ERROR ESTIMATE OF MULTISCALE FINITE ELEMENT METHOD FOR PERIODIC MEDIA REVISITED *

PINGBING MING† AND SIQI SONG†

Abstract. We derive the optimal energy error estimate for multiscale finite element method with oversampling technique applying to elliptic systems with rapidly oscillating periodic coefficients that are bounded measurable, which may admit rough microstructures. As a by-product of the energy error estimate, we derive the rate of convergence in $L^{d/(d-1)}-$norm with d the dimensionality.

Key words. Multiscale finite element method, homogenization, error estimate, oversampling

AMS subject classifications. 35J15, 65N12, 65N30

1. Introduction.

The multiscale finite element method (MsFEM) introduced by Hou an Wu [19] aims for solving the boundary value problems with rapidly oscillating coefficients without resolving the fine scale information. The main idea is to exploit the multiscale basis functions that capture the fine scale information of the underlying partial differential equations. MsFEM has been successfully applied to many problems such as two phase flows, nonlinear homogenization problems, convection-diffusion problems, elliptic interface problems with high-contrast coefficients and Poisson problem with rough and oscillating boundary, we refer to book [15] for a survey of MsFEM before 2009. More recent efforts for MsFEM focus on extending the method to deal with more general media; cf., [11, 7, 6]. We also refer to [31, 32, 2, 5] for a summary of recent progress for related methods.

In [20] and [16], the authors proved MsFEM converges for the scalar elliptic boundary value problem in two dimension with periodic oscillating coefficients in the energy norm, and the convergence rate is $\sqrt{\varepsilon}+h+\varepsilon/h$, where h is the mesh size of the triangulation, and ε is the period of the oscillation. The technical assumptions are

1. The coefficient matrix of the elliptical problem is symmetric, and each entry is a C^1 function;
2. The homogenized solution $u_0 \in W^{1,\infty}(\Omega)\cap H^2(\Omega)$;
3. The corrector χ defined in (3.2) belong to $W^{1,\infty}$.

The first assumption excludes the rough microstructures, which frequently appears in the realistic materials [36]; The second assumption is standard except that $u_0 \in W^{1,\infty}(\Omega)$, which may not be true even for Poisson equation posed on a ball [12]. The last assumption on the corrector is not realistic at all, though it may be true for certain special microstructures such as laminates [10] and for problems with piecewise Hölder continuous coefficients [25, 24]; We refer to [14] for an elaboration on this assumption.

Nevertheless, there are some subsequent endeavor on proving the error estimates for MsFEM under weaker assumptions; see, e.g., [8, 33, 9, 37], just name a few, most of them concern the second assumption, while it is still unknown whether the above assumptions may be removed or to what degree they may be weakened. Moreover,

*Submitted to the editors October 17, 2023.

Funding: The work of Ming was supported by the National Natural Science Foundation of China under the grants 11971467 and 12371438.

†LSEC, Institute of Computational Mathematics and Scientific/Engineering Computing, AMSS, Chinese Academy of Sciences, No. 55, East Road Zhong-Guan-Cun, Beijing 100190, China and School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China (mpb@lsec.cc.ac.cn,songsq@lsec.cc.ac.cn).

though MsFEM has been successfully applied to elliptic systems [15, 11], while it does not seem easy to extend the proof to elliptic systems because the maximum principle has been exploited, which may be invalid for elliptic systems [22].

The present work gives an affirmative answer to the above questions. Assuming that $u_0 \in W^{2,d}$ with the dimensionality $d = 2, 3$, we prove the optimal energy error estimate of MsFEM with/without oversampling for elliptic systems with bounded, measurable and symmetric periodical coefficients; cf. Theorem 4.1 and Theorem 4.10. The symmetry assumption may be dropped for MsFEM without oversampling, or for MsFEM with oversampling applying to the elliptic scalar problem. This means that MsFEM achieves optimal convergence rate for problems with rough microstructures.

As an application of the energy error estimate, we derive improved error estimate of MsFEM in $L^{d/(d-1)}-$norm by resorting to the Aubin-Nitsche dual argument [3, 30], naturally, this gives the L^2-error estimates for two-dimensional problem and the elliptic scalar problem in three dimension. Such estimate would be useful for analyzing MsFEM applying to the eigenvalue problems in composites [21].

There are two ingredients in our proof. The one is a local version of the multiplier estimates for periodic homogenization of elliptic systems [38, 35]; see Lemma 4.5, which helps us to remove the boundedness assumption on the gradient of the corrector. Another one is a local estimate of the gradient of the first order approximation of the solution; see Lemma 4.8, which bypasses the maximum principle in the proof, hence we may derive the error estimate for elliptic systems.

The remaining part of the paper is as follows. We formulate MsFEM with oversampling in § 2. In § 3, we recall some quantitative estimates of the periodic homogenization for elliptic systems. The energy error estimate will be given in § 4, from which we prove the error estimates in $L^{d/d-1}$ norm. As a direct consequence of these estimates, we prove the error estimates for MsFEM without oversampling. In the last section, we summarize our results and discuss certain extensions.

Throughout this paper, C is a generic constant that may be different at different occurrence, while it is independent of the mesh size h and the small parameter ε.

2. Multiscale Finite Element Method with Oversampling. We firstly fix some notations. Let Ω be a bounded Lipschitz domain in $\mathbb{R}^d$ (we focus on $d = 2, 3$). The standard Sobolev space $W^{k,p}(\Omega)$ will be used [1], which is equipped with the norm $\|\cdot\|_{W^{k,p}(\Omega)}$. We use the convention $H^k(\Omega) = W^{k,2}(\Omega)$. We denote by $W^{k,p}(\Omega;\mathbb{R}^m)$ the vector-valued function with each component belonging to $W^{k,p}(\Omega)$, and define $|D| := \text{mes} D$ for any measurable set D.

We consider the second order elliptic system in divergence form

$$\mathcal{L}_\varepsilon = -\,\text{div}\,(\,A(x/\varepsilon)\nabla\,)$$

with the coefficient A given by

$$A(y) = a_{ij}^{\alpha\beta}(y) \qquad i, j = 1, \cdots, d \text{ and } \alpha, \beta = 1, \cdots, m.$$

For $u = (u^1, \cdots, u^m)$,

$$(\,\mathcal{L}_\varepsilon(u)\,)^\alpha := -\frac{\partial}{\partial x_i}\left(a_{ij}^{\alpha\beta}\left(\frac{x}{\varepsilon}\right)\frac{\partial u^\beta}{\partial x_j}\right) \qquad \alpha = 1, \cdots, m.$$

We always assume that A is bounded measurable and satisfies the Legendre-Hadamard condition as

$$\lambda\,|\xi|^2\,|\eta|^2 \le a_{ij}^{\alpha\beta}(y)\xi_i\xi_j\eta_\alpha\eta_\beta \le \Lambda\,|\xi|^2\,|\eta|^2 \qquad \text{for a.e. } y \in \mathbb{R}^d, \tag{2.1}$$

where $\xi = (\xi_1, \cdots, \xi_d)$ and $\eta = (\eta_1, \cdots, \eta_m)$. The transpose of A is understood as $A^t(y) = a_{ji}^{\beta\alpha}(y)$. We assume that A is $1-$periodic; i.e., for all $z \in \mathbb{Z}^d$,

$$A(y+z) = A(y) \qquad \text{for a.e. } y \in \mathbb{R}^d.$$

Considering the following homogeneous boundary value problem: Given $f \in H^{-1}(\Omega;\mathbb{R}^m)$, we find $u^\varepsilon \in H_0^1(\Omega;\mathbb{R}^m)$ satisfying

$$\mathcal{L}_\varepsilon(u^\varepsilon) = f \quad \text{in} \quad \Omega \qquad \text{and} \quad u^\varepsilon = 0 \quad \text{on} \quad \partial\Omega \tag{2.2}$$

in the sense of distribution. The corresponding variational problem reads as: Find $u^\varepsilon \in H_0^1(\Omega;\mathbb{R}^m)$ such that

$$a_\Omega(u^\varepsilon, v) = \langle f, v\rangle_\Omega \quad \text{for all } v \in H_0^1(\Omega;\mathbb{R}^m), \tag{2.3}$$

where for any measurable subset $\widetilde{\Omega}$ of Ω,

$$a_{\widetilde{\Omega}}(u,v) := \int_{\widetilde{\Omega}} \nabla v \cdot A(x/\varepsilon)\nabla u \,\mathrm{d}x \qquad \text{and} \quad \langle f, v\rangle_{\widetilde{\Omega}} = \int_{\widetilde{\Omega}} f(x)\cdot v(x)\,\mathrm{d}x.$$

We shall drop the subscript when the subset is the whole domain Ω.

Ω is triangulated by $\mathcal{T}_h$ that consists of simplices τ with h_τ its diameter and $h = \max_{\tau\in\mathcal{T}_h} h_\tau$. We assume that $\mathcal{T}_h$ is shape-regular in the sense of Ciarlet-Raviart [13]: there exists a chunkiness parameter σ_0 such that $h_\tau/\rho_\tau \le \sigma_0$, where ρ_τ is the diameter of the largest ball inscribed into τ. We also assume that $\mathcal{T}_h$ satisfies the inverse assumption: there exists $\sigma_1 > 0$ such that $h/h_\tau \le \sigma_1$.

