the elliptic regularity (cf. [16] for example, and cf. [20] for the symmetric case) and by standard approximation theory [10]. However, the convergence rate of the multigrid method obtained in this way would depend on the singular perturbation ε, while our theory avoids such a dependence.

We remark that the effect of employing a small grid width in the x_1-direction resembles that of choosing a grid width smaller than ε in the following 1-D problem:

$$\varepsilon u'' + u' = f \quad \text{in } \Omega = (0,1), \qquad u(0) = u(1) = 0.$$

The central finite difference method and the linear finite element method produce the same stiffness matrix for the 1-D problem above. In [11] and [12], W. Hackbusch has shown for the 1-D problem that the contraction number of two-level multigrid methods for the centered finite difference problems is bounded away from 1 independent of the grid width if the grid width is smaller than ε. An economical way to extend Hackbusch's method to 2-D problems would be the one described above. However, if the perturbation parameter is too small, making the grid size smaller than ε would not be practical. One may use the streamline diffusion method then. A mesh with $h_k \le \varepsilon$ and $H_k \le \sqrt{\varepsilon}$ would

be effective in computation as it provides both stability and resolution of the boundary layer. A multigrid method was studied in [23] for solving Galerkin spectral equations arising from (7.1) with periodic boundary conditions.

The finite element formulation of (7.1) reads as follows: Find u_k such that

$$\varepsilon a(u_k, v) + b(u_k, v) = (f, v) \quad \forall v \in V_k, \tag{7.3}$$

where $a(u, v) = \int_\Omega \nabla u \cdot \nabla v \, dx$, $b(u, v) = \int_\Omega \partial_1 uv \, dx$, and $(u, v) = \int_\Omega uv \, dx$. The discrete inner product $(\cdot, \cdot)_k$ is again defined in (2.4). The multigrid method consists of two steps (cf. (2.9) and (2.10)). The fine-level smoothing will be defined later. The coarse-level residual problem (cf. (2.11)) is now defined as: Find $\overline{q} \in \widetilde{V}_{k-1}$ such that

$$\varepsilon a(\overline{q}, v) + b(\overline{q}, v) = (f, v) - \varepsilon a(w_m, v) - b(w_m, v) \quad \forall v \in \widetilde{V}_{k-1}, \tag{7.4}$$

where $\widetilde{V}_{k-1} \subset V_k$ is the image space $\mathrm{I}_k(V_{k-1})$ as in (2.6). In a two-level method, the corrected iterative solution (cf. (2.10)) is $w_{m+1} = w_m + \overline{q}$. As before, with the iteration error denoted by $e_j = u_k - w_j$, the two-level coarse-level correction generates $e_{m+1} = e_m - \overline{q}$, where $\overline{q}$ is defined by (7.4) and satisfies

$$\varepsilon a(\overline{q}, v) + b(\overline{q}, v) = \varepsilon a(e_m, v) + b(e_m, v) \quad \forall v \in \widetilde{V}_{k-1}. \tag{7.5}$$

Lemma 7.1. *The following estimates hold*:

$$|||e_{m+1}|||_{1,k} \le CH_k(1 + \delta_k)|||e_m|||_{2,k}, \tag{7.6}$$

$$|||e_{m+1}|||_{1,k} \le (1 + \rho_k^2/4 + C\delta_k)|||e_m|||_{1,k}, \tag{7.7}$$

where the constant C is independent of ε and k. Here the discrete norms $|||\cdot|||_{\cdot,k}$ are defined in (2.5), *and δ_k and ρ_k are defined in* (7.2).

