Numer. Math. (2002) 90: 459–486
Digital Object Identifier (DOI) 10.1007/s002110100274

Numerische Mathematik

Multiscale finite element for problems with highly oscillatory coefficients

Yalchin R. Efendiev[1], **Xiao-Hui Wu**[2,⋆]

[1] Institute for Mathematics and its Applications, University of Minnesota, 400 Lind Hall, 207 Church St. S.E., Minneapolis, MN 55455, USA
[2] Applied Mathematics 217-50, California Institute of Technology, Pasadena, CA 91125, USA

Received April 17, 1998 / Revised version received March 25, 2000 /
Published online June 7, 2001 –

Summary. In this paper, we study a multiscale finite element method for solving a class of elliptic problems with finite number of well separated scales. The method is designed to efficiently capture the large scale behavior of the solution without resolving all small scale features. This is accomplished by constructing the multiscale finite element base functions that are adaptive to the local property of the differential operator. The construction of the base functions is fully decoupled from element to element; thus the method is perfectly parallel and is naturally adapted to massively parallel computers. We present the convergence analysis of the method along with the results of our numerical experiments. Some generalizations of the multiscale finite element method are also discussed.

Mathematics Subject Classification (1991): 65N30

1 Introduction

Multiscale problems occur in many scientific and engineering disciplines, such as material science, earth and environmental science, petroleum engineering, just to name a few. These problems are characterized by the great number of spatial and time scales. They are difficult to analyze theoretically or solve numerically.

⋆ *Current address:* ExxonMobil Upstream Research Company, P. O. Box 2189, Houston, TX 77252, USA
Correspondence to: Y.R. Efendiev

When a standard finite element or finite difference method is applied to the multiscale problems, 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 beyond the reach of direct simulations. On the other hand, the large scale features of the solutions are often 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.

Here we analyze the multiscale finite element method (MsFEM) introduced in [13] for solving elliptic problems with oscillatory coefficients. The method has been successfully applied to flows and transport in random porous media [11, 12]. Similar methods for transport problems in oscillatory velocity fields and wave propagation through random media are also developed and will be reported separately. The purpose of this paper is to further establish the mathematical foundation of MsFEM. In particular, we provide estimates for MsFEM when it is applied to elliptic problems with finite number of well separated scales.

The idea of multiscale finite element method is to capture small scale information through the base functions constructed in the elements whose sizes are much larger than the the small scales of the problem and much smaller than the characteristic large scale of the problem. This is achieved by solving the finite element base functions from the leading order of homogeneous elliptic equation. In this way, the information at scales smaller than the mesh size is built into the base functions. These base functions are in general oscillatory. We remark that special base functions in finite element methods have been used by several authors in capturing multiscale solutions of PDE's. In particular, the works presented in [17, 5, 7, 14] are most relevant to the multiscale finite element method [13].

The advantage of MsFEM is its ability to reduce the size of a large scale computation. This offers significant savings in computer memory. For example, let N be the number of elements in each spatial direction, and let M be the number of subcell elements in each direction for solving the base functions. Then there are total $(MN)^n$ (n is the dimension) elements at the fine grid level. For a traditional FEM, the computer memory needed for solving the problem on the fine grid is $O(M^n N^n)$. In contrast, MsFEM requires only $O(M^n + N^n)$ amount of the memory. More discussions about the computational features of MsFEM can be found in [11].

In practice, many problems may have multiple scales ranging over a large interval. For example, in ground-water transport problems the scales can range from fine clay 10^{-4}mm to the coarse sand 1mm or gravels up to 60mm. Other problems may have continuous scales. In this paper, we

analyze the MsFEM for the problems with multiple but well separated scales. The latter problems are more difficult to analyze. We hope the present work could bridge the gap between the analyses of problems with and without scale separations.

Let ϵ_k $(k = 1, \ldots, n)$ be a sequence of decreasing well separated scales. We analyze the convergence rate of MsFEM for different choices of the mesh sizes, h. Thus, $h \ll \epsilon_n$ corresponds to well-resolved direct calculation, while $1 \gg h \gg \epsilon_1$ corresponds to the other extreme where only the largest scale feature are retained in the final solution. In most cases, however, h is between two neighboring small scales ϵ_k and ϵ_{k+1}. We establish H^1 and L_2 estimates for these three cases. For the most generic case with $\epsilon_k \gg h \gg \epsilon_{k+1}$, the leading order H^1 norm error of MsFEM is $C_1\sqrt{\epsilon_{k+1}/h} + C_2 h/\epsilon_k$. The first term in the estimate is due to the capturing the scales smaller than ϵ_{k+1} through the multiscale base functions, the second term is the error of resolving scales larger than ϵ_k with the homogenized part of the multiscale base functions.

As it can be noticed from the H^1 error estimate, a straightforward implementation of multiscale finite element method would fail if the mesh size is close to the small scales in the physical solution. This is an important phenomenon common in other upscaling methods, i.e., the resonance between the mesh scale and small physical scales. The improvement of the resonance error requires a new method of capturing the small scale effect [12, 9]. The second part of the error can also be improved. By changing the boundary conditions of the multiscale base functions, the homogenized part of the base function becomes higher order polynomials. Thus, we can obtain better convergence of order $h + (h/\epsilon_k)^m + \sqrt{\epsilon_{k+1}/h}$ with integer $m > 1$.

For the analysis of MsFEM we use the H^1 estimates for the first order correctors of partially homogenized solutions of problems with finite number of well separated scales. We present these estimates in the paper as they are not available in the literature. We would like to mention that in [2] a convergence result of homogenization (without estimate of convergence rate) has been obtained for the case with infinite number of scales.

We also present the L_2 estimates for MsFEM. The structure of the estimates are similar to that of the H^1 norm estimates. Because of some small terms in the H^1 estimate, which cannot be expressed through the L_2 norm of the source term of the problem, the L_2 norm estimate we derived in the Aubin-Nitsche fashion contains some overestimated terms. These terms, however, contain no resonance effect. Therefore, our L_2 estimate of the resonance error is tight. Moreover, the overestimated terms can indeed be eliminated by using the method introduced in [13], which compares the solution and its numerical counterpart at discrete nodal points.

Our error estimates are confirmed by our numerical experiments. The computations are extremely large and are done on parallel computers, e.g., the Intel Paragon computer. Even so, we can only test two scale problems. We confirm the L_2 estimate when the mesh size is between the physical scales ϵ_1 and ϵ_2, i.e. $\epsilon_1 \gg h \gg \epsilon_2$. The computations encounter difficulties because memory limitation prevents us from choosing well separated small scales in the tests. Consequently, it is difficult to separately verify each resonance error in the L_2 estimates, i.e., $C_1 h^2/\epsilon_1^2$ and $C_2\epsilon_2/h$. But different numerical examples demonstrate that in some cases either one of them can be dominating, whereas in other cases both may be important. Results for $h \ll \epsilon_2 \ll \epsilon_1$ are also presented.

The rest of the paper is organized as follows: The formulations of the 2-D problem and multiscale finite element method are introduced in the next section. In Sect. 3 we estimate the first order correctors for partially homogenized solutions. In Sects. 4 and 5 we derive H^1 and L_2 error estimates for MsFEM. The numerical results are presented in Sect. 6. The higher order MsFEM and other possible generalizations are discussed in the concluding remarks.

2 Formulations

In this section we introduce the model problem and the multiscale finite element method. In the paper 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:

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

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

$H_0^1(\Omega)$ consists of those functions in $H^1(\Omega)$ that vanish on $\partial\Omega$. $H^{-1}(\Omega)$ is dual space of $H_0^1(\Omega)$, i.e. the set of all continuous linear functionals on $H_0^1(\Omega)$. We define $H^{1/2}(\partial\Omega)$ 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. In the paper the space $C^k(\Omega)$, continuous functions along with their k^{th} derivatives is equipped with the norm

$$\|u\|_{C^k(\Omega)} = \sum_{\alpha=0}^{k} \max_{\Omega} |D^\alpha u|.$$

Throughout, C (with or without subscripts) denotes a generic constant, which is independent of ϵ and h (mesh size), unless otherwise stated and $C + C = C$, $C \cdot C = C$. When no confusion is possible the same symbol may denote different constants in different places.

Consider the following elliptic model problem

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

where $L_\epsilon = \nabla_i a_\epsilon^{ij} \nabla_j$ is the linear elliptic operator, ϵ is a small parameter, and a_{ij}^ϵ is symmetric and satisfies $\alpha|\xi|^2 \le \xi_i a_{ij}^\epsilon \xi_j \le \beta|\xi|^2$ for all $\xi \in \mathbf{R}^2$ with $0 < \alpha < \beta < \infty$, and $f \in C^0(\Omega)$.

Furthermore, we assume that $a^\epsilon(x)$ has the form:

$$a_\epsilon^{ij} = a^{ij}\left(\frac{x}{\epsilon_1}, \frac{x}{\epsilon_2}, \ldots, \frac{x}{\epsilon_n}\right) \tag{2.2}$$

where $\epsilon_1 \gg \epsilon_2 \gg \cdots \gg \epsilon_n$ is a set of n ordered length scales, which all depend on a single parameter ϵ. For example, ϵ_i, $i = 1, \ldots, n$ are some powers of ϵ, ϵ^{p_i} with $p_1 < p_2 < \cdots < p_n$. Moreover, for simplicity we assume $a^{ij}(y_1, y_2, \ldots, y_n)$ to be sufficiently smooth periodic functions in y_i ($i = 1, \ldots, n$) in a unit cube Y (e.g., $a^{ij} \in C^1$). This smoothness assumption is convenient but not crucial for our analysis here. We have analyzed MsFEM for the problems with discontinuous coefficients in [10]. In the following, we assume $\Omega = (0,1) \times (0,1) \subset \mathbf{R}^2$.

