

CONVERGENCE OF A MULTISCALE FINITE ELEMENT METHOD FOR ELLIPTIC PROBLEMS WITH RAPIDLY OSCILLATING COEFFICIENTS

THOMAS Y. HOU, XIAO-HUI WU, AND ZHIQIANG CAI

ABSTRACT. We propose a multiscale finite element method for solving second order elliptic equations with rapidly oscillating coefficients. The main purpose is to design a numerical method which is capable of correctly capturing the large scale components of the solution on a coarse grid without accurately resolving all the small scale features in the solution. This is accomplished by incorporating the local microstructures of the differential operator into the finite element base functions. As a consequence, the base functions are *adapted* to the local properties of the differential operator. In this paper, we provide a detailed convergence analysis of our method under the assumption that the oscillating coefficient is of two scales and is periodic in the fast scale. While such a simplifying assumption is *not* required by our method, it allows us to use homogenization theory to obtain a useful asymptotic solution structure. The issue of boundary conditions for the base functions will be discussed. Our numerical experiments demonstrate convincingly that our multiscale method indeed converges to the correct solution, independently of the small scale in the homogenization limit. Application of our method to problems with continuous scales is also considered.

1. INTRODUCTION

In this paper, we consider solving a class of two-dimensional, second order, elliptic boundary value problems with highly oscillatory coefficients. Such equations often arise in composite materials and flows in porous media. In practice, the oscillatory coefficients may contain many scales spanning over a great extent [6]. When a standard finite element or finite difference method is used to solve the equations, the degrees of freedom of the resulting discrete system can be extremely large, due to the necessary resolution for achieving meaningful (convergent) results. Limited by computing resources, many practical problems are still out of reach using direct simulations. On the other hand, it is often the large scale features of the solution and the averaged effect of small scales on large scales that are of the main interest. Thus, it is desirable to have a numerical method that can capture the effect of small scales on large scales without resolving the small scale details.

Received by the editor August 5, 1996 and, in revised form, November 10, 1997.

1991 *Mathematics Subject Classification*. Primary 65F10, 65F30.

Key words and phrases. Multiscale base functions, finite element, homogenization, oscillating coefficients.

This work is supported in part by ONR under the grant N00014-94-0310, by DOE under the grant DE-FG03-89ER25073, and by NSF under the grant DMS-9704976.

Here, we introduce a multiscale finite element method for solving partial differential equations with highly oscillating solutions. Our purpose is to capture the large scale structures of the solutions correctly and effectively, without the above restrictions. This is achieved by constructing the finite element base functions from the leading order homogeneous elliptic equation. Typically, the size of the element is larger than a certain cut-off scale of the oscillatory coefficient. Information at scales smaller than the mesh size is built into the base functions. In our multiscale method, the base functions can be very oscillatory. It is through these oscillatory base functions that we capture the small scale effect on the large scales. The small scale information within each element is brought into the large scale solution through the coupling of the global stiffness matrix. Thus, the large scale solution is correctly computed.

The main result of this paper is a sharp error estimate of our multi-scale finite element method for general 2-D elliptic problems with separable two-scale coefficients. More specifically, we consider a model equation with periodic coefficients, which depend on a small parameter ϵ determining the small scale of the problem. In this case, homogenization theory can be used to describe the structure of the solution. This multiscale solution structure plays a crucial role in our convergence analysis. We are particularly interested in the situation where the mesh size h is larger than ϵ . The convergence analysis shows clearly how the small scale information in the base functions leads to the correct large scale solution. Using homogenization theory, we show that the multiscale method approximates the homogenized solution with second order accuracy in the limit as $\epsilon \rightarrow 0$. This result cannot be obtained from the classical analysis alone, which only provides an overly pessimistic estimate $O(h^2/\epsilon^2)$.

Our analysis also reveals an important phenomenon common in many upscaling methods, i.e., the resonance between the mesh scale and the small scale in the physical solution. A straightforward implementation of the multiscale finite element method would fail to converge when the mesh scale is close to the small scale in the physical solution. A deeper analysis shows that the boundary layer in the first order corrector seems to be the main source of the resonance effect. By choosing the boundary conditions for the base function properly, it is possible to eliminate the boundary layer in the first order corrector. This would give rise to a nice conservative difference structure in the discretization. It can be shown that this conservative difference structure would lead to *cancellation of resonance errors* and give an improved rate of convergence independent of the small scales in the solution.

This improved convergence is essential for problems with continuous scales, since the mesh scale inevitably coincides with one of the small scales in the solution. Motivated by the analysis in this paper, we propose in a companion paper, [16], an over-sampling strategy which eliminates the resonance error and gives rise to a multiscale method truly independent of small scale features in the solution. Our extensive numerical experiments show that the multiscale method with the over-sampling strategy indeed gives convergent results even for problems with continuous scales. The issues of computational efficiency, parallel implementation, and comparison with other existing methods are also discussed in detail in [16].

It should be pointed out that the idea of using base functions governed by the differential equations, especially its application to the convection-diffusion equation with boundary layers, has been studied for some years (see, e.g., [5] and references therein). Similar idea applied to elliptic equations has also been considered by

Babuška et al. in [3] for 1-D problems and in [2] for a special class of 2-D problems with the coefficient varying locally in one direction using the approach outlined in [4]. However, most of these methods are based on the special properties of the harmonic average in one-dimensional elliptic problems. As indicated by our convergence analysis, there is a fundamental difference between one-dimensional problems and genuinely multi-dimensional problems. Special complications such as the resonance between the mesh scale and the physical scale never occur in the corresponding 1-D problems.

There are also other related methods based on homogenization theory and up-scaling arguments. When the coefficients are periodic or quasi-periodic with scale separation, the averaged solution can be obtained by solving the so-called cell problems, which are analytically derived from the homogenization theory (ref. [8, 17]). Such an approach has been successfully used to solve some practical problems in porous media and composite materials [11, 9]. However, it has a limited range of applications; it is not applicable to general problems without scale separation. It is also expensive to use when the number of separable scales is large [16]. Based on some simple physical and mathematical motivations, numerical upscaling methods have been devised and applied to more general problems with random coefficients (cf. [10, 20]). But the design principle is strongly motivated by the homogenization theory for periodic structures. Their applications to non-periodic structures are not always guaranteed to work. There has also been success in achieving numerical homogenization for some semi-linear hyperbolic systems, the incompressible Euler equations, and 1-D elliptic problems using the sampling technique (see e.g., [14, 12, 1]). However, the sampling technique still has its own limitations. Its application to general 2-D elliptic problems is still not satisfactory.

The rest of the paper is organized as follows. The formulations of the 2-D model problem and the multiscale finite element method are introduced in the next section. Some observations are made for the model problem and the multiscale method in order to motivate later analysis. The homogenization theory of the model equation is reviewed in §3. These results are used in §§4 and 5, where the convergence analyses of the 2-D problem for $h < \epsilon$ and $h > \epsilon$ cases are presented respectively. In §6, the asymptotic structure of the discrete linear system of equations is studied; it shows how the boundary conditions of the base functions may affect the convergence rate. Some numerical results are given in §7; they provide strong support for our analytical estimates. Derivation of conservative difference structures for the discrete 2-D system is provided in the Appendix.

2. FORMULATIONS

In this section, we introduce the model problem and the multiscale method. First, we establish some notation and conventions. In the following, the Einstein summation convention is used: summation is taken over repeated indices. Throughout the paper, we use the $L^2(\Omega)$ based Sobolev spaces $H^k(\Omega)$ equipped with norms and seminorms

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

$H_0^1(\Omega)$ consists of those functions in $H^1(\Omega)$ that vanish on $\partial\Omega$. $H^{-1}(\Omega)$ is the dual of $H_0^1(\Omega)$. We define $H^{1/2}(\partial\Omega)$ as the trace on $\partial\Omega$ of all functions in $H^1(\Omega)$;

a norm is given by $\|v\|_{1/2,\partial\Omega} = \inf \|u\|_{1,\Omega}$, where the infimum is taken over all $u \in H^1(\Omega)$ with trace v . Throughout, C (with and without a subscript) denotes a generic positive constant, which is independent of ϵ and h unless otherwise stated. *Throughout the paper, we assume that Ω is a unit square domain in R^2 , which satisfies the convexity assumption needed to obtain certain regularity properties for elliptic operators.*

2.1. Model problem and the multiscale method. Consider the following elliptic model problem:

$$(2.1) \quad L_\epsilon u_\epsilon = f \quad \text{in } \Omega, \quad u_\epsilon = 0 \quad \text{on } \partial\Omega,$$

where

$$L_\epsilon = \nabla \cdot a(\mathbf{x}/\epsilon) \nabla$$

is the linear elliptic operator, ϵ is a small parameter, and $a(\mathbf{x}) = (a_{ij}(\mathbf{x}))$ is symmetric and satisfies $\alpha|\xi|^2 \leq \xi_i a_{ij} \xi_j \leq \beta|\xi|^2$, for all $\xi \in R^2$ and with $0 < \alpha < \beta$. Furthermore, $a_{ij}(\mathbf{y})$ are periodic functions in \mathbf{y} in a unit cube Y and $a_{ij}(\mathbf{y}) \in W^{1,p}(Y)$ ($p > 2$). Below, for simplicity of notation we use u instead of u_ϵ (except in §3), keeping in mind that u depends on ϵ .

The variational problem of (2.1) is to seek $u \in H_0^1(\Omega)$ such that

$$(2.2) \quad a(u, v) = f(v), \quad \forall v \in H_0^1(\Omega),$$

where

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

It is easy to see that the linear form $a(\cdot, \cdot)$ is elliptic and continuous, i.e.,

$$(2.3) \quad \alpha |v|_{1,\Omega}^2 \leq a(v, v), \quad \forall v \in H_0^1,$$

and

$$(2.4) \quad |a(u, v)| \leq \beta |u|_{1,\Omega} |v|_{1,\Omega}, \quad \forall u, v \in H_0^1.$$

A finite element method is obtained by restricting the weak formulation (2.2) to a finite dimensional subspace of $H_0^1(\Omega)$. For $0 < h \leq 1$, let \mathcal{K}^h be a partition of Ω by rectangles K with diameter $\leq h$, which is defined by an axi-parallel rectangular mesh. In each element $K \in \mathcal{K}^h$, we define a set of nodal basis $\{\phi_K^i, i = 1, \dots, d\}$ with $d (= 4)$ being the number of nodes of the element. We will neglect the subscript K when working in one element. In our multiscale method, ϕ^i satisfies

$$(2.5) \quad L_\epsilon \phi^i = 0 \quad \text{in } K \in \mathcal{K}^h.$$

Let $\mathbf{x}_j \in \overline{K}$ ($j = 1, \dots, d$) be the nodal points of K . As usual, we require $\phi^i(\mathbf{x}_j) = \delta_{ij}$. One needs to specify the boundary condition of ϕ^i for well-posedness of (2.5), for which we refer to §5.1. For now, we assume that the base functions are continuous across the boundaries of the elements, so that

$$V^h = \text{span}\{\phi_K^i : i = 1, \dots, d; K \in \mathcal{K}^h\} \subset H_0^1(\Omega).$$

In the following, we study the approximate solution of (2.2) in V^h , i.e., $u^h \in V^h$ such that

$$(2.6) \quad a(u^h, v) = f(v), \quad \forall v \in V^h.$$

Remark 2.1. The above formulation of the multiscale method is not restricted to rectangular elements; it can be applied to triangular elements, which are more flexible in modeling more complicated geometries. In fact, in most of the analysis below, the shape of element is irrelevant, except in §6 where we find that the triangular elements have some advantages in discrete error cancellations.

2.2. General observations. As mentioned in the introduction, the purpose of the multiscale method is to capture the large scale solution. This general idea can be made more precise in the context of the above model problem. In this case, there are two distinct scales in the solution, which are characterized by 1 and ϵ . The large scale solution is nothing but the homogenized solution u_0 (see §3), which is the limit of u as $\epsilon \rightarrow 0$. In fact, u equals u_0 up to $O(\epsilon)$ perturbations. It can be shown that u_0 is the solution of a homogenized elliptic problem with constant coefficient; thus u_0 is smooth. The oscillations at the ϵ -scale are contained in the perturbations. Because the base functions ϕ^i are defined by the same operator L_ϵ , it is expected that they have local structure similar to that of u . Such a property of ϕ^i is the key to the present method (see §5.2).

The multiscale base functions are smooth if $h \ll \epsilon$ and can be well approximated by the standard continuous linear (bilinear) base functions. Thus, we expect the multiscale method to behave similarly to linear finite element methods. In §4.1, we apply the standard finite element analysis to the multiscale method with one particular choice of the base functions (ϕ^i can be chosen differently by selecting different boundary conditions on ∂K), and the result supports the expectation.

On the other hand, when $h \gg \epsilon$, ϕ^i contains a smooth part and an oscillatory part, which *cannot* be approximated by linear (bilinear) functions. In this case, the multiscale method is very different from conventional finite element methods. In fact, we will show that the multiscale method gives solutions that converge to u_0 in the limit as $\epsilon \rightarrow 0$, while the standard finite element method with piecewise polynomial base functions does not. An intuitive explanation is as follows. The effective coefficient a^* , which determines the homogenized operator, includes both the average of a and the averaged result of the interaction of small scale oscillations (see (3.7)). Polynomial base functions can only capture the first part (hence the wrong operator), because they do not characterize any oscillations. In contrast, multiscale base functions contain the small scale information *in the same fashion* as u does. Therefore, they are able to accurately capture a^* through the variational formulation. This informal argument is explored in more detail in §§5 and 6.

