SIAM J. NUMER. ANAL.
Vol. 37, No. 3, pp. 888–910
© 2000 Society for Industrial and Applied Mathematics

CONVERGENCE OF A NONCONFORMING MULTISCALE FINITE ELEMENT METHOD*

YALCHIN R. EFENDIEV†, THOMAS Y. HOU†, AND XIAO-HUI WU‡

Abstract. The multiscale finite element method (MsFEM) [T. Y. Hou, X. H. Wu, and Z. Cai, *Math. Comp.*, 1998, to appear; T. Y. Hou and X. H. Wu, *J. Comput. Phys.*, 134 (1997), pp. 169–189] has been introduced to capture the large scale solutions of elliptic equations with highly oscillatory coefficients. This is accomplished by constructing the multiscale base functions from the local solutions of the elliptic operator. Our previous study reveals that the leading order error in this approach is caused by the "resonant sampling," which leads to large error when the mesh size is close to the small scale of the continuous problem. Similar difficulty also arises in numerical upscaling methods. An oversampling technique has been introduced to alleviate this difficulty [T. Y. Hou and X. H. Wu, *J. Comput. Phys.*, 134 (1997), pp. 169–189]. A consequence of the oversampling method is that the resulting finite element method is no longer conforming. Here we give a detailed analysis of the nonconforming error. Our analysis also reveals a new cell resonance error which is caused by the mismatch between the mesh size and the wavelength of the small scale. We show that the cell resonance error is of lower order. Our numerical experiments demonstrate that the cell resonance error is generically small and is difficult to observe in practice.

Key words. multiscale, finite element, homogenization, resonant sampling

AMS subject classifications. 65F10, 65F30

PII. S0036142997330329

1. Introduction.

Multiscale problems in science and engineering are often described by partial differential equations (PDEs) with highly oscillatory coefficients. Typical examples include flows in porous media and turbulent transport problems. Solving these problems numerically is difficult because an accurate solution usually requires a very fine resolution and hence tremendous amount of computer memory and CPU time. Parallel computing relieves the difficulty to some degree, but the size of computation is not reduced in the traditional approaches which directly solve the equations on fine meshes.

Recently, a multiscale finite element method (MsFEM) has been developed [10, 8] for capturing the large scale solutions of multiscale problems on a coarse mesh (with mesh size larger than a certain cut-off scale of the problem). The main idea of the method is to build the local small scale information of the leading order differential operator into the finite element base functions. It is through these multiscale bases and the finite element formulation that the effect of small scales on the large scales are correctly captured. A key feature of MsFEM is that the construction of the base functions is a *local* operation within the elements. Thus, the construction in one element is decoupled from that in another element. In other words, a large scale computation is broken into many small and independent pieces. This results in many computational advantages [8], such as saving in computer memory and good parallel

*Received by the editors November 19, 1997; accepted for publication (in revised form) May 4, 1999; published electronically February 24, 2000. This work is supported in part by ONR under grant N00014-94-0310, by DOE under grant DE-FG03-89ER25073, and by NSF under grant DMS-9704976.

http://www.siam.org/journals/sinum/37-3/33032.html

†Applied Mathematics, Caltech, Pasadena, CA 91125 (yalchin@ama.caltech.edu, hou@ama.caltech.edu).

‡Applied Mathematics, Caltech, Pasadena, CA 91125. Current address: Exxon Production Research Company, P.O. Box 2189, Houston, TX 77252 (x9wu1@epr.exxon.com).

efficiency. We remark that special base functions in finite element methods have been used by several authors in capturing multiscale solutions of PDEs. In particular, the works presented in [16, 2, 5, 11] are most relevant to our previous conforming MsFEM [10].

The basic convergence property of the method has been established for a two-scale elliptic problem with periodic coefficients [10]. It is shown that the numerical solution converges to the homogenized solution in the limit of $\epsilon \to 0$ (ϵ is the small scale in the solution). The homogenization theory is used in the proof; however, it is neither required by the MsFEM formulation nor used in the computations. Our numerical experiments demonstrate that MsFEM, together with an oversampling method, is well applicable to general elliptic problems with many or continuous scales [9]. It is shown that numerical solutions computed using MsFEM on a coarse grid give accuracy comparable to that of well-resolved solutions computed using conventional methods on a fine grid. The application of MsFEM to practical problems such as two-phase flows in porous media and other types of equations is currently under study. It is worth mentioning that MsFEM also gives convergent solutions when $h \ll \epsilon$, just like the conventional methods.

Our analysis also reveals a resonance error between the grid scale and the scales of the continuous problem [10]. This is a common difficulty in numerical upscaling methods. For the two-scale problem, the error due to the resonance manifests as a ratio between the wavelength of the small scale oscillation and the grid size; the error becomes large when the two scales are close. The resonance represents a *fundamental* difficulty due to the mismatch between the local construction of the multiscale base functions and the global nature of the elliptic problems. This mismatch between the local solution and the global solution produces a boundary layer in the first order corrector of the local solution. Motivated by our analysis we propose an *oversampling* technique to overcome the difficulty due to scale resonance in [9]. The idea is quite simple and easy to implement. Since the boundary layer in the first order corrector is thin we can sample in a domain with size larger than h and use only the interior sampled information to construct the bases (h is the mesh size). By doing this, we greatly reduce the influence of the boundary layer in the larger domain on the base functions. In fact, we do not observe the resonance error due to the boundary layer in our computations.

Unfortunately, the oversampling technique results in a *nonconforming* MsFEM method. The analysis in [10] needs to be modified to take into account the nonconforming error. In this paper, we perform a careful estimate of the nonconforming errors in both the H^1 norm and the L^2 norm. Our analysis shows that the resonance error is not completely removed although its effect seems to be generically small from our extensive numerical experiments. Our analysis also reveals another source of resonance, which is the mismatch between the mesh size and the "perfect" sample size. In case of a periodic structure, the "perfect" sample size is the length of an integer multiple of the period. We call the new resonance the "cell resonance." In the error expansion, this resonance effect appears as a *higher* order correction. Similar resonance effect was observed also in [17]. Although the oversampling helps reduce the leading order resonance error, we find that oversampling alone does not remove the cell resonance error, which dominates the nonconforming error. In a subsequent paper, we will completely eliminate this cell resonance error by using the oversampling technique to construct the base functions but using piecewise linear functions as test functions [19]. This will reduce the nonconforming error and eliminate the resonance error completely.

We remark that from our computational experience, the cell resonance errors seem to be generically small and are rarely observed in computations. This may be due to some subtle error cancellation in the convolution with the discrete Green's function (i.e., the inverse of the stiffness matrix). The discrete Green's function can be highly oscillatory depending on the ratio h/ϵ. In fact, in our tests with the worst resonance case, $h/\epsilon = 1.5$, the solution appears to converge. Only when $h = 1/1024$ does the effect of cell resonance become strong enough to stop the convergence. For more general problems, such as problems with random coefficients, we did not find resonance errors through numerical tests (see [9]). These results strongly indicate that the chance of having significant cell resonance in practical computations is small.

The rest of the paper is organized as follows. The formulations of the two-dimensional (2-D) model problem and MsFEM along with the oversampling method are presented in section 2. In section 3 we give H^1 estimates of the error for the cases of $h \gg \epsilon$. In section 4, we present an L_2 analysis for the case $h \gg \epsilon$ using the discrete error analysis. This approach is unconventional but is quite effective in revealing the source of resonance error. It enables us to obtain tight estimates that cannot be reached through the use of the Nitsche trick. Section 5 contains some further discussion on the cell resonance error. Based on this analysis, we propose an averaging method to remove the cell resonance for the special case of periodic oscillation, i.e., $a_\epsilon(x) = a(x/\epsilon)$.

2. Model problem, MsFEM, and oversampling. In this section we introduce the model problem and review the MsFEM. First we state some notations and conventions. Throughout the paper, the Einstein summation convention is used: summation is taken over repeated indices. In what follows, we use the $L_2(\Omega)$ based Sobolev spaces $H^k(\Omega)$ equipped with norms and seminorms:

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

$H_0^1(\Omega)$ consists of those functions in $H^1(\Omega)$ that vanish on $\partial\Omega$. We define $H^{1/2}(\Omega)$ as the trace on $\partial\Omega$ of all functions in $H^1(\Omega)$ with the norm $\|v\|_{1/2,\partial\Omega} = \inf \|u\|_{1,\Omega}$, where the infimum is taken over all $u \in H^1(\Omega)$ with the trace v. Throughout, C (with or without subscripts) denotes a generic constant, which is independent of ϵ and h (mesh size), unless otherwise stated and $C + C = C$, $C \cdot C = C$.

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

$$L_\epsilon u_\epsilon = f, \qquad u_\epsilon = 0 \quad \text{on} \quad \partial\Omega, \tag{2.1}$$

where $L_\epsilon = -\nabla_i(a_\epsilon^{ij}\nabla_j)$ is the linear elliptic operator, ϵ is a small parameter, $f \in L_2(\Omega)$ is independent of ϵ, and a_ϵ^{ij} is symmetric and satisfies $\alpha|\xi|^2 \le \xi_i a_{ij}^\epsilon \xi_j \le \beta|\xi|^2$ for all $\xi \in \mathbf{R}^2$ and with $0 < \alpha < \beta < \infty$. Furthermore, we assume that $a_\epsilon^{ij}(x)$ has the form

$$a_\epsilon^{ij} = a^{ij}(x/\epsilon)$$

and $a^{ij}(y)$ is a sufficiently smooth periodic function in y in a unit cube Y. For our analysis it is sufficient to assume that $a^{ij}(y) \in C^1(Y)$. Throughout the paper,

we assume $\Omega = (0,1) \times (0,1) \subset \mathbf{R}^2$. The variational problem of (2.1) is to seek $u_\epsilon \in H_0^1(\Omega)$ such that

$$a(u_\epsilon, v) = f(v), \tag{2.2}$$

where

$$a(u,v) = \int_\Omega a_\epsilon^{ij} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} dx \tag{2.3}$$

and

$$f(v) = \int_\Omega f v dx. \tag{2.4}$$

It is easy to see that the linear form $a(\cdot,\cdot)$ is elliptic and continuous. 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 $\mathbf{K}^h$ be a partition of Ω of triangles K with diameter less than h. In each element $K \in \mathbf{K}^h$, we define a set of nodal basis $\{\phi^i_{\epsilon,K}\}$, $i = 1, \ldots, d$ with $d(= 3)$ 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_ϵ satisfies

$$L_\epsilon \phi^i_\epsilon = 0 \quad \text{in} \quad K \in \mathbf{K}^h. \tag{2.5}$$