For each element τ, we firstly choose an oversampling domain $S = S(\tau) \supset \tau$, which is also a simplex. Let λ_i be the ith barycentric coordinate of the simplex S and $e^\beta = (0, \cdots, 1 \cdots, 0)$ with 1 in the βth position. Denote $Q \in \mathbb{R}^{(d+1)\times m}$ with $Q_i^\beta = \lambda_i e^\beta$ for $i = 1, \cdots, d+1$ and $\beta = 1, \cdots, m$, we find $\psi_i^\beta - Q_i^\beta \in H_0^1(S;\mathbb{R}^m)$ such that

$$a_S(\psi_i^\beta, \varphi) = 0 \quad \text{for all} \quad \varphi \in H_0^1(S;\mathbb{R}^m). \tag{2.4}$$

Next, the basis function ϕ_i^β associated with the node x_i of τ is defined as

$$\phi_i^\beta = c_{ij}^\beta \psi_j^\beta \qquad i = 1, \cdots, d+1 \quad \text{and } \beta = 1, \cdots, m, \tag{2.5}$$

where the coefficients c_{ij}^β are determined by $c_{ik}^\beta Q_k^\beta(x_j) = \delta_{ij} e^\beta$ for any node x_j of τ. The matrix $c^\beta = (c_{ij}^\beta)$ is invertible because $\{\psi_i^\beta\}_{i=1}^{d+1}$ are linear independent over S. For $\phi_i = (\phi_i^1, \phi_i^2, \cdots, \phi_i^m)$, the multiscale finite element space is defined by

$$V_h := \text{Span}\{\phi_i \quad \text{for all nodes } x_i \text{ of } \mathcal{T}_h\}.$$

Note that $V_h \subsetneq H^1(\Omega;\mathbb{R}^m)$ because the functions in V_h may not be continuous across the element boundary. The bilinear form a_h is defined for any $v, w \in V_h$ in a piecewise manner as $a_h(v,w) := \sum_{\tau\in\mathcal{T}_h} a_\tau(v,w)$. The approximation problem reads as: Find $u_h \in V_h^0$ such that

$$a_h(u_h, v) = \langle f, v\rangle \qquad \text{for all} \quad v \in V_h^0, \tag{2.6}$$

where $V_h^0 \colon= \{v \in V_h \mid \text{the degrees of freedom of the nodes on } \partial\Omega \text{ are zero}\}$. It follows from [16, Appendix B] that

$$\| v \|_h := \left(\sum_{\tau \in \mathcal{T}_h} \|\nabla v\|^2_{L^2(\tau)} \right)^{1/2} \tag{2.7}$$

is a norm over V_h^0.

Remark 2.1. The authors in [18] introduced a new MsFEM that allows for the oversampling domain of more general shape, e.g. an element star, which facilitates the implementation of MsFEM, while it is equivalent to the original version [16] if the oversampling domain is a simplex.

3. Quantitative Estimates for Periodic Homogenization of Elliptic System. By the theory of H-convergence [29], the solution u^ε of (2.2) converges weakly to the homogenized solution u_0 in $H^1(\Omega;\mathbb{R}^m)$ as $\varepsilon \to 0$, and u_0 satisfies

$$\mathcal{L}_0(u_0) = f \quad \text{in} \quad \Omega, \qquad u_0 = 0 \quad \text{on} \quad \partial\Omega, \tag{3.1}$$

where $\mathcal{L}_0 = \operatorname{div}(\widehat{A}\nabla)$ with the homogenized coefficients $\widehat{A} = \widehat{a}_{ij}^{\alpha\beta}$ given by

$$\widehat{a}_{ij}^{\alpha\beta} = \fint_Y \left(a_{ij}^{\alpha\beta}(y) + a_{ik}^{\alpha\gamma} \frac{\partial \chi_j^{\gamma\beta}}{\partial y_k} \right) \mathrm{d}y,$$

where the unit cell $Y \colon= [0,1)^d$, and the corrector $\chi(y) = \left(\chi_j^\beta(y)\right) = \left(\chi_j^{\alpha\beta}\right)$ for $j = 1, \cdots, d$ and $\alpha, \beta = 1, \cdots, m$ satisfies the following cell problem: Find $\chi_j^\beta \in H^1_{\mathrm{per}}(Y;\mathbb{R}^m)$ such that $\int_Y \chi_j^\beta \,\mathrm{d}y = 0$ and

$$a_Y(\chi_j^\beta, \psi) = -a_Y(P_j^\beta, \psi) \qquad \text{for all} \quad \psi \in H^1_{\mathrm{per}}(Y;\mathbb{R}^m), \tag{3.2}$$

where $P_j^\beta = y_j e^\beta$, and for all $\phi, \psi \in H^1_{\mathrm{per}}(Y;\mathbb{R}^m)$,

$$a_Y(\phi, \psi) \colon= \int_Y a_{ij}^{\alpha\beta}(y) \frac{\partial \phi^\beta}{\partial y_j} \frac{\partial \psi^\alpha}{\partial y_i} \,\mathrm{d}y.$$

The existence and uniqueness of the solution of (3.2) follows from the ellipticity of A and the Lax-Milgram theorem. Moreover,

$$\left\| \nabla \chi_j^\beta \right\|_{L^2(Y)} \le \Lambda/\lambda \qquad \text{and} \qquad \left\| \chi_j^\beta \right\|_{H^1(Y)} \le C_p \Lambda/\lambda,$$

where C_p is the constant arising from Poincaré's inequality:

$$\| \psi \|_{H^1(Y)} \le C_p \| \nabla \psi \|_{L^2(Y)} \quad \text{for all} \quad \psi \in H^1_{\mathrm{per}}(Y) \quad \text{and} \quad \int_Y \psi \,\mathrm{d}y = 0.$$

By Meyers' regularity result [27, 28], there exists $p > 2$ such that

$$\left\| \nabla \chi_j^\beta \right\|_{L^p(Y)} \le C, \tag{3.3}$$

where the index p and the constant C depending only on λ and Λ. This inequality implies that χ is Hölder continuous when $d = 2$ by the Sobolev embedding theorem [1].

By the De Giorgi-Nash theorem, χ is also Hölder continuous when $d = 3$ and $m = 1$. Hence, for $m = 1, d = 2, 3$ and $m \geq 2, d = 2$, there exists C depending only on λ and Λ such that

$$\left\| \chi_j^\beta \right\|_{L^\infty(Y)} \leq C. \tag{3.4}$$

In case of $d = 3$ and $m \geq 2$, we only have

$$\left\| \chi_j^\beta \right\|_{L^q(Y)} \leq C \qquad \text{for certain} \quad q \geq 6, \tag{3.5}$$

which is a direct consequence of (3.3) and the Sobolev embedding theorem [1].

Another frequently used estimate for the corrector matrix is: For any measurable set D, and for $1 \leq p \leq \infty$, there exists C depends on d and p such that

$$\| \chi(x/\varepsilon) \|_{L^p(D)} \leq C \, |D|^{1/p} \, \| \chi \|_{L^p(Y)} . \tag{3.6}$$

Given the corrector χ, the first order approximation of u^ε is defined by

$$u_1^\varepsilon(x) \colon= u_0(x) + \varepsilon \chi(x/\varepsilon) \nabla u_0(x). \tag{3.7}$$

We summarize the convergence rate of u_1^ε in the following theorem.

THEOREM 3.1. *Assume that A is $1-$periodic and satisfies* (2.1). *Let Ω be a bounded Lipschitz domain in $\mathbb{R}^d$. Let u^ε and u_0 be the weak solutions of* (2.2) *and* (3.1)*, respectively.*

1. *If $u_0 \in W^{2,d}(\Omega; \mathbb{R}^m)$, then*

$$\| u^\varepsilon - u_1^\varepsilon \|_{H^1(\Omega)} \leq C\sqrt{\varepsilon} \, \| \nabla u_0 \|_{W^{1,d}(\Omega)} , \tag{3.8}$$

where C depends on λ, Λ and Ω.

2. *If the corrector χ is bounded and $u_0 \in H^2(\Omega; \mathbb{R}^m)$, then*

$$\| u^\varepsilon - u_1^\varepsilon \|_{H^1(\Omega)} \leq C\sqrt{\varepsilon} \, \| \nabla u_0 \|_{H^1(\Omega)} , \tag{3.9}$$

where C depends $\lambda, \Lambda, \| \chi \|_{L^\infty}$ and Ω.

The estimates (3.8) and (3.9) are taken from [35, Theorem 3.2.7].

We also need the following estimate in certain L^p-norm.

THEOREM 3.2. *Under the same assumption of Theorem 3.1, and assume that $A = A^t$ for $m \geq 2$. Suppose that $u_0 \in W^{2,q}(\Omega; \mathbb{R}^m)$ for $q = 2d/(d + 1)$. Then*

$$\| u^\varepsilon - u_0 \|_{L^p(\Omega)} \leq C\varepsilon \, \| \nabla u_0 \|_{W^{1,q}(\Omega)} , \tag{3.10}$$

where $p = 2d/(d - 1)$ and C depends only on λ, Λ and Ω.

This theorem was proved in [34]; See also [35, Theorem 3.4.3] with

$$\| u^\varepsilon - u_0 \|_{L^p(\Omega)} \leq C\varepsilon \, \| u_0 \|_{W^{2,q}(\Omega)} ,$$

which together with the Ponicaré's inequality leads to (3.10). Moreover, using a scaling argument, we rewrite (3.10) as

$$\| u^\varepsilon - u_0 \|_{L^p(\Omega)} \leq C\varepsilon \Big((\operatorname{diam} \Omega)^{-1} \, \| \nabla u_0 \|_{L^q(\Omega)} + \left\| \nabla^2 u_0 \right\|_{L^q(\Omega)} \Big) , \tag{3.11}$$

where C is independent of the diameter of Ω.

4. Error Estimates for the Periodic Media. Before stating the main result, we make an assumption on the size of the oversampling domain S [8].

Assumption A: There exist constants γ_1 and γ_2 independent of h such that

$$\operatorname{diam} S \le \gamma_1 h_\tau \qquad \text{and} \quad \operatorname{dist}(\partial\tau, \partial S) \ge \gamma_2 h_\tau.$$

Moreover, we always assume that $h > \varepsilon$.

4.1. $\mathbf{H}^1$ error estimate. The main result of this work is

THEOREM 4.1. *Assume that A is $1-$periodic and satisfies the Legendre-Hadamard condition (2.1). For $m \ge 2$, we assume $A = A^t$. Let Ω be a bounded Lipschitz domain in $\mathbb{R}^d$, and let u^ε and u_h be the solutions of Problems (2.3) and (2.6), respectively.*

For $m = 1, d = 2, 3$ or $m \ge 2, d = 2$, if $u_0 \in H^2(\Omega; \mathbb{R}^m)$, then

$$\| u^\varepsilon - u_h \|_{h,\Omega} \le C\left(\sqrt{\varepsilon} + \varepsilon/h + h\right)\Big(\| \nabla u_0 \|_{H^1(\Omega)} + \| f \|_{L^2(\Omega)}\Big), \tag{4.1}$$

where C depends on λ, Λ, Ω and the mesh parameters $\sigma_0, \sigma_1, \gamma_1, \gamma_2$.

