

MULTISCALE NUMERICAL METHODS FOR SINGULARLY PERTURBED CONVECTION-DIFFUSION EQUATIONS

PETER J. PARK* and THOMAS Y. HOU†

*Applied and Computational Mathematics
California Institute of Technology, Pasadena, CA 91125, USA*

**pjpark@acm.caltech.edu*

†*hou@acm.caltech.edu*

Received 8 January 2004

Accepted 6 February 2004

We present an efficient and robust approach in the finite element framework for numerical solutions that exhibit multiscale behavior, with applications to singularly perturbed convection-diffusion problems. The first type of equation we study is the convection-dominated convection-diffusion equation, with periodic or random coefficients; the second type of equation is an elliptic equation with singularities due to discontinuous coefficients and non-smooth boundaries. In both cases, standard methods for purely hyperbolic or elliptic problems perform poorly due to sharp boundary and internal layers in the solution.

We propose a framework in which the finite element basis functions are designed to capture the local small-scale behavior correctly. When the structure of the layers can be determined locally, we apply the multiscale finite element method, in which we solve the corresponding homogeneous equation on each element to capture the small scale features of the differential operator. We demonstrate the effectiveness of this method by computing the enhanced diffusivity scaling for a passive scalar in the cellular flow. We also carry out the asymptotic error analysis for its convergence rate and perform numerical experiments for verification. For a random flow with nonlocal layer structure, we use a variational principle to gain additional information in our attempt to design asymptotic basis functions. We also apply the same framework for elliptic equations with discontinuous coefficients or non-smooth boundaries. In that case, we construct local basis function near singularities using infinite element method in order to resolve extreme singularity. Numerical results on problems with various singularities confirm the efficiency and accuracy of this approach.

Keywords: Finite element method; multiple scales; effective diffusivity; discontinuous coefficients; infinite element method.

*Present address: Harvard-Partners Center for Genetics and Genomics, NRB 255, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115, USA.

1. Introduction

Many phenomena in science and engineering exhibit multiscale behavior, i.e., the solutions to the partial differential equations that model the phenomena are characterized by more than a single scale. A simple example is the problem with a boundary layer in fluid mechanics, whose solution contains sharp features characterized by $f(\mathbf{x}/\epsilon)$ in the layer while the rest of the domain is characterized by $f(\mathbf{x})$, for some smooth function f and a small parameter ϵ . Another example is an elliptic equation with a highly oscillatory coefficient arising in material science or flow through porous media. Numerical solutions of these problems present many challenges. The basic difficulty is that small scale information must be resolved in order to obtain the correct coarse scale solution but solving for the small features in a globally coupled manner results in a computationally infeasible problem.

In this work, we present a different approach in the finite element framework, originally motivated by the Multiscale Finite Element Method [Hou *et al.* (1999); Hou and Wu (1997)]. Whereas a conventional method requires a grid that is fine enough to resolve all the small scales of a problem, the main idea behind the Multiscale Finite Element Method is to build the local behavior of the differential operator into the basis functions to capture the small scale effect while having a relative coarse grid over the whole domain. This is done by solving the equation on each element to obtain the basis functions, rather than using the linear basis functions. For elliptic problems with multiple scale solutions that result from rapidly oscillating coefficients, this method has been shown to be effective in obtaining globally accurate solutions [Hou *et al.* (1999); Hou and Wu (1997)]. Other related multiscale methods include those presented in Babuška *et al.* [1994]; Brezzi and Russo [1994]; Hughes [1995]; Dorobantu and Engquist [1998]; Matache *et al.* [2000]; Cao *et al.* [2002]; and E and Engquist [2003]. Here, we apply the idea of building special basis functions to capture the correct local behavior to other more difficult situations. We first focus on the singularly perturbed convection-diffusion equation, with periodic as well as random coefficients. Then we consider elliptic equations with discontinuous coefficients and non-smooth boundaries.

1.1. Convection-diffusion equation

Solving the singularly perturbed convection-diffusion equation is made difficult due to its dual nature. Because the diffusive term is multiplied by a small parameter ϵ ($0 < \epsilon \ll 1$), the convective effect dominates and the equation behaves essentially as a hyperbolic one in a large part of the domain; in a small region, however, diffusion becomes important and the balance between the convective and diffusive effects creates boundary or internal layers. These layers are usually exponential and require a fine grid for an accurate solution.

General results for convection-diffusion problems are well summarized in Morton [1996]. For the singularly perturbed case, some specialized numerical methods are described in Roos *et al.* [1991]; Miller *et al.* [1996]. However, most of the work has been confined to the one-dimensional or simple two-dimensional cases in which

the solution is essentially one-dimensional in its behavior. One method of dealing with this problem is the modification of an underlying mesh. For simple problems, optimal meshing and its properties are well known [Miller *et al.* (1996)]. For more complicated problems, general adaptive meshing algorithms using *a posteriori* error estimates are needed [Bank (1990)]. However, one difficulty with this method is in finding the solution of the linear system, as the grid points, which are added for refinement in adaptive meshing algorithms, are globally coupled to the regular mesh. The resulting linear system may also lack the banded structure or the positive definiteness of the uniform grids [Golub *et al.* (1996)]. The second class of methods for dealing with the layers is based on a modification of the differential operator. The simplest approach in this category is the upwinding method. This method does produce a stable numerical solution, but at the loss of accuracy, often $\mathcal{O}(h)$ accurate away from the layers but only $\mathcal{O}(1)$ near the layer [Kellog and Tsan (1978)]. Another way of suppressing spurious oscillations is the method of artificial viscosity. In two dimensions, it is the streamline-diffusion method in which viscosity is added only in the direction of characteristics [Johnson (1987); Douglas, Jr. and Russell (1982)]. But, again, the fronts are smeared and the accuracy is degraded. The global estimates are also not uniform with respect to the diffusion parameter ϵ , and another parameter must be tuned in order to add the right amount of artificial viscosity. Since the layers are typically exponential, another approach is to modify the differencing scheme in such a way that an exponential function is captured exactly rather than a polynomial one. This idea of “exponential fitting” is very old, dating back to 1950s [Allen and Sournthwell (1955)], and there have been many variants on this theme [Stynes and O’Riordan (1991); Hegarty *et al.* (1993); Lube (1992); Sacco and Stynes (1998); Roos *et al.* (1996)]. In the finite element setting, the finite element space can be modified with basis functions that have the exponential behavior. The difficulty, however, has been that it is hard to construct such basis functions in genuinely two-dimensional problems.

In many cases, the correct behavior of a solution can be determined by examining the equation *locally*. We study this case in the context of the enhanced effective diffusivity problem from fluid mechanics [Avellaneda and Majda (1991); Avellaneda (1991); Isichenko (1992)]. This provides an interesting physical background for examining the performance of our method. Specifically, we study the transport properties of the highly oscillatory but periodic cellular velocity field. The fluid in this flow is rotating in each of the small cells of size δ , with opposite directions in adjacent cells. For this problem, we find that the basis functions based on the local solution of the homogeneous equation perform well, resulting in the correct scaling of the overall diffusivity property of the flow. We carry out the error analysis of the multiscale finite elements applied to this case, using homogenization theory [Jikov *et al.* (1991); Bensoussan *et al.* (1978); Sanchez-Palencia (1980)]. We derive sharp estimates for the error δ/h , where δ is the small scale to be homogenized, that does not depend on the layer parameter ϵ . In contrast, the standard finite element method has δ on the denominator and fails to converge unless we have a very fine mesh so that $h \ll \delta$. Since resolving the basis

functions numerically can be expensive, we also demonstrate the possibility of using an asymptotic expansion.

When the characteristic length scale that governs the singular behavior is larger than the element size, the information contained in the support of each element is not sufficient to determine the local behavior correctly. Thus, in order to capture the correct behavior of the multiscale solution, we need to incorporate some information about the solution structure into our computational basis. For the random velocity field we consider, theoretical understanding of the flow field based on percolation theory [Fannjiang and Papanicolaou (1997)] and the variational principle introduced in Fannjiang and Papanicolaou [1994] can provide additional information. This variational principle is nonlocal and a test function must be evaluated on the whole domain in order to compute the energy integral. While the test function that minimizes the integral has a complicated layer structure and is difficult to find, we can construct a test function that captures the essential behavior of the layers and use this to extract the location and the thickness of boundary layer. Once we have the information regarding the layers, we build exponential basis function along the streamlines on which the layers occur, while the bilinear basis functions of the underlying uniform grid are still kept for the smooth part of the solution.

1.2. *Elliptic equation with discontinuous coefficients*

Another class of problems that give rise to singular behavior are elliptic equations with coefficients that have large variations or discontinuities in the domain of interest. For example, in describing a flow through porous media, the coefficient of the equation corresponds to the permeability of the medium and can vary significantly in different regions. In solving for electrical properties of composite materials, conductivity of the adjacent materials can differ by many orders of magnitude. The problem of singular coefficients can also be compounded by non-smooth interfaces. Under these circumstances, it is known that conventional numerical methods do not perform well [Babuška and Osborn (1985); Leveque and Li (1994)]. For certain singularities, for example, the standard finite element method can even be arbitrarily slow in its convergence [Rüde and Zenger (1986)].

A variety of mesh refinement strategies are available for dealing with problems with rapidly-varying solutions, using, for instance, *a posteriori* error estimates to create finer meshes where more resolution is needed. However, for the type of singular behavior considered in this paper, a typical mesh refinement scheme cannot provide enough resolution. A great deal is known about the analytical properties of elliptic operators (see, e.g., Gilbarg and Trudinger [1983]) and for certain cases, incorporating analytic approximations of the solution at the singularities can result in an efficient method [Strang and Fix (1973); Borcea and Papanicolaou (1997)]. Unfortunately, it is in general difficult to obtain an analytic expression or to incorporate into a numerical method when it is known.

As before, we construct basis functions that capture the effect of singularities. The singularities are essentially local, and the finite element framework using these

basis functions then translates to an accurate global solution. However, there is an additional difficulty: the singular behavior is such that we are not able to solve the equation locally for the basis functions using a conventional finite element method. The remedy we present here is the construction of the local basis functions using the infinite element method [Ying (1995)]. This elegant technique takes advantage of the similarity of its special grid structure and has the effect of having infinite resolution at the singularity. This method has been used previously for singularities in simple situations, for example with one singularity near the center of a polygon, or in infinite domain problems. We therefore employ the infinite elements as a part of a general framework. Computationally, the method can be implemented efficiently. The number of unknowns in the solution process is proportional to the number of points on the element boundary, while having the effect of infinite resolution at the singularity. We also derive a way of computing the stiffness matrix efficiently. We examine the multiscale problem with a corner and checkerboard singularities within each cell and find that the method performs very well. As the contrast increases, the singularity gets more localized and even for a moderate mesh size in the infinite element, the method is not sensitive to the high contrast.

In the next section, we describe some preliminary material. In Sec. 3, we apply the multiscale finite element method [Hou *et al.* (1999); Hou and Wu (1997); Efendiev *et al.* (2000)] to the convection-diffusion problem. We demonstrate its effectiveness for the periodic problem by using it to obtain the correct diffusivity scaling for the cellular flow, and we obtain a sharp convergence rate of the multiscale method using asymptotic error analysis, as verified by numerical experiments. We also apply this method to the random coefficient case, in which additional information from a variational principle is needed to obtain a correct scaling for a basis function. In Sec. 4, we apply the idea to the elliptic equation with discontinuous coefficient on non-smooth boundary using infinite element method. Through these examples, we demonstrate the usefulness of the multiscale approach using local basis functions that capture the small scale features of the differential operator.

2. Formulation of the Problem

We introduce some basic spaces before stating the problem in the weak form for the finite element method. Let $\Omega \subset \mathcal{R}^2$ be a bounded domain with a Lipschitz continuous boundary $\partial\Omega$. We will use $L^2(\Omega)$ based Sobolev space $H^k(\Omega)$, which is the space of all functions $u \in L^2(\Omega)$ whose derivatives $D^\alpha u$ (in the sense of distributions) of order $|\alpha| \leq k$ are also in $L^2(\Omega)$. This space is equipped with norms and seminorms

$$\|u\|_{k,\Omega} = \left(\int_{\Omega} \sum_{|\alpha| \leq k} |D^\alpha u|^2 \right)^{1/2}, \quad |u|_{k,\Omega} = \left(\int_{\Omega} \sum_{|\alpha|=k} |D^\alpha u|^2 \right)^{1/2},$$

$$\|u\|_{\infty,\Omega} = \operatorname{ess\,sup}_{x \in \Omega} |u(x)|.$$

The space $H_0^1(\Omega)$ is the closure of the set $C_0^\infty(\Omega)$ in $H^1(\Omega)$. With the sufficiently smooth boundary $\partial\Omega$, $H_0^1(\Omega)$ is the set of all functions u in $H^1(\Omega)$ such that $u = 0$ on $\partial\Omega$. $H^{-1}(\Omega)$ is the dual space of $H_0^1(\Omega)$, the set of all continuous linear functions on $H_0^1(\Omega)$. $H^{1/2}(\Omega)$ is the trace on $\partial\Omega$ of all functions in $H^1(\Omega)$ with the norm $\|v\|_{1/2,\partial\Omega} = \inf \|u\|_{1,\Omega}$ where the infimum is taken over all $u \in H^1(\Omega)$ with trace v . C denotes a generic positive constant independent of small parameters unless otherwise stated, with $C + C = C$ and $C \cdot C = C$. We say that a quantity p is $\mathcal{O}(q)$ when $|p| \leq Cq$ for all q sufficiently small. Repeated indices indicate summation over them.

2.1. Weak formulation

The general form of our model problem is

$$-\nabla \cdot (a(\mathbf{x})\nabla u) = f, \quad u = 0 \text{ on } \partial\Omega \quad (\Omega \in \mathcal{R}^2), \quad (1)$$

where $a(\mathbf{x}) = (a_{ij}(\mathbf{x}))$ satisfies the uniform ellipticity condition $\alpha|\xi_i|^2 \leq a_{ij}\xi_i\xi_j \leq \beta|\xi_i|^2$ ($0 < \alpha < \beta$) for any $\xi \in \mathcal{R}^2$.

The weak formulation of (1) is to find $u \in H_0^1$ such that

$$a(u, v) = f(v), \quad \forall v \in H_0^1, \quad (2)$$

where

$$a(u, v) = \int_{\Omega} a_{ij} \frac{\partial v}{\partial x_i} \frac{\partial u}{\partial x_j} dx, \quad f(v) = \int_{\Omega} f v dx. \quad (3)$$

In the finite element method (FEM), we apply the differential operator exactly but restrict the weak form (2) to a finite dimensional subspace of $H_0^1(\Omega)$: Find $u^h \in V^h$ such that

$$a(u^h, v^h) = f(v^h), \quad \forall v^h \in W^h, \quad (4)$$

where $V^h = \text{span}\{\phi_1, \dots, \phi_N\}$ and $W^h = \text{span}\{\psi_1, \dots, \psi_N\}$ with *trial* and *test* functions ϕ_i and ψ_i , respectively. V^h is called the trial function space and W^h the test function space. When $V^h = W^h$ ($\phi_i = \psi_i$), we refer to the method as the Galerkin method; when $V^h \neq W^h$, we refer to it as the Petrov-Galerkin method. We choose $v = v^h$ in (2) and subtracting (4) from (2), we have

$$a(u - u^h, v^h) = 0, \quad (5)$$

an orthogonality property that becomes very useful in the error analysis. Let u be the weak solution of (2) and the approximation u^h of (4). Then we have, according to the C\`ea's Lemma [Brenner and Scott (1994)], which guarantees that the weak form picks out the optimal v^h from V^h :

$$\|u - u^h\|_{1,\Omega} \leq \frac{\beta}{\alpha} \min_{v^h \in V^h} \|u - v^h\|_{1,\Omega}. \quad (6)$$

For the convection-diffusion equation, we have non-symmetric $a_{ij} = \epsilon\delta_{ij} + H_{ij}$ as shown later in (17). We assume that a_{ij} are in $L^\infty(\Omega)$ and $f \in H^{-1}(\Omega)$. We only consider the homogeneous Dirichlet boundary condition for clarity, but other boundary

conditions can be easily incorporated. A nonhomogeneous Dirichlet condition can be easily translated into the forcing term; Neumann boundary condition is automatically enforced by the variational form if Dirichlet condition is not imposed. We fix the domain to be $\Omega = (0, 1) \times (0, 1) \subset \mathcal{R}^2$ for computation. The existence and uniqueness of the solution in the weak form can be proved by the Lax-Milgram Lemma (see, e.g., Brenner and Scott [1994]). The coercivity condition is satisfied by incompressible velocity fields. Unfortunately, in the singularly perturbed case, the C ea’s Lemma is not too helpful. Explicit calculation of the constants gives a $1/\epsilon$ factor, which makes the estimate less useful [Park (2000)]. We pay careful attention to the error structure in Sec. 3.4 to derive a sharper result.

2.2. Homogenization

Homogenization is a way of extracting an equation for the coarse scale behavior that takes into account the effect of small scales [Bensoussan *et al.* (1978); Jikov *et al.* (1991)]. We are often content with the coarse scale behavior of the solution, but we cannot simply ignore the fine scales because the fine scales interact with other scales to affect the coarse grid solution. The homogenized or the “effective” equation contains no small scales and is therefore much easier to solve. Physically speaking, it is a method by which a problem in a heterogeneous medium consisting of a large number of periodic cells is approximated by a problem in a homogeneous medium. Here, we review the basic homogenization theory that will be important later. We use δ to characterize the rapid variation in the coefficient; we reserve ϵ for diffusivity.

