ANALYSIS AND COMPARISON OF GEOMETRIC AND ALGEBRAIC MULTIGRID FOR CONVECTION-DIFFUSION EQUATIONS*

CHIN-TIEN WU† AND HOWARD C. ELMAN‡

Abstract. The discrete convection-diffusion equations obtained from streamline diffusion finite element discretization are solved on both uniform meshes and adaptive meshes. Estimates of error reduction rates for both geometric multigrid (GMG) and algebraic multigrid (AMG) are established on uniform rectangular meshes for a model problem. Our analysis shows that GMG with line Gauss–Seidel smoothing and bilinear interpolation converges if $h \gg \epsilon^{2/3}$, and AMG with the same smoother converges more rapidly than GMG if the interpolation constant β in the approximation assumption of AMG satisfies $\beta \ll (\frac{h}{\sqrt{\epsilon}})^\alpha$, where $\alpha = \begin{cases} 1, & h < \sqrt{\varepsilon}, \\ 2, & h \geq \sqrt{\varepsilon}. \end{cases}$ On unstructured triangular meshes, the performance of GMG and AMG, both as solvers and as preconditioners for GMRES, are evaluated. Numerical results show that GMRES with AMG preconditioning is a robust and reliable solver on both type of meshes.

1. Introduction. The purpose of this paper is to evaluate different solution strategies for the linear systems obtained from discretization of the convection-diffusion equation

$$(1.1) \qquad \begin{cases} -\epsilon \triangle u + b \cdot \nabla u = f & \text{on } \Omega, \\ u = g & \text{on } \partial\Omega, \end{cases}$$

where b and f are sufficiently smooth and the domain Ω is convex with Lipschitz boundary $\partial\Omega$. We are interested in the convection-dominated case, i.e., $|b| \gg \epsilon$. In this setting, the solution typically has steep gradients in some parts of the domain. These may take the form of boundary layers caused by Dirichlet conditions on the outflow boundaries or internal layers caused by discontinuities in the inflow boundaries.

Let $\Im_h$ be a given quasi-uniform mesh of triangles on Ω and let V_h be the linear finite element space on $\Im_h$. It is well known that the standard Galerkin finite element discretization on uniform grids produces inaccurate oscillatory solutions to convection-diffusion problems. Here, we will discretize (1.1) on V_h using the streamline diffusion finite element method (SDFEM) [7], a variant of the standard Galerkin method, where extra diffusion in the streamline direction is introduced. The SDFEM formulation entails finding $u_h \in V_h$ such that

$$(1.2) \qquad B_{sd}(u_h, v_h) = f_{sd}(v_h) \quad \text{for all } v_h \in V_h,$$

†Department of Computer Science and Engineering, National Taiwan Ocean University, Keelung Taiwan 20224, Republic of China (ctwu@cs.ntou.edu.tw).

‡Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742 (elman@cs.umd.edu). This author's work was supported in part by the National Science Foundation under grant DMS 0208015.

where

$$B_{sd}(u_h, v_h) = \epsilon(\nabla u_h, \nabla v_h) + (b \cdot \nabla u_h, v_h) + \sum_{T \in \Im_h} \delta_T (b \cdot \nabla u_h, b \cdot \nabla v_h)_T,$$

$$f_{sd}(v_h) = (f_h, v_h) + \sum_{T \in \Im_h} (f, \delta_T b \cdot \nabla v_h),$$

and δ_T is a stabilization parameter. With a carefully chosen value of δ_T, the streamline diffusion finite element discretization is able to eliminate most oscillations and produce accurate solutions in the regions where no layers are present [7]. As shown in [6], a good choice of δ_T is

(1.3)
$$\delta_T = \begin{cases} \frac{1}{2\|b\|_T} \left(1 - \frac{1}{P_{e_T}}\right) & \text{if } P_{e_T} > 1, \\ 0 & \text{if } P_{e_T} < 1, \end{cases}$$

where

$$P_{e_T} = \frac{\|b\|_T \, h_T}{2\epsilon} \quad \text{for } T \in \Im_h \text{ with diameter } h_T$$

is the mesh Peclét number.

This strategy does not produce accurate solutions in regions containing layers that are not resolved by the grid. Accuracy can be achieved at reasonable cost in such regions by adaptive mesh refinement, in which an a posteriori error estimation strategy is used to identify regions where errors are large, and a marking strategy is used to select elements to be refined. An adaptive solution strategy is to start with a coarse mesh, compute the discrete system, solve the linear system, and then locally refine the mesh using the information provided by an a posteriori error estimation. This adaptive solution process can be applied repeatedly until the a posteriori error estimator is less than a prescribed tolerance or the maximal number of mesh refinement steps is reached.

In this paper, we use the local error estimator proposed by Kay and Silvester [8], with the maximum marking strategy [11], and regular mesh refinement strategy for the adaptive refinement process. We are concerned with the costs of performing this computation with emphasis on the costs of solving the discrete systems that arise at each step of the mesh refinement. Let $A_h u_h = f_h$ denote the discrete system of equations derived from (1.2). Since adaptive mesh refinement produces a sequence of nested meshes, multigrid methods are natural candidates for solving this linear system. Here, we would like to explore the effectiveness of the geometric multigrid (GMG) [5] and algebraic multigrid (AMG) [17] methods. Since the linear system is nonsymmetric, Krylov subspace linear solvers such as the generalized minimal residual method (GMRES) [18] are also good choices. In order to accelerate the convergence of the Krylov subspace linear solvers, good preconditioners are needed. We would also like to investigate the performance of GMRES preconditioned by GMG and AMG.

This paper is organized as follows: Details of the error estimator and refinement strategy are given in section 2 where numerical studies that demonstrate the efficiency of the error indicator are presented. In section 3, multigrid algorithms and some estimates of two-level multigrid convergence are analyzed and numerical results that support our theoretical analysis are given. A comparison of the solvers is shown in section 4. This comparison examines both standalone versions of multilevel multigrid as well as versions in which multigrid is used as a preconditioner for GMRES. We also

compare these with unpreconditioned GMRES, and with a preconditioned GMRES algorithm in which the smoother used for multigrid is used as a preconditioner. These experiments are performed using the following two benchmark problems.

Problem 1: *Constant flow with characteristic and downstream layers.* Equation (1.1) is given with the coefficient $b = (\sin(\phi), \cos(\phi))$, where ϕ measures the angle of the flow direction from the y-axis, and the right-hand side $f = 0$ on the domain $\Omega = [-1, 1] \times [-1, 1]$. The Dirichlet boundary condition is set as $g = 1$ on the segments $y = -1 \cap x > 0$ and $x = 1$, and $g = 0$ elsewhere.

Problem 2: *Flow with closed characteristics:* Here, the coefficient vector (b_1, b_2) is $(2y(1 - x^2), 2x(1 - y^2))$ and the right-hand side $f = 0$ on the domain $\Omega = [-1, 1] \times [-1, 1]$. The Dirichlet boundary condition is $g = 1$ on the segments $y = 1$ and $g = 0$ elsewhere.

These examples are chosen because they exhibit many of the important features of flow problems so that their effect on discretization and multigrid solvers can be identified. In particular, Problem 1 has an exponential boundary layer at the outflow boundary and a characteristic internal layer induced by a discontinuity on the inflow boundary. Both of these can be treated using adaptive mesh refinement. Moreover, by varying the angle ϕ, we can assess the effect of flow direction on performance of solvers. Problem 2 also has a (characteristic) boundary layer caused by discontinuous boundary conditions, and its recirculating flow requires that care be taken in choosing smoothers for multigrid computations.

Finally, in section 5, we draw our conclusions.

2. Adaptive mesh refinement by a posteriori error estimation. One common technique to increase the accuracy of the finite element solution is mesh refinement. Unlike uniform mesh refinement, the adaptive mesh refinement process refines meshes only in the regions where errors between the weak solution of the partial differential equation and the corresponding finite element solution are large. In general, the adaptive mesh refinement process consists of loops of the following form:

$$\underbrace{Solve}_{1} \rightarrow \underbrace{Compute\ error\ indicator}_{2} \rightarrow \underbrace{Refine\ mesh}_{3}$$

In step 3, elements in which the value of the a posteriori error estimator is large are marked for refinement according to some element selection algorithm. On a given triangulation $\Im_h$, for any element $T \in \Im_h$, let η_T be the a posteriori error indicator of element T. A heuristic marking strategy is the *maximum marking strategy* where an element T^* is marked for refinement if

$$(2.1) \qquad\qquad \eta_{T^*} > \theta \max_{T \in \Im_h} \eta_T,$$

with a prescribed threshold $0 < \theta < 1$. For other marking strategies, we refer to [11]. Once decisions on where to refine are made, commonly used mesh refinement schemes are regular refinement, where a triangle is divided into four triangles equally, and longest side bisection [15], [16], where a triangle is divided into two triangles by adding a new node to the midpoint of the longest edge. The maximum marking strategy and regular refinement scheme are used in our experiments.

