MULTISCALE MODEL. SIMUL.
Vol. 12, No. 4, pp. 1424–1457

A COMBINED FINITE ELEMENT AND MULTISCALE FINITE ELEMENT METHOD FOR THE MULTISCALE ELLIPTIC PROBLEMS*

WEIBING DENG† AND HAIJUN WU†

Abstract. The oversampling multiscale finite element method (MsFEM) is one of the most popular methods for simulating composite materials and flows in porous media which may have many scales. But the method may be inefficient in some portions of the computational domain, e.g., near long narrow channels inside the domain due to the fact that the high-conductivity features cannot be localized within a coarse-grid block, or in a near-well region since the solution behaves like the Green function there. In this paper we develop a combined finite element and multiscale finite element method (FE-MsFEM), which deals with such portions by using the standard finite element method on a fine mesh and the other portions by the oversampling MsFEM. The transmission conditions across the FE-MSFE interface is treated by the penalty technique. To illustrate this idea, a rigorous error analysis for this FE-MsFEM is given under the assumption that the diffusion coefficient is periodic. Numerical experiments are carried out for the elliptic equations with periodic and random highly oscillating coefficients, as well as multiscale problems with high contrast channels or well singularities, to demonstrate the accuracy and efficiency of the proposed method.

Key words. multiscale problems, oversampling technique, interface penalty, combined finite element and multiscale finite element method

AMS subject classifications. 34E13, 35B27, 65N12, 65N15, 65N30

DOI. 10.1137/120898279

1. Introduction.

Let $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, be a polyhedral domain, and consider the elliptic equation

$$\begin{cases} -\nabla \cdot (\mathbf{a}^\epsilon(x)\nabla u_\epsilon(x)) = f(x) & \text{in } \Omega, \\ u_\epsilon(x) = 0 & \text{on } \partial\Omega, \end{cases} \tag{1.1}$$

where $0 < \epsilon \ll 1$ is a parameter that represents the ratio of the smallest and largest scales in the problem. We assume that $f \in L^2(\Omega)$, $a_{ij}^\epsilon \in L^\infty(\Omega)$, $1 \le i, j \le n$, and $\mathbf{a}^\epsilon(x) = (a_{ij}^\epsilon(x))$ is a symmetric, positive definite matrix:

$$\lambda|\xi|^2 \le a_{ij}^\epsilon(x)\xi_i\xi_j \le \Lambda|\xi|^2 \quad \forall \xi \in \mathbb{R}^n, \text{ a.e. in } \Omega \tag{1.2}$$

for some positive constants λ and Λ. The weak formulation of (1.1) is to find $u_\epsilon \in H_0^1(\Omega)$ such that

$$\int_\Omega \mathbf{a}^\epsilon(x)\nabla u_\epsilon(x)\nabla v \, \mathrm{d}x = \int_\Omega f v \, \mathrm{d}x \quad \forall v \in H_0^1(\Omega).$$

Problems of the type (1.1) are often used to describe the models arising from composite materials and flows in porous media, which contain many spatial scales.

*Received by the editors November 8, 2012; accepted for publication (in revised form) July 7, 2014; published electronically October 9, 2014.
http://www.siam.org/journals/mms/12-4/89827.html
†Department of Mathematics, Nanjing University, Jiangsu, 210093, People's Republic of China (wbdeng@nju.edu.cn, hjw@nju.edu.cn). The work of the first author was partially supported by the NSF of China under grants 10971096 and 11271187, by the PAPD, and by the Fundamental Research Funds for the Central Universities 1116020306. The work of the second author was partially supported by the National Magnetic Confinement Fusion Science Program under grant 2011GB105003 and by the NSF of China under grants 11071116, 91130004.

Solving these problems numerically is difficult because resolving the smallest scale usually requires very fine meshes and hence a tremendous amount of computer memory and CPU time. To overcome this difficulty, many methods have been designed to solve the problem on meshes that are coarser than the scale of oscillations. One of the most popular methods is the multiscale finite element method (MsFEM) [31, 46, 47], which takes its origin from the work of Babuška and Osborn [9, 8]. Two main ingredients of the MsFEM are the global formulation of the method such as various finite element methods and the construction of basis functions. The special basis functions are constructed from the local solutions of the elliptic operator that contains the small scale information within each element. By solving the problem (1.1) in the special basis function space, they get a good approximation of the full fine-scale solution. We remark that there have been many other methods proposed to solve this type of multiscale problem in the past several decades. See, for instance, wavelet homogenization techniques [21, 33], multigrid numerical homogenization techniques [38, 55], the subgrid upscaling method [2, 3], the heterogeneous multiscale method [24, 25, 26], the residual-free bubble method (or the variational multiscale method, discontinuous enrichment method) [14, 34, 39, 40, 49, 59], mortar multiscale methods [4, 58], and upscaling or numerical homogenization method [23, 35, 64]. We refer the reader to the book [29] for an overview and other references of multiscale numerical methods in the literature, especially a description of some intrinsic connections between most of these methods.

In this paper, we focus on the MsFEM. Many developments and extensions of the MsFEM have been done in the past ten years. See, for example, the mixed MsFEM [15, 1]; the MsFEMs for nonlinear problems [30, 32]; the Petro–Galerkin MsFEM [48]; the MsFEMs using limited global information [28, 56]; the adaptive MsFEM based on the a posteriori error estimate which splits into different contributions accounting for the coarse grid error, the fine grid error, and the oversampling error, respectively [45]; and the multiscale finite volume method [50]. In [47], it is shown that there is a resonance error between the grid scale and the scales of the continuous problem. Especially for the two-scale problem, the resonance error 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 scale resonance is a fundamental difficulty caused by the mismatch between the local construction of the multiscale basis functions and the global nature of the elliptic problems. This mismatch between the local solution and the global solution produces a thin boundary layer in the first order corrector of the local solution. To overcome the difficulty due to the scale resonance, an oversampling technique was proposed in [46, 31]. The basic idea is to compute the local problem in the domain with size larger than the mesh size H and use only the interior sampled information to construct the basis functions. By doing this, the influence of the boundary layer in the larger domain on the basis functions is greatly reduced. The other origin of scale resonance is the mismatch between the grid size and the "perfect" sample size (e.g., for periodic structures, the "perfect" sample size would be integer multiples of the period). The resonance thus created is called "cell resonance." It is shown that the Petro–Galerkin MsFEM can remove the cell resonance error by use of the conforming piecewise linear test space (see [48]).

However, the oversampling MsFEM may be inefficient for multiscale problems that have some singularities, such as the Dirac function singularities, which stem from the simulation of steady flow transport through highly heterogeneous porous media driven by extraction wells [17], or high-conductivity channels that connect the

boundaries of coarse-grid blocks [41, 42, 27, 57]. Noting that the FEM has many ways of dealing with the singularities—such as refining the mesh, enlarging the polynomial order of the finite element space, or even using the adaptive algorithm—a natural idea might be to combine the FEM and oversampling MsFEM and take advantage of both methods, thus improving the accuracy of the simulation.

In this paper, we introduce a combined finite element and multiscale finite element method (FE-MsFEM) by using the traditional FEM directly on a fine mesh of the problematic part of the domain and using the oversampling MsFEM on a coarse mesh of the other part. The difficulty in realizing this idea lies in joining the two methods together without losing the accuracy of both methods, i.e., dealing with the transmission condition on the interface between coarse and fine meshes efficiently. Thanks to the penalty techniques used in the interior penalty discontinuous (or continuous) Galerkin methods originating in the 1970s [10, 11, 22, 61, 5, 6], we may deal with the transmission condition on the interface by penalizing the jumps from the function values as well as the fluxes of the finite element solution on the fine mesh to those of the oversampling MsFE solution on the coarse mesh. We would like to remark that besides the applications of penalty technique to the interior penalty discontinuous (or continuous) Galerkin methods, this technique is also applied to the Helmholtz equation with high wave number to reduce the pollution error [36, 37, 62, 65] and applied to the interface problems to construct high order unfitted mesh methods [53, 63]. We remark also that [27] provided a different approach to dealing with long channels by introducing additional basis functions based on some localized eigenvalue problems to augment the classical MsFE space and [18] proposed an adaptive generalized multiscale finite element method (GMsFEM) for high-contrast flow problems.

In order to illustrate the accuracy of the FE-MsFEM, we first carry out a rigorous and careful analysis for the elliptic equation with periodic diffusion coefficient where the FEM is applied to a narrow region adjacent to the boundary of the computation domain, and the oversampling MsFEM is applied to the other part. It is shown that the H^1-error of our new method is just the sum of interpolation errors of both methods plus an error term of $O(\frac{H^2}{\sqrt{\epsilon}})$ introduced by the penalty terms, where H is the mesh size of the coarse mesh of the inner subdomain. In particular, if $H \leq C\sqrt{\epsilon}$, then this error term is dominated by the homogenization error (or order $\sqrt{\epsilon}$), that is, the FEM and oversampling MsFEM are jointed successfully by the introduced penalty technique. Second, we try to simulate multiscale elliptic problems with fine and long-ranged high-conductivity channels. These kinds of high-conductivity features cannot be localized within a coarse-grid block; hence it is difficult to be handled with standard or oversampling multiscale basis. By contrast, the numerical results show that the introduced FE-MsFEM can solve the high-contrast multiscale elliptic problems efficiently. Finally, we solve the elliptic multiscale problem with the Dirac function singularities inside the domain, which stems from the simulation of steady flow transport through highly heterogeneous porous media driven by extraction wells. It is well known that the discrete solution fails to give good approximation in the vicinity of the singularity. To overcome this difficulty, an oversampling MsFEM with new basis functions that locally resolve the well singularities is introduced in [17]. Our FE-MsFEM may solve such problems by using the traditional FEM on a fine mesh near the well singularities and using the oversampling MsFEM in the other part of the domain. The numerical results indicate our FE-MsFEM can also solve the well-singularity problems very efficiently. The convergence analysis for multiscale problems with singularities and applications to some practical problems such as two-phase flows

in porous media and other types of equations are currently under study.

The rest of this paper is organized as follows. In section 2, we formulate the FE-MsFEM for the model problem. In section 3, we review some classical homogenization results for the elliptic problems and give an interior H^2 norm error estimate between the multiscale solution and the homogenized solution with first order corrector. In section 4, we give some approximation properties for the oversampling MsFE space and the linear finite element space, respectively. The H^1-error estimate of the introduced FE-MsFEM is given in section 5. In section 6, we first give some numerical examples for both periodic and randomly generated coefficients to demonstrate the accuracy of the proposed method, and then we apply our method to multiscale elliptic problems which have fine and long-ranged high-conductivity channels as well as to multiscale problems with well singularities to demonstrate the efficiency of the method. Conclusions are drawn in the last section.

Before leaving this section, we fix some notation and conventions to be used in this paper. In the following, the Einstein summation convention is used: summation is taken over repeated indices. $L^2(\Omega)$ denotes the space of square integrable functions defined in domain Ω. We use the $L^2(\Omega)$-based Sobolev spaces $H^k(\Omega)$ equipped with norms and seminorms given by

$$\|u\|^2_{H^k(\Omega)} = \int_\Omega \sum_{|\alpha|\le k} |D^\alpha u|^2, \quad |u|^2_{H^k(\Omega)} = \int_\Omega \sum_{|\alpha|=k} |D^\alpha u|^2.$$

$\|u\|_{W^{k,\infty}(\Omega)}$ ($|u|_{W^{k,\infty}(\Omega)}$) is the $W^{k,\infty}$ norm (seminorm) of u in Ω. Throughout, $C, C_1, C_2, \ldots$ denote generic constants, which are independent of ϵ, H, and h unless otherwise stated. We also use the shorthand notation $A \lesssim B$ and $B \gtrsim A$ for the inequality $A \le CB$ and $B \ge CA$. The notation $A \eqsim B$ is equivalent to the statement $A \lesssim B$ and $B \lesssim A$.

2. FE-MsFEM formulation. In this section we present our FE-MsFEM. We describe the method only for the case of dealing with the difficulty of lack of information outside the domain in the oversampling MsFEM. Of course, the formulation can easily be extended to the case of dealing with singularities.

We first separate the research area Ω into two subdomains Ω_1 and Ω_2 such that $\Omega_2 \subset\subset \Omega$ and $\Omega = \Omega_1 \cup \Omega_2 \cup \Gamma$, where $\Gamma = \partial\Omega_1 \cap \partial\Omega_2$ is the interface of Ω_1 and Ω_2 (see Figure 1). For simplicity, we assume the following.

(H1) The length/area of Γ satisfies $|\Gamma| = O(1)$, and Γ is Lipschitz continuous.

Let $\mathcal{M}_h$ and $\mathcal{M}_H$ be the shape-regular and quasi-uniform triangulations of the domain Ω_1 and Ω_2, respectively, and denote by Γ_h and Γ_H the two partitions of the interface Γ induced by $\mathcal{M}_h$ and $\mathcal{M}_H$, respectively. For any element $K \in \mathcal{M}_h$ (or $K \in \mathcal{M}_H$), we define h_K (or H_K) as $\mathrm{diam}(K)$. Similarly, for each edge/face e of $K_e \in \mathcal{M}_h$ (or E of $K_E \in \mathcal{M}_H$), define h_e as $\mathrm{diam}(e)$ (or H_E as $\mathrm{diam}(E)$). Denote by $h = \max_{K\in\mathcal{M}_h} h_K$ and $H = \max_{K\in\mathcal{M}_H} H_K$. We assume the following for $h < H$.

(H2) On the interface Γ, $\mathcal{M}_H$ and $\mathcal{M}_h$ satisfy the matching condition that Γ_h is a refinement of Γ_H.

Clearly, each edge/face in Γ_H is composed of some edges/faces in Γ_h. Combining the two triangulations, we define $\mathcal{M}_{h,H}$ as the triangulation of Ω. (See Figure 1 for an illustration of triangulation.) Moreover, we assume the following.

(H3) For some constant C_0, $\mathrm{dist}\{\Gamma, \partial\Omega\} \ge C_0 H \ge h + 2\epsilon > 0$.

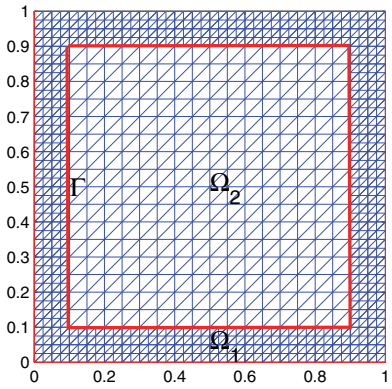

FIG. 1. *A separation of the domain and a sample mesh.*

Thus, we can define a narrow subdomain $\Omega_\Gamma \subset\subset \Omega$ surrounding Γ as

$$\begin{aligned}\Omega_\Gamma := \Gamma \cup \{x:\ x\in\Omega_1,\ \operatorname{dist}(x,\Gamma) < h+2\epsilon\}\\ \cup \{x:\ x\in\Omega_2,\ \operatorname{dist}(x,\Gamma) < H+2\epsilon\}.\end{aligned} \tag{2.1}$$

For any point on Γ, we associate a unit normal $\mathbf{n}$, which is oriented from Ω_1 to Ω_2. We also define the jump $[v]$ and average $\{v\}$ of v on the interface Γ as

$$[v] := v|_{\Omega_1} - v|_{\Omega_2}, \qquad \{v\} := \frac{v|_{\Omega_1} + v|_{\Omega_2}}{2}. \tag{2.2}$$

Introduce the "energy" space

$$V := \left\{v:\ v|_{\Omega_i} = v_i,\ \text{ where } v_i \in H_0^1(\Omega)\cap H^s(\Omega),\ i=1,2\right\} \text{ for some } s > \frac{3}{2}. \tag{2.3}$$

Testing the elliptic problem (1.1) by any $v\in V$, using integration by parts, and using the identity $[vw] = \{v\}[w] + [v]\{w\}$, we obtain

$$\sum_{i=1}^{2}\int_{\Omega_i} \mathbf{a}^\epsilon \nabla u_\epsilon\cdot\nabla v - \int_\Gamma \{\mathbf{a}^\epsilon\nabla u_\epsilon\cdot\mathbf{n}\}[v] = \int_\Omega fv.$$

Assuming further that $a_{ij}^\epsilon|_{\Omega_\Gamma} \in C(\Omega_\Gamma)$, $1\le i,j\le n$, we define the bilinear form $A_\beta(\cdot,\cdot)$ on $V\times V$ as follows:

$$\begin{aligned}A_\beta(u,v) := &\sum_{K\in\mathcal{M}_{h,H}}\int_K \mathbf{a}^\epsilon\nabla u\cdot\nabla v\\ &- \sum_{e\in\Gamma_h}\int_e \Big(\{\mathbf{a}^\epsilon\nabla u\cdot\mathbf{n}\}[v] + \beta[u]\{\mathbf{a}^\epsilon\nabla v\cdot\mathbf{n}\}\Big)\\ &+ J_0(u,v) + J_1(u,v),\end{aligned} \tag{2.4}$$

$$J_0(u,v) := \sum_{e\in\Gamma_h}\frac{\gamma_0}{\rho}\int_e [u][v], \tag{2.5}$$

$$J_1(u,v) := \sum_{e\in\Gamma_h}\gamma_1\rho\int_e [\mathbf{a}^\epsilon\nabla u\cdot\mathbf{n}][\mathbf{a}^\epsilon\nabla v\cdot\mathbf{n}], \tag{2.6}$$

where β is a real number such as $-1, 0, 1$, and $\gamma_0, \gamma_1, \rho > 0$ will be specified later. Define further the linear form $F(\cdot)$ on V:

$$F(v) := \int_\Omega fv.$$

It is easy to check that if the solution to the problem (1.1) satisfies $u_\epsilon \in H^2(\Omega)$, then it holds that

$$A_\beta(u_\epsilon, v) = F(v) \qquad \forall\, v \in V. \tag{2.7}$$

Remark 2.1. The regularity of $u_\epsilon \in H^2(\Omega)$ can be guaranteed by assuming that $a_{ij}^\epsilon \in W^{1,\infty}(\Omega)$. On the other hand, if $v \in H^s(\Omega)$ with $s > 3/2$, then $\partial v/\partial x_j$, $j = 1, \ldots, n$, has a trace in $L^2(\partial K)$ [13]. Noting that $a_{ij}^\epsilon \in L^\infty(\Omega)$ and $a_{ij}^\epsilon|_{\Omega_\Gamma} \in C(\Omega_\Gamma)$, $1 \le i, j \le n$, the bilinear form $A_\beta(\cdot, \cdot)$ is well-defined. It should be mentioned that the assumption $u_\epsilon \in H^2(\Omega)$ (or $a_{ij}^\epsilon \in W^{1,\infty}(\Omega)$) is unnecessary for the implementation of our discrete scheme (see (2.13)) but it will be used in the error analysis.