Proof. From $b(u, v) = 0$ and (7.5), we get

$$\begin{aligned} \varepsilon|||e_{m+1}|||_{1,k}^2 &= \varepsilon|||e_m - \overline{q}|||_{1,k}^2 = \varepsilon a(e_m - \overline{q}, e_m - \overline{q}) + b(e_m - \overline{q}, e_m - \overline{q}) \\ &= \varepsilon a(e_m - \overline{q}, e_m - \mathrm{I}_k\mathrm{I}_{k-1}e_m) + b(e_m - \overline{q}, e_m - \mathrm{I}_k\mathrm{I}_{k-1}e_m). \end{aligned} \tag{7.8}$$

By the triangle inequality, the first term in (7.8) is

$$a(e_m - \overline{q}, e_m - \mathrm{I}_k\mathrm{I}_{k-1}e_m) \le |||e_m - \overline{q}|||_{1,k}|||e_m - \mathrm{I}_k\mathrm{I}_{k-1}e_m|||_{1,k}. \tag{7.9}$$

By (4.1), Lemma 3.1, and Lemma 3.3, we get

$$|||e_m - \mathrm{I}_k\mathrm{I}_{k-1}e_m|||_{1,k} \le CH_k|||e_m|||_{2,k}. \tag{7.10}$$

By the definition of $b(\cdot, \cdot)$, we can estimate the second term in (7.8) as

$$\begin{aligned} |b(e_m - \overline{q}, \mathrm{I}_k\mathrm{I}_{k-1}e_m - e_m)| &\le |e_m - \overline{q}|_{H^1(\Omega)}\|\mathrm{I}_k\mathrm{I}_{k-1}e_m - e_m\|_{L^2(\Omega)} \\ &\le |||e_m - \overline{q}|||_{1,k}Ch_kH_k^{-1}|||e_m - \mathrm{I}_k\mathrm{I}_{k-1}e_m|||_{0,k}, \end{aligned} \tag{7.11}$$

where the following fact is used: $\|u\|_{L^2(\Omega)} \le Ch_kH_k^{-1}|||u|||_{0,k}$ for all $u \in V_k$, which can be proved easily by the definition (2.4). Applying (4.1) again, we conclude from (7.10) that

$$|||e_m - \mathrm{I}_k\mathrm{I}_{k-1}e_m|||_{0,k} \le CH_k|||e_m - \mathrm{I}_k\mathrm{I}_{k-1}e_m|||_{1,k} \le CH_k^2|||e_m|||_{2,k}. \tag{7.12}$$

Combining (7.8)–(7.12), we obtain (7.6). Applying definition (4.2) to the term $|||e_m\mathrm{I}_k\mathrm{I}_{k-1}e_m|||_{1,k}$ in (7.9) and (7.12), we can also prove (7.7) by (7.8)–(7.12). □

To obtain a constant rate of convergence, by (7.6), the fine-level smoothing needs to provide that

$$|||e_m|||_{2,k} \leq CH_k^{-1}|||e_0|||_{1,k}. \tag{7.13}$$

If one performs the fine-level smoothing on the normal system to deal with nonsymmetry as in [6], for example, (7.13) could be obtained, but the constant in (7.13) would depend on ε. In the following, we consider a multigrid method with one fine-level smoothing. J. Mandel [17] has considered such a multigrid method for nonsymmetric elliptic problems, but where the $b(\cdot, \cdot)$ is only a small perturbation to the system, since the mesh is required to be fine enough. We note that the $b(\cdot, \cdot)$ and the $\varepsilon a(\cdot, \cdot)$ are of the same order of magnitude for the present case. One iteration as follows will be performed in the fine-level smoothing (cf. (2.9)):

$$(w_1 - w_0, v)_k = \varepsilon^{-1}\lambda_k^{-1}((f, v) - \varepsilon a(w_0, v) - b(w_0, v)) \quad \forall v \in V_k, \tag{7.14}$$

where λ_k^{-1} ($\leq CH_k^2$, by Corollary 4.2) is as in (2.9). The error after one smoothing is

$$e_1 = (\mathrm{I} - \lambda_k^{-1}\mathrm{A}_k - \varepsilon^{-1}\lambda_k^{-1}\mathrm{B}_k)e_0, \tag{7.15}$$