A reliable computable a posteriori error estimator in step 2 is the key for the adaptive mesh refinement process to succeed. For the convection-diffusion equation discretized by SDFEM, the first a posteriori error estimation where the error is estimated by computing the residual was proposed by Verfürth [21], and the first a posteriori error estimation where the error is estimated by solving a local problem was

developed by Kay and Silvester [8]. In this study, we use the Kay and Silvester a posteriori error estimator, which we have found to be a more effective choice in [23]. Hereafter, we call this indicator the KS-indicator. First, let us introduce the following abbreviations. Let $\|\cdot\|_{0,\Omega}$ and $\|\cdot\|_{0,T}$ denote the L^2 norm on domain Ω and element T, respectively. Let $\mathcal{E}(T)$ be the set of edges of T and $\omega_T = \cup_{T' \cap T \in \mathcal{E}(T)} T'$, i.e., ω_T is the set of triangles T' that share a common edge with T. Let P_T^0 be the L^2 projection onto the space of polynomials of degree 0 on element T. The interior residual R_T of element T and the interelement flux jump R_E of edge E are defined as follows:

$$R_T = (f - b \cdot \nabla u_h)|_T,$$
$$R_T^0 = \mathcal{P}_T^0(R_T),$$
$$R_E = \begin{cases} [\frac{\partial u_h}{\partial n_E}]_E & \text{if } E \in \Omega, \\ 0 & \text{if } E \in \partial\Omega, \end{cases}$$

where $[\cdot]_E$ is the jump across edge E.[1] Let Φ be the element affine mapping from the physical domain to the computational domain and let χ_i be the nodal basis function of node i. The approximation space is denoted as $Q_T = Q_T \bigoplus B_T$, where

$$Q_T = span\{\psi_E \circ \Phi^{-1}| \ \psi_E = 4\chi_i\chi_j, \ i,j \text{ are the endpoints of } E \text{ and}$$
$$E \in \partial T \bigcap (\Omega \bigcup \Gamma_N)\}$$

is the space spanned by quadratic edge bubble functions and

$$B_T = span\left\{ \psi_T \circ \Phi^{-1}| \ \psi_T = 27\prod_{i=1}^{3} \chi_i \right\}$$

is the space spanned by cubic interior bubble functions. For details, see [8]. On each element T, the error estimator is then given by $\eta_{h,T} = \|\nabla e_T\|_{0,T}$, where $e_T \in Q_T$ satisfies

(2.2) $$\epsilon(\nabla e_T, \nabla v)_T = (R_T^0, v)_T - \frac{1}{2}\epsilon \sum_{E \in \partial T} (R_E, v)_E.$$

Let $e_h = u - u_h$. The a posteriori error estimation is specified as follows:

(global upper bound)

(2.3) $$\|\nabla(e_h)\|_{0,\Omega} \leq C \left(\sum_{T \in \Im_h} \eta_{h,T}^2 + \sum_{T \in \Im_h} \left(\frac{h}{\epsilon}\right)^2 \|R_T - R_T^0\|_{0,T}^2 \right)^{1/2}$$

(local lower bound)

(2.4) $$\eta_{h,T} \leq c \left(\|e_h\|_{0,\omega_T} + \sum_{T \subset \omega_T} \frac{h_T}{\epsilon} \|b \cdot \nabla e_h\|_{0,T} + \sum_{T \subset \omega_T} \frac{h_T}{\epsilon} \|R_T - R_T^0\|_{0,T} \right),$$

where constants C and c are independent of the diffusion parameter ϵ and mesh size h.

---

[1]This definition of R_E for $E \in \partial\Omega$ is for Dirichlet boundary conditions. We don't consider Neumann conditions in this study; in case $E \in \partial\Omega$ and a Neumann condition holds, the flux jump R_E would be set to $-2(\frac{\partial u_h}{\partial n_E})$ [8].

The sharpness of the error indicator can be revealed by computing the local effectivity index,

$$E_T = \max_{t \in \Im_h} \frac{\eta_{h,T}}{|u - u_h|_{1,T}},$$

and the global effectivity index,

$$E_\Omega = \frac{(\sum_{T \in \Im_h} \eta_{h,T}^2)^{1/2}}{|u - u_h|_{1,\Omega}},$$

where $|\cdot|_{1,\Omega}$ and $|\cdot|_{1,T}$ represent the H^1 seminorms on domain Ω and element T, respectively. If both indices are close to 1, the error indicator is reliable and efficient. The global upper bound provides an estimate of the error over the whole domain, and the local lower bound locates where the error is large. It has been shown that E_Ω grows in proportion to $\sqrt{Pe}$ and E_T grows in proportion to Pe for the KS-indicator [8]. The following numerical results support this conclusion.

Consider here equation (1.1) on the domain $\Omega = [0,1] \times [0,1]$ with $b = (\beta_1, \beta_2) = (\sin 75°, \cos 75°)$ and the Dirichlet boundary condition is obtained from the analytic solution

(2.5) $$u(x,y) = \frac{e^{\beta_1 x/\varepsilon} - 1}{e^{\beta_1/\varepsilon} - 1} + \frac{e^{\beta_2 y/\varepsilon} - 1}{e^{\beta_2/\varepsilon} - 1}.$$

Both global effectivity index E_Ω and local effectivity index E_T are computed. In order to see how the effectivity indices change in terms of the diffusion parameter ϵ and mesh size h, the problem is solved over uniform meshes with mesh size $h = \frac{1}{8}, \frac{1}{16}, \frac{1}{32}$, and $\frac{1}{64}$ for $\epsilon = \frac{1}{64}, \frac{1}{256}, \frac{1}{1024}$, and $\frac{1}{4096}$. For accurately approximating the true error $u - u_h$, a very fine mesh $\Im_f$, as shown in Figure 2.1, consisting of 11272 nodes and 21379 elements is generated by first solving the above test problem with $\epsilon = 10^{-3}$ on a 64×64 mesh followed by three mesh refinement steps where a large threshold $\theta = 0.75$ in the maximum marking strategy (2.1) is used to prevent unnecessary over refinement. An approximation of the exact error is then computed on $\Im_f$ by
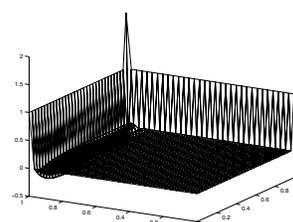
$$|u - u_h|_{1,\Omega^*} = \left(\sum_{\tau \in \Im_f \cup \Omega^*} |u - u_h|_{1,\tau}^2 \right)^{1/2} \quad \text{for } \Omega^* = \Omega \text{ or } T \in \Im_h,$$
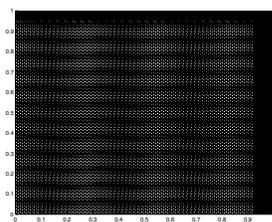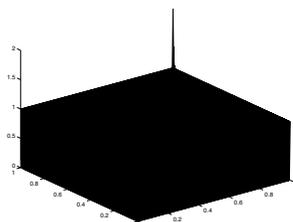
(a) Mesh $\Im_f$.   (b) Solution u on $\Im_f$.   (c) SDFEM solution u_h
                                                                    on $\Im_h$, where $h = \frac{1}{32}$.

FIG. 2.1.

TABLE 2.1
Effectivity indices of the KS-indicator.

ϵ	8×8	16×16	32×32	64×64
$\frac{1}{64}$	1.156	0.951	0.979	1.022
$\frac{1}{256}$	2.044	1.457	1.105	0.929
$\frac{1}{1024}$	4.016	2.772	1.952	1.422
$\frac{1}{4096}$	8.470	5.842	4.075	2.867

ϵ	8×8	16×16	32×32	64×64
$\frac{1}{64}$	2.504	1.714	1.687	1.557
$\frac{1}{256}$	9.242	4.637	2.536	1.750
$\frac{1}{1024}$	36.95	18.48	9.239	4.629
$\frac{1}{4096}$	147.8	73.90	36.95	18.48

(a) Global index E_Ω.   (b) Local index E_T.