We consider the divergence form of the equation,

$$-\nabla \cdot (a(\mathbf{x}, \mathbf{x}/\delta) \nabla u_\delta) = f \quad \text{in } \Omega, \quad u_\delta = g \quad \text{on } \partial\Omega, \quad (7)$$

where the coefficient $a(\mathbf{x}, \mathbf{y})$ is assumed to be periodic in \mathbf{y} , and $u_\delta \in H_0^1(\Omega)$ and $f \in H^{-1}(\Omega)$. The convection-diffusion equation can be written in this form, as described later. According to the homogenization theory, the solution u_δ has the following property of convergence,

$$u_\delta \rightarrow u_0 \quad \text{weakly in } H_0^1(\Omega), \quad (8)$$

$$a(\mathbf{x}, \mathbf{x}/\delta) \nabla u_\delta \rightarrow a^*(\mathbf{x}) \nabla u_0 \quad \text{weakly in } L^2(\Omega), \quad (9)$$

as $\delta \rightarrow 0$, where u_0 satisfies the homogenized equation

$$-\nabla \cdot a^*(\mathbf{x}) \nabla u_0 = f \quad \text{in } \Omega, \quad u_0 = g \quad \text{on } \partial\Omega. \quad (10)$$

The homogenized coefficient $a^*(\mathbf{x})$ does not contain the small scale δ .

It is natural to seek the first approximation in the form

$$u_\delta(\mathbf{x}) = u_0(\mathbf{x}) + \delta u_1(\mathbf{x}, \mathbf{x}/\delta) + \dots \quad (11)$$

where $u_1(\mathbf{x}, \mathbf{y})$ is periodic in \mathbf{y} with $\mathbf{y} = \mathbf{x}/\delta$. The justification for this expansion is given in Bensoussan *et al.* [1978]; Jikov *et al.* [1991]. Using multiple scale expansion, we find the homogenized coefficient to be

$$a_{ij}^*(\mathbf{x}) = \frac{1}{|Y|} \int_Y a_{ik} \left(\delta_{jk} + \frac{\partial \chi^j(\mathbf{x}, \mathbf{y})}{\partial y_k} \right) d\mathbf{y}, \quad (12)$$

where Y denotes the periodic cell of y variable and $|Y|$ is the volume of the cell, and where χ^k is the solution of the ‘‘cell problem’’

$$\frac{\partial}{\partial y_i} \left(a_{ij}(y) \frac{\partial \chi^k(\mathbf{x}, \mathbf{y})}{\partial y_j} \right) = - \frac{\partial}{\partial y_i} a_{ik}(\mathbf{x}, \mathbf{y}), \quad (13)$$

with periodic boundary condition in \mathbf{y} . The χ^k function is determined up to a constant; we impose $\langle \chi^k \rangle = 0$ to get a unique solution, where $\langle \cdot \rangle$ denotes the average over one period.

Now, since u_0 satisfies the boundary $u_\delta = g$ on the boundary $\partial\Omega$, $u_0 + \delta u_1$ does not. In order to enforce this boundary condition, we need to introduce a boundary ‘‘corrector’’ θ_δ , which satisfies

$$L_\delta \theta_\delta = 0 \quad \text{in } \Omega, \quad \theta_\delta = -u_1(\mathbf{x}, \mathbf{x}/\delta) \quad \text{on } \partial\Omega. \quad (14)$$

Putting this together, we have

$$u_\delta(\mathbf{x}) = u_0(\mathbf{x}) + \delta \chi^k \left(\mathbf{x}, \frac{\mathbf{x}}{\delta} \right) \frac{\partial u_0(\mathbf{x})}{\partial x_k} + \delta \theta_\delta + O(\delta^2). \quad (15)$$

This expansion is used in a later section.

3. The Convection-Diffusion Equation

3.1. Introduction

The convection-diffusion equation modeling the transport of the quantity $u(\mathbf{x}, t)$ is

$$\frac{\partial u(\mathbf{x}, t)}{\partial t} + \mathbf{b}(\mathbf{x}) \cdot \nabla u(\mathbf{x}, t) - \epsilon \Delta u(\mathbf{x}, t) = f(\mathbf{x}), \quad (16)$$

where $\mathbf{b}(\mathbf{x})$ is the given velocity field; ϵ ($0 < \epsilon \ll 1$) is the molecular diffusivity characterizing the Brownian motion; and f is the forcing on the system. The convection term can be written in a more general form as $\nabla \cdot (\mathbf{b}u)$ if we assume that the velocity fields are incompressible, $\nabla \cdot \mathbf{b} = 0$. This equation can be considered as the linearized version of the two-dimensional Navier-Stokes equations in vorticity formulation with viscosity $\epsilon = 1/Re$ (Re is the Reynolds number). We are interested in the convection-dominated case, with large $|\mathbf{b}|/\epsilon$; this is also referred to as the high Peclet number problem in fluid mechanics, where the Peclet number Pe is the ratio between the convective and the diffusive effects. If the velocity field is

obtained through a streamfunction ψ , that is, $\mathbf{b} = \nabla^\perp \psi = (-\psi_y, \psi_x)$, Eq. (16) can be written in the following divergence form:

$$\frac{\partial u(\mathbf{x}, t)}{\partial t} - \nabla \cdot (\epsilon \mathbf{I} + \mathbf{H}) \nabla u(\mathbf{x}, t) = f(\mathbf{x}), \quad (17)$$

with the skew-symmetric matrix

$$\mathbf{H} = \begin{pmatrix} 0 & -\psi \\ \psi & 0 \end{pmatrix}, \quad (18)$$

so that $\nabla \cdot \mathbf{H} = -\mathbf{b}$.

The limiting equation with $\epsilon \rightarrow 0$ is hyperbolic, and the equation qualitatively does not have many features that we would expect from an elliptic equation. In particular, solutions of (17) may have sharp boundary and internal layers. The boundary layer theory for the singularly perturbed problem has been studied for a long time, and there are some analytical techniques available, such as the matched asymptotics and multiple scale expansions. However, except for simple problems, the analytical approach is limited and we must turn to numerical schemes. Unfortunately, the standard numerical schemes have difficulties of their own. In the present work, we improve the numerical methods by combining them with some analytical understanding of the layer structure.

An important characteristic of (16) is that the operator is not self-adjoint, and we do not have the nice properties of the Sturm-Liouville type problems. The stiffness matrix of the finite element formulation is therefore nonsymmetric, requiring a different set of linear system solvers from the most commonly used ones. The matrix problem may already be ill-conditioned because of small ϵ . Also, the lack of self-adjointness creates additional difficulties in the formulation of a variational principle.

A common extension of successful one-dimensional methods to two-dimensional situations is the tensor product approach, in which the one-dimensional solution is used in each direction, i.e., $\phi_{ij}(x, y) = \phi_i(x)\phi_j(y)$. This works well for a small class of problems whose behavior is essentially one-dimensional, e.g., when $\mathbf{b}(\mathbf{x}) = (b_1(x), b_2(y))$ with $b_1(x) > b_0 > 0$ and $b_2(y) > b_0 > 0$ for some constant b_0 [Stynes and O'Riordan (1991); Hegarty *et al.* (1993)]. Other methods that work well in genuinely two-dimensional problems impose stringent conditions on the coefficients or require that much information is given about the layer structure in advance [Miller *et al.* (1996)]. For example, many methods require that the velocity field does not have a turning point where the coefficient switches the sign.

3.1.1. Rescaling

While Eq. (16) explicitly contains only one small parameter ϵ , another parameter enters when we want to consider the behavior of this equation in the large-domain and long-time limit. Rather than trying to solve the equation in successively large

domains for increasingly longer time, we introduce the rescaling parameter δ . We let

$$\mathbf{x}' = \mathbf{x}/\delta \quad \text{and} \quad t' = t/\delta^2, \quad (0 < \delta \ll 1), \quad (19)$$

which transforms (16) to, after dropping the apostrophe on \mathbf{x}' and t' ,

$$\frac{\partial u_\delta}{\partial t} + \frac{1}{\delta} \mathbf{b} \left(\frac{\mathbf{x}}{\delta} \right) \cdot \nabla u_\delta = \epsilon \Delta u_\delta. \quad (20)$$

We have also set $f = 0$, since there is no forcing in the problem we consider in this section. The analysis of this equation with periodic coefficients in the $\delta \rightarrow 0$ limit involves homogenization, which we reviewed in Sec. 2.2.

The homogenization theory applied to the convection-diffusion equation is the following [McLaughlin *et al.* (1985); Fannjiang and Papanicolaou (1994)]. As $\delta \rightarrow 0$, the solution to the rescaled equation u_δ converges to u , which satisfies the constant coefficient equation

$$\frac{\partial u}{\partial t} = \sigma_\epsilon \Delta u. \quad (21)$$

The convergence is in L^2 ,

$$\lim_{\delta \rightarrow 0} \sup_{0 \leq t \leq t_0} \int |u(\mathbf{x}, t) - u_\delta(\mathbf{x}, t)|^2 d\mathbf{x} = 0 \quad (22)$$

for any t_0 . Equation (21) means that in the large-domain, long-time limit, the overall behavior can be characterized as diffusive, with the convection term performing contributions of varying degree depending on the velocity field. How different velocity fields affect this “effective diffusivity” tensor σ_ϵ is of great interest and it is the test problem for the numerical methods we develop. Using homogenization theory, we can obtain an expression for this σ_ϵ [Fannjiang and Papanicolaou (1994)]:

$$\sigma_\epsilon(\mathbf{e}) = \epsilon \langle (\nabla \chi + \mathbf{e}) \cdot (\nabla \chi + \mathbf{e}) \rangle, \quad (23)$$

where $\langle \cdot \rangle$ denotes averaging over one period and \mathbf{e} is a unit vector. This comes directly from the expression for the homogenized coefficient (12). χ is the solution of the cell problem (13), which we can rewrite in a vector form as

$$\nabla \cdot [(\epsilon \mathbf{I} + \mathbf{H})(\nabla \chi + \mathbf{e})] = 0, \quad (24)$$

with \mathbf{H} as defined in (18). This can be simplified by using $\nabla \cdot \mathbf{H} = -\mathbf{b}$ to

$$-\epsilon \Delta \chi + \mathbf{b} \cdot \nabla \chi + \mathbf{b} \cdot \mathbf{e} = 0 \quad (25)$$

on the torus. σ_ϵ is generally a nonsymmetric matrix, but for the streamfunction we consider, there is a symmetry of the form $\mathbf{H}(\mathbf{x}) = -\mathbf{H}(-\mathbf{x})$ and this assures that the σ_ϵ tensor is symmetric [Fannjiang and Papanicolaou (1994)]. L^2 integrability of the streamfunction ψ is sufficient for the existence of this homogenization [Fannjiang and Papanicolaou (1994)].

3.1.2. Problems in one dimension

To gain insights into the behavior of the singularly perturbed equation, we first consider the simple one-dimensional problem,

$$-\epsilon u'' + a(x)u' = 0, \quad x \in [0, 1], \quad (26)$$

with Dirichlet boundary conditions $u(0) = 0$, $u(1) = 1$. If $a > 0$, the solution has an exponential layer of thickness $\mathcal{O}(\epsilon)$ near $x = 1$, in which the $\epsilon u''$ term becomes important. Intuitively, we can think of the u' term as the convective effect pushing the “fluid” to the right. If $a(x_0) = 0$ for some $x_0 \in [0, 1]$, x_0 is called a turning point and internal layers can occur.

Supposed $a(x) = a_0$. In the simple FEM, we expand $u(x) = \sum_1^N u_i \phi_i(x)$, where ϕ_i are the linear “hat” functions. The exact solution to the resulting discrete equations at the nodal points are

$$u_i = \frac{\gamma^i - 1}{\gamma^N - 1}, \quad \gamma = \frac{1 + ah/(2\epsilon)}{1 - ah/(2\epsilon)}.$$

We immediately see that unless $ah/\epsilon < 2$, oscillations will occur. The ratio ah/ϵ is often called the *mesh Peclet* number. This illustrates the common problem, that the mesh size h needs to be very small when ϵ is very small.

3.1.3. Green’s function approach

To deal with the present difficulty, many methods have been proposed. Among the most effective methods are the finite element methods with basis functions that contain the exponential behavior resembling the boundary layer of the solution [Roos *et al.* (1991)]. One can, for example, obtain the trial functions by solving the homogeneous equation modified by making the coefficient constant [Hegarty *et al.* (1993)]. In one dimension,

$$-\epsilon \phi_i'' + \bar{a} \phi_i' = 0, \quad \text{with } \bar{a} = (a(x_{i-1}) + a(x_i))/2, \quad \text{for } x \in (x_{i-1}, x_i). \quad (27)$$

The functions ϕ_i obtained this way have exponential layers. When these functions are summed up with correct weights, they provide a better approximation to the solution than the hat functions do. In fact, if the coefficient is constant, this scheme gives exact nodal values.

This can be understood in terms of Green’s functions. By definition, $u(\mathbf{x}) = \int_{\Omega} G(\mathbf{x}, \mathbf{x}_0) f(\mathbf{x}_0) dx_0$. Choosing $v(\mathbf{x}) = G(\mathbf{x}_0, \mathbf{x})$ in the bilinear form (2) and using the Green’s function, we have $a(u, G) = (f, G) = u$ for any choice of u . Therefore, letting $u \rightarrow u - u^h$, we can now write $a(u - u^h, G) = u - u^h$. Subtracting this from the orthogonality condition $a(u - u^h, v) = 0$,

$$u - u^h = a(u - u^h, G - v), \quad \forall v \in V^h. \quad (28)$$

This shows that the error $\|u - u^h\|$ can be minimized by selecting the test function space V^h to contain as much of the Green's functions as possible. Given this reasoning, it makes sense to choose V^h to include functions G_i that satisfy

$$L^*G_i(\mathbf{x}) = -\delta(\mathbf{x} - \mathbf{x}_i)$$

at each mesh point \mathbf{x}_i . L^* is the adjoint operator here.

Fortunately, in one dimension, G_i is a linear combination of local Green's functions $g_j(x)$, $x \in (x_{j-1}, x_{j+1})$, that satisfy

$$L^*g_j(x) = 0, \quad g_j(x_{j-1}) = g_j(x_{j+1}) = 0, \quad g_j(x_j) = 1. \quad (29)$$

For variable coefficient problems, previously proposed methods suggest solving a local problem with averaged coefficients. However, as we will also do in two dimensions, we solve the variable coefficient problem directly. It can be shown that for the one-dimensional equation (26) with a variable coefficient convection term, using linear trial functions and solutions to the local homogeneous adjoint equation as the test functions in the weak formulation results in exact nodal values, if the integrations are done exactly [Park (2000)]. If there is no forcing, adaptive trial functions and linear test functions also give exact nodal values.

In two dimensions, the global Green's function cannot be expressed as a linear combination of local Green's functions solved in each element. However, the adjoint equation is still useful. After integration by parts and some algebra, we can write

$$a(v, w) = \epsilon \sum_i \int_{\partial K} v \nabla w \cdot \mathbf{n} ds + \epsilon \sum_i \int_K (-\Delta w dx - \mathbf{b} \cdot \nabla w) v dx, \quad (30)$$

over elements denoted by K . Therefore, when we get w by solving the adjoint equation $-\epsilon \Delta w - \mathbf{b} \cdot \nabla w = 0$ on each element, the second term disappears. We then let $v = u - u^h$ and use the orthogonality property to get

$$\sum_i \int_{\partial K} (u - u^h) \nabla w \cdot \mathbf{n} ds = 0,$$

that is, the projection of the error $u - u^h$ onto the element boundary becomes zero.

3.2. The multiscale finite element method

In Hou *et al.* [1999], the multiscale FEM was introduced for elliptic equations with oscillatory coefficients. The main idea is to resolve the fine scales locally within each element of size $h \gg \delta$ by solving the homogeneous equation with some appropriate boundary conditions. This way, each basis function retains the oscillatory property of the differential operator. It is proved [Hou *et al.* (1999)] that when these elements are used to construct the global stiffness matrix, the averaged effects of the rapidly oscillating coefficients are correctly captured. By resolving the fine details inside the elements, problems that are prohibitively expensive with conventional finite element methods are broken down into smaller, manageable parts. Because the

elements are larger, the final solution is computed on a coarser grid; however, the large scale features are usually of interest in the first place, and sufficient information is obtained for that purpose. In this method, the *trial* functions are the solutions of the homogeneous equation and the test functions are some continuous functions, different from the approach suggested by the Green's functions. But both methods are similar and give comparable results in most cases.

Formally, the multiscale FEM formulation is the following. We let \mathbf{K}^h be a partition of Ω of triangles or rectangles with diameter less than h . We define a set of basis $\phi_{\delta,K}^i, i = 1, \dots, d$ ($d = 3$ for triangles and $d = 4$ for rectangles) such that $\phi_{\delta,K}^i$ satisfies

$$L_{\delta}\phi_{\delta}^i = 0 \quad \text{in } K. \quad (31)$$

At the nodal points $x_j \in K, j = 1, \dots, d$, we require $\phi_{\delta,K}^i(x_j) = \delta_{ij}$ as usual. The correct boundary conditions for (31) would match the global solution, but since we do not know this, we can impose linear boundary conditions for now; this issue is discussed later. We define

$$V^h = \text{span}\{\phi_{\delta,K}^i: i = 1, \dots, d, K \subset \mathbf{K}^h\} \subset H_0^1(\Omega),$$

and the Galerkin formulation is to find the solution $u_{\delta}^h \in V^h$ such that

$$a(u_{\delta}^h, v) = f(v) \quad \forall v \in V^h. \quad (32)$$

In order to understand the convergence of this method, the ideas from the homogenization procedure discussed in Sec. 2.2 play a critical role. Just as we expand u_{δ} in (11), we can also expand the basis function ϕ_{δ} as

$$\phi_{\delta}(\mathbf{x}) = \phi_0(\mathbf{x}) + \delta\chi^k \left(\mathbf{x}, \frac{\mathbf{x}}{\delta} \right) \frac{\partial\phi_0}{\partial x_k} + \delta\theta^{\delta}, \quad (33)$$

with ϕ_0^i, χ^k , and θ^{δ} defined similarly to (10), (13), and (14). Since ϕ_0 is a smooth function, we see that the oscillation in ϕ_{δ} comes from that of χ , which solves the cell problem (25). The error analysis is based on the fact that ϕ_{δ} and u_{δ} satisfy the same operator and therefore the two expansions (15) and (33) match, away from the small region near the boundary.