where $(\mathrm{A}_k u, v)_k := a(u, v)$ and $(\mathrm{B}_k u, v)_k := b(u, v)$. By expanding e_0 as a linear combination of eigenvectors of A_k, it is standard to get (cf. (4.14) and [17])

$$|||(\mathrm{I} - \lambda_k^{-1}\mathrm{A}_k)e_0|||_{1,k}^2 + \lambda_k^{-1}|||(\mathrm{I} - \lambda_k^{-1}\mathrm{A}_k)e_0|||_{2,k}^2 \leq |||e_0|||_{1,k}^2. \tag{7.16}$$

To estimate the effects of $b(\cdot, \cdot)$ to the smoothing (7.14), we can proceed as follows:

$$\begin{aligned}
|||\mathrm{B}_k e_0|||_{j,k} &= \sup_{v \in V_k, |||v|||_{j,k}=1} (\mathrm{B}_k e_0, \mathrm{A}_k^j v)_k = \sup_{|||v|||_{j,k}=1} b(e_0, \mathrm{A}_k^j v) \\
&\leq \sup_{|||v|||_{j,k}=1} \|\partial_1 e_0\|_{L^2(\Omega)} \|\mathrm{A}_k^j v\|_{L^2(\Omega)} \\
&\leq \sup_{|||v|||_{j,k}=1} |||e_0|||_{1,k} C h_k H_k^{-1} |||\mathrm{A}_k^j v|||_{0,k} \\
&\leq C h_k H_k^{-1-j} \sup_{|||v|||_{j,k}=1} |||e_0|||_{1,k} |||\mathrm{A}_k^{j/2} v|||_{0,k} \\
&= C h_k H_k^{-1-j} |||e_0|||_{1,k}, \qquad j = 1, 2.
\end{aligned}$$

Therefore, by (7.15), one gets

$$\begin{aligned}
|||e_1|||_{j,k}^2 &\leq (|||(\mathrm{I} - \lambda_k^{-1}\mathrm{A}_k)e_0|||_{j,k} + \varepsilon^{-1}\lambda_k^{-1}|||\mathrm{B}_k e_0|||_{j,k})^2 \\
&\leq |||(\mathrm{I} - \lambda_k^{-1}\mathrm{A}_k)e_0|||_{j,k}^2 + D_j |||e_0|||_{1,k}^2,
\end{aligned} \tag{7.17}$$

where $D_j = 2(C\delta_k H_k^{1-j}) + (C\delta_k H_k^{1-j})^2$, $j = 1, 2$.

Theorem 7.2. *The contraction number of the two-level multigrid method defined by* (7.4) *and* (7.14) *is bounded by* γ *defined by*

$$\gamma^2 = \frac{D_0 + C\rho_k^2 + C\delta_k}{D_0 + 1 + \rho_k^2/4 + C\delta_k}, \tag{7.18}$$

where ρ_k *and* δ_k *are as in* (7.2), C *as in* (7.6), *and* $D_0 = C^2H_k^2(1+\delta_k)^2\lambda_k$. *In fact,* $D_0 \leq 4$.

Proof. We need to show that $|||e_2|||_{1,k} \leq \gamma|||e_0|||_{1,k}$. By (7.17) and (7.16),

$$
\begin{aligned}
|||e_1|||_{2,k}^2 &\leq |||(I-\lambda_k^{-1}A_k)e_0|||_{2,k}^2 + D_2|||e_0|||_{1,k}^2 \\
&\leq \lambda_k(|||e_0|||_{1,k}^2 - |||(I-\lambda_k^{-1}A_k)e_0|||_{2,k}^2) + D_2|||e_0|||_{1,k}^2 .
\end{aligned} \tag{7.19}
$$