Let $x_j \in K$ $(j = 1, \ldots, d)$ be the nodal points of K. As usual we require $\phi^i_\epsilon(x_j) = \delta_{ij}$. One needs to specify the boundary condition of ϕ^i_ϵ for the well-posedness of (2.5). We will assume that the base functions are linear on the boundaries (unless otherwise stated). We have

$$V^h = \text{span}\{\phi^i_{\epsilon,K}; \quad i = 1, \ldots, d, \quad K \subset \mathbf{K}^h\} \subset H_0^1(\Omega)$$

and multiscale finite element consists of finding the solution of

$$a(u^h_\epsilon, v) = f(v) \quad \text{for all} \ \ v \in V^h. \tag{2.6}$$

2.2. Multiple scale expansion. The homogenization theory has been used in the analysis of the multiscale finite element method in [10]. The theory provides the detailed structures of the physical solution and the multiscale base function ϕ_ϵ. This information is crucial for obtaining precise error estimates. Here, we give a brief review of the multiple scale expansion of the solution and the base functions. It is known [3, 12, 1] that the solution of (2.1) can be expanded as

$$u_\epsilon = u_0 + \epsilon \chi^i \left(\frac{x}{\epsilon}\right) \nabla_i u_0 + \epsilon \theta^u. \tag{2.7}$$

In the expansion, $u_0 \in H^2(\Omega)$ is the solution of the homogenized equation

$$a^{ij}_* \nabla_i \nabla_j u_0 = f \quad \text{in} \quad \Omega, \tag{2.8}$$

satisfying $u_0 = 0$ on $\partial\Omega$. The constant homogenized coefficients a^{ij}_* are given by

$$a^{ik}_* = \frac{1}{|Y|} \int_Y a^{ij}(\delta_{jk} + \nabla^y_j \chi^k) dy, \tag{2.9}$$

where ∇^y is the gradient with respect to the "fast" variable $y = x/\epsilon$, and χ^k is the periodic solution in the unit cell Y of

$$\nabla_i^y a^{ij} \nabla_j^y \chi^k = -\nabla_i^y a^{ik}. \tag{2.10}$$

For θ^u we have

$$\begin{aligned} L_\epsilon \theta^u &= \frac{1}{\epsilon} L_\epsilon (u_\epsilon - u_0 - \epsilon \chi^i \nabla_i u_0) \quad \text{in} \quad \Omega, \\ \theta^u &= -\chi^i \nabla_i u_0 \quad \text{on} \quad \partial\Omega. \end{aligned}$$

Similarly, from (2.5) we have the expansion for the base functions:

$$\phi_\epsilon = \phi_0 + \epsilon \chi^i \left(\frac{x}{\epsilon}\right) \nabla_i \phi_0 + \epsilon \theta^\epsilon, \tag{2.11}$$

where ϕ_0, χ^i, and θ^ϵ are defined in the same way as above. The equation for θ^ϵ can be further simplified by taking into account the fact that ϕ_0 is linear along the boundaries of the triangular elements K. It follows that ϕ_0 is linear in K. Thus, by (2.5) and the equations for θ^ϵ and χ^k, we have

$$\begin{aligned} L_\epsilon \theta^\epsilon &= \frac{1}{\epsilon} L_\epsilon (\phi_\epsilon - \phi_0 - \epsilon \chi^k \nabla_k \phi_0) \\ &= -\frac{1}{\epsilon} \nabla_i a_\epsilon^{ij} \nabla_j (\phi_0 + \epsilon \chi^k \nabla_k \phi_0) \\ &= -\left(\nabla_i a_\epsilon^{ij} \nabla_j \chi^k + \frac{1}{\epsilon} \nabla_i a_\epsilon^{ik} \right) \nabla_k \phi_0 - \nabla_i \chi^k a_\epsilon^{ij} \nabla_j \nabla_k \phi_0 \\ &= 0. \end{aligned}$$

The expansions (2.7) and (2.11) form the basis for the convergence analysis in this paper. It enables us to examine the subtle details of the solution structures in both the continuous norms and the discrete norm. The expansion may not be explicitly available for more general problems; however, it has been the key to understanding the essential features of the multiscale method which do appear in the computations of more general problems.

2.3. The oversampling method. As mentioned in the introduction, the above MsFEM suffers from the resonance error. By using (2.11) we find that the boundary layer structure in the first order corrector, θ^ϵ, is a main cause of the resonance error (see [8] and also section 4). Since the boundary layer is thin we propose an oversampling method to reduce the resonance error. Specifically, we first construct base functions ψ_ϵ^j in a sampling domain $S \supset K$ by solving

$$\nabla_i a_{ik}^\epsilon \nabla_k \psi_\epsilon^j = 0 \quad \text{in} \quad S,$$

where ψ_ϵ^j is piecewise linear along ∂S and $\psi_\epsilon^j = \delta_{ij}$ at the ith nodal points of S. For simplicity, we assume S to be triangular and hence having the same number of nodal points as K. Moreover, we choose S sufficiently large such that $\text{diam}(S) = h_1 > h$ and ∂S is away from ∂K at a distance of order ϵ. Next, the base functions ϕ_ϵ on K are constructed from the linear superposition of ψ_ϵ:

$$\phi_{\epsilon,i}^k = \sum_{j=1}^d c_{ij}^k \psi_\epsilon^j, \tag{2.12}$$

where i is the index of the nodal point and constants c_{ij}^k are determined by the condition $\phi_{0,i}^k(x_j) = \sum_{j=1}^d c_{il}^k \psi_0^l(x_j) = \delta_{kj}$, x_j being the nodal points of K and ψ_0^j being the linear homogenized parts of ψ_ϵ^j (see below). By this procedure, the boundary layer structure near ∂S can be reduced. To make this more precise, we expand ψ_ϵ as

$$\psi_\epsilon = \psi_0 + \epsilon \chi^p \nabla_p \psi_0 + \epsilon \theta',$$

where ψ_0 is the homogenized part of ψ_ϵ, which is linear. For the corrector θ', we have $\theta' = \eta^p \nabla_p \psi_0$ with η^p being the solution of

$$\nabla_i a_\epsilon^{ij} \nabla_j \eta^p = 0 \quad \text{in} \quad S \quad \text{and} \quad \eta^p = -\chi^p \quad \text{on} \quad \partial S. \tag{2.13}$$

The analysis of the behavior of η^p is a complicated problem, which is yet to be rigorously carried out. In [4] the problem (2.13) in the half space $S = \mathbf{R}r_+$, whose boundary aligned with the period, was analyzed . Besides the fact that $\nabla \eta^p$ exponentially decays away from ∂S the authors found out that there exists a constant χ_*^p such that η^p, the solution of (2.13) with boundary condition $\eta^p = \chi^p - \chi_*^p$ on ∂S, decays exponentially too. Moskow and Vogelius [15] investigated some special features of the behavior of η^p for polygon domains with specially oriented sides. It was shown that η^p has exponential boundary layers away from the corners. We note that these boundary layers are the cause of large H^1 norm of η^p and the resonance in MsFEM [10]. Using the oversampling technique we reduce the effects of these boundary layers from the base functions. The behavior of η^p in the interior of S has not been rigorously carried out yet. But it was numerically demonstrated [8] that $\nabla \eta^p$ has small oscillations in the interior of S. Recently Prof. Z. Chen was able to prove (see Appendix C) the following fact.

LEMMA 2.1 (by Z. Chen). *Assume that $K \subset S$ is away from ∂S at least at a distance h. Then we have*

$$\|\nabla \eta^p\|_{L_\infty(K)} \leq \frac{C}{h}. \tag{2.14}$$

We would like to note that the estimate (2.14) cannot be improved in general (see [15]). Although the estimate (2.14) is sharp, we have observed from numerical experiments (later in the paper and in [9]) that with the oversampling technique the computational error caused by the terms containing η^p is usually negligible. An important consequence of the above construction is that the base functions ϕ are no longer continuous across the internal boundaries of the elements, nor are they zero on the external boundaries. Setting the base functions to be zero outside the external boundaries of the elements introduces first order discontinuities of ϕ along all sides of the elements. Thus, these base functions are nonconforming and V_h spanned by ϕ is no longer in H^1. This complicates the analysis of MsFEM.

In the next section, we estimate the nonconforming error. We remark that despite the complications, the nonconforming finite element methods have been widely used in practice when conforming elements that satisfy certain physical and numerical properties are too tedious or impossible to construct [6, 18]. In our computations, we usually choose a large sample domain, S, to contain many elements to improve the efficiency of computation [8]. Two implementations of the oversampling method have been tested. In [8], we choose S to be nonoverlapped, which is easy to implement. This reduces the boundary layers in the interior elements, but the elements adjacent to ∂S are still under the influence of the boundary layers. The accuracy of solutions

can be further improved by a fully implemented oversampling method [9], in which the sample domains have some overlaps (with width of $O(\epsilon)$). For the convenience and clarity in the derivations below, let us introduce some notations here. Let

$$\mathbf{K}_h(S) = \{K : K \in \mathbf{K}, K \subset S\}$$

be the set of all elements contained in the sample domain S and

$$\tilde{S} = \bigcup_{K \in \mathbf{K}_h(S)} K \subseteq S \tag{2.15}$$

be the union of these elements.