For $m \ge 2$ and $d = 3$, if $u_0 \in W^{2,3}(\Omega; \mathbb{R}^m)$, then

$$\| u^\varepsilon - u_h \|_{h,\Omega} \le C\left(\sqrt{\varepsilon} + \varepsilon/h + h\right)\Big(\| \nabla u_0 \|_{W^{1,3}(\Omega)} + \| f \|_{L^2(\Omega)}\Big), \tag{4.2}$$

where C depends on λ, Λ, Ω and the mesh parameters $\sigma_0, \sigma_1, \gamma_1, \gamma_2$.

The implication of the above theorem is as follows.

1. The convergence rate of MsFEM proved above is the same with that in [16] for the scalar elliptic problem in two dimension, while we remove the superfluous technical assumptions on the coefficient a^ε, the homogenized solution u_0 and the correctors χ.
2. The convergence rate of MsFEM is new for elliptic systems as well as problems in three dimension.
3. We clarify the dependence of the right-hand side of the energy error estimates on u_0 and f in the natural Sobolev norms, which together with the Aubin-Nitsche dual argument yields the convergence rate of MsFEM in $\mathrm{L}^{d/(d-1)}-$norm. In particular, we obtain the L^2 error estimate for problem in $d = 2$ and scalar elliptic problem in $d = 3$, cf. Theorem 4.9.
4. It would be interesting to know whether Assumption A can be removed or to what degree it can be weakened. One may start with making clear how the constants C in (4.1) and (4.2) depend on γ_1 and γ_2. Insightful discussion on this point may be found in [18].

The proof of Theorem 4.1 is based on *the second lemma of Strang* [4] because MsFEM with oversampling is a nonconforming method.

$$\| u^\varepsilon - u_h \|_h \le C\left(\inf_{v \in V_h^0} \| u^\varepsilon - v \|_h + \sup_{w \in V_h^0} \frac{|\langle f, w\rangle - a_h(u^\varepsilon, w)|}{\| w \|_h}\right), \tag{4.3}$$

where C depends on $\lambda, \Lambda, \gamma_1$ and γ_2. Therefore, the error estimate boils down to bounding the approximation error and the consistency error. To this end, we firstly define a MsFEM interpolant on each element $\tau \in \mathcal{T}_h$ as

$$\widetilde{u}(x)|_\tau := \sum_{i=1}^{d+1} u_0(x_i)\phi_i(x), \tag{4.4}$$

which may be written as $\widetilde{u}^\beta=\sum_{i=1}^{d+1}\sum_{k=1}^{d+1}u_0^\beta(x_i)c_{ik}^\beta\psi_k^\beta(x)$. It is well-defined over S, and

$$\mathcal{L}_\varepsilon(\widetilde{u})=0\quad\text{in}\quad S\qquad\text{and}\quad\widetilde{u}=\widetilde{u}_0\quad\text{on}\quad\partial S,$$

where $\widetilde{u}_0^\beta=\sum_{i=1}^{d+1}\sum_{k=1}^{d+1}u_0^\beta(x_i)c_{ik}^\beta Q_k^\beta(x)$. It is clear that the homogenization limit of $\widetilde{u}$ is $\widetilde{u}_0$. By definition, $\widetilde{u}_0|_\tau=\pi u_0$ with πu_0 the linear Lagrange interpolant of u_0 over τ. The first order approximation of $\widetilde{u}$ is defined as

$$\widetilde{u}_1^\varepsilon:=\widetilde{u}_0+\varepsilon(\chi\cdot\nabla)\widetilde{u}_0\quad\text{and}\quad\widetilde{u}_1^\varepsilon|_\tau=\pi u_0+\varepsilon(\chi\cdot\nabla)\pi u_0.$$

The approximation error of the MsFEM interpolant is given by

Lemma 4.2. *Under the same assumptions in Theorem 4.1, for $m=1,d=2,3$ or $m\ge 2,d=2$, there holds*

$$\|u^\varepsilon-\widetilde{u}\|_h\le C\left((\sqrt{\varepsilon}+h)\|\nabla u_0\|_{H^1(\Omega)}+\frac{\varepsilon}{h}\|\nabla u_0\|_{L^2(\Omega)}\right),\tag{4.5}$$

where C depends on λ,Λ,Ω and the mesh parameters $\sigma_0,\sigma_1,\gamma_1,\gamma_2$.

Furthermore, for $m\ge 2$ and $d=3$, there holds

$$\|u^\varepsilon-\widetilde{u}\|_h\le C\left((\sqrt{\varepsilon}+h)\|\nabla u_0\|_{W^{1,3}(\Omega)}+\frac{\varepsilon}{h}\|\nabla u_0\|_{L^2(\Omega)}\right),\tag{4.6}$$

where C depends on λ,Λ,Ω and the mesh parameters $\sigma_0,\sigma_1,\gamma_1,\gamma_2$.

Remark 4.3. The interpolation estimate (4.6) is new, while (4.5) with $m=1$ and $d=2$ was proved in [16] by assuming that $\nabla\chi$ is bounded. The proof therein does not apply to elliptic systems because the maximum principle used in the proof may fail for elliptic systems [26]. We shall use the local multiplier estimates in Lemma 4.5 to remove the boundedness assumption on $\nabla\chi$.

The next lemma concerns the estimate of the consistency error.

Lemma 4.4. *Under the same assumptions in Theorem 4.1, for $m=1,d=2,3$ or $m\ge 2,d=2$, there holds*

$$\sup_{w\in V_h^0}\frac{|\langle f,w\rangle-a_h(u^\varepsilon,w)|}{\|w\|_h}\le C\,(\varepsilon+\varepsilon/h)\left(\|\nabla u_0\|_{H^1(\Omega)}+\|f\|_{L^2(\Omega)}\right).\tag{4.7}$$

where C depends on λ,Λ,Ω and the mesh parameters $\sigma_0,\sigma_1,\gamma_1,\gamma_2$.

For $m\ge 2$ and $d=3$, there holds

$$\sup_{w\in V_h^0}\frac{|\langle f,w\rangle-a_h(u^\varepsilon,w)|}{\|w\|_h}\le C\,(\varepsilon+\varepsilon/h)\left(\|\nabla u_0\|_{W^{1,3}(\Omega)}+\|f\|_{L^2(\Omega)}\right),\tag{4.8}$$

where C depends on λ,Λ,Ω and the mesh parameters $\sigma_0,\sigma_1,\gamma_1,\gamma_2$.

Proof of Theorem 4.1 Substituting Lemma 4.2 and Lemma 4.4 into (4.3), we get Theorem 4.1.

4.1.1. Technical Results. The main ingredients in proving Lemma 4.2 and Lemma 4.4 are the following local multiplier estimate, which controls the L^2-norm of $(\nabla\chi)\psi$ for certain ψ, and a local estimate of ∇u_1^ε; cf. Lemma 4.8.

LEMMA 4.5. *Let χ be defined in (3.2) and suppose that D is a convex polyhedron. For any $\psi\in W^{1,d}(D;\mathbb{R}^m)$, there exists C independent of the size of D such that*

$$\varepsilon\,\|\,\nabla\chi(x/\varepsilon)\psi\,\|_{L^2(D)}\le C\,|D|^{1/2-1/d}\Big(\,\|\,\psi\,\|_{L^d(D)}+\varepsilon\,\|\,\nabla\psi\,\|_{L^d(D)}\Big).\tag{4.9}$$

If $\|\,\chi\,\|_{L^\infty}$ is bounded, then for any $\psi\in H^1(D;\mathbb{R}^m)$, there exists C independent of the size of D such that

$$\varepsilon\,\|\,\nabla\chi(x/\varepsilon)\psi\,\|_{L^2(D)}\le C(1+\|\,\chi\,\|_{L^\infty})\Big(\,\|\,\psi\,\|_{L^2(D)}+\varepsilon\,\|\,\nabla\psi\,\|_{L^2(D)}\Big).\tag{4.10}$$

The proof depends on the following multiplier estimates proved in [35, Lemma 3.2.8]: For any $\psi\in W^{1,d}(\Omega;\mathbb{R}^m)$,

$$\varepsilon\,\|\,\nabla\chi(x/\varepsilon)\psi\,\|_{L^2(\Omega)}\le C\Big(\,\|\,\psi\,\|_{L^d(\Omega)}+\varepsilon\,\|\,\nabla\psi\,\|_{L^d(\Omega)}\Big),\tag{4.11}$$

and for any $\psi\in H^1(\Omega;\mathbb{R}^m)$,

$$\varepsilon\,\|\,\nabla\chi(x/\varepsilon)\psi\,\|_{L^2(\Omega)}\le C(1+\|\,\chi\,\|_{L^\infty})\Big(\,\|\,\psi\,\|_{L^2(\Omega)}+\varepsilon\,\|\,\nabla\psi\,\|_{L^2(\Omega)}\Big),\tag{4.12}$$

where C depends on λ,Λ and Ω. These multiplier estimates are crucial to prove the error bounds (3.8) and (3.9). These estimates have been refined in Lemma 4.5 by tracing the dependence of the constant on the size of the domain.