We see that the boundary conditions are an important issue if the two expansions were to match better near the boundary. While we assumed linear boundary condition above, another possibility is to solve the one-dimensional version of the equation along each edge, as we do for the problems in this section. This usually results in some improvement in error. For very small ϵ , care must be taken: the layers inside the basis may have thickness of $\mathcal{O}(\delta\sqrt{\epsilon})$, but the reduced equation on the boundary may give layers of thickness $\mathcal{O}(\delta\epsilon)$. In that case, the mesh spacing that resolves the layers inside the basis may not resolve the layers on the boundary. The case in which solving the one-dimensional equation works the best is in Sec. 4, when coefficients have discontinuities. For many elliptic problems, the best solution turns out to be the oversampling method introduced in Hou and Wu [1997]. In that

case, the effect of a wrong boundary condition is restricted to a $\mathcal{O}(\delta)$ region near the boundary. The oversampling idea is then to solve for the basis function on a domain larger than the element and extract the information from the middle of the domain. This reduces the error due to the incorrect boundary conditions. A rigorous analysis has been carried out in Efendiev *et al.* [2000]. In the convection-diffusion case, the oversampling does not always work because the effect of the boundary condition may not be confined to such a small region.

3.3. Computation of the effective diffusivity

We apply the multiscale FEM discussed in the previous section to the time-dependent convection-diffusion problem with the periodic velocity field. In particular, we consider the “cellular” flow for which some analytical results exist. The problem we use to demonstrate the effectiveness of the method is the computation of the effective diffusivity property (23).

3.3.1. Effective diffusivity

The “cellular flow” is given by the streamfunction

$$\psi = \frac{1}{4\pi^2} \sin\left(\frac{2\pi x}{\delta}\right) \sin\left(\frac{2\pi y}{\delta}\right), \quad (x, y) \in [0, 1]^2, \quad (34)$$

where δ is the rescaling parameter that determines the number of cells. The scaling factor $1/(4\pi^2)$ is used to relate to other studies done on domains of size $[-\pi, \pi]^2$. This streamfunction with $\delta = 0.25$ is plotted in Fig. 1(a). The velocity vector is tangent to the streamlines shown.

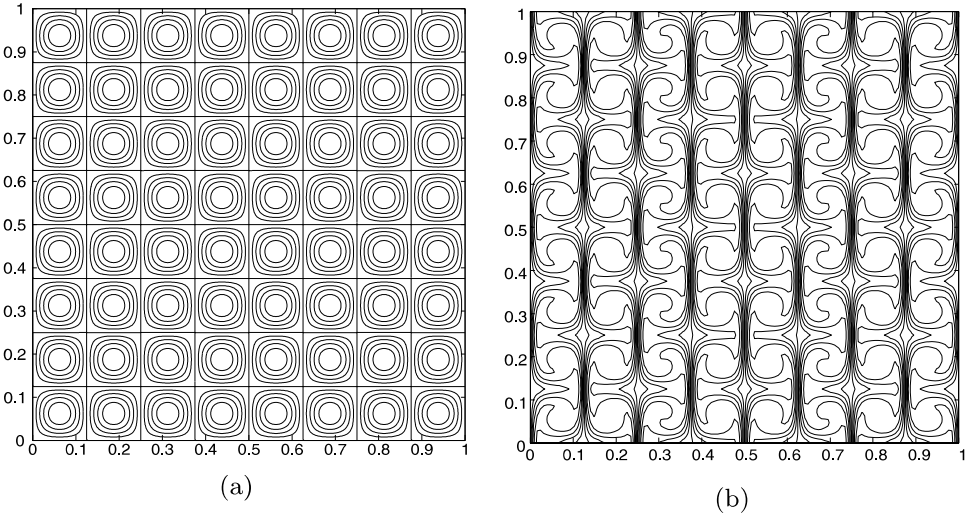


Fig. 1. (a) shows a periodic streamfunction ($\delta = 0.25$); (b) shows χ for the cellular flow ($\delta = 0.25$, $\epsilon = 0.001$).

A well-known result for this velocity field, described for example in Childress and Soward [1989]; Childress [1979]; Rosenbluth *et al.* [1987], concerns the long-time, large-distance diffusive behavior: the “effective diffusivity” σ_ϵ , in the limit $\epsilon \rightarrow 0$, scales as

$$\sigma_\epsilon \sim \sqrt{\epsilon}. \quad (35)$$

In the absence of convection, the effective diffusivity is simply ϵ . It is clear that convection can only increase the overall diffusivity, as it carries particles along the streamlines faster than diffusion does. This can be seen clearly when we expand the expression for σ_ϵ in (23) and simplify to get another expression,

$$\sigma_\epsilon = \epsilon + \epsilon \langle \nabla \chi \cdot \nabla \chi \rangle. \quad (36)$$

The result (35) can be understood through a simple scaling argument. The important feature in the solution is the formation of boundary layers near separatrices. It is these boundary layers that characterize the transport properties of particles. To determine the width of the layer along the separatrices, we set $\delta = 1$ and balance the diffusive flux across the layer with the convective flux along the layer. Comparing the time scales, $w^2/\epsilon \sim l/u_0$, where w is the width of the layer, l is the size of the cell, and u_0 is the magnitude of the velocity. Since l and u_0 are of $\mathcal{O}(1)$, we conclude that $w \sim \sqrt{\epsilon}$. Then we can use (23). Since the width of the cell boundary layer is of order $w \sim \sqrt{\epsilon}$, we obtain $\nabla \chi \sim 1/\sqrt{\epsilon}$. Substituting this in (23) and integrating over the width of the layer immediately gives (35).

In general, σ_ϵ is a tensor. It is the asymptotic rate of mean square displacement with different diffusivity depending on the direction. In the present case, however, the effective diffusivity is isotropic. This allows us to introduce a simple and intuitive definition, also useful for computations. With a slight modification from Fannjiang and Papanicolaou [1994], we can measure the mean square displacement by

$$\sigma_\epsilon = \lim_{t \rightarrow \infty} \frac{1}{4t} \iint (x^2 + y^2) u(x, y, t) dx dy \quad (37)$$

with the delta function at the origin as the initial function. We have inserted the factor 4 so that when there is no convection term, we get $\sigma_\epsilon = \epsilon$, which we can verify by putting in the Green’s function for the heat equation into (37).

3.3.2. Numerical results

In order to test the multiscale FEM, we compute the equation (20) in time, starting with a regularized delta function. An example of a multiscale basis function obtained by solving (31) is shown in Fig. 2, to illustrate the fine details within the basis. The basis function has the same layer structure as we would find in the global solution. The boundary conditions given are 1 at the lower left corner and 0 at the other corners; along the edges, the reduced one-dimensional equations are solved. Layers are strong at the lower left region because of this boundary condition.

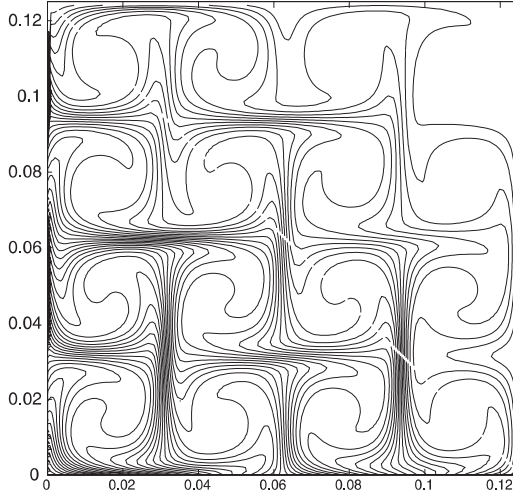


Fig. 2. A contour plot of the multiscale basis function, with 1 at the lower left corner and 0 at the other nodes. In this example, $\delta/h = 0.5$, where h is the mesh size.

Table 1. The diffusivity scaling for the cellular flow ($n = 16, m = 32$).

| ϵ | $\delta = 1$ | | $\delta = 0.1$ | | $\delta = 0.05$ | |
|------------|-------------------|--------|-------------------|--------|-------------------|--------|
| | σ_ϵ | Ratio | σ_ϵ | Ratio | σ_ϵ | Ratio |
| 0.04 | 0.03252 | — | 0.03265 | — | 0.03275 | — |
| 0.02 | 0.02092 | 0.6432 | 0.02316 | 0.7093 | 0.02355 | 0.7191 |
| 0.01 | 0.01141 | 0.5454 | 0.01643 | 0.7094 | 0.01665 | 0.7070 |
| 0.005 | 0.00619 | 0.5425 | 0.01300 | 0.7912 | 0.01233 | 0.7405 |
| 0.0025 | 0.00344 | 0.5557 | 0.01046 | 0.8046 | 0.00872 | 0.7072 |

For computation, we let $u(x, y, t) = \sum_j \xi_j(t) \phi_j(x, y)$ and $v = \phi_i$ in the weak formulation. A good discretization scheme for this problem is second order Adams-Bashforth on the advection term and Crank-Nicholson on the diffusion term. We can approximate the initial condition with a Gaussian or another smooth function such as

$$u(x, y) = \begin{cases} \frac{1}{16} (\sin(4\pi(r + 1/8)) + 1)^2, & r < 0.25 \\ 0, & r > 0.25 \end{cases},$$

where r is the radius. We restrict \mathcal{R}^2 to a finite domain and then scale it to the computational domain of the unit square using δ ; we take small Δt and compute up to some time T , depending on ϵ and δ . When σ_ϵ starts to decrease noticeably, we know that the fluid front has reached the boundary but is prevented from moving further. We compute the scaling by comparing the values of σ_ϵ at a fixed time for different ϵ . In Table 1, we show the rate at which σ_ϵ in (37) is changing. Given the scaling (35), we expect σ_ϵ to decrease by a factor of $1/\sqrt{2} \approx 0.7071$ when ϵ is

halved. We see in the table that when $\delta = 1$, σ_ϵ is nearly halved, meaning there is no enhancement. However, when $\delta = 0.05$, the ratio is close to the predicted value, as ϵ gets small. If the conventional bilinear elements are used, this behavior is missed completely.

We note that while we compute the effective diffusivity here, that is not the main objective of the method. The objective is to compute the transient state correctly. The fact that we get the correct scaling is just a consequence of obtaining the correct solution. For this convection-diffusion problem, the multiscale FEM offers a big computational saving over the conventional FEM. Since the velocity field is time independent, the cost in computing the multiscale base function is just an overhead at $t = 0$. Once we have computed the multiscale bases, the subsequent computations at later times can be performed on the coarse grid with a coarse time step. For the particular example considered here, the ratio between the coarse grid and the fine grid is equal to 32 in each dimension. Thus there is a factor of $32 \times 32 = 1024$ between the number of coarse grid and the number of fine grid. If we take into the account of saving in the number of time step, we can gain a factor 10 000s in computational cost comparing with the corresponding fine grid computation.

In the case of homogeneous periodic velocity field as the one considered here, we can further reduce the cost in computing the multiscale bases by solving for the cell problem $\chi(\mathbf{y})$ in a single periodic cell and use the first two terms in the asymptotic expansion of the base function in (33) as our multiscale basis, i.e.

$$\phi_\delta(\mathbf{x}) = \phi_0(\mathbf{x}) + \delta \chi^k \left(\frac{\mathbf{x}}{\delta} \right) \frac{\partial \phi_0}{\partial x_k}, \quad (38)$$

where $\phi_0(\mathbf{x})$ can be taken as the standard FEM base over a rectangular element, e.g. the bilinear base. In this case, the solution of the cell problem χ^k is independent of slow variable \mathbf{x} since the streamline is independent of \mathbf{x} . Therefore, by solving for the cell problem in a single periodic cell, we can generate the multiscale base functions for the entire computational domain.

Multigrid solver

One of the major difficulties for the singularly perturbed problems is solving the linear system. For small ϵ , the linear system may be indefinite and often ill-conditioned. This means some iterative techniques would not converge. It was shown in Golub *et al.* [1996] that the positive definiteness of the continuous differential operator is not always mirrored by some discretization schemes, even if the matrix is diagonally dominant. As a result, the standard multigrid algorithm, which is an $\mathcal{O}(N)$ algorithm where N is the total number of unknowns, often performs poorly. Fortunately, there is a modified multigrid algorithm developed specifically for convection-dominated problems [De Zeeuw (1990)]. It uses matrix-dependent prolongation and restriction operators that account for the character of the equation. We have found that this algorithm is very efficient and robust, usually converging under 10 or 20 iterations to the residual of 10^{-8} . Iterative methods such as BiCGSTAB or GMRES

for nonsymmetric matrices also work, but the number of iterations required may be large as ϵ gets small.

Parallel efficiency

A major advantage of the multiscale FEM algorithm is its parallel efficiency. The construction of one basis function is independent from that of any other; the elements can be divided evenly among all the processors first, and then collected to a master processor that performs the final calculations. Hence, the algorithm is almost perfectly parallelized: As the number of processors doubles, the computing time is halved. The time is not exactly halved since the master process has to do the final step after the global stiffness matrix is computed. Only when the number of processors is very large is the efficiency degraded, as the communication time among the processors takes a significant portion of the total time.

There is also a substantial saving in the memory required. We let n be the number of multiscale elements within each direction and let m be the number of grid points in each element. Then, the total memory required is $\mathcal{O}(n^2 + m^2)$, since $\mathcal{O}(m^2)$ operations for each element can be done in sequence. If the same resolution were to be achieved in direct simulation, $\mathcal{O}((nm)^2)$ would be required. If $n = 32, m = 32$, for example, there is a factor of 1000 saving in memory.

3.4. Asymptotic error analysis

We study the error of the multiscale method in this section. In the genuinely elliptic problem considered before [Hou *et al.* (1999); Efendiev *et al.* (2000)], there is one parameter δ , which characterizes the rapid oscillation. Now there is an additional parameter ϵ for thin layers that must be considered in examining the error terms.

3.4.1. Previous results

We first discuss the result for the symmetric coefficient, $a_{ij} = a_{ji}$, studied in detail in Hou *et al.* [1999], as it provides the framework for the analysis of this section. When the small scale of the problem δ is resolved by the mesh size h , the error estimate is

$$\|u - u^h\|_{0,\Omega} \leq C \left(\frac{h}{\delta}\right)^2 \|f\|_{0,\Omega}, \quad (h \ll \delta). \quad (39)$$

In this case, the multiscale basis functions do not have oscillations and they look similar to the bilinear functions. Thus, the convergence rate resembles that of the standard FEM. Note that we must have h small enough to resolve the δ scale for convergence. The two methods are different for the case of our interest, $h \gg \delta$. In Hou *et al.* [1999], estimates from homogenization theory are used to show that

$$\|u - u^h\|_{1,\Omega} \leq C_1 \left(\frac{\delta}{h}\right)^{1/2} + C_2 h \|f\|_{0,\Omega}, \quad (h \gg \delta). \quad (40)$$

Note that $\delta \ll 1$ now appears in the numerator of the first term. The method therefore converges as $\delta \rightarrow 0$ for a fixed h , unlike in the $h \ll \delta$ case (39). To get the L^2 estimate, the Aubin-Nitsche trick [Brenner and Scott (1994)] is usually employed to gain an extra order in h from the H^1 estimate. This method, however, does not work well for this problem and the $(\delta/h)^{1/2}$ term still remains.

In order to obtain the correct estimate as indicated by numerical experiments, subtle error cancellations in the discrete problem need to be examined. With this analysis, it is formally concluded that

$$\|u - u^h\|_{0,\Omega} \leq C_1 \frac{\delta}{h} + C_2 \delta + C_3 h^2 \|f\|_{0,\Omega}, \quad (h \gg \delta). \quad (41)$$

The leading order term is now (δ/h) and as long as h is large enough, the error is small, as verified by numerical examples [Hou *et al.* (1999)].

Discrete analysis

We review this analysis of the discrete problem [Hou *et al.* (1999); Efendiev *et al.* (2000)], so that we can understand the convection-diffusion case.

Using the triangular inequality, standard finite element estimates for bilinear elements, and the regularity and homogenization estimates [Hou *et al.* (1999)], we can write

$$\|u - u^h\|_{0,\Omega} \leq \|u - u_0\|_{0,\Omega} + \|u_0 - u_0^h\|_{0,\Omega} + \|u^h - u_0^h\|_{0,\Omega} \quad (42)$$

$$\leq C_1 \delta + C_2 h^2 \|f\|_{0,\Omega} + \|u^h - u_0^h\|_{0,\Omega}, \quad (43)$$

where u is the solution to the original continuous problem, u^h the numerical approximation to u given by (32); u_0 is the solution to the homogenized problem (10), u_0^h the numerical approximation to u_0 . We would obtain u_0^h by solving the bilinear form (2) with the homogenized coefficient.

