

ANALYSIS OF THE HETEROGENEOUS MULTISCALE METHOD  
FOR ELLIPTIC HOMOGENIZATION PROBLEMS

WEINAN E, PINGBING MING, AND PINGWEN ZHANG

CONTENTS

1. Introduction and main results	122
1.1. General methodology	122
1.2. Heterogeneous multiscale method	122
1.3. Main results	125
1.4. Recovering the microstructural information	126
1.5. Some technical background	127
2. Generalities	129
3. Estimating $e(\text{HMM})$	132
4. Reconstruction and compression	136
4.1. Reconstruction procedure	136
4.2. Compression operator	138
5. Nonlinear homogenization problems	139
5.1. Algorithms and main results	139
5.2. Estimating $e(\text{HMM})$	143
Appendix A. Estimating $e(\text{HMM})$ for problems with random coefficients	149
References	154

---

Received by the editors January 2, 2003.

2000 *Mathematics Subject Classification*. Primary 65N30, 74Q05; Secondary 74Q15, 65C30.

*Key words and phrases*. Heterogeneous multiscale method, homogenization problems.

The work of the first author was partially supported by ONR grant N00014-01-1-0674 and the National Natural Science Foundation of China through a Class B Award for Distinguished Young Scholars 10128102.

The work of the second author was partially supported by the Special Funds for the Major State Basic Research Projects G1999032804 and was also supported by the National Natural Science Foundation of China 10201033.

The work of the third author was partially supported by the Special Funds for the Major State Research Projects G1999032804 and the National Natural Science Foundation of China for Distinguished Young Scholars 10225103.

We thank Bjorn Engquist for inspiring discussions on the topic studied here.

## 1. INTRODUCTION AND MAIN RESULTS

1.1. **General methodology.** Consider the classical elliptic problem

$$(1.1) \quad \begin{cases} -\operatorname{div}(a^\varepsilon(x)\nabla u^\varepsilon(x)) = f(x) & x \in D \subset \mathbb{R}^d, \\ u^\varepsilon(x) = 0 & x \in \partial D. \end{cases}$$

Here  $\varepsilon$  is a small parameter that signifies explicitly the multiscale nature of the coefficient  $a^\varepsilon(x)$ . Several classical multiscale methodologies have been developed for the numerical solution of this elliptic problem, the most well known among which is the multigrid technique [8]. These classical multiscale methods are designed to resolve the details of the fine scale problem (1.1) and are applicable for general problems, i.e., no special assumptions are required for the coefficient  $a^\varepsilon(x)$ . In contrast modern multiscale methods are designed specifically for recovering partial information about  $u^\varepsilon$  at a sublinear cost, i.e., the total cost grows sublinearly with the cost of solving the fine scale problem [18]. This is only possible by exploring the special features that  $a^\varepsilon(x)$  might have, such as scale separation. The simplest example is when

$$(1.2) \quad a^\varepsilon(x) = a\left(x, \frac{x}{\varepsilon}\right),$$

where  $a(x, y)$  can either be periodic in  $y$ , in which case we assume the period to be  $I = [-1/2, 1/2]^d$ , or random but stationary under shifts in  $y$ , for each fixed  $x \in D$ . In both cases, it has been shown that [5, 36]

$$(1.3) \quad \|u^\varepsilon(x) - U(x)\|_{L^2(D)} \rightarrow 0,$$

where  $U(x)$  is the solution of a homogenized equation:

$$(1.4) \quad \begin{cases} -\operatorname{div}(\mathcal{A}(x)\nabla U(x)) = f(x) & x \in D, \\ U(x) = 0 & x \in \partial D. \end{cases}$$

The homogenized coefficient  $\mathcal{A}(x)$  can be obtained from the solutions of the so-called cell problem. In general, there are no explicit formulas for  $\mathcal{A}(x)$ , except in one dimension.

Several numerical methods have been developed to deal specifically with the case when  $a(x, y)$  is periodic in  $y$ . References [3, 4, 7] propose to solve the homogenized equations as well as the equations for the correctors. Schwab et al. [29, 38] use multiscale test functions of the form  $\varphi(x, x/\varepsilon)$  where  $\varphi(x, y)$  is periodic in  $y$  to extract the leading order behavior of  $u^\varepsilon(x)$ , extending an idea that was used analytically in the work of [2, 15, 34, 44] for the homogenization problems. These methods have the feature that their cost is independent of  $\varepsilon$ , hence sublinear as  $\varepsilon \rightarrow 0$ , but so far they are restricted to the periodic homogenization problem. An alternative proposal for more general problems but with much higher cost is found in [20, 25].

1.2. **Heterogeneous multiscale method.** HMM [16, 17, 18] is a general methodology for designing sublinear algorithms by exploiting scale separation and other special features of the problem. It consists of two components: selection of a macroscopic solver and estimating the missing macroscale data by solving locally the fine scale problem.

For (1.1) the macroscopic solver can be chosen as a conventional  $\mathcal{P}_k$  finite element method on a triangulation  $\mathcal{T}_H$  of element size  $H$  which should resolve the macroscale features of  $a^\varepsilon(x)$ . The missing data is the effective stiffness matrix at

this scale. This stiffness matrix can be estimated as follows. Assuming that the effective coefficient at this scale is  $\mathcal{A}_H(x)$ , if we knew  $\mathcal{A}_H(x)$  explicitly, we could have evaluated the quadratic form

$$\int_D \nabla V(x) \cdot \mathcal{A}_H(x) \nabla V(x) dx$$

by numerical quadrature: For any  $V \in X_H$ , the finite element space,

$$(1.5) \quad A_H(V, V) \simeq \sum_{K \in \mathcal{T}_H} |K| \sum_{x_\ell \in K} \omega_\ell (\nabla V \cdot \mathcal{A}_H \nabla V)(x_\ell),$$

where  $\{x_\ell\}$  and  $\{\omega_\ell\}$  are the quadrature points and weights in  $K$ ,  $|K|$  is the volume of  $K$ . In the absence of explicit knowledge of  $\mathcal{A}_H(x)$ , we approximate  $(\nabla V \cdot \mathcal{A}_H \nabla V)(x_\ell)$  by solving the problem:

$$(1.6) \quad \begin{cases} -\operatorname{div}(a^\varepsilon(x) \nabla v_\ell^\varepsilon(x)) = 0 & x \in I_\delta(x_\ell), \\ v_\ell^\varepsilon(x) = V_\ell(x) & x \in \partial I_\delta(x_\ell), \end{cases}$$

where  $I_\delta(x_\ell)$  is a cube of size  $\delta$  centered at  $x_\ell$ , and  $V_\ell$  is the linear approximation of  $V$  at  $x_\ell$ . We then let

$$(1.7) \quad (\nabla V \cdot \mathcal{A}_H \nabla V)(x_\ell) \simeq \frac{1}{\delta^d} \int_{I_\delta(x_\ell)} \nabla v_\ell^\varepsilon(x) \cdot a^\varepsilon(x) \nabla v_\ell^\varepsilon(x) dx.$$

(1.5) and (1.7) together give the needed approximate stiffness matrix at the scale  $H$ . For convenience, we will define the corresponding bilinear form: For any  $V, W \in X_H$

$$A_H(V, W) := \sum_{K \in \mathcal{T}_H} \frac{|K|}{\delta^d} \sum_{x_\ell \in K} \omega_\ell \int_{I_\delta(x_\ell)} \nabla v_\ell^\varepsilon(x) \cdot a^\varepsilon(x) \nabla w_\ell^\varepsilon(x) dx,$$

where  $w_\ell^\varepsilon$  is defined for  $W \in X_H$  in the same way that  $v_\ell^\varepsilon$  in (1.6) was defined for  $V$ .

In order to reduce the effect of the imposed boundary condition on  $\partial I_\delta(x_\ell)$ , we may replace (1.7) by

$$(1.8a) \quad (\nabla V \cdot \mathcal{A}_H \nabla V)(x_\ell) \simeq \frac{1}{(\delta')^d} \int_{I_{\delta'}(x_\ell)} \nabla v_\ell^\varepsilon(x) \cdot a^\varepsilon(x) \nabla v_\ell^\varepsilon(x) dx,$$

where  $\delta' < \delta$ . For example, we may choose  $\delta' = \delta/2$ . In (1.6), we used the Dirichlet boundary condition. Other boundary conditions are possible, such as Neumann and periodic boundary conditions. In the case when  $a^\varepsilon(x) = a(x, x/\varepsilon)$  and  $a(x, y)$  is periodic in  $y$ , one can take  $I_\delta(x_\ell)$  to be  $x_\ell + \varepsilon I$ , i.e.,  $\delta = \varepsilon$  and use the boundary condition that  $v_\ell^\varepsilon(x) - V_\ell(x)$  is periodic on  $I_\delta$ .

So far the algorithm is completely general. The savings compared with solving the full fine scale problem comes from the fact that we can choose  $I_\delta(x_\ell)$  to be smaller than  $K$ . The size of  $I_\delta(x_\ell)$  is determined by many factors, including the accuracy and cost requirement, the degree of scale separation, and the microstructure in  $a^\varepsilon(x)$ . One purpose for the error estimates that we present below is to give guidelines on how to select  $I_\delta(x_\ell)$ . As mentioned already, if  $a^\varepsilon(x) = a(x, x/\varepsilon)$  and  $a(x, y)$  is periodic in  $y$ , we can simply choose  $I_\delta(x_\ell)$  to be  $x_\ell + \varepsilon I$ , i.e.,  $\delta = \varepsilon$ . If  $a(x, y)$  is random, then  $\delta$  should be a few times larger than the local correlation

length of  $a^\varepsilon$ . In the former case, the total cost is independent of  $\varepsilon$ . In the latter case, the total cost depends only weakly on  $\varepsilon$  (see [31]).