Proof. Denote $L=\operatorname{diam}D$, and we apply the scaling $x'=x/L$ to D so that the rescaled element $\widehat{D}$ has diameter 1. Note that

$$x/\varepsilon=x'/\varepsilon'\qquad\text{with}\quad \varepsilon'=\varepsilon/L.$$

Hence $\varepsilon\nabla\chi(x/\varepsilon)=\varepsilon'\,\nabla_{x'}\chi(x'/\varepsilon')$ and $\psi(x)=\psi(Lx')=\widehat{\psi}(x')$. Applying (4.11) to $\widehat{D}$, we obtain that there exists C depends only on $\widehat{D}$ such that

$$\begin{aligned}\varepsilon\,\|\,\nabla\chi(x/\varepsilon)\psi\,\|_{L^2(D)}&\le\Big(|D|\,/\,\big|\widehat{D}\big|\Big)^{1/2}\varepsilon'\,\Big\|\,\nabla_{x'}\chi(x'/\varepsilon')\widehat{\psi}\,\Big\|_{L^2(\widehat{D})}\\&\le C\,|D|^{1/2}\Big(\,\Big\|\,\widehat{\psi}\,\Big\|_{L^d(\widehat{D})}+\varepsilon'\,\Big\|\,\nabla_{x'}\widehat{\psi}\,\Big\|_{L^d(\widehat{D})}\Big)\\&\le C\,|D|^{1/2-1/d}\Big(\,\|\,\psi\,\|_{L^d(D)}+\varepsilon\,\|\,\nabla\psi\,\|_{L^d(D)}\Big).\end{aligned}$$

This yields (4.9).

Replacing (4.11) by (4.12) and proceeding along the same line that leads to (4.9), we obtain (4.10). ∎

Another ingredient of the error estimate is the quantitative estimates for the MsFEM functions in V_h, which have been used in all the previous study. For any $w\in V_h$, we may write, on each element $\tau\in\mathcal{T}_h$,

$$w^\beta(x)|_\tau:=\sum_{i=1}^{d+1}w_i\phi_i(x)=\sum_{i=1}^{d+1}\sum_{k=1}^{d+1}w_i^\beta c_{ik}^\beta\psi_k^\beta(x)$$

for certain coefficients $w_i\in\mathbb{R}^m$. It is well-defined over S, and

$$\mathcal{L}_\varepsilon(w)=0\quad\text{in}\quad S\qquad\text{and}\quad w=w_0\quad\text{on}\quad\partial S,$$

where $w_0^\beta=\sum_{i=1}^{d+1}\sum_{k=1}^{d+1}w_i^\beta c_{ik}^\beta Q_k^\beta(x)$. It is clear that the homogenization limit of w is w_0, and there exists C depending on λ,Λ,γ_1 and γ_2, but independent of ε and h_τ, such that

$$\|\nabla w_0\|_{L^2(\tau)}\le C\|\nabla w\|_{L^2(\tau)}\qquad\text{for all}\quad \tau\in\mathcal{T}_h. \tag{4.13}$$

This inequality was proved in [16, Appendix B]. The first order approximation of w is defined by $w_1^\varepsilon := w_0+\varepsilon(\chi\cdot\nabla)w_0$.

LEMMA 4.6. *Suppose that* ***Assumption A*** *is true and $A=A^t$ for $m\ge 2$. For $w\in V_h$, there exists C such that*

$$\|w-w_0\|_{L^2(S)}\le C\varepsilon\|\nabla w_0\|_{L^2(S)}, \tag{4.14}$$

and

$$\|\nabla(w-w_1^\varepsilon)\|_{L^2(\tau)}\le C\frac{\varepsilon}{h_\tau}\|\nabla w_0\|_{L^2(S)}. \tag{4.15}$$

Proof. Applying Theorem 3.2 to w, using (3.11) and the fact that w_0 is linear over S, we obtain

$$\begin{aligned}\|w-w_0\|_{L^2(S)}&\le|S|^{1/2-1/p}\|w-w_0\|_{L^p(S)}\\&\le C\varepsilon|S|^{1/2-1/p}\Big((\operatorname{diam}S)^{-1}\|\nabla w_0\|_{L^q(S)}+\big\|\nabla^2w_0\big\|_{L^q(S)}\Big)\\&=C\frac{\varepsilon}{\operatorname{diam}S}|S|^{1/2-1/p+1/q}|\nabla w_0|\\&\le C\varepsilon\|\nabla w_0\|_{L^2(S)},\end{aligned}$$

where we have used $1/q-1/p=1/d$ in the last step. This gives (4.14).

Note that

$$a_S(w-w_1^\varepsilon,v)=0\qquad\text{for all}\quad v\in H_0^1(S;\mathbb{R}^m).$$

By the Caccioppoli inequality [17, Corollary 1.37] and **Assumption A**, there exists C that depends on λ,Λ,γ_1 and γ_2 such that

$$\|\nabla(w-w_1^\varepsilon)\|_{L^2(\tau)}\le\frac{C}{h_\tau}\|w-w_1^\varepsilon\|_{L^2(S)}. \tag{4.16}$$

Using the fact that ∇w_0 is a piecewise constant matrix and (3.6) with $p=2$, we obtain

$$\begin{aligned}\|w_1^\varepsilon-w_0\|_{L^2(S)}&=\varepsilon\|\chi(x/\varepsilon)\nabla w_0\|_{L^2(S)}=\varepsilon\|\chi(x/\varepsilon)\|_{L^2(S)}|\nabla w_0|\\&\le C\varepsilon|S|^{1/2}\|\chi\|_{L^2(Y)}|\nabla w_0|=C\varepsilon\|\chi\|_{L^2(Y)}\|\nabla w_0\|_{L^2(S)},\end{aligned}$$

which together with (4.14) and the triangle inequality gives

$$\|w-w_1^\varepsilon\|_{L^2(S)}\le\|w-w_0\|_{L^2(S)}+\|w_1^\varepsilon-w_0\|_{L^2(S)}\le C\varepsilon\|\nabla w_0\|_{L^2(S)}.$$

Substituting the above inequality into (4.16), we obtain (4.15). ∎

Another useful tool is the following inequality for a tubular domain defined below. Let $\tau\in\mathcal{T}_h$, for any $\delta>0$, we define

$$\tau_\delta:=\{\,x\in\tau\mid\operatorname{dist}(x,\partial\tau)\le\delta\,\}.$$

Lemma 4.7. *Let $1 \le p < \infty$, for any $v \in W^{1,p}(\tau)$, there exists C depending on p, d and σ_0 such that*

$$\| v \|_{L^p(\tau_\delta)} \le C(\delta/h_\tau)^{1/p} \| v \|_{W^{1,p}(\tau)} . \tag{4.17}$$

This inequality has appeared in many occurrences, and we give a proof for the readers' convenience.

Proof. For any $0 < s < \delta$, we let $\tau_s^c = \tau \backslash \tau_s$. It is clear that τ_s^c is also a simplex. For any face f of τ_s^c, we define a vector

$$m(x) = \frac{|f|}{d\,|\tau_s^c|}(x - a_f),$$

where a_f is the vertex opposite to f. A direct calculation gives that $m(x) \cdot n_f = 1$ for any $x \in f$, while $m(x) \cdot n_g$ vanishes on the remaining faces of τ_s^c, where n_g is the outward normal of the face g so that $x \in g$. Using the divergence theorem, we obtain

$$\begin{aligned}\int_f |v(x)|^p \, \mathrm{d}\sigma(x) &= \int_f |v(x)|^p \, m(x) \cdot n_f \mathrm{d}\sigma(x) = \int_{\tau_s^c} \operatorname{div} \left(|v(x)|^p \, m(x) \right) \, \mathrm{d}x \\ &= \int_{\tau_s^c} \left(\left(m(x) \cdot \nabla \right) |v(x)|^p + |v(x)|^p \operatorname{div} m(x) \right) \, \mathrm{d}x.\end{aligned}$$

A direct calculation gives

$$\max_{x \in \tau_s^c} |m(x)| \le \sigma_0 \qquad \operatorname{div} m(x) = \frac{|f|}{|\tau_s^c|} \le \frac{d\sigma_0}{h_\tau}.$$

A combination of the above two inequalities leads to

$$\begin{aligned}\int_f |v(x)|^p \, \mathrm{d}\sigma(x) &\le \sigma_0 \left(\frac{d}{h_\tau} \int_{\tau_s^c} |v(x)|^p \, \mathrm{d}x + p \int_{\tau_s^c} |v(x)|^{p-1} \, |\nabla v(x)| \, \mathrm{d}x \right) \\ &\le \frac{\sigma_0}{h_\tau} \left(d \int_\tau |v(x)|^p \, \mathrm{d}x + p h_\tau \int_\tau |v(x)|^{p-1} \, |\nabla v(x)| \, \mathrm{d}x \right).\end{aligned}$$

Summing up all faces $f \in \partial \tau_s^c$, we obtain

$$\int_{\partial \tau_s^c} |v(x)|^p \, \mathrm{d}\sigma(x) \le \frac{(d+1)\sigma_0}{h_\tau} \left(d \int_\tau |v(x)|^p \, \mathrm{d}x + p h_\tau \int_\tau |v(x)|^{p-1} \, |\nabla v(x)| \, \mathrm{d}x \right).$$

Integration with respect to s from 0 to δ, we obtain

$$\int_{\tau_\delta} |v(x)|^p \, \mathrm{d}\sigma(x) \le \frac{(d+1)\sigma_0 \delta}{h_\tau} \left(d \int_\tau |v(x)|^p \, \mathrm{d}x + p h_\tau \int_\tau |v(x)|^{p-1} \, |\nabla v(x)| \, \mathrm{d}x \right).$$

Using Hölder's inequality, we obtain

$$\| v \|_{L^p(\tau_\delta)} \le (\delta/h_\tau)^{1/p} ((d+1)\sigma_0)^{1/p} \left(d^{1/p} \| v \|_{L^p(\tau)} + (p h_\tau)^{1/p} \| v \|_{L^p(\tau)}^{1-1/p} \| \nabla v \|_{L^p(\tau)}^{1/p} \right).$$

This gives (4.17) for $p > 1$.

The proof for $p = 1$ is the same, we omit the details. ∎

To bound the consistency error, we need a local estimate of ∇u_1^ε, which helps us to remove the extra smoothness assumption on χ.