3. H^1 estimates. Because of the nonconformity of the base functions, we no longer enjoy Cea's lemma, but we have the following estimate between the exact solution and the numerical solution in the energy norm [18, 6]:

$$\|u_\epsilon - u_\epsilon^h\|_{h,\Omega} \leq C \left(\inf_{v_\epsilon^h \in \hat{H}^h} \|u_\epsilon - v_\epsilon^h\|_{h,\Omega} + \sup_{w_\epsilon^h \in \hat{H}^h} \frac{|f(w_\epsilon^h) - a(u_\epsilon, w_\epsilon^h)|}{\|w_\epsilon^h\|_h} \right), \tag{3.1}$$

where $\hat{H}^h$ is the finite dimensional space generated by the nonconforming basis functions (in general $\hat{H}^h \not\subset H^1$),

$$\|w_\epsilon^h\|_{h,\Omega} = \left(\sum_{K \in \mathbf{K}^h} \int_K |\nabla w_\epsilon^h|^2 dx \right)^{1/2}, \tag{3.2}$$

and

$$|f(w_\epsilon^h) - a(u_\epsilon, w_\epsilon^h)| = \left| \int_\Omega f w_\epsilon^h dx - \sum_{K \in \mathbf{K}^h} \int_K a_\epsilon^{ij} \nabla_i u_\epsilon \nabla_j w_\epsilon^h dx \right|. \tag{3.3}$$

3.1. Case of $h \gg \epsilon$.

THEOREM 3.1. *Let u_ϵ be the solution of* (2.1) *and u_ϵ^h be the numerical solution computed using MsFEM with oversampling. Assuming that u_0, the homogenized part of u_ϵ is in $W^{1,\infty}(\Omega)$, and the distance between K and ∂S is of order h, we have*

$$\|u_\epsilon - u_\epsilon^h\|_{h,\Omega} \leq C \frac{\epsilon}{h} + C_1 \sqrt{\epsilon} + C_2 h \qquad (h > \epsilon). \tag{3.4}$$

Remark 3.1. Using the inequality (B.1) it can be shown that (3.2) is a norm in $\hat{H}^h$.

We first consider the "conforming error" on the right-hand side (R.H.S.) of (3.1). Then, we analyze the nonconforming error given by the second term. For the analysis of nonconforming error we will need the following lemma.

LEMMA 3.2. *Let D be a unit box in $\mathbf{R}^2$ and $N(y) \in L^\infty(D)$ be a 1-periodic function and $\langle N \rangle = 0$. Then for all $f \in H^1(K) \bigcap L^\infty(K)$, $K \subset \mathbf{R}^2$, and* $\operatorname{diam}(K) = h$, *we have*

$$\left| \int_K f(x) N\left(\frac{x}{\epsilon}\right) dx \right| \leq C \epsilon h.$$

Proof. Define

$$f_i = \frac{1}{|Y_i|}\int_{Y_i} f dx,$$

where Y_i is a periodic cell of $N(x/\epsilon)$, $Y_i \subset K$. Then

$$\|f - f_i\|_{L_2(Y_i)} \le \epsilon \|\nabla f\|_{L_2(Y_i)}.$$

Denote $K^{'} = \cup_{Y_i \subset K} Y_i$. We have

$$\begin{aligned}\left|\int_K f(x) N\left(\frac{x}{\epsilon}\right) dx\right| &\le \left|\sum_{Y_i\subset K}\int_{Y_i}(f-f_i)N\left(\frac{x}{\epsilon}\right)dx\right| + \left|\sum_{Y_i\subset K}\int_{Y_i} f_i N\left(\frac{x}{\epsilon}\right)dx\right| \\ &\quad + \left|\int_{K\setminus K'} f(x)N\left(\frac{x}{\epsilon}\right)dx\right| \\ &\le \sum_{Y_i\subset K}\|f-f_i\|_{L_2(Y_i)}\|N\left(\frac{x}{\epsilon}\right)\|_{L_2(Y_i)} + \left|\int_{K\setminus K'} f(x)N\left(\frac{x}{\epsilon}\right)dx\right| \\ &\le C\epsilon\|\nabla f\|_{L_2(K')}\|N\|_{L_2(K')} + \left|\int_{K\setminus K'} f(x)N\left(\frac{x}{\epsilon}\right)dx\right| \\ &\le C\epsilon h.\end{aligned}$$

In the last two steps we have used the Cauchy–Schwarz inequality. □

Proof of Theorem 3.1. For the conforming part, we notice that in each element $K \in \mathbf{K}^h$, $v_\epsilon^h \in \hat{H}^h$ can be expanded as

$$v_\epsilon^h = v_0^h + \epsilon\chi^p \nabla_p v_0^h + \epsilon\theta^h, \tag{3.5}$$

where $v_0^h = v_k c_j^k \psi_0^j$ with $v_k = u_0(x_k)$ and $\theta^h = v_k c_j^k \theta_j^{'}$, ψ_ϵ^j are the base functions, $\theta_j^{'}$ are corresponding correctors in the sample domain $S \supset K$, and c_j^k are chosen in order to make $c_j^k \psi_\epsilon^j$ have the desirable values at the nodal points. So $\phi_0^k = c_j^k \psi_0^j$ are linear functions and $\phi_0^k(x_j) = \delta_{jk}$. Then by (2.7)

$$\|u_\epsilon - v_\epsilon^h\|_{h,K} \le \|u_0 - v_0^h\|_{h,K} + \|\epsilon\chi^p\nabla_p(u_0 - v_0^h)\|_{h,K} + \epsilon\|\theta^h\|_{h,K} + \epsilon\|\theta^u\|_{h,K}, \tag{3.6}$$

where $\|u\|_{h,K} = (\int_K (\nabla u)^2 dx)^{1/2}$. Taking into account that

$$\|u_0 - v_0^h\|_{h,K} \le Ch|u_0|_{2,K}, \tag{3.7}$$

and $\chi^p \in W^{1,\infty}(Y)$ (the latter follows from $a^{ij}(y) \in W^{1,p}$, $p > 2$; cf. Theorem 15.1 in [13]), we have

$$\|u_\epsilon - v_\epsilon^h\|_{h,K} \le Ch|u_0|_{2,K} + \epsilon\|\theta^h\|_{h,K} + \epsilon\|\theta^u\|_{h,K}, \tag{3.8}$$

where θ^h satisfies the following equation:

$$\nabla_i a_\epsilon^{ij} \nabla_j \theta^h = 0 \quad \text{in} \quad S \quad \text{and} \quad \theta^h = \chi^p \nabla_p v_0^h \quad \text{on} \quad \partial S.$$

Since v_0^h is linear in K, we have $\theta^h = \eta^p \nabla_p v_0^h$, where η^p is defined by (2.13) and $||\nabla\eta^p\|_{L_\infty(K)} \le C/h$ (see Lemma 2.1). Then using $\|\nabla v_0^h\|_{L_2(K)} \le C\|\nabla u_0\|_{L_2(K)}$ (see Appendix B) we get

$$\|\epsilon\theta^h\|_{h,K} \le C\frac{\epsilon}{h}\|\nabla v_0^h\|_{L_2(K)} \le C\frac{\epsilon}{h}\|\nabla u_0\|_{L_2(K)}.$$

Summing (3.8) over all $K \in \mathbf{K}^h$ we have

$$\|u_\epsilon - v_\epsilon^h\|_{h,\Omega} \leq Ch|u_0|_{2,\Omega} + C\frac{\epsilon}{h}\|\nabla u_0\|_{L_2(\Omega)} + C\sqrt{\epsilon} \leq C(h+\sqrt{\epsilon}) + C\frac{\epsilon}{h}.$$

Here we have used the fact that $|\epsilon\theta^u|_{1,\Omega} \leq C\sqrt{\epsilon}$ [12]. The nonconforming error in (3.1) can be written in the following way:

$$\begin{aligned} N.C. &\equiv \sup_{w_\epsilon^h \in \hat{H}^h} \frac{|f(w_\epsilon^h) - a(u_\epsilon, w_\epsilon^h)|}{\|w_\epsilon^h\|_h} = \sup_{w_\epsilon^h \in \hat{H}^h,\ \|w_\epsilon^h\|_h = 1} |f(w_\epsilon^h) - a(u_\epsilon, w_\epsilon^h)| \\ &\leq \sup_{w_\epsilon^h \in \hat{H}^h,\ \|w_\epsilon^h\|_h = 1} \{|f(w_\epsilon^h - l^h) - a(u_\epsilon, w_\epsilon^h - l^h)| + |f(l^h) - a(u_\epsilon, l^h)|\}, \end{aligned} \tag{3.9}$$

where l^h is the homogenized part of w_ϵ^h. Note that l^h is a linear function. Because of the conformity of the linear base functions, the second term on the R.H.S. of (3.9) is zero. Then (3.9) becomes

$$\begin{aligned} N.C. &\leq C \sup_{w_\epsilon^h \in \hat{H}^h,\ \|w_\epsilon^h\|_h = 1} \{|f(w_\epsilon^h - l^h) - a(u_\epsilon, w_\epsilon^h - l^h)|\} \\ &\leq C \sup_{w_\epsilon^h \in \hat{H}^h,\ \|w_\epsilon^h\|_h = 1} |f(w_\epsilon^h - l^h)| + C \sup_{w_\epsilon^h \in \hat{H}^h,\ \|w_\epsilon^h\|_h = 1} |a(u_\epsilon, w_\epsilon^h - l^h)|. \end{aligned} \tag{3.10}$$

First let us estimate the second term of the R.H.S. of (3.10). The first term can be estimated analogously. Since w_ϵ^h is a linear combination of the base functions, w_ϵ^h satisfies $\nabla_i a_\epsilon^{ij} \nabla_j w_\epsilon^h = 0$ in S. Using the expansion for $w_\epsilon^h = l^h + \epsilon\chi^p \nabla_p l^h + \epsilon\theta$, we have

$$\left| \int_K a_\epsilon^{ij} \frac{\partial u^\epsilon}{\partial x_j} \frac{\partial (w_\epsilon^h - l^h)}{\partial x_j} dx \right| \leq \left| \int_K a_\epsilon^{ij} \frac{\partial u^\epsilon}{\partial x_j} (\epsilon \nabla_i \chi^p) \nabla_p l^h dx \right| + \left| \int_K a_\epsilon^{ij} \frac{\partial u^\epsilon}{\partial x_j} \frac{\partial (\epsilon\theta)}{\partial x_j} dx \right|. \tag{3.11}$$