To formulate the FE-MsFEM, we need the oversampling MsFE space on $\mathcal{M}_H$ defined as follows (cf. [16, 46, 29]). For any $K \in \mathcal{M}_H$ with nodes $\{x_i^K\}_{i=1}^{n+1}$, let $\{\varphi_i^K\}_{i=1}^{n+1}$ be the basis of $P_1(K)$ satisfying $\varphi_i^K(x_j^K) = \delta_{ij}, 1 \le i, j \le n+1$, where δ_{ij} stands for Kronecker's symbol. For any $K \in \mathcal{M}_H$, we denote by $S = S(K)$ a macro-element (simplex) which contains K, and ∂S is away from ∂K at some distance $d_{S(K)} := \mathrm{dist}(K, \partial S)$ (to be specified later). Denote by $\{\varphi_i^S\}_{i=1}^{n+1}$ the nodal basis of $P_1(S)$ such that $\varphi_i^S(x_j^S) = \delta_{ij}, 1 \le i, j \le n+1$, where x_j^S are vertices of S. Let $\psi_i^S \in H^1(S)$, $i = 1, \ldots, n+1$, be the solution of the problem

$$-\nabla \cdot (\mathbf{a}^\epsilon \nabla \psi_i^S) = 0 \qquad \text{in } S, \qquad \psi_i^S|_{\partial S} = \varphi_i^S. \tag{2.8}$$

The oversampling MsFE basis functions over K is defined by

$$\bar{\psi}_i^K = c_{ij}^K \psi_j^S|_K \qquad \text{in } K, \tag{2.9}$$

with the constants c_{ij}^K determined by

$$\varphi_i^K = c_{ij}^K \varphi_j^S|_K \qquad \text{in } K. \tag{2.10}$$

The existence of the constants c_{ij}^K is guaranteed because $\{\varphi_j^S\}_{j=1}^{n+1}$ forms a basis of $P_1(K)$.

Let $\mathrm{OMS}(K) = \mathrm{span}\,\{\bar{\psi}_i^K\}_{i=1}^{n+1}$ be the set of space functions on K. Define the projection $\Pi_K : \mathrm{OMS}(K) \to P_1(K)$ as

$$\Pi_K \psi = c_i \varphi_i^K \qquad \text{if} \qquad \psi = c_i \bar{\psi}_i^K \in \mathrm{OMS}(K).$$

Introduce the space of discontinuous piecewise "OMS" functions and the space of discontinuous piecewise linear functions:

$$\bar{X}_H = \{\psi_H :\ \psi_H|_K \in \mathrm{OMS}(K)\ \ \forall K \in \mathcal{M}_H\},$$
$$\bar{W}_H = \{w_H :\ w_H|_K \in P_1(K)\ \ \forall K \in \mathcal{M}_H\}.$$

Define $\Pi_H : \bar{X}_H \to \bar{W}_H$ through the relation

$$\Pi_H \psi_H|_K = \Pi_K \psi_H \ \text{ for any } K \in \mathcal{M}_H, \psi_H \in \bar{X}_H.$$

The oversampling MsFE space on $\mathcal{M}_H$ is then defined as

$$X_H = \{\psi_H \in \bar{X}_H : \Pi_H \psi_H \in W_H \subset H^1(\Omega_2)\},$$

where $W_H = \bar{W}_H \cap H^1(\Omega_2)$ is the H^1-conforming linear finite element space over $\mathcal{M}_H$. In general, $X_H \not\subset H^1(\Omega_2)$ and the requirement $\Pi_H \psi_H \in W_H$ is to impose certain continuity of the functions $\psi_H \in X_H$ across the interelement boundaries. According to the definition of Π_K, we have $\Pi_K \bar{\psi}_i^K = \varphi_i^K$. Since φ_i^K is continuous across the element, the above requirement $\Pi_H \psi_H \in W_H$ is satisfied naturally due to the fact that in each node we only have one freedom (unknowns).

Denote by W_h the H^1-conforming linear finite element space over $\mathcal{M}_h$, and

$$W_h^0 := \{w_h \in W_h : \ w_h = 0 \text{ on } \partial\Omega_1/\Gamma\}. \tag{2.11}$$

We define the FE-MsFE approximation space $V_{h,H}$ as

$$V_{h,H} := \left\{v_{h,H} : \ v_{h,H}|_{\Omega_1} = v_h, \ v_{h,H}|_{\Omega_2} = v_H, \ \text{ where } v_h \in W_h^0, \ v_H \in X_H\right\}. \tag{2.12}$$

Note that $V_{h,H} \not\subset V$. We are now ready to define the FE-MsFEM inspired by the formulation (2.7): Find $u_{h,H} \in V_{h,H}$ such that

$$A_\beta(u_{h,H}, v_{h,H}) = F(v_{h,H}) \qquad \forall\, v_{h,H} \in V_{h,H}. \tag{2.13}$$

Remark 2.2. (a) If $\beta = 1$, then the bilinear form A_β is symmetric and, as a consequence, the stiffness matrix is symmetric as well. If $\beta \neq 1$, e.g., $\beta = -1$, then the method is nonsymmetric.

(b) The parameter $\rho > 0$ satisfies that $\rho \leq \epsilon$. In fact, it is chosen as ϵ in our later error analysis, while in practical computation it may be chosen as the mesh size h.

For further error analysis, we introduce several concepts related to the interface Γ and some discrete norms. Define the set of elements accompanying with the interface partition Γ_h (or Γ_H) as follows:

$$K_{\Gamma_h} := \{K \in \mathcal{M}_h : \ K \text{ has at least one edge/face in } \Gamma_h\}, \tag{2.14}$$

$$K_{\Gamma_H} := \{K \in \mathcal{M}_H : \ K \text{ has at least one edge/face in } \Gamma_H\}. \tag{2.15}$$

Clearly, from (H1), the number of elements in K_{Γ_h} is $O(\frac{1}{h^{n-1}})$ and the number of elements in K_{Γ_H} is $O(\frac{1}{H^{n-1}})$. Denote

$$\Omega_{\Gamma_H} = \cup\{K : \ K \in K_{\Gamma_H}\}, \qquad \Omega_{\Gamma_h} = \cup\{K : \ K \in K_{\Gamma_h}\}. \tag{2.16}$$

From the definition of Ω_Γ (see (2.1)), it is clear that $\Omega_{\Gamma_H}, \Omega_{\Gamma_h} \subset \Omega_\Gamma$ and $\operatorname{dist}(\Omega_{\Gamma_H}, \partial\Omega_\Gamma)$, $\operatorname{dist}(\Omega_{\Gamma_h}, \partial\Omega_\Gamma) \geq 2\epsilon$, respectively.

Denote

$$|v|_{1,H} = \left(\sum_{K\in\mathcal{M}_H} \|(\mathbf{a}^\epsilon)^{1/2}\nabla v\|^2_{L^2(K)}\right)^{1/2} \qquad \forall\, v \in \prod_{K\in\mathcal{M}_H} H^1(K).$$

We introduce the following energy norm and the broken norm on the space $V_{h,H}$:

$$\|v\|_E := \left(|v|^2_{1,H} + \left\|(\mathbf{a}^\epsilon)^{1/2}\nabla v\right\|^2_{L^2(\Omega_1)}\right)^{1/2}, \tag{2.17}$$

$$\begin{aligned}\|\!|v|\!\|_{1,h,H} := \Bigg(&\|v\|^2_E + \sum_{e\in\Gamma_h} \frac{\gamma_0}{\rho}\|[v]\|^2_{L^2(e)} + \sum_{e\in\Gamma_h} \gamma_1\,\rho\,\|[\mathbf{a}^\epsilon\nabla v\cdot\mathbf{n}]\|^2_{L^2(e)} \\ &+ \sum_{e\in\Gamma_h} \frac{\rho}{\gamma_0}\|\{\mathbf{a}^\epsilon\nabla v\cdot\mathbf{n}\}\|^2_{L^2(e)}\Bigg)^{1/2}.\end{aligned} \tag{2.18}$$

To conclude this section, we recall a trace inequality which will be used in this paper frequently. The proof is omitted here since it is a direct consequence of the standard trace inequality (cf. [13, Theorem 1.6.6, p. 39]) and the scaling argument (cf. [19]).

LEMMA 2.1. *Let K be an element of the triangulation $\{\mathcal{M}_H\}$ (or $\{\mathcal{M}_h\}$). Then, for any $v \in H^1(K)$, we have*

$$\|v\|_{L^2(\partial K)} \lesssim \operatorname{diam}(K)^{-1/2} \|v\|_{L^2(K)} + \|v\|_{L^2(K)}^{1/2} \|\nabla v\|_{L^2(K)}^{1/2}. \tag{2.19}$$

3. The homogenization results. In this section, we review some classical homogenization results for the elliptic problems and give an interior H^2 norm error estimate between the multiscale solution and the homogenized solution with first order corrector. Hereafter, we assume that $\mathbf{a}^\epsilon(x)$ has the form $\mathbf{a}(x/\epsilon)$, and $a_{ij}(y)$ is a periodic function in y in a unit cube Y. Moreover, we assume the following.

(H4) $a_{ij} \in W^{2,p}(Y)$ with $p > n$.

In our analysis, we need the facts that $\left|\nabla \chi^j(x/\epsilon)\right| \lesssim \epsilon^{-1}$, $\left|\nabla^2 \chi^j(x/\epsilon)\right| \lesssim \epsilon^{-2}$. Both follow from assumption (H4) (see [52, Theorem 15.1]). It is shown that under these assumptions (cf. [12, 51]), u_ϵ converges weakly in H^1 to the solution of the homogenized equation

$$\begin{cases} -\nabla \cdot (\mathbf{a}^* \nabla u_0(x)) = f(x) & \text{in } \Omega, \\ u_0(x) = 0 & \text{on } \partial\Omega, \end{cases} \tag{3.1}$$

where

$$a_{ij}^* = \frac{1}{|Y|} \int_Y a_{ik}(y) \left(\delta_{kj} + \frac{\partial \chi^j}{\partial y_k}(y) \right) \mathrm{d}y. \tag{3.2}$$

Here χ^j is the periodic solution of the cell problem

$$-\nabla_y \cdot (\mathbf{a}(y) \nabla_y \chi^j(y)) = \nabla_y \cdot (\mathbf{a}(y) \mathbf{e}_j), \quad j = 1, \ldots, n, \tag{3.3}$$

with zero mean, i.e., $\int_Y \chi^j dy = 0$, and $\mathbf{e}_j$ is the unit vector in the jth direction. The variational form of the problem (3.1) is to find $u_0 \in H_0^1(\Omega)$ such that

$$(\mathbf{a}^* \nabla u_0, \nabla v) = (f, v) \quad \forall\, v \in H_0^1(\Omega). \tag{3.4}$$

It can be shown that $\mathbf{a}^*$ is positive definite. Thus by the Lax–Milgram lemma, (3.4) has a unique solution. If $u_0 \in H^2(\Omega)$ and $f \in L^2(\Omega)$, from the regularity theory of elliptic equations, we have

$$\|u_0\|_{H^2(\Omega)} \leq C \|f\|_{L^2(\Omega)}. \tag{3.5}$$

Let θ_ϵ denote the boundary corrector which is the solution of

$$\begin{aligned} -\nabla \cdot (\mathbf{a}^\epsilon \nabla \theta_\epsilon) &= 0 && \text{in } \Omega, \\ \theta_\epsilon &= -\chi^j(x/\epsilon) \frac{\partial u_0(x)}{\partial x_j} && \text{on } \partial\Omega. \end{aligned} \tag{3.6}$$

From the maximum principle, we have

$$\|\theta_\epsilon\|_{L^\infty(\Omega)} \lesssim |u_0|_{W^{1,\infty}(\Omega)}. \tag{3.7}$$

In the following part, for the sake of convenience, we will set

$$u_1(x, x/\epsilon) = u_0(x) + \epsilon\chi^j(x/\epsilon)\frac{\partial u_0(x)}{\partial x_j}. \tag{3.8}$$

The following error estimates are known (cf., e.g., [15, 54, 17]).

THEOREM 3.1. *Assume that $u_0 \in H^2(\Omega)\cap W^{1,\infty}(\Omega)$. Then there exists a constant C independent of ϵ, the domain Ω, and the function f such that*

$$\|\nabla(u_\epsilon - u_1 - \epsilon\theta_\epsilon)\|_{L^2(\Omega)} \le C\epsilon\,|u_0|_{H^2(\Omega)}\,. \tag{3.9}$$

Moreover, if (H4) *holds, the boundary corrector θ_ϵ satisfies the estimate*

$$\|\epsilon\nabla\theta_\epsilon\|_{L^2(\Omega)} \le C\epsilon\,|u_0|_{H^2(\Omega)} + C\sqrt{\epsilon\,|\partial\Omega|}\,|u_0|_{W^{1,\infty}(\Omega)}\,, \tag{3.10}$$

where $|\partial\Omega|$ stands for the measure of the boundary $\partial\Omega$.

The following regularity estimate is an analogy of the classical interior estimate for elliptic equations in [43, Theorem 8.8, p. 183].

LEMMA 3.2. *Let $w \in H^1(D)$ be the weak solution of the elliptic equation*

$$-\nabla\cdot(\mathbf{B}(x)\nabla w) = g(x), \quad x \in D,$$

where $\mathbf{B}(x) = (b_{ij})_{n\times n}$ is uniformly Lipschitz continuous in D and satisfies

$$\alpha|\xi|^2 \le b_{ij}(x)\xi_i\xi_j \le \beta|\xi|^2 \quad \forall\xi \in \mathbb{R}^n,\, x \in \bar{D},$$

for some positive constants α and β, and $g \in L^2(D)$. Then for any subdomain $D' \subset\subset D$, we have $w \in H^2(D')$ and

$$|w|_{H^2(D')} \lesssim \left(\frac{1}{d'} + \frac{1}{r}\right)|w|_{H^1(D)} + \|g\|_{L^2(D)}\,, \tag{3.11}$$

where $d' = dist(D', \partial D)$ and r is a constant such that $|\mathbf{B}(x) - \mathbf{B}(y)| \le \frac{1}{r}|x - y|$.

Remark 3.1. The H^2 norm interior estimate for elliptic equations is well known in the literature. The importance in the above estimate is the explicit dependence of the bound on d' and r, which is crucial in our analysis.

Proof. First, we define the difference quotient as follows:

$$\Delta^h v(x) = \Delta_k^h v = \frac{v(x + h\mathbf{e}_k) - v(x)}{h},$$

where $\mathbf{e}_k$, $1 \le k \le n$, is the unit coordinate vector in the x_k direction. And then, following the proof presented in [43, Theorem 8.8, p. 183], we obtain

$$\begin{aligned}
&\alpha\int_D \left|\xi\nabla(\Delta^h w)\right|^2 \mathrm{d}x \\
&\quad\le \left(\frac{1}{r}\|\nabla w\|_{L^2(D)} + \|g\|_{L^2(D)}\right)\left(\left\|\xi\nabla(\Delta^h w)\right\|_{L^2(D)} + 2\left\|\Delta^h w\nabla\xi\right\|_{L^2(D)}\right) \\
&\qquad + C\left\|\xi\nabla(\Delta^h w)\right\|_{L^2(D)}\left\|\Delta^h w\nabla\xi\right\|_{L^2(D)},
\end{aligned}$$

where $\xi \in C_0^1(D)$ is the cut-off function such that $0 \le \xi \le 1$, $\xi = 1$ in D', and $|\nabla\xi| \le 2/d'$ in D. By use of Young's inequality and Lemma 7.23 in [43, p. 168], it

follows that

$$\begin{aligned}\left\|\xi\Delta^h(\nabla w)\right\|_{L^2(D)} &\le C\left(\frac{1}{r}\left\|\nabla w\right\|_{L^2(D)} + \left\|g\right\|_{L^2(D)} + \left\|\Delta^h w\nabla\xi\right\|_{L^2(D)}\right)\\ &\le C\left(\frac{1}{r}+\frac{1}{d'}\right)\left\|\nabla w\right\|_{L^2(D)} + C\left\|g\right\|_{L^2(D)}.\end{aligned}$$

By Lemma 7.24 in [43, p. 169] we obtain $\nabla w \in H^1(D')$, so that $w \in H^2(D')$ and the estimate (3.11) holds. This completes the proof. ∎

Utilizing the above H^2 interior estimate to equation $-\nabla\cdot(\mathbf{a}^*\nabla u_{0x_j}) = f_{x_j}$, $j = 1,\dots,n$, where x_j is the coordinate variable in the jth direction, we obtain the following lemma.

LEMMA 3.3. *Let D be a subdomain of Ω, and let u_0 be the solution to* (3.4). *Assume that $u_0 \in H^2(D)$ and $f \in H^1(D)$. Then for any subdomain $D' \subset\subset D$ with $d' = dist(D', \partial D)$, we have*

$$|u_0|_{H^3(D')} \lesssim \frac{1}{d'}\left|u_0\right|_{H^2(D)} + \left\|\nabla f\right\|_{L^2(D)}. \tag{3.12}$$

Further, by use of the H^2 interior estimate (3.11), we obtain an H^2 seminorm interior estimate of the error $u_\epsilon - u_1$ in the narrow domain Ω_Γ.