It is interesting to note that the convergence of the multiscale method when $h > \epsilon$ cannot be obtained from the standard finite element analysis, whose estimate is too pessimistic (like (4.14)). The reason is that the standard analysis does not take into account the structure of the solution nor that of the base functions. This situation is rectified in §§5 and 6.

3. HOMOGENIZATION AND RELATED ESTIMATES

In this section, we review the homogenization theory of equation (2.1) and recall some estimates obtained by Moskow and Vogelius [21]. These results reveal the structure of the solution and the multiscale functions. In §5, the estimates will be used both globally for u and element-wise for ϕ_K^i .

Let us consider the following elliptic problem:

$$(3.1) \quad L_\epsilon u_\epsilon = f, \quad \text{in } \Omega, \quad u_\epsilon = g, \quad \text{on } \partial\Omega,$$

where $f \in L^2(\Omega)$ and $g \in H^{\frac{1}{2}}(\partial\Omega)$. Following [8, 21], we may write the second order equation for u_ϵ as a first order system:

$$\begin{aligned} a(\mathbf{x}/\epsilon)\nabla u_\epsilon - \mathbf{v}_\epsilon &= 0, \\ -\nabla \cdot \mathbf{v}_\epsilon &= f, \end{aligned}$$

and look for a formal expansion of the form

$$\begin{aligned} u_\epsilon &= u_0(\mathbf{x}, \mathbf{x}/\epsilon) + \epsilon u_1(\mathbf{x}, \mathbf{x}/\epsilon) + \cdots, \\ \mathbf{v}_\epsilon &= \mathbf{v}_0(\mathbf{x}, \mathbf{x}/\epsilon) + \epsilon \mathbf{v}_1(\mathbf{x}, \mathbf{x}/\epsilon) + \cdots, \end{aligned}$$

where $u_j(\mathbf{x}, \mathbf{y})$ and $v_j(\mathbf{x}, \mathbf{y})$ are periodic in the “fast” variable $\mathbf{y} = \mathbf{x}/\epsilon$.

Introducing $\nabla = \nabla_{\mathbf{x}} + 1/\epsilon \nabla_{\mathbf{y}}$, substituting the expansion into the above system of equations, and collecting terms with the same power of ϵ , we get

$$(3.2) \quad O(\epsilon^{-1}) : \quad a(\mathbf{y})\nabla_{\mathbf{y}}u_0 = 0,$$

$$(3.3) \quad O(\epsilon^{-1}) : \quad -\nabla_{\mathbf{y}} \cdot v_0 = 0,$$

$$(3.4) \quad O(\epsilon^0) : \quad a(\mathbf{y})\nabla_{\mathbf{y}}u_1 + a(\mathbf{y})\nabla_{\mathbf{x}}u_0 - v_0 = 0,$$

$$(3.5) \quad O(\epsilon^0) : \quad -\nabla_{\mathbf{y}} \cdot v_1 - \nabla_{\mathbf{x}} \cdot v_0 = f_0.$$

From (2.1) and (3.2)–(3.5), we have $u_0 = u_0(\mathbf{x})$ satisfying

$$(3.6) \quad \nabla \cdot a^* \nabla u_0 = f \quad \text{in } \Omega, \quad u_0 = g \quad \text{on } \partial\Omega,$$

where a^* is a constant (effective coefficient) matrix, given by

$$(3.7) \quad a^*_{ij} = \frac{1}{|Y|} \int_Y a_{ik}(\mathbf{y}) (\delta_{kj} - \frac{\partial}{\partial y_k} \chi^j) d\mathbf{y};$$

χ^j is the periodic solution of

$$(3.8) \quad \nabla_{\mathbf{y}} \cdot a(\mathbf{y}) \nabla_{\mathbf{y}} \chi^j = \frac{\partial}{\partial y_i} a_{ij}(\mathbf{y})$$

with zero mean, i.e., $\int_Y \chi^j d\mathbf{y} = 0$. As shown in [8], a^* is symmetric and positive definite. Denote the homogenized operator as

$$L_0 = \nabla \cdot a^* \nabla;$$

thus $L_0 u_0 = f_0$. In addition, we have

$$(3.9) \quad u_1(\mathbf{x}, \mathbf{y}) = -\chi^j \frac{\partial u_0}{\partial x_j}.$$

Note that $u_0(\mathbf{x}) + \epsilon u_1(\mathbf{x}, \mathbf{y}) \neq u_\epsilon$ on $\partial\Omega$, due to the construction of u_1 . Thus, we introduce a first order correction term θ_ϵ , satisfying

$$(3.10) \quad L_\epsilon \theta_\epsilon = 0 \quad \text{in } \Omega, \quad \theta_\epsilon = u_1(\mathbf{x}, \mathbf{y}) \quad \text{on } \partial\Omega,$$

so that $u_0 + \epsilon(u_1(\mathbf{x}, \mathbf{y}) - \theta_\epsilon)$ satisfies the boundary condition of u_ϵ .

The following lemma is proved in Proposition 1 of [21] for convex polygonal domains and $a_{ij}(\mathbf{y}) \in C^\infty(Y)$:

Lemma 3.1. *Let u_0 denote the solution to (3.6) and suppose $u_0 \in H^2(\Omega)$; let $u_1(\mathbf{x}, \mathbf{y})$ be given by (3.9), and let $\theta_\epsilon \in H^1(\Omega)$ denote the solution to (3.10). There exists a constant C , independent of u_0 , ϵ and Ω , such that*

$$(3.11) \quad \|u_\epsilon - u_0 - \epsilon(u_1 - \theta_\epsilon)\|_{1,\Omega} \leq C\epsilon |u_0|_{2,\Omega}.$$

From the above lemma, one can easily deduce the following corollary [21].

Corollary 3.2. *Under the assumptions of Lemma 3.1, there exists a constant C , independent of u_0 and ϵ , such that*

$$(3.12) \quad \|u_\epsilon - u_0\|_{0,\Omega} \leq C\epsilon \|u_0\|_{2,\Omega}.$$

Remark 3.3. In obtaining the above estimates, we don't really need to use the strong regularity assumption, $a_{ij}(\mathbf{y}) \in C^\infty(Y)$. In fact, it is sufficient to assume that $a_{ij}(\mathbf{y}) \in W^{1,p}(Y)$ ($p > 2$). In this case, we still have $\chi^j(\mathbf{y}) \in W^{2,p}(Y) \cap C^{1,\alpha}(\bar{Y})$ with $\alpha = 1 - n/p$ (cf. Theorem 15.1 in [18]), which is sufficient to give the above results without modifying the proof in [21]. This improvement on the regularity assumption is important from practical considerations, since the a_{ij} in general are not very smooth. It is also of practical interest to study the case when the a_{ij} are only piecewise smooth and have jump discontinuities across certain interfaces. The result depends on the geometry of the jump interfaces. We conjecture that (3.11) and (3.12) still hold if the interfaces are sufficiently smooth and disjoint. This and related issues are currently under study in [13].

4. CONVERGENCE FOR $h < \epsilon$

As noted earlier, the multiscale method and the standard linear finite element method are closely related when $h \ll \epsilon$. This relation is explored in this section. With some modifications, the standard finite element analysis can be carried out for the multiscale method. For simplicity, we assume the ϕ^i are linear along ∂K .

4.1. Error estimates. Subtracting (2.6) from (2.2), we get the orthogonality property

$$(4.1) \quad a(u - u^h, v) = 0, \quad \forall v \in V^h.$$

First, we have Céa's lemma.

Lemma 4.1. *Let u and u^h be the solutions of (2.2) and (2.6) respectively. Then*

$$(4.2) \quad \|u - u^h\|_{1,\Omega} \leq C \frac{\beta}{\alpha} \|u - v\|_{1,\Omega}, \quad \forall v \in V^h.$$

Proof. (4.2) is an immediate consequence of the Poincaré-Friedrichs inequality, (2.3), (2.4), and (4.1). □

Next, we study the approximation property of the base functions. Define the local interpolant of a function u on $K \in \mathcal{K}^h$ as

$$(\mathcal{I}_K u)(\mathbf{x}) = \sum_{j=1}^d u(\mathbf{x}_j) \phi^j(\mathbf{x}),$$

where \mathbf{x}_j are the nodal points of K . The global interpolant of u is then defined by

$$\mathcal{I}_{\mathcal{K}^h} u|_K = \mathcal{I}_K u \quad \forall K \in \mathcal{K}^h.$$

For convenience, we denote both the local and global interpolant as u_I . equation (2.5) and linearity give that

$$(4.3) \quad \nabla \cdot a(\mathbf{x}/\epsilon)\nabla u_I = 0 \quad \text{in } K \in \mathcal{K}^h.$$

To obtain the interpolation error, we use the regularity estimate,

$$(4.4) \quad |u|_{2,\Omega} \leq (C/\epsilon)\|f\|_{0,\Omega},$$

which can be shown by following the proof of Lemma 7.1 in [18] (see also [13]). The $1/\epsilon$ factor in (4.4) is due to the small scale oscillations in u as shown by the multiple scale expansion of u in §3.

Lemma 4.2. *Let $u \in H^2(\Omega)$ be the solution of (2.1), and $u_I \in V^h$ be its interpolant varying linearly along ∂K . There exist constants $C_1 > 0$ and $C_2 > 0$, independent of h , such that*

$$(4.5) \quad \|u - u_I\|_{0,\Omega} \leq C_1(h^2/\epsilon)\|f\|_{0,\Omega},$$

$$(4.6) \quad \|u - u_I\|_{1,\Omega} \leq C_2(h/\epsilon)\|f\|_{0,\Omega}.$$

Proof. Let u_l be the standard bilinear interpolant of u in $K \in \mathcal{K}^h$. From approximation theory we have

$$(4.7) \quad \|u - u_l\|_{0,K} \leq C_1 h^2 |u|_{2,K},$$

$$(4.8) \quad |u - u_l|_{1,K} \leq C_2 h |u|_{2,K}.$$

In the following, we examine the difference between u_l and u_I on K . Since $u_l|_{\partial K} = u_I|_{\partial K}$, we have $u_l - u_I \in H_0^1(K)$, and by the Poincaré-Friedrichs inequality we get

$$(4.9) \quad \|u_l - u_I\|_{0,K} \leq C_3 h |u_l - u_I|_{1,K},$$

which, together with (2.1), (4.3), (4.8), and the boundedness of a_{ij} , yields

$$\begin{aligned} \alpha |u_l - u_I|_{1,K}^2 &\leq \int_K \nabla(u_l - u_I) \cdot a \nabla(u_l - u_I) d\mathbf{x} \\ &= - \int_K (u_l - u_I) \nabla \cdot a \nabla(u_l - u_I) d\mathbf{x} \\ &= - \int_K (u_l - u_I) (\nabla \cdot a \nabla(u_l - u) - f) d\mathbf{x} \\ &= \int_K \nabla(u_l - u_I) \cdot a \nabla(u_l - u) d\mathbf{x} + \int_K (u_l - u_I) f d\mathbf{x} \\ &\leq C |u_l - u_I|_{1,K} |u_l - u|_{1,K} + \|u_l - u_I\|_{0,K} \|f\|_{0,K} \\ &\leq |u_l - u_I|_{1,K} (CC_1 h |u|_{2,K} + C_3 h \|f\|_{0,K}). \end{aligned}$$

Thus, we obtain

$$(4.10) \quad |u_l - u_I|_{1,K} \leq \frac{h}{\alpha} (CC_1 |u|_{2,K} + C_3 \|f\|_{0,K}),$$

and hence by (4.9)

$$(4.11) \quad \|u_l - u_I\|_{0,K} \leq \frac{h^2}{\alpha} C_3 (CC_1 |u|_{2,K} + C_3 \|f\|_{0,K}).$$

Therefore, using the triangle inequality, (4.8), and (4.10), we get

$$|u - u_I|_{1,K} \leq |u - u_l|_{1,K} + |u_l - u_I|_{1,K} \leq h(C_1 |u|_{2,K} + C_2 \|f\|_{0,K}),$$

where the constants are redefined and depend on a . Thus,

$$\begin{aligned}
 |u - u_I|_{1,\Omega} &= \left(\sum_{K \in \mathcal{K}^h} |u - u_I|_{1,K}^2 \right)^{\frac{1}{2}} \\
 &\leq h \left(2 \sum_{K \in \mathcal{K}^h} (C_1^2 |u|_{2,K}^2 + C_2^2 \|f\|_{0,K}^2) \right)^{\frac{1}{2}} \\
 &\leq \sqrt{2} h (C_1 |u|_{2,\Omega} + C_2 \|f\|_{0,\Omega}) \\
 &\leq C(h/\epsilon) \|f\|_{0,\Omega},
 \end{aligned}
 \tag{4.12}$$

where in the last step (4.4) is used. equation (4.5) can be derived similarly. Then, (4.6) follows from (4.5) and (4.12). \square

From (4.2) and (4.6) we immediately have

Corollary 4.3. *Let u and u^h be the solutions of (2.1) and (2.6), respectively. Then there exists a constant C , independent of h and ϵ , such that*

$$\|u - u^h\|_{1,\Omega} \leq C(h/\epsilon) \|f\|_{0,\Omega}.
 \tag{4.13}$$

Moreover, using the standard Aubin-Nitsche trick, one obtains

Theorem 4.4. *Let u and u^h be the solutions of (2.1) and (2.6), respectively. Then there exists a constant C , independent of h and ϵ , such that*

$$\|u - u^h\|_{0,\Omega} \leq C(h/\epsilon)^2 \|f\|_{0,\Omega}.
 \tag{4.14}$$

4.2. Comments. In the above proof, u_l is used as an intermediate step towards the final results. As a consequence, the results rely on the H^2 regularity of u , which is the source of the h/ϵ terms in (4.14). There may be other ways of obtaining the estimates, but the results given above seem to be sharp in general, as supported by the numerical computations. From (4.8), it is easy to show that the same convergence holds for the linear finite element method, which confirms our previous observation. It should be noted that (4.14) is useful only when $h < \epsilon$ (see the next section).