The final problem is to solve

$$(1.8) \quad \min_{V \in X_H} \frac{1}{2} A_H(V, V) - (f, V).$$

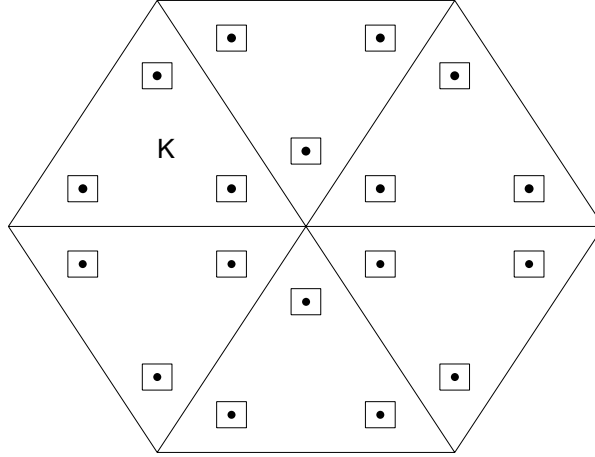


FIGURE 1. Illustration of HMM for solving (1.1). The dots are the quadrature points. The little squares are the microcell  $I_\delta(x_\ell)$ .

To summarize, HMM has the following features:

- (1) It gives a framework that allows us to maximally take advantage of the special features of the problem such as scale separation. For periodic homogenization problems, the cost of HMM is comparable to the special techniques

how to construct better approximations of  $u^\varepsilon$  from the HMM solutions. Our basic strategy is as follows. First we will prove a general statement that the error between the HMM solution and the solution of (1.4) is controlled by the standard error in the macroscale solver plus a new term, called  $e(\text{HMM})$ , due to the error in estimating the stiffness matrix. We then estimate  $e(\text{HMM})$ . This second part is only done for either periodic or random homogenization problems, since concrete results are only possible if the behavior of  $u^\varepsilon$  is well understood. We believe that this overall strategy will be useful for analyzing other multiscale methods.

We will always assume that  $a^\varepsilon(x)$  is smooth, symmetric and uniformly elliptic:

$$(1.9) \quad \lambda I \leq a^\varepsilon \leq \Lambda I$$

for some  $\lambda, \Lambda > 0$ . We will use the summation convention and standard notation for Sobolev spaces (see [1]). We will use  $|\cdot|$  to denote the absolute value of a scalar quantity, the Euclidean norm of a vector and the volume of a set  $K$ . For the quadrature formula (1.5), we will assume the following accuracy conditions for  $k$ th-order numerical quadrature scheme [11]:

$$(1.10) \quad \int_K p(x) dx = \frac{1}{|K|} \int_K p(x) dx = \sum_{\ell=1}^L \omega_\ell p(x_\ell) \quad \text{for all } p(x) \in \mathcal{P}_{2k-2}.$$

Here  $\omega_\ell > 0$ , for  $\ell = 1, \dots, L$ . For  $k = 1$ , we assume the above formula to be exact for  $p \in \mathcal{P}_1$ .

**1.3. Main results.** Our main results for the linear problem are as follows.

**Theorem 1.1.** *Denote by  $U_0$  and  $U_{\text{HMM}}$  the solution of (1.4) and the HMM solution, respectively. Let*

$$e(\text{HMM}) = \max_{\substack{x_\ell \in K \\ K \in \mathcal{T}_H}} \|\mathcal{A}(x_\ell) - \mathcal{A}_H(x_\ell)\|,$$

where  $\|\cdot\|$  is the Euclidean norm. If  $U_0$  is sufficiently smooth and (1.10) holds, then there exists a constant  $C$  independent of  $\varepsilon, \delta$  and  $H$ , such that

$$(1.11) \quad \|U_0 - U_{\text{HMM}}\|_1 \leq C(H^k + e(\text{HMM})),$$

$$(1.12) \quad \|U_0 - U_{\text{HMM}}\|_0 \leq C(H^{k+1} + e(\text{HMM})).$$

If there exists a constant  $C_0$  such that  $e(\text{HMM})|\ln H| < C_0$ , then there exists a constant  $H_0$  such that for all  $H \leq H_0$ ,

$$(1.13) \quad \|U_0 - U_{\text{HMM}}\|_{1,\infty} \leq C(H^k + e(\text{HMM}))|\ln H|.$$

At this stage, no assumption on the form of  $a^\varepsilon(x)$  is necessary.  $U_0$  can be the solution of an arbitrary macroscopic equation with the same right-hand side as in (1.1). Of course for  $U_{\text{HMM}}$  to converge to  $U_0$ , i.e.,  $e(\text{HMM}) \rightarrow 0$ ,  $U_0$  must be chosen as the solution of the homogenized equation, which we now assume exists. To obtain quantitative estimates on  $e(\text{HMM})$ , we must restrict ourselves to more specific cases.

**Theorem 1.2.** *For the periodic homogenization problem, we have*

$$e(\text{HMM}) \leq \begin{cases} C\varepsilon & \text{if } I_\delta(x_\ell) = x_\ell + \varepsilon I, \\ C\left(\frac{\varepsilon}{\delta} + \delta\right) & \text{otherwise.} \end{cases}$$

In the first case, we replace the boundary condition in (1.6) by a periodic boundary condition:  $v_\ell^\varepsilon - V_\ell$  is periodic with period  $\varepsilon I$ . For the second result we do not need to assume that the period of  $a(x, \cdot)$  is a cube: In fact it can be of arbitrary shape as long as its translation tiles up the whole space.

Another important case for which a specific estimate on  $e(\text{HMM})$  can be obtained is the random homogenization. In this case, using results in [43], we have

**Theorem 1.3.** *For the random homogenization problem, assuming that  $(\mathbf{A})$  in the*  
 $a$   $e(\mathbf{a})$

For this case, we can prove

**Theorem 1.5.** *Let  $\tilde{u}^\varepsilon$  be defined as above. Then*

$$(1.16) \quad \left( \sum_{K \in \mathcal{T}_H} \|\nabla(u^\varepsilon - \tilde{u}^\varepsilon)\|_{0,K}^2 \right)^{1/2} \leq C(\sqrt{\varepsilon} + H).$$

Similar results with some modification hold for nonlinear problems. The details are given in §5.

**1.5. Some technical background.** In this subsection, we will list some general results that will be frequently referred to later on.

Given a triangulation  $\mathcal{T}_H$ , it is called *regular* if there is a constant  $\sigma$  such that

$$\frac{H_K}{\rho_K} \leq \sigma \quad \text{for all } K \in \mathcal{T}_H$$

and if the quantity

$$H = \max_{K \in \mathcal{T}_H} H_K$$

approaches zero, where  $H_K$  is the diameter of  $K$  and  $\rho_K$  is the diameter of the largest ball inscribed in  $K$ .  $\mathcal{T}_H$  satisfies an *inverse assumption* if there exists a constant  $\nu$  such that

$$\frac{H}{H_K} \leq \nu \quad \text{for all } K \in \mathcal{T}_H.$$

A regular family of triangulation of  $\mathcal{T}_H$  satisfying the inverse assumption is called *quasi-uniform*.

The following interpolation result for the Lagrange finite element is adapted from [10]. Here and in what follows, for any  $k \geq 2$ ,  $\nabla^k v$  is understood in a piecewise manner.

**Theorem 1.6** ([10]). *Let  $\Pi$  be  $k$ th-order Lagrange interpolate operator, and assume that the following inclusions hold:*

$$(1.17) \quad W^{k+1,p}(\hat{K}) \hookrightarrow \mathcal{C}^0(\hat{K}) \quad \text{and} \quad W^{k+1,p}(\hat{K}) \hookrightarrow W^{m,q}(\hat{K}).$$

Then

$$(1.18) \quad |v - \Pi v|_{m,q,K} \leq C|K|^{1/q-1/p} \frac{H_K^{k+1}}{\rho_K^m} |v|_{k+1,p,K}.$$

If  $\mathcal{T}_H$  is regular, we have the global estimate

$$(1.19) \quad |v - \Pi v|_{m,q,D} \leq CH^{k+1-m+\min\{0,d(1/q-1/p)\}} |v|_{k+1,p,D}.$$

Inequality (1.18) is proven in [10, Theorem 3.1.6], and (1.19) is a direct consequence of (1.18) and the *inverse inequality* below.

Using (1.19) with  $p = q = 2$  and  $m = 2, k = 1$ , we have  $\|v - \Pi v\|_{2,D} \leq C\|v\|_{2,D}$ . Hence

$$(1.20) \quad \|\Pi v\|_{2,D} \leq \|v - \Pi v\|_{2,D} + \|v\|_{2,D} \leq C\|v\|_{2,D}.$$

We will also need the following form of the inverse inequality.

**Theorem 1.7** ([10, Theorem 3.2.6]). *Assume that  $\mathcal{T}_H$  is regular, and assume also that the two pairs  $(l, r)$  and  $(m, q)$  with  $l, m \geq 0$  and  $r, q \in [0, \infty]$  satisfy*

$$l \leq m \quad \text{and} \quad \mathcal{P}_k(\hat{K}) \subset W^{l,r}(\hat{K}) \cap W^{m,q}(\hat{K}).$$

Then there exists a constant  $C = C(\sigma, \nu, l, r, m, q)$  such that

$$(1.21) \quad |v|_{m,q,K} \leq CH_K^{l-m+d(1/q-1/r)} |v|_{l,r,K}$$

for any  $v \in \mathcal{P}_k(\hat{K})$ .