THEOREM 3.4. *Assume that $u_0 \in H^2(\Omega)\cap W^{1,\infty}(\Omega)$, $f|_{\Omega_\Gamma} \in H^1(\Omega_\Gamma)$, and* (H4) *hold. Then for any subdomain $\Omega' \subset\subset \Omega_\Gamma$ with dist$(\Omega', \partial\Omega_\Gamma) \ge 2\epsilon$, we have*

$$|u_\epsilon - u_1|_{H^2(\Omega')} \lesssim |u_0|_{H^2(\Omega)} + \epsilon\left\|\nabla f\right\|_{L^2(\Omega_\Gamma)} + \frac{1}{\sqrt{\epsilon}}\left|u_0\right|_{W^{1,\infty}(\Omega)}. \tag{3.13}$$

Proof. It is shown that, for any $\varphi \in H_0^1(\Omega)$ (see [15, p. 550] or [16, p. 125]),

$$\begin{aligned}&(\mathbf{a}(x/\epsilon)\nabla(u_\epsilon - u_1), \nabla\varphi)_\Omega\\ &= (\mathbf{a}^*\nabla u_0, \nabla\varphi)_\Omega - \left(\mathbf{a}(x/\epsilon)\nabla\left(u_0 + \epsilon\chi^k(x/\epsilon)\frac{\partial u_0}{\partial x_k}\right), \nabla\varphi\right)_\Omega\\ &= \epsilon\int_\Omega a_{ij}(x/\epsilon)\chi^k(x/\epsilon)\frac{\partial^2 u_0}{\partial x_j\partial x_k}\frac{\partial\varphi}{\partial x_i} - \epsilon\int_\Omega \alpha_{ij}^k(x/\epsilon)\frac{\partial^2 u_0}{\partial x_j\partial x_k}\frac{\partial\varphi}{\partial x_i},\end{aligned} \tag{3.14}$$

where $\alpha^k(x/\epsilon) = (\alpha_{ij}^k(x/\epsilon))$ are skew-symmetric matrices which satisfy that (see [51, p. 6])

$$\frac{\partial}{\partial y_j}(\alpha_{ij}^k(y)) = a_{ik}^* - a_{ij}\left(\delta_{kj} + \frac{\partial\chi^k(y)}{\partial y_j}\right), \qquad \int_Y \alpha_{ij}^k(y)\,\mathrm{d}y = 0.$$

Note that we have used the fact that $\left(\frac{\partial}{\partial x_j}\left(\alpha_{ij}^k(x/\epsilon)\frac{\partial}{\partial x_k}u_0(x)\right)\right)_{i=1}^n$ is divergence free to derive the last equality in (3.14). According to the assumption, there exists a subdomain $\Omega_c \subset\subset \Omega_\Gamma$ such that $\Omega' \subset\subset \Omega_c$ with $d_0 = \mathrm{dist}(\Omega', \partial\Omega_c) \ge \epsilon$ and $d_1 = \mathrm{dist}(\Omega_c, \partial\Omega_\Gamma) \ge \epsilon$. From (3.14), it follows that in Ω_c,

$$\nabla\cdot(\mathbf{a}(x/\epsilon)\nabla(u_\epsilon - u_1)) = \epsilon\frac{\partial}{\partial x_i}\left(a_{ij}(x/\epsilon)\chi^k(x/\epsilon)\frac{\partial^2 u_0}{\partial x_j\partial x_k} - \alpha_{ij}^k(x/\epsilon)\frac{\partial^2 u_0}{\partial x_j\partial x_k}\right).$$

Thus, from Lemma 3.2, it follows that

$$|u_\epsilon - u_1|_{H^2(\Omega')} \lesssim \left(\frac{1}{d_0}+\frac{1}{\epsilon}\right)|u_\epsilon - u_1|_{H^1(\Omega_c)} + |u_0|_{H^2(\Omega_c)} + \epsilon\,|u_0|_{H^3(\Omega_c)}.$$

Hence, from Theorem 3.1 and Lemma 3.3, the result (3.13) follows immediately. ∎

The following lemma gives a local H^2 seminorm estimate for u_1 in Ω_{Γ_h}, which will be used in the convergence analysis.

LEMMA 3.5. *Assume that* $u_0 \in H^2(\Omega) \cap W^{1,\infty}(\Omega)$, $f|_{\Omega_\Gamma} \in H^1(\Omega_\Gamma)$, *and* (H4) *hold. Then*

$$|u_1|_{H^2(\Omega_{\Gamma_h})} \lesssim |u_0|_{H^2(\Omega_\Gamma)} + \epsilon^{-1} h^{1/2} |u_0|_{W^{1,\infty}(\Omega_{\Gamma_h})} + \epsilon \|\nabla f\|_{L^2(\Omega_\Gamma)}, \tag{3.15}$$

where Ω_{Γ_h} *and* Ω_Γ *are defined in* (2.16) *and* (2.1), *respectively.*

Proof. It is easy to see that

$$\begin{aligned} |u_1|_{H^2(\Omega_{\Gamma_h})} &\le |u_0|_{H^2(\Omega_{\Gamma_h})} + \left| \epsilon \chi^j(x/\epsilon) \frac{\partial u_0}{\partial x_j} \right|_{H^2(\Omega_{\Gamma_h})} \\ &\lesssim |u_0|_{H^2(\Omega_{\Gamma_h})} + \epsilon^{-1} \left| \frac{\partial u_0}{\partial x_j} \right|_{L^2(\Omega_{\Gamma_h})} + \epsilon |u_0|_{H^3(\Omega_{\Gamma_h})}. \end{aligned}$$

Thus, from Lemma 3.3, it follows that

$$|u_1|_{H^2(\Omega_{\Gamma_h})} \lesssim |u_0|_{H^2(\Omega_\Gamma)} + \epsilon \|\nabla f\|_{L^2(\Omega_\Gamma)} + \epsilon^{-1} |\Omega_{\Gamma_h}|^{1/2} \|\nabla u_0\|_{L^\infty(\Omega_{\Gamma_h})},$$

which, combining with the fact that $|\Omega_{\Gamma_h}| = O(h)$, yields the result (3.15). ∎

We conclude this section with an interior gradient estimate whose proof follows easily from Avellaneda and Lin [7, Lemma 16] and is omitted (see also [31, Proposition C.1]).

LEMMA 3.6. *Let* $D' \subset\subset D$ *be bounded domains with* $d' = dist(D', \partial D) > 0$, *and let* w_ϵ *satisfy* $\nabla \cdot (\mathbf{a}^\epsilon \nabla w_\epsilon) = 0$ *in* D. *Then we have*

$$\|\nabla w_\epsilon\|_{L^\infty(D')} \lesssim \frac{1}{d'} \|w_\epsilon\|_{L^\infty(D)}.$$

4. Approximation properties of the FE-MsFE space $V_{h,H}$. In this section, we give some approximation properties for the oversampling MsFE space X_H and the linear finite element space W_h, respectively. In our analysis, we assume that the coarse mesh size H, the fine mesh size h, and the parameter ϵ satisfy the following condition.

(H5) $h < \epsilon < H$.

Hence when there is no influence on the dominance order of error estimates, we use frequently $\epsilon/H < 1$ and $h/\epsilon < 1$ to simplify the intermediate results.

Recall that ψ_i^S, $i = 1, \ldots, n+1$, are defined by (2.8). By the asymptotic expansion (cf. [31, 48]), we know that

$$\psi_i^S = \varphi_i^S + \epsilon \chi^j(x/\epsilon) \frac{\partial \varphi_i^S}{\partial x_j} + \epsilon \eta^j(x) \frac{\partial \varphi_i^S}{\partial x_j},$$

with η^j being the solution of

$$-\nabla \cdot (\mathbf{a}^\epsilon \nabla \eta^j) = 0 \quad \text{in } S, \qquad \eta^j\big|_{\partial S} = -\chi^j(x/\epsilon). \tag{4.1}$$

Therefore the oversampling MsFE basis function $\bar{\psi}_i^K$ (see (2.9)) has the following expansion:

$$\bar{\psi}_i^K = \varphi_i^K + \epsilon \chi^j(x/\epsilon) \frac{\partial \varphi_i^K}{\partial x_j} + \epsilon \eta^j(x) \frac{\partial \varphi_i^K}{\partial x_j}, \tag{4.2}$$

where $\{\varphi_i^K\}_{i=1}^{n+1}$ are the basis of $P_1(K)$ satisfying $\varphi_i^K(x_j^K) = \delta_{ij}$. Thus, the distance $d_{S(K)} = \operatorname{dist}(K, \partial S)$ is determined by the thickness of the boundary layer of η^j. Numerically, it has been observed that the boundary layer is about $O(\epsilon)$ thick [46]. In our analysis, we assume that the distance $d_{S(K)}$ satisfies the following condition.

(H6) $d_{S(K)} \geq C_0 H_K$, where C_0 is the constant from (H3).

By the maximum principle we have

$$\left\|\eta^j\right\|_{L^\infty(S)} \leq \left|\chi^j\right|_{L^\infty(S)} \lesssim 1, \tag{4.3}$$

which together with Lemma 3.6 implies that

$$\left\|\nabla\eta^j\right\|_{L^\infty(K)} \lesssim d_{S(K)}^{-1} \left\|\eta^j\right\|_{L^\infty(S)} \lesssim \frac{1}{d_{S(K)}}. \tag{4.4}$$

4.1. Approximation properties of oversampling MsFE space X_H. We first give some approximation properties of the oversampling MsFE space OMS(K).

LEMMA 4.1. *Assume that* (H5) *and* (H6) *hold. Then, for any $K \in \mathcal{M}_H$, there exists $\phi_H^K \in OMS(K)$ such that the following estimates hold:*

$$\left|u_1 - \phi_H^K\right|_{H^1(K)} \lesssim H_K \left|u_0\right|_{H^2(K)} + \epsilon H_K^{n/2} d_{S(K)}^{-1} \left|u_0\right|_{W^{1,\infty}(K)}, \tag{4.5}$$

$$\left\|u_1 - \phi_H^K\right\|_{L^2(K)} \lesssim H_K^2 \left|u_0\right|_{H^2(K)} + \epsilon H_K^{n/2} \left|u_0\right|_{W^{1,\infty}(K)}, \tag{4.6}$$

$$\left|u_1 - \phi_H^K\right|_{H^2(K)} \lesssim \epsilon^{-1} H_K \left|u_0\right|_{H^2(K)} + H_K^{n/2} d_{S(K)}^{-1} \left|u_0\right|_{W^{1,\infty}(K)} + \epsilon \left|u_0\right|_{H^3(K)}. \tag{4.7}$$

Proof. We take

$$\phi_H^K = \sum_{x_i^K \text{ node of } K} u_0(x_i^K)\bar{\psi}_i^K(x). \tag{4.8}$$

Then

$$\Pi_K \phi_H^K = I_H u_0,$$

where $I_H : C(\bar{\Omega}_2) \to W_H$ is the standard Lagrange interpolation operator over linear finite element space. By (4.2), we have the asymptotic expansion

$$\phi_H^K = I_H u_0 + \epsilon\chi^j(x/\epsilon)\frac{\partial(I_H u_0)}{\partial x_j} + \epsilon\eta^j \frac{\partial(I_H u_0)}{\partial x_j} \quad \text{on } K. \tag{4.9}$$

Therefore, from (4.4), we have

$$\begin{aligned}\left\|\nabla\Big(\phi_H^K - I_H u_0 - \epsilon\chi^j(x/\epsilon)\frac{\partial(I_H u_0)}{\partial x_j}\Big)\right\|_{L^2(K)} &= \left\|\epsilon\nabla\eta^j \frac{\partial(I_H u_0)}{\partial x_j}\right\|_{L^2(K)} \\ &\lesssim \epsilon d_{S(K)}^{-1} \left|u_0\right|_{W^{1,\infty}(K)} |K|^{1/2} \\ &\lesssim \epsilon H_K^{n/2} d_{S(K)}^{-1} \left|u_0\right|_{W^{1,\infty}(K)}.\end{aligned} \tag{4.10}$$

Further, since

$$\begin{aligned}&\left\|\nabla(u_0 - I_H u_0)\right\|_{L^2(K)} \lesssim H_K \left|u_0\right|_{H^2(K)}, \\ &\left\|\epsilon\nabla\Big(\chi^j(x/\epsilon)\frac{\partial(u_0 - I_H u_0)}{\partial x_j}\Big)\right\|_{L^2(K)} \lesssim (H_K + \epsilon)\left|u_0\right|_{H^2(K)},\end{aligned}$$

we obtain the result (4.5) immediately by using assumption (H5).

To prove the estimate (4.6), we first notice from (4.3) that

$$\left\|\phi_H^K - \left(I_H u_0 + \epsilon\chi^j(x/\epsilon)\frac{\partial(I_H u_0)}{\partial x_j}\right)\right\|_{L^2(K)} \lesssim \epsilon H_K^{n/2}\,|u_0|_{W^{1,\infty}(K)}\,. \tag{4.11}$$

Further, since

$$\begin{aligned}\|u_0 - I_H u_0\|_{L^2(K)} &\lesssim H_K^2\,|u_0|_{H^2(K)}\,,\\ \left\|\epsilon\chi^j(x/\epsilon)\frac{\partial(u_0 - I_H u_0)}{\partial x_j}\right\|_{L^2(K)} &\lesssim \epsilon H_K\,|u_0|_{H^2(K)}\,,\end{aligned}$$

we obtain the result (4.6) immediately by using assumption (H5).

To prove the estimate (4.7), it is easy to see that

$$|u_0 - I_H u_0|_{H^2(K)} = |u_0|_{H^2(K)}\,, \tag{4.12}$$

$$\left|\epsilon\chi^j(x/\epsilon)\frac{\partial(u_0 - I_H u_0)}{\partial x_j}\right|_{H^2(K)} \lesssim (1+\epsilon^{-1}H_K)\,|u_0|_{H^2(K)} + \epsilon\,|u_0|_{H^3(K)}\,. \tag{4.13}$$

From assumption (H6), there exists a simplex D such that

$$K \subset\subset D \subset\subset S \text{ and } \operatorname{dist}(K,\partial D) \gtrsim H_K,\ \operatorname{dist}(D,\partial S) \gtrsim d_{S(K)},\ \operatorname{diam}(D) \lesssim H_K.$$

From Lemma 3.2 and (4.4), we have

$$\left|\eta^j\right|_{H^2(K)} \lesssim \epsilon^{-1}\left|\eta^j\right|_{H^1(D)} \lesssim \epsilon^{-1}H_K^{n/2}\left\|\nabla\eta^j\right\|_{L^\infty(D)} \lesssim \epsilon^{-1}H_K^{n/2}d_{S(K)}^{-1}. \tag{4.14}$$

Hence

$$\left|\epsilon\eta^j\frac{\partial(I_H u_0)}{\partial x_j}\right|_{H^2(K)} \lesssim H_K^{n/2}d_{S(K)}^{-1}\,\|\nabla u_0\|_{L^\infty(K)}\,, \tag{4.15}$$

which, combining with (H5), (4.12), and (4.13), yields (4.7) immediately. This completes the proof. ∎

Denote by $d = \min_{K\in\mathcal{M}_H} d_{S(K)}$. From Lemma 4.1, we have the following local approximation estimates in K_{Γ_H}.

LEMMA 4.2. *There exists $\psi_H \in X_H$ such that*

$$\left(\sum_{K\in K_{\Gamma_H}} \|u_1 - \psi_H\|_{L^2(K)}^2\right)^{1/2} \lesssim H^2\,|u_0|_{H^2(\Omega_{\Gamma_H})} + \epsilon\sqrt{H}\,|u_0|_{W^{1,\infty}(\Omega_{\Gamma_H})}\,, \tag{4.16}$$

$$\left(\sum_{K\in K_{\Gamma_H}} |u_1 - \psi_H|_{H^1(K)}^2\right)^{1/2} \lesssim H\,|u_0|_{H^2(\Omega_{\Gamma_H})} + \frac{\epsilon}{d}\sqrt{H}\,|u_0|_{W^{1,\infty}(\Omega_{\Gamma_H})}\,, \tag{4.17}$$

$$\begin{aligned}&\left(\sum_{K\in K_{\Gamma_H}} |u_1 - \psi_H|_{H^2(K)}^2\right)^{1/2}\\ &\qquad\lesssim \frac{H}{\epsilon}\,|u_0|_{H^2(\Omega_\Gamma)} + \epsilon\,\|\nabla f\|_{L^2(\Omega_\Gamma)} + \frac{\sqrt{H}}{d}\,|u_0|_{W^{1,\infty}(\Omega_{\Gamma_H})}\,,\end{aligned} \tag{4.18}$$

where K_{Γ_H} and Ω_{Γ_H} are defined in (2.15) *and* (2.16)*, respectively.*

Proof. Let ϕ_H^K be defined in Lemma 4.1 by (4.8). We define ψ_H by

$$\psi_H|_K = \phi_H^K \quad \forall K \in \mathcal{M}_H. \tag{4.19}$$

Clearly, $\psi_H \in X_H$. From (4.6), it follows that

$$\begin{aligned}\sum_{K\in K_{\Gamma_H}} \|u_1-\psi_H\|^2_{L^2(K)} &\lesssim H^4\,|u_0|^2_{H^2(\Omega_{\Gamma_H})} + \epsilon^2 H^n\,|u_0|^2_{W^{1,\infty}(\Omega_{\Gamma_H})} \sum_{K\in K_{\Gamma_H}} 1\\ &\lesssim H^4\,|u_0|^2_{H^2(\Omega_{\Gamma_H})} + \epsilon^2 H\,|u_0|^2_{W^{1,\infty}(\Omega_{\Gamma_H})},\end{aligned}$$

which yields (4.16) immediately. Note that here we have used the fact that the number of elements in K_{Γ_H} is $O(\frac{|\Gamma|}{H^{n-1}})$. Similarly, (4.17) follows from (4.5).

It remains to prove (4.18). From (4.7) and Lemma 3.3, it follows that

$$\begin{aligned}&\sum_{K\in K_{\Gamma_H}} |u_1-\psi_H|^2_{H^2(K)}\\ &\lesssim \epsilon^{-2}H^2\,|u_0|^2_{H^2(\Omega_{\Gamma_H})} + \epsilon^2\,|u_0|^2_{H^3(\Omega_{\Gamma_H})} + H^n d^{-2}\,|u_0|^2_{W^{1,\infty}(\Omega_{\Gamma_H})} \sum_{K\in K_{\Gamma_H}} 1\\ &\lesssim \left(\epsilon^{-2}H^2+1\right)|u_0|^2_{H^2(\Omega_\Gamma)} + \epsilon^2\,\|\nabla f\|^2_{L^2(\Omega_\Gamma)} + Hd^{-2}\,|u_0|^2_{W^{1,\infty}(\Omega_{\Gamma_H})},\end{aligned}$$

which yields (4.18) immediately by using assumption (H5). The proof is complete. ∎

From Lemma 4.1 and Theorem 3.1, by taking the same ψ_H in Lemma 4.2, we also have the following result, which gives an approximation estimate of the space X_H (cf. [31, 29, 16]).

LEMMA 4.3. *There exists $\psi_H \in X_H$ such that*

$$|u_\epsilon-\psi_H|_{1,H} \lesssim H\,|u_0|_{H^2(\Omega)} + \frac{\epsilon}{d}\,|u_0|_{W^{1,\infty}(\Omega_2)} + \sqrt{\epsilon}\,|u_0|_{W^{1,\infty}(\Omega)}, \tag{4.20}$$

where $d = \min_{K\in\mathcal{M}_H} d_{S(K)}$.