The problem then is to estimate the $\|u^h - u_0^h\|_{0,\Omega}$ term in (43). To examine this term, we introduce the discrete l^2 norm. It is shown in Hou *et al.* [1999] that

$$\|u^h - u_0^h\|_{L^2} \leq C_1 \|u^h - u_0^h\|_{l^2} + C_2 \delta, \quad (44)$$

where

$$\|u^h - u_0^h\|_{l^2} = \left(\sum_{i \in \mathcal{N}} [u^h(\mathbf{x}_i) - u_0^h(\mathbf{x}_i)] h^2 \right)^{1/2}, \quad (45)$$

with \mathcal{N} containing all the nodal points of the mesh. Equations (43) and (44) show that $\|u^h - u_0^h\|_{l^2}$, the convergence of u^h to u_0^h at the nodal points, contains the crucial error term.

Let U^h be the vector containing the nodal points u^h . This is the solution to the discrete equation

$$A^h U^h = f^h, \quad (46)$$

where A^h and f^h are the global stiffness matrix and the load vector, obtained from (32) using $v = \phi_\delta^h$.

Similarly, we have for the homogenized problem,

$$A_0^h U_0^h = f_0^h, \quad (47)$$

where we obtain A_0^h and f_0^h from the bilinear form for (10) using $v = \phi_0^i$.

Since the basis function ϕ_i can be expanded as (33), we can also expand the stiffness matrix and load vector around the homogenized counterparts as

$$A^h = A_0^h + \delta A_1^h + \dots, \quad f^h = f_0^h + \delta f_1^h + \dots. \quad (48)$$

A_1^h is assembled from the local stiffness matrix of each element, which we denote by e . By substituting the expansion $\phi_\delta = \phi_0 + \delta \chi^k \partial \phi_0 / \partial x_k + \delta \theta^\delta$ (33) in the bilinear form (2), we find, after some algebra, that

$$A_{1_{kl}}^e = - \int_K \sigma^{ij} (\phi_{0,j}^k \theta_{,i}^l + \phi_{0,j}^l \theta_{,i}^k) dx + \int_K \delta a_\delta^{ij} \theta_{,i}^k \theta_{,j}^l dx + \frac{1}{\delta} \int_K \tilde{\sigma}^{ij} \phi_{0,j}^l \phi_{0,i}^k dx \quad (49)$$

and

$$f_{1_i}^e = - \int_K f (\chi^j \phi_{0,j}^i + \theta^i) dx, \quad (50)$$

where the comma is a shorthand for partial differentiation. Here,

$$\sigma^{ij} = a^{ik} \left(\delta_{jk} + \frac{\partial \chi^k}{\partial y_j} \right) \quad (51)$$

and

$$\tilde{\sigma}^{ij} = \sigma^{ij} - a_*^{ij} - \sigma^{kj} \chi_{,y_p}^i. \quad (52)$$

From (12) and (13), we find that $\langle \sigma^{ij} \rangle$ is the homogenized coefficient a_*^{ij} and $\sigma_{,y_i}^{ij} = 0$. Integrating by parts, we also obtain $\langle \sigma^{pj} \chi_{,y_p}^i \rangle = 0$ and hence $\langle \tilde{\sigma}^{ij} \rangle = 0$ [Efendiev *et al.* (2000)].

Given the expansions (48), we can deduce an expansion in U^h ,

$$U^h = U_0^h + \delta U_1^h + \delta^2 U_2^h + \dots. \quad (53)$$

In particular, U_1^h is given by

$$A_0^h U_1^h = f_1^h - A_1^h U_0^h. \quad (54)$$

Furthermore, we can express U_1^h as follows:

$$U_1^h = G_0^h f_1^h - G_0^h A_1^h U_0^h, \quad (55)$$

where $G_0^h = (A_0^h)^{-1}$.

Now the whole analysis reduces to estimating the order of U_1^h , since $\|U^h - U_0^h\| \leq \delta \|U_1^h\| + \delta^2 \|U_2\|^2 + \dots$. With $\|U_1^h\|$, (43) and (44) directly lead to the desired estimate of $\|u - u^h\|_{0,\Omega}$.

3.4.2. The convection-diffusion case

We now return to the convection-diffusion case, where we have, in place of (43),

$$\|u - u^h\|_{0,\Omega} \leq \|u - u_0\|_{0,\Omega} + \|u_0 - u_0^h\|_{0,\Omega} + \|u^h - u_0^h\|_{0,\Omega} \quad (56)$$

$$\leq C_1 \frac{\delta}{\sqrt{\epsilon}} + C_2 \frac{h^2}{\sqrt{\epsilon}} \|f\|_{0,\Omega} + \|u^h - u_0^h\|_{0,\Omega}. \quad (57)$$

The $1/\sqrt{\epsilon}$ is due to the fact that the norm of the solution grows at the rate of $\sqrt{\epsilon}$, as we will verify later. In the symmetric case, $a_{ij} = a_{ji}$, it was shown in Efendiev *et al.* [2000] that we can write the second term for U_1^h in (55) in a different form:

$$(G_0^h A_1^h U_0^h)_i = \sum_{k=1}^N \sum_{j=1}^{j < k} A_1^{kj} (U_0^j - U_0^k) (G_0^{ik} - G_0^{ij}); \quad (58)$$

$$U_1^h = G_0^h f_1^h - DG_0^h A_1^h DU_0^h. \quad (59)$$

D is a difference operator, details of which can be found in Efendiev *et al.* [2000]. The fact that we can write DG_0^h and DU_0^h here is crucial because this difference structure gives an additional $\mathcal{O}(h)$ in the convergence rate. With the estimates $\|G_0^h\|_{0,\Omega} \leq C/h^2$, $|f_1^h| \leq C_1\delta + C_2h$, and $\|A_1^h\|_{0,\Omega} \leq C/h$, this leads to $\|U_1^h\| \leq C_1 + C_2/h$ in the symmetric case and hence to (41).

In the convection-diffusion case, we do not have the same structure. After some algebra and discrete integration by parts, we have, instead of (59),

$$(G_0^h A_1^h U_0^h)_i = \sum_{k=1}^N \sum_{j=1}^{j < k} A_1^{kj} (U_0^j - U_0^k) (G_0^{ik} + G_0^{ij}). \quad (60)$$

Because of the plus sign between the G_0 terms, we can write this only as

$$U_1^h = G_0^h f_1^h - G_0^h A_1^h DU_0^h, \quad (61)$$

without the difference operator for G_0^h term. The fact that we lost $\mathcal{O}(h)$ means that the structure of (55) must be examined more carefully for possible additional cancellation.

3.4.3. Estimating A_1^h

We do this by carefully estimating the A_1^h term in (55). We denote the terms in (49) as follows:

$$\Lambda_1 = - \int_K \sigma^{ij} (\phi_{0,j}^k \theta_{,i}^l + \phi_{0,j}^l \theta_{,i}^k) dx, \quad (62)$$

$$\Lambda_2 = \delta \int_K a_\delta^{ij} \theta_{,i}^k \theta_{,j}^l dx, \quad (63)$$

$$\Lambda_3 = \frac{1}{\delta} \int_K \tilde{\sigma}^{ij} \phi_{0,j}^l \phi_{0,i}^k dx. \quad (64)$$

We now proceed to examine each of these terms.

(i) Estimating Λ_1 :

We first move the integral to the boundary to avoid estimating θ_i^l inside an element and use the fact that $\theta^l = \chi^l(\partial\phi_0/\partial x)$ on the boundary from (14) and that $\partial\phi_0/\partial x \sim 1/h$ in a cell of size h . Thus,

$$\int_K \sigma^{ij} \phi_{0,j}^k \theta_i^l dx = \int_{\partial K} \sigma^{ij} \phi_{0,j}^k \theta^l n_i ds \quad (65)$$

$$\sim \frac{1}{h} \int_{\partial K} (a^{ij} + a^{ip} \chi_{,y_p}^j) \chi^l n_i ds \quad (66)$$

$$= \frac{1}{h} \int_{\partial K} a^{ij} \chi^l n_i ds + \frac{1}{h} \int_{\partial K} a^{ip} \chi_{,y_p}^j \chi^l n_i ds. \quad (67)$$

The first term of (67) is $\mathcal{O}(1)$ since a^{ij} and χ^l are bounded. The second term is more difficult to estimate. Note that these terms are indexed by k, l and summed in other indices. Some terms are dropped once their sizes have been estimated, but the same indices should still be summed even though they may not be repeated. Now we group the boundary segments for the second term into two parts, with Γ_i denoting the four sides of an element K , consecutively numbered. Now (67) is the same order as

$$\begin{aligned} & \frac{1}{h} \left(\int_{\Gamma_1+\Gamma_3} a^{1k} \chi_{,y_k}^j \chi^l n_1 ds + \int_{\Gamma_2+\Gamma_4} a^{2k} \chi_{,y_k}^j \chi^l n_2 ds \right) \\ &= \frac{1}{h} \left(\int_{\Gamma_1+\Gamma_3} a^{11} \chi_{,y_1}^j \chi^l n_1 ds + \int_{\Gamma_1+\Gamma_3} a^{12} \chi_{,y_2}^j \chi^l n_1 ds \right. \\ & \quad \left. + \int_{\Gamma_2+\Gamma_4} a^{21} \chi_{,y_1}^j \chi^l n_2 ds + \int_{\Gamma_2+\Gamma_4} a^{22} \chi_{,y_2}^j \chi^l n_2 ds \right). \quad (68) \end{aligned}$$

We consider the $j = 1$ case now; the $j = 2$ case is analogous, except that the role of the second and third terms in (68) are reversed. The first and the last terms present little problem since the coefficient is $a^{11} = a^{22} = \epsilon$, which comes from the diffusion term and cancels out the effect of the χ term.

The second and third term appear more troublesome. χ satisfies the singularly perturbed problem (25) and so has layers. But we observe that $\nabla\chi$ is $\mathcal{O}(1/\sqrt{\epsilon})$, that is, the thickness of the layer in χ is $\mathcal{O}(\sqrt{\epsilon})$, as will be discussed soon. Since the streamfunction ψ is a smooth function and the layers occur around the separatrices $\psi = 0$, ψ is of $\mathcal{O}(\sqrt{\epsilon})$ in the layer. Thus, (68) can be estimated as

$$\epsilon \cdot \frac{1}{\sqrt{\epsilon}} + \int_{\Gamma_1+\Gamma_3} \psi \chi_{,y_2}^1 ds + \int_{\Gamma_2+\Gamma_4} \psi \chi_{,y_1}^1 ds + \epsilon \cdot \frac{1}{\sqrt{\epsilon}} \quad (69)$$

$$\sim \epsilon \cdot \frac{1}{\sqrt{\epsilon}} + \sqrt{\epsilon} + \sqrt{\epsilon} \cdot \frac{1}{\sqrt{\epsilon}} + \epsilon \cdot \frac{1}{\sqrt{\epsilon}} \quad (70)$$

$$\sim \mathcal{O}(1). \quad (71)$$

We plot the χ^1 and χ^2 functions in Fig. 3 to clarify the terms in (69). We note, however, that the domain in the figure is over one period and not over

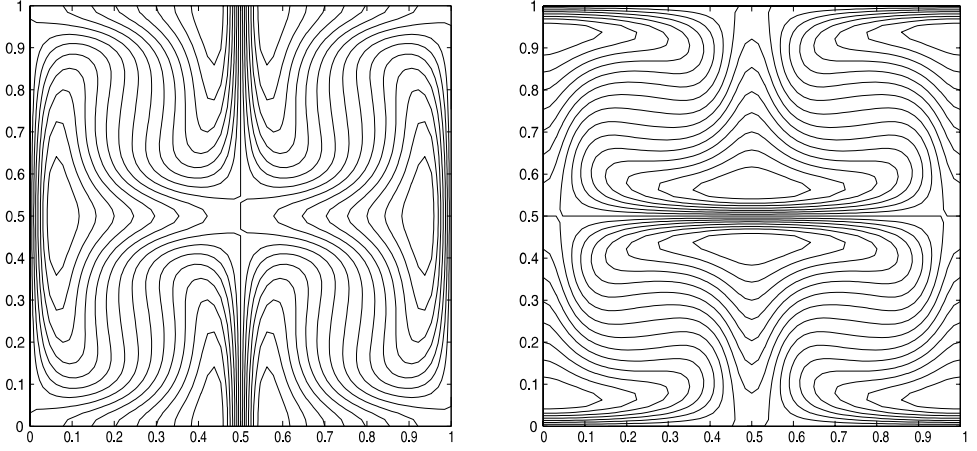


Fig. 3. Cell problem solutions χ_1 (left) and χ_2 (right), rescaled to the unit square.

some arbitrary element domain K of the integral. The key step in obtaining (71) is considering the interaction of χ and ψ . By examining where the layers occur, we obtain the $\mathcal{O}(1)$ estimate and avoid the $1/\sqrt{\epsilon}$ estimate, which is what we would get with a simpler procedure.

In (69) and (70), we have assumed that

$$\|\chi_{i,j}^l\|_{L^\infty(\partial K)} \sim \frac{1}{\sqrt{\epsilon}}. \quad (72)$$

This is what we expect from the argument for the effective diffusivity scaling, as the gradient of χ determines the $\sqrt{\epsilon}$ scaling. In fact, we see in (23) that χ directly determines the σ_ϵ scaling. With the layer of width $\mathcal{O}(\sqrt{\epsilon})$ and gradient $\mathcal{O}(1/\sqrt{\epsilon})$, χ gives the correct scaling, as verified numerically in Park [2000].

Estimate (72) can also be verified directly, as shown in Table 2, derivatives of χ_1 along each direction have the $1/\sqrt{\epsilon}$ scaling; χ_2 gives similar results.

(ii) Estimating Λ_2 :

We again transfer the integral to the boundary first.

$$\delta \int_K a^{ij} \theta_i^k \theta_j^l dx = \delta \int_{\partial K} a^{ij} \theta^k \theta_j^l n_i ds \quad (73)$$

$$\sim \frac{\delta}{h} \int_{\partial K} a^{ij} \chi^k \theta_j^l ds \quad (74)$$

$$\sim \frac{\delta}{h} \int_{\partial K} \epsilon \theta_j^l ds + \frac{\delta}{h} \int_{\partial K} \psi \theta_j^l ds \quad (75)$$

$$\sim \mathcal{O}(\sqrt{\epsilon}) + \mathcal{O}(1) = \mathcal{O}(1). \quad (76)$$

Table 2. $\|\chi_{,y_i}^1\|$ has $1/\sqrt{\epsilon}$ behavior for resolved solutions over one cell.

| ϵ | $\ \chi_{,y_1}^1\ _{L^\infty}$ | Ratio | $\ \chi_{,y_2}^1\ _{L^\infty}$ | Ratio |
|------------|--------------------------------|--------|--------------------------------|--------|
| 0.02 | 0.7903 | — | 0.5650 | — |
| 0.01 | 1.6755 | 2.1200 | 0.8479 | 1.5007 |
| 0.005 | 3.0552 | 1.8234 | 0.9482 | 1.1182 |
| 0.0025 | 4.7457 | 1.5533 | 1.2551 | 1.3236 |

Table 3. $\|\nabla\theta\|$ has $1/\sqrt{\epsilon}$ behavior, with resolved solutions over one cell.

| ϵ | $\ \theta_{,y_1}\ _{1^\infty}$ | Ratio | $\ \theta_{,y_2}\ _{1^\infty}$ | Ratio |
|------------|--------------------------------|--------|--------------------------------|--------|
| 0.02 | 0.7903 | — | 0.6099 | — |
| 0.01 | 1.6755 | 2.1200 | 1.1513 | 1.8877 |
| 0.005 | 3.0552 | 1.8234 | 1.9513 | 1.6949 |
| 0.0025 | 4.7457 | 1.5533 | 2.8846 | 1.4783 |

In order to obtain (76), we use the same argument as for Λ_1 , assuming that

$$\|\theta_{,j}^l\|_{L^\infty(\partial K)} \sim \frac{1}{\delta\sqrt{\epsilon}}. \quad (77)$$

In general, the structure of θ is very complicated, and it is hard to derive such estimates analytically. The effect of the boundary condition may travel past the immediate vicinity of the boundary. That is the reason we move the derivatives of θ from inside the element onto the boundary whenever possible. Since the estimate (77) is only on the boundary, the scaling is easier to understand. From (14), we know that θ behaves as χ on the boundary. So, the θ derivative along the boundary of the element scales as $1/\sqrt{\epsilon}$ with respect to ϵ , following (72). The derivative perpendicular to the boundary could be of order $1/(\delta\sqrt{\epsilon})$ since the $O(\sqrt{\epsilon})$ layer within the δ -size cells gives the absolute thickness of $O(\delta\sqrt{\epsilon})$. The numerical result in Table 3 appears to support this.

(iii) Estimating Λ_3 :

A rough estimate for Λ_3 gives

$$\frac{1}{\delta} \int_K \hat{\sigma}^{ij} \phi_{0,j}^l \phi_{0,i}^k dx = \frac{1}{\delta} \cdot \delta h \cdot \left(\frac{1}{h} \cdot \frac{1}{h} \right) \sim \frac{1}{h}. \quad (78)$$

However, we actually get an additional $\mathcal{O}(h)$ because we can write this term in a difference form $\Lambda_3 = D\lambda^h$ for some λ^h . Because this term does not involve χ or θ , it is exactly the same as in the elliptic case. In Efendiev *et al.* [2000], details on the difference structures of Λ_3 are described. We can also say the same thing for f_1^h , also described in Efendiev *et al.* [2000]. With the additional h , we conclude that $\Lambda_3 \sim \mathcal{O}(1)$.