On the other hand, if $u_I = u$ on ∂K , which is exactly the situation in 1-D, then $u - u_I \in H_0^1(K)$ and u_l is not needed. Following similar steps as above, it can be shown that

$$\|u - u^h\|_{1,\Omega} \leq Ch \|f\|_{0,\Omega},
 \tag{4.15}$$

which is independent of $|u|_{2,\Omega}$ in contrast to (4.13). Furthermore, the L^2 estimate becomes $\|u - u^h\|_{0,\Omega} \leq Ch^2 \|f\|_{0,\Omega}$, independent of ϵ .

An interesting superconvergence result, i.e. $u^h \equiv u$ at the nodal points, can also be obtained as follows. It is easy to check that for any $v \in V^h$

$$a(u_I, v) = a(u, v) = f(v) \quad (1\text{-D}).$$

Thus, by (2.6) and choosing $v = u_I - u^h$, we obtain $a(u_I - u^h, u_I - u^h) = 0$. This implies that $u^h = u_I$, the desired result.

However, in multi-dimensions, u_I and u are in general not equal on ∂K . This is the *essential* difference between one and multi-dimensional problems. For special quasi 2-D problems, such as those considered in [2], Babuška et al. were able to obtain (4.15) without using the H^2 regularity of u . This was done by carefully selecting some special base functions.

5. CONVERGENCE FOR $h > \epsilon$

From §4.1 we see that the multiscale method behaves similarly as the standard finite element method when $h < \epsilon$. However, in the following we show that the two methods behave very differently in the limit as $\epsilon \rightarrow 0$. We obtain the estimates for the case of $h > \epsilon$ by exploring the asymptotic structure in both u and ϕ^i , without which the standard analysis fails to give the correct estimate, e.g., the estimates given above imply that both methods fail to converge if $h > \epsilon$.

We note that for $h > \epsilon$, if a^* is a constant, then from the homogenization theory in §3, the multiscale base function ϕ^i consists of the standard linear function plus an order ϵ oscillatory part. As discussed later, the oscillations do not match with those of the solution in general. By Taylor expansion types of argument, it is easy to see that the interpolation accuracy of the multiscale base and the linear base functions are the same (see also §4.2). Thus, from the conventional approximation point of view, multiscale bases do not seem to be superior to linear bases. This paradox will be clarified in the analysis below, where we see that standard approximation theory alone is not sufficient for analyzing the multiscale method.

5.1. The boundary condition of base functions. In §4, we used a linear boundary condition for ϕ^i . Another choice of boundary conditions is to solve the reduced elliptic problems on each side of ∂K with boundary conditions 1 and 0 at the two end points, and use the resulting solution as the boundary condition for the base function. The reduced problems are obtained from (2.5) by deleting terms with partial derivatives in the direction normal to ∂K and having the coordinates normal to ∂K fixed as parameters. We call such boundary conditions for ϕ oscillatory boundary conditions. It is clear that the reduced problems are of the same form as (2.5). In case of a being separable in space, i.e., $a(\mathbf{x}) = a_1(\frac{x}{\epsilon})a_2(\frac{y}{\epsilon})$, the base function ϕ^i with oscillatory b.c., μ^i , can be computed analytically by forming a tensor product. It is easy to verify that the resulting elements are conforming and that $\sum_{i=1}^d \mu^i \equiv 1$ on ∂K , and hence

$$(5.1) \quad \sum_{i=1}^d \phi_K^i \equiv 1 \quad \forall K \in \mathcal{K}^h.$$

The same is true if μ^i is linear. We will see that (5.1) is important in obtaining tight L^2 estimates. For simplicity, we only present the proof in the case when μ^i is linear. The estimates also hold if the above oscillatory boundary condition is used. In §7 we will give numerical examples of using the oscillatory μ^i , which in some cases leads to significant improvement in the accuracy of numerical results.

5.2. H^1 estimates. We make the following observations. Due to (2.5), ϕ^i can be expanded as

$$(5.2) \quad \phi^i = \phi_0^i + \epsilon \phi_1^i - \epsilon \theta^i + \dots,$$

where

$$(5.3) \quad L_0 \phi_0^i = 0 \quad \text{in } K, \quad \phi_0^i = \mu^i \quad \text{on } \partial K,$$

$$(5.4) \quad \phi_1^i = -\chi^j \frac{\partial \phi_0^i}{\partial x_j},$$

$$(5.5) \quad L_\epsilon \theta^i = 0 \quad \text{in } K, \quad \theta^i = \phi_1^i \quad \text{on } \partial K.$$

Thus, the ϕ_0^i form a set of finite element bases for solving (3.6). We denote by $V_0^h \in H_0^1(\Omega)$ the space spanned by the ϕ_0^i . From the standard finite element analysis, we see that u_0 can be approximated by ϕ_0^i with first and second order accuracy in the H^1 and L^2 norms, respectively. Furthermore, by comparing (5.4) with (3.9), it can be seen that u_1 can be well approximated by ϕ^1 . Thus, we have a match between the solution and the base function up to the ϵ order. Our main result is

Theorem 5.1. *Let u and u^h be the solutions of (2.1) and (2.6), respectively. Then there exist constants C_1 and C_2 , independent of ϵ and h , such that*

$$(5.6) \quad \|u - u^h\|_{1,\Omega} \leq C_1 h \|f\|_{0,\Omega} + C_2 (\epsilon/h)^{\frac{1}{2}}.$$

Proof. The estimate of solution error (5.6) follows immediately from Cea's Lemma (Lemma 4.1) and following interpolation Lemma 5.3. \square

Remark 5.2. Throughout this section, we will use u_I as the interpolant of the homogenized solution, u_0 , using the multiscale base functions ϕ^i . This is different from the definition of u_I in the previous section.

Lemma 5.3. *Let u be the solution of (2.1) and $u_I \in V^h$ the interpolant of the homogenized solution u_0 , using the multiscale base functions ϕ^i . Then there exist constants C_1 and C_2 , independent of ϵ and h , such that*

$$(5.7) \quad \|u - u_I\|_{1,\Omega} \leq C_1 h \|f\|_{0,\Omega} + C_2 (\epsilon/h)^{\frac{1}{2}}.$$

To estimate $\|u - u_I\|_{1,\Omega}$, we match the expansions of u_I and u and use Lemma 3.1. By (2.5), we have $L_\epsilon u_I = 0$ in K . Thus u_I can be expanded in K as

$$u_I = u_{I0} + \epsilon u_{I1} - \epsilon \theta_{I\epsilon} + \dots,$$

where $L_0 u_{I0} = 0$ in K and

$$(5.8) \quad u_{I1} = -\chi^j \frac{\partial u_{I0}}{\partial x_j}.$$

Clearly, $u_{I0} \in V_0^h$ is an interpolant of u_0 :

$$(5.9) \quad u_{I0} = \sum_{i=1}^d u_0(\mathbf{x}_i) \phi_0^i.$$

Moreover, since $u_I = u_{I0}$ on ∂K , the first order correction $\theta_{I\epsilon}$ satisfies

$$(5.10) \quad L_\epsilon \theta_{I\epsilon} = 0 \quad \text{in } K, \quad \theta_{I\epsilon} = u_{I1} \quad \text{on } \partial K.$$

We have

Lemma 5.4. *Let $u_I \in V^h$ be the interpolant of u_0 using base functions ϕ^i . u_{I0} and u_{I1} are defined by (5.9) and (5.8), respectively. Denote by $\theta_{I\epsilon}$ the solution of (5.10). There exists a constant C , independent of ϵ and h , such that*

$$(5.11) \quad \|u_I - u_{I0} - \epsilon u_{I1} + \epsilon \theta_{I\epsilon}\|_{1,\Omega} \leq C \epsilon \|f\|_{0,\Omega}.$$

Proof. From (3.6) and elliptic regularity theory, we have

$$(5.12) \quad |u_0|_{2,\Omega} \leq C \|f\|_{0,\Omega}.$$

Moreover, by standard approximation theory,

$$|u_{0I}|_{2,K} \leq |u_0 - u_{0I}|_{2,K} + |u_{0I}|_{2,K} \leq C |u_0|_{2,K},$$

where u_{0l} is the bilinear interpolant of u_0 . Thus, from $L_0(u_{I0} - u_{0l}) = L_0u_{0l}$, we get

$$(5.13) \quad |u_{I0}|_{2,K} \leq C|u_0|_{2,K},$$

which together with Lemma 3.1 yields

$$\|u_I - u_{I0} - \epsilon u_{I1} + \epsilon \theta_{I\epsilon}\|_{1,K} \leq C\epsilon|u_{I0}|_{2,K} \leq C_1\epsilon|u_0|_{2,K}.$$

Taking the sum of the last inequality over \mathcal{K}^h and using (5.12), we get (5.11). \square

Now consider the expansion of u :

$$u = u_0 + \epsilon u_1 - \epsilon \theta_\epsilon + \cdots.$$

Lemma 3.1 and (5.12) imply that

$$(5.14) \quad \|u - u_0 - \epsilon u_1 + \epsilon \theta_\epsilon\|_{1,\Omega} \leq C\epsilon\|f\|_{0,\Omega}.$$

Thus, using the expansions of u and u_I , (5.11), (5.14), and the triangle inequality, we have

$$(5.15) \quad \begin{aligned} \|u - u_I\|_{1,\Omega} &\leq \|u_0 - u_{I0}\|_{1,\Omega} + \|\epsilon(u_1 - u_{I1})\|_{1,\Omega} \\ &\quad + \|\epsilon(\theta_\epsilon - \theta_{I\epsilon})\|_{1,\Omega} + C\epsilon\|f\|_{0,\Omega}. \end{aligned}$$

Therefore, we need to estimate the first three terms on the right hand side of (5.15). By (5.9), (4.6), and (5.12), we have

$$(5.16) \quad \|u_0 - u_{I0}\|_{1,\Omega} \leq Ch\|f\|_{0,\Omega}.$$

Noting that $\chi^j(\mathbf{y}) \in C^{1,\alpha}(\bar{Y})$, $\|\chi^j\|_{L^\infty(\Omega)} \leq C$, we get

$$\|\epsilon(u_1 - u_{I1})\|_{0,\Omega} \leq C\epsilon h\|f\|_{0,\Omega}.$$

Next, since $\|\nabla\chi^j\|_{L^\infty(K)} \leq C/\epsilon$, there exist C_1 and C_2 such that

$$\begin{aligned} \|\epsilon(u_1 - u_{I1})\|_{1,K} &\leq \epsilon(|u_0 - u_{I0}|_{1,K} \sum_{j=1}^d \|\nabla\chi^j\|_{L^\infty(K)} + |u_0 - u_{I0}|_{2,K} \sum_{j=1}^d \|\chi^j\|_{L^\infty(K)}) \\ &\leq C_1|u_0 - u_{I0}|_{1,K} + C_2\epsilon|u_0|_{2,K}. \end{aligned}$$

Thus, taking sum of the last inequality over \mathcal{K}^h and using (5.16) and (5.12), we have

$$(5.17) \quad \|\epsilon(u_1 - u_{I1})\|_{1,\Omega} \leq (C_1h + C_2\epsilon)\|f\|_{0,\Omega}.$$

For the last term on the right hand side of (5.15), we note that $\theta_{I\epsilon}$ does not match with θ_ϵ . This is because θ_ϵ is determined by the global boundary conditions imposed on u while $\theta_{I\epsilon}$ is determined locally. Thus, we cannot expect cancellation between $\theta_{I\epsilon}$ and θ_ϵ , and we need to estimate $\|\epsilon\theta_\epsilon\|_{1,\Omega}$ and $\|\epsilon\theta_{I\epsilon}\|_{1,\Omega}$ separately. Using a standard estimate for (3.10) (cf. [19]), we have

$$\|\epsilon\theta_\epsilon\|_{1,\Omega} \leq C\epsilon\|u_1\|_{1/2,\partial\Omega}.$$

From (3.9), $\|u_1\|_{1/2,\partial\Omega}$ can be calculated by using either the definition of $\|\cdot\|_{1/2,\partial\Omega}$ for a bounded domain or the interpolation inequality. We have $\|u_1\|_{1/2,\partial\Omega} \leq C\epsilon^{-1/2}$, and hence

$$(5.18) \quad \|\epsilon\theta_\epsilon\|_{1,\Omega} \leq C\sqrt{\epsilon}.$$

Similarly, for $\theta_{I\epsilon}$ in each $K \in \mathcal{K}^h$, we have

$$(5.19) \quad \|\epsilon \nabla \theta_{I\epsilon}\|_{0,K} \leq C\epsilon \|u_{I1}\|_{1/2,\partial K}.$$

equation (5.8) implies that

$$\|u_{I1}\|_{0,\partial K} \leq Ch^{\frac{1}{2}} \quad \text{and} \quad |u_{I1}|_{1,\partial K} \leq Ch^{\frac{1}{2}}\epsilon^{-1}.$$

Thus, using the interpolation inequality and summing over \mathcal{K}^h , we obtain

$$(5.20) \quad \|\epsilon \theta_{I\epsilon}\|_{1,\Omega} \leq h^{-\frac{1}{2}}(C_1\epsilon + C_2\epsilon^{\frac{1}{2}}).$$

Therefore, (5.7) follows from (5.15)–(5.20).