LEMMA 4.8. *There exists C independent of ε, δ and h_τ such that*

$$\|\nabla u_1^\varepsilon\|_{L^2(\tau_\delta)} \le C\left(\varepsilon + \sqrt{\delta/h_\tau}\right)|\tau|^{1/2-1/d}\|\nabla u_0\|_{W^{1,d}(\tau)}. \tag{4.18}$$

If χ is bounded, then

$$\|\nabla u_1^\varepsilon\|_{L^2(\tau_\delta)} \le C\left(\varepsilon + \sqrt{\delta/h_\tau}\right)\left(1 + \|\chi\|_{L^\infty(Y)}\right)\|\nabla u_0\|_{H^1(\tau)}. \tag{4.19}$$

Proof. Since τ is a simplex, we may decompose τ_δ into $d+1$ disjoint convex domains $\{\tau_\delta^i\}_{i=1}^{d+1}$. Over each τ_δ^i, using the local multiplier estimate (4.9), we obtain

$$\varepsilon\|\nabla\chi(x/\varepsilon)\nabla u_0\|_{L^2(\tau_\delta^i)} \le C\left|\tau_\delta^i\right|^{1/2-1/d}\left(\|\nabla u_0\|_{L^d(\tau_\delta^i)} + \varepsilon\left\|\nabla^2 u_0\right\|_{L^d(\tau_\delta^i)}\right).$$

Summing up the above estimate for $i = 1, \dots, d+1$ and using the scaled trace inequality (4.17) with $p = d$, we obtain

$$\begin{aligned}
\varepsilon\|\nabla\chi(x/\varepsilon)\nabla u_0\|_{L^2(\tau_\delta)} &\le C|\tau_\delta|^{1/2-1/d}\left(\|\nabla u_0\|_{L^d(\tau_\delta)} + \varepsilon\left\|\nabla^2 u_0\right\|_{L^d(\tau_\delta)}\right)\\
&\le C|\tau_\delta|^{1/2-1/d}(\delta/h_\tau)^{1/d}\|\nabla u_0\|_{W^{1,d}(\tau)}\\
&\quad + C\varepsilon|\tau|^{1/2-1/d}\left\|\nabla^2 u_0\right\|_{L^d(\tau)}\\
&\le C(\varepsilon + \sqrt{\delta/h_\tau})|\tau|^{1/2-1/d}\|\nabla u_0\|_{W^{1,d}(\tau)}.
\end{aligned}$$

Invoking the scaled trace inequality (4.17) with $p = 2$ and using Hölder's inequality, we obtain

$$\|\nabla u_0\|_{L^2(\tau_\delta)} \le C\sqrt{\delta/h_\tau}\|\nabla u_0\|_{H^1(\tau)} \le C\sqrt{\delta/h_\tau}|\tau|^{1/2-1/d}\|\nabla u_0\|_{W^{1,d}(\tau)}.$$

Using Hölder's inequality with $1/q = 1/2 - 1/d$ and (3.6) with $p = q$, we obtain

$$\begin{aligned}
\varepsilon\left\|\chi(x/\varepsilon)\nabla^2 u_0\right\|_{L^2(\tau_\delta)} &\le \varepsilon\left\|\chi(x/\varepsilon)\nabla^2 u_0\right\|_{L^2(\tau)} \le \varepsilon\|\chi(x/\varepsilon)\|_{L^q(\tau)}\left\|\nabla^2 u_0\right\|_{L^d(\tau)}\\
&\le C\varepsilon|\tau|^{1/2-1/d}\|\chi\|_{L^q(Y)}\left\|\nabla^2 u_0\right\|_{L^d(\tau)}.
\end{aligned}$$

A combination of the above three inequalities leads to (4.18).

If χ is bounded, then we sum up the local multiplier estimate (4.10) over τ_δ^i for $i = 1, \dots, d+1$ and obtain

$$\varepsilon\|\nabla\chi(x/\varepsilon)\nabla u_0\|_{L^2(\tau_\delta)} \le C(1 + \|\chi\|_{L^\infty(Y)})\left(\|\nabla u_0\|_{L^2(\tau_\delta)} + \varepsilon\left\|\nabla^2 u_0\right\|_{L^2(\tau_\delta)}\right).$$

Invoking the scaled trace inequality (4.17) again, we obtain

$$\begin{aligned}
\|\nabla u_1^\varepsilon\|_{L^2(\tau_\delta)} &\le \|\nabla u_0\|_{L^2(\tau_\delta)} + \varepsilon\|\nabla\chi(x/\varepsilon)\nabla u_0\|_{L^2(\tau_\delta)} + \varepsilon\left\|\chi\nabla^2 u_0\right\|_{L^2(\tau_\delta)}\\
&\le C(1 + \|\chi\|_{L^\infty(Y)})\left(\|\nabla u_0\|_{L^2(\tau_\delta)} + \varepsilon\left\|\nabla^2 u_0\right\|_{L^2(\tau)}\right)\\
&\le C\left(\varepsilon + \sqrt{\delta/h_\tau}\right)(1 + \|\chi\|_{L^\infty(Y)})\left\|\nabla^2 u_0\right\|_{L^2(\tau)}.
\end{aligned}$$

This gives (4.19) and finishes the proof. ∎

4.1.2. Proof of Lemma 4.2 and Lemma 4.4.

Proof for Lemma 4.2 Using the triangle inequality, we have

$$\begin{aligned}\|u^\varepsilon-\widetilde{u}\|_h&\le\|u^\varepsilon-u_1^\varepsilon\|_h+\|\widetilde{u}-\widetilde{u}_1^\varepsilon\|_h+\|u_1^\varepsilon-\widetilde{u}_1^\varepsilon\|_h\\&=\|\nabla(u^\varepsilon-u_1^\varepsilon)\|_{L^2(\Omega)}+\|\widetilde{u}-\widetilde{u}_1^\varepsilon\|_h+\|u_1^\varepsilon-\widetilde{u}_1^\varepsilon\|_h.\end{aligned}\tag{4.20}$$

Applying Lemma 4.6 to $\widetilde{u}$, using (4.15) and **Assumption A**, we obtain

$$\begin{aligned}\|\nabla(\widetilde{u}-\widetilde{u}_1^\varepsilon)\|_{L^2(\tau)}&\le C\frac{\varepsilon}{h_\tau}\|\nabla\widetilde{u}_0\|_{L^2(S)}=C\frac{\varepsilon}{h_\tau}|S|^{1/2}|\nabla\widetilde{u}_0|\\&=C\frac{\varepsilon}{h_\tau}|S|^{1/2}|\nabla\pi u_0|=C\frac{\varepsilon}{h_\tau}\frac{|S|^{1/2}}{|\tau|^{1/2}}\|\nabla\pi u_0\|_{L^2(\tau)}\\&\le C\frac{\varepsilon}{h_\tau}\|\nabla\pi u_0\|_{L^2(\tau)}.\end{aligned}$$

Summing up all $\tau\in\mathcal{T}_h$, using the shape-regular and inverse assumption of $\mathcal{T}_h$, we obtain

$$\begin{aligned}\|\widetilde{u}-\widetilde{u}_1^\varepsilon\|_h&\le C\frac{\varepsilon}{h}\|\nabla\pi u_0\|_{L^2(\Omega)}\le C\frac{\varepsilon}{h}\Big(\|\nabla(u_0-\pi u_0)\|_{L^2(\Omega)}+\|\nabla u_0\|_{L^2(\Omega)}\Big)\\&\le C\Big(\varepsilon\left\|\nabla^2u_0\right\|_{L^2(\Omega)}+\frac{\varepsilon}{h}\|\nabla u_0\|_{L^2(\Omega)}\Big).\end{aligned}\tag{4.21}$$

On each element $\tau, u_1^\varepsilon-\widetilde{u}_1^\varepsilon=u_0-\pi u_0+\varepsilon\chi(x/\varepsilon)\nabla(u_0-\pi u_0)$ and

$$\nabla(u_1^\varepsilon-\widetilde{u}_1^\varepsilon)=\nabla(u_0-\pi u_0)+\varepsilon\nabla\chi(x/\varepsilon)\nabla(u_0-\pi u_0)+\varepsilon\chi(x/\varepsilon)\nabla^2u_0.$$

For $m=1,d=2,3$ or $m\ge2,d=2$, χ is bounded by (3.4), using the local multiplier inequality (4.10), we obtain

$$\begin{aligned}\varepsilon\|\nabla\chi(x/\varepsilon)\nabla(u_0-\pi u_0)\|_{L^2(\tau)}&\le C\Big(\|\nabla(u_0-\pi u_0)\|_{L^2(\tau)}+\varepsilon\left\|\nabla^2u_0\right\|_{L^2(\tau)}\Big)\\&\le C(\varepsilon+h_\tau)\left\|\nabla^2u_0\right\|_{L^2(\tau)}.\end{aligned}$$

It follows from the above two equations that

$$\begin{aligned}\|\nabla(u_1^\varepsilon-\widetilde{u}_1^\varepsilon)\|_{L^2(\tau)}&\le\|\nabla(u_0-\pi u_0)\|_{L^2(\tau)}+\varepsilon\|\nabla\chi(x/\varepsilon)\nabla(u_0-\pi u_0)\|_{L^2(\tau)}\\&\quad+\varepsilon\left\|\chi(x/\varepsilon)\nabla^2u_0\right\|_{L^2(\tau)}\\&\le C\Big(1+\|\chi\|_{L^\infty(Y)}\Big)(\varepsilon+h_\tau)\left\|\nabla^2u_0\right\|_{L^2(\tau)}.\end{aligned}$$

Summing up all $\tau\in\mathcal{T}_h$, and using (3.4) again, we get

$$\|u_1^\varepsilon-\widetilde{u}_1^\varepsilon\|_h\le C(\varepsilon+h)\left\|\nabla^2u_0\right\|_{L^2(\Omega)}.\tag{4.22}$$

Substituting the above inequality, (3.9) and (4.21) into (4.20), we obtain (4.5).