Table 4. $\|G_0^h\|$ has $1/\sqrt{\epsilon}$ dependence ($n = 8, m = 256$).

| ϵ | $\ G_0^h\ _{l^2}$ | Ratio | $\ G_0^h\ _\infty$ | Ratio |
|------------|-------------------|--------|--------------------|--------|
| 0.04 | 80.6231 | — | 114.2264 | — |
| 0.02 | 144.2394 | 1.7891 | 204.3565 | 1.7890 |
| 0.01 | 218.5669 | 1.5153 | 309.6592 | 1.5153 |
| 0.005 | 292.8571 | 1.3399 | 414.9086 | 1.3399 |
| 0.0025 | 403.3811 | 1.3774 | 571.4950 | 1.3774 |

3.4.4. Error estimate

Putting the Λ_i estimates together, we obtain the following:

$$\|G_0^h A_1^h D U_0^h\| \leq C \left(\frac{1}{\sqrt{\epsilon} h} \right), \quad (79)$$

$$\|U^h - U_0^h\| \leq C_1 \left(\frac{\delta}{\sqrt{\epsilon} h} \right) + C_2 \left(\frac{h^2}{\sqrt{\epsilon}} \right) \|f\|_{0,\Omega}. \quad (80)$$

The $1/\sqrt{\epsilon}$ term comes from the definition $G_0^h \equiv (A_0^h)^{-1}$. This is the scaling we expect for G_0^h because the homogenized coefficient for A_0^h behaves as $\sqrt{\epsilon}$. We verify this numerically in Table 4.

Using (57) and the fact that $\|u\|_{0,\Omega} \sim 1/\sqrt{\epsilon}$, we finally have

$$\frac{\|u - u^h\|_{0,\Omega}}{\|u\|_{0,\Omega}} \leq C_1 \left(\frac{\delta}{h} \right) + C_2 h^2 \|f\|_{0,\Omega} \quad (81)$$

for the relative error. Due to the discrete analysis, we get the (δ/h) term as the leading order error. This is a sharp estimate, as will be verified in the next section. We note that the relative error is independent of ϵ . If we had used the conventional finite element analysis, the first step of using C  as's Lemma (6) by itself would give $1/\epsilon$, from which we cannot recover the current estimate.

3.4.5. Numerical results

We solve the problem numerically to verify the convergence rate given by (81). We use zero Dirichlet boundary condition and the forcing function

$$f = x(1 - y^2) + \sin(10(x - y)) + \cos(5(x + y)).$$

The well-resolved solution was computed on a 4096×4096 grid on the parallel computer (Intel Paragon).

In Table 5, we clearly see that the l^2 norm of the solution itself grows with increasing $1/\sqrt{\epsilon}$: Decreasing ϵ by 2 results in a $\sqrt{2}$ increase in the error. In order to verify the δ/h factor, we reduce δ and h at the same time. In Table 7, we clearly see that as δ and $h \rightarrow 0$ (with the ratio fixed at $\delta/h = 1/4$), the error stays about the same.

From (81), we expect the leading order error to have the $1/h$ term. We see in Table 6 that this is indeed the case, as halving h results in doubling of the error. In

Table 5. $\|u\|$ has $1/\sqrt{\epsilon}$ behavior ($n = 4096$).

| ϵ | δ | l^∞ norm | Ratio | l^2 norm | Ratio |
|------------|----------|-----------------|--------|------------|--------|
| 0.008 | 0.015625 | 3.7034 | — | 1.4633 | — |
| 0.004 | 0.015625 | 4.9398 | 1.3338 | 1.9515 | 1.3337 |
| 0.002 | 0.015625 | 6.9347 | 1.4038 | 2.7391 | 1.4036 |
| 0.001 | 0.015625 | 9.8403 | 1.4190 | 3.8858 | 1.4186 |

Table 6. $\|u - u^h\|$ has the $1/h$ term ($\epsilon = 0.008, \delta = 0.015625$).

| n | m | l^∞ error | Ratio | l^2 error | Ratio |
|-----|-----|------------------|--------|-------------|--------|
| 8 | 256 | 0.0675 | — | 0.0259 | — |
| 16 | 128 | 0.1605 | 2.3794 | 0.0628 | 2.4247 |
| 32 | 64 | 0.3358 | 2.0919 | 0.1297 | 2.0653 |
| 64 | 32 | 0.6134 | 1.8265 | 0.2397 | 1.8481 |

Table 7. $\|u - u^h\|$ has the δ/h term ($m = 256, \epsilon = 0.008$).

| n | δ | l^∞ error | Ratio | l^2 error | Ratio |
|-----|------------|------------------|--------|-------------|--------|
| 8 | 0.03125 | 0.1112 | — | 0.04566 | — |
| 16 | 0.015625 | 0.1517 | 0.7330 | 0.05954 | 0.7668 |
| 32 | 0.0078125 | 0.1661 | 0.9133 | 0.06457 | 0.9221 |
| 64 | 0.00390625 | 0.1625 | 1.0222 | 0.06365 | 1.0144 |

that computation, since ϵ and δ remain the same, we keep the overall resolution the same by keeping $n \times m$ constant. In all the numerical computations of this section, the results are in excellent agreement with the error estimate (81).

3.5. Asymptotic basis

Computing the basis functions numerically can itself be expensive if the fine structures are present as $\epsilon \rightarrow 0$. In some cases, it may be possible to use asymptotic expansion to generate the basis functions. In general, this is a difficult problem. There are many complicated internal layers along the separatrices of the streamfunction. Fortunately, we can simplify this one step by again using the asymptotic expansion:

$$\phi_\delta(\mathbf{x}) = \phi_0(\mathbf{x}) + \delta \chi^k \left(\mathbf{x}, \frac{\mathbf{x}}{\delta} \right) \frac{\partial \phi_0}{\partial x_k} + \delta \theta^\delta. \quad (82)$$

This is a multiple scale expansion for the basis function, just as we had for the global solution in (15). In the case of homogeneous stream function, χ is a function of the fast variable \mathbf{y} only. Thus, we can find the asymptotic expansion for χ instead of ϕ_δ , and then use (82) to get an expression for ϕ_δ . In general, constructing ϕ_δ by finding χ first is not a practical option, since there may not even be a periodic structure. However, we carry out this step here to gain insights into the solution structure.

The problem now is to find the asymptotic solution to the cell problem (25) for χ over the torus. This is still a difficult problem with possibly very complicated layer structures. For the cellular flow, however, we can make some progress. First, we can reduce (25) to a “quarter-cell problem” by using certain symmetry properties. A “cell” was shown in Fig. 1(a); it contains four square regions and so a quarter-cell has boundary layers along the edges and no internal layer. Looking at (25), when \mathbf{e} is unit vector in y , χ is even in the x direction and odd in the y direction. Then by defining $\rho = \chi + x$, the analysis of (25) is reduced to that of Fannjiang and Papanicolaou [1994],

$$-\epsilon \Delta \rho + \mathbf{b} \cdot \nabla \rho = 0, \tag{83}$$

$$\rho(x, 0) = 0, \quad \rho(x, 1) = 1, \quad \frac{\partial \rho}{\partial x}(0, y) = \frac{\partial \rho}{\partial x}(1, y) = 0. \tag{84}$$

To understand (83), we transform the problem to a new coordinate system based on the streamfunction $\psi(x, y)$. Because the boundary layers occur along the streamlines, $\psi(x, y)$ provides a good basis. We define $\psi(x, y)$ as one coordinate and then define a family of lines orthogonal to the streamlines [Childress (1979); Soward (1987)] by $\phi(x, y) = C \int \mathbf{b} \cdot d\mathbf{l}$, where $d\mathbf{l}$ is the tangent to the streamlines. This is an “angle” variable that measures the distance along a given streamline. $C = C(\psi)$ is a constant for the given level set of ψ . It is introduced here to calibrate $\phi(x, y)$, to make sure that it takes the same range of values on each streamline. We can pick $C = 1/C_1$, where C_1 is the circulation along the given streamline. For very small ϵ , we can also choose C_1 to be circulation along the cell boundary since the layers are close to it, and still have small errors [Anufriyev and Fishman (1982)]. This way, C is constant for all $\psi(x, y)$ and we avoid computing it for each streamline.

With these new variables, (83), after some rescaling, becomes [Childress (1979)],

$$\frac{\partial \rho}{\partial \phi} = \epsilon \frac{\partial^2 \rho}{\partial \psi^2}, \tag{85}$$

with appropriate boundary conditions [Childress (1979)]. Physically, this corresponds to the temperature distribution generated by periodic segments of heating, insulation, cooling, and insulation.

The solution to this problem can be obtained by the Wiener-Hopf method [Anufriyev and Fishman (1982)]. Solving the resulting integral equation and writing the solution as an expansion in temperature waves, we can find

$$\rho(\phi, \psi) = \frac{1}{2} + \frac{1}{2} \sum_{n=1}^{\infty} c_{2n-1} \exp\left(-\frac{\xi}{\delta_{2n-1}}\right) \sin\left(\frac{\xi}{\delta_{2n-1}} - (2n-1)\frac{\pi}{2}\phi - \theta_{2n-1}\right) \tag{86}$$

along with the values of all the constants [Anufriyev and Fishman (1982)]. This asymptotic expansion can now be converted numerically to (x, y) coordinates. In Figs. 4(a) and 4(b), we have plotted the two solutions of χ using the asymptotics and the finite element method, respectively. The asymptotic expansion gives a very good approximation in this case.

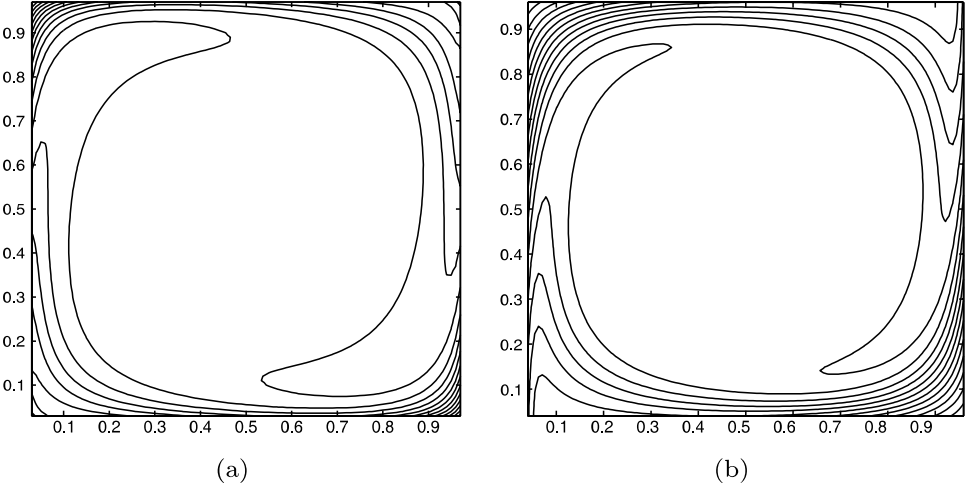


Fig. 4. (a) An asymptotic solution mapped to the (x, y) coordinates; (b) A quarter cell of χ solved numerically ($n = 128$).

In general, it is difficult to construct the multiscale basis function analytically since the multiscale basis functions may contain complicated, genuinely two-dimensional structures. However, when the flow has scale separation and a periodic structure, we can construct a semi-analytic multiscale basis function by using the first two terms in the multiscale expansion of the basis function, i.e.

$$\phi_\delta(\mathbf{x}) = \phi_0(\mathbf{x}) + \delta \chi^k \left(\mathbf{x}, \frac{\mathbf{x}}{\delta} \right) \frac{\partial \phi_0}{\partial x_k}. \quad (87)$$

The corrector χ can be computed numerically over a periodic cell. In particular, when the flow field is homogeneous, the corrector χ is independent of the slow variable, \mathbf{x} . Thus, we need only to compute the corrector numerically for one periodic cell and use it to construct the multiscale basis functions for all coarse grid cells. This will lead to significant computational savings in computing the multiscale basis functions. Moreover, it will reduce the effect of resonance errors introduced by the boundary corrector θ_δ [Efendiev *et al.* (2000)].

3.6. Random coefficients

When the characteristic length scale of the problem is large, the basis functions that sample only a small region cannot capture the correct behavior. An example is the case of random velocity field. The properties of random velocity fields have been studied for a long time, particularly in relation to turbulent flows [Kraichnan (1970); Fannjiang and Papanicolaou (1996); Avellaneda and Majda (1991)]. There are many ways of generating a random flow [Apelian *et al.* (1997)], including specifying the spectrum in the Fourier space, as is usually done in turbulence literature [Kraichnan (1970)], superposing shear flows in random directions [Balk and

McLaughlin (1999)], and randomly placing a large number of vortices [Avellaneda *et al.* (1991)]. We choose the random perturbation to the cellular flow streamfunction [Isichenko (1992)], as it allows for various scaling arguments and a prediction for the effect of randomness.

For a velocity field that is steady, incompressible, has mean zero, and comes from a stationary streamfunction $\psi(x, y)$, it was argued [Isichenko (1992); Isichenko and Kalda (1991)] that as $\epsilon \rightarrow 0$,

$$\sigma_\epsilon \sim \epsilon^{3/13}. \quad (88)$$

This claim (88) is due to a scaling argument based on some results from the percolation theory regarding the geometry of the streamlines of the flow [Fannjiang and Papanicolaou (1997)]. The straight lines in the cellular streamfunction Fig. 1(a) are the separatrices and their intersections are the hyperbolic stagnation points. Due to the nature of hyperbolic points, the cellular structure is unstable when these points are perturbed. This results in the cell boundaries reconnecting in some random manner, forming channels and islands of various sizes. One realization is shown in Fig. 5. Intuitively, the long loops near the critical level $\psi(x, y) = 0$ gives the dominant contribution to the overall diffusivity. This scaling behavior is difficult to verify computationally, as it is an asymptotic result in the limit that δ , the parameter that rescales the domain and time in (19), and viscosity ϵ are small.

We note that in the random flow case, the cell problem (24) is not as useful. In the periodic case, the cell problem allows us to understand the property of the flow by looking at a single period; in the random flow, that period becomes infinitely large. However, we find it convenient to work with the cell problem, even though the cell is the whole domain, since we have the convenient expression (23) for σ_ϵ in

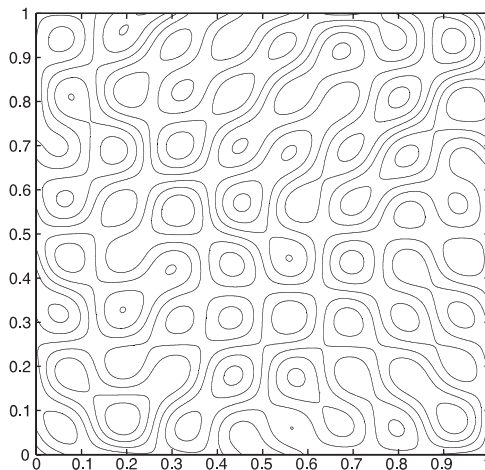


Fig. 5. A random streamfunction with $\delta = 0.25, \gamma = 0.5$.

terms of χ . We think of the random cell problem as the infinite volume limit of the periodic cell problem.

We could also solve the unsteady problem with a regularized delta function as an initial condition and compute σ_ϵ as we did for the cellular flow. However, this presents some complications in the numerical problem. We find that in order for any structures related to the topography of the streamlines to appear, we need to run the equation for a long period of time. However, because of the increase in the effective diffusivity, we have the fluid reaching the boundary more quickly. Instead, we solve for the steady-state solution $\rho(x, y)$ given by

$$-\epsilon\Delta\rho + \mathbf{b} \cdot \nabla\rho = 0, \quad (89)$$

$$\rho(x, 0) = 0, \quad \rho(x, 1) = 1, \quad \frac{\partial\rho}{\partial x}(0, y) = \frac{\partial\rho}{\partial x}(1, y) = 0, \quad (90)$$

just like (83) and then use

$$\sigma_\epsilon \approx \epsilon \iint |\nabla\rho|^2 dx dy. \quad (91)$$

This is equivalent to the cell problem (23) for $\chi = \rho - x$. (91) is an approximation because the problem is no longer periodic. Our computations using a multigrid solver [De Zeeuw (1990)] on a 4096×4096 grid gives an estimate very similar to the theoretical scaling, both as $\delta \rightarrow 0$ for small ϵ and $\epsilon \rightarrow 0$ for small δ [Park (2000)].

3.6.1. A variational principle

To overcome the problem caused by the nonlocality, we need to have some knowledge regarding the structure of the layers before solving the problem. This information can be derived from a variational principle in this problem [Avellaneda and Majda (1991); Fannjiang and Papanicolaou (1994)]. The main difficulty in obtaining a variational principle for the convection-diffusion equation comes from the convection term. Because of the skew-symmetric part \mathbf{H} in the coefficient ($\epsilon\mathbf{I} + \mathbf{H}$), we cannot easily formulate a functional that has σ_ϵ as the minimum. After a symmetrization procedure, Fannjiang and Papanicolaou [1994] found

$$\sigma_\epsilon(\mathbf{e}) = \inf_{\langle \nabla f \rangle = \mathbf{e}} \left(\epsilon \langle \nabla f \cdot \nabla f \rangle + \frac{1}{\epsilon} \langle \Gamma_g \mathbf{H} \nabla f \cdot \Gamma_g \mathbf{H} \nabla f \rangle \right), \quad (92)$$

at the cost of introducing the nonlocal operator Γ_g . $\Gamma_g = \nabla\Delta^{-1}\nabla\mathbf{f}$ which is the orthogonal projection operator onto the subspace of gradient fields. A major contribution to the integral comes from $|\nabla\chi|^2$ in the boundary layer. For computational purposes, (92) can be written in a different form involving a Poisson problem, which

can be solved fast [Fannjiang and Papanicolaou (1994)], in place of the nonlocal Γ_g operator.