Observe that $u^\epsilon = u_0 + \epsilon\chi^p \nabla_p u_0 + \epsilon\theta^u$ and

$$a_\epsilon^{ij} \nabla_j (u_0 + \epsilon\chi^p \nabla_p u_0) = a_*^{ij} \nabla_j u_0 + g^{ij} \nabla_j u_0 + \epsilon\chi^k a_\epsilon^{ij} \nabla_j \nabla_k u_0,$$

where $g^{ij} = a_\epsilon^{ij} + a_\epsilon^{ip} \nabla_p \chi^j - a_*^{ij}$. Then (3.11) becomes

$$\begin{aligned} &\left| \int_K a_\epsilon^{ij} \frac{\partial u^\epsilon}{\partial x_j} \frac{\partial (w_\epsilon^h - l^h)}{\partial x_j} dx \right| \\ &\quad \leq \left| \int_K a_\epsilon^{ij} \frac{\partial u^\epsilon}{\partial x_j} (\epsilon \nabla_i \chi^p) \nabla_p l^h dx \right| + \left| \int_K a_\epsilon^{ij} \frac{\partial u^\epsilon}{\partial x_j} \frac{\partial (\epsilon\theta)}{\partial x_j} dx \right| \\ &\quad \leq |\nabla_p l^h| \left| \int_K a_*^{ij} \frac{\partial u_0}{\partial x_j} \epsilon \nabla_i \chi^p dx \right| + |\nabla_p l^h| \left| \int_K g^{ij} \frac{\partial u_0}{\partial x_j} \epsilon \nabla_i \chi^p dx \right| \\ &\qquad + |\nabla_p l^h| \left| \int_K a_\epsilon^{ij} \frac{\partial^2 u_0}{\partial x_j \partial x_k} \epsilon \chi^k (\epsilon \nabla_i \chi^p) dx \right| \\ &\qquad + |\nabla_p l^h| \left| \int_K a_\epsilon^{ij} \frac{\partial (\epsilon\theta^u)}{\partial x_j} \epsilon \nabla_i \chi^p dx \right| + \left| \int_K a_\epsilon^{ij} \frac{\partial u^\epsilon}{\partial x_j} \frac{\partial (\epsilon\theta)}{\partial x_j} dx \right|. \end{aligned} \tag{3.12}$$

Using Lemma 3.2 along with the facts that $u_0 \in H_0^2(\Omega) \bigcap W^{1,\infty}(\Omega)$ and $\chi^p \in W^{1,\infty}(Y)$ (the latter follows from $a^{ij} \in W^{1,p}(Y)$, $p > 2$; cf. Theorem 15.1 in [13]),

we conclude that the first two terms on the R.H.S. of (3.12) are less than $C\epsilon h|\nabla l|$ because $\langle g^{ij}\nabla_i\chi^p\rangle = 0$ ($|\nabla l| = \max_{p=1,2}|\nabla_p l^h|$). It can be easily shown that the third term on the R.H.S. of (3.11) is bounded by $C\epsilon h|\nabla l^h||u_0|_{2,K}$. Then

$$\left|\int_K a_\epsilon^{ij}\frac{\partial u^\epsilon}{\partial x_j}\frac{\partial(w_\epsilon^h - l^h)}{\partial x_j}dx\right| \le C\epsilon h|\nabla l^h| + C\epsilon h|\nabla l^h||u_0|_{2,K} + C|\nabla l^h||\epsilon\theta^u|_{1,K}|\epsilon\chi|_{1,K} + C|u_\epsilon|_{1,K}|\epsilon\theta|_{1,K}.$$

Noting that $h|\nabla l^h| \le C\|\nabla l^h\|_{L_2(K)}$ and taking into account

$$\|\nabla l^h\|_{L_2(K)} \le C\|\nabla w_\epsilon^h\|_{L_2(K)},$$

we have

$$\left|\int_K a_\epsilon^{ij}\frac{\partial u_\epsilon}{\partial x_j}\frac{\partial(w_\epsilon^h - l^h)}{\partial x_j}dx\right| \le C\epsilon\|\nabla w_\epsilon^h\|_{L_2(K)} + C\epsilon\|\nabla w_\epsilon^h\|_{L_2(K)}|u_0|_{2,K} + C\|\nabla w_\epsilon^h\|_{L_2(K)}|\epsilon\theta^u|_{1,K} + C|u_\epsilon|_{1,K}|\epsilon\theta|_{1,K}. \tag{3.13}$$

By Lemma 2.1 for $\epsilon\theta$ one has

$$|\epsilon\theta|_{1,K} \le C\frac{\epsilon}{h}\|\nabla l^h\|_{L_2(K)} \le C\frac{\epsilon}{h}\|\nabla w_\epsilon^h\|_{L_2(K)}.$$

Summing (3.13) over all $K \in \mathbf{K}^h$ and using the Cauchy–Schwarz inequality, we obtain

$$\left|\int_\Omega a_\epsilon^{ij}\frac{\partial u^\epsilon}{\partial x_j}\frac{\partial(w_\epsilon^h - l^h)}{\partial x_j}dx\right| \le C\frac{\epsilon}{h} + C_3\epsilon|u_0|_{2,\Omega} + C_2|\epsilon\theta^u|_{1,\Omega} + C_1\epsilon \le C\frac{\epsilon}{h} + C_2\sqrt{\epsilon} + C_1\epsilon,$$

where we have used $\|\nabla w_\epsilon^h\|_h = 1$ and $|\epsilon\theta^u|_{1,\Omega} \le \sqrt{\epsilon}$. Analogously,

$$\begin{aligned}\left|\sum_{K\in\mathbf{K}^h}\int_K f(x)(w_\epsilon^h - l^h)dx\right| &\le \left|\sum_{K\in\mathbf{K}^h}\int_K f\epsilon\chi^p\nabla_p l^h dx\right| + \left|\sum_{K\in\mathbf{K}^h}\int_K f\epsilon\theta dx\right| \\ &\le \sum_{K\in\mathbf{K}^h}\epsilon\|\nabla l^h\|_{L_2(K)}\|f\|_{L_2(K)} + \sum_{K\in\mathbf{K}^h}\|\epsilon\theta\|_{L_2(K)}\|f\|_{L_2(K)}. \end{aligned} \tag{3.14}$$

The second term on the R.H.S. of (3.14) can be treated as in the previous case. It can be easily shown that this term is bounded by $C\epsilon$. The first term is bounded by $C\epsilon$. Combining the estimates for conforming and nonconforming errors we have (3.4). □

Remark 3.2. The estimate (3.4) is better than our previous estimate for MsFEM without oversampling [10]. The leading order error without oversampling is $(\epsilon/h)^{\frac{1}{2}}$ due to the boundary layers in θ. Here we notice that the resonance still exists. Inspecting the proof, we see that the ϵ/h term in (3.4) comes from the terms containing θ^h as well as from the integrals containing χ (see (3.11)). The latter is in the form $\int_K a(x)b(x/\epsilon)dx$, where $b(x/\epsilon)$ is a periodic function with average zero and $a(x)$ is

independent of ϵ. We call the error generated by these terms the cell resonance error. The cell resonance is caused by the mismatch between the mesh size and the "perfect" sample size. We see below (section 4) that a similar term gives rise to the resonance error in the L_2 estimates. We show that this resonance error is second order in the L_2 estimates.

4. L_2 estimates. In this section, we will derive the L_2 norm estimates for MsFEM-os (oversampling). Although the cell resonance has been shown in the H^1 analysis, the L_2 analysis reveals more clearly the source of this cell resonance. This enables us to design effective methods to remove the cell resonance (see section 5). Following [10], we use a discrete error analysis to overcome the difficulty. First, we show that the L_2 norm error can be determined from the discrete l^2 norm error. For the latter, we compare the discrete solution of (2.1) with that of the homogenized (2.8) at the nodal points. The reason for using the discrete analysis is to investigate the cell resonance in detail. The discrete analysis shows that the cell resonance is the second order resonance error. More specifically, let us denote by u_0^h the numerical solution of (2.8) which is calculated using the MsFEM-os method. Since the homogenized coefficients a_*^{ij} are constants, MsFEM-os reduces to the conventional finite element method with linear base functions. Then we have [10]

$$\begin{aligned} \|u_\epsilon - u_\epsilon^h\|_{L_2(\Omega)} &\le \|u_\epsilon - u_0\|_{L_2(\Omega)} + \|u_0 - u_0^h\|_{L_2(\Omega)} + \|u_0^h - u_\epsilon^h\|_{L_2(\Omega)} \\ &\le C\epsilon + C_1 h^2 + C_2 \|u_0^h - u_\epsilon^h\|_{l_2(\Omega)}. \end{aligned} \tag{4.1}$$

Let $N \sim 1/h^2$ be the number of nodal points. Denote $\|\cdot\|$ the standard maximum norm of matrices in $\mathbf{R}^{N\times N}$ and $|\cdot|$ the maximum norm of vectors in $\mathbf{R}^N$. The linear system of equations for U_ϵ^h is

$$A_\epsilon^h U_\epsilon^h = f_\epsilon^h, \tag{4.2}$$

where A_ϵ^h and f_ϵ^h are obtained from $a(u^h, v)$ and $f(v)$ by using $v = \phi_\epsilon^i$ for $i = 1, \dots, N$. Similarly, for U_0^h one has

$$A_0^h U_0^h = f_0^h, \tag{4.3}$$

where A_0^h and f_0^h are obtained by applying $v = \phi_0^i$ $(i = 1, \dots, N)$ to $a^*(u_0^h, v) = f(v)$ with

$$a^*(u_0^h, v) = \int_\Omega a_*^{ij} v_{,i} u_{0,j}^h dx.$$

The "comma" notation is used here and below for partial derivatives. Note that U_ϵ^h is not exactly the nodal values of u_ϵ^h since the base functions ϕ_ϵ^h have different values at the same node corresponding to triangular elements with the common vertex at this nodal point. On the other hand the values of the base functions in different triangles at a common node point differ from each other by order ϵ. So we sometimes still refer to U_ϵ^h as nodal values of the numerical solution. The main result of the section is summarized as follows.

L_2 error estimate. *Assuming that the distance between K and ∂S is of order h we have*

$$|U_\epsilon^h - U_0^h| \le C_r \epsilon^2/h^2 + C_1 \epsilon |\ln h| + C_\theta \epsilon/h \qquad (h \gg \epsilon). \tag{4.4}$$

Here the error $C_\theta \epsilon/h$ comes from the terms containing θ (see (4.5)) and $C_r\epsilon^2/h^2$ corresponds to the cell resonance error mentioned before. We would like to point out that (4.4) seems to be an overestimate for the actual computations (see [8, 9]). In our numerical experiments, we find $C_1\epsilon|\ln h|$ to be the dominant error in most cases, even when ϵ and h are of the same order. In particular, we do not observe the resonance error $C_\theta\epsilon/h$ in our computations. This indicates that C_θ in (4.4) is negligibly small. On the other hand we observe the presence of the resonance error $C_r\epsilon^2/h^2$ in our numerical experiments for very special choices of h/ϵ. Later in this section we are going to discuss the cell resonance error in more detail. It follows from the above estimate and (4.1) that for $h \gg \epsilon$,

$$\|u_\epsilon - u_\epsilon^h\|_{L_2(\Omega)} \le C\epsilon + C_1 h^2 + C_r\epsilon^2/h^2 + C_1\epsilon|\ln h| + CC_\theta\frac{\epsilon}{h}.$$

The above L_2 estimate can be proved rigorously. To illustrate the main ideas more clearly, we will present the error analysis through the following steps.