For $m\ge2$ and $d=3$, by (3.5), we have $\chi\in L^6(Y)$. Using the local multiplier estimate (4.9) and the standard interpolation estimate for πu_0, we obtain

$$\begin{aligned}\varepsilon\|\nabla\chi(x/\varepsilon)\nabla(u_0-\pi u_0)\|_{L^2(\tau)}&\le C|\tau|^{1/6}\Big(\|\nabla(u_0-\pi u_0)\|_{L^3(\tau)}+\varepsilon\left\|\nabla^2u_0\right\|_{L^3(\tau)}\Big)\\&\le C(\varepsilon+h_\tau)|\tau|^{1/6}\left\|\nabla^2u_0\right\|_{L^3(\tau)}.\end{aligned}$$

Using Hölder's inequality, the inequality (3.6) with $p=6, D=\tau$ and (3.5), we obtain

$$\varepsilon \left\| \chi(x/\varepsilon)\nabla^2 u_0 \right\|_{L^2(\tau)} \le \varepsilon \left\| \chi(x/\varepsilon) \right\|_{L^6(\tau)} \left\| \nabla^2 u_0 \right\|_{L^3(\tau)} \le C\varepsilon \left|\tau\right|^{1/6} \left\| \nabla^2 u_0 \right\|_{L^3(\tau)}.$$

Proceeding along the same line that leads to (4.22), we obtain

$$\left\| \nabla(u_1^\varepsilon - \widetilde{u}_1^\varepsilon) \right\|_{L^2(\tau)} \le C(\varepsilon + h_\tau) \left|\tau\right|^{1/6} \left\| \nabla^2 u_0 \right\|_{L^3(\tau)}.$$

Summing up all $\tau \in \mathcal{T}_h$ and using Hölder's inequality, we get

$$\left\| u_1^\varepsilon - \widetilde{u}_1^\varepsilon \right\|_h \le C(\varepsilon + h) \left\| \nabla^2 u_0 \right\|_{L^3(\Omega)}.$$

Substituting the above inequality, (3.8) and (4.21) into (4.20), we obtain (4.6).

Proof for Lemma 4.4 For $w \in V_h^0$, over each oversampling domain S, let w_0 be its homogenized part over S. By $w_0 \in H_0^1(\Omega;\mathbb{R}^m)$, there holds

$$a_h(u^\varepsilon, w_0) = \langle f, w_0 \rangle.$$

Therefore, we write the consistency error functional as

$$\begin{aligned} \langle f, w\rangle - a_h(u^\varepsilon, w) &= \langle f, w - w_0\rangle - a_h(u^\varepsilon, w - w_0) \\ &= \langle f, w - w_0\rangle - a_h(u^\varepsilon, w - w_1^\varepsilon) - a_h(u^\varepsilon, w_1^\varepsilon - w_0). \end{aligned}$$

Using Lemma 4.6, (4.14), (4.13) and **Assumption** A, we obtain

$$\begin{aligned} \left\| w - w_0 \right\|_{L^2(\tau)} &\le \left\| w - w_0 \right\|_{L^2(S)} \le C\varepsilon \left\| \nabla w_0 \right\|_{L^2(S)} \\ &\le C\varepsilon \left\| \nabla w_0 \right\|_{L^2(\tau)} \le C\varepsilon \left\| \nabla w \right\|_{L^2(\tau)}, \end{aligned}$$

which immediately implies

$$|\langle f, w - w_0\rangle| \le C\varepsilon \left\| f \right\|_{L^2(\Omega)} \left\| w \right\|_h. \tag{4.23}$$

Using (4.15), (4.13) again, and the inverse assumption of $\mathcal{T}_h$, we obtain

$$\begin{aligned} |a_h(u^\varepsilon, w - w_1^\varepsilon)| &\le \Lambda \sum_{\tau\in\mathcal{T}_h} \left\| \nabla u^\varepsilon \right\|_{L^2(\tau)} \left\| \nabla(w - w_1^\varepsilon) \right\|_{L^2(\tau)} \\ &\le C \sum_{\tau\in\mathcal{T}_h} \frac{\varepsilon}{h_\tau} \left\| \nabla u^\varepsilon \right\|_{L^2(\tau)} \left\| \nabla w_0 \right\|_{L^2(\tau)} \\ &\le C\frac{\varepsilon}{h} \sum_{\tau\in\mathcal{T}_h} \left\| \nabla u^\varepsilon \right\|_{L^2(\tau)} \left\| \nabla w \right\|_{L^2(\tau)} \\ &\le C\frac{\varepsilon}{h} \left\| \nabla u^\varepsilon \right\|_{L^2(\Omega)} \left\| w \right\|_h. \end{aligned}$$

Combining the above two estimates, we obtain

$$|\langle f, w - w_0\rangle - a_h(u^\varepsilon, w - w_1^\varepsilon)| \le C\,(\,\varepsilon + \varepsilon/h\,) \left\| f \right\|_{L^2(\Omega)} \left\| w \right\|_h, \tag{4.24}$$

where we have used the a-priori estimate $\left\| \nabla u^\varepsilon \right\|_{L^2(\Omega)} \le C \left\| f \right\|_{L^2(\Omega)}$.

It remains to bound $a_h(u^\varepsilon, w_1^\varepsilon - w_0)$. On each element τ, we introduce a cut-off function $\rho_\varepsilon \in C_0^\infty(\tau)$ such that $0 \le \rho_\varepsilon \le 1$ and $|\nabla\rho_\varepsilon| \le C/\varepsilon$, moreover,

$$\rho_\varepsilon = \begin{cases} 1 & \operatorname{dist}(x, \partial\tau) \ge 2\varepsilon, \\ 0 & \operatorname{dist}(x, \partial\tau) \le \varepsilon. \end{cases}$$

Denote $\widehat{w}^\varepsilon=(w_1^\varepsilon-w_0)(1-\rho_\varepsilon)$, which is the oscillatory part of w_1^ε supported inside the strip $\tau_{2\varepsilon}$. We write

$$\begin{aligned}a_\tau(u^\varepsilon,w_1^\varepsilon-w_0)=&a_\tau(u^\varepsilon,(w_1^\varepsilon-w_0)\rho_\varepsilon)+a_\tau(u^\varepsilon,\widehat{w}^\varepsilon)\\=&\langle f,(w_1^\varepsilon-w_0)\rho_\varepsilon\rangle_\tau+a_\tau(u^\varepsilon,\widehat{w}^\varepsilon).\end{aligned}$$

Using (3.6) with $p=2$, we obtain

$$\begin{aligned}|\langle f,(w_1^\varepsilon-w_0)\rho_\varepsilon\rangle_\tau|&\le\varepsilon\,\|\,f\,\|_{L^2(\tau)}\,\|\,\chi(x/\varepsilon)\,\|_{L^2(\tau)}\,|\nabla w_0|\\&\le C\varepsilon\,|\tau|^{1/2}\,\|\,f\,\|_{L^2(\tau)}\,\|\,\chi\,\|_{L^2(Y)}\,|\nabla w_0|\\&=C\varepsilon\,\|\,f\,\|_{L^2(\tau)}\,\|\,\chi\,\|_{L^2(Y)}\,\|\,\nabla w_0\,\|_{L^2(\tau)}\,.\end{aligned}\tag{4.25}$$

A direct calculation gives[1]

$$\|\,\nabla\widehat{w}^\varepsilon\,\|_{L^2(\tau_{2\varepsilon})}\le C\sqrt{\varepsilon/h_\tau}\,\|\,\nabla w_0\,\|_{L^2(\tau)}\,,\tag{4.26}$$

which together with the local estimate (4.18) implies that, for $m\ge2$ and $d=3$, there holds

$$\begin{aligned}&|a_\tau(u^\varepsilon,\widehat{w}^\varepsilon)|\le|a_\tau(u_1^\varepsilon,\widehat{w}^\varepsilon)|+|a_\tau(u^\varepsilon-u_1^\varepsilon,\widehat{w}^\varepsilon)|\\&\le C\left(\left(\varepsilon+\frac{\varepsilon}{h_\tau}\right)|\tau|^{1/6}\,\|\,\nabla u_0\,\|_{W^{1,3}(\tau)}+\sqrt{\frac{\varepsilon}{h_\tau}}\,\|\,\nabla(u^\varepsilon-u_1^\varepsilon)\,\|_{L^2(\tau)}\right)\|\,\nabla w_0\,\|_{L^2(\tau)}\,.\end{aligned}$$

This estimate together with (4.25) implies

$$\begin{aligned}|a_\tau(u^\varepsilon,w_1^\varepsilon-w_0)|\le C\Big(&\left(\varepsilon+\frac{\varepsilon}{h_\tau}\right)|\tau|^{1/6}\,\|\,\nabla u_0\,\|_{W^{1,3}(\tau)}+\sqrt{\frac{\varepsilon}{h_\tau}}\,\|\,\nabla(u^\varepsilon-u_1^\varepsilon)\,\|_{L^2(\tau)}\\&+\varepsilon\,\|\,f\,\|_{L^2(\tau)}\Big)\,\|\,\nabla w_0\,\|_{L^2(\tau)}\,.\end{aligned}$$

Summing up the above estimates for all $\tau\in\mathcal{T}_h$, using (4.13), (3.9), the inverse assumption of $\mathcal{T}_h$ and Hölder's inequality, we obtain

$$\begin{aligned}|a_h(u^\varepsilon,w_1^\varepsilon-w_0)|\le C\Big(&\left(\varepsilon+\frac{\varepsilon}{h}\right)\|\,\nabla u_0\,\|_{W^{1,3}(\Omega)}+\sqrt{\frac{\varepsilon}{h}}\,\|\,\nabla(u^\varepsilon-u_1^\varepsilon)\,\|_{L^2(\Omega)}\\&+\varepsilon\,\|\,f\,\|_{L^2(\Omega)}\Big)\,\|\,w\,\|_h\\\le& C\left(\varepsilon+\frac{\varepsilon}{h}\right)\left(\|\,\nabla u_0\,\|_{W^{1,3}(\Omega)}+\|\,f\,\|_{L^2(\Omega)}\right)\|\,w\,\|_h\,.\end{aligned}$$

This inequality together with (4.24) implies (4.8).

For $m=1,d=2,3$ or $m\ge2,d=2$, χ is bounded. Replacing (4.18) by (4.19) and proceeding along the same line that leads to (4.8), we obtain (4.7).

4.2. $\mathbf{L}^{d/(d-1)}$ error estimate.

We exploit the Aubin-Nitsche trick to obtain the error estimate of MsFEM in $L^{d/(d-1)}-$norm with $d=2,3$.