4.2. Approximation properties of linear finite element space W_h. Since $|\Omega_1|$, the area/volume of Ω_1, may be small, we prefer estimates with explicit dependence on it. To attain this aim, we use the Scott–Zhang interpolation instead of the standard finite element interpolation in this subsection.

We first introduce the Scott–Zhang interpolation operator $Z_h : H^1(\Omega_1) \to W_h$. For any node z in $\mathcal{M}_h$, let $\phi_z(x)$ be the nodal basis function associated with z and let e_z be an edge/face with one vertex at z. Then the Scott–Zhang interpolation operator is defined as [60]

$$Z_h v = \sum_{\text{node } z \text{ in} \mathcal{M}_h} \left(\int_{e_z} \psi_z v\right)\phi_z \quad \forall\, v \in H^1(\Omega_1), \tag{4.21}$$

where $\psi_z(x)$ is a linear function that satisfies $\int_{e_z}\psi_z(x)w(x) = w(z)$ for any linear function $w(x)$ on e_z. Suppose $e_z \subset \partial\Omega_1$ for $z \in \partial\Omega_1$ and $e_z \subset \Omega_{\Gamma_h}$ for $z \in \Omega_{\Gamma_h}$, where Ω_{Γ_h} is defined in (2.16). It is easy to check that $\|\psi_z\|_{L^\infty(\Omega_z)} \lesssim h_{e_z}^{-(n-1)}$, $\|\nabla\psi_z\|_{L^\infty(\Omega_z)} \lesssim h_{e_z}^{-n}$, where $\Omega_z := \text{supp}(\phi_z)$, and

$$Z_h v = v \quad \forall\, v \in W_h. \tag{4.22}$$

This operator enjoys the following stability and interpolation estimates (see [60]).

LEMMA 4.4. *For any $K \in \mathcal{M}_h$, we have*

$$\|Z_h v\|_{L^\infty(K)} \lesssim \|v\|_{L^\infty(\tilde{K})}, \ \|\nabla Z_h v\|_{L^p(K)} \lesssim \|\nabla v\|_{L^p(\tilde{K})}, \quad p = 2, \infty. \tag{4.23}$$

$$\|v - Z_h v\|_{L^2(K)} + h_K \|v - Z_h v\|_{H^1(K)} \lesssim h_K^2 |v|_{H^2(\tilde{K})}, \tag{4.24}$$

$$\|v - Z_h v\|_{L^\infty(K)} \lesssim h_K^2 |v|_{W^{2,\infty}(\tilde{K})}, \tag{4.25}$$

where $\tilde{K}$ is the union of all elements in $\mathcal{M}_h$ having nonempty intersection with K.

Moreover, we need the following error estimate between u_1 and its Scott–Zhang interpolant, which uses only the regularity of the homogenization solution u_0.

LEMMA 4.5. *For any $K \in \mathcal{M}_h$, we have*

$$|u_1 - Z_h u_1|_{H^1(K)} \lesssim \left(h_K + \epsilon + \frac{h_K^2}{\epsilon}\right) |u_0|_{H^2(\tilde{K})} + \frac{h_K^{n/2+1}}{\epsilon} |u_0|_{W^{1,\infty}(K)}, \tag{4.26}$$

where $\tilde{K}$ is defined in Lemma 4.4.

Proof. Denote by $v_j := \frac{\partial u_0}{\partial x_j}$. It is easy to see that

$$\begin{aligned}
|u_1 - Z_h u_1|_{H^1(K)} &\lesssim |u_0 - Z_h u_0|_{H^1(K)} + \epsilon \left|\chi^j(x/\epsilon) v_j - Z_h\left(\chi^j(x/\epsilon) v_j\right)\right|_{H^1(K)} \\
&\lesssim |u_0 - Z_h u_0|_{H^1(K)} + \epsilon \left|(\chi^j(x/\epsilon) - Z_h \chi^j(x/\epsilon)) v_j\right|_{H^1(K)} \\
&\quad + \epsilon \left|Z_h \chi^j(x/\epsilon)\left(v_j - \langle v_j\rangle_K\right)\right|_{H^1(K)} \\
&\quad + \epsilon \left|Z_h \chi^j(x/\epsilon) \langle v_j\rangle_K - Z_h\left(\chi^j(x/\epsilon) v_j\right)\right|_{H^1(K)} \\
&:= \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathrm{IV},
\end{aligned}$$

where $\langle\cdot\rangle_K = \frac{1}{|K|}\int_K (\cdot)\,\mathrm{d}x$. From (4.24), we have

$$\mathrm{I} \lesssim h_K |u_0|_{H^2(\tilde{K})}. \tag{4.27}$$

From the facts that $\left|\nabla \chi^j(x/\epsilon)\right| \lesssim \epsilon^{-1}$, $\left|\nabla^2 \chi^j(x/\epsilon)\right| \lesssim \epsilon^{-2}$ (both of which follow from assumption (H4); see [52, Theorem 15.1]), and (4.23)–(4.25), we have

$$\mathrm{II} \lesssim \frac{h_K^{n/2+1}}{\epsilon} |u_0|_{W^{1,\infty}(K)} + \frac{h_K^2}{\epsilon} |u_0|_{H^2(K)}, \tag{4.28}$$

$$\mathrm{III} \lesssim (h_K + \epsilon) |u_0|_{H^2(K)}, \tag{4.29}$$

where we have used the Poincaré inequality to derive the second inequality. It remains to estimate IV. According to the definition of the Scott–Zhang interpolation, we have

$$\begin{aligned}
\mathrm{IV} &= \epsilon \left| \sum_{\text{node } z \in K} \phi_z(x) \int_{e_z} \psi_z \chi^j(x/\epsilon) \left(\langle v_j\rangle_K - v_j\right) \right|_{H^1(K)} \\
&\lesssim \epsilon h_K^{n/2} \left\| \sum_{\text{node } z \in K} \nabla\phi_z(x) \int_{e_z} \psi_z \chi^j(x/\epsilon) \left(\langle v_j\rangle_K - v_j\right) \right\|_{L^\infty(K)} \\
&\lesssim \epsilon h_K^{-n/2} \sum_{\text{node } z \in K} \sum_{j=1}^{n} \int_{e_z} \left|\langle v_j\rangle_K - v_j\right|.
\end{aligned}$$

For each node $z \in K$, there exist a number $M_z \lesssim 1$ and a sequence of elements $K_{z,m} \subset \tilde{K}$, $m = 1, \ldots, M_z$, such that $K_{z,1} = K$; $K_{z,i}$ and $K_{z,i+1}$ have a common

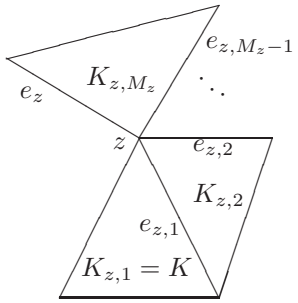

FIG. 2. *Illustration of the elements* $K_{z,m}$, $m = 1, \ldots, M_z$.

edge/face $e_{z,i}$, $i = 1, \ldots, M_z - 1$; and $e_{z,M_z} := e_z$ is an edge/face of K_{z,M_z}. See Figure 2 for an illustration of the elements $K_{z,m}, m = 1, \ldots, M_z$. Clearly, we have

$$\begin{aligned}
&\int_{e_z} \left|\langle v_j\rangle_K - v_j\right| \le \int_{e_{z,M_z}} \left|\langle v_j\rangle_{K_{z,M_z}} - v_j\right| + h_{e_{z,M_z}} \left|\langle v_j\rangle_{K_{z,1}} - \langle v_j\rangle_{K_{z,M_z}}\right| \\
&= \int_{e_{z,M_z}} \left|\langle v_j\rangle_{K_{z,M_z}} - v_j\right| \\
&\quad + h_{e_{z,M_z}} \left| \sum_{m=1}^{M_z-1} \left(\frac{1}{h_{e_{z,m}}} \int_{e_{z,m}} \left(\langle v_j\rangle_{K_{z,m}} - v_j\right) - \frac{1}{h_{e_{z,m}}} \int_{e_{z,m}} \left(\langle v_j\rangle_{K_{z,m+1}} - v_j\right) \right) \right| \\
&\lesssim \sum_{m=1}^{M_z} \int_{\partial K_{z,m}} \left|\langle v_j\rangle_{K_{z,m}} - v_j\right| \\
&\lesssim h_K^{(n-1)/2} \sum_{m=1}^{M_z} \left\|\langle v_j\rangle_{K_{z,m}} - v_j\right\|_{L^2(\partial K_{z,m})}.
\end{aligned}$$

Thus, from Lemma 2.1 and the Poincaré inequality, we obtain

$$\begin{aligned}
\mathrm{IV} &\lesssim \epsilon h_K^{-1/2} \sum_{\text{node } z\in K} \sum_{j=1}^{n} \sum_{m=1}^{M_z} \Bigg(h_K^{-1/2} \left\|\langle v_j\rangle_{K_{z,m}} - v_j\right\|_{L^2(K_{z,m})} \\
&\qquad + \left\|\langle v_j\rangle_{K_{z,m}} - v_j\right\|_{L^2(K_{z,m})}^{1/2} \left\|\nabla\left(\langle v_j\rangle_{K_{z,m}} - v_j\right)\right\|_{L^2(K_{z,m})}^{1/2} \Bigg) \\
&\lesssim \epsilon \sum_{j=1}^{n} \|\nabla v_j\|_{L^2(\tilde{K})} \lesssim \epsilon \, |u_0|_{H^2(\tilde{K})},
\end{aligned}$$

which, combining with (4.27)–(4.29), yields the result immediately. ∎

Further, we have the following result, which gives H^1 approximation estimates of the space W_h.

LEMMA 4.6. *The following estimates hold for* $h \lesssim \epsilon$*:*

$$|u_1 - Z_h u_1|_{H^1(\Omega_1)} \lesssim \epsilon \, |u_0|_{H^2(\Omega_1)} + \frac{h}{\epsilon} \, |\Omega_1|^{1/2} \, |u_0|_{W^{1,\infty}(\Omega_1)}, \tag{4.30}$$

$$|u_1 - Z_h u_1|_{H^1(\Omega_{\Gamma_h})} \lesssim h \, |u_0|_{H^2(\Omega_\Gamma)} + \frac{h^{3/2}}{\epsilon} \, |u_0|_{W^{1,\infty}(\Omega_{\Gamma_h})} + \epsilon h \|\nabla f\|_{L^2(\Omega_\Gamma)}, \tag{4.31}$$

where Ω_Γ *is defined in* (2.1) *and* Ω_{Γ_h} *is defined in* (2.16).

Proof. Inequality (4.30) is a direct consequence of Lemma 4.5. From Lemma 4.4, it follows that

$$|u_1 - Z_h u_1|_{H^1(\Omega_{\Gamma_h})} \lesssim h\,|u_1|_{H^2(\Omega_{\Gamma_h})},$$

which together with Lemma 3.5 yields (4.31) immediately. ∎

5. Error estimates for the FE-MsFEM. In this section we derive the H^1-error estimate for the FE-MsFEM in the case where $\beta = 1$. For other cases such that $\beta = 0, -1$, the analysis is similar and is omitted here. Since the convergence analysis is only done for the periodic coefficient case, we will fix $\rho = \epsilon$ in the later analysis.

We first recall a stability estimate for Π_H which can be proved by the method in [15, Lemma 4.10] (see also [16, Lemma 9.8]).

LEMMA 5.1. *Assume that* (H6) *holds. Then there exists a constant γ independent of H and ϵ such that if $\epsilon/H_K \le \gamma$ for all $K \in \mathcal{M}_H$, then the following estimate is valid:*

$$\|\nabla \Pi_H w_H\|_{L^2(K)} \lesssim \|\nabla w_H\|_{L^2(K)} \qquad \forall\, w_H \in X_H.$$

Proof. By (4.2), we have the asymptotic expansion

$$w_H = w_H^0 + \epsilon \chi^k(x/\epsilon)\frac{\partial w_H^0}{\partial x_k} + \epsilon \eta^k \frac{\partial w_H^0}{\partial x_k} \quad \text{on } K, \tag{5.1}$$

where $w_H^0 = \Pi_H w_H \in W_H$. By simple calculations

$$a_{ij}(x/\epsilon)\frac{\partial w_H}{\partial x_j} = a_{ij}^* \frac{\partial w_H^0}{\partial x_j} - G_i^k(x/\epsilon)\frac{\partial w_H^0}{\partial x_k} + \epsilon a_{ij}(x/\epsilon)\frac{\partial \eta^k}{\partial x_j}\frac{\partial w_H^0}{\partial x_k}, \tag{5.2}$$

where $G_i^k(y) = a_{ik}^* - a_{ij}(y)\big(\delta_{kj} + \frac{\partial \chi^k(y)}{\partial y_j}\big)$ satisfies

$$\int_Y G_i^k(y)\,\mathrm{d}y = 0 \qquad \text{and} \qquad \frac{\partial G_i^k}{\partial y_i} = 0.$$

Multiplying (5.2) by ∇w_H^0 and integrating over K, we obtain

$$\begin{aligned}\int_K a_{ij}^* \frac{\partial w_H^0}{\partial x_j}\frac{\partial w_H^0}{\partial x_i}\,\mathrm{d}x = &\int_K a_{ij}(x/\epsilon)\frac{\partial w_H}{\partial x_j}\frac{\partial w_H^0}{\partial x_i}\,\mathrm{d}x + \int_K G_i^k(x/\epsilon)\frac{\partial w_H^0}{\partial x_k}\frac{\partial w_H^0}{\partial x_i}\,\mathrm{d}x \\ &- \epsilon \int_K a_{ij}(x/\epsilon)\frac{\partial \eta^k}{\partial x_j}\frac{\partial w_H^0}{\partial x_k}\frac{\partial w_H^0}{\partial x_i}\,\mathrm{d}x.\end{aligned}$$

From (4.4) and (H6), we have

$$\left|\epsilon \int_K a_{ij}(x/\epsilon)\frac{\partial \eta^k}{\partial x_j}\frac{\partial w_H^0}{\partial x_k}\frac{\partial w_H^0}{\partial x_i}\,\mathrm{d}x\right| \lesssim \frac{\epsilon}{d_{S(K)}}\left\|\nabla w_H^0\right\|_{L^2(K)}^2 \lesssim \frac{\epsilon}{H_K}\left\|\nabla w_H^0\right\|_{L^2(K)}^2.$$

By the result in [15, Lemma 4.9] (see also [31, Lemma 3.2]), we have

$$\left|\int_K G_i^k(x/\epsilon)\frac{\partial w_H^0}{\partial x_k}\frac{\partial w_H^0}{\partial x_i}\,\mathrm{d}x\right| \lesssim \epsilon H_K \left\|\nabla w_H^0\right\|_{L^\infty(K)}^2 \lesssim \frac{\epsilon}{H_K}\left\|\nabla w_H^0\right\|_{L^2(K)}^2.$$

Note that $\mathbf{a}^*$ is positive definite; that is, there exists a constant α^* such that $\xi_i a^*_{ij} \xi_j \geq \alpha^* |\xi|^2$ for all $\xi \in \mathbb{R}^n$ (see [51]). Thus

$$\alpha^* \left\|\nabla w_H^0\right\|^2_{L^2(K)} \lesssim \|\nabla w_H\|_{L^2(K)} \left\|\nabla w_H^0\right\|_{L^2(K)} + \frac{\epsilon}{H_K} \left\|\nabla w_H^0\right\|^2_{L^2(K)},$$

which implies the result by assuming that $\frac{\epsilon}{H_K}$ is sufficiently small. ∎

The following lemma gives an inverse estimate for the function in space $\mathrm{OMS}(K)$.

LEMMA 5.2. *Assume that* (H4) *holds. Then under the assumptions of Lemma* 5.1 *we have the following estimate:*

$$|v_H|_{H^2(K)} \lesssim \epsilon^{-1} |v_H|_{H^1(K)} \quad \forall v_H \in OMS(K). \tag{5.3}$$

Proof. Assume that $v_H = c_i \bar{\psi}_i^K$. By (4.2), we have the asymptotic expansion

$$v_H = v_H^0 + \epsilon \chi^j(x/\epsilon) \frac{\partial v_H^0}{\partial x_j} + \epsilon \eta^j \frac{\partial v_H^0}{\partial x_j}, \tag{5.4}$$

where $v_H^0 = \Pi_K v_H \in P_1(K)$. Thus, by use of the fact that $\left|\nabla^2 \chi^j(x/\epsilon)\right| \lesssim \epsilon^{-2}$, from (4.14), we obtain

$$\begin{aligned} |v_H|_{H^2(K)} &\leq \epsilon \left| \chi^j(x/\epsilon) \frac{\partial v_H^0}{\partial x_j} \right|_{H^2(K)} + \epsilon \left| \eta^j \frac{\partial v_H^0}{\partial x_j} \right|_{H^2(K)} \\ &\lesssim (\epsilon^{-1} + d_{S(K)}^{-1}) \left\|\nabla v_H^0\right\|_{L^2(K)}, \end{aligned}$$

which implies (5.3) by using the assumption $\epsilon \lesssim H_K \lesssim d_{S(K)}$ and Lemma 5.1. This completes the proof of the lemma. ∎

The following lemma gives the continuity and coercivity of the bilinear form $A_\beta(\cdot,\cdot)$ for the FE-MsFEM.

LEMMA 5.3. *We have*

$$|A_\beta(v,w)| \leq 2 |||v|||_{1,h,H} |||w|||_{1,h,H} \quad \forall\, v, w \in V_{h,H}. \tag{5.5}$$

Further, let the assumptions of Lemma 5.2 *be fulfilled. Then for any* $0 < \gamma_1 \lesssim 1$, *there exists a constant* α_0 *independent of* h, H, ϵ, *and the penalty parameters such that if* $\gamma_0 \geq \alpha_0/\gamma_1$, *then*

$$A_\beta(v_{h,H}, v_{h,H}) \geq \frac{1}{2} |||v_{h,H}|||^2_{1,h,H} \quad \forall\, v_{h,H} \in V_{h,H}. \tag{5.6}$$

Proof. Inequality (5.5) is a direct consequence of the definitions (2.4)–(2.6), (2.17), and (2.18) and the Cauchy–Schwarz inequality.