Variational problem of (2.1) is to seek $u \in H_0^1(\Omega)$ s.t.

$$a(u, v) = f(v), \quad \forall v \in H_0^1(\Omega), \tag{2.3}$$

where

$$a(u, v) = \int_\Omega a_\epsilon^{ij} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} dx, \quad f(v) = \int_\Omega f v dx. \tag{2.4}$$

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

A finite element method is obtained by restricting the weak formulation (2.3) to a finite dimensional subspace of $H_0^1(\Omega)$. For $0 < h \le 1$, let $\mathbf{K}^h$ be a partition of Ω of rectangles K with diameter less h, which is defined by an axi-parallel rectangular mesh. In each element $K \in \mathbf{K}^h$, we define a set of nodal basis $\{\phi_K^i\}$, $i = 1, \ldots, d$, 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

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

Let $x_j \in K$ ($j = 1, \ldots, d$) be the nodal points of K. As usual we require $\phi^i(x_j) = \delta_{ij}$. One needs to specify the boundary condition of ϕ^i for

the well-poseness of (2.5). We will assume that the base functions are linear on each side of on the boundary (unless otherwise stated). So we have:

$$V^h = \text{span}\{\phi_K^i;\ \ i = 1, \ldots, d,\ \ K \subset \mathbf{K}^h\} \subset H_0^1(\Omega).$$

In the following we study the approximate solution of (2.3) in V^h, i.e., $u^h \in V^h$ such that

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

3 Estimates for first order correctors

In this section we review the homogenization theory of (2.1) [6,2] and estimate the first order correctors, namely the difference between the solution and its H^1 approximater. The main difficulty in this estimate is to express the H^1 norm of the first order corrector through right-hand side of (2.1), i.e. $\|f\|_{0,\Omega}$. This is essential for the use of Aubin-Nitsche trick in the L_2 analysis of MsFEM. In Lemma 3.1 we prove such a result for smooth domains. However, for convex polygons we could not obtain the similar estimate. As shown later, this leads to a slightly overestimated L_2 norm error.

We consider the case with two scale cases:

$$\nabla_i a^{ij}\left(\frac{x}{\epsilon_1}, \frac{x}{\epsilon_2}\right)\nabla_j u_\epsilon = f \ \text{ in } \Omega, \qquad u^\epsilon = 0 \ \text{ on } \partial\Omega \tag{3.1}$$

where $f \in L_2(\Omega)$ and $\epsilon_1 \gg \epsilon_2$. For further convenience we take $y = \frac{x}{\epsilon_1}$ and $z = \frac{x}{\epsilon_2}$. Fixing $\frac{x}{\epsilon_1} = \lambda$ as a parameter we consider $a^{ij}(\frac{x}{\epsilon_1}, \frac{x}{\epsilon_2})$ as a family of functions $a^{ij}(\lambda, z)$ where λ is a parameter. By the assumption made in the previous section, $a(\lambda, z)$ is z periodic for any λ and $\alpha|\xi|^2 \le \xi_i a^{ij}(\lambda, z)\xi_j \le \beta|\xi|^2$ for all $\xi \in \mathbf{R}^2$ with $0 < \alpha < \beta < \infty$. Then the family of operators,

$$A_\lambda^\epsilon = \frac{\partial}{\partial x_i} a^{ij}\left(\lambda, \frac{x}{\epsilon_2}\right)\frac{\partial}{\partial x_j}, \tag{3.2}$$

can be homogenized by the standard homogenization rule, λ being a parameter. Furthermore we homogenize A_λ^ϵ with respect to ϵ_1. Thus

$$\frac{\partial}{\partial x_i} a^{ij}\left(\frac{x}{\epsilon_1}, \frac{x}{\epsilon_2}\right)\frac{\partial}{\partial x_j}, \tag{3.3}$$

is homogenized in two steps by the reiterated homogenization. More specifically, we define $\chi_\lambda^k(z)$ on $Z = (0,1) \times (0,1)$ as the periodic solution of

$$\frac{\partial}{\partial z_i} a^{ij}(\lambda, z)\frac{\partial(\chi_\lambda^k(z) + z_k)}{\partial z_j} = 0, \tag{3.4}$$

such that $\int_Z \chi_\lambda^k(z) = 0$. Then the homogenized operator for A_λ^ϵ is given by

$$A_\lambda = \frac{\partial}{\partial x_i} a_\lambda^{ij} \frac{\partial}{\partial x_j} \tag{3.5}$$

where

$$a_\lambda^{ij} = \frac{1}{|Z|} \int_Z a_{kl}(\lambda, z) \frac{\partial(\chi_\lambda^i + z_i)}{\partial z_k} \frac{\partial(\chi_\lambda^j + z_j)}{\partial z_l} dz. \tag{3.6}$$

Now taking into account that $\lambda = y = \frac{x}{\epsilon_1}$, A_λ can be homogenized as

$$A_0 = \frac{\partial}{\partial x_i} a_0^{ij} \frac{\partial}{\partial x_j}$$

where

$$a_0^{ij} = \frac{1}{|Y|} \int_Y a^{kl}(y) \frac{\partial(\chi^i + y_i)}{\partial y_k} \frac{\partial(\chi^j + y_j)}{\partial y_l} dy \tag{3.7}$$

and χ^j is the periodic solution of $A_\lambda(\chi^j - y_j) = 0$, such that $\int_Y \chi^j = 0$.

Remark 3.1. The reiterated homogenization procedure can be used for the n-scale case. Denoting H as the partial homogenization operator we have

$$a_0 = H \circ H \circ \cdots \circ H \circ a \left(\frac{x}{\epsilon_1}, \frac{x}{\epsilon_2}, \ldots, \frac{x}{\epsilon_n} \right).$$

Following the homogenization steps represented above we can approximate the solution of (3.1) as

$$u_\epsilon = u_0^\lambda + \epsilon_2 \chi_\lambda^m \left(\frac{x}{\epsilon_2} \right) \nabla_m u_0^\lambda + \theta_\epsilon \tag{3.8}$$

where u_0^λ is the solution of

$$A_\lambda u_0^\lambda = f \quad \text{in } \Omega, \qquad u_0^\lambda = 0 \quad \text{on } \partial\Omega. \tag{3.9}$$

Lemma 3.1. *Let θ_ϵ be the solution of (3.8), then under the assumption that $\partial\Omega$ is sufficiently smooth we have*

$$\|\theta_\epsilon\|_{H^1(\Omega)} \leq \left(C_1 \epsilon_2^{1/(2+\eta)} + C_2 \frac{\epsilon_2}{\epsilon_1} \right) \|f\|_{0,\Omega} \tag{3.10}$$

where $\eta > 0$ is an arbitrary positive number.

Proof. The proof of this lemma uses the notations introduced earlier. Denoting $u_\epsilon^1 = u_0^\lambda + \epsilon_2 \chi_\lambda^m(\frac{x}{\epsilon_2}) \nabla_m u_0^\lambda$ we can write $a_\epsilon^{ij} \nabla_j u_\epsilon^1$ as:

$$\begin{aligned} a_\epsilon^{ij} \nabla_j u_\epsilon^1 &= (a_\epsilon^{ij} + a_\epsilon^{ik} \nabla_k^z \chi_\lambda^j) \nabla_j u_0^\lambda + \epsilon_2 a_\epsilon^{ij} \chi_\lambda^m \nabla_j \nabla_m u_0^\lambda \\ &= a_\lambda^{ij} \nabla_j u_0^\lambda + g_i^j \nabla_j u_0^\lambda + \epsilon_2 a_\epsilon^{ij} \chi_\lambda^m \nabla_j \nabla_m u_0^\lambda \end{aligned} \tag{3.11}$$

where $g_i^j = a_\epsilon^{ij} + a_\epsilon^{ik} \nabla_k^z \chi_\lambda^j - a_\lambda^{ij}$ with ∇_k^z denoting $\partial / \partial z_k$. Using (3.4) and the fact that $\nabla_i a_\lambda^{ij} \nabla_j$ is a partially homogenized operator for (3.3) we have

$$\int_Z g_i^j dz = 0 \quad \text{and} \quad \frac{\partial}{\partial z_i} g_i^j = 0. \tag{3.12}$$

Thus, g_i^k is a periodic solenoidal vector with average zero. It can be expressed as [15]

$$g_i^k = \frac{\partial}{\partial z_j} \alpha_{ij}^k(y, z)$$

where $\alpha_{ij}^k = -\alpha_{ji}^k$ and $\alpha_{ij}^k \in (L_2(Y), H^1(Z))$. Using this representation we can write (3.11) as

$$\begin{aligned} a_\epsilon^{ij} \nabla_j u_\epsilon^1 = a_\lambda^{ij} \nabla_j u_0^\lambda + \epsilon_2 \frac{\partial}{\partial x_j} \left(\alpha_{ij}^k(y, z) \frac{\partial}{\partial x_k} u_0^\lambda \right) - \frac{\epsilon_2}{\epsilon_1} \frac{\partial}{\partial y_j} \left(\alpha_{ij}^k \frac{\partial}{\partial x_k} u_0^\lambda \right) \\ - \epsilon_2 \alpha_{ij}^k(y, z) \frac{\partial^2}{\partial x_j \partial x_k} u_0^\lambda + \epsilon_2 a_\epsilon^{ij} \chi_\lambda^m \nabla_j \nabla_m u_0^\lambda. \end{aligned} \tag{3.13}$$

Denoting the last two terms on the r.h.s. by r_ϵ^i, from (3.1), (3.8), and (3.9) we have

$$\begin{aligned} &\nabla_i a_\epsilon^{ij} \nabla_j \theta_\epsilon = -\text{div } \mathbf{r}_\epsilon \quad in\ \Omega, \\ &\theta_\epsilon = \epsilon_2 \chi_\lambda^m \left(\frac{x}{\epsilon_2} \right) \nabla_m u_0^\lambda \quad on\ \partial\Omega. \end{aligned} \tag{3.14}$$

Equation (3.14) along with the regularity properties for the solution of the elliptic PDE gives

$$\|\theta_\epsilon\|_{H^1(\Omega)} \le C \|\text{div } \mathbf{r}_\epsilon\|_{-1,\Omega} + C \|\epsilon_2 \chi_\lambda^m \left(\frac{x}{\epsilon_2} \right) \nabla_m u_0^\lambda\|_{H^{1/2}(\partial\Omega)}. \tag{3.15}$$

Furthermore using the fact that $\|\text{div } \mathbf{p}\|_{-1,\Omega} \le C \|p\|_{0,\Omega}$ for any $p \in L_2(\Omega)$ we have

$$\begin{aligned} \|\text{div } \mathbf{r}_\epsilon\|_{-1,\Omega} \le C \|\mathbf{r}_\epsilon\|_{0,\Omega} &\le C \left\| \epsilon_2 \alpha_{ij}^k(y, z) \frac{\partial^2}{\partial x_j \partial x_k} u_0^\lambda \right\|_{0,\Omega} \\ &\quad + C \|\epsilon_2 a_\epsilon^{ij} \nabla_j \nabla_m u_0^\lambda\|_{0,\Omega} \\ &\le C \epsilon_2 |u_0^\lambda|_{2,\Omega} \le C \frac{\epsilon_2}{\epsilon_1} \|f\|_{0,\Omega}. \end{aligned} \tag{3.16}$$

In the last step we have used

$$|u_0^\lambda|_{2,\Omega} \leq \frac{C}{\epsilon_1}\|f\|_{0,\Omega} \tag{3.17}$$

which can be shown following [16].