Step one: Asymptotic expansion of the discrete operator. Using the asymptotic expansion of the base functions

$$\phi_\epsilon^j = \phi_0^j + \epsilon\chi^p\nabla_p\phi_0^j + \epsilon\theta_\epsilon^j, \tag{4.5}$$

we can expand the stiffness matrix of the problem around the stiffness matrix of the homogenized problem as follows:

$$A_\epsilon^h = A_0^h + \epsilon A_1^h, \qquad f_\epsilon^h = f_0^h + \epsilon f_1^h, \tag{4.6}$$

where A_1^h and f_1^h are assembled from

$$A_{1_{kl}}^e = -\int_K \sigma^{ij}(\phi_{0,j}^k\theta_{,i}^l + \phi_{0,j}^l\theta_{,i}^k)dx + \int_K \epsilon a_\epsilon^{ij}\theta_{,i}^k\theta_{,j}^l dx + \frac{1}{\epsilon}\int_K \tilde{\sigma}^{ij}\phi_{0,j}^l\phi_{0,i}^k dx \tag{4.7}$$

and

$$f_{1_i}^e = -\int_K f(x)\chi^p\phi_{0,p}^i dx - \int_K f(x)\theta^i dx, \tag{4.8}$$

respectively, where $K \in \mathbf{K}^h$; σ^{ij} is the integrand in (2.9), i.e.,

$$\sigma^{ij} = a^{ij}(\delta_{jk} + \nabla_j^y\chi^k),$$

and $\tilde{\sigma}^{ij} = \sigma^{ij} - a_*^{ij} - \sigma^{pj}\chi_{,y_p}^i$. Note that from (2.9) and (2.10) we have $\langle\sigma^{ij}\rangle = a_*^{ij}$ and $\sigma_{,y_i}^{ij} = 0$, respectively. Using integration by parts we further obtain $\langle\sigma^{pj}\chi_{,y_p}^i\rangle = 0$, and hence $\langle\tilde{\sigma}^{ij}\rangle = 0$. Later, we will use this property to extract a conservative structure in A_1^h operator.

Step two: Asymptotic expansion of the discrete solution. Next we will obtain an asymptotic expansion of U_ϵ^h in powers of ϵ. This asymptotic expansion can be shown to be convergent later in this section.

$$U_\epsilon^h = U_0^h + \sum_{i\ge 1}\epsilon^i U_i^h, \tag{4.9}$$

where U_i^h $(i \ge 1)$ are given by

$$A_0^h U_1^h = f_1^h - A_1^h U_0^h, \tag{4.10}$$

$$A_0^h U_i^h = -A_1^h U_{i-1}^h \qquad (i > 1). \tag{4.11}$$

We note that the expansion enables us to avoid dealing with the inverse of A_ϵ^h whose entries are oscillatory. In contrast, A_0^h and its inverse are well understood. Furthermore, we note that because the homogenized base functions are linear, the expansions of A_ϵ^h and f_ϵ^h have only two terms. This simplifies the equations for U_i^h $(i > 1)$ (i.e., (4.11)) and their analysis.

Step three: Recognizing the conservative structure in A_1^h. Let $G_0^h = (A_0^h)^{-1}$. Then we have

$$U_1^h = G_0^h f_1^h - G_0^h A_1^h U_0^h, \tag{4.12}$$

$$U_i^h = -G_0^h A_1^h U_{i-1}^h \qquad (i > 1). \tag{4.13}$$

In [10], it was found that the linear system (4.12) has a conservative structure which leads to cancellation of the resonance error by using "summation by parts." Here we give a simpler proof which is the content of the following lemma.

LEMMA 4.1. *Let $G_0^h = (A_0^h)^{-1}$. Then for any $V \in \mathbf{R}^N$*

$$(G_0^h A_1^h V)_i = \sum_{k=1}^{N} \sum_{j=1}^{j<k} A_1^{kj} (V^j - V^k)(G_0^{ik} - G_0^{ij}). \tag{4.14}$$

Proof. First, we note that

$$\sum_{k=1}^{3} \phi_\epsilon^k = 1, \qquad \sum_{k=1}^{3} \phi_0^k = 1$$

for all $K \in \mathbf{K}^h$. Then it follows from (4.7) [10] that

$$\sum_{j=1}^{N} A_1^{ij} = 0.$$

Thus,

$$\sum_{j=1}^{N} A_1^{ij} V^j = \sum_{j=1}^{N} A_1^{ij} V^j - \sum_{j=1}^{N} A_1^{ij} V^i = \sum_{j=1,\ j\neq i}^{N} A_1^{ij} (V^j - V^i).$$

By the symmetry of A_1^{ij} (cf. (4.7)), we can combine $A_1^{kj}(V^j - V^k)$ and $A_1^{jk}(V^k - V^j)$ in the sum of $G_0^{ik} A_1^{kj} V^j$ and get

$$\sum_{k=1}^{N} G_0^{ik} \sum_{j=1,\ j\neq k}^{N} A_1^{kj} (V^j - V^k) = \sum_{k=1}^{N} \sum_{j=1}^{j<k} A_1^{kj} (V^j - V^k)(G_0^{ik} - G_0^{ij}).$$

Thus, (4.14) follows immediately. ∎

We note that A_1^h is sparse. The following definitions are helpful to describe the sparsity. We say two nodal points are neighbors if they are the vertices of the same element. For a nodal point with index k, we denote the set of the indices of its neighbors by O_k (see Figure 4.1). We have $A_1^{ij} = 0$ if $j \notin O_i$. Thus, the R.H.S. of (4.14) can be written as

$$\sum_{k=1}^{N} \sum_{j\in O_k, j<k} A_1^{kj} (V^j - V^k)(G_0^{ik} - G_0^{ij}).$$

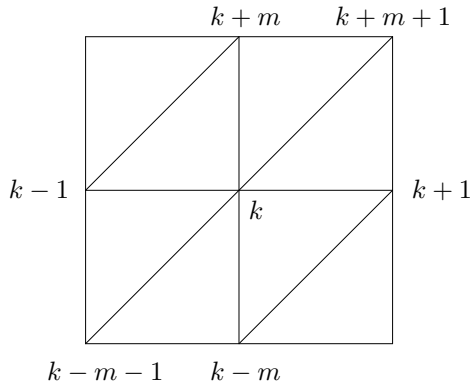

FIG. 4.1. *Example of neighboring nodes in a triangulation of a rectangular mesh with $m \times n$ nodal points. $O_k = \{k-m-1, k-m, k-1, k+1, k+m, k+m+1\}$.*

Since the summation of j depends only on the geometry of the mesh and is independent of N (or h), we can safely ignore it when estimating the above expression. This will greatly simplify the presentation. Introduce a difference operator D such that for any $V \in \mathbf{R}^N$, $(DV)_k \equiv V^k - V^j$ for some $j \in O_k$. Note that j is not specified in the definition. For a matrix, D applies to its row vectors unless otherwise stated. With this definition, from (4.14), and Lemma 4.1 we get

$$U_1^h = G_0^h f_1^h - DG_0^h A_1^h DU_0^h. \tag{4.15}$$

Here, the important features to notice are the differences on G_0^h and U_0^h; the details of these differences, however, are not important and are hidden in D. For example, for an interior nodal point k (see Figure 4.1) $DG_0^h A_1^h DU_0^h$ becomes

$$\begin{aligned} DG_0^h A_1^h DU_0^h = {} & D_{x+y} G_0^{ik} A_1^{kk-m-1} D_{x+y} U_0^k + D_y G_0^{ik} A_1^{kk-m} D_y U_0^k \\ & + D_x G_0^{ik} A_1^{kk-1} D_x U_0^k, \end{aligned} \tag{4.16}$$

where $D_{x+y} G_0^{ik} = G_0^{ik-m-1} - G_0^{ik}$, $D_y G_0^{ik} = G_0^{ik-m} - G_0^{ik}$, $D_x G_0^{ik} = G_0^{ik-1} - G_0^{ik}$, and similarly for U_0^k. We would like to note that in DG_0 the difference operator is applied to the second index of G_0.

Step four: Additional conservative structure in A_1^h and error cancellations. To further exploit the conservative structure in A_1^h, we break A_1^h into two parts, Λ_1^e and Λ_2^e, where Λ_1^e are those integrals in (4.7) containing θ_ϵ and Λ_2^e are the integral without θ_ϵ, i.e.,

$$\begin{aligned} \Lambda_{1_{kl}}^e &= -\int_K \sigma^{ij} (\phi_{0,j}^k \theta_{,i}^l + \phi_{0,j}^l \theta_{,i}^k) dx + \int_K \epsilon a_\epsilon^{ij} \theta_{,i}^k \theta_{,j}^l dx, \\ \Lambda_{2_{kl}}^e &= \frac{1}{\epsilon} \int_K \tilde{\sigma}^{ij} \phi_{0,j}^l \phi_{0,i}^k dx. \end{aligned} \tag{4.17}$$

Similarly, let F_1^e and F_2^e be the first and second integrals in (4.8). Moreover, let Λ_i^h be the matrix whose entries are assembled from Λ_i^e $(i = 1, 2)$ such that

$$A_1^h = \Lambda_1^h + \Lambda_2^h. \tag{4.18}$$

F_1^h and F_2^h are defined similarly.