(a) $\Im_1$.   (b) $\Im_2$.   (c) $\Im_3$.   (d) $\Im_4$.
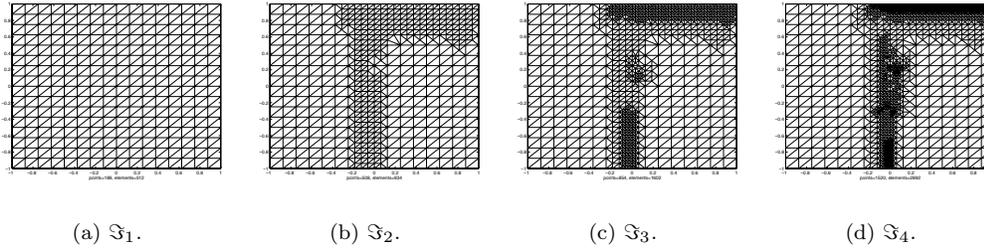
FIG. 2.2. *Adaptively refined mesh with threshold value $\theta = 0.01$.*

(a) Solution on $\Im_1$.   (b) Solution on $\Im_2$.   (c) Solution on $\Im_3$.   (d) Solution on $\Im_4$.
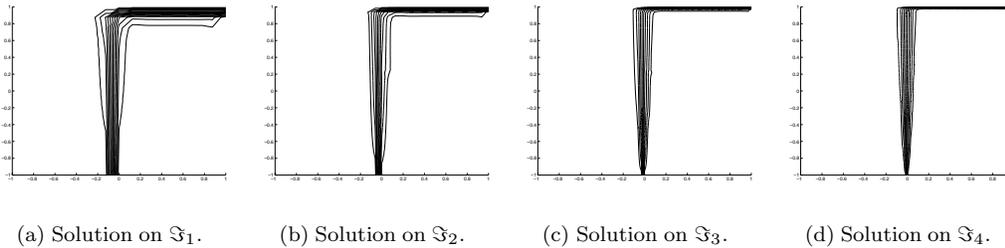
FIG. 2.3. *Contour plots of the solutions of Problem 1 on adaptively refined meshes starting with a uniform 16×16 mesh.*

where the discrete true solution u is evaluated directly from (2.5) on $\Im_f$ and the SDFEM solution u_h is prolonged from $\Im_h$ onto $\Im_f$ by standard bilinear interpolation. In Table 2.1, for $h \gg \epsilon$, E_Ω and E_T decrease by factors approximately of size $\sqrt{2}$ and 2, respectively, as the mesh size is halved. Also, E_Ω and E_T double and quadruple, respectively, when ϵ is reduced by a factor of four. Thus, $E_T = O(P_e)$ and $E_\Omega = O(\sqrt{P_e})$, as shown in [8].

Although there is deterioration in reliability and efficiency for these error indicators, the regions with large errors can still be located when a good threshold value θ is chosen. Next, we give a representative picture of how the adaptive solution process improves the solution quality. We use the KS-indicator for mesh refinement in solving Problem 1 with the diffusion parameter $\epsilon = 10^{-3}$ and $\phi = 0$. From the local lower bound (2.4), we suspect that the value of the indicator $\eta_{h,T}$ in the outflow boundary layer regions is $O(\frac{h}{\epsilon})$ larger than its value in the internal layer regions. Therefore, for the case considered here, a threshold value $\theta = 0.01 \approx \frac{\epsilon}{h}$ is chosen so that both the boundary and internal layer regions can be refined by the maximum marking strategy (2.1). A sequence of adaptively refined meshes is shown in Figure 2.2 in which ele-

ments in both boundary layer and internal layer regions are refined. It can be seen in Figure 2.3 that both the characteristic and downstream layers are resolved accurately by the mesh refinement process.

3. The multigrid algorithms. Given a sequence of meshes $\Im_1, \Im_2, \ldots, \Im_n$, let V_k be the piecewise linear finite element subspace associated with $\Im_k$, let A_k be the matrix defined on $\Im_k$, and let w_k be the initial guess. Let MG_1 represent the direct solver on the coarsest grid $\Im_1$. The general form of the k-level multigrid algorithm is shown in Algorithm 1, where f_k is the right-hand side obtained from discretization of data function f on $\Im_k$, M_k^{-1} represents the smoothing operator, and the operators I_k^{k-1} and I_{k-1}^k represent the grid transfers, restriction and prolongation, between $\Im_{k-1}$ and $\Im_k$. The multigrid V-cycle and W-cycle are defined by choosing $m = 1$ and $m = 2$, respectively.

Unlike in solving self-adjoint elliptic problems, GMG methods usually do not possess mesh-independent convergence for the convection-diffusion equation with dominant convection unless the mesh size $h \ll \epsilon$ or special treatments are employed for the meshes, interpolations, and relaxation [4] when $h \gg \epsilon$. For $h \ll \epsilon$ (and coarse grid size also $< \epsilon$), Bramble et al. [1], [2] and Wang [22] have shown GMG converges with a rate independent of mesh size by using a *compact perturbation technique*. To achieve mesh-independent convergence for $h \gg \epsilon$, Reusken and Olshanskii have recently shown L^2-convergence by using semicoarsening, matrix-dependent prolongation, and line relaxation for Problem 1 discretized by an upwind finite difference method [9], [14], and Szepessy proved L^1-convergence on R^2 by residual damping through large smoothing steps [12]. On the other hand, the convergence of AMG is essentially based on an M-matrix property of the discrete system [17]. No convergence result of algebraic multigrid for the convection-diffusion equation exists to our knowledge. In this section, we describe the versions of GMG and AMG considered here. We restrict our attention to the 2-level V-cycle multigrid and show some convergence results for Problem 1 on uniform rectangular meshes. For this problem, the matrix from SDFEM discretization with stabilization parameter $\delta_T = \frac{h}{2}$ has the form

$$(3.1) \qquad A_k = \begin{bmatrix} D & -U & & & \\ -L & D & -U & & 0 \\ & -L & \ddots & \ddots & \\ & & \ddots & \ddots & -U \\ & 0 & & -L & D & -U \\ & & & & -L & D \end{bmatrix},$$

Algorithm 1 Multigrid algorithm.

1. set $x_k = w_k$,
2. (presmoothing) $x_k = w_k + M_k^{-1}(g_k - A_k w_k)$,
3. (restriction) $\bar{g}_k = I_k^{k-1}(g_k - A_k x_k)$,
4. (correction) $q_i = MG_{k-1}(q_{i-1}, \bar{g}_k)$ for $1 \leq i \leq m$, $m = 1$ or 2 and $q_0 = 0$,
5. (prolongation) $\bar{q}_m = I_{k-1}^k q_m$,
6. set $x_k = x_k + \bar{q}_m$,
7. (postsmoothing) $x_k = x_k + M_k^{-1}(g_k - A_k x_k)$,
8. set $y_k = MG_k(w_k, g_k) = x_k$.

where the block tridiagonal matrices

$$D = h \times \text{tridiag} \left[\frac{1}{6} - \frac{1}{3}\frac{\epsilon}{h}, \frac{2}{3} + \frac{8}{3}\frac{\epsilon}{h}, \frac{1}{6} - \frac{1}{3}\frac{\epsilon}{h} \right],$$

(3.2)
$$L = h \times \text{tridiag} \left[\frac{1}{6} + \frac{1}{3}\frac{\epsilon}{h}, \frac{2}{3} + \frac{1}{3}\frac{\epsilon}{h}, \frac{1}{6} + \frac{1}{3}\frac{\epsilon}{h} \right],$$

$$U = \epsilon \times \text{tridiag} \left[\frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right].$$

In the following, we describe each component of the multigrid methods used here. First, we use the same smoother in both GMG and AMG. On uniform rectangular meshes, one step of a horizontal line Gauss–Seidel (HLGS) smoother is used. The HLGS method uses the block lower triangle of A_k as a splitting operator. (Thus, the HLGS sweep can be viewed as proceeding along horizontal lines, one line at a time, in the vertical direction.) Second, in GMG, (1) the coarse grids are obtained directly from the mesh refinement process; (2) the grid transfer is via linear interpolation for prolongation, and the restriction operator I_k^{k-1} is then taken to be the transpose of the prolongation operator I_{k-1}^k; and (3) the matrix A_k is obtained directly from SDFEM discretization in the refinement process. For AMG, we use the method developed by Ruge and Stüben [17] (AMG): (1) a coarse grid at level $k - 1$ is generated by coarsening the fine grid along the so-called strong connection direction of matrix graph $\mathcal{G}$ of $A_k = (a_{i,j})$ where a directive edge $\overrightarrow{e}_{i,j} \in \mathcal{G}$ is defined to be strongly connected from node i to node j if $\mu \leq \frac{-a_{i,j}}{\max_{m \neq i} (-a_{i,m})}$ for a given parameter $0 < \mu < 1$; (2) the interpolation operator I_{k-1}^k is defined dynamically by a sophisticated algebraic formula during the AMG coarsening process [17]; and (3) the matrix A_{k-1} for a coarse grid is computed by $I_k^{k-1} A_k I_{k-1}^k$. The above process is repeated until all coarse grids are generated.

Next, we show the convergence behavior of 2-level V-cycle multigrid for Problem 1. Let E_k^s be the error reduction operator of the HLGS smoother and let $E_k^c = I - I_{k-1}^k A_{k-1}^{-1} I_k^{k-1} A_k$ be the coarse grid correction operator. The error reduction operators of multigrid methods can then be written as $E_k^s E^c E_k^s$. The notation E_k^{mg} and E_k^{amg} denotes the error reduction operators of geometric multigrid and the algebraic multigrid, respectively. The notation $\|\cdot\|$ denotes the usual discrete L^2 norm for any vector and matrix. The notation $\|\cdot\|_0$ and $\|\cdot\|_1$ represents the continuous L^2 and H^1 norms. Variants of these norms are specified using matrix notation in (3.17) below.