It remains to prove (5.6). We have

$$\begin{aligned} A_\beta(v_{h,H}, v_{h,H}) &= \sum_{K \in \mathcal{M}_{h,H}} \left\|(\mathbf{a}^\epsilon)^{1/2} \nabla v_{h,H}\right\|^2_{L^2(K)} - 2 \sum_{e \in \Gamma_h} \int_e \{\mathbf{a}^\epsilon \nabla v_{h,H} \cdot \mathbf{n}\} [v_{h,H}] \\ &\quad + \sum_{e \in \Gamma_h} \left(\frac{\gamma_0}{\epsilon} \|[v_{h,H}]\|^2_{L^2(e)} + \gamma_1 \epsilon \|[\mathbf{a}^\epsilon \nabla v_{h,H} \cdot \mathbf{n}]\|^2_{L^2(e)} \right) \\ &= |||v_{h,H}|||^2_{1,h,H} - \sum_{e \in \Gamma_h} \frac{\epsilon}{\gamma_0} \|\{\mathbf{a}^\epsilon \nabla v_{h,H} \cdot \mathbf{n}\}\|^2_{L^2(e)} - 2 \sum_{e \in \Gamma_h} \int_e \{\mathbf{a}^\epsilon \nabla v_{h,H} \cdot \mathbf{n}\} [v_{h,H}]. \end{aligned}$$

It is obvious that

$$2\sum_{e\in\Gamma_h}\int_e\{\mathbf{a}^\epsilon\nabla v_{h,H}\cdot\mathbf{n}\}\,[v_{h,H}]\le 2\sum_{e\in\Gamma_h}\|\{\mathbf{a}^\epsilon\nabla v_{h,H}\cdot\mathbf{n}\}\|_{L^2(e)}\,\|[v_{h,H}]\|_{L^2(e)}$$
$$\le\sum_{e\in\Gamma_h}\frac{\gamma_0}{2\epsilon}\,\|[v_{h,H}]\|^2_{L^2(e)}+\sum_{e\in\Gamma_h}\frac{2\epsilon}{\gamma_0}\,\|\{\mathbf{a}^\epsilon\nabla v_{h,H}\cdot\mathbf{n}\}\|^2_{L^2(e)}.$$

Therefore,

$$A_\beta(v_{h,H},v_{h,H})\ge|||v_{h,H}|||^2_{1,h,H}-\sum_{e\in\Gamma_h}\frac{\gamma_0}{2\epsilon}\,\|[v_{h,H}]\|^2_{L^2(e)}-\sum_{e\in\Gamma_h}\frac{3\epsilon}{\gamma_0}\,\|\{\mathbf{a}^\epsilon\nabla v_{h,H}\cdot\mathbf{n}\}\|^2_{L^2(e)}.$$

It is clear that, for any $e\in\Gamma_h$,

$$\{\mathbf{a}^\epsilon\nabla v_{h,H}\cdot\mathbf{n}\}\Big|_e=(\mathbf{a}^\epsilon\nabla v_H)\cdot\mathbf{n}+\frac{1}{2}\,[\mathbf{a}^\epsilon\nabla v_{h,H}\cdot\mathbf{n}]\Big|_e,$$

where $v_H:=(v_{h,H})|_{\Omega_2}$. Hence, we have

$$A_\beta(v_{h,H},v_{h,H})\ge|||v_{h,H}|||^2_{1,h,H}-\sum_{e\in\Gamma_h}\frac{\gamma_0}{2\epsilon}\,\|[v_{h,H}]\|^2_{L^2(e)}$$
$$-\frac{6\epsilon}{\gamma_0}\sum_{e\in\Gamma_h}\left(\|(\mathbf{a}^\epsilon\nabla v_H)\cdot\mathbf{n}\|^2_{L^2(e)}+\frac{1}{4}\,\|[\mathbf{a}^\epsilon\nabla v_{h,H}\cdot\mathbf{n}]\|^2_{L^2(e)}\right).\tag{5.7}$$

From Lemma 2.1, the inverse estimate (5.3), and $\epsilon\lesssim H$, we have

$$\sum_{e\in\Gamma_h,\,e\subset E}\|(\mathbf{a}^\epsilon\nabla v_H)\cdot\mathbf{n}\|^2_{L^2(e)}=\|(\mathbf{a}^\epsilon\nabla v_H)\cdot\mathbf{n}\|^2_{L^2(E)}$$
$$\le C\left(\frac{1}{H_E}\,\|\nabla v_H\|^2_{L^2(K_E)}+\|\nabla v_H\|_{L^2(K_E)}\,\|\nabla^2 v_H\|_{L^2(K_E)}\right)$$
$$\le\frac{C_1}{\epsilon}\left\|(\mathbf{a}^\epsilon)^{1/2}\nabla v_H\right\|^2_{L^2(K_E)},\tag{5.8}$$

where $K_E\in\mathcal{M}_H$ is the element containing E. Therefore, from (5.7) and (5.8), it follows that

$$A_\beta(v_{h,H},v_{h,H})\ge|||v_{h,H}|||^2_{1,h,H}-\sum_{e\in\Gamma_h}\frac{\gamma_0}{2\epsilon}\,\|[v_{h,H}]\|^2_{L^2(e)}$$
$$-\frac{6C_1}{\gamma_0}\,|v_H|^2_{1,H}-\frac{3}{2\gamma_0\gamma_1}\sum_{e\in\Gamma_h}\gamma_1\epsilon\,\|[\mathbf{a}^\epsilon\nabla v_{h,H}\cdot\mathbf{n}]\|^2_{L^2(e)}$$
$$\ge|||v_{h,H}|||^2_{1,h,H}-\max\left(\frac{1}{2},\frac{6C_1\gamma_1}{\gamma_0\gamma_1},\frac{3}{2\gamma_0\gamma_1}\right)|||v_{h,H}|||^2_{1,h,H}.$$

Noting that $\gamma_1\lesssim 1$, there exists a constant $\alpha_0>0$ independent of h, H, ϵ such that if $\gamma_0\gamma_1\ge\alpha_0$, then $\max\left(\frac{6C_1\gamma_1}{\gamma_0\gamma_1},\frac{3}{2\gamma_0\gamma_1}\right)\le\frac{1}{2}$. This completes the proof of the lemma. ∎

The following lemma is an analogue of Strang's lemma for nonconforming FEMs.

LEMMA 5.4. *Let the assumptions of Lemma 5.2 be fulfilled. Then there exists a constant α_0 independent of ϵ, h, H, and the penalty parameters such that for $0<\gamma_1\lesssim 1$, $\gamma_0\ge\alpha_0/\gamma_1$, the following error estimate holds:*

$$|||u_\epsilon-u_{h,H}|||_{1,h,H}$$
$$\lesssim\inf_{v_{h,H}\in V_{h,H}}|||u_\epsilon-v_{h,H}|||_{1,h,H}+\sup_{w_{h,H}\in V_{h,H}}\frac{\left|\int_\Omega f w_{h,H}\,\mathrm{d}x-A_\beta(u_\epsilon,w_{h,H})\right|}{|||w_{h,H}|||_{1,h,H}}.\tag{5.9}$$

Proof. For any $v_{h,H} \in V_{h,H}$, from Lemma 5.3, (5.6), (2.13), and (5.5), we have, for $0 < \gamma_1 \lesssim 1$ and $\gamma_0 \geq \alpha_0/\gamma_1$,

$$\begin{aligned}
&|||u_{h,H} - v_{h,H}|||^2_{1,h,H} \lesssim A_\beta(u_{h,H} - v_{h,H}, u_{h,H} - v_{h,H}) \\
&\quad = (f, u_{h,H} - v_{h,H}) - A_\beta(v_{h,H}, u_{h,H} - v_{h,H}) \\
&\quad = (f, u_{h,H} - v_{h,H}) - A_\beta(u_\epsilon, u_{h,H} - v_{h,H}) + A_\beta(u_\epsilon - v_{h,H}, u_{h,H} - v_{h,H}) \\
&\quad \lesssim (f, u_{h,H} - v_{h,H}) - A_\beta(u_\epsilon, u_{h,H} - v_{h,H}) + |||u_\epsilon - v_{h,H}|||_{1,h,H} \, |||u_{h,H} - v_{h,H}|||_{1,h,H}.
\end{aligned}$$

Hence

$$\begin{aligned}
|||u_\epsilon - u_{h,H}|||_{1,h,H} &\leq |||u_\epsilon - v_{h,H}|||_{1,h,H} + |||u_{h,H} - v_{h,H}|||_{1,h,H} \\
&\lesssim |||u_\epsilon - v_{h,H}|||_{1,h,H} + \frac{|(f, u_{h,H} - v_{h,H}) - A_\beta(u_\epsilon, u_{h,H} - v_{h,H})|}{|||u_{h,H} - v_{h,H}|||_{1,h,H}},
\end{aligned}$$

which yields the error estimate (5.9). This completes the proof. ∎

Now, we are ready to present the main result of the paper, which gives the error estimate in the norm $|||\cdot|||_{1,h,H}$ for the FE-MsFEM.

THEOREM 5.5. *Let u_ϵ be the solution of* (1.1)*, and let $u_{h,H}$ be the numerical solution computed using FE-MsFEM defined by* (2.13)*. Assume that $u_0 \in H^2(\Omega) \cap W^{1,\infty}(\Omega)$, $f \in L^2(\Omega)$, and $f|_{\Omega_\Gamma} \in H^1(\Omega_\Gamma)$, that assumptions* (H1)–(H6) *hold, and that the penalty parameters satisfy $0 < \gamma_1 \lesssim 1$ and $\gamma_0 \geq \alpha_0/\gamma_1$. Then there exists a constant γ independent of H and ϵ such that if $\epsilon/H_K \leq \gamma$ for all $K \in \mathcal{M}_H$, the following error estimate holds:*

$$\begin{aligned}
|||u_\epsilon - u_{h,H}|||_{1,h,H} \lesssim &\left(\sqrt{\epsilon} + \frac{\epsilon}{H} + \frac{\epsilon}{d} + \frac{h}{\epsilon}\,|\Omega_1|^{1/2}\right)|u_0|_{W^{1,\infty}(\Omega)} + H\,\|f\|_{L^2(\Omega)} \\
&+ \frac{H^2}{\sqrt{\epsilon}}\frac{|u_0|_{H^2(\Omega_\Gamma)}}{\sqrt{|\Omega_\Gamma|}} + \epsilon^2\,\|\nabla f\|_{L^2(\Omega_\Gamma)},
\end{aligned}$$

where Ω_Γ is defined in (2.1) *and $d = \min_{K \in \mathcal{M}_H} d_{S(K)}$.*

Remark 5.1. (a) The error bound consists of three parts: the first part of order $O\left(\sqrt{\epsilon} + \frac{\epsilon}{d} + \frac{\epsilon}{H} + H\right)$ from the oversampling MsFE approximation in Ω_2, the second part of order $O\left(\frac{h}{\epsilon}\,|\Omega_1|^{1/2}\right)$ from the finite element approximation in Ω_1, and the third part $\frac{H^2}{\sqrt{\epsilon}}\frac{|u_0|_{H^2(\Omega_\Gamma)}}{\sqrt{|\Omega_\Gamma|}} + \epsilon^2\,\|\nabla f\|_{L^2(\Omega_\Gamma)}$ from the penalizations on Γ.

(b) Suppose that the interface Γ is chosen such that $\mathrm{dist}(\Gamma, \partial\Omega) = O(H)$. If the average value

$$\frac{|u_0|_{H^2(\Omega_\Gamma)}}{\sqrt{|\Omega_\Gamma|}} \lesssim 1, \tag{5.10}$$

then we have

$$|||u_\epsilon - u_{h,H}|||_{1,h,H} \lesssim \sqrt{\epsilon} + \frac{\epsilon}{d} + \frac{\epsilon}{H} + H + \frac{hH^{1/2}}{\epsilon} + \frac{H^2}{\sqrt{\epsilon}}.$$

In this case, we may choose $H \approx \sqrt{\epsilon}$ and $h \approx \epsilon^{5/4}$ to ensure that $|||u_\epsilon - u_{h,H}|||_{1,h,H} \lesssim \sqrt{\epsilon}$. The condition (5.10) may be checked by using the standard singularity decomposition results for elliptic problems on polygonal domains [44, 20]. For example, we

may show for the two-dimensional case ($n = 2$) that if the inner angles of the polygon Ω are less than $\frac{2}{3}\pi$, then (5.10) holds.

(c) In [31], it is shown that the H^1-error between u_ϵ and the oversampling MsFEM solution is $(\sqrt{\epsilon} + \frac{\epsilon}{H} + \frac{\epsilon}{d}) |u_0|_{W^{1,\infty}(\Omega)} + H \|u_0\|_{H^2(\Omega)}$. Here we have extra two parts due to the finite element approximation in Ω_1 and penalization on Γ. We remark that the oscillating term $\frac{\epsilon}{H}$ can be removed by the Petro–Galerkin MsFEM (see [48]).

(d) The term $\frac{H^2}{\sqrt{\epsilon}}$ in our analysis comes from the penalty term which joints the two methods. It does not come from either the FEM part or the oversampling MsFEM part. In order to achieve the homogenization error (which is $O(\sqrt{\epsilon})$), our method requires $H \eqsim \sqrt{\epsilon}$ (which implies $\frac{H^2}{\sqrt{\epsilon}} \lesssim \sqrt{\epsilon}$), so is the classical oversampling MsFEM (since its error bound contains the terms $H + \frac{\epsilon}{H}$). In this sense the accuracy of our method is not bad. On the other hand, if $H \ll \sqrt{\epsilon}$ or $H \gg \sqrt{\epsilon}$, then both methods give error bounds greater than the homogenization error. The error bound of our method is of the same order as that of the oversampling MsFEM in the former case ($H \ll \sqrt{\epsilon}$) since $\frac{H^2}{\sqrt{\epsilon}} \lesssim \sqrt{\epsilon}$, but is worse in the latter case. Note that the Petrov–Galerkin MsFEM also has a homogenization error of $O(\sqrt{\epsilon})$.

Proof. According to Lemma 5.4, the proof is divided into two parts. The first part is devoted to estimating the interpolation error, and the second part to estimating the nonconforming error.

Part 1. *Interpolation error estimate.* We set $v_{h,H}$ as $v_{h,H}|_{\Omega_1} = \hat{u}_h, v_{h,H}|_{\Omega_2} = \psi_H$, where $\hat{u}_h := Z_h(u_1 + \epsilon\theta_\epsilon)$, and ψ_H is defined in Lemma 4.2 (see (4.19)). We are going to estimate $|||u_\epsilon - v_{h,H}|||_{1,h,H}$, i.e., to estimate each term in its definition (see (2.18)).

First, noting that $\|\nabla Z_h\theta_\epsilon\|_{L^2(\Omega_1)} \lesssim \|\nabla\theta_\epsilon\|_{L^2(\Omega)}$ (see Lemma 4.4), from Lemmas 4.3 and 4.6, Theorem 3.1, and (H5), we have

$$\begin{aligned}(5.11)\quad &\left(|u_\epsilon - v_{h,H}|_{1,H}^2 + \left\|(\mathbf{a}^\epsilon)^{1/2}\nabla(u_\epsilon - v_{h,H})\right\|_{L^2(\Omega_1)}^2\right)^{1/2}\\ &\lesssim H\,|u_0|_{H^2(\Omega)} + \sqrt{\epsilon}\,|u_0|_{W^{1,\infty}(\Omega)} + \frac{\epsilon}{d}\,|u_0|_{W^{1,\infty}(\Omega_2)} + \frac{h}{\epsilon}\,|\Omega_1|^{1/2}\,|u_0|_{W^{1,\infty}(\Omega_1)}.\end{aligned}$$

Further, since $[u_\epsilon] = 0$ and $[u_1] = 0$, it is easy to see that

$$\begin{aligned}\sum_{e\in\Gamma_h}\frac{\gamma_0}{\epsilon}\int_e [u_\epsilon - v_{h,H}]^2 &= \sum_{e\in\Gamma_h}\frac{\gamma_0}{\epsilon}\int_e [u_1 - v_{h,H}]^2\\ &\lesssim \sum_{E\in\Gamma_H}\frac{\gamma_0}{\epsilon}\int_E (u_1 - \psi_H)^2 + \sum_{e\in\Gamma_h}\frac{\gamma_0}{\epsilon}\int_e (u_1 - Z_h u_1)^2 + \sum_{e\in\Gamma_h}\frac{\gamma_0}{\epsilon}\int_e \epsilon^2(-Z_h\theta_\epsilon)^2.\end{aligned}$$

We estimate the three terms on the right-hand side. From Lemma 2.1, we have

$$\int_E (u_1 - \psi_H)^2 \lesssim H^{-1}\|u_1 - \psi_H\|_{L^2(K)}^2 + \|u_1 - \psi_H\|_{L^2(K)}\,\|\nabla(u_1 - \psi_H)\|_{L^2(K)}.$$

Hence, taking a summation over Γ_H yields

$$\begin{aligned}\sum_{E\in\Gamma_H}\frac{\gamma_0}{\epsilon}\int_E (u_1 - \psi_H)^2 &\lesssim \frac{1}{H\epsilon}\sum_{K\in K_{\Gamma_H}}\|u_1 - \psi_H\|_{L^2(K)}^2\\ &\quad + \frac{1}{\epsilon}\Big(\sum_{K\in K_{\Gamma_H}}\|u_1 - \psi_H\|_{L^2(K)}^2\Big)^{1/2}\Big(\sum_{K\in K_{\Gamma_H}}\|\nabla(u_1 - \psi_H)\|_{L^2(K)}^2\Big)^{1/2}.\end{aligned}$$