Remark 5.5. It should be noted that when $h < \epsilon$, the asymptotic expansion from homogenization theory is no longer useful. In this case, the expansion should be done in terms of h , returning to standard approximation theory (§4.1). It is interesting to see that since ϕ^i has the same asymptotic structure as u , the leading order term in $\|u - u^h\|_{1,\Omega}$ becomes ϵ/h , in contrast to h/ϵ from the standard analysis (see §4.1). Moreover, (5.6) indicates that as $\epsilon \rightarrow 0$, $\|u - u^h\|_{1,\Omega} = O(h)$. Therefore, the multiscale base solution *converges* to the correct solution in the homogenized limit.

Remark 5.6. The above analysis can be applied to the standard linear finite element method. It can be shown that there is an $O(1)$ term in $\|e_I\|_{1,\Omega}$ coming from $\|\epsilon(u_1 - u_{I1})\|_{1,\Omega}$ ($u_{I1} = 0$ in this case), which is independent of h and ϵ . Thus the linear finite element method does not converge as $\epsilon \rightarrow 0$ when $\epsilon < h$.

5.3. L^2 estimates. The Aubin-Nitsche trick can be used again to obtain the L^2 estimate. It follows from (5.7) and (5.6) that

$$(5.21) \quad \|u - u^h\|_{0,\Omega} \leq C_1 h^2 \|f\|_{0,\Omega} + C_2 (\epsilon/h)^{\frac{1}{2}}.$$

Note that the convergence rate is still dominated by the $\sqrt{\epsilon/h}$ term; thus no improvement is gained in (5.21) and the result is not sharp. Nevertheless, (5.21) shows that $u - u^h = O(h^2)$ as $\epsilon \rightarrow 0$. This observation points to another way of looking at the problem.

Let $u_0^h \in V_0^h$ be the Galerkin solution of (3.6). Since the μ^i are linear (see §5.1), by Theorem 4.4 one has

$$(5.22) \quad \|u_0 - u_0^h\|_{0,\Omega} \leq Ch^2 \|f\|_{0,\Omega}.$$

Moreover, (3.12), (5.22), and the triangle inequality imply

$$(5.23) \quad \begin{aligned} \|u - u^h\|_{0,\Omega} &\leq \|u - u_0\|_{0,\Omega} + \|u_0 - u_0^h\|_{0,\Omega} + \|u^h - u_0^h\|_{0,\Omega} \\ &\leq C_1 \epsilon \|u_0\|_{2,\Omega} + C_2 h^2 \|f\|_{0,\Omega} + \|u^h - u_0^h\|_{0,\Omega}. \end{aligned}$$

Therefore the problem becomes that of finding the L^2 -convergence from u^h to u_0^h . Note that intuitively the convergence in $L^2(\Omega)$ cannot be better than $O(\epsilon)$, due to the mismatch between θ_ϵ and $\theta_{I\epsilon}$ (see §5.2). Thus, in general, using (5.23) should not amplify the error bound significantly.

Next, we show that the order of convergence is determined by how well u^h approximates u_0^h at the discrete nodal points, i.e., the convergence analysis is transformed to a discrete one. For any $K \in \mathcal{K}^h$, consider

$$\begin{aligned}
 \|u^h - u_0^h\|_{0,K} &= \left\| \sum_{i=1}^d (u^h(\mathbf{x}_i)\phi^i - u_0^h(\mathbf{x}_i)\phi_0^i) \right\|_{0,K} \\
 (5.24) \quad &= \left\| \sum_{i=1}^d [(u^h(\mathbf{x}_i) - u_0^h(\mathbf{x}_i))\phi^i + u_0^h(\mathbf{x}_i)(\phi^i - \phi_0^i)] \right\|_{0,K} \\
 &\leq \sum_{i=1}^d |u^h(\mathbf{x}_i) - u_0^h(\mathbf{x}_i)| \|\phi^i\|_{0,K} + \left\| \sum_{i=1}^d u_0^h(\mathbf{x}_i)(\phi^i - \phi_0^i) \right\|_{0,K}.
 \end{aligned}$$

We will show that the global contribution of the second term on the right hand side of the inequality is $O(\epsilon)$. Denote $\tilde{u} = \sum_{i=1}^d u_0^h(\mathbf{x}_i)\phi^i$; then $L_\epsilon \tilde{u} = 0$. Hence \tilde{u} has a multiple scale expansion and, by (3.11),

$$\|\tilde{u} - \tilde{u}_0 - \epsilon \tilde{u}_1 + \epsilon \tilde{\theta}_\epsilon\|_{0,K} \leq C\epsilon |\tilde{u}_0|_{2,K},$$

where $\tilde{u}_0 = \sum_{i=1}^d u_0^h(\mathbf{x}_i)\phi_0^i$, $\tilde{u}_1 = -\chi^j \partial \tilde{u}_0 / \partial x_j$, and $\tilde{\theta}_\epsilon$ is the correction term, defined similarly as θ_{T_ϵ} (see (5.10)). It follows that

$$(5.25) \quad \|\tilde{u} - \tilde{u}_0\|_{0,\Omega} \leq \epsilon(C|\tilde{u}_0|_{2,\Omega} + \|\tilde{u}_1\|_{0,\Omega} + \|\tilde{\theta}_\epsilon\|_{0,\Omega}),$$

where $|\tilde{u}_0|_{2,\Omega}$ is defined by

$$|\tilde{u}_0|_{2,\Omega}^2 = \sum_{K \in \mathcal{K}^h} |\tilde{u}_0|_{2,K}^2.$$

Since $u_0^h(\mathbf{x}_i)$ is a smooth solution on the grid, i.e., its divided difference is bounded, one can verify that the right hand side of (5.25) is bounded by $C\epsilon$, where C is a constant independent of ϵ and h .

From (5.24), (5.25), and the fact that $\|\phi_K^i\|_{0,K} \leq Ch$, we have

$$\|u^h - u_0^h\|_{0,\Omega} \leq C \left(\sum_{i \in \mathcal{N}} (u^h(\mathbf{x}_i) - u_0^h(\mathbf{x}_i))^2 h^2 \right)^{\frac{1}{2}} + C\epsilon,$$

where \mathcal{N} is the set of indices of all nodal points on the mesh. Thus, $\|u^h - u\|_{0,\Omega}$ is $O(h^2 + \epsilon)$ plus the discrete l^2 norm of $(u^h - u_0^h)(\mathbf{x}_i)$. In the next section, we show that to the leading order the error in discrete l^2 norm is in general $O(\epsilon/h)$ by formally applying the multiple scale expansion technique to the discrete linear system of equations for $u^h(\mathbf{x}_i)$. The approach is not rigorous, but it reveals the insight of subtle error cancellation and points to possible ways of improving the multiscale method. Therefore, from (5.23) and (5.24) we may formally conclude that

$$(5.26) \quad \|u - u^h\|_{0,\Omega} \leq C_1 h^2 \|f\|_{0,\Omega} + C_2 \epsilon + C_3 \frac{\epsilon}{h} \quad (\epsilon < h).$$

Numerical tests are supplied in §7 to confirm (5.26).

Remark 5.7. The estimate (5.26) is a clear improvement over (5.21). It is interesting to note that in both cases, $h < \epsilon$ and $h > \epsilon$, the ratio between ϵ and h is the dominating factor in the error estimates. Indeed, there are two scales in the discrete problem, ϵ and h , and hence their ratio is an intrinsic parameter for the

problem. In case of $h > \epsilon$, the above results indicate that to capture the averaged behavior of L_ϵ , h needs to be sufficiently large compared to ϵ in order to get enough samples of the small scales. On the other hand, our numerical tests in §7 indicate that results better than (5.26) can be obtained in certain cases. This is due to some further error cancellation, which is analyzed in §6.

6. DISCRETE ERROR ANALYSIS

As shown in §5.3, the L^2 estimate of convergence is determined by the convergence of u^h to u_0^h at the nodal points. Denote by U^h the nodal point value of u^h . The linear system of equations for U^h is

$$(6.1) \quad A^h U^h = f^h,$$

where A^h and f^h are obtained from $a(u^h, v)$ and $f(v)$ by using $v = \phi^i$ for $i \in \mathcal{N}$. Our central aim in this section is to estimate the rate of convergence from U^h to the nodal values of u_0^h . and to investigate possible ways to improve the convergence rate by constructing better boundary conditions for the base functions.

The key ideas of our analysis can be best illustrated by studying the 1-D model problem, which is given in §6.1. As mentioned above (see §4.2), the 1-D problem is rather special, and conventional finite element analysis gives sharp estimates. The estimates, in fact, do *not* rely on homogenization theory or the solution structure. However, such an analysis in higher dimensions is impossible. Here, we use a formal asymptotic expansion to analyze (6.1). This approach reveals the subtle structure of the discrete system, which is otherwise unattainable from a conventional analysis. More important, the expansion is fully general and applicable to multi-dimensional problems.

After studying the 1-D problem, we briefly analyze the 2-D problem in §6.2. We will emphasize the new difficulty due to the 2-D effect and describe how to overcome it.

6.1. One-dimensional case. In this subsection, we present the analysis for the 1-D problem to show the structure of U^h and demonstrate the discrete error cancellation. These results are valid for 2-D problems with separable coefficients, where the base functions are the tensor products of the 1-D bases.

Consider the solution of

$$\frac{d}{dx} a\left(\frac{x}{\epsilon}\right) \frac{du}{dx} = f \quad \text{in } I = (0, 1),$$

with boundary conditions $u(0) = u(1) = 0$. Subdivide the interval I by a uniform partition

$$P : 0 = x_0 < x_1 < \dots < x_N = 1$$

with mesh size $h = 1/N$. Let $\mathcal{N} = \{0, 1, \dots, N\}$, $\mathcal{I} = \{x_i | i \in \mathcal{N}\}$, and $I_i = [x_i, x_{i+1}]$. The multiscale base function is defined as in 2-D by

$$(6.2) \quad -\frac{d}{dx} a \frac{d\phi^i}{dx} = 0, \quad \forall x \in I \setminus \mathcal{I}; \quad \phi^i(x_j) = \delta_{ij}, \quad \forall j \in \mathcal{N}.$$

Thus, the stiffness matrix is given by

$$(6.3) \quad A_{ij}^h = \int_{x_{i-1}}^{x_{i+1}} a(x/\epsilon) \frac{d\phi^i}{dx} \frac{d\phi^j}{dx} dx \quad (j = i - 1, i, i + 1).$$

Noting that $\phi^{i-1} + \phi^i \equiv 1$ in I_{i-1} for all i , we have

$$A_{ii}^h = -(A_{ii-1}^h + A_{ii+1}^h).$$

Moreover, since $A_{ij}^h = A_{ji}^h$, we may define $B_i^h = A_{ii-1}^h$; thus, $A_{ii+1}^h = A_{i+1i}^h = B_{i+1}^h$ and $A^h U^h$ can be written in a *conservative* form as

$$(6.4) \quad (A^h U^h)_i = D^+(B_i^h D^- U_i^h) \quad (i = 1, \dots, N-1),$$

where D^+ and D^- are the forward and backward difference operators. As shown below, this conservative form is a key in the discrete analysis. In addition, $A_0^h U_0^h$ can be written in the conservative form, and we can define $B_{0i}^h = A_{0ii-1}^h$ ($i = 1, \dots, N-1$).

We note that since ϵ is a small parameter in (6.1), we may formally expand A^h , f^h , and U^h with respect to ϵ :

$$(6.5) \quad A^h = A_0^h + \epsilon A_1^h + \dots, \quad f^h = f_0^h + \epsilon f_1^h + \dots, \quad U^h = U_0^h + \epsilon U_1^h + \dots.$$

These expansions are determined by using the expansion of ϕ^i , i.e., (5.2). It is straightforward to derive

$$(6.6) \quad B_i^h = -\frac{a^*}{h} - \epsilon \frac{a^*}{h^2} (\chi_i - \chi_{i-1}) + \dots,$$

where χ is the solution of the cell problem (see (3.8)) and $\chi_i = \chi(x_i/\epsilon)$. From the first term on the right hand side we see that A_0^h is the stiffness matrix for the homogenized equation. Similarly, expanding f^h we find that f_{0i}^h is given by $\int_0^1 f \phi_0^i dx$. Thus, $U_{0i}^h = u_0^h(x_i)$.

To determine the rate of convergence, we estimate the leading order error U_1^h . Let $G^h = (A_0^h)^{-1}$. The $O(\epsilon)$ terms in (6.5) give

$$U_1^h = G^h f_1^h - G^h A_1^h U_0^h.$$

Let us consider the second term here; the first term can be estimated similarly. To calculate $G^h A_1^h U_0^h$, we note that expanding (6.4) preserves the conservative structure at all orders. In particular, we have

$$(6.7) \quad (A_1^h U_0^h)_i = D^+(B_{1i}^h D^- U_{0i}^h),$$

where B_{1i}^h is given by the second term on the right hand side of (6.6) (without ϵ). It follows that

$$(6.8) \quad \begin{aligned} (G^h A_1^h U_0^h)_i &= \sum_{j=1}^{N-1} G_{ij}^h D^+(B_{1j}^h D^- U_{0j}^h) \quad (i = 1, \dots, N-1) \\ &= -\sum_{j=1}^N B_{1j}^h (D^- G_{ij}^h) (D^- U_{0j}^h) \quad (G_{i0} = G_{iN} = 0) \\ &= O(1/h). \end{aligned}$$

In the derivation, we have used the (piecewise) smoothness of G^h and U_{0j}^h , implying that $D^- G_{ij}^h$ and $D^- U_{0j}^h$ are $O(h)$. From this and a similar estimate for $G^h f_1^h$, we obtain

$$(6.9) \quad \|\epsilon U_1^h\|_{l^2} = O(\epsilon/h).$$

Remark 6.1. In the above derivation, besides utilizing the conservative structure and the summation by parts, which are independent of the spatial dimension, we used the properties of the 1-D discrete Green's function G^h . The 2-D G^h has a different structure. However, the $O(\epsilon/h)$ estimate still holds in 2-D (see §6.2).