In order to show convergence, the following auxiliary lemmas are needed.

LEMMA 3.1. *Given two symmetric matrices B_1 and B_2, assume that $B_1, B_2 \geq 0$, B_1 is irreducible, and B_2 is positive definite. The following properties hold.*

1. *There exists a positive eigenvector x^+ of $B_2^{-1}B_1$ such that*

(3.3)
$$B_2^{-1} B_1 x^+ = \rho(B_2^{-1} B_1) x^+.$$

2. *If $\alpha I - B_2^{-1}B_1$ is nonsingular and $(\alpha I - B_2^{-1}B_1)^{-1} \geq 0$, then $\rho(B_2^{-1}B_1) < \alpha$.*

Proof. The existence of a positive eigenvector satisfying (3.3) is essentially a generalization of the well-known Perron–Frobenius theorem [20, Theorem 2.7]. Using (3.3), one can prove the second result using a standard argument of the Perron–Frobenius theory; see Theorem 3.16 in [20].   □

LEMMA 3.2. *Let L and D be the matrices defined in (3.2). For any $\delta \geq (1 + \frac{2\epsilon}{h})\frac{h}{\epsilon}$, the matrix $\delta(D - L) - D$ is an M-matrix.*

Proof. Let us choose $\delta = \frac{h\gamma}{\epsilon}$ for some $\gamma > 0$. From (3.2), $D - L = \frac{\epsilon}{3} \times$ tridiag$[-2, 7, -2]$. Therefore, we have

$$\delta(D - L) - D = h \times \text{tridiag} \left[\frac{-2\gamma}{3}, \frac{7\gamma}{3}, \frac{-2\gamma}{3} \right] - h \times \text{tridiag} \left[\frac{1}{6} - \frac{\epsilon}{3h}, \frac{2}{3} + \frac{8\epsilon}{h}, \frac{1}{6} - \frac{\epsilon}{3h} \right]$$

$$= h \times \text{tridiag} \left[-\left(\frac{2\gamma}{3} + \frac{1}{6} - \frac{\epsilon}{3h} \right), \frac{7\gamma}{3} - \frac{2}{3} - \frac{8\epsilon}{3h}, -\left(\frac{2\gamma}{3} + \frac{1}{6} - \frac{\epsilon}{3h} \right) \right].$$

Since

$$\frac{7\gamma}{3} - \frac{2}{3} - \frac{8\epsilon}{3h} - 2 \left(\frac{2\gamma}{3} + \frac{1}{6} - \frac{\epsilon}{3h} \right) = \gamma - 1 - \frac{2\epsilon}{h},$$

clearly, for $\gamma \geq 1 + \frac{2\epsilon}{h}$, the matrix $\delta(D - L) - D$ is irreducible and weakly diagonal dominant. This implies that the matrix $\delta(D - L) - D$ and $(\delta(D - L) - D)^{-1}$ are positive definite. Moreover, since the off-diagonal entries of $(\delta(D - L) - D)$ are all negative, the matrix $\delta(D - L) - D$ is an M-matrix for $\delta \geq (1 + \frac{2\epsilon}{h}) \frac{h}{\epsilon}$. □

LEMMA 3.3. *For $h \gg \epsilon$, the error reduction matrix E_k^s of the HLGS iterative method for the matrix obtained from SDFEM discretization of Problem 1 satisfies the following inequality:*

(3.4) $$\|E_k^s\| \leq 3 \frac{\epsilon}{h^2} \min \left\{ \frac{3h}{\sqrt{\epsilon}}, 1 \right\}.$$

Proof. With A_k from (3.1), E_k^s is the product of the inverse of the block lower triangular part of A_k and the strict block upper triangular part of A_k. By direct computation, one has $E_k^s = G_1 G_2$, where

$$G_1 = \begin{bmatrix} 0 & I & & & \\ 0 & D^{-1}L & \ddots & & \\ \vdots & \vdots & & \ddots & I \\ 0 & (D^{-1}L)^{n-2} & \cdots & D^{-1}L & I \\ 0 & (D^{-1}L)^{n-1} & \cdots & (D^{-1}L)^2 & D^{-1}L \end{bmatrix}, G_2 = \begin{bmatrix} 0 & & & \\ & D^{-1}U & & \\ & & \ddots & \\ & & & D^{-1}U \end{bmatrix}.$$

From the Gerschgorin circle theorem, the following inequalities hold for $h \gg \epsilon$:

$$\|U\| = \rho(U) \leq \epsilon,$$

$$\|D^{-1}\| = \rho(D^{-1}) = \frac{1}{\lambda_{min}(D)} < \frac{1}{\frac{h}{3} + \frac{10\epsilon}{3}} \leq \frac{3}{h}.$$

Therefore, we have

(3.5) $$\|G_2\| \leq 3 \frac{\epsilon}{h}.$$

Next, we estimate $\|G_1\|$. First, let us estimate $\|D^{-1}L\|$. Considering $D^{-1}L = I - D^{-1}(D - L)$, we have

(3.6)

$$\alpha I - D^{-1}L = D^{-1}(D - L) - (1 - \alpha)I = (1 - \alpha) \left\{ D^{-1} \left[\frac{1}{1 - \alpha}(D - L) - D \right] \right\}.$$

Let us choose α satisfying $\frac{1}{1-\alpha} = \delta = (1 + \frac{2\epsilon}{h})\frac{h}{\epsilon}$. Lemma 3.2 implies that the matrix $\frac{1}{1-\alpha}(D - L) - D$ is an M-matrix. Consequently, $(\frac{1}{1-\alpha}(D - L) - D)^{-1} \geq 0$. Then using (3.6) and the fact that $D \geq 0$, it follows that the matrix $\alpha I - D^{-1}L > 0$. Since D is also positive definite, by Lemma 3.1 we can conclude that

$$(3.7) \qquad \rho(D^{-1}L) < \alpha = 1 - \frac{1}{\delta} = 1 - \frac{\epsilon}{h}\left(\frac{1}{1 + 2\frac{\epsilon}{h}}\right) < 1 - \frac{\epsilon}{3h}.$$

Let $x = (x_1, x_2, \ldots, x_N) \in V_h$, where $x_i \in R^N$ and $\sum_{i=1}^{N} \|x_i\|^2 = 1$. We have that $\|G_1 x\| < \|G_1 y\|$ for $y = (\|x_1\| x^+, \|x_2\| x^+, \ldots, \|x_N\| x^+)$, where x^+ is the Perron–Frobenius eigenvector of $D^{-1}L$. Therefore, the eigenvector corresponding to the maximum eigenvalue has the following form:

$$y = (0, \beta_1 x^+, \beta_2 x^+, \ldots, \beta_{N-1} x^+), \text{ where } \sum_{i=1}^{N-1} \beta_i^2 = 1.$$

By direct computation,

$$\|G_1 y\| \leq \left\{ \sum_{i=1}^{N} \left\| \sum_{k=1}^{i} \beta_k (D^{-1}L)^{i-k} x^+ \right\|^2 \right\}^{1/2}$$

$$= \left\{ \sum_{i=1}^{N} \left(\sum_{k=1}^{i} \beta_k l^{i-k} \right)^2 \right\}^{1/2},$$

where $l = \rho(D^{-1}L)$. Since $l < 1$ and $\sum_{i=1}^{N-1} \beta_i^2 = 1$, the inequalities

$$(3.8) \qquad \|G_1\| \leq \left\{ \sum_{i=1}^{N-1} \left(\sum_{k=1}^{i} \beta_k \right)^2 \right\}^{1/2} \leq \left(\sum_{i=1}^{N-1} i \right)^{1/2} \leq N$$

and

$$(3.9) \qquad \|G_1\| \leq \left\{ \sum_{i=1}^{N-1} \left[\sum_{k=1}^{i} \beta_k^2 \right] \left[\sum_{k=1}^{i} (l^{i-k})^2 \right] \right\}^{1/2} \leq \left\{ \sum_{i=1}^{N-1} \sum_{k=1}^{i} (l^{i-k})^2 \right\}^{1/2}$$

$$\leq \left\{ \frac{1}{1-l^2} \left(\sum_{i=1}^{N-1} 1 - l^{2i} \right) \right\}^{1/2} \leq \left(\frac{N}{1-l} \right)^{1/2}$$

hold. Recall that $l < 1 - \frac{\epsilon}{3h}$ from (3.7). By combining (3.8) and (3.9), we have

$$(3.10) \qquad \|G_1\| \leq \frac{1}{h} \min\left\{ \frac{3h}{\sqrt{\epsilon}}, 1 \right\}.$$