If in addition  $\mathcal{T}_H$  satisfies the inverse assumption, then there exists a constant  $C = C(\sigma, \nu, l, r, m, q)$  such that

$$(1.22) \quad \left( \sum_{K \in \mathcal{T}_H} |v|_{m,q,K}^q \right)^{1/q} \leq CH^{l-m+\min\{0, d(1/q-1/r)\}} \left( \sum_{K \in \mathcal{T}_H} |v|_{l,r,K}^r \right)^{1/r}$$

for any  $v \in X_H$  and  $r, q < \infty$ , with

$$\begin{aligned} \max_{K \in \mathcal{T}_H} |v|_{m,\infty,K} \quad \text{replacing} \quad & \left( \sum_{K \in \mathcal{T}_H} |v|_{m,q,K}^q \right)^{1/q}, \quad \text{if } q = \infty, \\ \max_{K \in \mathcal{T}_H} |v|_{l,\infty,K} \quad \text{replacing} \quad & \left( \sum_{K \in \mathcal{T}_H} |v|_{l,r,K}^r \right)^{1/r}, \quad \text{if } r = \infty. \end{aligned}$$

The following simple result will be used repeatedly.

**Lemma 1.8.** *Let  $A_1(x)$  and  $A_2(x)$  be symmetric matrices satisfying (1.9). Let  $\varphi_1$  be the solution of*

$$(1.23) \quad -\operatorname{div}(A_1(x)\nabla\varphi_1(x)) = \operatorname{div}(\tilde{A}_1(x)\nabla F_1(x)) \quad x \in \Omega,$$

with either the Dirichlet or periodic boundary condition on  $\partial\Omega$ . Let  $\varphi_2$  be a solution of (1.23) with  $A_1, \tilde{A}_1$  and  $F_1$  replaced by  $A_2, \tilde{A}_2$  and  $F_2$ , respectively, and let  $\varphi_2$  satisfy the same boundary condition as  $\varphi_1$ . Then

$$(1.24) \quad \begin{aligned} \lambda \|\nabla(\varphi_1 - \varphi_2)\|_{0,\Omega} &\leq \max_{x \in \Omega} |(\tilde{A}_1 - \tilde{A}_2)(x)| \|\nabla F_1\|_{0,\Omega} + \max_{x \in \Omega} |(A_1 - A_2)(x)| \|\nabla \varphi_2\|_{0,\Omega} \\ &+ \max_{x \in \Omega} |\tilde{A}_2(x)| \|\nabla(F_1 - F_2)\|_{0,\Omega}. \end{aligned}$$

*Proof.* Inequality (1.24) is a direct consequence of

$$\lambda \|\nabla(\varphi_1 - \varphi_2)\|_{0,\Omega}^2 \leq \int_{\Omega} \nabla(\varphi_1 - \varphi_2) \cdot ((\tilde{A}_2 - \tilde{A}_1)\nabla F_1 + (A_2 - A_1)\nabla \varphi_2 + \tilde{A}_2\nabla(F_2 - F_1)).$$

□

The following simple result underlies the stability of HMM for problem (1.1).

**Lemma 1.9.** *Let  $\varphi$  be the solution of*

$$(1.25) \quad \begin{cases} -\operatorname{div}(a\nabla\varphi) = 0 & \text{in } \Omega \subset \mathbb{R}^d, \\ \varphi = V_\ell & \text{on } \partial\Omega, \end{cases}$$

where  $V_\ell$  is a linear function and  $a = (a_{ij})$  satisfies

$$\lambda I \leq a \leq \Lambda I.$$

Then we have

$$(1.26) \quad \|\nabla V_\ell\|_{0,\Omega} \leq \|\nabla\varphi\|_{0,\Omega} \quad \text{and} \quad \left( \int_{\Omega} \nabla\varphi \cdot a\nabla\varphi \right)^{1/2} \leq \left( \int_{\Omega} \nabla V_\ell \cdot a\nabla V_\ell \right)^{1/2}.$$



*Proof.* Notice that  $\varphi = V_\ell$  on the edges of  $\Omega$ , using the fact that  $\nabla V_\ell$  is a constant in  $\Omega$ , and integration by parts leads to

$$\int_{\Omega} \nabla(\varphi - V_\ell)(x) \nabla V_\ell(x) \, dx = 0,$$

which implies

$$\int_{\Omega} |\nabla \varphi(x)|^2 \, dx = \int_{\Omega} |\nabla V_\ell(x)|^2 \, dx + \int_{\Omega} |\nabla(\varphi - V_\ell)(x)|^2 \, dx.$$

This gives the first result in (1.26). Multiplying (1.25) by  $\varphi(x) - V_\ell(x)$  and integrating by parts, we obtain

$$\begin{aligned} \int_{\Omega} \nabla \varphi(x) \cdot a \nabla \varphi(x) \, dx + \int_{\Omega} \nabla(\varphi - V_\ell)(x) \cdot a \nabla(\varphi - V_\ell)(x) \, dx \\ = \int_{\Omega} \nabla V_\ell(x) \cdot a \nabla V_\ell(x) \, dx. \end{aligned}$$

This gives the second part of (1.26).  $\square$

*Remark 1.10.* For this result, the coefficient  $a = (a_{ij})$  may depend on the solution, i.e., (1.25) may be nonlinear.

*Remark 1.11.* The same result holds if we use instead a periodic boundary condition:  $\varphi - V_\ell$  is periodic with period  $\Omega$ .

## 2. GENERALITIES

Here we prove Theorem 1.1. We will let  $U_H = U_{\text{HMM}}$  for convenience.

Since  $U_H$  is the numerical solution associated with the quadratic form  $A_H$ ,  $U_0$  is the exact solution associated with the quadratic form  $A$ , defined for any  $V \in H_0^1(D)$  as

$$A(V, V) = \int_D \nabla V(x) \cdot \mathcal{A}(x) \nabla V(x) \, dx.$$

To estimate  $U_0 - U_H$ , we view  $A_H$  as an approximation to  $A$ , and we use *Strang's first lemma* [10].

Using (1.26) with  $\Omega = I_\delta(x_\ell)$  and (1.9), for any  $V \in X_H$ , we have

$$\begin{aligned} A_H(V, V) &\geq \lambda \sum_{K \in \mathcal{T}_H} |K| \sum_{x_\ell \in K} \omega_\ell \int_{I_\delta(x_\ell)} |\nabla V_\ell(x)|^2 \, dx \\ &= \lambda \sum_{K \in \mathcal{T}_H} |K| \sum_{x_\ell \in K} \omega_\ell |\nabla V(x_\ell)|^2 \\ (2.1) \qquad &= \lambda \|\nabla V\|_0^2. \end{aligned}$$

Similarly, for any  $V, W \in X_H$ , we obtain

$$\begin{aligned}
|A_H(V, W)| &\leq \sum_{K \in \mathcal{T}_H} |K| \sum_{x_\ell \in K} \omega_\ell \left( \int_{I_\delta(x_\ell)} \nabla V_\ell \cdot a^\varepsilon \nabla V_\ell \right)^{\frac{1}{2}} \left( \int_{I_\delta(x_\ell)} \nabla W_\ell \cdot a^\varepsilon \nabla W_\ell \right)^{\frac{1}{2}} \\
&\leq \Lambda \sum_{K \in \mathcal{T}_H} |K| \sum_{x_\ell \in K} \omega_\ell |\nabla V(x_\ell)| |\nabla W(x_\ell)| \\
&= \Lambda \sum_{K \in \mathcal{T}_H} \int_K |\nabla V(x)| |\nabla W(x)| dx \\
(2.2) \quad &\leq \Lambda \|\nabla V\|_0 \|\nabla W\|_0.
\end{aligned}$$

The existence and the uniqueness of the solutions to (1.8) follow from (2.1) and (2.2) via the Lax-Milgram lemma and the *Poincaré inequality*.