To estimate of the second term on the r.h.s. of (3.15) we use

$$\left\|\epsilon_2\chi_\lambda^m\left(\frac{x}{\epsilon_2}\right)\nabla_m u_0^\lambda\right\|_{H^{1/2}(\partial\Omega)} = \inf\|\phi\|_{1,\Omega}, \tag{3.18}$$

where the $\inf$ is over all ϕ satisfying $\phi = \epsilon_2\chi_\lambda^m(\frac{x}{\epsilon_2})\nabla_m u_0^\lambda$ on the boundary $\partial\Omega$. For the construction of the continuation of $\epsilon_2\chi_\lambda^m(\frac{x}{\epsilon_2})\nabla_m u_0^\lambda$ on to Ω we introduce a family of boundary functions τ^ϵ satisfying the following conditions [15].

1. $\tau^\epsilon \in C_0^\infty(\Omega)$, $0 \leq \tau^\epsilon \leq 1$, $\tau^\epsilon = 1$ outside the ϵ_2 neighborhood of $\partial\Omega$.
2. $\epsilon_2|\nabla\tau^\epsilon| \leq C$ in Ω, where the constant C does not depend on ϵ_i (i=1,2). Such functions can be constructed for any domain with Lipschitz boundary. Then we have

$$\begin{aligned}
&\left\|\epsilon_2\chi_\lambda^m\left(\frac{x}{\epsilon_2}\right)\nabla_m u_0^\lambda\right\|_{H^{1/2}(\partial\Omega)}\\
&\quad\leq \left\|(1-\tau^\epsilon)\,\epsilon_2\chi_\lambda^m\left(\frac{x}{\epsilon_2}\right)\nabla_m u_0^\lambda\right\|_{H^1(\Omega)}\\
&\quad\leq C\left\|(1-\tau^\epsilon)\,\epsilon_2\chi_\lambda^m\left(\frac{x}{\epsilon_2}\right)\nabla_m u_0^\lambda\right\|_{0,\Omega}\\
&\qquad+C\left\|(\nabla\tau^\epsilon)\,\epsilon_2\chi_\lambda^m\left(\frac{x}{\epsilon_2}\right)\nabla_m u_0^\lambda\right\|_{0,\Omega}\\
&\qquad+C\left\|(1-\tau^\epsilon)\,\epsilon_2\nabla\chi_\lambda^m\left(\frac{x}{\epsilon_2}\right)\nabla_m u_0^\lambda\right\|_{0,\Omega}\\
&\qquad+C\left\|(1-\tau^\epsilon)\,\epsilon_2\chi_\lambda^m\left(\frac{x}{\epsilon_2}\right)\nabla\nabla_m u_0^\lambda\right\|_{0,\Omega}\\
&\quad\leq C\epsilon_2\|u_0^\lambda\|_{1,\Omega}+C\left\|(\nabla\tau^\epsilon)\,\epsilon_2\chi_\lambda^m\left(\frac{x}{\epsilon_2}\right)\nabla_m u_0^\lambda\right\|_{0,\Omega}\\
&\qquad+C\left\|(1-\tau^\epsilon)\,\epsilon_2\nabla\chi_\lambda^m\left(\frac{x}{\epsilon_2}\right)\nabla_m u_0^\lambda\right\|_{0,\Omega}+C\epsilon_2\|u_0^\lambda\|_{2,\Omega}\\
&\quad\leq C\epsilon_2\|u_0^\lambda\|_{1,\Omega}+\|\epsilon_2\nabla\tau^\epsilon\|_{L^{2+\eta}(\Omega)}|u_0^\lambda|_{W^{1,2+4/\eta}(\Omega)}\\
&\qquad+C\|1-\tau^\epsilon\|_{L^{2+\eta}(\Omega)}|u_0^\lambda|_{W^{1,2+4/\eta}(\Omega)}+C\epsilon_2|u_0|_{2,\Omega}
\end{aligned} \tag{3.19}$$

where $\eta > 0$. Furthermore using the inequalities

$$\begin{aligned} \|u_0^\lambda\|_{2,\Omega} &\le \frac{C}{\epsilon_1}\|f\|_{0,\Omega}, \\ \|\epsilon_2 \nabla \tau^\epsilon\|_{L^{2+\eta}(\Omega)} &\le C\epsilon_2^{1/(2+\eta)} \\ \|1-\tau^\epsilon\|_{L^{2+\eta}(\Omega)} &\le C\epsilon_2^{1/(2+\eta)}, \\ \|\nabla u_0^\lambda\|_{L^{2+4/\eta}(\Omega)} &\le C\|f\|_{0,\Omega} \end{aligned} \tag{3.20}$$

we conclude (3.10). Note that the last inequality follows from Theorem 4(i) of [3]. □

Remark 3.2. For convex polygon domains we assume that u_ϵ (or any partially homogenized part of it) is in $C^2(\Omega)$ for the fixed ϵ and

$$\|u_\epsilon\|_{C^1(\Omega)} \le C, \quad \|u_\epsilon\|_{C^2(\Omega)} \le \frac{C}{\epsilon_n} \tag{3.21}$$

where ϵ_n is the smallest scale in (2.2) (or in the partially homogenized problem). This assumptions is true for fixed ϵ_i ($i = 1, \dots, n$) under the compatibility conditions stated in [4]. Under the assumptions (3.21), for convex polygonal domains we have

$$\|\theta_\epsilon\|_{H^1(\Omega)} \le C\frac{\epsilon_2}{\epsilon_1} + C_1\sqrt{\epsilon}_2. \tag{3.22}$$

In the analysis of MsFEM the assumptions (3.21) will be used. In the last section of the paper we present the error analysis for MsFEM without using the assumptions (3.21).

Remark 3.3. For further convenience the quantities which depend on $\frac{\epsilon_i}{\epsilon_{i-1}}$ with $2 \le i \le n$ and ϵ_i, $i \ge 1$, we denote $O(\epsilon)$. It indicates that $O(\epsilon)$ is an asymptotically small quantity independent of the mesh size.

The above procedure can be applied for homogenization of many scale problems. Instead of (3.8), using the reiterated homogenization we find

$$u_\epsilon = u_0\left(x, \frac{x}{\epsilon_1}, \dots, \frac{x}{\epsilon_m}\right) + u_\epsilon^1 + \theta_\epsilon, \tag{3.23}$$

where u_0 is the partially homogenized part of u_ϵ over the scales $\epsilon_{m+1}, \dots, \epsilon_n$, u_ϵ^1 is given by

(3.24)

$$u_\epsilon^1 = \sum_{i=1}^{n-m+1} \left(\sum_{m_k, m<m_1\le\cdots\le m_i\le n} \epsilon_{m_i} \chi^{k_i}\left(\frac{x}{\epsilon_{m_i}}\right)\right.$$

$$\times \left(\epsilon_{m_{i-1}} \nabla_{k_i} \chi^{k_{i-1}} \left(\frac{x}{\epsilon_{m_{i-1}}}\right)\right) \cdots \left(\epsilon_{m_1} \nabla_{k_2} \chi^{k_1} \left(\frac{x}{\epsilon_{m_1}}\right)\right) \nabla_{k_1} u_0 \Bigg),$$

and θ_ϵ is the remaining part. The details of the derivation of (3.24) is omitted here. Furthermore, it can be shown that for smooth domains Ω

$$\|\theta_\epsilon\|_{H^1(\Omega)} \leq O(\epsilon)\|f\|_{0,\Omega}$$

with

$$O(\epsilon) = C_1 \epsilon_{m+1}^{1/(2+\eta)} + C_2 \max_{i \geq m+1} \frac{\epsilon_i}{\epsilon_{i-1}} \quad (\eta > 0).$$

For convex polygonal domains, under the assumptions (3.21) the estimate for θ_ϵ is similar to (3.22),

$$\|\theta_\epsilon\|_{H^1(\Omega)} \leq C_1 \epsilon_{m+1}^{1/2} + C_2 \max_{i \geq m+1} \frac{\epsilon_i}{\epsilon_{i-1}}. \tag{3.25}$$

4 H^1 estimates

For simplicity, we first present the estimates for problems with two scales in detail. The multiscale case can be analyzed in a similar manner.

4.1 H^1 estimates for two scale case

In this section we analyze the MsFEM for three different cases: (1) $\epsilon_1 \gg h \gg \epsilon_2$, (2) $h \gg \epsilon_1 \gg \epsilon_2$, and (3) $\epsilon_1 \gg \epsilon_2 \gg h$. As in the standard FEM we have Cea's Lemma [13]:

Lemma 4.1. *Let u and u^h be the solutions of (2.1) and (2.6) respectively. Then*

$$\|u - u^h\|_{1,\Omega} \leq C \frac{\beta}{\alpha} \|u - v\|_{1,\Omega}, \ \ \forall v \in V^h \tag{4.1}$$

Next we formulate two lemmas which will be used in the analysis of MsFEM. In the formulations of these lemmas the domain K is a finite element (e.g., rectangular), $K \in \mathbf{K}^h$, whose diameter is of order h.

Lemma 4.2. *For any $f \in H^1(K)$ we have*

$$\|f\|_{L_2(\partial K)} \leq C h^{1/2} |f|_{H^1(K)} + C h^{-1/2} \|f\|_{L_2(K)}. \tag{4.2}$$

This lemma can be derived from the standard trace inequality [1] using the scaling argument. We omit the proof of the lemma.

The next lemma is an interpolation inequality (see e.g. [8]):

Lemma 4.3. *If $u \in H_0^1(K) \bigcap H^2(K)$, then*

$$\|\nabla u\|_{L_2(K)} \leq C h \|\Delta u\|_{L_2(K)}.$$

4.1.1 Case 1: $\epsilon_1 \gg h \gg \epsilon_2$

Theorem 4.4. *Let* u_ϵ *and* u_ϵ^h *be the solution and MsFEM solution of (2.1) respectively. Then*

$$\|u_\epsilon - u_\epsilon^h\|_{1,\Omega} \leq C_1 \frac{h}{\epsilon_1} + C_2 \left(\frac{\epsilon_2}{h}\right)^{1/2}. \tag{4.3}$$

Proof. Define $v_\epsilon^h \in V^h$ such that in each $K \in \mathbf{K}^h$

$$v_\epsilon^h(x) = (I_k u)(x) = \sum_{j=1}^{d} \alpha_j \phi^j(x) \tag{4.4}$$

with $\alpha_j = u_0(x_j)$ where u_0 is the partially homogenized part of u_ϵ over the scale ϵ_2 and x_j are the nodal points of K. Then in any element $K \in \mathbf{K}^h$ we have

$$\nabla_i a_\epsilon^{ij} \nabla_j (u_\epsilon - v_\epsilon^h) = f \quad \text{in } K, \tag{4.5}$$

where v_ϵ^h on ∂K is a piecewise linear function whose values at the nodal points are $u_0(x_j)$.