Therefore, from (3.5) and (3.10), inequality (3.4) holds.   □

LEMMA 3.4 (smoothing property). *Let E_k^s be the error reduction operator of the HLGS iterative method on V_k. The following inequality holds:*

$$(3.11) \qquad \|A_k(E_k^s)\| \leq \epsilon \left(1 + 3\frac{\epsilon}{h_k^2} \min\left\{ \frac{3h_k}{\sqrt{\epsilon}}, 1 \right\} \right).$$

Proof. By directly multiplying A_k and E_k^s, we have

$$A_k E_k^s = \begin{bmatrix} 0 & U - U(D^{-1}L)D^{-1}U & -UD^{-1}U \\ 0 & -U(D^{-1}L)^2 D^{-1}U & \ddots & & \ddots \\ \vdots & \vdots & & \ddots & -U(D^{-1}L)D^{-1}U & -UD^{-1}U \\ 0 & -U(D^{-1}L)^{n-1}D^{-1}U & \cdots & -U(D^{-1}L)^2 D^{-1}U & -U(D^{-1}L)D^{-1}U \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

$$= \mathrm{diag}[U](T_1 - T_2 G_2),$$

where

$$T_1 = \begin{bmatrix} 0 & I & & & \\ 0 & 0 & \ddots & & \\ \vdots & \vdots & \ddots & I & \\ 0 & 0 & \cdots & 0 & I \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \quad \text{and} \quad T_2 = \begin{bmatrix} 0 & D^{-1}L & I & & \\ 0 & (D^{-1}L)^2 & \ddots & \ddots & \\ \vdots & \vdots & \ddots & D^{-1}L & I \\ 0 & (D^{-1}L)^{n-1} & \cdots & (D^{-1}L)^2 & D^{-1}L \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Using the same argument as in the proof of Lemma 3.3, one can show

$$\|T_2 G_2\| \le 3\frac{\epsilon}{h_k^2} \min\left\{\frac{3h_k}{\sqrt{\epsilon}}, 1\right\}.$$

Therefore,

$$\|A_k E_k^s\| \le \epsilon\left(1 + 3\frac{\epsilon}{h_k^2} \min\left\{\frac{3h_k}{\sqrt{\epsilon}}, 1\right\}\right)$$

$$= \epsilon\left(1 + 3\frac{\epsilon}{h_k^2} \min\left\{\frac{3h_k}{\sqrt{\epsilon}}, 1\right\}\right). \qquad \square$$

Using the inequalities in Lemmas 3.3 and 3.4, our multigrid convergence results can be shown. In the following, estimates of the error reduction operators of GMG and AMG are given in Theorems 3.5 and 3.7, respectively.

THEOREM 3.5. *Let E_k^{mg} be the error reduction operator of the 2-level V-cycle geometric multigrid in Algorithm 1 with the HLGS smoother and bilinear interpolation prolongation on a uniform rectangular mesh $\Im_k$ for Problem 1. The following estimate for E_k^{mg} holds:*

(3.12) $$\|E_k^{mg}\| \le c\frac{\epsilon}{h^{3/2}} \quad \text{for some constant } c.$$

Proof. First, in general, the error reduction operator of GMG can be written as

(3.13) $$E_k^{mg} = E_k^s E^c E_k^s = E_k^s(I - I_{k-1}^k(I - E_{k-1}^{mg})A_{k-1}^{-1}I_k^{k-1}A_k)E_k^s.$$

For 2-level GMG, since $E_{k-1}^{mg} = 0$, we have

(3.14) $$\|E_k^{mg}\| = \left\|E_k^s(I - I_{k-1}^k A_{k-1}^{-1}I_k^{k-1}A_k)E_k^s\right\|$$

$$\le \|E_k^s\|\left\|(I - I_{k-1}^k A_{k-1}^{-1}I_k^{k-1}A_k)E_k^s\right\|.$$

Let $P_k : H^1 \mapsto V_k$ be the projection operator defined by $B_{sd}(P_k u, v) = B_{sd}(u, v)$ for all $v \in V_k$ for all k. We have $I_k^{k-1} A_k = A_{k-1} P_{k-1}$ on V_k. Therefore, let $e^s = E_k^s(e)$ for any $e \in V_k \subset H^1$,

$$\left\| (I - I_{k-1}^k A_{k-1}^{-1} I_k^{k-1} A_k) E_k^s(e) \right\| = \left\| (I - I_{k-1}^k P_{k-1}) e^s \right\|$$

$$\leq c h_k^{-2} \left\| (I - P_{k-1}) e^s \right\|_0 \leq c h_k^{-2} \frac{1}{\sqrt{\epsilon}} \left\| (I - P_{k-1}) e^s \right\|_1$$

$$\leq c h_k^{-2} \sqrt{\frac{h_k}{\epsilon}} |e^s|_1 \quad \text{by the a priori error estimation,}$$

$$\leq c h_k^{-2} \sqrt{\frac{h_k}{\epsilon^2}} \left\| A_k E_k^s(e) \right\|_0 \quad \text{by the regularity estimate,}$$

$$\leq c \sqrt{h_k} \left(1 + 3 \frac{\epsilon}{h_k^2} \min \left\{ \frac{3 h_k}{\sqrt{\epsilon}}, 1 \right\} \right) \|e\| \quad \text{by (3.11),}$$

where c is a constant independent of h and ϵ. By combining the above estimate and (3.4), inequality (3.14) implies

$$\|E_k^{mg}\| \leq c \frac{\epsilon}{h^{3/2}} \min \left\{ \frac{h}{\sqrt{\epsilon}}, 1 \right\} \left(1 + \frac{\epsilon}{h^2} \min \left\{ \frac{h}{\sqrt{\epsilon}}, 1 \right\} \right) \leq c \frac{\epsilon}{h^{3/2}}. \qquad \square$$

Remark 3.6. From inequality (3.12), it is clear that the geometric multigrid converges when $h_k \gg \epsilon^{2/3}$. The above 2-level analysis can be generalized by mathematical induction and estimation of (3.13). In [23], we have found that the multigrid V-cycle converges when $h \gg \sqrt{\epsilon}$.

Next, we analyze AMG convergence. The framework to prove the convergence of algebraic multigrid is based on the *smoothing assumption*,

$$(3.15) \qquad \exists\, \alpha > 0 \text{ such that } \|E_k^s e_h\|_1^2 \leq \|e_h\|_1^2 - \alpha \|e_h\|_2^2 \text{ for any } e_h \in V_k,$$

and the *approximation assumption*,

$$(3.16) \qquad \min_{e_H \in V_{k-1}} \left\| e_h - I_H^h e_H \right\|_0^2 \leq \beta \|e_h\|_1^2 \text{ with } \beta > 0 \text{ independent with } e_h \in V_k,$$

where

$$(3.17) \qquad \|v\|_0^2 = \langle Dv, v \rangle, \quad \|v\|_1^2 = \langle A_k v, v \rangle, \quad \text{and} \quad \|v\|_2^2 = \langle D^{-1} A_k v, A_k v \rangle$$

for any $v \in V_k$. For problems where the coefficient matrix is an M-matrix, Ruge and Stüben have proved that 2-level AMG converges when both the smoothing assumption and the approximation assumption hold for M-matrices. It has also been shown that the smoothing assumption holds for the usual point Gauss–Seidel relaxation, and the approximation assumption holds for some simple coarse grid selection algorithms

with algebraic interpolation formula [17]. The classical algebraic multigrid coarsening algorithm is designed to satisfy the approximation assumption based on the property of algebraically smooth error, whereby the smooth errors vary slowly along the strong connection directions.

For the discrete convection-diffusion Problem 1, the matrix obtained from SDFEM discretization is nonsymmetric and not an M-matrix. As a result, the usual framework of AMG analysis cannot be applied directly. In Lemma 3.2 of [14], Reusken has shown that there is a constant c independent of ϵ and h such that $\|E^c\| \leq c$ for a two-grid method with a coarse grid from semicoarsening and a matrix-dependent prolongation. In the following, we take this as an assumption and estimate the error reduction rate of AMG in terms of the parameter β in the approximation assumption (3.16) by using (3.4) and (3.11).

THEOREM 3.7. *Assume the approximation assumption* (3.16) *holds. Then the following inequality holds:*

$$(3.18) \qquad \|E_k^{amg}\| \leq \sqrt{3}c\frac{\epsilon^2}{h^3} \min\left\{\frac{h}{\sqrt{\epsilon}}, 1\right\} \frac{\beta}{\sqrt{h}},$$

where c is some constant independent of ϵ and h.