To streamline the proof of Theorem 1.1, we introduce the following auxiliary bilinear form  $\hat{A}_H$ .

$$\hat{A}_H(V, W) = \sum_{K \in \mathcal{T}_H} \hat{A}_K(V, W) \quad \text{with} \quad \hat{A}_K(V, W) = |K| \sum_{x_\ell \in K} \omega_\ell (\nabla W \cdot \mathcal{A} \nabla V)(x_\ell).$$

Classical results on numerical integration [11, Theorem 6] give for any  $V, W \in X_H$ ,

$$(2.3) \quad \left| \hat{A}_K(V, W) - \int_K \nabla W \cdot \mathcal{A} \nabla V dx \right| \leq CH^m \|V\|_{m,K} \|\nabla W\|_{0,K} \quad 1 \leq m \leq k.$$

Moreover, for any  $V, W \in X_H$ , if  $\|V\|_{k+1}$  and  $\|W\|_2$  are bounded, we have [11, Theorem 8],

$$(2.4) \quad |\hat{A}_H(V, W) - A(V, W)| \leq CH^{k+1} \|V\|_{k+1} \|W\|_2.$$

*Proof of Theorem 1.1.* Using the *first Strang lemma* [10, Theorem 4.1.1], we have

$$\|U_0 - U_H\|_1 \leq C \inf_{V \in X_H} \left( \|U_0 - V\|_1 + \sup_{W \in X_H} \frac{|A_H(V, W) - A(V, W)|}{\|W\|_1} \right).$$

Let  $V = \Pi U_0$  and using (1.19) with  $m = 1, p = q = 2$ , we have

$$(2.5) \quad \inf_{V \in X_H} \|U_0 - V\|_1 \leq \|U_0 - \Pi U_0\|_1 \leq CH^k.$$

It remains to estimate  $|A_H(V, W) - A(V, W)|$  for  $V = \Pi U_0$  and  $W \in X_H$ . Using (2.3), we get

$$\begin{aligned}
|A_H(V, W) - A(V, W)| &\leq |A_H(V, W) - \hat{A}_H(V, W)| + |\hat{A}_H(V, W) - A(V, W)| \\
(2.6) \quad &\leq (e(\text{HMM})) \|\nabla V\|_0 + CH^k \|V\|_k \|\nabla W\|_0.
\end{aligned}$$

This gives (1.11)

To get the  $L^2$  estimate, we use the Aubin-Nitsche dual argument [10]. To this end, consider the following auxiliary problem: Find  $w \in H_0^1(D)$  such that

$$(2.7) \quad A(v, w) = (U_0 - U_H, v) \quad \text{for all } v \in H_0^1(D).$$

The standard regularity result reads [24]

$$(2.8) \quad \|w\|_2 \leq C \|U_0 - U_H\|_0.$$

Putting  $v = U_0 - U_H$  into the right-hand side of (2.7), we obtain

$$\begin{aligned}
\|U_0 - U_H\|_0^2 &= A(U_0 - U_H, w - \Pi w) + (A_H(U_H, \Pi w) - A(U_H, \Pi w)) \\
&= A(U_0 - U_H, w - \Pi w) \\
&\quad + (A_H(U_H - \Pi U_0, \Pi w) - A(U_H - \Pi U_0, \Pi w)) \\
(2.9) \quad &\quad + (A_H(\Pi U_0, \Pi w) - A(\Pi U_0, \Pi w)).
\end{aligned}$$

Using (2.6) with  $k = 1$ , we bound the first two terms in the right-hand side of the above identity as

$$|A(U_0 - U_H, w - \Pi w)| \leq C \|U_0 - U_H\|_1 \|w - \Pi w\|_1 \leq CH \|U_0 - U_H\|_1 \|w\|_2$$

and

$$|A_H(U_H - \Pi U_0, \Pi w) - A(U_H - \Pi U_0, \Pi w)| \leq (e(\text{HMM}) + CH) \|U_0 - U_H\|_1 \|\Pi w\|_1.$$

The last term in the right-hand side of (2.9) may be decomposed into

$$\begin{aligned}
A_H(\Pi U_0, \Pi w) - A(\Pi U_0, \Pi w) &= (A_H(\Pi U_0, \Pi w) - \hat{A}_H(\Pi U_0, \Pi w)) \\
&\quad + (\hat{A}_H(\Pi U_0, \Pi w) - A(\Pi U_0, \Pi w)).
\end{aligned}$$

It follows from (2.4) that

$$|\hat{A}_H(\Pi U_0, \Pi w) - A(\Pi U_0, \Pi w)| \leq CH^{k+1} \|U_0\|_{k+1} \|w\|_2.$$

By definition of  $e(\text{HMM})$  and using (1.20), we get

$$|A_H(\Pi U_0, \Pi w) - \hat{A}_H(\Pi U_0, \Pi w)| \leq Ce(\text{HMM}) \|\nabla \Pi U_0\|_0 \|w\|_2.$$

Combining the above estimates and using (2.8) lead to (1.12).

It remains to prove (1.13). As in [37], for any point  $z \in D$ , we define the regularized Green's function  $G^z \in H_0^1(D)$  and the discrete Green's function  $G_H^z \in X_H$  as

$$\begin{aligned}
(2.10) \quad A(G^z, V) &= (\delta_z, \partial V) \quad \text{for all } V \in H_0^1(D), \\
A(G_H^z, V) &= (\delta_z, \partial V) \quad \text{for all } V \in X_H,
\end{aligned}$$

where  $\delta_z$  is the regularized Dirac- $\delta$  function defined in [37]. It is well known that

$$(2.11) \quad \|G^z - G_H^z\|_{1,1} \leq C \quad \text{and} \quad \|G_H^z\|_{1,1} \leq C |\ln H|.$$

A proof for (2.11) can be obtained by using the weighted-norm technique [37]. We refer to [9, Chapter 7] for details. Using the definition of  $G^z$  and  $G_H^z$ , a simple manipulation gives

$$\begin{aligned}
\partial(U_0 - U_H)(z) &= A(G^z, U_0 - \Pi U_0) + A(G^z, \Pi U_0 - U_H) \\
&= A(G^z - G_H^z, U_0 - \Pi U_0) + A(G_H^z, U_0 - U_H) \\
&= A(G^z - G_H^z, U_0 - \Pi U_0) + A_H(U_H, G_H^z) - A(U_H, G_H^z) \\
&= A(G^z - G_H^z, U_0 - \Pi U_0) + (A_H(\Pi U_0, G_H^z) - A(\Pi U_0, G_H^z)) \\
&\quad + (A_H(U_H - \Pi U_0, G_H^z) - A(U_H - \Pi U_0, G_H^z)).
\end{aligned}$$

Using (2.11), we obtain

$$\begin{aligned}
\|U_0 - U_H\|_{1,\infty} &\leq C \|U_0 - \Pi U_0\|_{1,\infty} + |A(\Pi U_0, G_H^z) - A_H(\Pi U_0, G_H^z)| \\
&\quad + |A(U_H - \Pi U_0, G_H^z) - A_H(U_H - \Pi U_0, G_H^z)|.
\end{aligned}$$

Using (2.6), we get

$$\begin{aligned}
|A(\Pi U_0, G_H^z) - A_H(\Pi U_0, G_H^z)| &\leq (e(\text{HMM}) + CH^k) \sum_{K \in \mathcal{T}_H} \|\Pi U_0\|_{k,K} \|\nabla G_H^z\|_{0,K} \\
&\leq C(e(\text{HMM}) + H^k) \sum_{K \in \mathcal{T}_H} \|\Pi U_0\|_{k,\infty,K} \|\nabla G_H^z\|_{L^1(K)} \\
&\leq C(e(\text{HMM}) + H^k) |\ln H| \|U_0\|_{k+1,\infty},
\end{aligned}$$

where we have used the inverse inequality (1.21).

Similarly, we have

$$\begin{aligned}
&|A(U_H - \Pi U_0, G_H^z) - A_H(U_H - \Pi U_0, G_H^z)| \\
&\leq (e(\text{HMM}) + CH) \sum_{K \in \mathcal{T}_H} \|U_H - \Pi U_0\|_{1,K} \|\nabla G_H^z\|_{0,K} \\
&\leq C(e(\text{HMM}) + H) \sum_{K \in \mathcal{T}_H} \|U_H - \Pi U_0\|_{1,\infty,K} \|\nabla G_H^z\|_{0,1,K} \\
&\leq C(e(\text{HMM}) + H) |\ln H| \|U_0 - U_H\|_{1,\infty} \\
&\quad + C(e(\text{HMM}) + H) |\ln H| H^k \|U_0\|_{k+1,\infty}.
\end{aligned}$$

A combination of the above three estimates yields

$$\begin{aligned}
\|U_0 - U_H\|_{1,\infty} &\leq CH^k + C(e(\text{HMM}) + H) |\ln H| \|U_0 - U_H\|_{1,\infty} \\
&\quad + C(e(\text{HMM}) + H^k) |\ln H| \|U_0\|_{k+1,\infty}.
\end{aligned}$$

If  $e(\text{HMM})|\ln H| < C_0 := 1/(2C)$ , then there exists a constant  $H_0$  such that for all  $H \leq H_0$ ,

$$C(e(\text{HMM}) + H) |\ln H| \leq 1/2 + CH |\ln H| < 1.$$

We thus obtain (1.13) and this completes the proof.  $\square$

Combining the foregoing proof for the  $L^2$  and  $W^{1,\infty}$  estimates, using the Green's function defined in [39], we obtain

*Remark 2.1.* Under the same condition for the  $W^{1,\infty}$  estimate in Theorem 1.1, we have

$$\|U_0 - U_H\|_{L^\infty} \leq C(e(\text{HMM}) + H^{k+1}) |\ln H|^2.$$

### 3. ESTIMATING $e(\text{HMM})$

In this section, we estimate  $e(\text{HMM})$  for problems with locally periodic coefficients. The estimate of  $e(\text{HMM})$  for problems with random coefficients can be found in the Appendix.