We note that (6.8) can be further improved, since B_1^h has a difference structure (see (6.6)) and hence further summation by parts is possible. Indeed, we have

$$\begin{aligned} \sum_{j=1}^N B_{1j}^h(D^- G_{ij}^h)(D^- U_{0j}^h) &= -(a^*/h^2) \sum_{j=1}^N D^- \chi_j(D^- G_{ij}^h)(D^- U_{0j}^h) \\ &= (a^*/h^2)[\chi_0(D^- G_{i1}^h)(D^- U_{01}^h) \\ &\quad + \sum_{j=1}^{N-1} \chi_j D^+((D^- G_{ij}^h)(D^- U_{0j}^h)) \\ &\quad - \chi_N(D^- G_{iN}^h)(D^- U_{0N}^h)] \\ &= O(1). \end{aligned}$$

Note that in the last step we have used

$$D^+ D^- G_{ij}^h = h^2 \delta_{ij} / a^* \quad \text{and} \quad D^+ D^- U_{0j}^h = O(h^2).$$

Thus, we see that *the difference form is essential for improving the estimate*. A better estimate for $G^h f_1^h$ can be derived similarly. Here, we just mention that f_1^h has difference form due to the fact that the slope of ϕ_0^i has opposite signs in the intervals I_{i-1} and I_i .

Thus, (6.9) now improves to

$$(6.10) \quad \|\epsilon U_1^h\|_{l_2} = O(\epsilon),$$

which is *independent* of ϵ and h . In practice this is very important, since for problems with many scales, h is inevitably close to one of the small scales. We will see in the next subsection that a similar estimate (6.14) holds also in 2-D under certain conditions.

6.2. Two-dimensional error cancellation. As remarked above, (6.9) is valid in 2-D. Here we just outline the proof, since it is essentially the same as the one presented in the previous subsection. The details are given in Appendix A.

Following the 1-D argument, one can show that the leading order contribution of U^h in 2-D is U_0^h . Moreover, due to (5.1), $A_1^h U_0^h$ can be written in a generalized conservative form:

$$(6.11) \quad (A_1^h U_0^h)_{ij} = \sum_{s=1}^4 (D_s^+ B_{ij}^s D_s^-) U_{0ij}^h,$$

where (i, j) is the 2-D index for grid points, and D_s^+ and D_s^- are forward and backward difference operators in i, j , and the two diagonal directions (see (A.10)). The role of B^s is similar to B_1^h in (6.7). Using (6.11) and summation by parts along the directions according to D_s^+ , one obtains (6.9), provided that the divided differences of the discrete Green's function G^h , i.e., $D_s^- G^h$ ($s = 1, \dots, 4$), are absolutely summable.

To show that $D_s^- G^h$ are indeed summable, we need to use some results for the regularized Green's function \mathcal{G} for the homogenized operator L_0 , obtained by

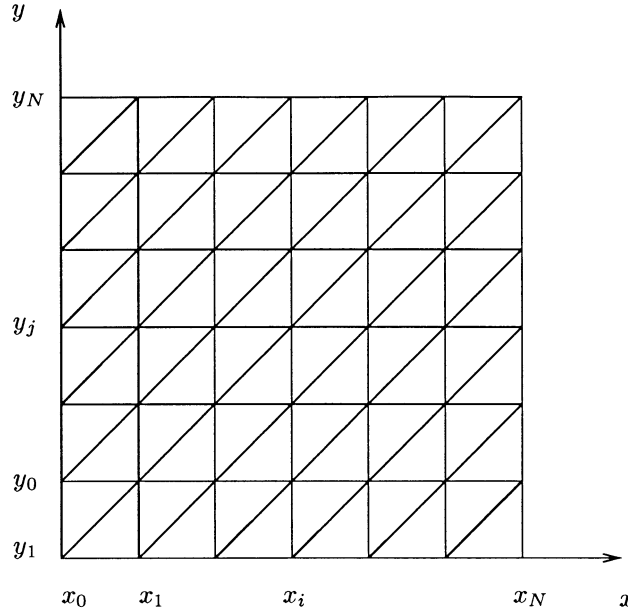


FIGURE 1. Rectangular mesh with triangulation.

Frehse and Rannacher [15]. Their results were obtained for triangular elements. For simplicity, we consider a triangulation given by Figure 1.

In this case, V_0^h is identical to the linear finite element space considered in [15]. There should be no difficulty in extending the results to rectangular elements. Without getting into the details, we only cite the important estimates that are relevant for our purpose; for details of the definitions and derivations, see [15]. One of the main results in [15] is the following two estimates:

$$(6.12) \quad \|\mathcal{G} - \mathcal{G}^h\|_{1,1,\Omega} \leq Ch|\log(h)| \quad \text{and} \quad \|D^2\mathcal{G}\|_{0,1,\Omega} \leq C|\log(h)|,$$

where \mathcal{G}^h is the Galerkin projection of \mathcal{G} on V_0^h ; $\|\cdot\|_{0,1,\Omega}$ and $\|\cdot\|_{1,1,\Omega}$ are 0 and 1 norms based on $L^1(\Omega)$ respectively (indicated by the second subscripts). We see that G^h consists of the nodal values of \mathcal{G}^h . From (6.12), the fact that $|\nabla\mathcal{G}|$ is integrable (independently of h), and the equivalence between $D_s^-G^h/h$ and $D_s\mathcal{G}^h$ (D_s being the directional derivative) in each K , one easily deduces that the divided differences of G^h are absolutely summable.

The more interesting but also more challenging result for 2-D is (6.10). Here, the effect of dimensionality shows up. We have seen that the key to improving (6.9) is to explore the difference structures in f_1^h and B_{ij}^s . The method of finding the difference forms in B^s and f_1^h is outlined in Appendix A. The idea is to convert volume integrals over K into boundary integrals over ∂K . Here we only consider the troublesome term,

$$(6.13) \quad \epsilon' \int_{\partial K'} \theta^k n_i a_{ij} \frac{\partial \theta^l}{\partial x'_j} ds' \quad (\epsilon' = \epsilon/h, x'_j = x_j/h),$$

which in general cannot be written in difference forms (see (A.12)).

To find a resolution to this problem, we need to further understand the structure of θ^k . Recall that $L'_\epsilon \theta^k = 0$ and $\theta^k = g(\mathbf{x}'/\epsilon')$ is highly oscillatory on $\partial K'$ (see (5.10)). This gives θ^k a boundary layer structure, which has been analyzed in [7] for the half space problem and further in [21] for problems in convex polygons. The analysis is complicated. The main result is that there exist a constant $\gamma > 0$ such that $\nabla \theta^k$ decays like $\exp(-\gamma x_n/\epsilon')$, where x_n is the coordinate along the inward normal to the boundary. From this, we may conclude that θ^k has a boundary layer with thickness $O(\epsilon')$. This observation is consistent with the $O(1/\sqrt{\epsilon'})$ estimate of $\|\theta^k\|_{1,K'}$.

In the 1-D problem, the θ^k have no boundary layers. In fact, to the leading order, they are linear functions. Therefore, $d\theta^k/dx$ do not appear in the A_1^h . The main difference between 1-D and 2-D problems is that the boundary condition on θ^k in 1-D is given at two nodal points and hence there is no oscillation in the boundary condition, whereas in 2-D it is given along the line segments of ∂K and the oscillation occurs.

The structure of θ^k is solely determined by its boundary condition, which in turn is determined by the boundary conditions of ϕ^k . Therefore, a judicious choice of μ^k may remove the boundary layer of θ^k . In this case, we would have $\theta_{,j}^k = O(1)$ on $\partial K'$. Thus (6.13) would become $O(\epsilon')$ and does not enter A_1^h . Such boundary conditions do exist, e.g., we may set $\phi^k = \phi_0^k + \epsilon \phi_1^k$ on the boundary of K , which enforces $\theta^k = 0$ in the element. We do not advocate such an approach, since it requires solving the cell problems to obtain ϕ_1^k . The point is that finding the appropriate boundary conditions is a *local* problem, which is determined by the properties of L_ϵ . In [16], we introduce an over-sampling technique to overcome the difficulty associated with the boundary layer of θ^k . This technique does not rely on the homogenization assumptions and can be applied to general multiple scale problems. Our extensive numerical experiments demonstrate that the multi-scale method with the over-sampling technique indeed gives convergent results for problems with continuous scales.

Furthermore, if a is separable in 2-D, the base functions can be constructed from the tensor products of the above 1-D bases. Such a construction corresponds to using the oscillatory μ^i (see §5.1) as the boundary condition for ϕ^i . In this case, it is easy to show that the 2-D θ^i do not have boundary layers. This is a special example of obtaining the appropriate boundary condition without solving the cell problem. On the other hand, if the linear boundary conditions are used, the θ^i will have boundary layers.

It is worth mentioning that in some cases, the special base functions used in [2] satisfy (2.5), although they were constructed differently. It can be shown that these base functions have the desirable boundary conditions in the sense that the θ^i have no boundary layers.

By using (6.12) it can be shown that, similarly to \mathcal{G} , the second order divided differences of G^h , i.e., $D_s^+ D_t^- G^h/h^2$ ($s, t = 1, \dots, 4$), are absolutely summable and the results are bounded by $C|\log(h)|$, where $C > 0$ is a constant independent of h . It follows that if B^s can be written in difference forms, using summation by parts one more time as in the 1-D case would give $G^h A_1^h U_0^h = O(h|\log(h)|)$, and hence we have

$$(6.14) \quad \|\epsilon U_1^h\|_{l^2} = O(\epsilon|\log(h)|) \quad (\text{in 2-D}).$$

This estimate is slightly weaker than (6.10), but it is still a great improvement over (6.9).

7. NUMERICAL EXPERIMENTS

In this section, we study the convergence and accuracy of the multiscale method through numerical computations. The model problem is solved using the multiscale method with base functions defined by both linear and oscillatory boundary conditions (see §5.1). Since it is very difficult to construct a test problem with exact solution and sufficient generality, we use resolved numerical solutions in place of exact solutions. The numerical results are compared with the theoretical analysis. In addition, we give a numerical example of applying the method to more general problems with nonseparable scales. For more practical applications of the method and other numerical issues, such as the parallel implementation and performance of the method, see [16].

7.1. Implementation. The implementation of the multiscale method is fairly straightforward. We outline the implementation here and define some notation to be frequently used below. We consider problems in a square domain with unit length. Let N be the number of elements in the x and y directions. The size of mesh is thus $h = 1/N$. To compute the base functions, each element is discretized into $M \times M$ subcell elements with size $h_s = h/M$. Rectangular elements are used in all numerical tests.

To solve the subcell problem, we use the standard linear finite element method. After solving the base functions, the local stiffness matrix A^e and the right hand side vector f^e are computed from (A.1) using numerical quadrature rules. When the base is solved by using the finite element method, we compute the gradient of a base function at the center of a subcell element and use the two-dimensional centered trapezoidal rule for the volume integration. This procedure ensures that the entries of A^e (hence of A^h) are computed with second order accuracy. In our computations, we only solve three base functions, i.e., ϕ^i ($i = 1, 2, 3$). The fourth one is obtained from $\phi^4 = 1 - \sum_{i=1}^3 \phi^i$.

The amount of computation in the volume integral for A^e can be reduced by recasting it as a boundary integral using (2.5). However, it turns out that this approach may yield a global stiffness matrix that is not positive definite when the subcell resolution is not sufficiently high; large errors result from integration by parts.

We use the multigrid method with matrix dependent prolongation [22] to solve for the base functions and the linear system resulting from the multiscale finite element discretization. We also use this multigrid method to solve for a well-resolved solution of the model problem by the linear finite element method. The method is very robust for general 2-D second order elliptic equations (see [22]). Our numerical tests show that the rate of convergence for this multigrid method is almost independent of ϵ .

To facilitate the comparison among different schemes, we use the following shorthands: LFEM stands for the standard linear finite element method and MFEM stands for the multiscale finite element method. An addition letter appended to these shorthands, “L” or “O”, indicates that a linear or an oscillatory boundary condition, respectively, is specified for the base functions.

TABLE 1. Results of Example 7.1 ($h < \epsilon$, $M = 8$, $u_{max} \approx 0.02$).

| ϵ | N | MFEM-O | | MFEM-L | | LFEM | |
|------------|------|---------------|------|---------------|------|---------------|------|
| | | $\ E\ _{l_2}$ | rate | $\ E\ _{l_2}$ | rate | $\ E\ _{l_2}$ | rate |
| 0.08 | 64 | 5.60e-4 | | 6.90e-5 | | 2.55e-4 | |
| | 128 | 2.32e-4 | 1.3 | 1.58e-5 | 2.1 | 6.65e-5 | 1.9 |
| | 256 | 7.11e-5 | 1.7 | 3.58e-6 | 2.1 | 1.66e-5 | 2.0 |
| | 512 | 1.87e-5 | 1.9 | 7.09e-7 | 2.3 | 3.92e-6 | 2.1 |
| 0.04 | 128 | 5.82e-4 | | 5.65e-5 | | 2.39e-4 | |
| | 256 | 2.39e-4 | 1.3 | 1.23e-5 | 2.2 | 6.20e-5 | 1.9 |
| | 512 | 7.33e-5 | 1.7 | 2.71e-6 | 2.2 | 1.55e-5 | 2.0 |
| 0.02 | 256 | 5.92e-4 | | 5.10e-5 | | 2.32e-4 | |
| | 512 | 2.42e-4 | 1.3 | 1.08e-5 | 2.2 | 5.98e-5 | 2.0 |
| | 1024 | 7.42e-5 | 1.7 | 2.32e-6 | 2.2 | 1.50e-5 | 2.0 |

7.2. Numerical results. In all the examples below, the resolved solutions are obtained using LFEM. Given the wave length of the small scale ϵ , we solve the model problem twice on two meshes with one mesh size being twice the other. Then the Richardson extrapolation is used to approximate the exact solutions from the numerical solutions on the two meshes. Throughout our numerical experiments, both of the mesh sizes used to compute the well-resolved solution are less than $\epsilon/10$, so that the error of the extrapolated solutions is less than 10^{-7} . In Tables 1–9 we present the absolute error; the maxima of the solutions are given so as to provide a measure of the relative error. All computations are performed on the unit square domain $\Omega = (0, 1) \times (0, 1)$.