A key observation made in [10, 8] is that the MsFEM can be improved if f_1^h and A_1^h can be written in difference forms. Such difference structures enable further error cancellation and hence reduce the error. It was found [10] that Λ_2^h can be written in difference forms; however, the boundary layer structures in θ_ϵ^j prevent Λ_1^h from having a difference structure. Therefore, the idea of oversampling is introduced to reduce the boundary layers [8]. In Appendix A, we will show Λ_2^h and f_1^h have a difference structure, i.e.,

$$\Lambda_2^h = D\lambda^h \quad \text{and} \quad f_1^h = D\hat{f}^h, \tag{4.19}$$

where λ^h is a matrix and D applies to its column vectors. Consequently, for the first term on the R.H.S. of (4.15), we have

$$|G_0^h f_1^h| = |G_0^h D\hat{f}^h| = |DG_0^h \hat{f}^h| \le \|DG_0^h\|\,|\hat{f}^h| \le C\epsilon/h + C_1. \tag{4.20}$$

It should be noted that in obtaining the second equality, we have used the summation by parts. The details in this step are not shown. Notably, we ignored the sums of the boundary terms generated by the summation by parts (analogous to the boundary integrals produced by the integration by parts). These sums are at a lower dimension and do not exceed the interior sums. Below, we use the summation by parts in the above symbolic fashion. We remark that it can be justified rigorously (cf. [10, 8]). Using (4.22), (4.24), and (4.25) we obtain

$$|DG_0^h \Lambda_2^h DU_0^h| = |D^2 G_0^h \lambda^h DU_0^h + DG_0^h \lambda^h D^2 U_0^h| \le C|\ln h| + C_1. \tag{4.21}$$

Step five: Estimate for U_1^h. By the oversampling, the error due to θ_ϵ^j is reduced to lower order. Using $\|\nabla\theta_\epsilon^j\|_{L_\infty(K)} \le C/h$ (which follows from Assumption A) and the sparsity of A_1^h, we can show that

$$\|\Lambda_1^h\| \le C/h, \qquad \|\Lambda_2^h\| \le C/h, \tag{4.22}$$

$$|F_1^h| \le C\epsilon, \qquad |F_2^h| \le Ch. \tag{4.23}$$

Moreover, it can be shown that

$$|DU_0^h| \le Ch, \qquad |D^2 U_0^h| \le Ch^2, \tag{4.24}$$

$$\|G_0^h\| \le C/h^2, \qquad \|DG_0^h\| \le C/h, \qquad \|D^2 G_0^h\| \le C\ln|h|. \tag{4.25}$$

Combining (4.21), (4.18), (4.20), and our estimates for Λ_1^h, we obtain the estimate for U_1^h:

$$|\epsilon U_1^h| \le C\epsilon|\ln h| + C_1\epsilon + C_\theta \frac{\epsilon}{h}. \tag{4.26}$$

Estimate (4.26) shows that U_1^h does not contain the cell resonance error.

Step six: Estimate for higher order terms, U_i^h ($i > 1$), and the cell resonance. To find the effect of the cell resonance in the L_2 estimate for U_ϵ^h, we need to estimate U_i^h ($i > 1$). It turns out that the higher order terms behave differently from the leading order term, U_1^h. To illustrate this point, let us consider U_2^h. From (4.11) we have

$$U_2^h = DG_0^h A_1^h DU_1^h = DG_0^h(\Lambda_1^h + \Lambda_2^h)DU_1^h.$$

TABLE 5.1
$\|U_\epsilon^h - U_0^h\|_{l_2}$. *Note:* $\alpha = h/\epsilon$ *is irrational in the middle column.*

h	$\alpha = 1.5$		$\alpha \approx 1.5396$		$\alpha = 1.5625$	
	l_2	rate	l_2	rate	l_2	rate
1/16	1.85e-4		1.75e-4		1.58e-4	
1/32	1.01e-4	0.87	9.70e-5	0.85	8.53e-5	0.89
1/64	5.60e-5	0.85	4.57e-5	1.09	4.44e-5	0.94
1/128	3.51e-5	0.68	2.45e-5	0.90	2.26e-5	0.97
1/256	2.76e-5	0.34	1.30e-5	0.91	1.14e-5	0.99
1/512	2.57e-5	0.10	6.74e-6	0.95	5.75e-6	0.99
1/1024	2.54e-5	0.02	3.31e-6	1.02	2.91e-6	0.98

From the definition of A_1^h, it is clear that the leading order error comes from Λ_1^h as well as from Λ_2^h. The error from Λ_2^h we call the cell resonance error. The difference structure of Λ_2^h and summation by parts do not help reduce the cell resonance error. The reason is that U_1^h, unlike U_0^h, is oscillatory and its change over a distance of h is no longer small. In fact, DU_1^h and the higher order differences $D^pU_1^h$ $(p > 1)$ are all $O(1)$. Thus $|U_2^h| \le C/h^2$. Similarly, it can be shown that U_i^h $(i > 1)$ are of order $1/h^i$ or $1/h^i \ln(|h|)$. Consequently, the formal series (4.9) converges for small ϵ/h. Using estimates for A_ϵ^{-1} it can be shown that the formal series (4.9) converges to the solution.

5. A remark on the cell resonance and averaged oversampling method. It is evident from the derivation that the resonance error, $C_r\epsilon^2/h^2$, in (4.4) is due to Λ_2^h, which is not improved by the oversampling. In this section, we further investigate the effect of the cell resonance through numerical experiments. To better understand Λ_2^h, we consider

$$\Lambda_{2_{kl}}^e = \frac{1}{\epsilon}\int_K \tilde{\sigma}^{ij}\phi_{0,j}^l\phi_{0,i}^k dx$$

(see (4.7)). Because $\langle\tilde{\sigma}^{ij}\rangle = 0$, applying Lemma A.1 to our triangulation (Figure 4.1), we deduce that $\Lambda_2^h = 0$ if the $\alpha \equiv h/\epsilon$ is an integer. It follows immediately that $|U_\epsilon^h - U_0^h| = O(\epsilon)$ and there is no cell resonance. In general, $\|\Lambda_2^h\|$ depends on the value of α; the norm is smaller if α is closer to an integer. This is illustrated in Table 5.1. This table shows the discrete l_2 norm of $U_\epsilon^h - U_0^h$ for three fixed α's for the test problem used in [8, 9]:

$$a_\epsilon = \frac{2 + P\sin(2\pi x/\epsilon)}{2 + P\cos(2\pi y/\epsilon)} + \frac{2 + \sin(2\pi y/\epsilon)}{2 + P\sin(2\pi x/\epsilon)}, \tag{5.1}$$

$(P = 1.8)$, and $f = -1$. We see degrading convergence as h decreases when $\alpha = 1.5$. Moreover, the error stagnates at the bottom of the column. This result shows that the constant C_r in front of the resonance error (see (4.4)) is quite small. In this case, it is less than 6×10^{-5} (or 0.3% of the maximum value of U_0^h). For the other two α's, however, the convergence appears to be first order. The C_r in both cases must be extremely small if not zero. Note that the differences in the values of α are small. Thus, C_r is sensitive to α and depends on α in a subtle way. In [7] we investigate the behavior of $G_\epsilon^h = (A_\epsilon^h)^{-1}$ with respect to integer and fractional parts of α. Our numerical observations reveal that the cell resonance term is dominant in G_ϵ^h. In fact, the effect of the cell resonance term increases as the fractional part of α approaches 0.5 and the integer part of α decreases to 1. These observations (see also [7]) are

consistent with the fact that the cell resonance is the dominant resonance error in the above computation for the case $\alpha = 1.5$. We do not observe any resonance error for other values of α. In fact, for other choices of a_ϵ, we do not find any trace of the resonance error due to θ. The fact that C_r is generically small has been demonstrated before [9, 8] by various examples including some random cases. Another important observation from this example (and many others we have tested) is that there is no trace of the resonance error, $C_\theta \epsilon/h$, due to θ. This seems to suggest that our error estimate, $C_\theta \epsilon/h$, is an overestimate. The resonance error due to θ is negligibly small. This may be due to some subtle error cancellation that has not been taken into account in our error analysis. This will be studied further in the future. There could be several factors contributing to the subtle dependence of C_r on α. Among them we notice that the expansion of U_ϵ^h is strictly valid for $\epsilon \ll h$, while the cancellations among terms of different orders can be important when h and ϵ are close. Therefore, we study below the difference between U_ϵ^h and U_0^h directly. Following the derivations in the previous section, we have

$$\begin{aligned} U_\epsilon^h - U_0^h &= \epsilon(DG_\epsilon^h A_1^h DU_0^h + G_\epsilon^h f_1^h) \\ &= \epsilon[DG_\epsilon^h \Lambda_1^h DU_0^h + D^2 G_\epsilon^h \lambda^h DU_0^h + DG_\epsilon^h \lambda^h D^2 U_0^h + DG_\epsilon^h \hat{f}^h], \end{aligned}$$

where $G_\epsilon^h = (A_\epsilon^h)^{-1}$. Note that we have used (4.19). Having obtained (4.22) and (4.23), our remaining problem is to estimate DG_ϵ^h and $D^2 G_\epsilon^h$. Using the asymptotic expansion of G_ϵ^h with respect to ϵ we can show that $\|DG_\epsilon^h\| = O(1/h)$ for $\epsilon \ll h$. Our numerical tests indicate that this estimate is indeed valid even for $\alpha = h/\epsilon \sim 1$, e.g., those values in Table 5.1. If this is the case, then we have

$$U_\epsilon^h - U_0^h = \epsilon D^2 G_\epsilon^h \lambda^h DU_0^h + O(\epsilon) + C_\theta \epsilon/h. \tag{5.2}$$

The cell resonance error is then contained in first term on the R.H.S. As usual, we can estimate $\|D^2 G_\epsilon^h\|$, which gives the upper bound of the error. Indeed, noting that $\|\lambda^h\|$ has the same order as $\|\Lambda_2^h\|$, i.e., $1/h$, and that $|DU_0^h| \sim h$, we have

$$|U_\epsilon^h - U_0^h| \le C\epsilon \|D^2 G_\epsilon^h\| + O(\epsilon) + C_\theta \epsilon/h. \tag{5.3}$$

It should be stressed that in this upper bound, we have ignored the subtle cancellation in $D^2 G_\epsilon^h \lambda^h DU_0^h$. Therefore (5.3) tends to be an overestimate. According to (5.3) the $|\ln h|$ behavior of $\|D^2 G_\epsilon^h\|$ implies first order convergence of the solution but *not* vice versa. In fact, for $\alpha = 1.5625$ our numerical tests indicate that $\|D^2 G_\epsilon^h\|$ behaves almost the same as in the case of $\alpha = 1.5$ (the $1/h$ behavior); however, the solutions converge with first order (see Table 5.1). We attribute this delicate situation to the complicated error cancellation in $D^2 G_\epsilon^h \lambda^h DU_0^h$, which depends on the coefficients a_ϵ and α. In [7] we discuss the behavior of $\|D^2 G_\epsilon^h\|$ with respect to α and the regions of α for which $\|D^2 G_\epsilon^h\|$ is of order $|\ln h|$. In particular, we show that the regions of the $|\ln h|$ behavior of $\|D^2 G_\epsilon^h\|$ are rapidly extending as the fractional part of α approaches to zero or the integer part of α increases.