Given this information, it is reasonable to try test functions of the form

$$f(x, y) = \exp\left(-\frac{|\psi(x, y) - c|}{\epsilon^\alpha}\right) \quad (93)$$

for approximating the behavior of χ , where α is the unknown exponent we wish to find and c is the unknown level set at which the longest streamline occurs. We expect this c to be near 0 if the other parameters are such that the percolation results hold. Since $f(x, y)$ has a layer of thickness proportional to ϵ^α and $\nabla f \sim 1/\epsilon^\alpha$, we expect the α of (93) that achieves the minimum of the variational integral to be close to the true value of the exponent. Variational principles for the convection-diffusion problem has been used mostly as a theoretical tool for deriving estimates of the effective diffusivity for many flows, but we use it as a numerical tool here. Numerically, the minimization of the integral using the test function (93) gives estimates close to the theoretical ones for small δ and ϵ , both for cellular and random flows. In Figs. 6(a) and 6(b), we plot the values of the functional for the exponent α and the level set value c .

With the scaling information provided by the variational principle, we attempt to design a more efficient numerical method. The idea is similar to that of the previous section. We would like to have the trial or test functions that contain the same behavior as the solution. The idea presented in this section can be applied more generally. Given a problem and its variational principle, if the dominant contribution to the integral comes from a layer structure of the solution, we can use this idea to find out the thickness and the location of the layer. Then we can use this information to design a more efficient method to solve the original problem.

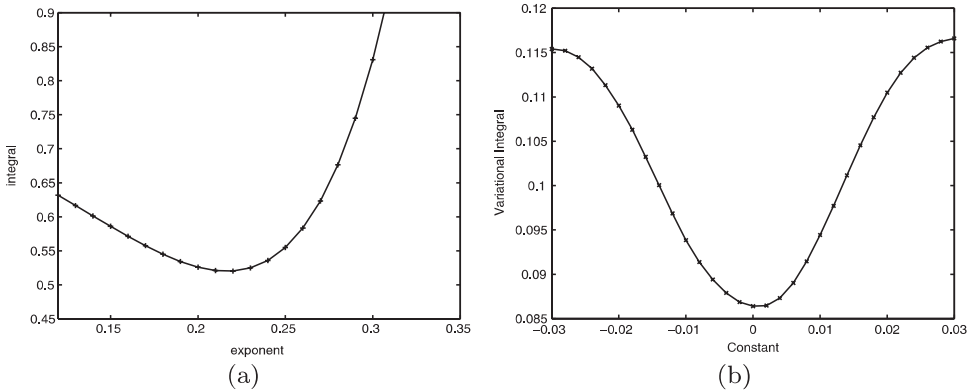


Fig. 6. (a) Variational integral for the random velocity field as a function of the exponent α ($\epsilon = 0.0001$, $\delta = 0.125$); (b) Variational integral for the random velocity field as a function of the constant c ($\alpha = 0.23$, $\epsilon = 0.0001$, $\delta = 0.25$).

3.6.2. A numerical method

One-dimensional problem

We noted previously that the test function space should be composed of exponentials that solve the local adjoint problem in order to minimize the error for a given mesh. This idea of the exponential basis has been studied extensively mostly for the one-dimensional case, assuming a positive coefficient, $a(x) \geq a_0 > 0$ [Roos *et al.* (1991); O’Riordan and Stynes (1991); Miller *et al.* (1996); Stynes and O’Riordan (1991)]. In this section, we study how basis functions should be constructed for a more general case when the coefficient has a turning point. Understanding the turning point case in one dimension is essential for the two-dimensional problems studied in the next section, since a similar phenomenon occurs all along the layers.

We consider the simple case

$$-\epsilon u'' + a(x)u' = 0, \quad a(x) = \begin{cases} 1, & 0 < x < 1/2 \\ 0, & x = 1/2 \\ -1, & 1/2 < x < 1 \end{cases} \quad (94)$$

with boundary conditions $u(0) = 0$ and $u(1) = 1$. The exact solution can be found easily [Park (2000)]. The solution has a sharp internal layer at $x = 1/2$. Intuitively, the “fluid” on the left of $x = 1/2$ is being pushed to the right while the opposite is true on the right of $x = 1/2$.

For this problem, simply replacing the linear basis function at the layer with the local solutions of the homogeneous equation does not work. We need to solve the equation locally, but we must separate the exponential behavior on each side of the turning point into two different bases as well as including the linear basis. The shape and the direction of the layer cannot be determined locally, so three basis functions should be present over the same support. The problem is well-posed as long as the three bases remain linearly independent.

With the two types of bases, ϕ_i and ψ_i , that are linear and exponential respectively, we can write

$$u = \sum_i a_i \phi_i + \sum_i b_i \psi_i, \quad (95)$$

including ψ_i only at the turn points where the layer will occur.

One way to solve this is as a minimization for the variational principle, using the conjugate gradient method for instance [Park (2000)]. A better way is through the finite element formulation, with modifications for the extra unknowns b_i . With ψ_i , the size of the local stiffness matrix is larger. Because we include the extra bases only near the possible layer locations, the matrix problem increases only by a small number of rows and columns. While the new matrix does not have a banded structure anymore, it can be solved efficiently, as will be described later in this section. With the extra bases, oscillations do not occur even with a coarse grid and the performance is similar as $\epsilon \rightarrow 0$.

Two-dimensional problem

In two dimensions, we would also like to design basis functions that capture the exponential behavior. The difficulty, as mentioned previously, is that solving the equation locally does not give the correct behavior for the velocity fields we consider in this section. This is the reason we introduced the variational principle in Sec. 3.6.1. In this section, we examine how to use our knowledge of the layer structure, based on the percolation theory and the variational principle, to design basis functions that will capture a global layer accurately. We find that the framework developed in this section can be used more broadly as we will see in the next section.

Ideally, we would like to construct special basis functions $\Psi_i, i = 1, \dots, s$ such that $u - \sum_1^s b_i \Psi_i$ is smooth when b_i are chosen optimally. If we can do that, the smooth part can be approximated well by the bilinear functions Φ_i . Then, for $u^h = \sum_i^n a_i \Phi_i + \sum_i^s b_i \Psi_i$, we would have [Strang and Fix (1973)]

$$\|u - u^h\|_l \leq C h^{k-l} \left\| u - \sum_1^s c_i \Psi_i \right\|_k. \quad (96)$$

In general, it is too hard to find all these Ψ_i . It is equivalent to finding the location and behavior of all the layers in the solution exactly. For the random flow problem, we are trying to capture just the main feature of the solution which we think has the dominant contribution. We see in the numerical examples that the large errors in fact come from few sharp layers and that the use of a limited number of Ψ_i improves the result significantly.

Numerical results

The extra basis functions are constructed based on an insight gained from the one-dimensional example. We define the support of the new basis functions along the layer and then embed the correct exponential behavior. At each location, there are two bases, each with the predicted layer shape on one side, and linear function on the other. The underlying grid is uniform and the bilinear bases on this grid take care of the smooth part of the solution. This method applied to the cellular flow appears to work well, with exponential bases picking up most of the weight near the layers in the solution [Park (2000)]. In fact, the minimum l^2 error for the solution is achieved when the basis functions are constructed with $\alpha \approx 1/2$ and it is robust with respect to a slight error in the estimate of α . This confirms that the value of α obtained through the variational principle can be used to approximate the exponential behavior.

The streamlines along which the layers occur are generally not aligned with the underlying uniform grid. The computer implementation in this case becomes more involved. For example, we need to segment evenly the separatrices on which layers occur, define the element support appropriately, and use a mapping from a quadrilateral support to a square to construct modified linear bases. While the variational principle of the last section gives information about a dominant layer,

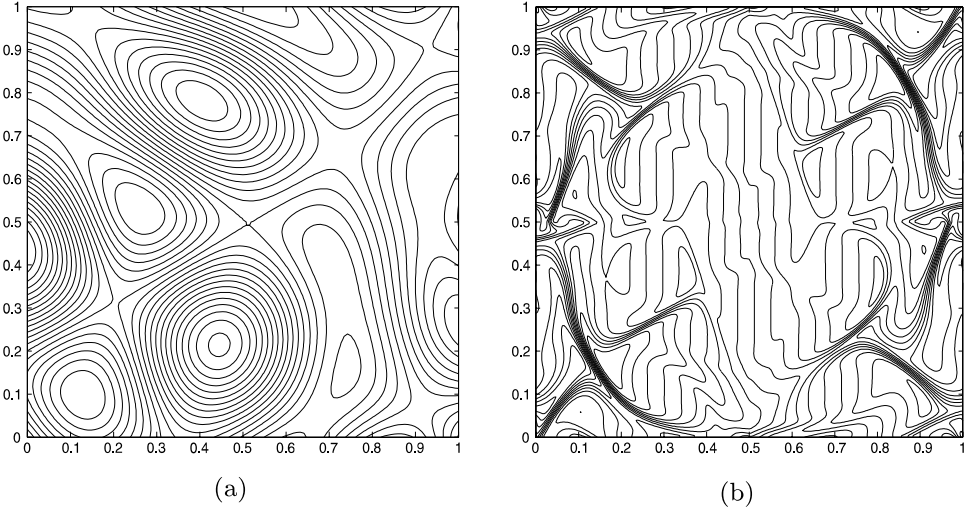


Fig. 7. (a) A random streamfunction for the use of exponential basis in a quarter of the domain; (b) solution to the cell problem for a random streamfunction, generated by extending the random streamfunction (a) to all four quadrants and then rescaling it to the unit square, with periodic boundary conditions.

this may not be sufficient. If δ is not small, for example, there may not be one such layer; there may be other layers we need to capture. For instance, the internal layers will also occur along the separatrices that contain the hyperbolic points, because they are where the flows in opposite directions meet. Therefore, we search for these hyperbolic points near the $\psi = 0$ and place additional extra bases functions nearby.

We first generate some random streamfunction, such as the one shown in Fig. 7(a). Since the formulation of the cell problem requires periodic boundary conditions, we extend the original streamfunction in the first quadrant to the other quadrants by reflections $\psi(x, y) = \psi(x, -y)$ and $\psi(x, y) = \psi(-x, y)$ [Fannjiang and Papanicolaou (1997)]. We then rescale the domain to $[0, 1]^2$ before solving the problem. A resolved solution for the velocity field given by the streamfunction of Fig. 7(a) extended in this way is shown in Fig. 7(b). To use the exponential basis functions, we lay down a mesh along the streamlines. The boxes along the curved streamlines indicated the support of the special basis functions. Along this curved mesh, we place a set of exponential basis functions, in a manner analogous to the one-dimensional example of Sec. 3.6.2.

The results for the $\epsilon = 0.001$ case are shown in Table 8. We see that the error is decreased by 60% for both in the l^2 and l^∞ norm when exponential functions are used. Since changing the resolution in the interior of the basis function results in only small changes, most of the error is due to other weak layers not approximated by the exponential form or due to the differences in the exponential behavior on

Table 8. Improvement due to the use of exponential functions ($\epsilon = 0.001$).

| | n | m | l^2 error | l^∞ error |
|-------------|-----|-----|-------------|------------------|
| Linear | 16 | 16 | 0.089987 | 0.28294 |
| Exponential | 16 | 8 | 0.038651 | 0.11858 |
| | 16 | 16 | 0.037504 | 0.11340 |
| | 16 | 32 | 0.037553 | 0.11117 |

different layers. For smaller ϵ , the difference is larger. For $\epsilon = 0.0005$, we see that the multigrid method that we found to be very robust does not converge in the bilinear case unless n is at least 64. For $\epsilon = 0.00025$, we need at least $n = 256$ for its convergence. With the exponential basis, smaller ϵ cases are handled more effectively. Although the error does increase each time ϵ is halved (roughly 40%), we obtained stable solutions with relatively small error. For example, for $\epsilon = 0.00025$, the l^2 error is 0.080120 (the l^∞ error is 0.21643).

It is not always possible to determine the shape of the layers exactly. The widths also vary a little on different layers. Therefore, it is desirable to have low sensitivity to small inaccuracies in the estimates for α , the exponent that determines the layer thickness given some ϵ . Our simulations show that the method is tolerant to small errors in the estimate of α [Park (2000)].

Ideally, we would like to use an efficient method such as the one developed in this section to compute the diffusivity scaling for the random flow accurately. However, in the steady-state problem, we find that the layer does not form a long smooth curve. That may be true in the time-dependent problem but still only in the long-time limit. Instead, due to the nonhomogeneous boundary condition, the solution develops more small scale features. We see that the layer locations are very complicated. We find that the resulting complexity of these layers is too burdensome for the method of this section. The fact that the method seems to be insensitive to small errors in the layer thickness estimate works to our advantage, but the location of the layers are not clearly known. As δ gets small, there are many potential sites for these layers and it becomes difficult to account for all of them. For less complex problems, the exponential basis approach works well, as we demonstrated above with numerical examples.

There is one important difference between conventional methods and the exponential approach we consider here: because the latter is based on asymptotic theory, its performance should improve as the parameters get small. For example, if we can manage to evolve the time-dependent problem for a long time, there should be more coherent structure in the solution, as predicted by the percolation results. Then the present method should capture more of the important features in the solution. In contrast, other methods do not take advantage of any solution structure and would deteriorate quickly for smaller parameters.

Solution of the augmented linear system

The global stiffness matrix of the conventional finite element method in two dimensions is a banded matrix with nine diagonals. With the extra basis functions, however, the stiffness matrix contains added rows and columns, as shown in (97).

$$\left(\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}. \quad (97)$$

The matrix A represents the interaction of the bilinear elements, B and C the interactions between the bilinear and the exponential, and D the interaction between the exponentials; x_1 is the weight of the bilinear elements and x_2 that of the special elements. The support of the exponential basis function does not have to be the same size as that of a bilinear element; thus one exponential basis function can interact with more than eight bilinear elements. As a result, B , C , and D in general do not have a regular pattern and must be treated as full matrices.

This system can be solve efficiently using a generalization of the Bordering algorithm [Keller (1987)] to take advantage of the banded structure of A . Motivated by the Schur complement form, we have

$$(D - CA^{-1}B)x_2 = g - CA^{-1}f.$$

Thus we can solve the system by (i) solve for η in $A\eta = B$; (ii) let $L = D - C\eta$; (iii) solve for ξ in $A\xi = f$; (iv) let $r = g - C\xi$; (v) solve $Lx_2 = r$; and (vi) solve $\xi - \eta x_2$. Therefore, we can still use the fast multigrid solver [De Zeeuw (1990)] for the banded matrix A . Step (i) is the most expensive, since we need to solve the matrix problem for each column of B . However, the multigrid iteration can be supplied with a good initial guess and hence is more efficient after the first solve. Many columns of B are also zeros and do not require the iteration. One difficulty still arises for the cell problem (25) if the periodic boundary condition is required for the cell problem. We then store all the matrices in a sparse form and supply the matrix-vector multiplication routine to the GMRES algorithm for a reasonable good performance [Barrett *et al.* (1994)].

4. Elliptic Equation with Discontinuous Coefficients

In the previous section, we considered the singularly perturbed elliptic equation that has the characteristics of a hyperbolic equation. Now we consider the genuinely elliptic case but with degeneracies coming from highly variable, possibly even discontinuous, coefficients or non-smooth boundaries. In general, the coefficient may be a matrix (a_{ij}) , but here we will deal with the isotropic case, $a_{ij} = a \delta_{ij}$. The coefficient $a(x, y)$, for instance, may represent permeability or electrical conductivity in a medium and its contrast, $\max a(x, y) / \min a(x, y)$, may be high.

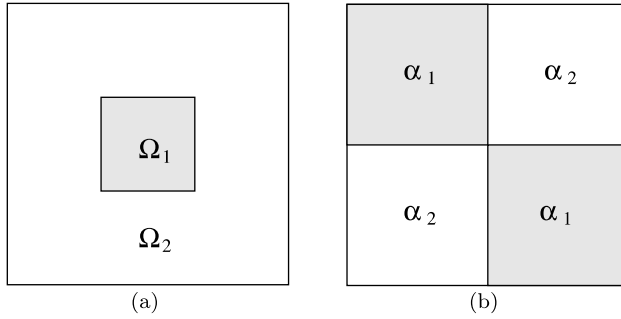


Fig. 8. An interface with sharp corners: (a) different coefficients in regions Ω_1 and Ω_2 ; (b) the checkerboard singularity at the center.

When the coefficient is discontinuous, we require the governing equation (1) to hold in each domain and also that the solution and the normal flux be continuous across the interface. Suppose we have

$$a(x, y) = \begin{cases} \alpha_1 & \text{in } \Omega_1, \\ \alpha_2 & \text{in } \Omega_2, \end{cases}$$

where Ω is a convex polygon in \mathcal{R}^2 , $\Omega_1 \subset \Omega$ is an open domain with boundary $\Gamma = \partial\Omega_1 \subset \Omega$, and $\Omega_2 = \Omega \setminus \Omega_1$. An example with a non-smooth interface is shown in Fig. 8(a). The continuity and jump conditions are

$$u_1|_{\partial\Omega_1} = u_2|_{\partial\Omega_1}, \quad \alpha_1 \frac{\partial u_1}{\partial n} \Big|_{\partial\Omega_1} = \alpha_2 \frac{\partial u_2}{\partial n} \Big|_{\partial\Omega_1}, \quad (98)$$

where n is the outward normal to the interface $\partial\Omega_1$.

When the interface is smooth, the singularity is not severe, with smooth solutions inside each region. Many methods have been developed for this case and they work well, at least for moderately large contrast. A method that is the simplest conceptually but nontrivial in implementation is aligning the grid with the discontinuity. In Chen and Zou [1998], it is proved that if the boundary is at least C^2 , this finite element method converges nearly in the same optimal way, in both the L^2 and energy norms, as in the problems without interfaces. There are also many methods that use regular grids, such as the Immersed Interface Method [Leveque and Li (1994)], but they require the interface to be smooth.