We assume that  $a^\varepsilon(x) = a(x, x/\varepsilon)$ , where  $a^\varepsilon$  is smooth in  $x$  and periodic in  $y$  with period  $I$ . Define  $\kappa = \lfloor \delta/\varepsilon \rfloor$ , and we introduce  $\hat{V}_\ell$  as

$$(3.1) \quad \hat{V}_\ell(x) = V_\ell(x) + \varepsilon \chi^k \left( x_\ell, \frac{x}{\varepsilon} \right) \frac{\partial V_\ell}{\partial x_k}(x),$$

where  $\{\chi^j\}_{j=1}^d$  is defined as: For  $j = 1, \dots, d$ ,  $\chi^j(x, y)$  is periodic in  $y$  with period  $I$  and satisfies

$$(3.2) \quad -\frac{\partial}{\partial y_i} \left( a_{ik} \frac{\partial \chi^j}{\partial y_k} \right) (x, y) = \frac{\partial}{\partial y_i} a_{ij}(x, y) \quad \text{in } I, \quad \int_I \chi^j(x, y) dy = 0.$$

Given  $\{\chi^j\}_{j=1}^d$ , the homogenized coefficient  $\mathcal{A} = (\mathcal{A}_{ij}(x))$  is given by

$$\mathcal{A}_{ij}(x) = \int_I \left( a_{ij} + a_{ik} \frac{\partial \chi^j}{\partial y_k} \right) (x, y) dy.$$

Note that  $\{\chi^j\}_{j=1}^d$  is smooth and bounded in all norms.

First let us consider the case when  $I_\delta(x_\ell) = x_\ell + \varepsilon I$ , and (1.6) is solved with the periodic boundary condition. Denote by  $\hat{v}_\ell^\varepsilon$  the solution of (1.6) with the coefficients  $a^\varepsilon(x)$  replaced by  $a(x_\ell, x/\varepsilon)$ .  $\hat{v}_\ell^\varepsilon$  may be viewed as a perturbation of  $v_\ell^\varepsilon$ . Using Lemma 1.8, we get

$$(3.3) \quad \|\nabla(v_\ell^\varepsilon - \hat{v}_\ell^\varepsilon)\|_{0, I_\varepsilon} \leq C\varepsilon \|\nabla V_\ell\|_{0, I_\varepsilon}.$$

Observe that  $\hat{v}_\ell^\varepsilon = \hat{V}_\ell$ . A direct calculation yields

$$\begin{aligned} (\nabla W \cdot (\mathcal{A}_H - \mathcal{A}) \nabla V)(x_\ell) &= \int_{I_\varepsilon} \nabla w_\ell^\varepsilon \cdot \left[ a\left(x, \frac{x}{\varepsilon}\right) - a\left(x_\ell, \frac{x}{\varepsilon}\right) \right] \nabla v_\ell^\varepsilon dx \\ &\quad + \int_{I_\varepsilon} \nabla w_\ell^\varepsilon \cdot a\left(x_\ell, \frac{x}{\varepsilon}\right) \nabla(v_\ell^\varepsilon - \hat{v}_\ell^\varepsilon) dx. \end{aligned}$$

Using (3.3), we get

$$(3.4) \quad e(\text{HMM}) \leq C\varepsilon.$$

Next we consider the more general case when  $I_\delta$  is a cube of size  $\delta$  not necessarily equal to  $\varepsilon$ . The following analysis applies equally well to the case when the period of  $a(x, \cdot)$  is of general and even nonpolygonal shape. This situation arises in some examples of composite materials [30]. We will show that if  $\delta$  is much larger than  $\varepsilon$ , then the averaged energy density for the solution of (1.6) closely approximates the energy density of the homogenized problem. We begin with the following observation:

$$(3.5) \quad \begin{aligned} (\nabla W \cdot \mathcal{A} \nabla V)(x_\ell) &= \nabla W_\ell(x) \cdot \mathcal{A}(x_\ell) \nabla V_\ell(x) \\ &= \int_{I_{\kappa\varepsilon}(x_\ell)} \nabla W_\ell \cdot a\left(x_\ell, \frac{x}{\varepsilon}\right) \nabla \hat{V}_\ell dx. \end{aligned}$$

We first establish some estimates on the solution of the cell problem (1.6). We will write  $I_\delta$  instead of  $I_\delta(x_\ell)$  if there is no risk of confusion.

**Lemma 3.1.** *There exists a constant  $C$  independent of  $\varepsilon$  and  $\delta$  such that for each  $\ell$ ,*

$$(3.6) \quad \|\nabla v_\ell^\varepsilon\|_{0, I_\delta \setminus I_{\kappa\varepsilon}} \leq C \left( \left( \frac{\varepsilon}{\delta} \right)^{1/2} + \delta \right) \|\nabla V_\ell\|_{0, I_\delta}.$$

*Proof.* We still denote by  $\hat{v}_\ell^\varepsilon$  the solution of (1.6) with the coefficient  $a^\varepsilon(x)$  replaced by  $a(x_\ell, x/\varepsilon)$ . Using Lemma 1.8, we get

$$(3.7) \quad \|\nabla(v_\ell^\varepsilon - \hat{v}_\ell^\varepsilon)\|_{0, I_\delta} \leq C\delta \|\nabla V_\ell\|_{0, I_\delta}.$$

Define  $\theta_\ell^\varepsilon = \hat{v}_\ell^\varepsilon - \hat{V}_\ell$ , which obviously satisfies

$$(3.8) \quad \begin{cases} -\operatorname{div}\left(a\left(x_\ell, \frac{x}{\varepsilon}\right) \nabla \theta_\ell^\varepsilon(x)\right) = 0 & x \in I_\delta(x_\ell), \\ \theta_\ell^\varepsilon(x) = -\varepsilon \chi^k\left(x_\ell, \frac{x}{\varepsilon}\right) \frac{\partial V_\ell}{\partial x_k}(x) & x \in \partial I_\delta(x_\ell). \end{cases}$$

Note that  $\theta_\ell^\varepsilon$  is simply the *boundary layer correction* for the cell problem (1.6) [5]. It is proved in [45, (1.51) in §1.4], using the rescaling  $x' = x/\delta$  over  $I_\delta$  and  $\varepsilon' = \varepsilon/\delta$ .

$$(3.9) \quad \|\nabla\theta_\ell^\varepsilon\|_{0,I_\delta} \leq C\left(\frac{\varepsilon}{\delta}\right)^{1/2} \|\nabla V_\ell\|_{0,I_\delta}.$$

This together with (3.7) gives

$$(3.10) \quad \|\nabla(v_\ell^\varepsilon - \hat{V}_\ell)\|_{0,I_\delta} \leq C\left(\left(\frac{\varepsilon}{\delta}\right)^{1/2} + \delta\right) \|\nabla V_\ell\|_{0,I_\delta}.$$

A straightforward calculation gives

$$(3.11) \quad \|\nabla\hat{V}_\ell\|_{0,I_\delta \setminus I_{\kappa\varepsilon}} \leq C\left(\frac{\varepsilon}{\delta}\right)^{1/2} \|\nabla V_\ell\|_{0,I_\delta},$$

which together with (3.10) leads to

$$\begin{aligned} \|\nabla v_\ell^\varepsilon\|_{0,I_\delta \setminus I_{\kappa\varepsilon}} &\leq \|\nabla\hat{V}_\ell\|_{0,I_\delta \setminus I_{\kappa\varepsilon}} + \|\nabla(v_\ell^\varepsilon - \hat{V}_\ell)\|_{0,I_\delta \setminus I_{\kappa\varepsilon}} \\ &\leq \|\nabla\hat{V}_\ell\|_{0,I_\delta \setminus I_{\kappa\varepsilon}} + \|\nabla(v_\ell^\varepsilon - \hat{V}_\ell)\|_{0,I_\delta} \\ &\leq C\left(\left(\frac{\varepsilon}{\delta}\right)^{1/2} + \delta\right) \|\nabla V_\ell\|_{0,I_\delta}. \end{aligned}$$

This gives (3.6). □

As in (3.6), we also have

$$(3.12) \quad \|\nabla v_\ell^\varepsilon\|_{0,I_\delta \setminus I_{(\kappa-2)\varepsilon}} \leq C\left(\left(\frac{\varepsilon}{\delta}\right)^{1/2} + \delta\right) \|\nabla V_\ell\|_{0,I_\delta}.$$

**Theorem 3.2.**

$$(3.13) \quad e(\text{HMM}) \leq C\left(\frac{\varepsilon}{\delta} + \delta\right).$$

*Proof.* Note that  $v_\ell^\varepsilon = (v_\ell^\varepsilon - \hat{v}_\ell^\varepsilon) + \theta_\ell^\varepsilon + \hat{V}_\ell$ . We have

$$(\nabla W \cdot (\mathcal{A}_H - \mathcal{A})\nabla V)(x_\ell) = :I_1 + I_2 + I_3,$$

where

$$\begin{aligned} I_1 &= \int_{I_\delta} \nabla w_\ell^\varepsilon \cdot a\left(x, \frac{x}{\varepsilon}\right) \nabla(v_\ell^\varepsilon - \hat{v}_\ell^\varepsilon) dx, \quad I_2 = \int_{I_\delta} \nabla w_\ell^\varepsilon \cdot a\left(x, \frac{x}{\varepsilon}\right) \nabla\theta_\ell^\varepsilon dx, \\ I_3 &= \int_{I_\delta} \nabla w_\ell^\varepsilon \cdot a\left(x, \frac{x}{\varepsilon}\right) \nabla\hat{V}_\ell dx - \nabla W_\ell \cdot \mathcal{A}(x_\ell) \nabla V_\ell. \end{aligned}$$

Using (3.7) and (2.2), we bound  $I_1$  as

$$\begin{aligned} |I_1| &\leq \Lambda\delta^{-d} \|\nabla(v_\ell^\varepsilon - \hat{v}_\ell^\varepsilon)\|_{0,I_\delta} \|\nabla w_\ell^\varepsilon\|_{0,I_\delta} \\ &\leq C\delta^{1-d} \|\nabla V_\ell\|_{0,I_\delta} \|\nabla W_\ell\|_{0,I_\delta} = C\delta \|\nabla V_\ell\| \|\nabla W_\ell\|. \end{aligned}$$