Thus, from (4.16) and (4.17), it follows that

$$\sum_{E\in\Gamma_H}\frac{\gamma_0}{\epsilon}\int_E(u_\epsilon-\psi_H)^2\lesssim\epsilon\,|u_0|^2_{W^{1,\infty}(\Omega_{\Gamma_H})}+\frac{H^3}{\epsilon}\,|u_0|^2_{H^2(\Omega_{\Gamma_H})}\,,$$

where we have used $H\lesssim d$ and

$$H^{\frac{3}{2}}\,|u_0|_{H^2(\Omega_{\Gamma_H})}\,|u_0|_{W^{1,\infty}(\Omega_{\Gamma_H})}\lesssim\epsilon\,|u_0|^2_{W^{1,\infty}(\Omega_{\Gamma_H})}+\frac{H^3}{\epsilon}\,|u_0|^2_{H^2(\Omega_{\Gamma_H})}$$

to derive this inequality. Further, from Lemmas 2.1 and 4.4, it follows that

$$\begin{aligned}&\sum_{e\in\Gamma_h}\frac{\gamma_0}{\epsilon}\int_e(u_1-Z_hu_1)^2\\&\quad\lesssim\frac{1}{\epsilon}\sum_{e\in\Gamma_h}\left(h^{-1}\,\|u_1-Z_hu_1\|^2_{L^2(K_e)}+\|u_1-Z_hu_1\|_{L^2(K_e)}\,\|\nabla(u_1-Z_hu_1)\|_{L^2(K_e)}\right)\\&\quad\lesssim\frac{h^3}{\epsilon}\,|u_1|^2_{H^2(\Omega_{\Gamma_h})}\,,\end{aligned}$$

where Ω_{Γ_h} is defined in (2.16). In addition, from (4.21) and noting that $e_z\subset\Gamma$ for $z\in\Gamma$, we have

$$\sum_{e\in\Gamma_h}\frac{\gamma_0}{\epsilon}\int_e\epsilon^2(Z_h\theta_\epsilon)^2\lesssim\epsilon\,\|Z_h\theta_\epsilon\|^2_{L^\infty(\Gamma)}\lesssim\epsilon\,\|\theta_\epsilon\|^2_{L^\infty(\Gamma)}\lesssim\epsilon\,|u_0|^2_{W^{1,\infty}(\Omega)}\,,$$

where we have used (3.7) to derive the last inequality. Therefore, from the above three estimates and Lemma 3.5 (noting that $h<\epsilon$), we have

$$\begin{aligned}\sum_{e\in\Gamma_h}\frac{\gamma_0}{\epsilon}\int_e[u_\epsilon-v_{h,H}]^2\lesssim\frac{H^3}{\epsilon}\,|u_0|^2_{H^2(\Omega_\Gamma)}+h^2\,|u_0|^2_{H^2(\Omega)}\\+\epsilon\,|u_0|^2_{W^{1,\infty}(\Omega)}+\epsilon^4\,\|\nabla f\|^2_{L^2(\Omega_\Gamma)}\,.\end{aligned}\tag{5.12}$$

Next, we estimate the term

$$\sum_{e\in\Gamma_h}\frac{\epsilon}{\gamma_0}\,\|\{\mathbf{a}^\epsilon\nabla(u_\epsilon-v_{h,H})\cdot\mathbf{n}\}\|^2_{L^2(e)}\,.$$

It is easy to see that

$$\begin{aligned}\sum_{e\in\Gamma_h}\frac{\epsilon}{\gamma_0}\,\|\{\mathbf{a}^\epsilon\nabla(u_\epsilon-v_{h,H})\cdot\mathbf{n}\}\|^2_{L^2(e)}\lesssim\sum_{E\in\Gamma_H}\frac{\epsilon}{\gamma_0}\,\|\mathbf{a}^\epsilon\nabla(u_\epsilon-\psi_H)\cdot\mathbf{n}\|^2_{L^2(E)}\\+\sum_{e\in\Gamma_h}\frac{\epsilon}{\gamma_0}\,\|\mathbf{a}^\epsilon\nabla(u_\epsilon-\hat{u}_h)\cdot\mathbf{n}\|^2_{L^2(e)}\,.\end{aligned}$$

From Lemma 2.1, we have

$$\begin{aligned}&\|\mathbf{a}^\epsilon\nabla(u_\epsilon-\psi_H)\cdot\mathbf{n}\|^2_{L^2(E)}\lesssim\|\nabla(u_\epsilon-u_1)\cdot\mathbf{n}\|^2_{L^2(E)}+\|\nabla(u_1-\psi_H)\cdot\mathbf{n}\|^2_{L^2(E)}\\&\quad\le H^{-1}\,\|\nabla(u_\epsilon-u_1)\|^2_{L^2(K)}+\|\nabla(u_\epsilon-u_1)\|_{L^2(K)}\left\|\nabla^2(u_\epsilon-u_1)\right\|_{L^2(K)}\\&\qquad+H^{-1}\,\|\nabla(u_1-\psi_H)\|^2_{L^2(K)}+\|\nabla(u_1-\psi_H)\|_{L^2(K)}\left\|\nabla^2(u_1-\psi_H)\right\|_{L^2(K)}\,.\end{aligned}$$

Hence, a summation over Γ_H gives us

$$\begin{aligned}
&\sum_{E\in\Gamma_H}\frac{\epsilon}{\gamma_0}\left\|\mathbf{a}^\epsilon\nabla(u_\epsilon-\psi_H)\cdot\mathbf{n}\right\|^2_{L^2(E)}\\
&\quad\lesssim \epsilon H^{-1}\left\|\nabla(u_\epsilon-u_1)\right\|^2_{L^2(\Omega_{\Gamma_H})}+\epsilon H^{-1}\sum_{K\in K_{\Gamma_H}}\left\|\nabla(u_1-\psi_H)\right\|^2_{L^2(K)}\\
&\qquad+\epsilon\left\|\nabla(u_\epsilon-u_1)\right\|_{L^2(\Omega_{\Gamma_H})}\left\|\nabla^2(u_\epsilon-u_1)\right\|_{L^2(\Omega_{\Gamma_H})}\\
&\qquad+\epsilon\Big(\sum_{K\in K_{\Gamma_H}}\left\|\nabla(u_1-\psi_H)\right\|^2_{L^2(K)}\Big)^{1/2}\Big(\sum_{K\in K_{\Gamma_H}}\left\|\nabla^2(u_1-\psi_H)\right\|^2_{L^2(K)}\Big)^{1/2}.
\end{aligned}$$

Therefore, it follows from Theorems 3.1 and 3.4, Lemma 4.2, and the assumption $\epsilon<H\lesssim d$ that

$$\begin{aligned}
&\sum_{E\in\Gamma_H}\frac{\epsilon}{\gamma_0}\left\|\mathbf{a}^\epsilon\nabla(u_\epsilon-\psi_H)\cdot\mathbf{n}\right\|^2_{L^2(e)}\\
&\lesssim \epsilon H^{-1}\big(\epsilon^2\left|u_0\right|^2_{H^2(\Omega)}+\epsilon\left|u_0\right|^2_{W^{1,\infty}(\Omega)}\big)\\
&\quad+\epsilon H^{-1}\Big(H^2\left|u_0\right|^2_{H^2(\Omega_{\Gamma_H})}+\frac{\epsilon^2}{d^2}H\left|u_0\right|^2_{W^{1,\infty}(\Omega_{\Gamma_H})}\Big)\\
&\quad+\big(\epsilon\left|u_0\right|_{H^2(\Omega)}+\sqrt{\epsilon}\left|u_0\right|_{W^{1,\infty}(\Omega)}\big)\\
&\qquad\times\Big(\epsilon\left|u_0\right|_{H^2(\Omega)}+\sqrt{\epsilon}\left|u_0\right|_{W^{1,\infty}(\Omega)}+\epsilon^2\left\|\nabla f\right\|_{L^2(\Omega_\Gamma)}\Big)\\
&\quad+\Big(H\left|u_0\right|_{H^2(\Omega_{\Gamma_H})}+\frac{\epsilon}{d}\sqrt{H}\left|u_0\right|_{W^{1,\infty}(\Omega_{\Gamma_H})}\Big)\\
&\qquad\times\Big(H\left|u_0\right|_{H^2(\Omega_\Gamma)}+\frac{\epsilon}{d}\sqrt{H}\left|u_0\right|_{W^{1,\infty}(\Omega_{\Gamma_H})}+\epsilon^2\left\|\nabla f\right\|_{L^2(\Omega_\Gamma)}\Big)\\
&\lesssim H^2\left|u_0\right|^2_{H^2(\Omega)}+\epsilon\left|u_0\right|^2_{W^{1,\infty}(\Omega)}+\epsilon^4\left\|\nabla f\right\|^2_{L^2(\Omega_\Gamma)},
\end{aligned}\tag{5.13}$$

where we have used $\frac{\epsilon}{\sqrt{H}}<\sqrt{\epsilon}$ and Young's inequality to derive the last inequality. Similarly, from Lemmas 2.1 and 4.4, we have

$$\begin{aligned}
&\sum_{e\in\Gamma_h}\frac{\epsilon}{\gamma_0}\left\|\mathbf{a}^\epsilon\nabla(u_\epsilon-\hat{u}_h)\cdot\mathbf{n}\right\|^2_{L^2(e)}\lesssim\epsilon\sum_{E\in\Gamma_H}\left\|\mathbf{a}^\epsilon\nabla(u_\epsilon-u_1)\cdot\mathbf{n}\right\|^2_{L^2(E)}\\
&\qquad+\epsilon\sum_{e\in\Gamma_h}\left\|\mathbf{a}^\epsilon\nabla(u_1-Z_hu_1)\cdot\mathbf{n}\right\|^2_{L^2(e)}+\epsilon^3\sum_{e\in\Gamma_h}\left\|\mathbf{a}^\epsilon\nabla(-Z_h\theta_\epsilon)\cdot\mathbf{n}\right\|^2_{L^2(e)}\\
&\quad\lesssim\epsilon H^{-1}\left\|\nabla(u_\epsilon-u_1)\right\|^2_{L^2(\Omega_{\Gamma_H})}+\epsilon\left\|\nabla(u_\epsilon-u_1)\right\|_{L^2(\Omega_{\Gamma_H})}\left\|\nabla^2(u_\epsilon-u_1)\right\|_{L^2(\Omega_{\Gamma_H})}\\
&\qquad+\epsilon h^{-1}\left\|\nabla(u_1-Z_hu_1)\right\|^2_{L^2(\Omega_{\Gamma_h})}+\epsilon\left\|\nabla(u_1-Z_hu_1)\right\|_{L^2(\Omega_{\Gamma_h})}\left\|\nabla^2u_1\right\|_{L^2(\Omega_{\Gamma_h})}\\
&\qquad+\epsilon^3\left\|\nabla(Z_h\theta_\epsilon)\right\|^2_{L^\infty(\Gamma)}.
\end{aligned}$$

From Lemma 4.4 and 3.6, the definition of Ω_{Γ_h} in (2.16), and (3.7), we have

$$\epsilon^3\left\|\nabla(Z_h\theta_\epsilon)\right\|^2_{L^\infty(\Gamma)}\lesssim\epsilon^3\left\|\nabla\theta_\epsilon\right\|^2_{L^\infty(\Omega_{\Gamma_h})}\lesssim\epsilon^3H^{-2}\left\|\theta_\epsilon\right\|^2_{L^\infty(\Omega)}\lesssim\epsilon\left|u_0\right|^2_{W^{1,\infty}(\Omega)}.$$

Thus, from the above two estimates, Theorems 3.1 and 3.4, and Lemmas 3.5 and 4.6, it follows that

$$\begin{aligned}
&\sum_{e\in\Gamma_h}\frac{\epsilon}{\gamma_0}\left\|\mathbf{a}^\epsilon\nabla(u_\epsilon-\hat{u}_h)\cdot\mathbf{n}\right\|^2_{L^2(e)}\\
&\qquad\lesssim\epsilon^2\left|u_0\right|^2_{H^2(\Omega)}+\epsilon\left|u_0\right|^2_{W^{1,\infty}(\Omega)}+\epsilon^4\left\|\nabla f\right\|^2_{L^2(\Omega_\Gamma)}.
\end{aligned}\tag{5.14}$$

It is obvious that a similar argument to that above can be used to get the same error bound for the term

$$\sum_{e\in\Gamma_h}\gamma_1\epsilon\left\|[\mathbf{a}^\epsilon\nabla(u_\epsilon-v_{h,H})\cdot\mathbf{n}]\right\|^2_{L^2(e)}.$$

Thus, it follows from (5.11)–(5.14) that

$$\begin{aligned}\inf_{v_{h,H}\in V_{h,H}}|||u_\epsilon-v_{h,H}|||_{1,h,H}&\lesssim H\,|u_0|_{H^2(\Omega)}+\left(\sqrt{\epsilon}+\frac{\epsilon}{d}\right)|u_0|_{W^{1,\infty}(\Omega)}\\&\quad+\frac{h}{\epsilon}\,|\Omega_1|^{1/2}\,|u_0|_{W^{1,\infty}(\Omega_1)}+\frac{H^{3/2}}{\sqrt{\epsilon}}\,|u_0|_{H^2(\Omega_\Gamma)}+\epsilon^2\,\|\nabla f\|_{L^2(\Omega_\Gamma)}.\end{aligned}\tag{5.15}$$

Part 2. *The nonconforming error estimate.* Define

$$\mathcal{E}_H^{\mathrm{I}}:=\text{set of all interior edges/faces of }\mathcal{M}_H\ .$$

For any $w_{h,H}\in V_{h,H}$, noticing that $\partial\Omega_2/\Gamma$ is empty, it is easy to see

$$\begin{aligned}(f,w_{h,H})&=(f,w_h)_{\Omega_1}+(f,w_H)_{\Omega_2}=(-\nabla\cdot(\mathbf{a}^\epsilon\nabla u_\epsilon),w_h)_{\Omega_1}+(-\nabla\cdot(\mathbf{a}^\epsilon\nabla u_\epsilon),w_H)_{\Omega_2}\\&=(\mathbf{a}^\epsilon\nabla u_\epsilon,\nabla w_h)_{\Omega_1}-\int_\Gamma\mathbf{a}^\epsilon\nabla u_\epsilon\cdot\mathbf{n}w_h\\&\qquad+\sum_{K\in\mathcal{M}_H}\left((\mathbf{a}^\epsilon\nabla u_\epsilon,\nabla w_H)_K-\int_{\partial K}\mathbf{a}^\epsilon\nabla u_\epsilon\cdot\mathbf{n}_K w_H\right)\\&=(\mathbf{a}^\epsilon\nabla u_\epsilon,\nabla w_h)_{\Omega_1}-\sum_{e\in\Gamma_h}\int_e\mathbf{a}^\epsilon\nabla u_\epsilon\cdot\mathbf{n}(w_h-w_H)\\&\qquad+\sum_{K\in\mathcal{M}_H}(\mathbf{a}^\epsilon\nabla u_\epsilon,\nabla w_H)_K-\sum_{E\in\mathcal{E}_H^{\mathrm{I}}}\int_E\mathbf{a}^\epsilon\nabla u_\epsilon\cdot\mathbf{n}_E[w_H].\end{aligned}$$

Here the unit normal vector $\mathbf{n}_E$ is oriented from K to K', and the jump $[v]$ of v on an interior side $E=\partial K\cap\partial K'\in\mathcal{E}_H^{\mathrm{I}}$ is defined as $[v]:=v|_K-v|_{K'}$. Furthermore, by the definition of A_β, it follows that

$$A_\beta(u_\epsilon,w_{h,H})=(\mathbf{a}^\epsilon\nabla u_\epsilon,\nabla w_h)_{\Omega_1}-\sum_{e\in\Gamma_h}\int_e\mathbf{a}^\epsilon\nabla u_\epsilon\cdot\mathbf{n}(w_h-w_H)+\sum_{K\in\mathcal{M}_H}(\mathbf{a}^\epsilon\nabla u_\epsilon,\nabla w_H)_K.$$

Thus, we have

$$\begin{aligned}(f,w_{h,H})-A_\beta(u_\epsilon,w_{h,H})&=-\sum_{E\in\mathcal{E}_H^{\mathrm{I}}}\int_E\mathbf{a}^\epsilon\nabla u_\epsilon\cdot\mathbf{n}_E\,[w_H]\\&=-\sum_{E\in\mathcal{E}_H^{\mathrm{I}}}\int_E\mathbf{a}^\epsilon\nabla u_\epsilon\cdot\mathbf{n}_E\,[w_H-\Pi_H w_H]\\&=-\sum_{K\in\mathcal{M}_H}\int_{\partial K}\mathbf{a}^\epsilon\nabla u_\epsilon\cdot\mathbf{n}_K(w_H-\Pi_H w_H)-\sum_{E\in\Gamma_H}\int_E\mathbf{a}^\epsilon\nabla u_\epsilon\cdot\mathbf{n}(w_H-\Pi_H w_H)\\&:=\mathrm{R}_1+\mathrm{R}_2.\end{aligned}$$

Since

$$\mathrm{R}_1=\int_{\Omega_2}f(w_H-\Pi_H w_H)-\sum_{K\in\mathcal{M}_H}\int_K\mathbf{a}^\epsilon\nabla u_\epsilon\nabla(w_H-\Pi_H w_H),$$

we can estimate R_1 by following [29, Chapter 6], or [16, Chapter 9], or the proof presented in [31, Theorem 3.1], and we obtain

$$|\mathrm{R}_1| \le C\Big(\epsilon\,|u_0|_{H^2(\Omega_2)} + \Big(\sqrt{\epsilon} + \frac{\epsilon}{H}\Big)\,|u_0|_{W^{1,\infty}(\Omega_2)} + \epsilon\,\|f\|_{L^2(\Omega_2)}\Big)\,|w_H|_{1,H}\,. \tag{5.16}$$

Next, we consider the second term R_2. For any $w_H \in X_H$, from (4.2), we have the asymptotic expansion

$$w_H = \Pi_K w_H + \epsilon\chi^j(x/\epsilon)\frac{\partial(\Pi_K w_H)}{\partial x_j} + \epsilon\eta^j\frac{\partial(\Pi_K w_H)}{\partial x_j} \quad \text{in } K.$$

Thus, from (4.3) and Lemma 5.1, we obtain

$$\begin{aligned}\|w_H - \Pi_K w_H\|_{L^\infty(\partial K)} &\lesssim \epsilon\|\nabla\Pi_K w_H\|_{L^\infty(K)} \lesssim \epsilon H^{-n/2}\|\nabla\Pi_K w_H\|_{L^2(K)}\\ &\lesssim \epsilon H^{-n/2}\|\nabla w_H\|_{L^2(K)},\end{aligned}$$

which yields

$$\begin{aligned}|\mathrm{R}_2| &\lesssim \sum_{E\in\Gamma_H}\|\nabla u_\epsilon\|_{L^2(E)}\,\|w_H - \Pi_H w_H\|_{L^2(E)}\\ &\lesssim \|\nabla u_\epsilon\|_{L^2(\Gamma)}\Big(\sum_{E\in\Gamma_H} H^{n-1}\,\|w_H - \Pi_H w_H\|^2_{L^\infty(E)}\Big)^{1/2}\\ &\lesssim \|\nabla u_\epsilon\|_{L^2(\Gamma)}\Big(\sum_{K\in K_{\Gamma_H}} \epsilon^2 H^{-1}\|\nabla w_H\|^2_{L^2(K)}\Big)^{1/2}\\ &\lesssim \epsilon H^{-1/2}\,\|\nabla u_\epsilon\|_{L^2(\Gamma)}\,|w_H|_{1,H} \lesssim \sqrt{\epsilon}\,\|\nabla u_\epsilon\|_{L^2(\Gamma)}\,|w_H|_{1,H}\,.\end{aligned}$$