We divide the solution of (4.5) into two parts: $(u_\epsilon - v_\epsilon^h) = (u_\epsilon - v_\epsilon^h)_1 + (u_\epsilon - v_\epsilon^h)_2$, where $(u_\epsilon - v_\epsilon^h)_1$ and $(u_\epsilon - v_\epsilon^h)_2$ satisfy

$$\begin{aligned} \nabla_i a_{ij}^\epsilon \nabla_j (u_\epsilon - v_\epsilon^h)_1 &= f \quad \text{in } K, \\ (u_\epsilon - v_\epsilon^h)_1 &= 0 \quad \text{on } \partial K; \end{aligned} \tag{4.6}$$

$$\begin{aligned} \nabla_i a_\epsilon^{ij} \nabla_j (u_\epsilon - v_\epsilon^h)_2 &= 0 \ \text{ in } K, \\ (u_\epsilon - v_\epsilon^h)_1 &= (u_\epsilon - v_\epsilon^h) \ \text{ on } \partial K; \end{aligned} \tag{4.7}$$

respectively. $(u_\epsilon - v_\epsilon^h)_1$ can be estimated from (4.6) using the Lemma 4.3,

$$\|(u_\epsilon - v_\epsilon^h)_1\|_{H^1(K)} \leq Ch \|f\|_{L_2(K)}. \tag{4.8}$$

As for $(u_\epsilon - v_\epsilon^h)_2$, using the expansion over the scale ϵ_2, $u_\epsilon = u_0 + \epsilon_2 \chi_\lambda^m \nabla_m u_0 + \theta_\epsilon$, and (4.7) we have

$$\begin{aligned} &\left\| \left(u_\epsilon - v_\epsilon^h\right)_2 \right\|_{H^1(K)} \\ &\quad \leq \|u_\epsilon - v_\epsilon^h\|_{H^{1/2}(\partial K)} \\ &\quad \leq C \left\| u_0 + \epsilon_2 \chi_\lambda^m \left(\frac{x}{\epsilon_2}\right) \nabla_m u_0 - v_\epsilon^h \right\|_{H^{1/2}(\partial K)} + C \|\theta_\epsilon\|_{H^{1/2}(\partial K)} \\ &\quad \leq C \|u_0 - v_\epsilon^h\|_{H^{1/2}(\partial K)} + C \left\| \epsilon_2 \chi_\lambda^m \left(\frac{x}{\epsilon_2}\right) \nabla_m u_0 \right\|_{H^{1/2}(\partial K)} \end{aligned}$$

(4.9) $+C\|\theta_\epsilon\|_{H^1(K)}.$

Combination of (4.8) and (4.9) gives

$$\|u_\epsilon - v_\epsilon^h\|_{H^1(K)} \le C_1 h\|f\|_{0,K} + C\|u_0 - v_\epsilon^h\|_{H^{1/2}(\partial K)}$$
$$+C\|\epsilon_2\chi^m(\frac{x}{\epsilon_2})\nabla_m u_0\|_{H^{1/2}(\partial K)} + C\|\theta_\epsilon\|_{H^1(K)}. \tag{4.10}$$

The third term on the r.h.s. of this inequality can be estimated as

(4.11)
$$\begin{aligned}
&\left\|\epsilon_2\chi^m\left(\frac{x}{\epsilon_2}\right)\nabla_m u_0\right\|_{H^{1/2}(\partial K)} \\
&\le \left\|\epsilon_2\chi^m\left(\frac{x}{\epsilon_2}\right)\nabla_m u_0\right\|^{1/2}_{L_2(\partial K)} \left\|\epsilon_2\chi^m\left(\frac{x}{\epsilon_2}\right)\nabla_m u_0\right\|^{1/2}_{H^1(\partial K)} \\
&\le C\epsilon_2^{1/2}\|\nabla u_0\|^{1/2}_{L_2(\partial K)}\left(\|\nabla u_0\|^{1/2}_{L_2(\partial K)} + \epsilon_2^{1/2}|\nabla u_0|^{1/2}_{H^1(\partial K)}\right) \\
&= C\epsilon_2^{1/2}\|\nabla u_0\|_{L_2(\partial K)} + C\epsilon_2\|\nabla u_0\|^{1/2}_{L_2(\partial K)}|\nabla u_0|^{1/2}_{H^1(\partial K)} \\
&\le C\epsilon_2^{1/2}\|\nabla u_0\|_{L_2(K)}h^{-1/2} + C\left(\epsilon_2 h\right)^{1/2}|\nabla u_0|_{H^1(K)} + C\frac{\epsilon_2}{\epsilon_1^{1/2}}h^{1/2}.
\end{aligned}$$

In the last step we have used Lemma 4.2 for the first term and the assumption that $\|u_0\|_{C^2(\Omega)} \le C/\epsilon_1$ (see (3.21)) for the second term.

For the second term on the r.h.s. of (4.10) we can write

$$\|u_0 - v_\epsilon^h\|_{H^{1/2}(\partial K)} = \inf_{\phi|_{\partial K} = u_0 - v_\epsilon^h} \|\phi\|_{H^1(K)}. \tag{4.12}$$

Since v_ϵ^h is the linear function on ∂K we can extend it as a bilinear function $\tilde{v}_\epsilon^h$ onto K. Noticing the fact that because of (4.4) u_0 and $\tilde{v}_\epsilon^h$ coincide at the nodal points we have

$$\|u_0 - v_\epsilon^h\|_{H^{1/2}(\partial K)} \le \|u_0 - \tilde{v}_\epsilon^h\|_{H^1(K)} \le Ch|u_0|_{2,K}. \tag{4.13}$$

Combining (4.11) and (4.13) we have

$$\|u_\epsilon - v_\epsilon^h\|_{1,K} \le Ch\|f\|_{0,K} + Ch|u_0|_{2,K} + C\epsilon_2^{1/2}h^{-1/2}\|\nabla u_0\|_{L_2(K)}$$
$$+C\epsilon_2^{1/2}h^{1/2}|\nabla u_0|_{H^1(K)} + C\frac{\epsilon_2}{\epsilon_1^{1/2}}h^{1/2} + C\|\theta_\epsilon\|_{H^1(K)}. \tag{4.14}$$

Summing (4.14) over all $K \subset \mathbf{K}$ and using Cea's lemma we get

$$\|u_\epsilon - u_\epsilon^h\|_{1,\Omega} \le Ch\|f\|_{0,\Omega} + Ch|u_0|_{2,\Omega} + C\epsilon_2^{1/2}h^{-1/2}\|\nabla u_0\|_{L_2(\Omega)}$$
$$+C\epsilon_2^{1/2}h^{1/2}|\nabla u_0|_{H^1(\Omega)} + C\frac{\epsilon_2}{\epsilon_1^{1/2}h^{1/2}} + C_4\|\theta_\epsilon\|_{H^1(\Omega)}$$

$$\leq \left(Ch + C\frac{h}{\epsilon_1} + C_2\sqrt{\frac{\epsilon_2}{h}} + C_3\frac{\sqrt{\epsilon_2 h}}{\epsilon_1} \right) \|f\|_{0,\Omega}$$

$$+C\frac{\epsilon_2}{\sqrt{\epsilon_1 h}} + C\frac{\epsilon_2}{\epsilon_1} + C\sqrt{\epsilon_2}. \tag{4.15}$$

In the last step we have used (3.17) and the estimate for $\|\theta_\epsilon\|_{H^1(\Omega)}$ (see (3.22)). Note in (4.15) $\sqrt{\epsilon_2 h}/\epsilon_1$, $\epsilon_2/\sqrt{\epsilon_1 h}$, and $\sqrt{\epsilon_2}$ are much smaller than $\sqrt{\epsilon_2/h}$, and $\epsilon_2/\epsilon_1 \ll h/\epsilon_1$. Thus, dropping the lower order terms and using Cea's lemma we get (4.3). □

4.1.2 Case 2: $h \gg \epsilon_1 \gg \epsilon_2$

Theorem 4.5. *Let u_ϵ and u_ϵ^h be the solution and MsFEM solution of (2.1) respectively. Then*

$$\|u_\epsilon - u_\epsilon^h\|_{1,\Omega} \leq C_1 h + C_2 \left(\frac{\epsilon_1}{h}\right)^{1/2} + C_3\frac{\epsilon_2}{\epsilon_1}. \tag{4.16}$$

Proof. We define v_ϵ^h as in (4.4) with u_0 denoting the fully homogenized solution over the scales ϵ_1 and ϵ_2. Using the partition $u_\epsilon - v_\epsilon^h = (u_\epsilon - v_\epsilon^h)_1 + (u_\epsilon - v_\epsilon^h)_2$ as in case 1, and taking into account the inequality (4.8) for $(u_\epsilon - v_\epsilon^h)_1$ and the expansion of u_ϵ over the scales ϵ_1 and ϵ_2,

$$u_\epsilon = u_0 + \epsilon_1\chi^m\left(\frac{x}{\epsilon_1}\right)\nabla_m u_0 + \epsilon_2\chi_\lambda^m\left(\frac{x}{\epsilon_2}\right)\nabla_m u_0$$
$$+\epsilon_2\chi_\lambda^m\left(\frac{x}{\epsilon_2}\right)\left(\epsilon_1\nabla_m\chi^l\left(\frac{x}{\epsilon_1}\right)\nabla_l u_0\right) + \theta',$$

we have

$$\begin{aligned}
&\|u_\epsilon - v_\epsilon^h\|_{H^1(K)} \\
&\leq \left\|\left(u_\epsilon - v_\epsilon^h\right)_1\right\|_{H^1(K)} + \left\|\left(u_\epsilon - v_\epsilon^h\right)_2\right\|_{H^1(K)} \\
&\leq Ch\|f\|_{0,K} + C\|u_\epsilon - v_\epsilon^h\|_{H^{1/2}(\partial K)} \\
&\leq C\|f\|_{0,\Omega} + C\Bigg(\|u_0 - v_\epsilon^h\|_{H^{1/2}(\partial K)} + \left\|\epsilon_1\chi^m\left(\frac{x}{\epsilon_1}\right)\nabla_m u_0\right\|_{H^{1/2}(\partial K)} \\
&\quad + \left\|\epsilon_2\chi_\lambda^m\left(\frac{x}{\epsilon_2}\right)\nabla_m u_0\right\|_{H^{1/2}(\partial K)} \\
&\quad + \left\|\epsilon_2\chi_\lambda^m\left(\frac{x}{\epsilon_2}\right)\left(\epsilon_1\nabla_m\chi^l\left(\frac{x}{\epsilon_1}\right)\nabla_l u_0\right)\right\|_{H^{1/2}(\partial K)} + \|\theta'\|_{H^1(K)}\Bigg).
\end{aligned} \tag{4.17}$$