Example 7.1. In this example, we solve (2.1) with

$$(7.1) \quad a(\mathbf{x}/\epsilon) = \frac{1}{2 + P \sin(2\pi(x - y)/\epsilon)},$$

where P is a parameter controlling the magnitude of the oscillation. We take $P = 1.8$ in this example. The right hand side function $f(x, y)$ is given by

$$(7.2) \quad f(x, y) = -\frac{1}{2}[(6x^2 - 1)(y^4 - y^2) + (6y^2 - 1)(x^4 - x^2)].$$

On $\partial\Omega$, we impose $u = 0$. Here the effective coefficient a^* is a full 2×2 matrix. In Table 1, the results of convergence tests for $h < \epsilon$ are given, where $E = U - U^h$ is the error at the nodal points. Clearly, the MFEM-L is $O(h^2/\epsilon^2)$ as LFEM, which is consistent with (4.14); on the other hand, MFEM-O converges slower but the rate increases as h decreases. We note that the boundary condition for the base functions has a strong influence on the convergence of the method. Here, the linear boundary condition leads to significantly better accuracy. We will see more of such an effect below; see e.g., Table 4. The result of LFEM is listed in the table just for the purpose of reference. It should be emphasized that the multiscale method is not designed for resolved computations.

In Tables 2 and 3, the convergence of the method with $h > \epsilon$ is shown. Table 2 indicates that the error is proportional to h^{-1} , which together with results of Table 3 confirms the $O(\epsilon/h)$ estimates given in §6.1 and (5.26). In this case, the linear and oscillatory boundary conditions of the ϕ^i lead to similar accuracy.

TABLE 2. Results of Example 7.1 ($\epsilon = 0.005$).

| Mesh | | MFEM-O | | | MFEM-L | | |
|------|-----|----------------|---------------|-------|----------------|---------------|-------|
| N | M | $\ E\ _\infty$ | $\ E\ _{l^2}$ | rate | $\ E\ _\infty$ | $\ E\ _{l^2}$ | rate |
| 32 | 64 | 2.01e-4 | 8.29e-5 | | 2.73e-4 | 1.15e-4 | |
| 64 | 32 | 5.07e-4 | 2.07e-4 | -1.3 | 5.81e-4 | 2.44e-4 | -1.1 |
| 128 | 16 | 9.38e-4 | 3.87e-4 | -0.90 | 1.12e-3 | 4.74e-4 | -0.96 |
| 256 | 8 | 2.20e-3 | 9.13e-4 | -1.2 | 2.02e-3 | 8.63e-4 | -0.86 |

TABLE 3. Results of Example 7.1 ($\epsilon/h = 0.32$, $M = 32$).

| N | ϵ | MFEM-O | | MFEM-L | | LFEM | |
|-----|------------|---------------|-------|---------------|-------|------|---------------|
| | | $\ E\ _{l^2}$ | rate | $\ E\ _{l^2}$ | rate | MN | $\ E\ _{l^2}$ |
| 8 | 0.04 | 1.30e-4 | | 1.36e-4 | | 256 | 6.20e-5 |
| 16 | 0.02 | 1.63e-4 | -0.33 | 1.98e-4 | -0.54 | 512 | 5.98e-5 |
| 32 | 0.01 | 1.94e-4 | -0.25 | 2.31e-4 | -0.22 | 1024 | 5.88e-5 |
| 64 | 0.005 | 2.07e-4 | -0.09 | 2.44e-4 | -0.08 | 2048 | 5.83e-5 |

Moreover, we note that in Table 3 the ratio ϵ/h is of order 1, well beyond the validity regime for the asymptotic expansion. Still, the error is rather small. We repeat the test by doubling the ratio. We observe that the error also becomes twice as large. Therefore, we may estimate that the constant in front of ϵ/h is about 2×10^{-4} , which is rather small. In fact, from the table we see that the error of the multiscale solution is comparable to that of the LFEM solution, which is obtained on a fine mesh with $N^2 M^2$ elements of size h_s . We stress that the resolutions of LFEM and MFEM are different, although they use the same total number of elements at the fine level. The resolution of MFEM is h ; scales smaller than h are captured rather than resolved. In contrast, the LFEM resolves the smallest scale with h_s . The comparison in Table 3 indicates that the MFEM can produce fairly good results when the well resolved solutions are not obtainable due to practical limitations of computer memory. This makes the MFEM very useful in practice.

Example 7.2. Here, we use

$$a(\mathbf{x}/\epsilon) = \frac{1}{4 + P(\sin(2\pi x/\epsilon) + \sin(2\pi y/\epsilon))} \quad (P = 1.8),$$

$$f(x, y) = -1, \quad u = g(x, y) = \frac{\sqrt{4 - P^2}}{2}(x^2 + y^2) \quad \text{on } \partial\Omega.$$

In this case, a^* is a scalar constant and ϕ_0^i are bilinear functions. As shown in Table 4, MFEM-O becomes much more accurate than MFEM-L in the resolved convergence test. In addition, it can be seen clearly that MFEM-O converges faster than h^2/ϵ^2 (approximately $O(h^2/\epsilon)$), which can be attributed to the oscillatory boundary conditions for the base functions. The change of the behavior of MFEM-O is more clearly seen in Tables 5 and 6, which include the results for $h > \epsilon$. As shown by Table 6, MFEM-O converges like $O(\epsilon)$ in the maximum norm and even faster in the l^2 norm. In contrast, MFEM-L does not converge as we decrease the mesh size.

To better understand the above results, we examine how the structure of θ^k may affect the rate of convergence. For this purpose, we solve the base functions in the

TABLE 4. Results of Example 7.2 ($h < \epsilon$, $M = 8$, $u_{max} \approx 0.87$).

| ϵ | N | MFEM-O | | MFEM-L | | LFEM | |
|------------|------|---------------|------|---------------|------|---------------|------|
| | | $\ E\ _{l_2}$ | rate | $\ E\ _{l_2}$ | rate | $\ E\ _{l_2}$ | rate |
| 0.08 | 64 | 3.62e-5 | | 3.73e-4 | | 6.90e-4 | |
| | 128 | 8.94e-6 | 2.0 | 9.74e-5 | 1.9 | 1.79e-4 | 1.9 |
| | 256 | 2.24e-6 | 2.0 | 2.43e-5 | 2.0 | 4.48e-5 | 2.0 |
| 0.04 | 128 | 1.67e-5 | | 3.88e-4 | | 7.14e-4 | |
| | 256 | 4.11e-6 | 2.0 | 1.01e-4 | 1.9 | 1.85e-4 | 1.9 |
| | 512 | 8.62e-7 | 2.3 | 2.52e-5 | 4.0 | 4.62e-5 | 2.0 |
| 0.02 | 256 | 1.11e-5 | | 3.85e-4 | | 7.12e-4 | |
| | 512 | 2.74e-6 | 2.0 | 1.00e-4 | 1.9 | 1.84e-4 | 1.9 |
| | 1024 | 3.97e-7 | 2.8 | 2.49e-5 | 2.0 | 4.60e-5 | 2.0 |

TABLE 5. Results of Example 7.2 ($\epsilon = 0.005$).

| Mesh | | MFEM-O | | | MFEM-L | | |
|------|-----|----------------|---------------|-------|----------------|---------------|-------|
| N | M | $\ E\ _\infty$ | $\ E\ _{l_2}$ | rate | $\ E\ _\infty$ | $\ E\ _{l_2}$ | rate |
| 32 | 64 | 5.73e-5 | 1.09e-5 | | 5.73e-4 | 3.01e-4 | |
| 64 | 32 | 1.52e-5 | 2.09e-5 | -0.94 | 1.22e-3 | 6.68e-4 | -1.1 |
| 128 | 16 | 5.51e-5 | 2.46e-5 | -0.24 | 2.27e-3 | 1.25e-3 | -0.90 |
| 256 | 8 | 1.02e-4 | 5.34e-5 | -1.1 | 4.65e-3 | 2.57e-3 | -1.0 |

TABLE 6. Results of Example 7.2 ($\epsilon/h = 0.32$, $M = 32$).

| N | ϵ | MFEM-O | | | | MFEM-L | |
|-----|------------|----------------|------|---------------|------|----------------|---------------|
| | | $\ E\ _\infty$ | rate | $\ E\ _{l_2}$ | rate | $\ E\ _\infty$ | $\ E\ _{l_2}$ |
| 8 | 0.04 | 9.07e-4 | | 5.14e-4 | | 1.15e-3 | 2.60e-4 |
| 16 | 0.02 | 5.12e-4 | 0.82 | 1.39e-4 | 1.9 | 1.11e-3 | 5.82e-4 |
| 32 | 0.01 | 2.88e-4 | 0.83 | 4.63e-5 | 1.6 | 1.12e-3 | 6.50e-4 |
| 64 | 0.005 | 1.52e-4 | 0.92 | 2.09e-5 | 1.1 | 1.22e-3 | 6.68e-4 |

rescaled domain $K' = (0, 1) \times (0, 1)$. Specifically, we let the base function ϕ equal 1 at point $(1, 1)$ and 0 at other corners. We compute the base functions obtained by using linear and oscillatory boundary conditions, μ_l and μ_o , respectively. The contour plots are given in Figure 2, where $\epsilon' = 0.2$, the solid lines are for μ_o and the dash lines for μ_l . The two solutions are almost indistinguishable. In Figure 3, we plot the contours of the correctors, θ_l and θ_o , obtained from $\phi - \phi_0 + \epsilon' \chi^j \partial \phi_0 / \partial x_j'$. Here $\phi_0 = x'y'$ (bilinear) and the χ^j are obtained from numerical solutions of the cell problem (3.8). Clearly, the θ_l , corresponding to the boundary condition μ_l , has a boundary layer with a thickness of $O(\epsilon')$; whereas the θ_o , corresponding to the boundary condition μ_o , has a rather weak boundary layer. This test confirms our idea of obtaining error cancellation by removing the boundary layers in θ . On the other hand, numerically we found also cases in which the boundary layer does not hinder the convergence of MFEM. However, it is a general trend that weaker boundary layers correspond to the more accurate MFEM solutions.

It should be noted that in our convergence analysis the base functions are assumed to be exact. This is not the case in the above examples, nor in practice.

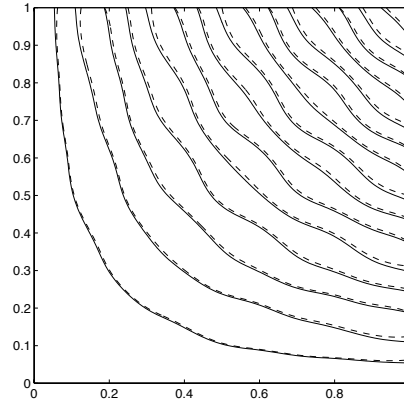


FIGURE 2. Multiscale base functions used in MFEM-O and MFEM-L for Example 7.2. Solid line: MFEM-O; dash line: MFEM-L.

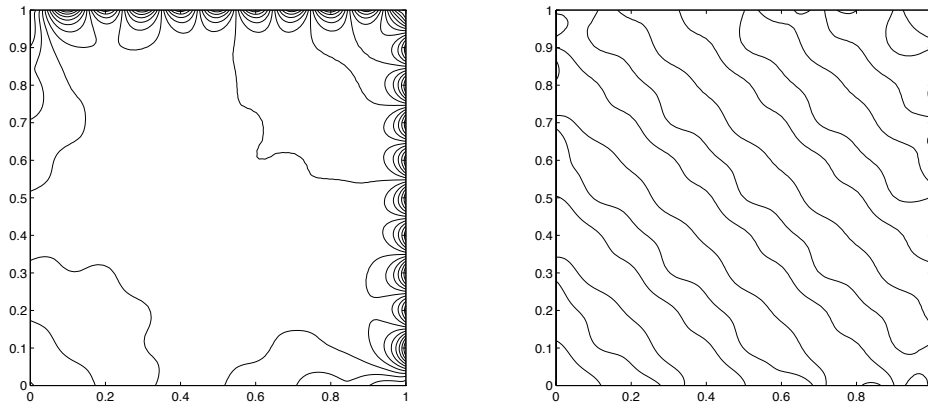


FIGURE 3. First order correctors of the base functions shown in Figure 2. Left: θ_l ; right: θ_o .

Therefore, it is desirable to know how the accuracy of base functions influences the final results. To this end, a Chebyshev spectral method was used for solving the base functions and computing the quadratures in A^h and f^h . We find that the accuracy of the final results is relatively insensitive to the accuracy of the base functions. In particular, second order accuracy in the base functions and in the numerical quadratures seems to be sufficient for obtaining accurate results. This is further demonstrated below.