Note that in the last inequality of the above equation we used $\frac{\epsilon}{\sqrt{H}} < \sqrt{\epsilon}$. On the other hand, it follows from Theorems 3.1 and 3.4 and Lemmas 2.1 and 3.3 that

$$\begin{aligned}\|\nabla u_\epsilon\|^2_{L^2(\Gamma)} &\le \|\nabla(u_\epsilon - u_1)\|^2_{L^2(\Gamma)} + \|\nabla u_0\|^2_{L^2(\Gamma)} + \Big\|\nabla\Big(\epsilon\chi^i(x/\epsilon)\frac{\partial u_0}{\partial x_j}\Big)\Big\|^2_{L^2(\Gamma)}\\ &\lesssim \sum_{K\in K_{\Gamma_H}}\Big(H^{-1}\,\|\nabla(u_\epsilon - u_1)\|^2_{L^2(K)} + \|\nabla(u_\epsilon - u_1)\|_{L^2(K)}\,|u_\epsilon - u_1|_{H^2(K)}\Big)\\ &\quad + |u_0|^2_{W^{1,\infty}(\Omega)} + \epsilon^2\sum_{K\in K_{\Gamma_H}}\Big(H^{-1}\,|u_0|^2_{H^2(K)} + |u_0|_{H^2(K)}\,|u_0|_{H^3(K)}\Big)\\ &\lesssim \epsilon\,|u_0|^2_{H^2(\Omega)} + |u_0|^2_{W^{1,\infty}(\Omega)} + \epsilon^3\,\|\nabla f\|^2_{L^2(\Omega_\Gamma)}\,,\end{aligned}$$

where we have used the assumption $\epsilon/H < 1$ to derive the last inequality. Therefore

$$|\mathrm{R}_2| \lesssim \Big(\epsilon\,|u_0|_{H^2(\Omega)} + \sqrt{\epsilon}\,|u_0|_{W^{1,\infty}(\Omega)} + \epsilon^2\,\|\nabla f\|_{L^2(\Omega_\Gamma)}\Big)\,|w_H|_{1,H}\,. \tag{5.17}$$

It follows from (5.16) and (5.17) that the nonconforming error in Lemma 5.4 is

$$\begin{aligned}&\sup_{w_{h,H}\in V_{h,H}}\frac{\big|\int_\Omega f w_{h,H} - A_\beta(u, w_{h,H})\big|}{|||w_{h,H}|||_{1,h,H}}\\ &\quad\lesssim \epsilon\,|u_0|_{H^2(\Omega)} + \Big(\sqrt{\epsilon} + \frac{\epsilon}{H}\Big)\,|u_0|_{W^{1,\infty}(\Omega)} + \epsilon\,\|f\|_{L^2(\Omega)} + \epsilon^2\,\|\nabla f\|_{L^2(\Omega_\Gamma)}\,,\end{aligned}$$

which, combined with (5.15), (5.9), and (3.5), completes the proof. ∎

6. Numerical tests. In this section, we first demonstrate the accuracy and efficiency of the proposed FE-MsFEM by solving the model problem (1.1) with periodic and randomly generated coefficients, respectively, and then show the ability of the FE-MsFEM to solve two kinds of multiscale elliptic problems with high-contrast channels and well singularities, respectively. In all computations we do not assume that the diffusion coefficient values are available outside of the research domain. In order to illustrate the performance of our method, we also implement two other kinds of methods. The first is the standard MsFEM. The second one is a mixed basis MsFEM, which uses the oversampling multiscale basis inside the domain but away from the boundary, while using the standard MsFEM basis near the boundary. This way, the mixed basis MsFEM doesn't need to use the outside information of the coefficient. We also show the results of the standard linear FEM on the corresponding coarse grid to get a feeling for the accuracy of the multiscale methods.

For the FE-MsFEM and mixed basis MsFEM, the triangulation may be done by the following three steps.

- First, we triangulate the domain Ω with a coarse mesh whose mesh size H is much bigger than ϵ.
- Second, we choose the union of coarse-grid elements adjacent to the boundary $\partial\Omega$ (and the channels or wells, if they exist) as Ω_1 and denote $\Omega \setminus \overline{\Omega_1}$ by Ω_2. For example, in our tests, we choose two layers of coarse-grid elements (and the coarse-grid elements containing the channels or wells, if they exist) to form the domain Ω_1. Hence the distance of Γ away from $\partial\Omega$ is $2H$.
- Finally, in Ω_2 we use the oversampling MsFEM basis on coarse-grid elements, while in Ω_1 we use the traditional linear FEM basis on a fine mesh for the FE-MsFEM, or use the standard MsFEM basis on coarse-grid elements for the mixed basis MsFEM. In our tests, we fix the mesh size of the fine mesh $h = 1/1024$, which is small enough to resolve the smallest scale of oscillations.

See Figure 1 for a sample of the triangulation used in this paper. The coarse mesh triangulation follows the method introduced in [15, 17]. In that method, a uniform $L \times L$ finite element mesh of Ω is constructed by first dividing the domain Ω into $L \times L$ subrectangles and then connecting the lower-left and the upper-right vertices of each subrectangle. The finite elements of the mesh is divided into two groups: the lower-right and upper-left elements. For each triangle K, an oversampling element $S(K)$ is created according to whether K is a lower-right element or an upper-left element, as shown in Figure 3.

Since there are no exact solutions to the problems considered here, we will solve

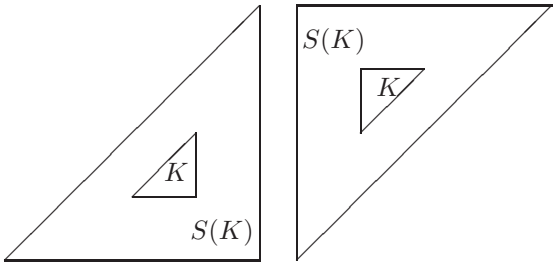

FIG. 3. *The element K and its oversampling element $S(K)$: lower-right elements (left) and upper-left elements (right). The length of the horizontal and vertical edges of $S(K)$ is four times the corresponding length of the edges of K.*

them on a very fine mesh with mesh size $h_f = 1/4096$ by use of the traditional linear finite element method, and consider their numerical solutions as the "exact" solutions, denoted by u_e. Denoting by u_h the numerical solutions computed by the methods considered in this section, we measure the relative errors in the L^2, L^∞, and energy norms as follows:

$$\frac{\|u_h - u_e\|_{L^2}}{\|u_e\|_{L^2}}, \quad \frac{\|u_h - u_e\|_{L^\infty}}{\|u_e\|_{L^\infty}}, \quad \frac{\|u_h - u_e\|_E}{\|u_e\|_E}.$$

We also compare the CPU times T_1 and T_2 spent by the four methods mentioned above, where T_1 is the CPU time of computing the multiscale basis functions, and T_2 is the CPU time of assembling the stiffness matrix and solving the discretized system of algebraic equations. In all tests, for simplicity, the penalty parameters in our FE-MsFEM are chosen as $\gamma_0 = 20$ and $\gamma_1 = 0.1$. The coefficient $\mathbf{a}^\epsilon$ is chosen as the form $\mathbf{a}^\epsilon = a^\epsilon I$, where a^ϵ is a scalar function and I is the 2×2 identity matrix.

6.1. Application to elliptic problems with highly oscillating coefficients. We first consider the model problem (1.1) in the squared domain $\Omega = (0,1) \times (0,1)$. Assume that $f = 1$ and the coefficient $\mathbf{a}^\epsilon(x_1, x_2)$ has the periodic form

$$a^\epsilon(x_1, x_2) = \frac{2 + 1.8\sin(2\pi x_1/\epsilon)}{2 + 1.8\cos(2\pi x_2/\epsilon)} + \frac{2 + 1.8\sin(2\pi x_2/\epsilon)}{2 + 1.8\sin(2\pi x_1/\epsilon)}, \tag{6.1}$$

where we fix $\epsilon = 1/100$. In our FE-MsFEM, we consider two choices of the parameter ρ. The first choice is $\rho = \epsilon$, as stated in our theoretical analysis, while the other is $\rho = h$, the size of the fine mesh. The second choice is useful when the scales are nonseparable. We first choose $H = 1/32$ and report the relative errors in the L^2, L^∞, and energy norms in Table 1. We can see that the FE-MsFEMs give the most accurate results among the methods considered here; even when we take $\rho = h$, the FE-MsFEM still works well. The CPU time T_2 of our FE-MsFEM for assembling the stiffness matrix and solving the linear system is a little longer than that of the mixed basis MsFEM, while the CPU time T_1 for computing the multiscale basis functions is shorter. We remark that the CPU time T_2 can be shortened further by applying more efficient FEMs in the problematic region.

TABLE 1

Relative errors in the L^2, L^∞, and energy norms for the model problem with periodic coefficient given by (6.1). $\epsilon = 1/100$, $H = 1/32$, $h = 1/1024$, $\gamma_0 = 20$, $\gamma_1 = 0.1$.

Relative error	L^2	L^∞	Energy norm	CPU time (s)	
				T_1	T_2
FEM $H = 1/32$	0.1150e-00	0.2311e-00	0.8790e-00	0	0.012
MsFEM	0.7454e-01	0.7355e-01	0.2929e-00	32.26	1.255
Mixed basis MsFEM	0.3609e-01	0.3713e-01	0.2256e-00	419.34	1.255
FE-MsFEM $\rho = \epsilon$	0.1432e-01	0.1538e-01	0.1587e-00	358.38	2.766
FE-MsFEM $\rho = h$	0.1446e-01	0.1546e-01	0.1578e-00	358.38	2.766

The following numerical experiment shows that the coarse mesh size H plays a role as that described in Theorem 5.5. We fix $h = 1/1024$ and $\epsilon = 1/100$. Three kinds of coarse mesh size are chosen. The first one, $H = 1/64$, is denoted as 64×16; the second one, $H = 1/32$, is denoted as 32×32; the last one, $H = 1/16$, is denoted as 16×64. The results are shown in Table 2.

TABLE 2
Relative errors of the FE-MsFEM with $H = 1/64, 1/32$, and $1/16$, respectively, for the model problem with periodic coefficient given by (6.1). $\rho = \epsilon = 1/100$, $h = 1/1024$, $\gamma_0 = 20$, $\gamma_1 = 0.1$.

Relative error	L^2	L^∞	Energy norm	CPU time (s)	
				T_1	T_2
64×16	0.1291e-01	0.1610e-01	0.1766e-00	409.51	0.670
32×32	0.1432e-01	0.1538e-01	0.1587e-00	358.38	2.766
16×64	0.1431e-01	0.1872e-01	0.1640e-00	269.40	2.972

From the table, it is easy to see that as H gets larger, the relative error in the energy norm first gets lower and then gets higher, which coincides with the theoretical results in Theorem 5.5.

Next we simulate the model problem with the random log-normal permeability field $\mathbf{a}^\epsilon(x)$, which is generated by using the moving ellipse average technique [23] with the variance of the logarithm of the permeability $\sigma^2 = 1.5$, and the correlation lengths $l_1 = l_2 = 0.01$ (isotropic heterogeneity) in x_1 and x_2 directions, respectively. One realization of the resulting permeability field in our numerical experiments is depicted in Figure 4, where $\frac{a_{\max}(x)}{a_{\min}(x)} = 1.6137e + 05$. We also compare three kinds of methods, including the standard MsFEM, the mixed basis MsFEM, and the FE-MsFEM. In this test, we set $H = 1/32$ and $\rho = h$ since there is no explicit ϵ in this example. The relative errors for the methods are listed in Table 3. From the table, we can also see that the FE-MsFEM gives the most accurate results among the methods considered here, and the direct coarse mesh FEM gives a wrong approximation to the gradient of solution.

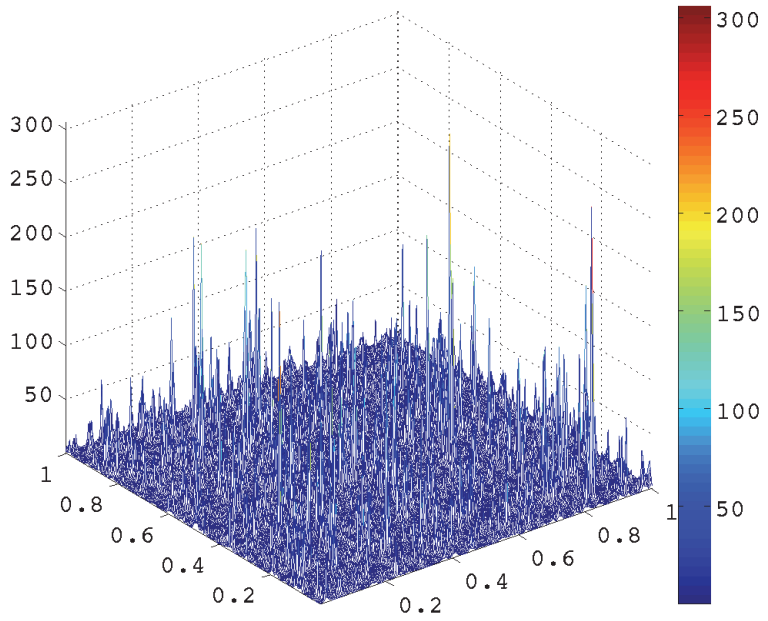

FIG. 4. *The random log-normal permeability field $a^\epsilon(x)$. $\frac{a_{\max}(x)}{a_{\min}(x)} = 1.6137e + 05$.*

TABLE 3

Relative errors in the L^2, L^∞, and energy norms for the model problem with random coefficient as shown in Figure 4. $H = 1/32, \rho = h = 1/1024$, $\gamma_0 = 20$, $\gamma_1 = 0.1$.

Relative error	L^2	L^∞	Energy norm	CPU time (s)	
				T_1	T_2
FEM $H = 1/32$	0.2145e-00	0.2618e-00	0.1339e+01	0	0.012
MsFEM	0.3707e-00	0.3747e-00	0.6051e-00	25.61	0.015
Mixed basis MsFEM	0.1126e-00	0.1832e-00	0.4844e-00	338.54	0.015
FE-MsFEM	0.2535e-01	0.8149e-01	0.3153e-00	286.01	1.430

6.2. Application to multiscale problems with high-contrast channels. In this subsection, we use the introduced FE-MsFEM to solve two elliptic multiscale problems which have high-contrast channels inside the domain.

In the first example, the coefficient $\mathbf{a}^\epsilon$ is characterized by a fine and long-ranged high-permeability channel, which is set in the following way. The example utilizes the periodic coefficient a^ϵ in (6.1) as the background, while changing the values on a narrow and long channel that are defined from $(90/1024, 498/1024)$ to $(934/1024, 518/1024)$ with the new value $a^\epsilon = 10^5$ (see Figure 5). We set $H = 1/32$ and $\rho = \epsilon = 1/100$. The results are presented in Table 4, where the relative errors in the L^2, L^∞, and energy norms are shown. We observe that the FE-MsFEM performs better than the other methods.

In the second example, we use the coefficient depicted in Figure 6 that corresponds to a coefficient with background one and high-permeability channels and inclusions

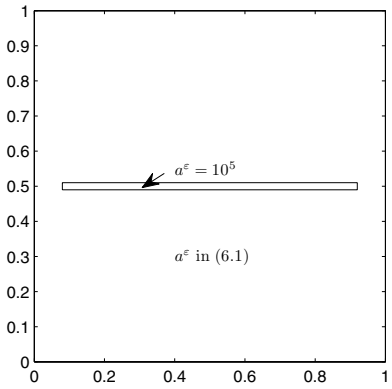

FIG. 5. *Permeability field.*

TABLE 4

Relative errors for the model problem with the permeability depicted in Figure 5. $\rho = \epsilon = 1/100$, $H = 1/32$, $h = 1/1024$, $\gamma_0 = 20$, $\gamma_1 = 0.1$.

Relative error	L^2	L^∞	Energy norm	CPU time (s)	
				T_1	T_2
FEM $H = 1/32$	0.1196e-00	0.2720e-00	0.1319e+02	0	0.016
MsFEM	0.1795e-00	0.2484e-00	0.4151e-00	32.61	1.259
Mixed basis MsFEM	0.6914e-01	0.2393e-00	0.3264e-00	428.68	1.268
FE-MsFEM	0.1434e-01	0.2622e-01	0.1641e-00	366.47	3.254

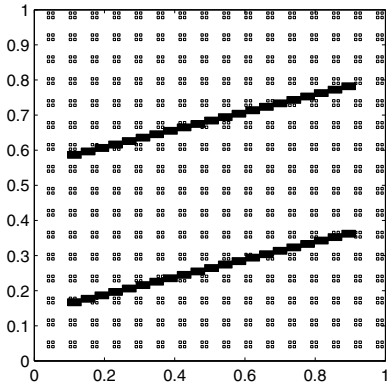

FIG. 6. *Permeability field:* $a^\epsilon = 10^5$ *in two channels consisting of dark small rectangles;* $a^\epsilon = 8 \times 10^4$ *in small square inclusions;* $a^\epsilon = 1$ *otherwise.*

TABLE 5

Relative errors for the model problem with the permeability depicted in Figure 6. $H = 1/32$, $\rho = h = 1/1024$, $\gamma_0 = 20$, $\gamma_1 = 0.1$.

Relative error	L^2	L^∞	Energy norm	CPU time (s)	
				T_1	T_2
FEM $H = 1/32$	0.3339e-00	0.4615e-00	0.8511e+02	0	0.012
MsFEM	0.3552e-00	0.4011e-00	0.5946e-00	26.32	0.907
Mixed basis MsFEM	0.2222e-00	0.2623e-00	0.5026e-00	351.97	0.915
FE-MsFEM	0.4786e-02	0.1362e-01	0.9322e-01	236.03	6.108

with permeability values equal to 10^5 and 8×10^4, respectively. The results are listed in Table 5. We observe that our FE-MsFEM gives better results than the other three methods.

6.3. Application to multiscale problems with the Dirac function singularities. In this subsection, we consider the elliptic multiscale problem with the Dirac function singularities inside the domain, which stems from the simulation of steady flow transport through highly heterogeneous porous media driven by extraction wells [17].