Treating the third, fourth, and fifth terms on the r.h.s of (4.17) similarly as in (4.11) (i.e., applying the interpolation inequality and Lemma 4.2), we get

$$\begin{aligned}
&\|u_\epsilon - v_\epsilon^h\|_{H^1(K)} \\
&\quad \le Ch\|f\|_{0,K} + C\|u_\epsilon - v_\epsilon^h\|_{H^{1/2}(\partial K)} + C\epsilon_1 h^{-1/2}\|\nabla u_0\|_{0,K} \\
&\qquad + C\epsilon_1^{1/2}h^{1/2}|\nabla u_0|_{1,K} + C\epsilon_1 h^{1/2} + C\sqrt{\frac{\epsilon_2}{h}}\|\nabla u_0\|_{0,K} \\
&\qquad + C\epsilon_2^{1/2}h^{1/2}|\nabla u_0|_{1,K} + C\epsilon_2 h^{1/2} + C\|\theta'\|_{1,K}. \qquad (4.18)
\end{aligned}$$

For the second term on the r.h.s. of (4.18), since v_ϵ^h is a linear function on ∂K we can continue it as a bilinear function $\tilde{v}_\epsilon^h$ onto K. Noticing that the values of u_0 and $\tilde{v}_\epsilon^h$ coincide at the nodal points we have

$$\|u_0 - v_\epsilon^h\|_{H^{1/2}(\partial K)} \le \|u_0 - \tilde{v}_\epsilon^h\|_{H^1(K)} \le Ch\|u_0\|_{2,K} \qquad (4.19)$$

Finally, summing (4.18) over $K \in \mathbf{K}^h$ and using (3.21), (4.19), and the fact $|u_0|_{2,\Omega} \le C\|f\|_{0,\Omega}$ we get

$$\begin{aligned}
\|u_\epsilon - v_\epsilon^h\|_{1,\Omega} &\le C\left(h + \sqrt{\frac{\epsilon_1}{h}} + (\epsilon_1 h)^{1/2} + \sqrt{\frac{\epsilon_2}{h}} + (\epsilon_2 h)^{1/2}\right)\|f\|_{0,\Omega} \\
&\quad + C\epsilon_1 h^{-1/2} + C\epsilon_2 h^{-1/2} + C\frac{\epsilon_2}{\epsilon_1} + C\sqrt{\epsilon_1}. \qquad (4.20)
\end{aligned}$$

Here we have used the estimate (3.25) for $\|\theta'\|_{1,\Omega}$. In (4.20) we may neglect $\sqrt{\epsilon_1 h}$, $\epsilon_1/\sqrt{h}$, and $\sqrt{\epsilon_2}$ since they are lower order comparing with $\sqrt{\epsilon_1/h}$. Thus, (4.16) follows from (4.20) and Cea's lemma. □

4.1.3 Case 3: $\epsilon_1 \gg \epsilon_2 \gg h$

Theorem 4.6. *Let u_ϵ and u_ϵ^h be the solution and MsFEM solution of (2.1) respectively. Then*

$$\|u_\epsilon - u_\epsilon^h\|_{1,\Omega} \le C\frac{h}{\epsilon_2}\|f\|_{0,\Omega} \qquad (4.21)$$

Proof. Define v_ϵ^h as in (4.4) and take α_j to be the nodal value of u_ϵ at x_j. Using the partition of $u_\epsilon - v_\epsilon^h$ we obtain

$$\|u_\epsilon - v_\epsilon^h\|_{H^1(K)} \le Ch\|f\|_{0,K} + \|u_\epsilon - v_\epsilon^h\|_{H^{1/2}(\partial K)}. \qquad (4.22)$$

The continuation of v_ϵ^h from ∂K onto K as a bilinear interpolant of u_ϵ in K, $\tilde{v}_\epsilon^h$, yields

$$\|u_\epsilon - v_\epsilon^h\|_{H^{1/2}(\partial K)} \le \|u_\epsilon - \tilde{v}_\epsilon^h\|_{H^1(K)} \le Ch\|u_\epsilon\|_{2,K}. \qquad (4.23)$$

Summing (4.22) over all K, using (4.23), (3.17), and Cea's lemma we get (4.21). □

4.2 H^1 error estimates for many scale case

Assuming the order of h is between the scales ϵ_{m+1} and ϵ_m for $m \in [1, n]$, i.e.,

$$\epsilon_1 \gg \cdots \gg \epsilon_m \gg h \gg \epsilon_{m+1} \cdots \gg \epsilon_n$$

we have the following result.

Theorem 4.7. *Let u_ϵ and u_ϵ^h be the solution and MsFEM solution of (2.1) respectively. Then*

$$\|u_\epsilon - u_\epsilon^h\|_{1,\Omega} \leq C_1 \frac{h}{\epsilon_m} + C_2 \left(\frac{\epsilon_{m+1}}{h}\right)^{1/2} + C_3 \max_{i \geq m+1} \frac{\epsilon_i}{\epsilon_{i-1}}. \tag{4.24}$$

Proof. The proof of this theorem is similar to the proof of the Theorem 4.4. First we define v_ϵ^h as in (4.4) with $\alpha_j = u_0(x_j)$ where u_0 is the partially homogenized part of u_ϵ over the scales $\epsilon_{m+1}, \ldots, \epsilon_n$. In this case the expansion of u_ϵ is given by (3.23),

$$u_\epsilon = u_0(x, x/\epsilon_1, \cdots, x/\epsilon_m) + u_\epsilon^1 + \theta,$$

where u_ϵ^1 is defined by (3.24). Furthermore for $\|u_\epsilon - v_\epsilon^h\|_{H^1(K)}$ we have

(4.25)

$$\begin{aligned}
&\|u_\epsilon - v_\epsilon^h\|_{H^1(K)} \\
&\quad \leq Ch\|f\|_{0,K} + C\|u_\epsilon - v_\epsilon^h\|_{H^{1/2}(\partial K)} \\
&\quad \leq Ch\|f\|_{0,K} + C\|u_0 - v_\epsilon^h\|_{H^{1/2}(\partial K)} + C\|u_\epsilon^1\|_{H^{1/2}(\partial K)} + C\|\theta\|_{1,K} \\
&\quad \leq Ch\|f\|_{0,K} + Ch|u_0|_{2,K} + C\|u_\epsilon^1\|_{H^{1/2}(\partial K)} + C\|\theta\|_{1,K}.
\end{aligned}$$

In the last step we extended v_ϵ^h on to K as the bilinear interpolant of u_0. This gives us the following estimate

$$\|u_0 - v_\epsilon^h\|_{H^{1/2}(\partial K)} \leq Ch|u_0|_{2,K}. \tag{4.26}$$

Next to estimate the third term on the r.h.s. of (4.25) we use the following inequality

$$|u_\epsilon^1|_{1,\partial K} \leq C\|\nabla u_0\|_{0,\partial K} + C\epsilon_{m+1}|\nabla u_0|_{1,\partial K}. \tag{4.27}$$

In fact, by (3.24) u_ϵ^1 can be written in the form of $C_\epsilon(x) \cdot \nabla u_0$, where $C_\epsilon(x)$ contains the linear combinations of products of $\chi_k(\frac{x}{\epsilon_k})$ $(k = m+1, \ldots, n)$ and their gradients. Furthermore, it can be verified that

$$\max_x |\nabla C_\epsilon(x)| \leq C + \max_{i \geq m+1} (\epsilon_i/\epsilon_{i-1}) \leq C \quad \text{and}$$

$$\max_x |C_\epsilon(x)| \le C\epsilon_{m+1},$$

and hence ∇u_ϵ^1 is bounded. Therefore,

$$\begin{aligned}\|\nabla u_\epsilon^1\|_{0,\partial K} &\le \|\nabla C_\epsilon(x)\cdot\nabla u_0\|_{0,\partial K} + \|C_\epsilon(x)\cdot\nabla^2 u_0\|_{0,\partial K}\\ &\le \max_x |\nabla C_\epsilon(x)|\|\nabla u_0\|_{0,\partial K} + \max_x |C_\epsilon(x)|\|\nabla^2 u_0\|_{0,\partial K}\\ &\le C\|\nabla u_0\|_{0,\partial K} + C\epsilon_{m+1}|\nabla u_0|_{1,\partial K}.\end{aligned}$$

From (4.27), the interpolation inequality, the trace inequality (4.2), and (3.21) it follows that

$$\begin{aligned}&\|u_\epsilon^1\|_{H^{1/2}(\partial K)}\\ &\quad\le \|u_\epsilon^1\|_{L_2(\partial K)}^{1/2}\|u_\epsilon^1\|_{H^1(\partial K)}^{1/2}\\ &\quad\le C\epsilon_{m+1}^{1/2}\|\nabla u_0\|_{L_2(\partial K)}^{1/2}\left(\|\nabla u_0\|_{L_2(\partial K)}^{1/2} + \epsilon_{m+1}^{1/2}|\nabla u_0|_{H^1(\partial K)}\right)\\ &\quad\le C\epsilon_{m+1}^{1/2}\|\nabla u_0\|_{L_2(\partial K)} + C\epsilon_{m+1}{\epsilon_m}^{-1/2}h^{1/2}\\ &\quad\le C\epsilon_{m+1}^{1/2}h^{-1/2}\|\nabla u_0\|_{0,K} + C\epsilon_{m+1}^{1/2}h^{1/2}|\nabla u_0|_{1,K}\\ &\quad\quad + C\epsilon_{m+1}h^{1/2}\epsilon_m^{-1/2}. \qquad (4.28)\end{aligned}$$

Substituting (4.28) into (4.25) and then summing (4.25) over all $K \in \mathbf{K}^h$, we have

$$\begin{aligned}\|u_\epsilon - v_\epsilon^h\|_{1,\Omega} &\le C\left(h + \sqrt{\frac{\epsilon_{m+1}}{h}} + \frac{h}{\epsilon_m} + \frac{\epsilon_{m+1}^{1/2}h^{1/2}}{\epsilon_m}\right)\|f\|_{0,\Omega}\\ &\quad + C\frac{\epsilon_{m+1}}{\sqrt{\epsilon_m h}} + C\|\theta_\epsilon\|_{1,\Omega}\\ &\le C\left(h + \sqrt{\frac{\epsilon_{m+1}}{h}} + \frac{h}{\epsilon_m} + \frac{\epsilon_{m+1}^{1/2}h^{1/2}}{\epsilon_m}\right)\|f\|_{0,\Omega}\\ &\quad + C\frac{\epsilon_{m+1}}{\sqrt{\epsilon_m h}} + C\max_{i\ge m+1}\frac{\epsilon_i}{\epsilon_{i-1}}. \qquad (4.29)\end{aligned}$$

The estimate (4.24) follows from (4.29), (3.25) (estimate of $\|\theta_\epsilon\|_{1,\Omega}$), and Cea's lemma. The lower order terms with h ($\ll h/\epsilon_m$) and with $\sqrt{\epsilon_{m+1}h}/\epsilon_m$, $\epsilon_{m+1}/\sqrt{\epsilon_m h}$, and $\sqrt{\epsilon_{m+1}}$ (which are asymptotically smaller than $\sqrt{\epsilon_{m+1}/h}$) are neglected. □

Remark 4.1. The estimate (4.24) shows that the error becomes larger as h approaches either ϵ_m or ϵ_{m+1}. This is the resonance phenomenon mentioned in the introduction. Furthermore, we see from (4.29) that both resonance errors dominate when $h \sim \epsilon_m$ or $h \sim \epsilon_{m+1}$.