The more difficult problem is with the singularities that arise due to a non-smooth interface. Here, we would like to carry out a similar strategy as in the convection-diffusion case, attempting to capture the effect of the singularity in the basis function and thus obtain a correct coarse grid solution. However, the singularities due to non-smooth interfaces can be very severe. Therefore, unlike in the previous section, it may not be possible to resolve singularity numerically even in the basis function if one uses a conventional finite element method. We overcome

this problem by constructing the basis using the infinite element method, which, as described below, can be very effective in capturing the most severe singularities.

4.1. *Corner and checkerboard singularities*

In general, an effective numerical method requires an analytic form of the singularity [Strang and Fix (1973)]. The degree of singularity caused by a corner in the interface depends on the shape of the corner and the contrast in the coefficient. When f is sufficiently smooth on the boundary, the solution in the neighborhood of the singularity can be expanded in polar coordinates [Han (1982)] as

$$u(r, \theta) = \sum_{i=0}^{\infty} \gamma_i r^{\nu_i} \Theta_i(\theta),$$

where a steep gradient due to small ν_i requires many grid points. The correct form for the singularity can be obtained by solving the Sturm-Liouville problem for the exponents ν_i : $(a\Theta')' + \nu^2\Theta = 0$, with the associated eigenfunctions Θ_i and appropriate jump conditions at the interface, $\Theta(\theta_i^+) = \Theta(\theta_i^-)$, $a(\theta_i^+)\Theta'(\theta_i^+) = a(\theta_i^-)\Theta'(\theta_i^-)$. Knowing ν_i , it is possible, for example, to use a conformal mapping of type $z = \xi^\alpha$ to map a region around the singularity to a new domain.

For simple problems, this process can be carried out and we can obtain exact formulas for eigenvalues and eigenfunctions. But in general, this must be solved numerically. In the new method we introduce, there is no need to solve this eigenvalue problem. The local grid will have sufficient resolution, even if ν_i are small.

The more difficult case is that of the “checkerboard” configuration, shown in Fig. 8(b) with the singularity at the center. Sometimes it is referred to as the “four-corner juncture” problem. In general, the four regions may have four different conductivities, $\alpha_i, i = 1, \dots, 4$. In the neighborhood of the singularity, the analytic solution behaves as [Rüde and Zenger (1986)]

$$u(r, \theta) = w_1(\theta)r^\gamma + w_2(\theta)r^{2-\gamma} + \mathcal{O}(r^2), \quad (99)$$

where $\gamma = \gamma(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ and $0 \leq \gamma \leq 1$. When three of the four regions have the same α , as in a corner of a rectangle, we get $\gamma = 2/3$. For our numerical example, we consider the case when $\alpha_1 = \alpha_3 = 1$ and $\alpha_2 = \alpha_4 = \alpha$. In that case, $\gamma = \sqrt{\alpha}$ [Rüde and Zenger (1986)]; when the contrast is 10 000, the leading order behavior at the singularity is $r^{0.01}$. We note that as $\alpha \rightarrow 0$, we also have $\gamma \rightarrow 0$. In the limit of $\alpha \rightarrow 0$, the solution becomes discontinuous at the singularity.

Since the singularity is local, we can use regular finite elements in the smooth region away from the singularities. Near the singularity, we employ the “infinite elements.” The infinite element method is an elegant technique that has the advantages of having essentially infinite mesh refinement at the singularity.

4.2. The infinite element method

First appeared in early 1970s, the infinite element method has been most useful in infinite domain problems and singularity problems. Its many properties and some applications are described in Ying [1995] and the references therein. The main idea is to lay down a grid with certain similarity structure and then utilize the self-similarity in reducing the size of the problem. For example, consider a square domain with a singularity at the center. We lay down a grid that looks like the one in Fig. 9(a). The dimensions of the elements decrease geometrically by a factor of $0 < \xi < 1$ as they get closer to the center. Close to the center, the elements get arbitrarily small and there is essentially an infinite number of refinements. We call the region between two successive nested squares a “layer” (see Fig. 9(b)). In general, the element can be any polygon and the singularity does not have to be located at the center. The only requirement is that the similarity is preserved between the layers and in the triangulation of each layer. In partitioning inside a layer, we do it in such a way as to avoid small angles at the element corners.

For the following discussion, we label the nodes by letting $y_k^{(i)}$, ($k = 0, \dots, \infty$; $i = 1, \dots, n$) be the i th point around the center on the k th layer. This is illustrated in Fig. 9(b). Now, to create the global stiffness matrix, we add up the local stiffness matrix for each element to get a “layer stiffness matrix.” It can be written as

$$\begin{pmatrix} K & -A^T \\ -A & K' \end{pmatrix}, \tag{100}$$

with K and K' being positive definite. This is a $2n \times 2n$ matrix that relates $y_k^{(i)}$, ($i = 1, \dots, n$) and $y_{k+1}^{(i)}$, ($i = 1, \dots, n$) for any k . The key observation is

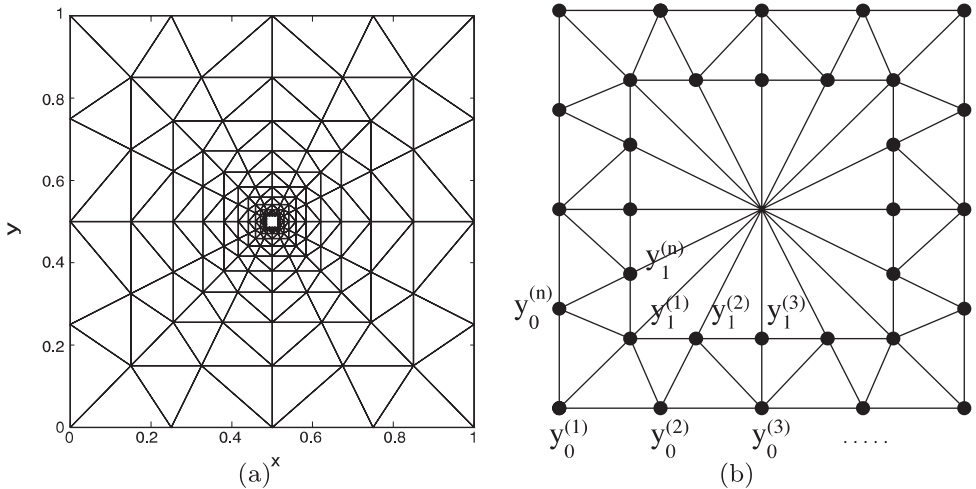


Fig. 9. (a) An example of the infinite element grid (further refinements at the center are not shown in the picture); (b) The node numbering scheme for the first layer.

that this matrix is the same for all layers due to the self-similarity of the grid. For the global stiffness matrix, we put the layer matrices together and obtain

$$\begin{pmatrix} K & -A^T & & \\ -A & K' + K & -A^T & \\ & -A & K' + K & \ddots \\ & & \ddots & \ddots \end{pmatrix} \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} b_0 \\ 0 \\ 0 \\ \vdots \end{pmatrix}. \quad (101)$$

This is an infinite set of equations, with b_0 from the boundary condition. In order to solve this system, we first assume that there is a matrix X , called the *transfer matrix*, such that

$$y_{k+1} = X y_k. \quad (102)$$

The existence of this matrix is proved in Ying [1995]. This transfer matrix plays a crucial role: Given the boundary condition y_0 , we can compute the values at all the interior grid points y_k , ($k = 1, \dots, \infty$) from X through (102). This means all the information about the singularity is contained in X . With $y_{k+1} = X y_k$ and $y_{k+2} = X^2 y_k$, each block line gives $-A + (K' + K)X - A^T X^2 = 0$. While solving for X can be formulated as an eigenvalue problem, the following simpler iterative method can be used instead [Han (1982)]. We generate a sequence of matrices K_i, K'_i , and A_i by

$$K_i = K_{i-1} - A_{i-1}^T (K_{i-1} + K'_{i-1})^{-1} A_{i-1}, \quad (103)$$

$$K'_i = K'_{i-1} - A_{i-1} (K_{i-1} + K'_{i-1})^{-1} A_{i-1}, \quad (104)$$

$$A_i = A_{i-1} (K_{i-1} + K'_{i-1})^{-1} A_{i-1}, \quad (105)$$

$$K_z^i = K_i - \frac{A_i^T g g^T A_i}{g^T K'_i g}, \quad (106)$$

with $K_0 = K, K'_0 = K'$ and $A_0 = A$ from (100). In Ying [1995], it is proved that the sequence of K_z^i converges:

$$\lim_{i \rightarrow \infty} K_z^i = K_z.$$

With K_z , we can then compute X using [Ying (1995)]

$$X = (K_z + K')^{-1} A. \quad (107)$$

The motivation for this procedure is the minimization of energy [Ying (1995)]. The most expensive part of the algorithm is inverting the $2n \times 2n$ matrix $(K_{i-1} + K'_{i-1})$ at each iteration; but for relatively small n , e.g. $n = 16, 32, 64$, the algorithm is still fast. Using this algorithm, we define $X^{(k)} \equiv (K_z^k + K'_k)^{-1} A_k$ and iterate until $X^{(k)}$ converges to some specified tolerance. For the problems we discuss in the next section, the number of iterations till $\|X^{(k)} - X^{(k+1)}\| < 10^{-10}$ is fewer than 12 and is not sensitive to the severity of the singularity.

4.3. Infinite element as a basis function

We utilize the resolution of the infinite elements by employing them at the singularities; away from the singularities, we can use the standard bilinear elements. In an uniform grid, the only requirement is that the singular part of the solution is contained in one element, which is not difficult to do since the worst singular behavior is confined to a local neighborhood.

Once we compute all the basis functions, we need to put them together into the global stiffness matrix. The integration over the linear elements can be done with the two-dimensional trapezoidal rule as usual. But over an infinite element, it is not immediately clear how to sum up the contribution of all the layers. The shapes of the local elements are different and there are infinitely many of them. Fortunately, we find that a simple expression can be obtained just in terms of the boundary condition and the transfer matrix. Integrating over an element K ,

$$\int_K \alpha \nabla \phi_j \cdot \nabla \phi_i \, d\mathbf{x} = \sum_{\substack{\text{layers} \\ k=1, \dots, \infty}} \int_{k\text{th layer}} a \nabla \phi_i^{(k)} \cdot \nabla \phi_j^{(k)} \, d\mathbf{x} \quad (108)$$

where, for each layer k , the infinite element solutions $\phi_i^{(k)}, \phi_j^{(k)}$ are made up of linear elements. After some algebra in carrying out the summation, there is cancellation of terms between succeeding layers and it can be shown [Park (2000)] that

$$\int_K a \nabla \phi_i^{(k)} \cdot \nabla \phi_j^{(k)} \, d\mathbf{x} = \tilde{\alpha}^{(1)T} (K - A^T X) \tilde{\beta}^{(1)}, \quad (109)$$

where $\tilde{\alpha}^{(1)}$ is the boundary condition for ϕ_i and $\tilde{\beta}^{(1)}$ for ϕ_j . Therefore, the sum over an element has been reduced to multiplying the boundary condition by some matrix. This is not unexpected since the transfer matrix X should contain all the interior information.

As in the multiscale FEM case, the correct boundary condition for the infinite element basis function is not known. However, good results are obtained when the reduced one-dimensional problem is used on the boundary. This is especially simple for a piecewise constant coefficient because an exact solution can be obtained easily that satisfies the continuity and the jump conditions. Details can be found in Park [2000].

4.4. Numerical results

When the singularity is severe, the standard FEM converges very slowly as described later in the section. This makes it hard to obtain the highly-resolved “exact” solution to which we can compare other computational results. However, as long as there are no pointwise oscillations, we are able to estimate the convergence rate β using a sequence of grids by $\beta = (\log 2)^{-1} \log (\|u_n - u_{2n}\| / \|u_{2n} - u_{4n}\|)$, where u_n, u_{2n} , and u_{4n} denote the solutions at the grids of dimensions $n, 2n$, and $4n$, respectively.

Corner problem

We first consider singularities that arise from sharp corners in the interface, such as the one shown in Fig. 8(a). Suppose we have the coefficient

$$a(x, y) = \begin{cases} 1 & \text{if } x \in \Omega_1, \\ \alpha & \text{if } x \in \Omega_2. \end{cases}$$

For this problem, the standard FEM converges with $\mathcal{O}(h)$, when the grids are lined up with the interface [Han (1982)]. This is easy to do if the interface is simple; if the interface is complicated, fitting the mesh around the interface is not a trivial process. If the singularity happens to be inside an element, the standard FEM results in a large error.

With the infinite elements at the four corners of the square, the convergence rate is still first order in term of h . However, the error is much smaller because the singularities are now well-resolved. We see in Table 9 that the l^2 error from the infinite element method is about 10 times smaller than that produced by the standard FEM, even with moderate $m = 16$ for the infinite element. A major advantage of our new method is also the following: Since the main source of the error is the singularity at the corners, we can keep the uniform coarse grid in the whole domain and simply refine the few infinite element bases locally. Unlike the standard adaptive mesh methods in which new refinement points are coupled globally to all the others, the refinement inside an infinite element is decoupled from the rest of the domain. The size of the global stiffness matrix does not change. We can see in Table 9 that the reduction in error with this refinement is nearly first order in h_m in the l^∞ norm, where $h_m = h/m$ is the mesh size along the boundary of the element. With this strategy, we can obtain very small errors.

Table 9. Improvements in error with the infinite element bases at the corners.

| n | m | l^2 error | Rate | l^∞ error | Rate |
|-------|-----|-------------|--------|------------------|--------|
| FEM 8 | 32 | 0.07619 | — | 0.17850 | — |
| IEM 8 | 16 | 0.00733 | — | 0.02946 | — |
| IEM 8 | 32 | 0.00430 | 0.7695 | 0.01515 | 0.9594 |
| IEM 8 | 64 | 0.00299 | 0.5242 | 0.00768 | 0.9801 |

Table 10. Sensitivity of the infinite element basis method for the corner problem ($n = 8, m = 32$).

| α_1 | α_2 | l^2 error | Ratio | l^∞ error | Ratio |
|------------|------------|-------------|--------|------------------|--------|
| 1 | 0.001 | 0.0042230 | — | 0.015367 | — |
| 1 | 0.0001 | 0.0041871 | 0.9915 | 0.015252 | 0.9925 |
| 1 | 0.00001 | 0.0039059 | 0.9328 | 0.013827 | 0.9066 |
| 1 | 0.000001 | 0.0038436 | 0.9840 | 0.012988 | 0.9393 |

Table 11. Sensitivity to the contrast with the infinite element basis ($n = 16, m = 16$).

| α | l^2 error | Ratio | l^∞ error | Ratio |
|----------|-------------|-------|------------------|-------|
| E-1 | 0.0006797 | — | 0.0010896 | — |
| E-2 | 0.0048499 | 7.14 | 0.0078976 | 7.24 |
| E-3 | 0.0144204 | 2.97 | 0.0239461 | 3.03 |
| E-4 | 0.0172374 | 1.20 | 0.0300191 | 1.25 |
| E-5 | 0.0166001 | 0.96 | 0.0309556 | 1.03 |
| E-6 | 0.0166416 | 1.00 | 0.0311084 | 1.00 |
| E-7 | 0.0177008 | 1.06 | 0.0311447 | 1.00 |

We note that increasing m (an element has m points along each edge) here is different from the refinement of multiscale basis of the previous sections. In the infinite element case, the number of unknowns is only doubled when m is doubled, since only the boundary points are involved in the computation. In the multiscale and most other methods, the number of unknowns is squared when m is doubled. Because the infinite element grid can resolve severe singularities, the error should also be insensitive to high contrast. Indeed, as we vary the contrast of the coefficients, the error stays basically the same up to $\alpha = 10^{-6}$ or so, as shown in Table 10.

Checkerboard problem

For the more difficult problem of the checkerboard type shown in Fig. 8(b), the leading order behavior at the singularity is r^γ with $\gamma = \sqrt{\alpha}$, as described earlier. For a fixed point in the domain, the convergence of the standard finite element or finite difference method is predicted to be $\mathcal{O}(h^{2\gamma})$ [Rüde and Zenger (1986)]. This means that, even with the assumption that the gridlines are lined up with the discontinuities and the four-corner is on a grid point, as the contrast in the conductivities increases, the method is arbitrarily slow. We have verified this numerically [Park (2000)]. In Fig. 10, we plot an example of the infinite element basis for the checkerboard case, with the contrast of 10^5 . The values on the boundary are obtained by solving the one-dimensional equation. The severe singularity at the center makes it obvious why an infinite element is needed.

We verify the performance of the infinite element bases by considering the problem with multiple junctures, with the coefficient alternating between α_1 and α_2 many times at regular intervals in both directions. For convenience, we again fix $\alpha_1 = 1$ and let α_2 be the parameter α . We lay down a grid in such a way that each juncture is placed in the middle of an element. The singular behavior is usually confined to the immediate vicinity of the singular point; as long as the element covers this area, the location of the singularity within the element is not important.

The convergence for the standard FEM is very poor with respect to the increase in contrast. In Table 11, we show the behavior of the error for the infinite element basis method as the contrast is increased, for the case α alternating eight times in each direction. We see that the error increases initially, but grows less sensitive

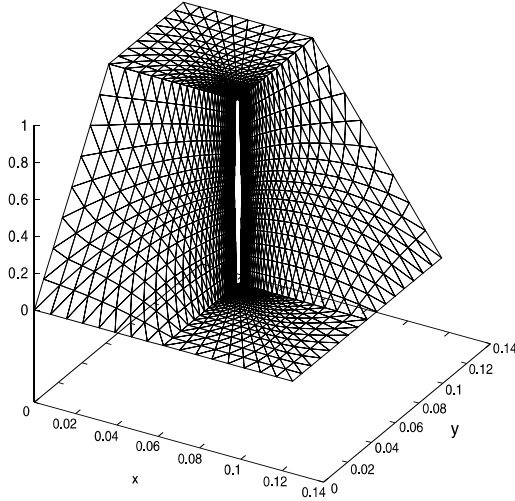


Fig. 10. An infinite element for the checkerboard pattern, with conductivities $\alpha_1 = 1$ and $\alpha_2 = 0.00001$ ($m = 16$); first 30 layers are plotted.