Using the symmetry of  $a^\varepsilon$ ,  $I_2 = \int_{I_\delta} \nabla(\theta_\ell^\varepsilon + \varepsilon\chi^k(x_\ell, \frac{x}{\varepsilon}) \frac{\partial V_\ell}{\partial x_k}(1 - \rho^\varepsilon)) \cdot a\left(x, \frac{x}{\varepsilon}\right) \nabla w_\ell^\varepsilon dx$  and

$$\begin{aligned} I_2 &= \int_{I_\delta} \nabla\left(\theta_\ell^\varepsilon + \varepsilon\chi^k\left(x_\ell, \frac{x}{\varepsilon}\right) \frac{\partial V_\ell}{\partial x_k}(1 - \rho^\varepsilon)\right) \cdot a\left(x, \frac{x}{\varepsilon}\right) \nabla w_\ell^\varepsilon dx \\ &\quad - \int_{I_\delta} \nabla\left(\varepsilon\chi^k\left(x_\ell, \frac{x}{\varepsilon}\right) \frac{\partial V_\ell}{\partial x_k}(1 - \rho^\varepsilon)\right) \cdot a\left(x, \frac{x}{\varepsilon}\right) \nabla w_\ell^\varepsilon dx, \end{aligned}$$

where  $\rho^\varepsilon(x) \in C_0^\infty(I_\delta)$ ,  $|\nabla \rho^\varepsilon| \leq C/\varepsilon$ , and

$$(3.14) \quad \rho^\varepsilon(x) = \begin{cases} 1 & \text{if } \text{dist}(x, \partial I_\delta) \geq 2\varepsilon, \\ 0 & \text{if } \text{dist}(x, \partial I_\delta) \leq \varepsilon. \end{cases}$$

Using (1.6) and  $\theta_\ell^\varepsilon + \varepsilon \chi^k(x_\ell, \frac{x}{\varepsilon}) \frac{\partial V_\ell}{\partial x_k} (1 - \rho^\varepsilon) \in H_0^1(I_\delta)$ , integrating by parts makes the first term in the right-hand side of  $I_2$  vanish; therefore we write  $I_2$  as

$$I_2 = - \int_{I_\delta} a_{ij} \left( x, \frac{x}{\varepsilon} \right) \frac{\partial w_\ell^\varepsilon}{\partial x_i} \frac{\partial \chi^k}{\partial y_j} \frac{\partial V_\ell}{\partial x_k} (1 - \rho^\varepsilon) dx + \varepsilon \int_{I_\delta} a_{ij} \left( x, \frac{x}{\varepsilon} \right) \frac{\partial w_\ell^\varepsilon}{\partial x_i} \chi^k \frac{\partial V_\ell}{\partial x_k} \frac{\partial \rho^\varepsilon}{\partial x_j} dx.$$

Using (3.12), we bound  $I_2$  as

$$|I_2| \leq C \delta^{-d} \|\nabla w_\ell^\varepsilon\|_{0, I_\delta \setminus I_{(\kappa-2)\varepsilon}} \|\nabla V_\ell\|_{0, I_\delta \setminus I_{(\kappa-2)\varepsilon}} \leq C \left( \frac{\varepsilon}{\delta} + \delta^2 \right) |\nabla W_\ell| |\nabla V_\ell|.$$

Using (3.2) and integrating by parts, we obtain

$$\int_{I_\delta} \nabla w_\ell^\varepsilon \cdot a \left( x_\ell, \frac{x}{\varepsilon} \right) \nabla \hat{V}_\ell dx = \int_{I_\delta} \nabla W_\ell \cdot a \left( x_\ell, \frac{x}{\varepsilon} \right) \nabla \hat{V}_\ell dx,$$

which together with (3.5) gives

$$\begin{aligned} I_3 &= \int_{I_\delta} \nabla w_\ell^\varepsilon \left[ a \left( x, \frac{x}{\varepsilon} \right) - a \left( x_\ell, \frac{x}{\varepsilon} \right) \right] \nabla \hat{V}_\ell dx \\ &\quad + \frac{1}{\delta^d} \int_{I_\delta \setminus I_{\kappa\varepsilon}} \nabla W_\ell \cdot a \left( x_\ell, \frac{x}{\varepsilon} \right) \nabla \hat{V}_\ell dx + (|\kappa\varepsilon/\delta|^d - 1) \nabla W_\ell \cdot \mathcal{A}(x_\ell) \nabla V_\ell. \end{aligned}$$

The last term of  $I_3$  is bounded by

$$| |\kappa\varepsilon/\delta|^d - 1 | |\nabla W_\ell \cdot \mathcal{A}(x_\ell) \nabla V_\ell| \leq C \frac{\varepsilon}{\delta} |\nabla V_\ell| |\nabla W_\ell|,$$

where we have used  $| |\kappa\varepsilon/\delta|^d - 1 | \leq C\varepsilon/\delta$ . Using (3.11), we get

$$\begin{aligned} \delta^{-d} \left| \int_{I_\delta \setminus I_{\kappa\varepsilon}} \nabla W_\ell \cdot a \left( x_\ell, \frac{x}{\varepsilon} \right) \nabla \hat{V}_\ell dx \right| &\leq C \frac{1}{\delta^d} \|\nabla \hat{V}_\ell\|_{0, I_\delta \setminus I_{\kappa\varepsilon}} \|\nabla W_\ell\|_{0, I_\delta \setminus I_{\kappa\varepsilon}} \\ &\leq C \left( \frac{\varepsilon}{\delta} + \delta^2 \right) |\nabla V_\ell| |\nabla W_\ell|. \end{aligned}$$

Consequently, we obtain

$$\begin{aligned} |I_3| &\leq C \delta^{1-d} \|\nabla w_\ell^\varepsilon\|_{0, I_\delta} \|\nabla \hat{V}_\ell\|_{0, I_\delta} + C \left( \frac{\varepsilon}{\delta} + \delta^2 \right) |\nabla V_\ell| |\nabla W_\ell| \\ &\leq C \left( \frac{\varepsilon}{\delta} + \delta \right) |\nabla V_\ell| |\nabla W_\ell|. \end{aligned}$$

Combining the estimates for  $I_1, I_2$  and  $I_3$  gives the desired result (3.13).  $\square$

*Remark 3.3.* An explicit expression for  $v_\ell^\varepsilon$  is available in one dimension, from which we may show that the bound for  $e(\text{HMM})$  is sharp.

## 4. RECONSTRUCTION AND COMPRESSION

**4.1. Reconstruction procedure.** Next we consider how to construct better approximations to  $u^\varepsilon$  from  $U_H$ . We will restrict ourselves to the case when  $k = 1$ .

*Proof of Theorem 1.4.* Subtracting (1.1) from (1.14), we obtain

$$\begin{cases} -\operatorname{div}(a^\varepsilon(x)\nabla(\tilde{u}^\varepsilon - u^\varepsilon)(x)) = 0 & x \in \Omega_\eta, \\ (\tilde{u}^\varepsilon - u^\varepsilon)(x) = U_H(x) - u^\varepsilon(x) & x \in \partial\Omega_\eta. \end{cases}$$

Using classical interior estimates for elliptic equation [24], we have

$$\|\nabla(\tilde{u}^\varepsilon - u^\varepsilon)\|_{0,\Omega} \leq \frac{C}{\eta} \|\tilde{u}^\varepsilon - u^\varepsilon\|_{0,\Omega_\eta}.$$

Using the Hopf maximum principle, we get

$$\begin{aligned} \frac{1}{\eta^2} \int_{\Omega_\eta} |(\tilde{u}^\varepsilon - u^\varepsilon)(x)|^2 dx &\leq \frac{C}{\eta^2} \|\tilde{u}^\varepsilon - u^\varepsilon\|_{L^\infty(\Omega_\eta)}^2 \leq \frac{C}{\eta^2} \|u^\varepsilon - U_H\|_{L^\infty(\partial\Omega_\eta)}^2 \\ &\leq \frac{C}{\eta^2} \left( \|U_0 - U_H\|_{L^\infty(\Omega_\eta)}^2 + \|u^\varepsilon - U_0\|_{L^\infty(\Omega_\eta)}^2 \right). \end{aligned}$$

A combination of the above two results implies Theorem 1.4.  $\square$

*Proof of Theorem 1.5.* Denote  $I_\varepsilon(x_K) = x_K + \varepsilon I$  and define  $\hat{u}^\varepsilon$  as the solution of

$$(4.1) \quad -\operatorname{div}\left(a\left(x_K, \frac{x}{\varepsilon}\right)\nabla\hat{u}^\varepsilon(x)\right) = 0 \quad \text{in } I_\varepsilon(x_K),$$

with the boundary condition that  $\hat{u}^\varepsilon - U_H$  is periodic on  $\partial I_\varepsilon(x_K)$  and

$$\int_{I_\varepsilon(x_K)} (\hat{u}^\varepsilon - U_H) dx = 0,$$

where  $x_K$  is the barycenter of  $K$ .

It is easy to verify that  $\hat{u}^\varepsilon$  takes the explicit form

$$(4.2) \quad \hat{u}^\varepsilon(x) = U_H(x) + \varepsilon \chi^k\left(x_K, \frac{x}{\varepsilon}\right) \frac{\partial U_H}{\partial x_k}(x).$$

Note that the periodic extension of  $\hat{u}^\varepsilon - U_H$  is still  $\varepsilon \chi^k\left(x_K, \frac{x}{\varepsilon}\right) \frac{\partial U_H}{\partial x_k}(x)$ . This means that  $\hat{u}^\varepsilon$  is also well defined for the whole of  $K$  and takes the same explicit form as (4.2).