We first recall the problems as stated in [17]. Let $\Omega \in \mathbf{R}^2$ be a bounded domain with Lipschitz boundary $\partial\Omega$. $B(x_0, r)$ will be denoted as the disk centered at x_0 with radius $r > 0$. Let $B_j = B(\bar{x}_j, \delta), 1 \le j \le N$, be mutually disjoint subdomains inside Ω that are occupied by the wells. Denote $\Omega_\delta = \Omega \setminus (\cup_{j=1}^N \bar{B}_j)$. We consider the following single-phase pressure equation which is formed by combining Darcy's law with the conservation of mass:

$$-\nabla \cdot (a^\epsilon(x) \nabla u_\delta) = 0 \qquad \text{in } \Omega_\delta, \tag{6.2}$$

where u_δ is pressure and a^ϵ is the permeability, which is typically highly variable in space. We will impose the homogeneous Dirichlet boundary condition $u_\delta|_{\partial\Omega} = 0$ on the exterior boundary. On the well boundary ∂B_j, two quantities are of particular importance in practical applications: the well bore pressure $u_\delta|_{\partial B_j}$ and the well flow rate $q_j = \int_{\partial B_j} a^\epsilon \nabla u_\delta \cdot \mathbf{n} ds$, where $\mathbf{n}$ is the unit outer normal to $\partial\Omega_\delta$. Here we fix the well flow rate q_j and try to find the well bore pressure $u_\delta|_{\partial B_j}$.

Since the size of the wells δ is negligible in practical situations, the first approximation to be made is to replace (6.2) by the following problem:

$$\begin{cases} -\nabla \cdot (a^\epsilon(x)\nabla u) = \sum_{j=1}^{N} q_j \delta_{\bar{x}_j} & \text{in } \Omega, \\ u(x) = 0 & \text{on } \partial\Omega, \end{cases} \tag{6.3}$$

where $\delta_{\bar{x}_j}$ is the Dirac measure at $\bar{x}_j$. In this paper, we try to compute the approximate well bore pressure via solving the approximation problem (6.3). The computations are performed on the unit square $\Omega = (0,1) \times (0,1)$. In this example, we assume the coefficient is

$$a^\epsilon(x_1, x_2) = \frac{1}{(2 + 1.5 \sin \frac{2\pi x_1}{\epsilon})(2 + 1.5 \sin \frac{2\pi x_2}{\epsilon})}. \tag{6.4}$$

We consider two wells $B_j = B(\bar{x}_j, \delta)$ with $\bar{x}_1 = (\frac{1}{4}, \frac{3}{4}), \bar{x}_2 = (\frac{3}{4}, \frac{1}{4})$ and radius $\delta = 10^{-5}$. In the computations we take $q_1 = -1$ and $q_2 = 1$, which corresponds to the situation that the well B_1 is an extraction well and B_2 is an injection well. In our test, we choose $\epsilon = 1/64 = 0.015625$. The "exact" well bore pressures, obtained using the method introduced in [17, section 6] based on the well-resolved solution on a uniform 2048×2048 mesh, are $\alpha_1 = -5.3884973$ in the first well and $\alpha_2 = 5.3884973$ in the second well (see [17, Example 7.1]).

In this test, the FE-MsFEM is implemented via refining the coarse-grid elements around the wells as well as the union of coarse-grid elements adjacent to the boundary $\partial\Omega$. We also show the results obtained by Algorithm 7.1 in [17], where the method needs to compute the discrete Green functions in a very fine mesh. We referred to this method as G-MsFEM. Based on the numerical solution of (6.3), we use the developed new Peaceman method in [17, section 6] to compute the well bore pressures on each well. The results are listed in Table 6. We observe that the introduced FE-MsFEM provides a better approximation of the well bore pressure than G-MsFEM, and a much better approximation than the MsFEM and mixed basis MsFEM in this example. More extensive numerical experiments will be reported elsewhere.

TABLE 6

Results for the approximate problem (6.3). $H = 1/64$, $h = 1/2048$, $\rho = \epsilon = 1/64$, $\gamma_0 = 20$, $\gamma_1 = 0.1$.

Methods	Well no.	Well bore pressure	Relative error
MsFEM	1	−7.3333601	0.3609e-00
Mixed basis MsFEM	1	−7.3340404	0.3611e-00
G-MsFEM	1	−5.3838442	0.8635e-03
FE-MsFEM	1	−5.3877810	0.1329e-03
MsFEM	2	7.3333601	0.3609e-00
Mixed basis MsFEM	2	7.3340404	0.3611e-00
G-MsFEM	2	5.3739254	0.2704e-02
FE-MsFEM	2	5.3877810	0.1329e-03

7. Conclusions. In this paper, we have developed a new numerical scheme for the elliptic multiscale problems which joins the oversampling MsFEM and the standard FEM using penalty techniques. The idea is first to separate the research domain

into two parts, Ω_1 and $\Omega_2 = \Omega \setminus \overline{\Omega_1}$, in which Ω_1 contains singular points (or regions) where the oversampling MsFEM is inefficient. Then we apply the standard FEM on a fine mesh of Ω_1 and the oversampling MsFEM on a coarse mesh of Ω_2. The two methods are joined on the interface $\Gamma = \partial\Omega_1 \cap \partial\Omega_2$ of the fine and coarse meshes by penalizing the jumps of the function values as well as the fluxes of discrete solutions.

A rigorous and careful analysis has been given for the elliptic equation with periodic diffusion coefficient to show that, under some mild assumptions, if Γ is so chosen that $\mathrm{dist}(\Gamma, \partial\Omega) \gtrsim H$, then the H^1-error of our new method is of the order

$$O\Big(\sqrt{\epsilon} + \frac{\epsilon}{d} + \frac{\epsilon}{H} + H + \frac{h}{\epsilon}|\Omega_1|^{\frac{1}{2}} + \frac{H^2}{\sqrt{\epsilon}}\Big),$$

which exactly consists of the oversampling MsFE approximation error in Ω_2, the FE approximation error in Ω_1, and the error contributed by the penalizations on Γ. Note that, for simplicity, we have only analyzed the linear version of the FEM for the discretization on Ω_1.

Numerical experiments are carried out for the elliptic equations with periodic oscillating or random coefficients, as well as the multiscale problems with high-contrast channels or well singularities, to verify the theoretical findings and compare the performance of our FE-MsFEM with the standard MsFEM or mixed basis MsFEM. It is shown that the FE-MsFEM gives much better accuracy than the other two methods in the problems with singularities.

There are several ways to improve further the performance of our FE-MsFEM. First, the linear FEM on Ω_1 can be apparently extended to higher order FEMs to reduce the error term related to Ω_1. Second, since Ω_1 may contain singularities, another interesting project is to consider the combination of an adaptive FEM on local refined meshes on Ω_1 and an oversampling MsFEM on Ω_2. Third, based on existing numerical results for oversampling MsFEMs [46], we conjecture that the theoretical assumption of $\mathrm{dist}(\Gamma, \partial\Omega) \gtrsim H$ may be weakened to $\mathrm{dist}(\Gamma, \partial\Omega) \geq C\epsilon$ (at least in practice) for some constant C. These will be left as future studies.

Acknowledgment. The authors would like to thank the referees for their careful reading and constructive comments, which improved the paper.

REFERENCES

[1] J. E. Aarnes, *On the use of a mixed multiscale finite element method for greater flexibility and increased speed or improved accuracy in reservoir simulation*, Multiscale Model. Simul., 2 (2004), pp. 421–439.

[2] T. Arbogast, *Numerical subgrid upscaling of two-phase flow in porous media*, in Numerical Treatment of Multiphase Flows in Porous Media, Lecture Notes in Phys. 552, Z. Chen, R. E. Ewing, and Z. C. Shi, eds., Springer-Verlag, New York, 2000, pp. 35–49.

[3] T. Arbogast, *Implementation of a locally conservative numerical subgrid upscaling scheme for two-phase Darcy flow*, Comput. Geosci., 6 (2002), pp. 453–481.

[4] T. Arbogast, G. Pencheva, M. F. Wheeler, and I. Yotov, *A multiscale mortar mixed finite element method*, Multiscale Model. Simul., 6 (2007), pp. 319–346.

[5] D. N. Arnold, *An interior penalty finite element method with discontinuous elements*, SIAM J. Numer. Anal., 19 (1982), pp. 742–760.

[6] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal., 39 (2002), pp. 1749–1779.

[7] M. Avellaneda and F. Lin, *Compactness methods in the theory of homogenization*, Comm. Pure Appl. Math., 40 (1987), pp. 803–847.

[8] I. Babuška, G. Caloz, and J. 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.

[9] I. Babuška and J. E. Osborn, *Generalized finite element methods: Their performance and their relation to mixed methods*, SIAM J. Numer. Anal., 20 (1983), pp. 510–536.
[10] I. Babuška and M. Zlámal, *Nonconforming elements in the finite element method with penalty*, SIAM J. Numer. Anal., 10 (1973), pp. 863–875.
[11] G. Baker, *Finite element methods for elliptic equations using nonconforming elements*, Math. Comp., 31 (1977), pp. 44–59.
[12] A. Bensoussan, J. L. Lions, and G. Papanicolaou, *Asymptotic Analysis for Periodic Structure*, Stud. Math. Appl. 5, North-Holland, Amsterdam, 1978.
[13] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, New York, 2002.
[14] F. Brezzi, L. P. Franca, T. J. R. Hughes, and A. Russo, $b = \int g$, Comput. Methods Appl. Mech. Engrg., 145 (1997), pp. 329–339.
[15] Z. Chen and T. Y. Hou, *A mixed multiscale finite method for elliptic problems with oscillating coefficients*, Math. Comp., 72 (2002), pp. 541–576.
[16] Z. Chen and H. Wu, *Selected Topics in Finite Element Method*, Science Press, Beijing, 2010.
[17] Z. Chen and X. Y. Yue, *Numerical homogenization of well singularities in the flow transport through heterogeneous porous media*, Multiscale Model. Simul., 1 (2003), pp. 260–303.
[18] E. T. Chung, Y. Efendiev, and G. Li, *An adaptive GMsFEM for high-contrast flow problems*, J. Comput. Phys., 273 (2014), pp. 54–76.
[19] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North Holland, Amsterdam, 1978.
[20] M. Dauge, *Elliptic Boundary Value Problems in Corner Domains: Smoothness and Asymptotics of Solutions*, Lecture Notes in Math. 1341, Springer-Verlag, Berlin, 1988.
[21] M. Dorobantu and B. Engquist, *Wavelet-based numerical homogenization*, SIAM J. Numer. Anal., 35 (1998), pp. 540–559.
[22] J. Douglas, Jr., and T. Dupont, *Interior Penalty Procedures for Elliptic and Parabolic Galerkin Methods*, Lecture Notes in Phys. 58, Springer-Verlag, Berlin, 1976.
[23] L. Durlofsky, *Numerical calculation of equivalent grid block permeability tensors for heterogeneous porous media*, Water Resources Res., 27 (1991), pp. 699–708.
[24] W. E and B. Engquist, *The heterogeneous multiscale methods*, Commun. Math. Sci., 1 (2003), pp. 87–132.
[25] W. E and B. Engquist, *Multiscale modeling and computation*, Notices Amer. Math. Soc., 50 (2003), pp. 1062–1070.
[26] W. E, P. Ming, and P. Zhang, *Analysis of the heterogeneous multiscale method for elliptic homogenization problems*, J. Amer. Math. Soc., 18 (2005), pp. 121–156.
[27] Y. Efendiev, J. Galvis, and X. H. Wu, *Multiscale finite element methods for high-contrast problems using local spectral basis functions*, J. Comput. Phys., 230 (2011), pp. 937–955.
[28] Y. Efendiev, V. Ginting, T. Y. Hou, and R. Ewing, *Accurate multiscale finite element methods for two-phase flow simulations*, J. Comput. Phys., 220 (2006), pp. 155–174.
[29] Y. Efendiev and T. Y. Hou, *Multiscale Finite Element Methods Theory and Applications*, Springer, Lexington, KY, 2009.
[30] Y. Efendiev, T. Y. Hou, and V. Ginting, *Multiscale finite element methods for nonlinear partial differential equations*, Commun. Math. Sci., 2 (2004), pp. 553–589.
[31] Y. R. Efendiev, T. Y. Hou, and X.-H. Wu, *Convergence of a nonconforming multiscale finite element method*, SIAM J. Numer. Anal., 37 (2000), pp. 888–910.
[32] Y. Efendiev and A. Pankov, *Numerical homogenization of nonlinear random parabolic operators*, Multiscale Model. Simul., 2 (2004), pp. 237–268.
[33] B. Engquist and O. Runborg, *Wavelet-based numerical homogenization with applications*, in Multiscale and Multiresolution Methods: Theory and Applications, Lect. Notes Comput. Sci. Eng. 20, T. Barth, T. Chan, and R. Heimes, eds., Springer-Verlag, Berlin, 2002, pp. 97–148.
[34] C. Farhat, I. Harari, and L. P. Franca, *The discontinuous enrichment method*, Comput. Meth. Appl. Mech. Eng., 190 (2001), pp. 6455–6479.
[35] C. L. Farmer, *Upscaling: A review*, in Proceedings of the Institute of Computational Fluid Dynamics Conference on Numerical Methods for Fluid Dynamics, Oxford, UK, 2001.
[36] X. Feng and H. Wu, *Discontinuous Galerkin methods for the Helmholtz equation with large wave number*, SIAM J. Numer. Anal., 47 (2009), pp. 2872–2896; preprint available from arXiv:0810.1475.
[37] X. Feng and H. Wu, *hp-discontinuous Galerkin methods for the Helmholtz equation with large wave number*, Math. Comp., 80 (2011), pp. 1997–2024.
[38] J. Fish and V. Belsky, *Multigrid method for a periodic heterogeneous medium, part* I: *Multiscale modeling and quality in multidimensional case*, Comput. Meth. Appl. Mech. Eng.,

126 (1995), pp. 17–38.
[39] J. Fish and Z. Yuan, *Multiscale enrichment based on partition of unity*, Internat. J. Numer. Methods Engrg., 62 (2005), pp. 1341–1359.
[40] L. P. Franca and A. Russo, *Deriving upwinding, mass lumping and selective reduced integration by residual-free bubbles*, Appl. Math. Lett., 9 (1996), pp. 83–88.
[41] J. Galvis and Y. Efendiev, *Domain decomposition preconditioners for multiscale flows in high-contrast media*, Multiscale Model. Simul., 8 (2010), pp. 1461–1483.
[42] J. Galvis and Y. Efendiev, *Domain decomposition preconditioners for multiscale flows in high-contrast media: Reduced dimension coarse spaces*, Multiscale Model. Simul., 8 (2010), pp. 1621–1644.
[43] D. Gilbarg and N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, Berlin, 2001.
[44] P. Grisvard, *Elliptic Problems on Nonsmooth Domains*, Pitman, Boston, 1985.
[45] P. Henning, M. Ohlberger, and B. Schweizer, *An adaptive multiscale finite element method*, Multiscale Model. Simul., 12 (2014), pp. 1078–1107.
[46] 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.
[47] T. Y. Hou, X. H. Wu, and Z. Cai, *Convergence of a multiscale finite element method for elliptic problems with rapidly oscillation coefficients*, Math. Comp., 68 (1999), pp. 913–943.
[48] T. Y. Hou, X. H. Wu, and Y. Zhang, *Removing the cell resonance error in the multiscale finite element method via a Petrov-Galerkin formulation*, Commun. Math. Sci., 2 (2004), pp. 185–205.
[49] T. Hughes, *Multiscale phenomena: Green's functions, the Dirichlet to Neumann formulation, subgrid scale models, bubbles and the origin of stabilized methods*, Comput. Methods Appl. Mech. Engrg., 127 (1995), pp. 387–401.
[50] P. Jenny, S. Lee, and H. Tchelepi, *Multi-scale finite-volume method for elliptic problems in subsurface flow simulation*, J. Comput. Phys., 187 (2003), pp. 47–67.
[51] V. V. Jikov, S. M. Kozlov, and O. A. Oleinik, *Homogenization of Differential Operators and Integral Functionals*, Springer-Verlag, Berlin, 1994.
[52] O. A. Ladyzhenskaia and N. N. Uraltseva, *Linear and Quasilinear Elliptic Equations*, Academic Press, New York, 1968.
[53] R. Massjung, *An hp-Error Estimate for an Unfitted Discontinuous Galerkin Method Applied to Elliptic Interface Problems*, RWTH 300, IGPM Report, 2009.
[54] S. Moskow and M. Vogelius, *First-order corrections to the homogenised eigenvalues of a periodic composite medium. A convergence proof*, Proc. Roy. Soc. Edinburgh Sect. A, 127 (1997), pp. 1263–1300.
[55] J. D. Moulton, J. E. Dendy, and J. M. Hyman, *The black box multigrid numerical homogenization algorithm*, J. Comput. Phys., 141 (1998), pp. 1–29.
[56] H. Owhadi and L. Zhang, *Metric-based upscaling*, Comm. Pure Appl. Math., 60 (2007), pp. 675–723.
[57] H. Owhadi and L. Zhang, *Localized bases for finite-dimensional homogenization approximations with nonseparated scales and high contrast*, Multiscale Model. Simul., 9 (2011), pp. 1373–1398.
[58] M. Peszyńska, M. Wheeler, and I. Yotov, *Mortar upscaling for multiphase flow in porous media*, Comput. Geosci., 6 (2002), pp. 73–100.
[59] G. Sangalli, *Capturing small scales in elliptic problems using a residual-free bubbles finite element method*, Multiscale Model. Simul., 1 (2003), pp. 485–503.
[60] R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.
[61] M. F. Wheeler, *An elliptic collocation-finite element method with interior penalties*, SIAM J. Numer. Anal., 15 (1978), pp. 152–161.
[62] H. Wu, *Pre-asymptotic error analysis of CIP-FEM and FEM for Helmholtz equation with high wave number. Part* I: *Linear version*, IMA J. Numer. Anal., 2013; doi:10.1093/imanum/drt033.
[63] H. Wu and Y. Xiao, *An unfitted hp-interface penalty finite element method for elliptic interface problems*, submitted.
[64] X. H. Wu, Y. Efendiev, and T. Hou, *Analysis of upscaling absolute permeability*, Discrete Contin. Dyn. Syst. Ser. B, 2 (2002), pp. 185–204.
[65] L. Zhu and H. Wu, *Preasymptotic error analysis of CIP-FEM and FEM for Helmholtz equation with high wave number. Part* II: *hp version*, SIAM J. Numer. Anal., 51 (2013), pp. 1828–1852.