Table 12. Error for multiple juncture problem as the infinite element basis is refined ($\alpha = 1/256$).

| n | m | l^2 error | Rate | l^∞ error | Rate |
|-----|-----|-------------|--------|------------------|--------|
| 16 | 8 | 0.01290978 | — | 0.02175406 | — |
| | 16 | 0.00888366 | 0.5392 | 0.01456563 | 0.5787 |
| | 32 | 0.00398064 | 1.1582 | 0.00644239 | 1.1769 |
| | 64 | 0.00129888 | 1.6157 | 0.00209229 | 1.6225 |
| | 128 | 0.00037668 | 1.7859 | 0.00060590 | 1.7879 |

as the contrast is increased. Past $\alpha = 0.0001$ or so, the error essentially remains the same. We can explain this in terms of the area affected by the singularity. As the contrast grows larger, the singular behavior gets more severe, but it also gets more localized. We can see this from the analytical estimate (99) and from Fig. 10. Therefore, once the singular behavior is sufficiently localized due to high enough contrast, the infinite element captures this with very little error. The insensitivity to contrast due to the infinite element mesh is the great advantage of this method.

As described for the corner problem, we can again refine the local mesh inside the element, rather than the global mesh, in order to gain a substantial error reduction. In Table 12, we show how the overall error changes when the basis functions are refined. We see that the convergence rate with respect to the local grid size h_m actually increases and approaches 2. This shows that the overall error is indeed coming from the lack of refinement at the singularity and that the infinite element method can deal with the severe singularities very well.

We already mentioned that, in general, it is hard to obtain a well-resolved solution when there is a strong singularity. However, for the checkerboard problem,

Table 13. Comparisons to the homogenized solution obtained using $a^* = \sqrt{\alpha_1 \alpha_2}$.

| | n | m | l^2 error | % error | l^∞ error | % error |
|--------|-----|-----|-------------|---------|------------------|---------|
| Linear | 64 | 16 | 0.950445 | 93 | 2.183628 | 93 |
| IEM | 64 | 8 | 0.146707 | 14 | 0.379787 | 16 |
| | 64 | 16 | 0.064785 | 6.4 | 0.189092 | 8.1 |
| | 64 | 32 | 0.032905 | 3.2 | 0.106049 | 4.5 |
| | 64 | 64 | 0.023659 | 2.3 | 0.067518 | 2.9 |

we know that the homogenized coefficient is $\sqrt{\alpha_1 \alpha_2}$, where α_1 and α_2 are the alternating coefficients [Keller (1987)]. The final example uses this result to verify the performance of the infinite element method. We pick some boundary condition and forcing, e.g., $f(x, y) = 1$ in the domain and $u(x, y) = \sin(1.5\pi x) \sin(\pi(y + 0.5))$ on the boundary. In Table 13, we show how different methods perform. We let $n = 64$, which corresponds to $\delta = 1/64$, and $\alpha_1 = 1, \alpha_2 = 0.001$ in this case. We see that the linear method completely fails to capture the correct behavior, and the error is $\mathcal{O}(1)$. On the other hand, using the infinite elements gives the correct averaging, and the error is very small, down to 2.3% when $m = 64$. As before, the solution improves as the mesh inside the element is refined. As α_2 decreases, the size of the solution grows, but the relative error stays nearly the same. The relative errors are 8.9%, 8.1%, 7.3% for $\alpha_2 = 0.01, 0.001, 0.0001$, respectively. Again, piecewise constant boundary conditions from the one-dimensional problem appear to work well.

5. Conclusions

The approach we propose in the present work is to construct the basis functions in the finite element setting that contain the correct local behavior. We studied two important examples in detail to demonstrate that this framework can lead to efficient numerical methods: the boundary and internal layer structure in the convection-diffusion equation and the singular behavior in the elliptic equation. Simple extensions of successful one-dimensional approaches are not capable of capturing complicated two-dimensional effects, and general adaptive mesh algorithms lead to large, often ill-conditioned matrix problems.

When the characteristic length scale of the problem is small enough to be contained in an element basis, the multiscale finite element method works well, as the solution of the homogeneous equation solved locally reflects the global behavior correctly. This is the case for the periodic cellular flow we considered. To obtain the correct estimate of the convergence rate, we considered the subtle cancellation of errors in the discrete equation. For the cellular flow, the leading order term of the relative error was found to be $\mathcal{O}(\delta/h)$, where δ is the spatial oscillation and h the coarse grid mesh size. Thus the method converges as $\delta \rightarrow 0$ for a fixed h , regardless of the small diffusivity ϵ . We then try to improve the multiscale method further

by constructing asymptotic basis functions, since solving them numerically may become expensive in the limit $\epsilon \rightarrow 0$. Although we can construct such bases successfully for the cellular flow, numerical methods have to be employed to construct the multiscale basis in general. When the flow has scale separation and periodic structure, one can construct semi-analytic multiscale bases by solving the periodic cell problem numerically. This would reduce the overhead in computing the multiscale bases. For the random flow obtained by a perturbation of the cellular structure, the layer structure is nonlocal and the multiscale basis function cannot produce the necessary basis functions by sampling a small region. But we found that using a variational principle, exponential basis functions with correct scaling can be constructed and added to the finite element space. When these extra functions are incorporated correctly, we can capture the correct leading order behavior in the physical solution with a relatively coarse grid.

The same approach of adaptive basis functions provides a natural setting for dealing with the singular behavior of an elliptic problem. For the discontinuous coefficient or non-smooth interface problems with high contrast, the basis functions can be constructed efficiently using the infinite element method. Based on the local self-similarity of a specialized grid, it provides an accurate way of capturing the behavior near a singular point. Because most of the error comes from the neighborhood of singularities, we found that the local refinement of the infinite element results in the reduction of the global error and that this can be done efficiently involving only the points on the element boundary. We applied this method to problems with corners and the checkerboard singularities and showed that the method is efficient and robust.

Throughout this work, the main idea has been that when the finite element space is enriched with the basis functions that capture the correct local behavior, we can use a relatively coarse grid and yet still capture the leading order singular behavior in a robust manner. This idea should be directly applicable to many other problems in which the effect of the singularity is local. Further research would involve obtaining better predictions regarding the location of the layers for the problems with longer characteristic length scale using such tools as variational principles, as well as dealing with layers that evolve in time.

Acknowledgments

We would like to thank Prof. Albert Fannjian, Prof. Yalchin Efendiev, and Dr. Xiao-Hui Wu for several stimulating discussions during the preparation of this work. This research was in part support by an NSF grant DMS-0073916 and an NSF ITR grant ACI-0204932.

References

- Allen, D. and Southwell, R. (1955). Relaxation methods applied to determining the motion in two dimensions of a viscous fluid past a fixed cylinder. *Quarterly Journal of Mechanics and Applied Mathematics*, **8**: 129–145.

- Anufriyev, A. P. and Fishman, V. M. (1982). Magnetic field structure in the two-dimensional motion of a conducting field. *Geomagnetism and Aeronomy*, **22**: 245–248.
- Apelian, C., Holmes, R. L. and Avellaneda, M. (1997). A turbulent transport model: Streamline results for a class of random velocity fields in the plane. *Communications on Pure and Applied Mathematics*, **50**: 1053–1088.
- Avellaneda, M. (1991). Enhanced diffusivity and intercell transition layers in 2-d models of passive advection. *Journal of Mathematical Physics*, **32**: 3209–3212.
- Avellaneda, M. and Majda, A. J. (1991). An integral-representation and bounds on the effective diffusivity in passive advection by laminar and turbulent flows. *Communications in Mathematical Physics*, **138**: 339–391.
- Avellaneda, M., Torquato, S. and Kim, I. C. (1991). Diffusion and geometric effects in passive advection by random arrays of vortices. *Physics of Fluids*, **3**: 1880–1891.
- Babuška, I., Caloz, I. and Osborn, J. E. (1994). Special finite element methods for a class of second order elliptic problems with rough coefficients. *SIAM J. Numer. Anal.*, **31**: 945–981.
- Babuška, I. and Osborn, J. E. (1985). Finite element methods for the solution of problems with rough input data. In Grisvard, P., Wendland, W. and Whiteman, J. R., editors, *Singularities and Constructive Methods for Their Treatment: Proceedings of the Conference Held in Oberwolfach, West Germany, November 20–26, 1983*, pp. 1–18. Springer-Verlag, Berlin.
- Balk, A. M. and McLaughlin, R. M. (1999). Passive scalar in a random wave field: The weak turbulence approach. *Physics Letters A*, **256**: 299–306.
- Bank, R. E. (1990). *PLTMG: A Software Package for Solving Elliptic Partial Differential Equations*. SIAM Books, Philadelphia.
- Barrett, R., Berry, M., Chan, T. F., Demmel, J., Donato, J., Dongarra, J., Eijkhout, V., Pozo, R., Romine, C., and Van der Vorst, H. (1994). *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, 2nd Edition*. SIAM, Philadelphia, PA.
- Bensoussan, A., Lions, J.-L. and Papanicolaou, G. C. (1978). *Asymptotic Analysis for Periodic Structures*. North Holland, Amsterdam.
- Borcea, L. and Papanicolaou, G. C. (1997). A hybrid numerical method for high contrast conductivity problems. *Journal of Computational and Applied Mathematics*, **87**: 61–77.
- Brenner, S. C. and Scott, L. R. (1994). *The Mathematical Theory of Finite Element Methods*. Springer-Verlag, New York.
- Brezzi, F. and Russo, A. (1994). Choosing bubbles for advection-diffusion problems. *Math. Models Methods Appl. Sci.*, **4**: 571–587.
- Cao, L. Q., Cui, J. Z. and Zhu, D. C. (2002). Multiscale asymptotic analysis and numerical simulation for the second order Helmholtz equations with rapidly oscillating coefficients over general convex domains. *SIAM J. Numer. Anal.*, **40**: 543–577.
- Chen, Z. and Zou, J. (1998). Finite element methods and their convergence for elliptic and parabolic interface problems. *Numerische Mathematik*, **79**: 175–202.
- Childress, S. (1979). Alpha-effect in flux ropes and sheets. *Physics of the Earth and Planetary Interiors*, **20**: 172–180.
- Childress, S. and Soward, A. M. (1989). Scalar transport and alpha-effect for a family of cats-eye flows. *Journal of Fluid Mechanics*, **205**: 99–133.
- De Zeeuw, P. M. (1990). Matrix-dependent prolongations and restrictions in a blackbox multigrid solver. *Journal of Computational and Applied Mathematics*, **33**: 1–27.
- Dorobantu, M. and Engquist, B. (1998). Wavelet-based numerical homogenization. *SIAM J. Numer. Anal.*, **35**: 540–559.

- Douglas, Jr., J. and Russell, T. F. (1982). Numerical methods for convection-dominated diffusion problems based on combining the method of characteristics with finite element or finite difference procedures. *SIAM J. Numer. Anal.*, **19**: 871–885.
- E, W. and Engquist, B. (2003). The heterogeneous multi-scale methods. *Comm. Math. Sci.*, **1**: 87–133.
- Efendiev, Ya. E., Hou, T. Y. and Wu, X. H. (2000). Convergence of a nonconforming multiscale finite element method. *SIAM J. Numer. Anal.*, **37**: 888–910.
- Fannjiang, A. and Papanicolaou, G. C. (1994). Convection enhanced diffusion for periodic flows. *SIAM Journal on Applied Mathematics*, **54**: 333–408.
- Fannjiang, A. and Papanicolaou, G. C. (1996). Diffusion in turbulence. *Prob. Theory and Rel. Fields*, **105**: 279–334.
- Fannjiang, A. and Papanicolaou, G. C. (1997). Convection-enhanced diffusion for random flows. *Journal of Statistical Physics*, **88**: 1033–1076.
- Gilbarg, D. and Trudinger, N. S. (1983). *Elliptic Partial Differential Equations of Second Order*. Springer-Verlag, Berlin.
- Golub, G. H., Silvester, D. J. and Wathen, A. J. (1996). Diagonal dominance and positive definiteness of upwind approximations for advection diffusion problems. In Griffiths, D. F. and Watson, G. A., editors, *Numerical Analysis: A.R. Mitchell 75th Birthday Volume*, pp. 125–131. World Scientific, Singapore.
- Han, H. (1982). The numerical-solutions of interface problems by infinite element method. *Numerische Mathematik*, **39**: 39–50.
- Hegarty, A. F., O’Riordan, E. and Stynes, M. (1993). A comparison of uniformly convergent difference-schemes for two-dimensional convection diffusion-problems. *Journal of Computational Physics*, **105**: 24–32.
- Hou, T. Y. and Wu, X. H. (1997). A multiscale finite element method for elliptic problems in composite materials and porous media. *Journal of Computational Physics*, **134**: 169–189.
- Hou, T. Y., Wu, X. H. and Cai, Z. Q. (1999). Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients. *Mathematics of Computation*, **68**: 913–943.
- Hughes, T. J. R. (1995). Multiscale phenomena: Green’s functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods. *Comput. Methods Appl. Mech. Engrg.*, **127**: 387–401.
- Isichenko, M. B. (1992). Percolation, statistical topography, and transport in random-media. *Reviews of Modern Physics*, **64**: 961–1043.
- Isichenko, M. B. and Kalda, Ya. L. (1991). Statistical topography 2: Two-dimensional transport of a passive scalar. *Journal of Nonlinear Science*, **1**: 375–396.
- Jikov, V. V., Kozlov, S. M. and Oleinik, O. A. (1991). *Homogenization of Differential Operators and Integral Functionals*. Springer-Verlag, Berlin.
- Johnson, C. (1987). *Numerical Solution of Partial Differential Equations by the Finite Element Method*. Cambridge University Press, Cambridge, Great Britain.
- Keller, H. B. (1987). *Lectures on Numerical Methods in Bifurcation Problems*. Springer-Verlag, Berlin, for the Tata Institute of Fundamental Research.
- Keller, J. B. (1987). Effective conductivity of periodic composites composed of two very unequal conductors. *J. Math. Phys.*, **28**: 2516–2520.
- Kellogg, R. B. and Tsan, A. (1978). Analysis of some difference approximations for a singular perturbation problem without turning points. *Mathematics of Computation*, **32**: 1025–1039.
- Kraichnan, R. H. (1970). Diffusion by a random velocity field. *The Physics of Fluids*, **13**: 22–31.

- Leveque, R. J. and Li, Z. L. (1994). The immersed interface method for elliptic equations with discontinuous coefficients and singular sources. *SIAM J. Numer. Anal.*, **31**: 1019–1044.
- Lube, G. (1992). An asymptotically fitted finite element method for convection dominated convection-diffusion-reaction problems. *Zeitschrift für Angewandte Mathematik und Mechanik*, **72**: 189–200.
- Matache, A. M., Babuska, I. and Schwab, C. (2000). Generalized p -FEM in homogenization. *Numer. Math.*, **86**: 319–375.
- McLaughlin, D. W., Papanicolaou, G. C. and Pironneau, O. R. (1985). Convection of microstructure and related problems. *SIAM Journal of Applied Mathematics*, **45**: 780–797.
- Miller, J. J. H., O’Riordan, E. and Shishkin, G. I. (1996). *Fitted Numerical Methods for Singular Perturbation Problems*. World Scientific, Singapore.
- Morton, K. W. (1996). *Numerical Solution of Convection-Diffusion Problems*. Chapman and Hall, London.
- O’Riordan, E. and Stynes, M. (1991). A globally uniformly convergent finite element method for a singularly perturbed elliptic problem in two dimensions. *Mathematics of Computation*, **57**: 47–62.
- Park, P. J. (2000). *Multiscale Numerical Methods for the Singularly-Perturbed Convection-Diffusion Equation*. PhD thesis, California Institute of Technology.
- Roos, H. G., Adam, D. and Felgenhauer, A. (1996). A novel nonconforming uniformly convergent finite element method in two dimensions. *Journal of Mathematical Analysis and Applications*, **201**: 715–755.
- Roos, H.-G., Stynes, M. and Tobiska, L. (1991). *Numerical Methods for Singularly Perturbed Differential Equations*. Springer-Verlag, Berlin.
- Rosenbluth, M. N., Berk, H. L., Doxas, I. and Horton, W. (1987). Effective diffusion in laminar convective flows. *Physics of Fluids*, **80**: 2636–2647.
- Rüde, U. and Zenger, C. (1986). On the treatment of singularities in the multigrid method. In W. Hackbusch and U. Trottenberg, editors, *Multigrid Methods II*, pp. 261–271. Springer-Verlag, Berlin.
- Sacco, R. and Stynes, M. (1998). Finite element methods for convection-diffusion problems using exponential splines on triangles. *Computers and Mathematics with Applications*, **35**: 35–45.
- Sanchez-Palencia, E. (1980). *Non-homogenous Media and Vibration Theory*. Springer-Verlag, New York.
- Soward, A. M. (1987). Fast dynamo action in a steady flow. *Journal of Fluid Mechanics*, **180**: 267–295.
- Strang, G. W. and Fix, G. J. (1973). *An Analysis of the Finite Element Method*. Prentice-Hall, Englewood Cliffs, N. J.
- Stynes, M. and O’Riordan, E. (1991). An analysis of a singularly perturbed two-point boundary value problems using only finite element techniques. *Mathematics of Computation*, **56**: 663–675.
- Ying, L. (1995). *Infinite Element Methods*. Peking University Press, Beijing.