Using  $\int_I \chi^k(x_K, y) dy = 0$  for  $k = 1, \dots, d$  and that  $\nabla U_H$  is a piecewise constant on  $K$ , we obtain

$$(4.3) \quad \int_{I_\varepsilon(x_K)} (\hat{u}^\varepsilon - U_H)(x) dx = \int_{I_\varepsilon(x_K)} \varepsilon \chi^k\left(x_K, \frac{x}{\varepsilon}\right) \frac{\partial U_H}{\partial x_k}(x) dx = 0.$$

As in (3.7), we have

$$\|\nabla(\tilde{u}^\varepsilon - \hat{u}^\varepsilon)\|_{0,I_\varepsilon(x_K)} \leq C\varepsilon \|\nabla U_H\|_{0,I_\varepsilon(x_K)}.$$

From the construction of  $\tilde{u}^\varepsilon$ , we have for any  $x_1 \in K$ ,

$$\|\nabla(\tilde{u}^\varepsilon - \hat{u}^\varepsilon)\|_{0,I_\varepsilon(x_1)} = \|\nabla(\tilde{u}^\varepsilon - \hat{u}^\varepsilon)\|_{0,I_\varepsilon(x_K)}.$$

Since  $\nabla U_H$  is constant over  $K$ , we get

$$(4.4) \quad \|\nabla(\tilde{u}^\varepsilon - \hat{u}^\varepsilon)\|_{0,K} \leq C\varepsilon \|\nabla U_H\|_{0,K}.$$



Adding up for all  $K \in \mathcal{T}_H$  and using the a priori estimate  $\|\nabla U_H\|_0 \leq C\|f\|_0$ , we obtain

$$(4.5) \quad \left( \sum_{K \in \mathcal{T}_H} \|\nabla(\tilde{u}^\varepsilon - \hat{u}^\varepsilon)\|_{0,K}^2 \right)^{1/2} \leq C\varepsilon \|\nabla U_H\|_0 \leq C\varepsilon.$$

Using (4.2), a straightforward calculation gives

$$\frac{\partial \hat{u}^\varepsilon}{\partial x_i} = \frac{\partial U_H}{\partial x_i} + \frac{\partial \chi^k}{\partial y_i} \left( x_K, \frac{x}{\varepsilon} \right) \frac{\partial U_H}{\partial x_k}.$$

Define the first order approximation of  $u^\varepsilon$  as

$$u_1^\varepsilon(x) = U_0 + \varepsilon \chi^k \left( x, \frac{x}{\varepsilon} \right) \frac{\partial U_0}{\partial x_k},$$

where  $\{\chi^k\}_{k=1}^d$  is the solutions of (3.2). Obviously,

$$\frac{\partial u_1^\varepsilon}{\partial x_i} = \frac{\partial U_0}{\partial x_i} + \left( \varepsilon \frac{\partial \chi^k}{\partial x_i} + \frac{\partial \chi^k}{\partial y_i} \right) \left( x, \frac{x}{\varepsilon} \right) \frac{\partial U_0}{\partial x_i} + \varepsilon \chi^k \left( x, \frac{x}{\varepsilon} \right) \frac{\partial^2 U_0}{\partial x_i \partial x_k}.$$

A combination of the above estimates leads to

$$\begin{aligned} \|\nabla(\hat{u}^\varepsilon - u_1^\varepsilon)\|_{0,K} &\leq C\|\nabla(U_H - U_0)\|_{0,K} + C\varepsilon|U_0|_{1,K} \\ &\quad + \left\| \left( \frac{\partial \chi^k}{\partial y_i} \left( x, \frac{x}{\varepsilon} \right) - \frac{\partial \chi^k}{\partial y_i} \left( x_K, \frac{x}{\varepsilon} \right) \right) \frac{\partial U_0}{\partial x_k} \right\|_{0,K} + C\varepsilon|U_0|_{2,K} \\ &\leq C|U_0 - U_H|_{1,K} + C(\varepsilon + H)\|U_0\|_{2,K}. \end{aligned}$$

Summing up for all  $K \in \mathcal{T}_H$  and using Theorem 1.1 for  $k = 1$  and Theorem 1.2 for the case  $I_\delta = x_K + \varepsilon I$ , we get

$$\left( \sum_{K \in \mathcal{T}_H} \|\nabla(\hat{u}^\varepsilon - u_1^\varepsilon)\|_{0,K}^2 \right)^{1/2} \leq C(\varepsilon + H),$$

which together with (4.5) and the classical estimate for  $u^\varepsilon - u_1^\varepsilon$  [5, 32, 45], i.e.,

$$\|u^\varepsilon - u_1^\varepsilon\|_1 \leq C\sqrt{\varepsilon},$$

gives

$$\begin{aligned} \left( \sum_{K \in \mathcal{T}_H} \|\nabla(u^\varepsilon - \tilde{u}^\varepsilon)\|_{0,K}^2 \right)^{1/2} &\leq \|u^\varepsilon - u_1^\varepsilon\|_1 + \left( \sum_{K \in \mathcal{T}_H} \|\nabla(u_1^\varepsilon - \hat{u}^\varepsilon)\|_{0,K}^2 \right)^{1/2} \\ &\quad + \left( \sum_{K \in \mathcal{T}_H} \|\nabla(\hat{u}^\varepsilon - \tilde{u}^\varepsilon)\|_{0,K}^2 \right)^{1/2} \\ &\leq C(\sqrt{\varepsilon} + H). \end{aligned}$$

□

#### Corollary 4.1.

$$(4.6) \quad \|\tilde{u}^\varepsilon - u^\varepsilon\|_0 \leq C(\varepsilon + H^2).$$

*Proof.* Using the definition of  $\tilde{u}^\varepsilon$ , we have  $\int_{I_\varepsilon(x_K)} (\tilde{u}^\varepsilon - U_H)(x) dx = 0$ . Together with (4.3), we have

$$\int_{I_\varepsilon(x_K)} (\tilde{u}^\varepsilon - \hat{u}^\varepsilon)(x) dx = 0.$$

An application of the *Poincaré inequality* gives

$$\|\tilde{u}^\varepsilon - \hat{u}^\varepsilon\|_{0, I_\varepsilon(x_K)} \leq C\varepsilon \|\nabla(\tilde{u}^\varepsilon - \hat{u}^\varepsilon)\|_{0, I_\varepsilon(x_K)} \leq C\varepsilon^2 \|\nabla U_H\|_{0, I_\varepsilon(x_K)}.$$

As before for any  $I_\varepsilon(x_1)$ ,  $\|\tilde{u}^\varepsilon - \hat{u}^\varepsilon\|_{0, I_\varepsilon(x_1)} = \|\tilde{u}^\varepsilon - \hat{u}^\varepsilon\|_{0, I_\varepsilon(x_K)}$ , note that  $\nabla U_H$  is a constant on  $K$ . We obtain

$$(4.7) \quad \|\tilde{u}^\varepsilon - \hat{u}^\varepsilon\|_{0, K} \leq C\varepsilon^2 \|\nabla U_H\|_{0, K}.$$

On each element  $K$ , we have

$$\|\hat{u}^\varepsilon - U_H\|_{0, K} \leq C\varepsilon \|\nabla U_H\|_{0, K}.$$

Combining the above and summing up for all  $K \in \mathcal{T}_H$ , we get

$$\|\tilde{u}^\varepsilon - U_H\|_0 \leq C\varepsilon \|\nabla U_H\|_0 \leq C\varepsilon,$$

which together with

$$\|u^\varepsilon - U_H\|_0 \leq \|u^\varepsilon - U_0\|_0 + \|U_0 - U_H\|_0 \leq C(\varepsilon + H^2)$$

leads to (4.6), where we have used the estimate for  $U_0$  [5, 32, 45], i.e.,

$$\|u^\varepsilon - U_0\|_0 \leq C\varepsilon. \quad \square$$

**4.2. Compression operator.** The compression operator (denoted by  $Q$ ) maps the microvariables to the macrovariables [16]. It plays an important role in the general framework of HMM, even though for the present problem HMM can be formulated without explicitly specifying the compression operator beforehand. Typically the compression operator is some spatial/temporal averaging, or projection to some slow manifolds. It is of interest to consider the error bound for  $Qu^\varepsilon - U_H$ . We first list some natural properties of the *compression operator*.