The exact account of the error cancellation in $D^2 G_\epsilon^h \lambda^h DU_0^h$ is not well understood. Here, we propose an averaged oversampling method to eliminate this cell resonance error in the special case of periodic oscillation, i.e., $a_\epsilon = a(x/\epsilon)$. A more general approach will be presented in a subsequent paper which completely eliminates the cell resonance error for general oscillatory coefficients. To demonstrate the main idea of the averaged oversampling method, we assume that the homogenized coefficients

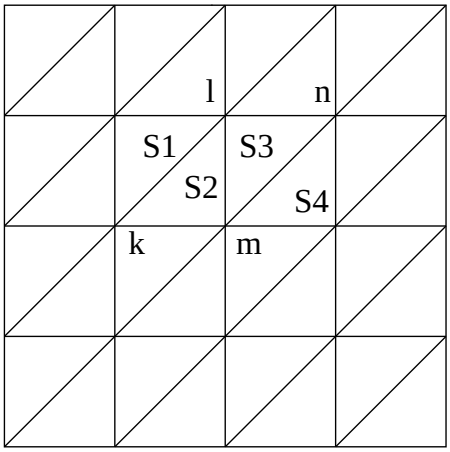

FIG. 5.1.

are constants and the mesh is uniform (Figure 5.1). As we showed above the stiffness matrix corresponding to the oversampling method can be written as

$$A = A_0 + \epsilon A_1, \quad A_1 = A_1^{cell} + A_1^{\theta},$$

where A_1^{cell} and A_1^{θ} consist of the cell term and of terms depending on θ, respectively. To simplify the presentation, we introduce the following notation.

Definition. Denote Δ_{kl} the union of two triangular elements with the common side whose vertices are k and l. For example, Δ_{kl} (Figure 5.1) is the union of elements S_1 and S_2.

We assume that the nodal points are numerated lexicographically from 1 to N. Then the entries of A_0^{ij} $(i, j = 1, \ldots, N)$ are the same on the 2-D mesh along each diagonal with i being the index of interior nodal point. Indeed, it can be easily checked that $\xi^{kl} = a_*^{ij} \nabla_i \phi_0^k \nabla_j \phi_0^l$ is the same for k and l for which Δ_{kl} can be translated to each other (see example below). For the same reason, the elements of the cell matrix A_1^{cell} has the same integrands for those k and l for which Δ_{kl} can be translated to each other, especially the elements on the same diagonals corresponding to the interior nodal points (but notice that they are integrated over different domains). For example, Δ_{kl} and Δ_{mn} (Figure 5.1) can be translated to each other and

$$A_{cell}^{kl} = \frac{1}{h^2} \int_{S_1+S_2} \sigma^{12} dx, \qquad A_{cell}^{mn} = \frac{1}{h^2} \int_{S_3+S_4} \sigma^{12} dx.$$

Consequently, taking the arithmetic average over the diagonal entries of A^{ij} (i being the index of the interior nodal point) and replacing them with the average, we obtain a new stiffness matrix with the same A_0. But by doing so we reduce ϵA_1^{cell} and get rid of resonance. Indeed, the arithmetic average of the above mentioned elements in the cell matrix with the same integrand $h^{-2} \int_K \zeta(x/\epsilon) dx$ will be averaged to $\int_\Omega \zeta(x/\epsilon) dx$, where $\mathrm{diam}(\Omega)$ is of order 1. The value of this integral is of order ϵ. It can be shown that the domain of integration will sum up to a single connected domain with order one size. Thus using averaged oversampling, we actually approximate the effective equation accurately without suffering from the cell resonance error. Based on the above observation, it is easy to show that the convergence of MsFEM with averaged oversampling does not suffer from the cell resonance and the convergence rate for

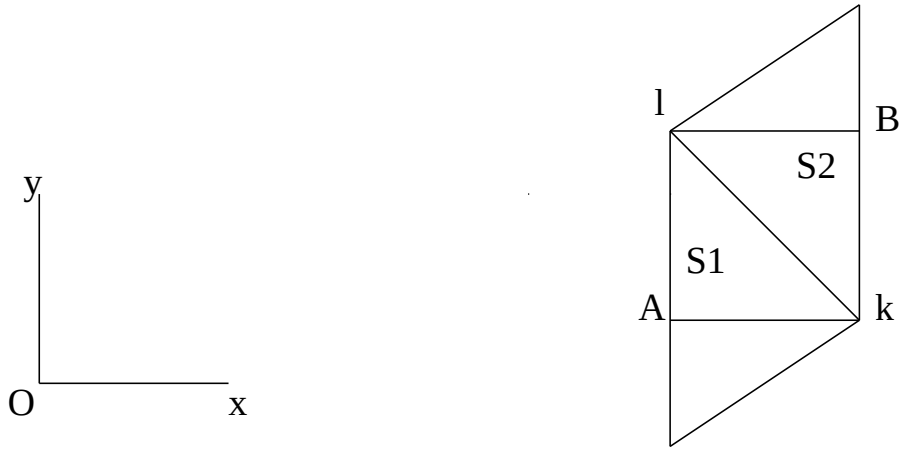

FIG. A.1.

problems with periodic coefficients is $C_\theta \epsilon/h + C(\epsilon \log(h) + h)$. It is conceivable that such technique may be also helpful for more general multiscale coefficients if we apply this averaging technique locally.

Appendix A. Difference form of the cell term. To show that Λ_2^h can be written in the difference form, we need just to prove that $\phi_{0,1}^k \phi_{0,1}^l$, $\phi_{0,2}^k \phi_{0,2}^l$, and $\phi_{0,1}^k \phi_{0,2}^l + \phi_{0,2}^k \phi_{0,1}^l$ are the same constants in each of two triangular elements with the common side kl. In this way the integrand is the same in the union of these two elements. Since the average of the integrand of Λ_2^h is zero, it is the divergence of a periodic field. Consequently, Λ_2^h can be written as a difference of integrals over the boundaries. To show that these constants are the same for our configuration, we need the following lemma.

LEMMA A.1. *If $S_1 \bigcup S_2$ is a parallelogram, then $\phi_{0,j}^k \phi_{0,p}^l + \phi_{0,j}^l \phi_{0,p}^k$ is the same in S_1 and S_2.*

Proof. In this case we have

$$\phi_{0,j}^l = -\phi_{0,j}^k \quad (j = 1, 2).$$

Note that the line segment kA is parallel to the line segment lB, which is also parallel to the line segment Ox; see Figure A.1. Moreover, since $|kA| = |lB|$ and $\phi_0^k|_{S_1}(k) = 1$, $\phi_0^k|_{S_1}(A) = 0$, $\phi_0^l|_{S_2}(l) = 1$, $\phi_0^l|_{S_2}(B) = 0$, we have

$$\phi_{0,1}^k|_{S_1} = -\phi_{0,1}^l|_{S_2}.$$

The same can be shown for $j = 2$. ∎

Remark A.1. It can be shown that this difference structure leads to the summation by parts in (4.21).

Now let us show that $\int_K f\theta^k dx$ can be written in the difference form ($\int_K f\chi^p \phi_{0,p}^k dx$ is similar). Consider the configuration illustrated in Figure A.2. By reordering the terms, we obtain

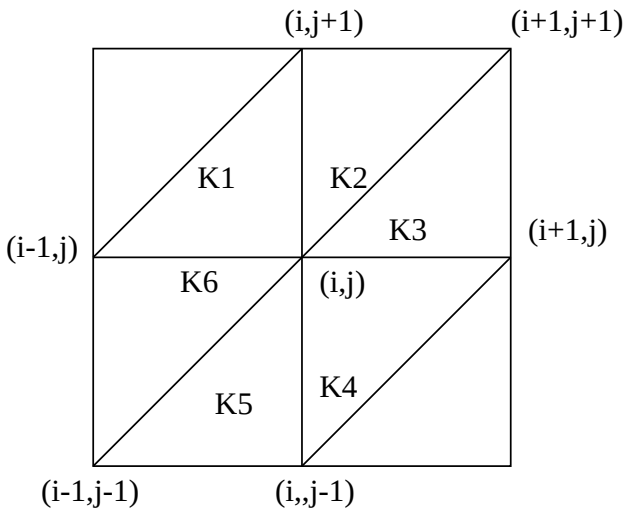

FIG. A.2.

$$\int f\theta^{ij}dx = \sum_{i=1}^{6}\int_{K_i} f\theta^{ij}dx = \left(\int_{K_5} f\theta^{ij}dx - \int_{K_1} f\theta^{ij+1}dx\right) + \left(\int_{K_3} f\theta^{ij}dx - \int_{K_1} f\theta^{i-1j}dx\right) + \left(\int_{K_4} f\theta^{ij}dx - \int_{K_6} f\theta^{i-1j}dx\right) + \left(\int_{K_2} f\theta^{ij}dx - \int_{K_6} f\theta^{i-1j-1}dx\right). \tag{A.1}$$

Here we have used $\sum_{k=1}^{3}\theta^k = 0$ on $K \in \mathbf{K}^h$. This difference structure leads to the summation by parts in $G^h f_1^h$.