Applying (7.17) for $j=1$ again to the above estimate, we obtain that

$$
\begin{aligned}
|||e_1|||_{2,k}^2 &\leq \lambda_k(|||e_0|||_{1,k}^2 + D_1|||e_0|||_{1,k}^2 - |||e_1|||_{1,k}^2) + D_2|||e_0|||_{1,k}^2 \\
&\leq \lambda_k(1+D_1+CH_k^2D_2)|||e_0|||_{1,k}^2 - (1+\rho^2/4+C\delta_k)^1|||e_2|||_{1,k}^2 ,
\end{aligned} \tag{7.20}
$$

where (7.7) with $m=1$ is used in the last step. Combining (7.20) and (7.6) with $m=1$, we obtain that $|||e_2|||_{1,k} \leq \gamma|||e_0|||_{1,k}$. We omit the numerical estimate of D_0 and of the other constants in (7.18). □

By letting ρ_k in (7.18) be small enough, the two-level multigrid method converges at a rate $\gamma < 1$. One can easily analyze the multilevel method by Theorem 7.2 following the work of Mandel [17]. The convergence rates of two-level multigrid methods are listed in Table 4. We can see by column 7 in Table 4 that Theorem 7.2 is verified. In fact, the multigrid method converges more rapidly in the presence of the nonsymmetric term $b(\cdot\,,\cdot)$, compared with the data in column 6 of Table 1. From column 5 of Table 4, we notice that the convergence of the nested multigrid method deteriorates. We are able to extend our theory above to cover the nested case (see column 2 of Table 4), similarly as we did in §§5–6.

We now discuss briefly the multigrid solver for the finite element equations arising from the streamline diffusion discretization of (7.1). In the streamline diffusion method, we quasiuniformly triangulate the domain and we replace the test function v in (7.3) by $v+\alpha h\partial_1 v$, where h is the mesh size. The resulting finite element equations are

$$
\begin{aligned}
&(\varepsilon+\alpha h)(\partial_1 u_k\,,\,\partial_1 v) + \varepsilon(\partial_2 u_k\,,\,\partial_2 v) + (\partial_1 u_k\,,\,v) \\
&\quad = (f\,,\,v+\alpha h\partial_1 v) \quad \forall v \in V_k .
\end{aligned} \tag{7.21}
$$

The stiffness matrices of (7.21) and (7.3) are of the same type. The diffusion in the direction of the streamline (x_1-direction) is much bigger than that in the other direction. This is the same situation as for finite elements on the above

TABLE 4. Spectral radii of 2-level multigrid iterative operators for (7.1)

ρ(CS) (nested meshes)					ρ(CS) (nonnested meshes)		
grid	$\varepsilon = h_k$	$\varepsilon = h_k^2$	grid	$\varepsilon = h_k$	grid	$\varepsilon = h_k$	$\varepsilon = h_k^2$
10×10	0.7089	1.4015	10×8	0.7619	10×8	0.5851	2.2447
14×14	0.7164	4.0869	14×8	0.8411	14×8	0.4385	2.3346
18×18	0.7191	8.1342	18×8	0.8847	18×8	0.4241	3.0278
22×22	0.7204	$\gg 1$	22×8	0.9102	22×8	0.4472	3.9692
26×26	0.7209	$\gg 1$	26×8	0.9259	26×8	0.4596	4.6948

high-aspect-ratio meshes. Thus, the coarse-level correction needs to correct the low-frequency error in the x_2-direction completely, and we need the special multigrid method. We can repeat the above analysis for (7.3) to get the following theorem.

Theorem 7.3. *The two-level multigrid method defined by* (7.4) *and* (7.14) *converges at a constant rate independent of* ε *when solving the streamline diffusion finite element equations* (7.21).