Proof. For any $e \in V_k$, we have

$$(3.19) \qquad \|E_k^{amg}(e)\| = \|E_k^s E^c E_k^s(e)\| \leq \frac{\epsilon}{h^2} \min\left\{\frac{h}{\sqrt{\epsilon}}, 1\right\} \|E^c E_k^s(e)\|.$$

Let $\bar{e} = E_k^s(e)$. Since $\mathrm{Ker}(E^c) = Rang(I_{k-1}^k)$, we have

$$\|E^c(\bar{e})\| \leq c \min_{e_H \in V_{k-1}} \left\|\bar{e} - I_{k-1}^k e_H\right\| \leq c\sqrt{\frac{3}{h}} \min_{e_H \in V_{k-1}} \left\|\bar{e} - I_{k-1}^k e_H\right\|_0$$

$$\leq c\sqrt{\frac{3}{h}}\beta \|\bar{e}\|_1 \text{ by approximation assumption (3.16)},$$

$$(3.20) \qquad \leq \sqrt{\frac{3}{h}}\beta(\|A_k E_k^s(e)\| \|E_k^s(e)\|)^{1/2}$$

$$\leq c\sqrt{\frac{3}{h}}\beta \left(\epsilon\left[1 + \frac{\epsilon}{h^2}\min\left\{\frac{h}{\sqrt{\epsilon}}, 1\right\}\right]\left[\frac{\epsilon}{h^2}\min\left\{\frac{h}{\sqrt{\epsilon}}, 1\right\}\right]\right)^{1/2} \|e\|$$

$$\text{by (3.4) and (3.11)}$$

$$\leq c\sqrt{\frac{3}{h}}\frac{\epsilon}{h}\beta \|e\|.$$

By combining (3.19) and (3.20), inequality (3.18) holds. $\square$

Remark 3.8. Inequality (3.18) can be rewritten as

$$\|E_k^{amg}\| \leq c\frac{\epsilon}{h^{3/2}}\min\left\{\frac{h}{\sqrt{\epsilon}}, 1\right\}\frac{\beta\epsilon}{h^2} = c\frac{\epsilon}{h^{3/2}}\left(\frac{\sqrt{\epsilon}}{h}\right)^\alpha \beta,$$

where $\alpha = \begin{cases} 1, & h < \sqrt{\varepsilon}, \\ 2, & h \geq \sqrt{\varepsilon}. \end{cases}$ Comparison with the estimate of E_k^{mg} in (3.12) suggests that algebraic multigrid will converge more rapidly than geometric multigrid if the parameter β in the approximation assumption is less than $(\frac{h}{\sqrt{\epsilon}})^\alpha$. Although we have not

TABLE 3.1
Comparison of GMG and AMG for Problem 1 on uniform rectangular meshes.

ϵ	10^{-1}	10^{-2}	10^{-3}	10^{-4}
$h = \frac{1}{8}$	5	4	2	2
$h = \frac{1}{16}$	5	4	3	2
$h = \frac{1}{32}$	5	5	4	2

ϵ	10^{-1}	10^{-2}	10^{-3}	10^{-4}
$h = \frac{1}{8}$	4	3	2	2
$h = \frac{1}{16}$	4	3	3	2
$h = \frac{1}{32}$	5	4	3	2

(a) GMG iterative steps.   (b) AMG iterative steps.

yet estimated β for the interpolation operator from AMG coarsening, AMG converges more rapidly than GMG, as shown in Table 3.1. Moreover, in Table 3.1, the facts that both GMG and AMG converge more rapidly as ϵ becomes smaller and converge more slowly as the mesh size h decreases are also consistent with our theoretical analysis. In these numerical studies, the stopping tolerance is set to be

$$\|r_m\| \leq 10^{-6} \|r_0\|, \tag{3.21}$$

where r_0 is the initial residual and r_m is the residual at the mth iteration.

Remark 3.9. Since the row-sum $\sum_j (a_{i,j})$ and column-sum $\sum_i (a_{i,j})$ are equal for the matrix A_k arising from Problem 1, the following equality holds:

$$\|e\|_1^2 = (A_k e, e) = \sum_{i,j} a_{i,j} e_i e_j = \frac{1}{2} \left(\sum_{i,j} (-a_{i,j})(e_i - e_j)^2 \right) + \sum_i s_i e_i^2, \tag{3.22}$$

where $s_i = \sum_j (a_{i,j})$. Recall that the algebraically smooth error e_s is characterized by $E_k^s(e_s) \approx e_s$. Clearly, inequalities (3.4) and (3.11) imply

$$\langle A_k e_s, e_s \rangle \approx \langle A_k E_k^s e_s, E_k^s e_s \rangle \leq \frac{c}{h} \|A_k E_k^s\| \|E_k^s\| \|e_s\|_0^2$$

$$\leq c \frac{\varepsilon^2}{h^3} \min\left\{ \frac{3h}{\sqrt{\varepsilon}}, 1 \right\} \left(1 + \frac{3\varepsilon}{h^2} \min\left\{ \frac{3h}{\sqrt{\varepsilon}}, 1 \right\} \right) \|e_s\|_0^2$$

$$\leq c \frac{\varepsilon^2}{h^3} \|e_s\|_0^2.$$

For $h \gg \epsilon^{2/3}$, $\|e_s\|_1^2 \ll \|e_s\|_0^2$ follows from the above estimate. Moreover, by (3.22), this is equivalent to

$$\frac{1}{2} \sum_i \sum_{j \neq i-1, i+1} (-a_{i,j})(e_i - e_j)^2 + \sum_i s_i e_i^2$$

$$\ll \sum_i a_{i,i} e_i^2 + a_{i,i-1}(e_i - e_{i-1})^2 + a_{i,i+1}(e_i - e_{i+1})^2$$

$$\Rightarrow \frac{1}{2} \sum_i \sum_{j \neq i-1, i+1} \left(\frac{-a_{i,j}}{a_{i,i}} \right) (e_i - e_j)^2 \ll \sum_i e_i^2 + \frac{1}{4}(e_i - e_{i-1})^2 + \frac{1}{4}(e_i - e_{i+1})^2.$$

It can be seen from (3.1) that the weights $\frac{a_{i,j}}{a_{i,i}} = O(1)$ for downstream indices j and $O(\frac{\epsilon}{h})$ for upstream. Therefore, we can conclude that the smooth error tends to vary more slowly in downstream directions than in upstream directions. Since the strong
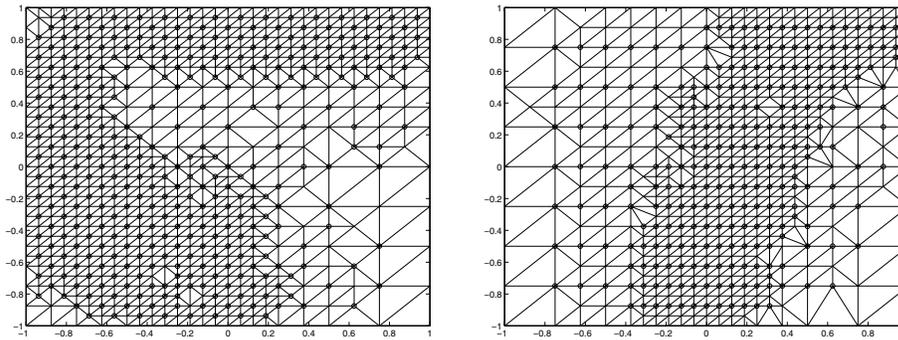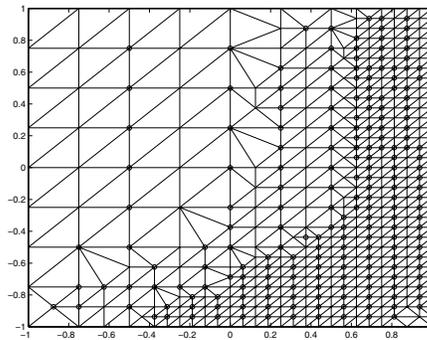
(a) $\phi = -45°$ on an adaptive mesh.

(b) $\phi = 15°$ on an adaptive mesh.

(c) $\phi = 75°$ on an adaptive mesh.

FIG. 3.1. *Coarse grid points selected by AMG coarsening.*

connectivity of the matrix A_k is also along the downstream direction, we expect the AMG coarsening to select a large number of downstream nodes, and the resulting strategy tends to resemble one of semicoarsening.

In the following, we show the AMG coarse grids for the Problem 1 with $\epsilon = 10^{-3}$ and $\phi = -45°$, 15°, and 75°, where the connection parameter is $\mu = 0.25$. Each fine grid is generated by two adaptive mesh refinement steps from an initial 8×8 mesh and the threshold value θ in (2.1) is set to 0.01. In Figure 3.1, the fine grid is plotted and the coarse grid points, selected by AMG coarsening, are marked by ○. Evidently, the AMG coarsening process is sensitive to the flow direction mentioned in Remark 3.9 and generates meshes by means of semicoarsening. The numerical results in Table 3.1, and Table 4.3 and 4.4 of the next section, shows that AMG converges more quickly than GMG for Problem 1. This leads us to conjecture that the algebraically smooth errors are well in the range of the AMG interpolation operator, i.e., the parameter β in the approximation assumption is small. Analysis of the AMG interpolation and approximation assumption will be considered in a future study.