Appendix B. The estimate for linear function. In this section we show that on any triangular element K

$$\|\nabla l\|_{L_2(K)} \le C\|\nabla u\|_{L_2(K)}, \tag{B.1}$$

where l is a linear function and u satisfies the following equation on a triangular domain $S \supset K$:

$$\begin{aligned} \nabla_i a^{ij}\nabla_j u &= 0 \quad \text{in } S, \\ u &= l \quad \text{on } \partial S. \end{aligned} \tag{B.2}$$

Let us consider the difference of u and l: $\phi = u - l$. Clearly ϕ satisfies the following equation:

$$\begin{aligned} \nabla_i a^{ij}\nabla_j \phi &= -\nabla_i a^{ij}\nabla_j l \quad \text{in } S, \\ \phi &= 0 \quad \text{on } \partial S. \end{aligned} \tag{B.3}$$

Introducing the auxiliary function v_i $(i = 1, 2)$ defined by

$$\begin{aligned} \nabla_i a^{ij}\nabla_j v_k &= -\nabla_i a^{ik} \quad \text{in } S, \\ v_k &= 0 \quad \text{on } \partial S. \end{aligned} \tag{B.4}$$

We can express the solution of (B.3) as a linear combination of v_i. This gives

$$u = l + v_i \nabla_i l = l + v_i \alpha_i,$$

where $\alpha_i = \nabla_i l$ are constants. Then we obtain

$$\begin{aligned}\int_K (\nabla u)^2 dx = \alpha_1^2 \int_K [(1+\nabla_1 v_1)^2 + (\nabla_2 v_1)^2] dx \\ + \alpha_2^2 \int_K [(1+\nabla_2 v_2)^2 + (\nabla_1 v_2)^2] dx \\ + 2\alpha_1\alpha_2 \int_K [(1+\nabla_1 v_1)\nabla_1 v_2 + (1+\nabla_2 v_2)\nabla_2 v_1] dx = (A\alpha, \alpha),\end{aligned}$$

where A is the matrix with the following elements:

$$\begin{aligned}A^{11} &= \int_K [(1+\nabla_1 v_1)^2 + (\nabla_2 v_1)^2] dx, \\ A^{12} &= \int_K [(1+\nabla_1 v_1)\nabla_1 v_2 + (1+\nabla_2 v_2)\nabla_2 v_1] dx, \\ A^{21} &= \int_K [(1+\nabla_1 v_1)\nabla_1 v_2 + (1+\nabla_2 v_2)\nabla_2 v_1] dx, \\ A^{22} &= \int_K [(1+\nabla_2 v_2)^2 + (\nabla_1 v_2)^2] dx.\end{aligned}$$

It can be checked that

$$\begin{aligned}A^{11} &= \int_K [(1+\nabla_1 v_1)^2 + (\nabla_2 v_1)^2] dx > 0, \\ A^{22} &= \int_K [(1+\nabla_2 v_2)^2 + (\nabla_1 v_2)^2] dx > 0,\end{aligned}$$

and

$$(A\alpha, \alpha) > 0$$

are the sufficient conditions for $(A\alpha, \alpha) \geq d\alpha^2$, for some $d > 0$. For example under these conditions d can be chosen to be $2d = A^{11} + A^{22} - \sqrt{(A^{11} - A^{22})^2 + 4(A^{12})^2}$. Note that if $d > 0$, then $d > Ch^2$, from which it follows that $d\alpha^2 \geq C\|\nabla l\|_{L_2(K)}$. Assuming the opposite, i.e., any of the inequalities does not hold, we have that $v_1 + x_1$ or $v_2 + x_2$ or u is constant in K. Let's note that $v_i + x_i$ $(i = 1, 2)$ satisfy

$$\begin{aligned}\nabla_i a^{ij} \nabla_j (v_k + x_k) = 0 \quad \text{in} \quad S, \\ v_k + x_k = x_k \quad \text{on} \quad \partial S.\end{aligned} \tag{B.5}$$

Therefore $v_i + x_i$ $(i = 1, 2)$ cannot be constant in K [14]. Consequently, A^{11} and A^{22} are strictly positive. Also $\int_K (\nabla u)^2 dx > 0$ if $|\nabla l| > 0$ [14] which guarantees $(A\alpha, \alpha) > 0$. If $|\nabla l| = 0$ then (B.1) satisfies. This completes the proof of (B.1).

Appendix C. The proof of Lemma 2.1 (by Z. Chen[1]). Let S be a triangle in $\mathbf{R}^2$. We consider the following Dirichlet problem over S:

$$L_\epsilon \theta_\epsilon := -\frac{\partial}{\partial x_i}\left(a_{ij}\left(\frac{x}{\epsilon}\right)\frac{\partial \theta_\epsilon}{\partial x_j}\right) = 0 \quad \text{in } S, \tag{C.1}$$

[1]Institute of Mathematics, Academia Sinica, Beijing, 100080, People's Republic of China. Current address: Applied Mathematics, 217-50, Caltech, Pasadena, CA 91125 (zchen@ama.caltech.edu).

$$\theta_\epsilon = \chi^j\left(\frac{x}{\epsilon}\right)\frac{\partial u_0}{\partial x_j} \quad \text{on } \partial S, \tag{C.2}$$

where ϵ is assumed to be a small parameter, and $a(x/\epsilon) = (a_{ij}(x/\epsilon))$ is symmetric and satisfies the ellipticity condition $\lambda|\xi|^2 \le a_{ij}\xi_i\xi_j \le \lambda^{-1}|\xi|^2$ for all $\xi \in \mathbf{R}^2$ and for some positive constant λ. Furthermore, we assume that $a_{ij}(y)$ are $C^1(\mathbf{R}^2)$ periodic functions with respect to the unit cube Y, $\chi^j(y)$ are $C(\mathbf{R}^2)$ periodic functions, and u_0 is a function defined in S such that $\nabla u_0 \in L^\infty(\partial S)$. Let K be a subdomain in S such that $\operatorname{dist}(K, \partial S) \ge \delta_0 h_K$, where $\delta_0 > 0$ and h_K is the diameter of K. The purpose of this appendix is to show the following proposition.

PROPOSITION C.1. *There exists a constant C depending only on δ_0, λ, $\|a_{ij}\|_{C^1(Y)}$, and $\|\chi^j\|_{C(Y)}$ such that for all $\epsilon > 0$,*

$$\|\nabla\theta_\epsilon\|_{L^\infty(K)} \le \frac{C}{h_K}\|\nabla u_0\|_{L^\infty(\partial S)}. \tag{C.3}$$

The proof of this result is based on the following interior gradient estimate due to Avellaneda and Lin [1, Lemma 16].

LEMMA C.2. *There exists a constant C depending only on λ and $\|a_{ij}\|_{C^1(Y)}$ such that for all $\epsilon > 0$, u_ϵ satisfying $L_\epsilon u_\epsilon = 0$ in $B(0,r)$ and $\|u_\epsilon\|_{L^\infty(B(0,r))} < +\infty$, where $B(0,r) = \{x \in \mathbf{R}^2 : |x| \le r\}$, the following estimate is valid:*

$$\|\nabla u_\epsilon\|_{L^\infty(B(0,r/2))} \le C\, r^{-1}\|u_\epsilon\|_{L^\infty(B(0,r))}.$$

Proof of Proposition C.1. For any $x \in K$, since $B(x, \delta_0 h_K/2) \subset S$, we can apply Lemma C.2 to deduce that

$$\|\nabla\theta_\epsilon\|_{L^\infty(K)} \le \frac{C}{h_K}\|\theta_\epsilon\|_{L^\infty(S)}.$$

On the other hand, by using the maximum principle, we know from (C.1)–(C.2) that

$$\|\theta_\epsilon\|_{L^\infty(S)} \le \left\|\chi^j\left(\frac{x}{\epsilon}\right)\frac{\partial u_0}{\partial x_j}\right\|_{L^\infty(\partial S)} \le C\|\nabla u_0\|_{L^\infty(\partial S)}.$$

This completes the proof. □

Finally we remark that Proposition C.1 is also valid for elliptic system of equations. In this case, one should use the boundary estimate in [1, Lemma 21] for the Poisson kernel instead of the maximum principle to obtain the L^∞ bound for θ_ϵ.

REFERENCES

[1] M. AVELLANEDA AND F.-H. LIN, *Compactness methods in the theory of homogenization*, Comm. Pure Appl. Math., 40 (1989), pp. 803–843.

[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), pp. 945–981.

[3] A. BENSOUSSAN, J. L. LIONS, AND G. PAPANICOLAOU, *Asymptotic Analysis for Periodic Structure*, Stud. Math. Appl. 5, North–Holland, Amsterdam, 1978.

[4] A. BENSOUSSAN, J. L. LIONS, AND G. PAPANICOLAOU, *Boundary layer analysis in homogenization of diffusion equations with Dirichlet conditions in the half space*, in Proceedings of the International Symposium on Stochastic Differential Equations, K. Ito, ed., John Wiley, New York, Chichester, Brisbane, 1978, pp. 21–40.

[5] F. BREZZI, L. FRANCA, T. HUGHES, AND A. RUSSO, *b is equal to integral of g*, Comput. Methods Appl. Mech. Engrg., 145 (1997), pp. 329–339.

[6] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North–Holland, Amsterdam, 1978.
[7] Y. R. Efendiev, *The Multiscale Finite Element Method (MsFEM) and Its Applications*, Ph.D. thesis, Applied Mathematics, Caltech, Pasadena, CA, 1999.
[8] 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), pp. 169–189.
[9] T. Y. Hou and X. H. Wu, *A multiscale finite element method for PDEs with oscillatory coefficients*, in Numerical Treatment of Multi-Scale Problems, Notes Numer. Fluid Mech. 70, Vieweg-Verlag, Wiesbaden, 1999, pp. 58–69.
[10] T. Y. Hou, X. H. Wu, and Z. Cai *Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients*, Math. Comp., 68 (1999), pp. 913–943.
[11] T. Hughes, G. Feijoo, L. Mazzei, and J. Quincy, *The variational multiscale method: A paradigm for computational mechanics*, Comput. Methods Appl. Mech. Engrg., 166 (1998), pp. 3–24.
[12] V. V. Jikov, S. M. Kozlov, and O. A. Oleinik, *Homogenization of Differential Operators and Integral Functionals*, Springer-Verlag, Berlin, 1994.
[13] O. A. Ladyzhenskaia and N. N. Uraltseva, *Linear and Quasilinear Elliptic Equations*, Academic Press, New York, 1968.
[14] F.-H. Lin, *Nodal sets of solutions of elliptic and parabolic equations*, Comm. Pure Appl. Math., 44 (1991), pp. 287–308.
[15] 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, 127A (1997), pp. 1263–1299.
[16] E. Oriordan and M. Stynes, *A globally uniformly convergent finite element method for a singularly perturbed elliptic problem in* 2 *dimensions*, Math. Comp., 57 (1991), pp. 47–62.
[17] F. Santosa and M. Vogelius, *First-order corrections to the homogenized eigenvalues of a periodic composite medium*, SIAM J. Appl. Math., 53 (1993), pp. 1636–1668.
[18] G. Strang and G. J. Fix, *An Analysis of the Finite Element*, Prentice–Hall, Englewood Cliffs, NJ, 1973.
[19] Y. Zhang, T. Y. Hou, and X. H. Wu, *Convergence of a Petrov–Galerkin nonconforming multiscale finite element method*, SIAM J. Numer. Anal., to appear.