THEOREM 4.9. *Under the same assumption of Theorem 4.1, and suppose that* $\varphi\in H_0^1(\Omega;\mathbb{R}^m)$ *satisfying*

$$\int_\Omega\nabla\varphi\cdot\widehat{A}\nabla\psi\,\mathrm{d}x=\langle F,\psi\rangle\qquad\textit{for all}\quad\psi\in H_0^1(\Omega;\mathbb{R}^m).$$

[1] We may also refer to [14, Lemma 3.1] for a proof of (4.26).

For $m=1, d=2,3$ or $m\ge 2, d=2$, if the shift estimate

$$\| \varphi \|_{H^2(\Omega)} \le C \| F \|_{L^2(\Omega)} \tag{4.27}$$

holds true, then for $m=1, d=2,3$, there holds

$$\| u-u_h \|_{L^2(\Omega)} \le C(\varepsilon+h^2+\varepsilon/h)\Big(\| \nabla u_0 \|_{H^1(\Omega)} + \| f \|_{L^2(\Omega)} \Big). \tag{4.28}$$

For $m\ge 2, d=2$, there holds

$$\| u-u_h \|_{L^2(\Omega)} \le C(\varepsilon+h^2+\varepsilon/h) \| f \|_{L^2(\Omega)}. \tag{4.29}$$

For $m\ge 2$ and $d=3$, if the shift estimate

$$\| \varphi \|_{W^{2,3}(\Omega)} \le C \| F \|_{L^3(\Omega)} \tag{4.30}$$

holds true, then

$$\| u-u_h \|_{L^{3/2}(\Omega)} \le C(\varepsilon+h^2+\varepsilon/h) \| f \|_{L^3(\Omega)}. \tag{4.31}$$

Except the resonance error ε/h, the other two items in the above error estimates are *optimal*. For scalar elliptic equation and elliptic systems in two dimension, we obtain the L^2 error estimate.

Proof. For any $g\in L^2(\Omega;\mathbb{R}^m)$, we find $v^\varepsilon\in H^1_0(\Omega;\mathbb{R}^m)$ such that

$$\int_\Omega \nabla w\cdot (A(x/\varepsilon))^t \nabla v^\varepsilon \,\mathrm{d}x = \int_\Omega g\cdot w\,\mathrm{d}x \qquad \text{for all}\quad w\in H^1_0(\Omega;\mathbb{R}^m). \tag{4.32}$$

Let v_h be the MsFEM approximation of v^ε defined by

$$a_h(w,v_h) = \int_\Omega g\cdot w\,\mathrm{d}x \qquad \text{for all}\quad w\in V_h^0. \tag{4.33}$$

It follows from (4.32) and (4.33) that

$$\begin{aligned}\int_\Omega g\cdot (u^\varepsilon-u_h)\,\mathrm{d}x &= a(u^\varepsilon,v^\varepsilon) - a_h(u_h,v_h)\\ &= a_h(u^\varepsilon-u_h, v^\varepsilon-v_h) + a_h(u^\varepsilon-u_h,v_h) + a_h(u_h,v^\varepsilon-v_h)\\ &= a_h(u^\varepsilon-u_h,v^\varepsilon-v_h)\\ &\quad + \big[a_h(u^\varepsilon,v_h) - \langle f,v_h\rangle + a_h(u_h,v^\varepsilon) - \langle g,u_h\rangle\big].\end{aligned}$$

For $m=1, d=2,3$ or $m\ge 2, d=2$, using the energy error estimate (4.1) and the regularity assumption (4.27), we obtain

$$\begin{aligned}|a_h(u^\varepsilon-u_h,v^\varepsilon-v_h)| &\le \Lambda \| u^\varepsilon-u_h \|_h \| v^\varepsilon-v_h \|_h\\ &\le C(\varepsilon+h^2+\varepsilon^2/h^2)\Big(\| \nabla u_0 \|_{H^1(\Omega)} + \| f \|_{L^2(\Omega)} \Big) \| g \|_{L^2(\Omega)}.\end{aligned}$$

Using (4.7) and (4.27), we bound the consistency error functional as

$$\begin{aligned}&|a_h(u^\varepsilon,v_h) - \langle f,v_h\rangle + a_h(u_h,v^\varepsilon) - \langle g,u_h\rangle|\\ &\le C(\varepsilon+\varepsilon/h)\Big(\| \nabla u_0 \|_{H^1(\Omega)} + \| f \|_{L^2(\Omega)} \Big) \| g \|_{L^2(\Omega)}.\end{aligned}$$

A combination of the above three estimates yields (4.28).

For $m \ge 2, d = 2$, noting that $A = A^t$ and the shift estimate (4.27) is also valid for u_0, this gives (4.29).

For $m \ge 2$ and $d = 3$, χ is unbounded. Replacing (4.27), (4.1) and (4.7) by (4.30), (4.2) and (4.8), respectively, and proceeding along the same line that leads to (4.28), we obtain

$$\| u - u_h \|_{L^{3/2}(\Omega)} \le C(\varepsilon + h^2 + \varepsilon/h) \Big(\| \nabla u_0 \|_{W^{1,3}(\Omega)} + \| f \|_{L^3(\Omega)} \Big).$$

Noting that $A^t = A$ and the shift estimate (4.30) is also valid for u_0, this gives (4.31).□

4.3. Error estimates for MsFEM without oversampling. We visit the error estimates of MsFEM without oversampling [19]. The multiscale basis function is $\phi^\beta = \{\phi_i^\beta\}_{i=1}^{d+1}$ is constructed as (2.4) with $S(\tau)$ replaced by τ.

$$V_h := \text{Span}\{\phi_i \quad \text{for all nodes } x_i \text{ of } \mathcal{T}_h\},$$

and $V_h^0 := \{\, v \in V_h \mid v = 0 \quad \text{on } \partial\Omega \,\}$. The approximation problem reads as: Find $u_h \in V_h^0$ such that

$$a(u_h, v) = \langle f, v \rangle \qquad \text{for all} \quad v \in V_h^0. \tag{4.34}$$

Under the same assumptions of Theorem 4.1 except that A is not necessarily symmetric when $m \ge 2$, we prove the energy error estimate for MsFEM without oversampling.

THEOREM 4.10. *Assume A is 1−periodic and satisfies the Legendre-Hadamard condition (2.1). Let Ω be a bounded Lipschitz domain in $\mathbb{R}^d$. Let u^ε and u_h be the solutions of (2.3) and (4.34), respectively.*

For $m = 1, d = 2, 3$ or $m \ge 2, d = 2$, if $u_0 \in H^2(\Omega; \mathbb{R}^m)$, then

$$\| \nabla(u^\varepsilon - u_h) \|_{L^2(\Omega)} \le C \Big((\sqrt{\varepsilon} + h) \| \nabla u_0 \|_{H^1(\Omega)} + \sqrt{\varepsilon/h} \| \nabla u_0 \|_{L^2(\Omega)} \Big), \tag{4.35}$$

where C depends on λ, Λ, Ω and the mesh parameters σ_0 and σ_1.

For $m \ge 2$ and $d = 3$, if $u_0 \in W^{2,3}(\Omega; \mathbb{R}^m)$, then

$$\| \nabla(u^\varepsilon - u_h) \|_{L^2(\Omega)} \le C \Big((\sqrt{\varepsilon} + h) \| \nabla u_0 \|_{W^{1,3}(\Omega)} + \sqrt{\varepsilon/h} \| \nabla u_0 \|_{L^2(\Omega)} \Big), \tag{4.36}$$

where C depends on λ, Λ, Ω and the mesh parameters σ_0 and σ_1.

As a direct consequence of the above theorem, we obtain the $\mathrm{L}^{d/(d-1)}$ error estimate for MsFEM without oversampling. The proof follows the same line that leads to Theorem 4.9, we omit the proof.

Corollary 4.11. Under the same assumption of Theorem 4.9 except that A is not necessarily symmetric for $m \ge 2$. Let u^ε and u_h be the solutions of (2.3) and (4.34), respectively. For $m = 1, d = 2, 3$ or $m \ge 2, d = 2$, there holds

$$\| u - u_h \|_{L^2(\Omega)} \le C(\varepsilon + h^2 + \varepsilon/h) \| \nabla u_0 \|_{H^1(\Omega)}.$$

For $m \ge 2$ and $d = 3$, there holds

$$\| u - u_h \|_{L^{3/2}(\Omega)} \le C(\varepsilon + h^2 + \varepsilon/h) \| \nabla u_0 \|_{W^{1,3}(\Omega)}.$$

The proof of Theorem 4.10 relies on Theorem 3.1 and Lemma 4.5. We only sketch the main steps because the details are the same with the line leading to Theorem 4.1.