Example 7.3. We choose the coefficient $a(x, y)$ to be separable in space, i.e.,

$$a(\mathbf{x}) = \frac{1}{(2 + P \sin(2\pi x/\epsilon))(2 + P \sin(2\pi y/\epsilon))},$$

for which we have $a^* = 1/(2\sqrt{4 - P^2})$; $f(x, y)$ and $g(x, y)$ are the same as in Example 7.2. In this case, the exact base functions are known analytically; we use MFEM-E to indicate MFEM with the exact bases. The comparison in Table 7

TABLE 7. Results of Example 7.3 ($\epsilon/h = 0.32$, $M = 32$, $u_{max} \approx 0.87$).

| N | ϵ | MFEM-O | | | | MFEM-E | | | |
|-----|------------|---------------|------|-----------------|------|---------------|------|-----------------|------|
| | | $\ E\ _{l_2}$ | rate | $\ E_0\ _{l_2}$ | rate | $\ E\ _{l_2}$ | rate | $\ E_0\ _{l_2}$ | rate |
| 16 | 0.04 | 1.03e-4 | | 4.37e-3 | | 6.92e-5 | | 4.30e-4 | |
| 32 | 0.02 | 4.73e-5 | 1.1 | 2.26e-3 | 1.0 | 2.72e-5 | 1.3 | 2.22e-4 | 1.0 |
| 64 | 0.01 | 2.24e-5 | 1.1 | 1.15e-3 | 1.0 | 1.03e-5 | 1.4 | 1.13e-4 | 1.0 |
| 128 | 0.005 | 1.13e-5 | 1.0 | 5.80e-4 | 1.0 | 4.19e-6 | 1.3 | 5.69e-4 | 1.0 |

TABLE 8. Results of Example 7.4 ($M = 32$, $u_{max} \approx 0.16$).

| N | ϵ | MFEM-O | | MFEM-L | | LFEM | |
|-----|------------|---------------|------|---------------|------|------|---------------|
| | | $\ E\ _{l_2}$ | rate | $\ E\ _{l_2}$ | rate | MN | $\ E\ _{l_2}$ |
| 8 | 0.04 | 1.04e-3 | | 9.44e-4 | | 256 | 2.73e-4 |
| 16 | 0.02 | 2.55e-4 | 2.0 | 1.67e-4 | 2.5 | 512 | 2.72e-4 |
| 32 | 0.01 | 4.86e-5 | 2.4 | 4.61e-5 | 1.9 | 1024 | 2.72e-4 |
| 64 | 0.005 | 1.05e-5 | 2.2 | 9.59e-5 | -1.1 | 2048 | 2.73e-4 |

clearly shows that the convergence and accuracy of the final results are not sensitive to the accuracy of the base functions. Note that E_0 in the table is the error compared with the homogenized solution u_0 .

Example 7.4. In this example, a non-smooth coefficient ($C^{0,\alpha}$)

$$a(x, y) = \frac{1}{1 + |\sin(2\pi x/\epsilon)| + |\sin(2\pi y/\epsilon)|}$$

is chosen. Moreover, $f = -1$ and $g = 0$. The convergence of MFEM solutions are shown in Table 8 for fixed $\epsilon/h = 0.32$ ($M = 32$). The LFEM solutions obtained on equivalent fine meshes are shown for comparison. We see that the multiscale finite element method handles non-smooth coefficients very well. In this case, we obtain a second order convergence for MFEM-O. At present, we don't have a full explanation for this "super" convergence. It is probably due to some additional cancellation of errors. We believe MFEM can also be applied to problems with discontinuous coefficients, as long we can obtain accurate approximation of the multiscale base functions. For interface problems, the geometry of the jump interfaces often has singularities (e.g., checkerboard problem). These singularities pose challenges to both conventional FEM and the multiscale finite element method. In the context of MFEM, the singularities need to be handled carefully in order to obtain accurate base functions. We are currently investigating effective methods for computing the multiscale base functions with complicated geometric singularities. This will be reported in a future paper.

Example 7.5. In this last example, we show some results of solving problems with nonseparable scales. The coefficient is chosen as

$$a(x, y) = 4 + P(\cos(2\pi \tanh(5(x - 0.5)))/\epsilon) + \sin(2\pi \tanh(5(y - 0.5)))/\epsilon),$$

and $f(x, y)$ and $g(x, y)$ are the same as in Example 7.2. Here the parameter ϵ is used to control the smallest scale of the problem. It can be seen that $a(x, y)$ oscillates more rapidly as x or y becomes close to 0.5. In the calculations presented in Table 9, we have used $P = 1$, and $\epsilon = 0.025$, so the smallest scale of the problem is

TABLE 9. Results of Example 7.5: l^2 norm error ($u_{max} \approx 1.73$).

| N | M | MFEM-O | MFEM-L | LFEM |
|-----|----|---------|---------|---------|
| 64 | 32 | 6.66e-5 | 3.71e-4 | 4.22e-3 |
| 128 | 16 | 2.54e-5 | 6.07e-4 | 1.59e-3 |
| 256 | 8 | 1.11e-5 | 7.82e-4 | 1.65e-3 |
| 512 | 4 | 4.53e-6 | 3.39e-4 | 6.58e-4 |

about 0.005. Again, we use LFEM to obtain a well resolved solution. As in the previous examples, the convergence and accuracy of MFEM depend on the boundary conditions of the base functions. In this case, the oscillatory boundary condition leads to better results. The under-resolved solutions using LFEM are also shown for comparison. As expected, the errors of LFEM solutions on a coarse grid are larger than those of MFEM solutions. One reason is that the linear finite element method tends to average out the small scale information and cannot correctly capture the scale interaction. We note that the difference between LFEM and MFEM-L is most significant when $N = 64$, which corresponds to $\epsilon/h = 0.32$. The discrepancy becomes smaller as h decreases, since the small scales are eventually resolved by the fine grid.

We remark that the purpose of the multiscale method is to provide a systematic approach to capture the small scale effect on the large scales when we cannot afford to resolve all the small scale features in the physical solution. In this regard, it is most relevant to test the performance of the method when h is sufficiently large compared to ϵ . In the computations shown in Table 9, the $N = 64$ case roughly satisfies this requirement. We clearly see that MFEM gives a superior performance than the corresponding LFEM. As before, we note that the boundary conditions have a great effect on the accuracy of the method. In [16] we study this important issue further and propose an over-sampling technique to eliminate the boundary layers in the first order corrector. This technique can be applied to general elliptic problems with many scales, and gives an improved rate of convergence for MFEM.

7.3. Remarks. The results of the above numerical experiments can be summarized as follows. The numerical results confirm our analysis of the multiscale method applied to the model problem (2.1). In particular, our numerical results confirm that the discrete error analysis is correct and sharp. Furthermore, the boundary condition on the base functions can have a significant effect on the convergence and accuracy of the multiscale method. At present, we do not have the optimum boundary condition, which is the target of our future research, but the oscillatory boundary condition devised in §5.1 seems to be useful for many problems. In any case, further study of the *local* boundary conditions for the base functions is very important for improving the multiscale method. The numerical results also indicate that the accuracy of the solution is insensitive to the accuracy of the base functions. It has been demonstrated that second order accurate base functions are sufficient for practical purposes.

APPENDIX A. EXPANSIONS OF THE 2-D DISCRETE LINEAR SYSTEM

In this appendix, we derive an asymptotic expansion for (6.1) in the form of (6.5). From (5.2) and $\partial\phi_0^i/\partial x_j = O(1/h)$, it is evident that the natural parameter for the expansion of ϕ^i is ϵ/h . From another point of view, ϕ^i is defined on elements

with size h . Thus we can obtain the same result by a scaling argument. For all $K \in \mathcal{K}^h$, define

$$K' = \{\mathbf{x}' : \mathbf{x}' = \mathbf{x}/h, \mathbf{x} \in K\}$$

and $\epsilon' = \epsilon/h$. Thus, $\text{diam}(K') = 1$, and $\mathbf{y} = \mathbf{x}/\epsilon = \mathbf{x}'/\epsilon'$ is the fast variable. We have $L'_\epsilon = \nabla_{\mathbf{x}'} \cdot a \nabla_{\mathbf{x}'}$ and $L'_\epsilon \phi^i = 0$. Therefore, ϕ^i can be expanded as

$$\phi^i = \phi_0^i + \epsilon' \phi_1^i - \epsilon' \theta^i + \dots,$$

where ϕ_0^i , ϕ_1^i , and θ^i are defined by (5.3), (5.4), and (5.5), respectively, with L_ϵ , x_j , and K being replaced by their rescaled counterparts.

To simplify the presentation, we use the ‘‘comma’’ notation below for the derivatives, e.g., $\phi_{,j}^i = \partial \phi^i / \partial x'_j$. Furthermore, we use superscript ‘‘e’’ to denote the local stiffness matrix and variables in an element. Thus, for $K \in \mathcal{K}^h$, we write

$$(A.1) \quad A_{kl}^e = \int_{K'} a_{ij} \phi_{,i}^k \phi_{,j}^l d\mathbf{x}', \quad f_k^e = h^2 \int_{K'} f \phi^k d\mathbf{x}' \quad (k, l = 1, \dots, d).$$

Substituting the expansions of ϕ^k and ϕ^l into (A.1) and noting that

$$\chi_{,j}^i = \frac{1}{\epsilon'} \frac{\partial \chi^i}{\partial y_j} = \frac{1}{\epsilon'} \chi_{,y_j}^i,$$

we have the zeroth order terms as below:

$$\bar{A}_{kl}^e = \int_{K'} a_{ij} (\phi_{0,i}^k \phi_{0,j}^l - \chi_{,y_i}^p \phi_{0,p}^k \phi_{0,j}^l - \chi_{,y_j}^q \phi_{0,q}^l \phi_{0,i}^k + \chi_{,y_i}^p \phi_{0,p}^k \chi_{,y_j}^q \phi_{0,q}^l) d\mathbf{x}'.$$

Let $\sigma_{ij} = a_{ik} - a_{ik} \chi_{,y_k}^j$. Then we have $a_{ij}^* = \langle \sigma_{ij} \rangle$, where $\langle \cdot \rangle$ denotes average over Y . We may recast the above expression as

$$(A.2) \quad \bar{A}_{kl}^e = \int_{K'} (\sigma_{ij} \phi_{0,i}^k \phi_{0,j}^l - \sigma_{pj} \chi_{,y_p}^i \phi_{0,i}^k \phi_{0,j}^l) d\mathbf{x}'.$$

Since $\langle \sigma_{ij} - a_{ij}^* \rangle = 0$, it can be shown [14] that

$$\int_{K'} (\sigma_{ij} - a_{ij}^*) \phi_{0,i}^k \phi_{0,j}^l d\mathbf{x}' \leq C \epsilon',$$

where C is independent of ϵ (and h). Moreover, (3.8) implies that

$$(A.3) \quad \sigma_{ij, y_i} = 0,$$

from which integration by parts yields

$$\langle \sigma_{pj} \chi_{,y_p}^i \rangle = \int_Y \sigma_{pj} \chi_{,y_p}^i d\mathbf{y} = - \int_Y \sigma_{pj, y_p} \chi^i d\mathbf{y} = 0.$$

Therefore, we have

$$\int_{K'} \sigma_{pj} \chi_{,y_p}^i \phi_{0,i}^k \phi_{0,j}^l d\mathbf{x}' \leq C \epsilon',$$

and hence

$$\begin{aligned} \bar{A}_{kl}^e &= \int_{K'} a_{ij}^* \phi_{0,i}^k \phi_{0,j}^l d\mathbf{x}' + \int_{K'} (\sigma_{ij} - a_{ij}^*) \phi_{0,i}^k \phi_{0,j}^l d\mathbf{x}' - \int_{K'} \sigma_{pj} \chi_{,y_p}^i \phi_{0,i}^k \phi_{0,j}^l d\mathbf{x}' \\ &= A_0^e + O(\epsilon'), \end{aligned}$$

which implies that $\bar{A}^h = A_0^h + O(\epsilon/h)$.

Now, we collect $O(\epsilon')$ terms in the expansion, and get

$$(A.4) \quad \begin{aligned} A_{1kl}^e &= - \int_{K'} \sigma_{ij} [\phi_{0,j}^k (\chi^q \phi_{0,iq}^l + \theta_{,i}^l) + \phi_{0,j}^l (\chi^p \phi_{0,ip}^k + \theta_{,i}^k)] d\mathbf{x}' \\ &\quad + \int_{K'} \epsilon' a_{ij} \theta_{,i}^k \theta_{,j}^l d\mathbf{x}' + \frac{1}{\epsilon'} \int_{K'} (\sigma_{ij} - a_{ij}^* - \sigma_{pj} \chi_{,y_p}^i) \phi_{0,i}^k \phi_{0,j}^l d\mathbf{x}'. \end{aligned}$$

Some explanations need to be made for the terms containing $\theta_{,i}^k$ (or $\theta_{,j}^l$). We note that $L'_\epsilon \theta^k = 0$ and θ^k is oscillatory on ∂K (see (5.10)); hence the derivatives of θ^k can be large. However, from the earlier analysis we know that $|\theta^k|_{1,K'} = O(1/\sqrt{\epsilon'})$. Thus by the Cauchy-Schwarz inequality, the second integral on the right hand side of (A.4) is $O(1)$. Moreover, note that θ^l is bounded by the maximum principle. Using (A.3) and integration by parts, we have

$$(A.5) \quad \int_{K'} \sigma_{ij} \phi_{0,j}^k \theta_{,i}^l d\mathbf{x}' = \int_{\partial K'} n_i \sigma_{ij} \phi_{0,j}^k \theta^l ds' - \int_{K'} \sigma_{ij} \phi_{0,ij}^k \theta^l d\mathbf{x}' = O(1),$$

where n_i is the i th component of the unit outward normal vector on $\partial K'$. Obviously, the estimate holds for $\int_{K'} \sigma_{ij} \phi_{0,j}^l \theta_{,i}^k d\mathbf{x}'$ and for the other terms in (A.4). Thus, we obtain

$$A^h = A_0^h + \frac{\epsilon}{h} A_1^h + O\left(\frac{\epsilon^2}{h^2}\right).$$

The expansion can be carried to higher order terms, but for our purpose, it suffices to stop at the first order.