For Problem 2, multigrid convergence is much harder to prove. In fact, the presence of a stagnation point leads to poor performance of multigrid methods when the equation is discretized by a first order upwind scheme [10], [4]. However, many

numerical studies have shown that GMG- and AMG-accelerated Krylov space methods, such as GMRES and BiCGSTAB, achieve nearly mesh-independent convergence [10], [19]. Recently, Ramage has demonstrated that GMRES with GMG preconditioning achieves mesh-independent convergence when SDFEM with an optimal stabilization parameter is employed for discretization on uniform rectangular mesh [13]. In the next section, GMG and AMG are used as a preconditioner for GMRES in some numerical tests.

4. Numerical results. In this section, the performances of different linear solvers, including GMG, AMG, and preconditioned GMRES, for the discrete convection-diffusion equation are compared on both adaptively refined meshes and uniform meshes. We use both Problem 1 with $\phi = 0$ and Problem 2. In both problems, (1.1) is discretized by SDFEM with the stabilization parameter δ_T defined in (1.3) on both a uniform 32×32 mesh and an adaptively refined mesh for $\epsilon = 10^{-2}, 10^{-3}$, and 10^{-4}. Here, a 32×32 mesh is generated from 3 uniform refinements starting with a 4×4 initial mesh, and the adaptively refined mesh is generated by refining an initial 8×8 mesh 4 times, where full multigrid is used with those coarse meshes. The threshold value θ in the maximum marking strategy is chosen such that elements in the regions containing large errors can be refined for both problems.

For Problem 1, $\theta = 0.1$, 0.01, and 0.001 for $\epsilon = 10^{-2}, 10^{-3}$, and 10^{-4}, respectively. The adaptive meshes are shown in Figure 4.1. For Problem 2, $\theta = 0.1$ for all ϵ and the adaptive meshes are shown in Figure 4.2.

Since there are no natural "lines" in unstructured meshes or in the coarse meshes of AMG, only point-versions of the Gauss–Seidel method are used for smoothing. Grid points are numbered in lexicographical order where the y-coordinate is the primary

(a) Mesh: $\epsilon = 10^{-2}$.   (b) Mesh: $\epsilon = 10^{-3}$.   (c) Mesh: $\epsilon = 10^{-4}$.
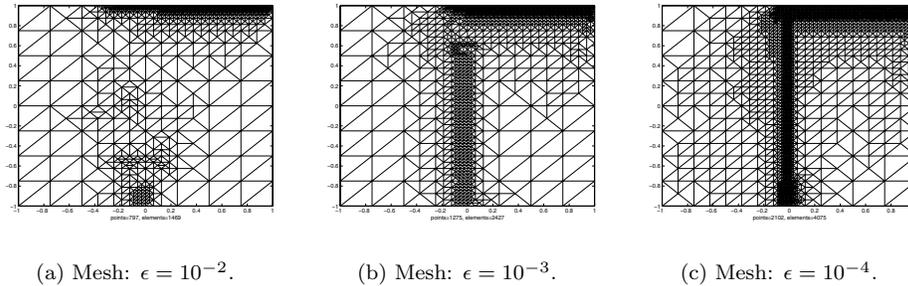
FIG. 4.1. *Problem 1: Adaptively refined meshes after 4 refinements starting from an 8×8 grid for various ϵ.*

(a) Mesh: $\epsilon = 10^{-2}$.   (b) Mesh: $\epsilon = 10^{-3}$.   (c) Mesh: $\epsilon = 10^{-4}$.
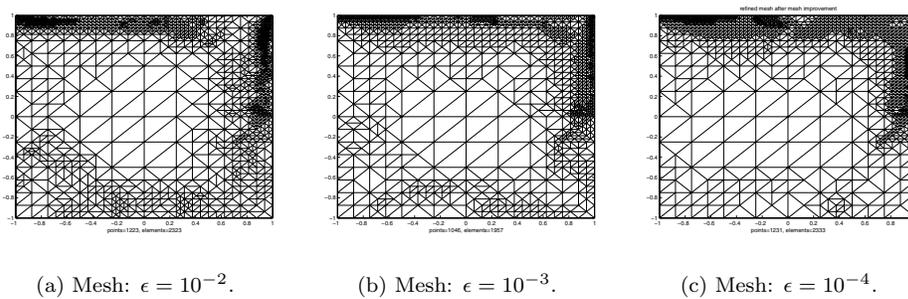
FIG. 4.2. *Problem 2: Adaptively refined meshes for various ϵ.*

TABLE 4.2
Number of coarse grid points from adaptive refinement and AMG coarsening.

$\log_{10} \frac{1}{\epsilon}$	GMG			AMG		
	2	3	4	2	3	4
Level = 5	797	1275	2102	797	1275	2102
Level = 4	410	649	1047	359	586	997
Level = 3	215	320	528	193	328	547
Level = 2	122	176	239	112	179	308
Level = 1	81	81	81	55	101	153

(a) Problem 1.

$\log_{10} \frac{1}{\epsilon}$	GMG			AMG		
	2	3	4	2	3	4
Level = 5	1233	946	922	1233	946	922
Level = 4	625	632	823	597	496	445
Level = 3	305	373	387	338	291	258
Level = 2	150	196	199	205	168	165
Level = 1	81	81	81	121	95	97

(b) Problem 2.

TABLE 4.3
Problem 1: Iteration counts for GMG and AMG on uniform meshes.

Level	GMG	AMG
3	13	7
2	13	6
1	12	6

(a) $\epsilon = 10^{-2}$.

Level	GMG	AMG
3	27	8
2	26	7
1	16	6

(b) $\epsilon = 10^{-3}$.

Level	GMG	AMG
3	51	11
2	35	8
1	17	6

(c) $\epsilon = 10^{-4}$.

TABLE 4.4
Problem 1: Iteration counts for GMG and AMG on adaptive meshes.

Level	GMG	AMG
4	9	6
3	8	6
2	7	5
1	7	5

(a) $\epsilon = 10^{-2}$.

Level	GMG	AMG
4	22	6
3	24	6
2	18	6
1	17	6

(b) $\epsilon = 10^{-3}$.

Level	GMG	AMG
4	59	11
3	57	9
2	47	7
1	34	6

(c) $\epsilon = 10^{-4}$.

dependence on ϵ and mesh size. Problem 2, with a recirculating flow, is more difficult for both solvers. On a uniform grid, GMG diverges for $\epsilon = 10^{-4}$, but AMG manages to converge, as shown in Table 4.5. On adaptive meshes, as shown in Table 4.6, textbook-like performance is achieved for both methods. However, AMG is considerably slower than GMG, especially for small $\epsilon = 10^{-4}$ on the finest mesh. Our speculation

TABLE 4.5
Problem 2: *Iteration counts for GMG and AMG on uniform meshes.*

Level	GMG	AMG
3	25	9
2	25	8
1	20	7

Level	GMG	AMG
3	179	11
2	135	21
1	43	11

Level	GMG	AMG
3	-	52
2	-	43
1	57	13

(a) $\epsilon = 10^{-2}$.   (b) $\epsilon = 10^{-3}$.   (c) $\epsilon = 10^{-4}$.

TABLE 4.6
Problem 2: *Iteration counts for GMG and AMG on adaptive meshes.*

Level	GMG	AMG
4	9	12
3	9	11
2	7	10
1	6	7

Level	GMG	AMG
4	27	26
3	21	26
2	24	21
1	20	17

Level	GMG	AMG
4	30	39
3	32	38
2	40	44
1	30	36

(a) $\epsilon = 10^{-2}$.   (b) $\epsilon = 10^{-3}$.   (c) $\epsilon = 10^{-4}$.

TABLE 4.7
Problem 1: *Iteration counts for various GMRES methods on finest grids.*

ϵ	10^{-2}	10^{-3}	10^{-4}
GMG	13	27	51
AMG	7	8	11
GMRES	56	76	95
GMRES-HGS	24	26	36
GMRES-GMG	11	16	22
GMRES-AMG	6	7	9

ϵ	10^{-2}	10^{-3}	10^{-4}
GMG	9	22	59
AMG	6	6	11
GMRES	83	146	-
GMRES-HGS	19	25	48
GMRES-GMG	8	13	27
GMRES-AMG	5	5	9

(a) Iterative steps on uniform mesh.   (b) Iterative steps on adaptive mesh.

is that this is caused by inadequacy of the grid transfer operator used in AMG and the fact that the number of coarse grid points generated by AMG coarsening is only about half the number of coarse grid points used in GMG, as shown in Table 4.2(b). AMG interpolation can be improved by the recently developed AMGe methods [3]. As we will show below, performance can also be remedied using a Krylov space acceleration.