Proof of Theorem 4.10 Noting that MsFEM without oversampling is conforming, i.e., $V_h^0 \subset H_0^1(\Omega;\mathbb{R}^m)$, we obtain

$$\|\nabla(u^\varepsilon - u_h)\|_{L^2(\Omega)} \le (1+\Lambda/\lambda) \inf_{v\in V_h^0} \|\nabla(u^\varepsilon - v)\|_{L^2(\Omega)}. \tag{4.37}$$

Define MsFEM interpolant $\widetilde{u}(x)$ as (4.4). Using the triangle inequality, we obtain

$$\|\nabla(u^\varepsilon - \widetilde{u})\|_{L^2(\Omega)} \le \|\nabla(u^\varepsilon - u_1^\varepsilon)\|_{L^2(\Omega)} + \|\nabla(\widetilde{u} - \widetilde{u}_1^\varepsilon)\|_{L^2(\Omega)} + \|\nabla(u_1^\varepsilon - \widetilde{u}_1^\varepsilon)\|_{L^2(\Omega)}.$$

The estimate of $\|\nabla(u^\varepsilon - u_1^\varepsilon)\|_{L^2(\Omega)}$ follows from Theorem 3.1, and the estimate of $\|\nabla(u_1^\varepsilon - \widetilde{u}_1^\varepsilon)\|_{L^2(\Omega)}$ is the same with the corresponding term in Lemma 4.2. Note that $\widetilde{u}_1^\varepsilon$ is the first order approximation of $\widetilde{u}$ over τ. For $m=1, d=2,3$ or $m\ge 2, d=2$, using (3.9), we get

$$\begin{aligned}\|\nabla(\widetilde{u} - \widetilde{u}_1^\varepsilon)\|_{L^2(\tau)} &\le C\sqrt{\varepsilon/h_\tau}\,\|\nabla\pi u_0\|_{L^2(\tau)}\\ &\le C\Big(\sqrt{\varepsilon/h_\tau}\,\|\nabla u_0\|_{L^2(\tau)} + \sqrt{\varepsilon h_\tau}\,\|\nabla u_0\|_{H^1(\tau)}\Big).\end{aligned}$$

Summing up the above estimate for all $\tau\in\mathcal{T}_h$, and using the inverse assumption of $\mathcal{T}_h$, we obtain

$$\|\nabla(\widetilde{u} - \widetilde{u}_1^\varepsilon)\|_{L^2(\Omega)} \le C\Big(\sqrt{\varepsilon/h}\,\|\nabla u_0\|_{L^2(\Omega)} + \sqrt{\varepsilon h}\,\|\nabla u_0\|_{H^1(\Omega)}\Big). \tag{4.38}$$

For $m\ge 2$ and $d=3$, using (3.8) and the fact that $\nabla\pi u_0$ is a piecewise constant matrix over τ, we get

$$\|\nabla(\widetilde{u} - \widetilde{u}_1^\varepsilon)\|_{L^2(\tau)} \le C\sqrt{\varepsilon/h_\tau}\,|\tau|^{1/6}\,\|\nabla\pi u_0\|_{L^3(\tau)} = C\sqrt{\varepsilon/h_\tau}\,\|\nabla\pi u_0\|_{L^2(\tau)}.$$

Proceeding along the same line that leads to (4.38), we obtain

$$\|\nabla(\widetilde{u} - \widetilde{u}_1^\varepsilon)\|_{L^2(\Omega)} \le C\Big(\sqrt{\varepsilon/h}\,\|\nabla u_0\|_{L^2(\Omega)} + \sqrt{\varepsilon h}\,\|\nabla u_0\|_{H^1(\Omega)}\Big).$$

A combination of all the above estimates completes the proof.

Remark 4.12. We have used Theorem 3.1 to bound $\|\nabla(\widetilde{u} - \widetilde{u}_1^\varepsilon)\|_{L^2(\tau)}$ instead of Lemma 4.6, we need not assume the symmetry of A when $m\ge 2$.

5. Conclusion. Under suitable regularity assumptions on the homogenized solution, we proved the optimal energy error estimates for MsFEM with or without oversampling applying to elliptic systems with bounded measurable periodic coefficients. The present work may be extended to elliptic system with locally periodic coefficients, i.e., $A^\varepsilon = A(x, x/\varepsilon)$ with the aid of a new local multiplier estimate. The extension to elliptic system for the coefficients with stratified structure is also very interesting. We believe that the machineries developed in the present work may be useful to analyze other MsFEM such as the mixed MsFEM [8], Crouzeix-Raviart MsFEM [23], or MsFEM with different oversampling techniques [16]. We shall leave these for further pursuit.

Acknowledgments. We gratefully acknowledge the helpful suggestions made by the anonymous referees, which greatly improved the presentation of the paper.

REFERENCES

[1] R. Adams and J. Fournier, *Sobolev Spaces*, Academic Press, 2nd ed., 2003.

[2] R. Altmann, P. Henning, and D. Peterseim, *Numerucal homogenization beyond scale separation*, Acta Numerica, (2021), pp. 1–86.

[3] J. Aubin, *Behaviour of the error of the approximation solution of boundary value problems for linear elliptic operators and by Galerkin's and finite difference methods*, Ann. Scuols. Norm. Sup. Pisa, 21 (1967), pp. 599–637.

[4] A. Berger, R. Scott, and G. Strang, *Approximate boundary conditions in the finite element method*, Symposia Mathematica, X (1972), pp. 295–313.

[5] K. Chen, S. Chen, Q. Li, J. Lu, and J. S. Wright, *Low-rank approximation for multiscale pdes*, Notice Amer. Math. Soc., 69 (2022), pp. 901–913.

[6] Y. Chen and T. Y. Hou, *Multiscale elliptic PDEs upscaling and function approximation via subsampled data*, Multiscale Model. Simul., 20 (2022), pp. 188–219.

[7] Y. Chen, T. Y. Hou, and Y. Wang, *Exponential convergence for multiscale linear elliptic PDEs via adaptive edge basis functions*, Multiscale Model. Simul., 19 (2021), pp. 980–1010.

[8] Z. Chen and T. Y. Hou, *A mixed multiscale finite element method for elliptic problems with oscillating coefficients*, Math. Comp., 72 (2003), pp. 541–576.

[9] Z. Chen and H. Wu, *Selected Topics in Finite Element Method*, Science Press, 2010.

[10] M. Chipot, D. Kinderlehrer, and G. Vergara-Caffarelli, *Smoothness of linear laminates*, Arch. Rational Mech. Anal., 96 (1986), pp. 81–96.

[11] E. T. Chung, Y. Efendiev, and W. T. Leung, *Generalized multiscale finite element methods for wave propagation in heterogeneous media*, Multiscale Model. Simul., 12 (2014), pp. 1691–1721.

[12] A. Cianchi, *Maximizing the L^∞ norm of the gradient of solutions to the Poisson equation*, J. Geom. Anal., 2 (1992), pp. 213–227.

[13] P. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.

[14] R. Du and P. Ming, *Convergence of the heterogeneous multiscale finite element method for elliptic problem with nonsmooth microstructures*, Multiscale Model. Simul., 8 (2010), pp. 1770–1783.

[15] Y. Efendiev and T. Y. Hou, *Multiscale Finite Element Methods*, vol. 4 of Surveys and Tutorials in the Applied Mathematical Sciences, Springer, New York, 2009. Theory and applications.

[16] Y. Efendiev, T. Y. Hou, and X. Wu, *Convergence of a nonconforming multiscale finite element method*, SIAM J. Numer. Anal., 37 (2000), pp. 888–910.

[17] Q. Han and F. Lin, *Ellipitic Partial Differential Equations*, American Mathematical Society, 1997.

[18] P. Henning and D. Peterseim, *Oversampling for the multiscale finite element method*, Multiscale Model. Simul., 11 (2013), pp. 1149–1175.

[19] T. Y. Hou and X. Wu, *A multiscale finite element method for elliptic problems in composite materials and porous media*, J. Comput. Phys., 134 (1997), pp. 169–189.

[20] T. Y. Hou, X. Wu, and C. Z., *Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients*, Math. Comp., 68 (1999), pp. 913–943.

[21] S. Kesavan, *Homogenization of elliptic eigenvalue problems*, Appl. Math. Opt., 5 (1979), pp. Part I, 153–167; Part II, 197–216.

[22] G. Kresin and V. Maz'ya, *Maximum Principles and Sharp Constants for Solutions of Elliptic and Parabolic Systems*, vol. 183 of Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 2012.

[23] C. Le Bris, F. Legoll, and A. Lozinski, *MsFEM à la Crouzeix-Raviart for highly oscillatory elliptic problems*, in Partial differential equations: theory, control and approximation, Springer, Dordrecht, 2014, pp. 265–294.

[24] Y. Li and L. Nirenberg, *Estimates for elliptic systems from composite material*, Comm. Pure Appl. Math., 56 (2003), pp. 892–925. Dedicated to the memory of Jürgen K. Moser.

[25] Y. Li and M. Vogelius, *Gradient estimates for solutions to divergence form elliptic equations with discontinuous coefficients*, Arch. Ration. Mech. Anal., 153 (2000), pp. 91–151.

[26] V. Maz'ya and J. Rossmann, *Elliptic Equations in Polyhedral Domains*, vol. 162 of Mathe-

matical Surveys and Monographs, American Mathematical Society, Providence, RI, 2010.
[27] N. Meyers, *An L^p-estimate for the gradient of solutions of second order elliptic divergence equations*, Ann. Scuola Norm. Sup. Pisa, 17 (1963), pp. 189–206.
[28] N. Meyers, *Some results on regularity for solutions of non-linear elliptic systems and quasi-regular functions*, Duke Math. J., 42 (1975), pp. 121–136.
[29] F. Murat and L. Tartar, *H-convergence*, in Topics in the mathematical modelling of composite materials, vol. 31 of Progr. Nonlinear Differential Equations Appl., Birkhäuser Boston, Boston, MA, 1997, pp. 21–43.
[30] J. Nitsche, *Ein Kriterium für die Quasioptimalität des Ritzschen Verfahrens*, Numer. Math., 11 (1968), pp. 346–348.
[31] H. Owhadi, *Multigrid with rough coefficients and multiresolution operator decomposition from hierarichical information games*, SIAM Rev., 59 (2017), pp. 99–149.
[32] H. Owhadi and C. Scovel, *Operator-Adapted Wavelets, Fast Solvers, and Numerical Homogenization*, Cambridge University Press, 2019.
[33] M. Sarkis and H. Versieux, *Convergence analysis for the numerical boundary corrector for elliptic equations with rapidly oscillating coefficients*, SIAM J. Numer. Anal., 46 (2008), pp. 545–576.
[34] Z. Shen, *Boundary estimates in elliptic homogenization*, Anal. PDE, 10 (2017), pp. 653–694.
[35] Z. Shen, *Periodic Homogenization of Elliptic Systems*, Spriger Nature Switzerland AG, 2018.
[36] S. Torquato, *Random Heterogeneous Materials: Microstructure and Macroscopic Properties*, Springer-Verlag, New York, Inc., 2002.
[37] C. Ye, H. Dong, and J. Cui, *Convergence rate of multiscale finite element method for various boundary problems*, J. Comput. Appl. Math., 374 (2020), p. 112754.
[38] V. Zhikov and S. Pastukhova, *On the operator estimates for some problems in homogenization theory*, Russ. J. Math. Phys., 12 (2005), pp. 515–524.