5 L_2 estimates

In this section we derive the L_2 estimates using the H^1 estimates along with the duality argument. Because some terms in the H^1 estimate (4.29) can not be expressed through $\|f\|_{0,\Omega}$, the estimates we obtain from the duality argument is not optimum when h is not comparable with the physical scales. The estimates do capture the correct order of the resonance error, which are the leading order error when h is comparable with the physical scales. Employing the method introduced in [13], which compares the discrete values of the solution and the numerical solution, we can show that the order of the method is not affected by those small terms not expressed through $\|f\|_{0,\Omega}$. The present approach, however, is more concise to present. We use the following abstract lemma.

Lemma 5.1. *Let u and u^h be the solutions of (2.3) and (2.6) respectively with V^h consisting of conforming base functions. If*

$$\|u - u^h\|_{1,\Omega} \leq C_1\gamma\|f\|_{0,\Omega} + C_2\delta \tag{5.1}$$

where γ and δ are small positive quantities, $C_1 \geq 0$, $C_2 \geq 0$, then

$$\|u - u^h\|_{0,\Omega} \leq C_3\gamma^2\|f\|_{0,\Omega} + C_4\sqrt{\gamma\delta\|f\|_{0,\Omega}} + C_5\delta$$

Proof. We use Aubin-Nitshe trick as follows: Let w be the solution of (2.3) with $f = u - u^h$, i.e. $w \in H_0^1(\Omega)$ satisfies

$$a(w, v) = (u - u^h, v), \quad \forall v \in H_0^1(\Omega). \tag{5.2}$$

Let $w_h \in V^h$ be the interpolant of w. Then (5.1) implies

$$\|w - w^h\|_{1,\Omega} \leq C_1\gamma\|u - u^h\|_{0,\Omega} + C_2\delta. \tag{5.3}$$

Choosing $v = u - u^h$ in (5.2) we have

$$\begin{aligned}
\|u - u^h\|_{0,\Omega}^2 &= a(u - u^h, w) \\
&= a(u - u^h, w - w_h) \\
&\leq \beta\|u - u^u\|_{1,\Omega}\|w - w_I\|_{1,\Omega} \\
&\leq \beta(C_1\gamma\|f\|_{0,\Omega} + C_2\delta)(C_1\gamma\|u - u^h\|_{0,\Omega} + C_2\delta) \\
&\leq \|u - u^h\|_{0,\Omega}(C_1\gamma^2\|f\|_{0,\Omega} + C_2\delta\gamma) + C\delta\gamma\|f\|_{0,\Omega} \\
&\quad + C_1\delta^2.
\end{aligned} \tag{5.4}$$

Therefore,

$$\|u - u^h\|_{0,\Omega} \leq |C_1\gamma^2\|f\|_{0,\Omega} + C_1\delta\gamma|$$

$$
\begin{aligned}
&+\sqrt{(C_1\gamma^2\|f\|_{0,\Omega}+C_1\delta\gamma)^2+4(C\delta\gamma\|f\|_{0,\Omega}+C_1\delta^2)}\\
&\leq C\gamma^2\|f\|_{0,\Omega}+C_1\delta\gamma+C\gamma^2\|f\|_{0,\Omega}+C_1\delta\gamma\\
&\quad +C_2\sqrt{\delta\gamma\|f\|_{0,\Omega}}+C_3\delta\\
&\leq C\gamma^2\|f\|_{0,\Omega}+C_1\sqrt{\gamma\delta\|f\|_{0,\Omega}}+C_3\delta. \qquad (5.5)
\end{aligned}
$$

□

In our H^1 estimates γ contains the resonance errors which may become $O(1)$ depending on the mesh size h. On the other hand, δ is an asymptotically small quantity independent of the mesh size. In particular, $\gamma = C(h + \epsilon_{m+1}^{1/2}/h^{1/2} + h/\epsilon_m + \epsilon_{m+1}^{1/2}h^{1/2}/\epsilon_m)$ and $\delta = C\epsilon_{m+1}\epsilon_m^{-1/2}h^{-1/2} + O(\epsilon)$ for our problem (see (4.29)). Note that when h becomes comparable with any physical scale γ is of order one while δ remains asymptotically small. Consequently, the term $\sqrt{\gamma\delta}$ is an asymptotically small quantity which does not resonate at any scale of the problem, i.e. $\sqrt{\gamma\delta} = O(\epsilon)$ (see Remark 3.3). Thus γ^2 is the dominating resonance error when h becomes comparable with the physical scales. It is worth to note that some terms in γ and δ (see(4.29)) in our problem do not change when we decrease the mesh size and the physical scales at the same time but remain asymptotically negligible. These terms can be called lower order resonance terms.

Applying Lemma 5.1 to the cases analyzed in Sect. 4 we have the following resonance errors:

Theorem 5.2. *Let u_ϵ and u_ϵ^h be the solution and MsFEM solution of (2.1) respectively. Then we have the following resonance errors in the L_2 error estimates:*

1. *Case $\epsilon_1 \gg h \gg \epsilon_2$.*

$$\|u_\epsilon - u_\epsilon^h\|_{0,\Omega} \leq C_1\left(\frac{h}{\epsilon_1}\right)^2 + C_2\frac{\epsilon_2}{h} \qquad (5.6)$$

2. *Case $h \gg \epsilon_1 \gg \epsilon_2$.*

$$\|u_\epsilon - u_\epsilon^h\|_{0,\Omega} \leq C_1h^2 + C_2\frac{\epsilon_1}{h} \qquad (5.7)$$

3. *Case $\epsilon_1 \gg \epsilon_2 \gg h$.*

$$\|u_\epsilon - u_\epsilon^h\|_{0,\Omega} \leq C\left(\frac{\epsilon_2}{h}\right)^2 \qquad (5.8)$$

4. *Case $\epsilon_1 \gg \cdots \gg \epsilon_m \gg h \gg \epsilon_{m+1} \cdots \gg \epsilon_n$.*

$$\|u_\epsilon - u_\epsilon^h\|_{0,\Omega} \leq C_1\left(\frac{h}{\epsilon_m}\right)^2 + C_2\frac{\epsilon_{m+1}}{h} \qquad (5.9)$$

6 Numerical experiments

In this section we study the convergence and the accuracy of the multiscale method through numerical experiments. The model problem is solved by using the multiscale method with the base functions defined by (2.5) and the linear boundary conditions. Since it is very difficult to construct a test problem with both exact solution and sufficient generality, we use resolved numerical solutions in place of exact solutions. The numerical results are compared with the theoretical analysis.

The implementations of the multiscale method has been given in [11]. Here we outline the implementation and define some notations to be used below. All computations are performed on a unit square domain, i.e., $\Omega = (0,1) \times (0,1)$. Let N be the number of elements in x and y directions. The mesh size is thus $h = 1/N$. To compute the base functions, each element discretized into $M \times M$ subcell elements with size of $h_s = h/M$. 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 and the right hand side are computed using numerical quadrature rules. We compute the gradients of a base function at the center of a subcell element and use two-dimensional centered trapezoidal rule for the volume integration. This procedure ensures that the entries of the stiffness matrix 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 $\phi^4 = 1 - \phi^1 - \phi^2 - \phi^3$.

In all examples below, the resolved solutions are obtained using linear FEM. Given the wave length of small scales ϵ_1 and ϵ_2, we solve the model problem twice on two meshes with one mesh size being twice of other. Then the Richardson extrapolation is used to approximate the exact solutions from numerical solutions on two meshes. Throughout our numerical experiments, both of the mesh sizes used to compute the well resolved solution are less than $\epsilon/8$, so that the error of the extrapolated solutions is less that 10^{-6}.

The main difficulty in our tests is that the choices of well separated physical small scales for test problems are severely limited by the available computer memory. Our parallel implementation of the multiscale method on an Intel Paragon computer with 512 processors enables us to test the two scale problem. Even so, we are limited to testing the cases with $\epsilon_1 \gg h \gg \epsilon_2$ and $\epsilon_1 \gg \epsilon_2 \gg h$. The case with $h \gg \epsilon_1 \gg \epsilon_2$ with small mesh size h requires the value of ϵ_2 to be very small in order for the resonance error to be noticeable in numerical tests (recall that except the resonance errors there are errors associated with ϵ_2/ϵ_1, see (4.20) and (5.5)). This (the smallness of ϵ_2) makes it impossible to calculate the fully resolved (benchmark) solution in the whole domain using the traditional FEM with

our computer resources. Thus, it is difficult calculate the numerical error or the convergence of MsFEM.

The case with h in between the two physical scales is generic and important for practical purposes. As indicated by Theorem 6.2, the L_2 norm error is given by

$$C_1 \left(\frac{h}{\epsilon_1}\right)^2 + C_2 \frac{\epsilon_2}{h}, \tag{6.1}$$

which consists of two resonance error. One would expect that when h is close to ϵ_1 the first term dominates and when h is close to ϵ_2 the second term dominates. This asymptotic observation, however, is not always reflected by the numerical results shown below. There are two reasons. Firstly, the constants, C_1 and C_2, may differ by a large factor which is problem dependent. Secondly, it is difficult to choose well separated ϵ_1 and ϵ_2 in the numerical computations. Therefore, the two error may interact with each other. Instead of verifying each of them, in the following, we show that the numerical error does follow the estimate in whole. Furthermore, we use a least square fitting to obtain the constants C_1 and C_2 in (6.1). These constants indicate the relative magnitude of the two terms. More discussions along with numerical examples can be found in [10].

Again due to limitation of computing resources, in all tests below, ϵ_1 and ϵ_2 are fixed during the tests and we only allow h to vary.

Example 1. In this example, we solve (2.1) with $f = -1$, $u|_{\partial\Omega} = 0$, and

$$a^{ij} = \frac{\delta^{ij}}{3 + \sin(\frac{2\pi x}{\epsilon_1}) + \cos(\frac{2\pi y}{\epsilon_2})}, \tag{6.2}$$

where δ^{ij} are the Kroneker's symbols.