Next, we compare the performance of preconditioned GMRES methods on both the finest uniform mesh and the finest adaptive mesh. Tables 4.7 and 4.8 examine the iteration counts for various versions of preconditioned GMRES. The Krylov space acceleration of the convergence of GMG and AMG is observed in most cases. The numerical results indicate that the AMG-preconditioned GMRES requires the fewest iterations to converge, compared to GMG, AMG, and other versions of GMRES. Particularly in Problem 2, despite the fact that AMG converges more slowly than GMG, GMRES-AMG outperforms GMRES-GMG. This observation suggests that AMG can be a good preconditioner for Krylov space solvers in solving complex flow problems. Moreover, in Table 4.9, iterations of GMRES-AMG on each level of the adaptive meshes are shown for $\epsilon = 10^{-2} \cdots 10^{-4}$. It is clear that the convergence of GMRES-AMG depends only weakly on the mesh size and ϵ on adaptive meshes.

Remark 4.1. Notice a difference between performance of GMG for Problem 1

TABLE 4.8
Problem 2: *Iteration counts for various GMRES methods on finest grids.*

ϵ	10^{-2}	10^{-3}	10^{-4}
GMG	25	179	-
AMG	9	11	52
GMRES	392	-	-
GMRES-ADGS	33	54	73
GMRES-GMG	11	29	42
GMRES-AMG	6	8	15

ϵ	10^{-2}	10^{-3}	10^{-4}
GMG	9	27	30
AMG	11	26	39
GMRES	199	-	-
GMRES-ADGS	30	37	35
GMRES-GMG	7	12	12
GMRES-AMG	6	8	9

(a) Iterative steps on uniform mesh.　　(b) Iterative steps on adaptive mesh.

TABLE 4.9
Iteration counts of GMRES-AMG for various ϵ on adaptive meshes.

Level	$\epsilon = 10^{-2}$	10^{-3}	10^{-4}
4	5	5	9
3	5	5	8
2	5	5	6
1	4	5	6

Level	$\epsilon = 10^{-2}$	10^{-3}	10^{-4}
4	6	8	9
3	6	7	9
2	5	7	9
1	4	6	7

(a) Iterative steps in Problem 1.　　(b) Iterative steps in Problem 2.

here and the bounds shown in section 3 which suggest convergence is essentially independent of ϵ. The latter result holds only in special settings and depends on having a very accurate smoother. Similar ϵ-independent results for this problem can be found in [9], [12], [14]. In contrast, the better performance of AMG is due to matrix-dependent grid transfer operators and the superior choice of coarse mesh identified by AMG coarsening, for which the smoother is better able to mimic the flow characteristics.

5. Conclusion. In this work, we have explored solution strategies for the convection-diffusion equation using adaptive gridding techniques and a multigrid algorithm. In section 3, the convergence analysis for both GMG with bilinear interpolation and line Gauss–Seidel smoother and AMG with the same smoother not only shows that GMG converges for Problem 1 when $h \gg \epsilon^{2/3}$, but also suggests that AMG converges faster than GMG on uniform rectangular meshes. Although the AMG interpolation and approximation assumption remain to be analyzed, our numerical results support this observation on both triangular and rectangular uniform meshes. In addition, numerical studies in section 4 show that the convergence rates of both GMG and AMG deteriorates as the diffusion parameter ϵ decreases. To overcome this drawback, we have found that Krylov space acceleration significantly improves the convergence of both GMG and AMG, especially when ϵ is small. Moreover, GMRES-AMG converges more rapidly than GMRES-GMG in all of our test cases on both uniform and adaptive meshes.

With adaptive mesh refinement, GMG seems to be a natural choice because the coarse grids, correction operators, and interpolations are ready to be used in GMG. However, GMG convergence is slower than AMG for some examples of both tested problems even though more coarse grid points are employed in GMG. On the other hand, the most important strength of AMG is that it can be treated as a black-box solver. There is no need to compute coarse grids and interpolations explicitly in AMG. The facts that AMG is applicable to a wider range of applications, AMG

converges for all of our test cases, and GMRES-AMG outperforms all other solvers in
our numerical studies make AMG very attractive for solving the SDFEM-discretized
convection-diffusion problems.

REFERENCES

[1] J. H. BRAMBLE, D. A. KWAK, AND J. E. PASCIAK, *Uniform convergence of multigrid V-cycle iterations for indefinite and nonsymmetric problems*, SIAM J. Numer. Anal., 31 (1994), pp. 1746–1763.
[2] J. H. BRAMBLE, J. E. PASCIAK, AND J. XU, *The analysis of multigrid algorithms for nonsymmetric and indefinite elliptic problems*, Math. Comp., 51 (1988), pp. 398–414.
[3] M. BREZINA, A. J. CLEARY, R. D. FALGOUT, V. E. HENSON, J. E. JONES, T. A. MANTEUFFEL, S. F. McCORMICK, AND J. W. RUGE, *Algebraic multigrid based on element interpolation (AMGe)*, SIAM J. Sci. Comput., 22 (2000). pp. 1570–1592.
[4] P. M. DE ZEEUW, *Matrix-dependent prolongations and restrictions in a blackbox multigrid solver*, J. Comput. Appl. Math., 33 (1990), pp. 1–27.
[5] R. P. FEDORENKO, *A relaxation method for solving elliptic difference equations*, USSR Comput. Math. Phys., 1 (1961), pp. 1092–1096.
[6] B. FISCHER, A. RAMAGE, D. SILVESTER, AND A. J. WATHEN, *On parameter choice and iterative convergence for stabilised discretisations of advection-diffusion problems*, Comput. Methods Appl. Mech. Engrg., 179 (1999), pp. 179–195.
[7] T. J. R. HUGHES, M. MALLET, AND A. MIZUKAMI, *A new finite element formulation for computational fluid dynamics:* II. *Beyond SUPG*, Comput. Methods Appl. Mech. Engrg., 54 (1986), pp. 485–501.
[8] D. KAY AND D. SILVESTER, *The reliability of local error estimators for convection-diffusion equations*, IMA J. Numer. Anal, 21 (2001), pp. 107–122.
[9] M. A. OLSHANSKII AND A. REUSKEN, *Convergence analysis of a multigrid method for a convection-dominated model problem*, SIAM J. Numer. Anal., 42 (2004), pp. 1261–1291.
[10] C. W. OOSTERLEE AND T. WASHIO, *An evaluation of parallel multigrid as a solver and a preconditioner for singularly perturbed problems*, SIAM J. Sci. Comput., 19 (1998), pp. 87–110.
[11] A. PAPASTAVROU AND R. VERFÜRTH, *A posteriori error estimators for stationary convection-diffusion problems: A computational comparison*, Comput. Methods Appl. Mech. Engrg., 189 (2000), pp. 449–462.
[12] I. PERSSON, K. SAMUELSSON, AND A. SZEPESSY, *On the convergence of multigrid methods for flow problems*, Electron. Trans. Numer. Anal., 8 (1999), pp. 46–87.
[13] A. RAMAGE, *A multigrid preconditioner for stabilised discretizations of advection-diffusion problems*, J. Comput. Appl. Math., 110 (1999), pp. 187–203.
[14] A. REUSKEN, *Convergence analysis of a multigrid method for convection-diffusion equations*, Numer. Math., 91 (2002), pp. 323–349.
[15] M. C. RIVARA, *Algorithms for refining triangular grids suitable for adaptive and multigrid techniques*, Internat. J. Numer. Methods Engrg., 20 (1984), pp. 745–756.
[16] M. C. RIVARA, *Using longest-side bisection techniques for the automatic refinement of Delaunay triangulation*, Internat. J. Numer. Methods Engrg., 40 (1997), pp. 581–597.
[17] J. W. RUGE AND K. STÜBEN, *Algebraic multigrid*, in Multigrid Methods, Frontiers Appl. Math. 3, S. F. McCormick, ed., SIAM, Philadelphia, 1987, pp. 73–130.
[18] Y. SAAD AND M. H. SCHULTZ, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.
[19] U. TROTTENBERG, C. W. OOSTERLEE, AND A. SCHÜLLER, *Multigrid*, Academic Press, San Diego, CA, 2001.
[20] R. S. VARGA, *Matrix Iterative Analysis*, Springer-Verlag, Berlin, Heidelberg, 2000.
[21] R. VERFÜRTH, *A posteriori error estimators for convection-diffusion equations*, Numer. Math., 80 (1998), pp. 641–663.
[22] J. WANG, *Convergence analysis of multigrid algorithms for nonselfadjoint and indefinite elliptic problems*, SIAM J. Numer. Anal., 30 (1993), pp. 275–285.
[23] C.-T. WU, *On the Implementation of an Accurate and Efficient Solver for Convection-Diffusion Equations*, Ph.D. thesis, University of Maryland, College Park, MD, 2003.