Using the expansion of ϕ^i , we can express f^h as follows:

$$f^h = f_0^h + \epsilon' f_1^h + \dots.$$

In each $K \in \mathcal{K}^h$, we have

$$(A.6) \quad f_{0i}^e = h^2 \int_{K'} f \phi_0^i d\mathbf{x}' \quad \text{and} \quad f_{1i}^e = -h^2 \int_{K'} f (\chi^p \phi_{0,p}^i + \theta^i) d\mathbf{x}'.$$

It follows that U^h of (6.1) can be expanded as

$$U^h = U_0^h + \epsilon' U_1^h + \dots,$$

where $A_0^h U_0^h = f_0^h$ and

$$(A.7) \quad A_0^h U_1^h = f_1^h - A_1^h U_0^h.$$

We will show that $A_1^h U_0^h$ can be written in a conservative form. By (5.1), one has

$$\sum_{k=1}^d \phi_{,k}^i \equiv 0, \quad \sum_{k=1}^d \phi_{,ij}^k \equiv 0.$$

Moreover, from

$$\sum_{k=1}^d \phi_1^k = \sum_{k=1}^d \sum_{p=1}^n \chi^p \phi_{0,p}^k = \sum_{p=1}^n \chi^p \sum_{k=1}^d \phi_{0,p}^k \equiv 0$$

and (5.5) we get

$$\sum_{k=1}^d \theta^k \equiv 0 \quad \text{and} \quad \sum_{k=1}^d \theta_{,i}^k \equiv 0.$$

From the above identities and (A.4), we have for all $K \in \mathcal{K}^h$

$$(A.8) \quad \sum_{j=1}^d A_{1ij}^e U_{0j}^h = \sum_{j=1, j \neq i}^d A_{1ij}^e (U_{0j}^h - U_{0i}^h) \quad (i = 1, \dots, d).$$

This equation and the symmetry of A_1^e then imply that $A_1^h U_0^h$ can be written in a conservative form. More precisely, assuming that the rectangular mesh consists of $N \times N$ elements (hence $h = 1/N$) and that the vertical and horizontal grid lines are labeled by i and j ($i, j = 0, \dots, N$), we may write $A_1^h U_0^h$ in the stencil format centered at node (i, j) as

$$(A.9) \quad (A_1^h U_0^h)_{ij} = \sum_{s=1}^4 (D_s^+ B_{ij}^s D_s^-) U_{0ij}^h.$$

Here, the B_{ij}^s ($s = 1, \dots, 4$) are assembled from the A_{1kl}^e ($k \neq l$); k and l are local indices in an element. The difference operators D_s^+ and D_s^- are defined as follows: for any v defined on the nodal points,

$$(A.10) \quad \begin{aligned} D_1^+ v_{ij} &= v_{i+1j} - v_{ij}, & D_1^- v_{ij} &= v_{ij} - v_{i-1j}; \\ D_2^+ v_{ij} &= v_{ij+1} - v_{ij}, & D_2^- v_{ij} &= v_{ij} - v_{ij-1}; \\ D_3^+ v_{ij} &= v_{i+1j+1} - v_{ij}, & D_3^- v_{ij} &= v_{ij} - v_{i-1j-1}; \\ D_4^+ v_{ij} &= v_{i+1j-1} - v_{ij}, & D_4^- v_{ij} &= v_{ij} - v_{i-1j+1}. \end{aligned}$$

Equation (A.9) also holds for the triangulation of the rectangular mesh (see e.g., Figure 1), except that there are fewer terms. The above calculation can also be applied to more general triangulations with more complexity.

In the following, we outline the method of finding the difference forms in B^s and f_1^h without going into the complicated details of algebra. We would like to point out that the triangular elements have some advantages over the rectangular ones in studying such difference forms. The main difference between the two types of elements is that the ϕ_0^i are *always* linear for the former, while they are in general unknown functions for the latter. Thus, the analysis of the triangular elements is much simpler. In addition, there are cases in which B^s can be written in difference forms on triangular elements but not on the rectangular ones with bilinear ϕ_0^i . Thus, in the following, we consider the triangulation as depicted in Figure 1.

First, we consider B^s assembled from A_1^e as described above. Since the ϕ_0^k are linear, from (A.4) we have

$$(A.11) \quad \begin{aligned} A_{1kl}^e &= - \int_{K'} \sigma_{ij} (\phi_{0,j}^k \theta_{,i}^l + \phi_{0,j}^l \theta_{,i}^k) d\mathbf{x}' + \int_{K'} \epsilon' a_{ij} \theta_{,i}^k \theta_{,j}^l d\mathbf{x}' \\ &\quad + \frac{1}{\epsilon'} \int_{K'} \tilde{\sigma}_{ij} \phi_{0,i}^k \phi_{0,j}^l d\mathbf{x}', \end{aligned}$$

where $\tilde{\sigma}_{ij} = \sigma_{ij} - a_{ij}^* - \sigma_{pj} \chi_{,y_p}^i$ and $\langle \tilde{\sigma}_{ij} \rangle = 0$. Using the Fourier expansion, it can be shown that there exists a periodic third rank tensor κ_{mij} such that $\langle \kappa_{mij} \rangle = 0$ and $\tilde{\sigma}_{ij} = \kappa_{mij} y_m$. Then from (A.11), (A.5), and the fact that $\phi_{0,ij}^k = 0$, we obtain,

using integration by parts,

$$(A.12) \quad \begin{aligned} A_{1kl}^e = & - \int_{\partial K'} n_i \sigma_{ij} (\phi_{0,j}^k \theta^l + \phi_{0,j}^l \theta^k) ds' + \epsilon' \int_{\partial K'} \theta^k n_i a_{ij} \theta_{,j}^l ds' \\ & + \int_{\partial K'} n_m \kappa_{mij} \phi_{0,i}^k \phi_{0,j}^l ds' + O(\epsilon). \end{aligned}$$

Therefore, to the leading order B^s can be expressed by the sum of boundary integrals in (A.12) from relevant elements. The integrands may be further simplified by using the fact that the ϕ_0^k are linear base functions. Note that for two adjacent elements, the outward normal vectors for the two elements point in the opposite directions at the interface of the elements. This implies that the values of the boundary integrals along that interface have opposite signs if the (simplified) integrands are continuous at the interface. From this observation and (A.12), one can verify that for the triangulation in Figure 1 the contributions to B^s from the first and third integrals in (A.12) can be written in difference forms. The derivation is rather tedious and is omitted here. However, we also find that the second term in (A.12) prevents B^s from having a difference structure. In general, B_{ij}^s has the following expression:

$$B_{ij}^s = D_1^- C_{ij}^s + D_2^- D_{ij}^s + B_\theta^s$$

where C_{ij}^s and D_{ij}^s are uniquely defined functions for i and j , and B_θ^s is due to $\int_{\partial K'} n_i a_{ij} \theta^k \theta_{,j}^l$. When θ has boundary layer structures, B_θ^s in general cannot be written in difference form. However, as we pointed out in §6.2, if we can eliminate the boundary layer in θ by choosing a proper boundary condition for ϕ_i , then B_θ^s would only contribute to the next order terms. Therefore, B^s admits a conservative form.

For f_{1ij}^h , we need to consider the sum of f_1^e (see (A.6)) in the six elements centered with the common node (i, j) (Figure 1). It is easy to verify that the sum of $\int_{K'} f \chi^p \phi_{0,p}^i dx'$ over the six elements can be written in difference forms in the i and j directions. Moreover, from the above discussion of the leading order solution of θ^k , it can be shown that the sum of $\int_{K'} f \theta^k dx'$ consists of terms in difference forms and some $O(h)$ terms.

ACKNOWLEDGMENTS

We are grateful to Professor Bjorn Engquist, Dr. Michael Holst, and Mr. Yalchin Efendiev for many interesting and helpful discussions. We also thank the referees for their comments and suggestions, which have lead to a significant improvement of this paper.

REFERENCES

- [1] M. Avellaneda, T. Y. Hou, and G. Papanicolaou, *Finite difference approximations for partial differential equations with rapidly oscillating coefficients*, Mathematical Modelling and Numerical Analysis **25** (1991), 693–710. MR **92j**:65152
- [2] I. Babuška, G. Caloz, and E. Osborn, *Special finite element methods for a class of second order elliptic problems with rough coefficients*, SIAM J. Numer. Anal. **31** (1994), 945–981. MR **95g**:65146
- [3] I. Babuška and E. Osborn, *Generalized finite element methods: Their performance and their relation to mixed methods*, SIAM J. Numer. Anal. **20** (1983), 510–536. MR **84h**:65076

- [4] ———, *Finite element methods for the solution of problems with rough data*, Singularities and Constructive Methods and Their Treatment (P. Grisvard, W. Wendland, and J. R. Whiteman, eds.), Springer-Verlag Lecture Notes in Mathematics 1121, Berlin-New York, 1985, pp. 1–18. MR **86m**:65138
- [5] I. Babuška and W. G. Szymczak, *An error analysis for the finite element method applied to convection-diffusion problems*, Comput. Methods Appl. Math. Engrg. **31** (1982), 19–42. MR **83g**:76012
- [6] J. Bear, *Use of models in decision making*, Transport and Reactive Processes in Aquifers (T. H. Dracos and F. Stauffer, eds.), Balkema, Rotterdam, 1994, pp. 3–9.
- [7] A. Bensoussan, J. L. Lion, and G. Papanicolaou, *Boundary layer analysis in homogenization of diffusion equations with dirichlet conditions in the half space*, Proceedings of the International Symposium on Stochastic Differential Equations (Kyoto) (K. Ito, ed.), Wiley, 1976, pp. 21–40. MR **80g**:35127
- [8] ———, *Asymptotic analysis for periodic structure*, Studies in Mathematics and Its Applications, vol. 5, North-Holland Publ., 1978. MR **82h**:35001
- [9] M. E. Cruz and A. Petera, *A parallel Monte-Carlo finite-element procedure for the analysis of multicomponent random media*, Int. J. Numer. Methods Eng. **38** (1995), 1087–1121.
- [10] L. J. Durlofsky, *Numerical-calculation of equivalent grid block permeability tensors for heterogeneous porous media*, Water Resour. Res. **27** (1991), 699–708.
- [11] B. B. Dykaar and P. K. Kitanidis, *Determination of the effective hydraulic conductivity for heterogeneous porous media using a numerical spectral approach: 1. Method*, Water Resour. Res. **28** (1992), 1155–1166.
- [12] W. E and T. Y. Hou, *Homogenization and convergence of the vortex method for 2-d euler equations with oscillatory vorticity fields*, Comm. Pure and Appl. Math. **43** (1990), 821–855. MR **91h**:35263
- [13] Y. R. Efendiev, Ph.D. thesis, Caltech, 1998.
- [14] B. Engquist and T. Y. Hou, *Particle method approximation of oscillatory solutions to hyperbolic differential equations*, SIAM J. Numer. Anal. **26** (1989), 289–319. MR **90f**:65160
- [15] J. Frehse and R. Rannacher, *Eine l^1 -fehlerabschätzung für diskrete Grundlösungen in der Methode der finiten Elemente*, Finite Elemente (J. Frehse, ed.), Bonn. Math. Schrift., no. 89, 1975, pp. 92–114. MR **57**:11104
- [16] T. Y. Hou and X. H. Wu, *A multiscale finite element method for elliptic problems in composite materials and porous media*, J. Comput. Phys. **134** (1997), 169–189. MR **98e**:73132
- [17] S. M. Kozlov, *Averaging differential operators with almost periodic, rapidly oscillating coefficients*, Math. USSR Sbornik **35** (1978), 481–498. MR **81m**:35017
- [18] O. A. Ladyzhenskaya and N. N. Ural'tseva, *Linear and quasilinear elliptic equations*, Academic Press, New York, 1968. MR **39**:5941
- [19] J. L. Lions and E. Magenes, *Problèmes aux limites non homogènes et applications*, vol. I, Paris, Dunod, 1968, English translation, Springer-Verlag, 1972. MR **40**:512; MR **50**:2670
- [20] J. F. Mccarthy, *Comparison of fast algorithms for estimating large-scale permeabilities of heterogeneous media*, Transport in Porous Media **19** (1995), 123–137.
- [21] S. Moskow and M. Vogelius, *First order corrections to the homogenized eigenvalues of a periodic composite medium. A convergence proof*, Proc. Roy. Soc. Edinburgh Sect. A. **127** (1997), 1263–1299. CMP 98:06
- [22] P. M. De Zeeuw, *Matrix-dependent prologation and restrictions in a blackbox multigrid solver*, J. Comput. Applied Math. **33** (1990), 1–27. MR **92c**:65152

APPLIED MATHEMATICS, 217-50, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CA 91125

E-mail address: hou@ama.caltech.edu

APPLIED MATHEMATICS, 217-50, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CA 91125

Current address: Exxon Production Research Company, P. O. Box 2189, Houston, TX 77252

E-mail address: xwu@ama.caltech.edu

DEPARTMENT OF MATHEMATICS, PURDUE UNIVERSITY, WEST LAFAYETTE, IN 47907-1395

E-mail address: zcai@math.purdue.edu