BIBLIOGRAPHY

1. R. E. Alcouffe, A. Brandt, J. E. Dendy, and J. W. Painter, *The multi-grid method for the diffusion equation with strongly discontinuous coefficients*, SIAM J. Sci. Statist. Comput. **2** (1981), 430–454.
2. I. Babuška, *The finite element method for elliptic equations with discontinuous coefficients*, Computing **5** (1970), 207–213.
3. I. Babuška and A. K. Aziz, *The angle condition in the finite element method*, SIAM J. Numer. Anal. **13** (1976), 214–226.
4. I. Babuška and J. E. Osborn, *Finite element methods for the solution of problems with rough input data*, Technical Note BN-1017, Laboratory for Numerical Analysis, University of Maryland, 1984.
5. R. Bank and T. Dupont, *An optimal order process for solving finite element equations*, Math. Comp. **36** (1981), 35–51.
6. R. Bank, *A comparison of two multilevel iterative methods for nonsymmetric and indefinite elliptic finite element equations*, SIAM J. Numer. Anal. **18** (1981), 724–743.
7. D. Braess, *The contraction number of a multigrid method for solving the Poisson equation*, Numer. Math. **37** (1981), 387–404.
8. 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), 1–31.
9. J. H. Bramble, J. E. Pasciak, J. Wang, and J. Xu, *Convergence estimates for multigrid algorithms without regularity assumptions*, Math. Comp. **57** (1991), 23–45.
10. P. G. Ciarlet, *The finite element method for elliptic problems*, North-Holland, Amsterdam, New York, and Oxford, 1978.
11. W. Hackbusch, *Multi-grid methods and applications*, Springer-Verlag, Berlin and Heidelberg, 1985.
12. ———, *Multigrid convergence for a singular perturbation problem*, Linear Algebra Appl. **58** (1984), 125–145.
13. P. Jamet, *Estimations d'erreur pour des éléments finis droits presque dégénérés*, RAIRO Sér. Rouge Anal. Numér. **10** (1976), 43–60.
14. C. Johnson, A. H. Schatz, and L. B. Wahlbin, *Crosswind smear and pointwise errors in streamline diffusion finite element methods*, Math. Comp. **49** (1987), 25–38.
15. R. B. Kellogg, *On the Poisson equation with intersecting interfaces*, Appl. Anal. **4** (1975), 101–129.
16. G. Lube, *Uniform in* ε *discretization error estimates for convection dominated convection-diffusion problems*, RAIRO Modél. Math. Anal. Numér. **22** (1988), 477–498.
17. J. Mandel, *Multigrid convergence for nonsymmetric, indefinite variational problems and one smoothing step*, Appl. Math. Comput. **19** (1986), 201–216.
18. S. F. McCormick (ed.), *Multigrid methods*, Frontiers in Applied Mathematics, Vol. 3, SIAM, Philadelphia, 1987.
19. J. E. Osborn, *Special finite element methods for elliptic problems with rough coefficients*, lecture notes, Workshop on Mathematics of Computation in Partial Differential Equations, Cornell University, 1991.

20. A. H. Schatz and L. B. Wahlbin, *On the finite element method for singularly perturbed reaction-diffusion problems in two and one dimensions*, Math. Comp. **49** (1987), 25–38.
21. R. Verfürth, *The contraction number of a multigrid method with mesh ratio 2 for solving Poisson's equation*, Linear Algebra Appl. **60** (1984), 113–128.
22. J. Xu, *Theory of multilevel methods*, Ph.D. thesis, Cornell University, 1989. Also Report AM48, Department of Mathematics, Pennsylvania State University, 1989.
23. S. Zhang, *Multi-level iterative techniques*, Ph.D. thesis, Pennsylvania State University, 1988.
24. ———, *Optimal-order nonnested multigrid methods for solving finite element equations* I: *On quasi-uniform meshes*, Math. Comp. **55** (1990), 23–36.
25. ———, *Optimal-order nonnested multigrid methods for solving finite element equations* II: *On non-quasi-uniform meshes*, Math. Comp. **55** (1990), 439–450.

Department of Mathematical Sciences, University of Delaware, Newark, Delaware 19716

E-mail address: szhang@math.udel.edu