First we present the numerical examples to demonstrate the convergence rate (6.1). The error estimate (6.1) corresponds to the case $\epsilon_1 \gg h \gg \epsilon_2$. In our numerical examples we fix the values of ϵ_1 and ϵ_2, $\epsilon_1 = 0.5$ and $\epsilon_2 = 0.005$, and vary h between ϵ_1 and ϵ_2. In the Table 6.1 we compare the discrete l_2 errors between the MsFEM solution and the refined solution. The presence of both h^2/ϵ_1^2 and ϵ_2/h can be observed from this table. Indeed if h is close to ϵ_1 the convergence rate of the method is the second order with respect to h, and if h is close to ϵ_2 the convergence rate of the method is negative first order (see (6.1)). Notice that because there are lower order errors (see (4.15) and (5.5)) except the leading resonance errors we observe some deviation from (6.1). Furthermore, in Fig. 6.1 we try to find the best fit in the form of (6.1) to the l_2 errors given in the Table 6.1. Using the least square method we compute the values of C_1 and C_2 in (6.1) for this best fit. The computed values of $C_1 = 5e-4$ and $C_2 = 1.5e-3$ indicate that

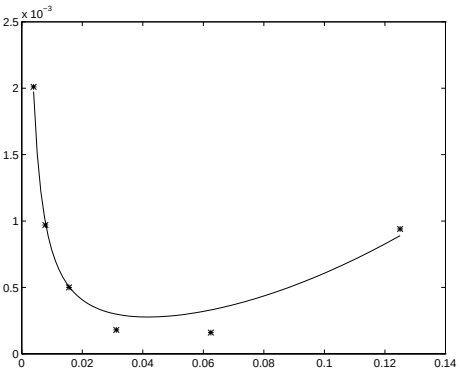

Fig. 6.1.

Table 6.1. $\|u_\epsilon - u_\epsilon^h\|_{l_2}$ for $\epsilon_1 \gg h \gg \epsilon_2$

h	l_2	rate
1/8	9.42e-4	
1/16	1.61e-4	2.5
1/32	1.87e-4	
1/64	5.08e-4	-1.44
1/128	9.77e-4	-0.94
1/256	2.01e-3	-1.04

Table 6.2. $\|u_\epsilon - u_\epsilon^h\|_{l_2}$ and $\|u_\epsilon - u_\epsilon^h\|_{l_\infty}$ for $h \ll \epsilon_2 \ll \epsilon_1$

h	l_2	rate	l_∞	rate
1/32	1.10e-3		2.04e-3	
1/64	3.22e-4	1.77	5.89e-4	1.79
1/128	8.40e-5	1.98	1.52e-4	1.95
1/256	2.09e-5	2.00	3.79e-5	2.00
1/512	4.90e-6	2.09	8.93e-6	2.08
1/1024	1.22e-6	2.00	2.22e-6	2.00

the second term in (6.1) has larger weight than the first term. Recall that the second term in (6.1) is due to capturing the small scales on the large scales. The discrepancy between the best fit in the form (6.1) and the discrete l_2 errors in the Fig. 6.1 is due to the lower order errors in the L_2 error estimate.

The result of the numerical experiments for the case $h \ll \epsilon_2 \ll \epsilon_1$ is shown in Table 6.2. In this table we compare the discrete l_2 errors between the MsFEM solution and the refined solution for fixed values of ϵ_1 and ϵ_2, $\epsilon_1 = 0.2$ and $\epsilon_2 = 0.08$. The theoretical results of the previous sections predict the second order convergence rate in h for the resonance error. As it can be observed that the convergence rate of MsFEM is the second order with respect to h.

Table 6.3. $\|u_\epsilon - u_\epsilon^h\|_{l_2}$ for $\epsilon_1 \gg h \gg \epsilon_2$

h	l_2	rate
1/8	2.15e-3	
1/16	5.28e-4	2.02
1/32	3.28e-4	
1/64	1.75e-4	
1/128	1.60e-4	

Example 2. In our next example, we consider the equation (2.1) with $f = 0$ and linear boundary conditions, $u|_{\partial\Omega} = x$, and

$$a^{ij} = \left(\frac{2+\sin(2\pi x/\epsilon_1)}{2+\cos(2\pi y/\epsilon_1)} + \frac{2+\cos 2\pi(y/\epsilon_1)}{2+\sin(2\pi x/\epsilon_1)}\right) \times \left(\frac{2+\sin(2\pi x/\epsilon_2)}{2+\cos(2\pi y/\epsilon_2)} + \frac{2+\cos(2\pi y/\epsilon_2)}{2+\sin(2\pi x/\epsilon_2)}\right) \delta^{ij} \tag{6.3}$$

where δ^{ij} are Kronoker's symbols.

As in the first example we fix $\epsilon_1 = 0.125$ and $\epsilon_2 = 0.0078125$ and vary the mesh size h between ϵ_1 and ϵ_2. In the Table 6.3 the discrete l_2 errors between the computed solution using MsFEM and the refined solution are presented. As we see from this table the resonance error ϵ_2/h does not appear explicitly for the mesh size $h = O(\epsilon_2)$. The reason for this is the large weight of C_1 in (6.1) which suppresses the term ϵ_2/h. As in the previous example using the least square method we find the best fit in the form (6.1) to the discrete l_2 error data given in the Table 6.3. The constants C_1 and C_2 for this best fit are $C_1 = 2.1 \times 10^{-3}$ and $C_2 = 2.0 \times 10^{-4}$. Consequently, the h^2/ϵ_1^2 term in (6.1) has larger weight than the ϵ_2/h term. The fitted curve as well as the data points for the numerical error are plotted in Fig. 6.2. It shows that the error varies consistently with the estimate (6.1). The deviation between the computed l_2 errors and the best fit of the form (6.1) is due to the lower order errors in the L_2 error estimate (see (4.15) and (5.5)).

Example 3. In our last example we take $f = -1$, and $u|_{\partial\Omega} = 0$ while defining a^{ij} by (6.3). In the Table 6.4 we present the discrete l_2 error between the refined solution and the solution computed with the use of MsFEM for the fixed $\epsilon_1 = 0.125$ and $\epsilon_2 = 0.0078125$. As in the previous examples the mesh size h varies between ϵ_1 and ϵ_2. As we can see from this table that the presence of both h^2/ϵ_1^2 and ϵ_2/h terms (see (6.1)) appear as h becomes close to ϵ_1 and ϵ_2, respectively. Using the least square method we can compute the values of C_1 and C_2 ($C_1 = 0.79 \times 10^{-4}$, $C_2 = 0.29 \times 10^{-4}$) for which the expression (6.1) approximates the l_2 error data given in the Table 6.4. The discrete l_2 error data between the solution computed using MsFEM and

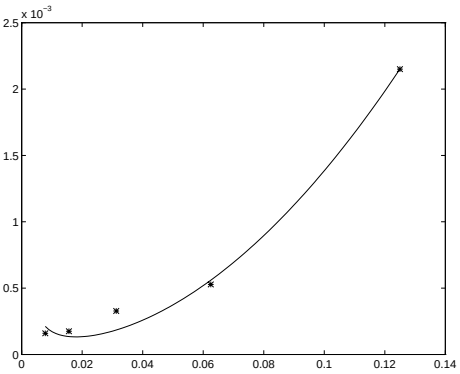

Fig. 6.2.

Table 6.4. $\|u_\epsilon - u_\epsilon^h\|_{l_2}$ for $\epsilon_1 \gg h \gg \epsilon_2$

h	l_2	rate
1/8	8.25e-5	
1/16	1.92e-5	2.10
1/32	1.77e-5	
1/64	1.17e-5	
1/128	3.04e-5	-1.38

Table 6.5. $\|u_\epsilon - u_\epsilon^h\|_{l_2}$ and $\|u_\epsilon - u_\epsilon^h\|_{l_\infty}$ for $h \ll \epsilon_2 \ll \epsilon_1$

h	l_2	rate	l_∞	rate
1/64	1.14e-5		2.84e-5	
1/128	2.81e-6	2.02	5.79e-6	2.29
1/256	7.13e-7	1.97	1.56e-6	1.89
1/512	1.69e-7	2.07	3.91e-7	2.00
1/1024	3.33e-8	2.35	1.01e-7	1.95

the refined solution for the case $h \ll \epsilon_2 \ll \epsilon_1$ is shown in Table 6.5, where the second order convergence of MsFEM can be observed.

7 Concluding remarks and generalizations

The purpose of the multiscale method is to provide a systematic approach to capture the small scale effect on large scales when we cannot afford to resolve all the small scale features in the physical solution. Our study shows that MsFEM is a robust method for practical multiple scale problems. In particular, the method works for multiple scale problems and the mesh size can be chosen to be between the two physical scales. We note that there are two types of resonance error, one from resolving the large scales, the other from capturing the small scales. The second type of error is caused by the

artificial boundary layers in our base functions. This important issue and its numerical resolution, e.g., the over-sampling method, has been analyzed for problems with one small scale [9]. They will be further studied in the multiple scale cases in our subsequent work.

The first type of resonance error is common among traditional finite difference and finite element methods. The traditional approaches, however, cannot capture the small scales. To reduce this error, a natural idea is to generalize MsFEM to higher order in the sense that the large scales are more accurately resolved. The idea is to construct base functions such that their homogenized parts consist of higher order polynomials than linear (bilinear) functions. This can be achieved by changing the linear boundary condition of the base functions to higher order polynomials.

Denoting S^h as a finite dimensional subspace of $H_0^1(\Omega)$ such that for any

$$b(u,v) = \int_\Omega b^{ij}(x/\epsilon_1, \ldots, x/\epsilon_k)\nabla_i u \nabla_j v dx$$

with $b^{ij}(y_1, \ldots, y_k)$ are sufficiently smooth periodic functions in y_i $(i = 1, \ldots, k)$ in a unit cube, $\alpha|\xi|^2 \le b^{ij}\xi_i\xi_j \le \beta|\xi|^2$, and $\epsilon_1 \gg \epsilon_2 \gg \cdots \gg \epsilon_k \gg h$ we have

$$\|u - u^h\|_{1,\Omega} \le C\left(\frac{h}{\epsilon_k}\right)^n, \tag{7.1}$$

where $u \in H_0^1(\Omega)$ is the solution of $b(u,v) = (f,v)$ for any $v \in H_0^1(\Omega)$ and $u^h \in S^h$ is the solution of $b(u^h, v^h) = (f, v^h)$ for any $v^h \in S^h$. For example S^h can be the space spanned by high order polynomials of degree n [8]. In each element $K \in \mathbf{K}^h$ we denote $P_K^i \in S^h$ $(i = 1, \ldots, d)$ a set of nodal basis of S^h. Then the high order multiscale base functions are given by ϕ_K^i $(i = 1, \ldots, d)$, which satisfy (2.5) and the boundary conditions

$$\phi_K^i|_{\partial K} = P_K^i|_{\partial K}. \tag{7.2}$$

It can be shown that with these high order multiscale base functions we have

$$\|u_\epsilon - u_\epsilon^h\|_{1,\Omega} \le Ch + C\left(\frac{h}{\epsilon_m}\right)^n + C_3\sqrt{\frac{\epsilon_{m+1}}{h}} + C\max_{i \ge m+1}\frac{\epsilon_i}{\epsilon_{i-1}} \tag{7.3}$$

instead of (4.24). The outline of the proof is the following. We continue v_ϵ^h on to K as in (4.26). The nodal values of the interpolant α_i in (4.4) can be defined as those of u^h satisfying (7.1). Then taking into account the estimates for u_ϵ^1 and θ of Sect. 4.2 we can conclude (7.3). We note S^h needs not to be a subspace of H^1, e.g., S^h can consists of nonconforming multiscale base functions.

The estimate (7.3) can also be obtained for the smooth domains Ω. Note that in this case we have the smoothness of u_ϵ (and any partially homogenized part of it) for fixed ϵ. The difficulty in deriving H^1 norm estimate for MsFEM in this case is in the calculation of the multiscale base functions near the boundary $\partial\Omega$. There are various ways to treat curve boundaries in the finite element methods. In fact, following the triangulation of Ω in [18] and using the nodal base functions constructed in that paper to provide the boundary conditions for the multiscale base functions, we can show that (7.3) holds on the smooth domain.

Finally we would like to note that the assumption (3.21) which requires some compatibility conditions for the problem is not necessary for deriving the H^1 estimates. Without this assumption it can be shown that

$$\begin{aligned}&\|u_\epsilon - u_\epsilon^h\|_{1,\Omega}\\ &\quad\le C\left(h + \frac{h}{\epsilon_m} + \left(\frac{\epsilon_{m+1}}{h}\right)^{1/2-1/p} + \sqrt{p}\frac{\epsilon_{m+1}^{1/2-1/p}h^{1/2+1/p}}{\epsilon_m}\right)\\ &\qquad\times\|f\|_{0,\Omega} + \|\theta_\epsilon\|_{1,\Omega}\end{aligned} \tag{7.4}$$

where $p > 2$ is an arbitrary constant. Because of the insufficient smoothness of the homogenized parts of the solution near the corner points of the domain Ω, the rate of convergence of $\|\theta_\epsilon\|_{1,\Omega}$ to zero as $\epsilon \to 0$ deteriorates depending on the parameters of the problem.

In order to show the estimate (7.4) we only need to reestimate $\|u_\epsilon^1\|_{H^{1/2}(\partial K)}$ in (4.25),

$$\begin{aligned}\|u_\epsilon - v_\epsilon^h\|_{H^1(K)} &\le Ch\|f\|_{0,K} + Ch|u_0|_{2,K} + C\|u_\epsilon^1\|_{H^{1/2}(\partial K)}\\ &\quad + C\|\theta\|_{1,K}.\end{aligned} \tag{7.5}$$

Here we derive this estimate only for the two scale case $\epsilon_1 \gg h \gg \epsilon_2$. The derivation for the general n scale case is similar. In the case $\epsilon_1 \gg h \gg \epsilon_2$, $u_\epsilon^1 = \epsilon_2\chi^m(\frac{x}{\epsilon_2})\nabla_m u_0^\lambda$.

Introducing the family of boundary function τ_ϵ in K as in Sect. 3 with the properties 1 and 2 we have

$$\begin{aligned}\|u_\epsilon^1\|_{H^{1/2}(\partial K)} &= \left\|\epsilon_2\chi^m\left(\frac{x}{\epsilon_2}\right)\nabla_m u_0^\lambda\right\|_{H^{1/2}(\partial K)}\\ &\le \left\|\epsilon_2\left(1-\tau_\epsilon\right)\chi^m\left(\frac{x}{\epsilon_2}\right)\nabla_m u_0^\lambda\right\|_{H^1(K)}\\ &\le \left\|\epsilon_2\left(1-\tau_\epsilon\right)\chi^m\left(\frac{x}{\epsilon_2}\right)\nabla_m u_0^\lambda\right\|_{L_2(K)}\\ &\quad + \left\|\epsilon_2\nabla\tau_\epsilon\chi^m\left(\frac{x}{\epsilon_2}\right)\nabla_m u_0^\lambda\right\|_{L_2(K)}\end{aligned}$$

$$+\left\|\epsilon_2\left(1-\tau_\epsilon\right)\nabla\left(\chi^m\left(\frac{x}{\epsilon_2}\right)\right)\nabla_m u_0^\lambda\right\|_{L_2(K)}$$
$$+\left\|\epsilon_2\left(1-\tau_\epsilon\right)\chi^m\left(\frac{x}{\epsilon_2}\right)\nabla\nabla_m u_0^\lambda\right\|_{L_2(K)}. \tag{7.6}$$

Denoting the support of $1-\tau_\epsilon$ by S we have $\text{meas}(S)\le C\epsilon_2 h$. Furthermore taking into account that $|\epsilon_2\nabla\chi|\le C$ and $|\epsilon_2\nabla\tau|\le C$ we can estimate the r.h.s of (7.6) as

$$\begin{aligned}\|u_\epsilon^1\|_{H^{1/2}(\partial K)} &\le C\epsilon_2|u_0|_{H^1(K)}+C|u_0|_{H^1(S)}\\ &\quad +C|u_0|_{H^1(S)}+C\epsilon_2|u_0|_{H^2(K)}.\end{aligned} \tag{7.7}$$

For the estimate of $\|\nabla u_0\|_{L_2(S)}$ we use the following inequalities [19]:

$$\|v\|_{L_2(S)}\le(\text{meas}(S))^{1/2-1/p}\|v\|_{L_p(S)},\quad p>2,$$

and

$$\|v\|_{L_p(K)}\le Ch^{-1+2/p}\|v\|_{L_2(K)}+Ch^{2/p}\sqrt{p}|v|_{H^1(K)},\quad p\ge 1.$$

Then

$$\begin{aligned}|u_0|_{H^1(S)} &\le C(\epsilon_2 h)^{1/2-1/p}\|\nabla u_0\|_{L_p(S)}\\ &\le C(\epsilon_2 h)^{1/2-1/p}\left(h^{-1+2/p}\|\nabla u_0\|_{L_2(K)}+h^{2/p}\sqrt{p}|\nabla u_0|_{H^1(K)}\right)\\ &\le C(\epsilon_2/h)^{1/2-1/p}\|\nabla u_0\|_{L_2(K)}\\ &\quad +C(\epsilon_2 h)^{1/2-1/p}h^{2/p}\sqrt{p}|\nabla u_0|_{H^1(K)}.\end{aligned} \tag{7.8}$$

Furthermore combining the estimates (7.7) and (7.8) we have

$$\begin{aligned}\|u_\epsilon^1\|_{H^{1/2}(\partial K)} \le C\Bigg(&(\epsilon_2/h)^{1/2-1/p}\|\nabla u_0\|_{L_2(K)}\\ &+(\epsilon_2^{1/2-1/p}h^{1/2+1/p}\sqrt{p}+\epsilon_2)|u_0|_{H^2(K)}\Bigg).\end{aligned} \tag{7.9}$$

Finally summing (7.5) over all K we obtain (7.4). As we see the estimate (7.4) is slightly weaker than (4.29). Note that $\epsilon_2^{1/2-1/p}h^{1/2+1/p}\sqrt{p}/\epsilon_1$ is asymptotically small.

Acknowledgements. We are grateful to Professor Thomas Y. Hou and Dr. Yu Zhang for reading the manuscript and for many interesting and helpful discussions. We also would like to thank Caltech HPCC for providing the parallel computing resource. 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.

References

1. S. Agmon: Lectures on elliptic boundary value problems. Princeton, N.J.: Van Nostrand 1965
2. G. Allaire, M. Briane: Multiscale convergence and reiterated homogenization. Proc. Roy. Soc. Edinburgh A Math. **126**, 297–342 (1996)
3. M. Avellaneda, F-H. Lin: Compactness methods in the theory of homogenization. Commun Pure Appl. Math. **XL**, 803–843 (1989)
4. A. Azzam: Smoothness properties of bounded solution of dirichlet problem for elliptic equations in regions with corners on the boundaries. Canad. Math. Bull. **23** (1980)
5. I. Babuška, G. Caloz, E. Osborn: Special finite element methods for a class of second order elliptic problems with rough coefficients. SIAM J. Numer. Anal. **31**, 945–981 (1994)
6. A. Bensoussan, J. L. Lions, G. Papanicolaou: Asymptotic Analysis for Periodic Structure, volume 5 of Studies in Mathematics and Its Applications. Amsterdam: North-Holland 1978
7. F. Brezzi, L. Franca, T. Hughes, A. Russo: b is equal to integral of g. Comput. Methods Appl. Mech. Engrg. **145**, 329–339 (1997)
8. P. G. Ciarlet: The finite element method for elliptic problems. Amsterdam: North-Holland 1978
9. Y. R. Efendiev, T. Y. Hou, X. H. Wu: The convergence of non-conforming multiscale finite element methods. SIAM in Num. Anal. **37**, (3) 888–910 (1998)
10. Y.R. Efendiev: The multiscale finite element method and its applications. Ph.D thesis, California Institute of Technology, 1999
11. T. Y. Hou, X. H. Wu: A multiscale finite element method for elliptic problems in composite materials and porous media. J. Computat. Phy. **134**, 169–189 (1997)
12. T. Y. Hou, X. H. Wu: A multiscale finite element method for PDEs with oscillatory coefficients. Notes on Numerical Fluid Mechanics **70**, 58–69 (1999)
13. T. Y. Hou, X. H. Wu, Z. Cai: Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients. Math. Comp. **68** (227), 913–943 (1999)
14. T. Hughes, G. Feijoo, L. Mazzei, J. Quincy: The variational multiscale method - a paradigm for computational mechanics. Comput. Methods Appl. Mech. Engrg. **166**, 3–24 (1998)
15. V. V. Jikov, S. M. Kozlov, O. A. Oleinik: Homogenization of differential operators and integral functionals. Berlin Heidelberg New York: Springer 1994
16. O. A. Ladyzhenskaia, N. N. Uraltseva: Linear and quasilinear elliptic eqautions. New York: Academic Press 1968
17. E. Oriordan, M. Stynes: A globally uniformly convergent finite element method for a singularly perturbed elliptic problem in 2 dimensions. Math. Comp. **57**, 47–62 (1991)
18. R. Scott: Interpolated boundary conditions in the finite element method. SIAM J. Numer. Anal. **12**, 404–427 (1975)
19. L. Ying: Infinite Element Methods. Beijing: Peking University Press 1995