- For any  $\phi \in X$ ,  $Q\phi \in X_H$ .
- There exists a constant  $C$  such that

$$\|Q\phi\|_0 \leq C\|\phi\|_0.$$

- For any  $k \geq 1$ , if  $\phi \in H^{k+1}(\Omega) \cap H_0^1(\Omega)$ , then

$$\|\phi - Q\phi\|_0 \leq CH^{k+1}|\phi|_{k+1}.$$

**Theorem 4.2.** *Assume that  $Q$  satisfies all three requirements and  $U_0 \in H^{k+1}(D)$  for any  $k \geq 1$ . Then*

$$(4.8) \quad \|Qu^\varepsilon - U_H\|_0 \leq C(\varepsilon + H^{k+1}).$$

Moreover, if  $\mathcal{T}_H$  is quasi-uniform, then

$$(4.9) \quad \|Qu^\varepsilon - U_H\|_1 \leq C\left(\frac{\varepsilon}{H} + H^k\right).$$

*Proof.* We decompose  $Qu^\varepsilon - U_H$  into

$$(4.10) \quad Qu^\varepsilon - U_H = Q(u^\varepsilon - U_0) + (QU_0 - U_0) + (U_0 - U_H).$$

Using the fact that  $Q$  is bounded in  $L_{\text{norm}}^2$ , we obtain

$$\|Q(u^\varepsilon - U_0)\|_0 \leq C\|u^\varepsilon - U_0\|_0 \leq C\varepsilon.$$

Using the third property of  $Q$ , we have

$$\|QU_0 - U_0\|_0 \leq CH^{k+1}.$$

Using Theorem 1.1 and the first estimate in Theorem 1.2, we have

$$\|U_0 - U_H\|_0 \leq C(\varepsilon + H^{k+1}).$$

A combination of these three estimates implies (4.8), which together with the inverse inequality (cf. Theorem 1.7) leads to (4.9).  $\square$

It remains to give some examples of the *compression operator*. The following two types of operators meet all three requirements:

- the  $L^2$ -projection operator onto  $X_H$ ,
- the Clément-type interpolation operator [12].

*Remark 4.3.* Notice that in one dimension, the standard Lagrange interpolant does not meet the second requirement. However, it is still possible to derive (4.9) via another approach. Moreover, a careful study of one dimensional examples shows that the term  $\varepsilon/H$  in (4.9) is sharp.

## 5. NONLINEAR HOMOGENIZATION PROBLEMS

**5.1. Algorithms and main results.** We consider the following nonlinear problem which has been discussed in [6, 23]:

$$(5.1) \quad \begin{cases} -\operatorname{div}(a^\varepsilon(x, u^\varepsilon(x)) \nabla u^\varepsilon(x)) = f(x) & x \in D, \\ u^\varepsilon(x) = 0 & x \in \partial D. \end{cases}$$

In this section, we define  $X := W_0^{1,p}(D)$  with  $p > 1$  and  $X_H$  is defined as the  $\mathcal{P}_k$  finite element subspace of  $X$ .

We assume that  $a^\varepsilon(x, u^\varepsilon)$  satisfies

$$\lambda |\xi|^2 \leq a_{ij}^\varepsilon \xi_i \xi_j \leq \Lambda |\xi|^2 \quad \text{for all } \xi \in \mathbb{R}^d,$$

with  $0 < \lambda \leq \Lambda$ . Moreover, we assume that  $a^\varepsilon(x, z)$  is equi-continuous in  $z$  uniformly with respect to  $x$  and  $\varepsilon$ .

The homogenized problem, if it exists, is of the following form:

$$\begin{cases} \mathcal{L}U_0 := -\operatorname{div}(\mathcal{A}(x, U_0(x)) \nabla U_0(x)) = f(x) & x \in D, \\ U_0(x) = 0 & x \in \partial D. \end{cases}$$

If we let

$$A(v, w) = (\mathcal{A}(x, v) \nabla v, \nabla w) \quad \text{for all } v, w \in X,$$

then

$$(5.2) \quad A(U_0, v) = (f, v) \quad \text{for all } v \in X',$$

where  $X'$  is the dual space of  $X$ .

The linearized operator of  $\mathcal{L}$  at  $U_0$  is defined for any  $v \in H_0^1(D)$  by

$$\mathcal{L}_{\text{lin}}(U_0)v = -\operatorname{div}(\mathcal{A}(x, U_0) \nabla v + \mathcal{A}_p(x, U_0) \nabla U_0 v),$$

where  $\mathcal{A}_p(x, z) := \nabla_z \mathcal{A}(x, z)$ .  $\mathcal{L}_{\text{lin}}$  induces a bilinear form through

$$\hat{A}(u; v, w) := (\mathcal{A}(x, u) \nabla v, \nabla w) + (\mathcal{A}_p(x, u) \nabla u v, \nabla w) \quad \text{for all } v, w \in H_0^1(D).$$

Our basic assumption is that the linearized operator  $\mathcal{L}_{\text{lin}}$  is an isomorphism from  $H_0^1(D)$  to  $H^{-1}(D)$ , so  $U_0$  must be an isolated solution of (5.2).

To formulate HMM, for each quadrature point  $x_\ell$ , define  $v_\ell^\varepsilon$  to be the solutions of

$$(5.3) \quad \begin{cases} -\operatorname{div}(a^\varepsilon(x, v_\ell^\varepsilon) \nabla v_\ell^\varepsilon(x)) = 0 & x \in I_\delta(x_\ell), \\ v_\ell^\varepsilon(x) = V_\ell(x) & x \in \partial I_\delta(x_\ell). \end{cases}$$

We can define  $w_\ell^\varepsilon$  similarly.

For any  $V, W \in X_H$ , define

$$\nabla W(x_\ell) \cdot \mathcal{A}_H(x_\ell, V(x_\ell)) \nabla V(x_\ell) = \int_{I_\delta(x_\ell)} \nabla w_\ell^\varepsilon(x) \cdot a^\varepsilon(x, v_\ell^\varepsilon(x)) \nabla v_\ell^\varepsilon(x) dx$$

and

$$A_H(V, W) := \sum_{K \in \mathcal{T}_H} |K| \sum_{x_\ell \in K} \omega_\ell \nabla W(x_\ell) \cdot \mathcal{A}_H(x_\ell, V(x_\ell)) \nabla V(x_\ell).$$

The HMM solution is given by the problem:

*Problem 5.1.* Find  $U_H \in X_H$  such that

$$(5.4) \quad A_H(U_H, V) = (f, V) \quad \text{for all } V \in X_H.$$

For any  $v, v_H, w \in X$ , define

$$(5.5) \quad R(v, v_H, w) := A(v_H, w) - A(v, w) - \hat{A}(v; v_H - v, w).$$

It is easy to see that for any  $v$  and  $v_H$  satisfying  $\|v\|_{1, \infty} + \|v_H\|_{1, \infty} \leq M$ ,

$$(5.6) \quad |R(v, v_H, w)| \leq C(M) (\|e_H\|_{0, 2p}^2 + \|e_H \nabla e_H\|_{0, p}) \|\nabla w\|_{0, q}$$

for  $e_H := v - v_H$  and  $\frac{1}{p} + \frac{1}{q} = 1, p, q \geq 1$  (see [42, Lemma 3.1] for a similar result). Therefore we have

**Lemma 5.2.**  $U_H \in X_H$  is the solution of Problem 5.1 if and only if

$$(5.7) \quad \begin{aligned} \hat{A}(U_0; U_0 - U_H, V) &= R(U_0, U_H, V) \\ &+ A_H(U_H, V) - A(U_H, V) \quad \text{for all } V \in X_H. \end{aligned}$$

For any  $V, W \in X_H$ , define

$$(5.8) \quad E(V, W) := \nabla W(x_\ell) \cdot (\mathcal{A}_H - \mathcal{A})(x_\ell, V(x_\ell)) \nabla V(x_\ell).$$

Define  $e(\text{HMM})$  as

$$(5.9) \quad e(\text{HMM}) = \max_{\substack{x_\ell \in K, K \in \mathcal{T}_H \\ V \in X_H \cap W^{1, \infty}(D), W \in X_H}} \frac{|E(V, W)|}{|\nabla V_\ell| |\nabla W_\ell|}.$$

The existence and uniqueness of the solution of Problem 5.1 are proved in the following lemma.

**Lemma 5.3.** Assume that  $U_0 \in W^{2, p}(D)$  with  $p > d$  and  $\mathcal{L}_{\text{lin}}$  is an isomorphism from  $H_0^1(D)$  to  $H^{-1}(D)$ . If  $e(\text{HMM})$  is uniformly bounded and there exist constants  $H_0 > 0$  and  $\mathcal{M}_1 > 0$  such that for  $0 < H \leq H_0$  and

$$(5.10) \quad e(\text{HMM})^{1/2} |\ln H| \leq \mathcal{M}_1,$$

then Problem 5.1 has a solution  $U_H$  satisfying

$$(5.11) \quad \begin{aligned} \|U_H - P_H U_0\|_{1, \infty} &\leq e(\text{HMM})^{1/2} + H^{1-d/p}, \\ \|U_0 - U_H\|_{1, \infty} &\leq C(e(\text{HMM})^{1/2} + H^{1-d/p}), \end{aligned}$$

where  $P_H U_0 \in X_H$  is defined as

$$(5.12) \quad \hat{A}(U_0; P_H U_0, V) = \hat{A}(U_0; U_0, V) \quad \text{for all } V \in X_H.$$

Moreover, if there exists a constant  $\eta(M)$  with  $0 < \eta(M) < 1$  such that

$$(5.13) \quad \sum_{K \in \mathcal{T}_H} |K| \sum_{x_\ell \in K} \omega_\ell |E(V, Z) - E(W, Z)| \leq \eta(M) \|V - W\|_1 \|Z\|_1$$

for all  $V, W \in X_H \cap W^{1,\infty}(D)$  and  $Z \in X_H$ , satisfying  $\|V\|_{1,\infty} + \|W\|_{1,\infty} \leq M$ , then there exists a constant  $H_1 > 0$  such that for  $0 < H \leq H_1$ , the HMM solution  $U_H$  satisfying (5.11)<sub>1</sub> is locally unique.

*Proof.* Since  $\mathcal{L}_{\text{in}}$  is an isomorphism from  $H_0^1(D)$  to  $H^{-1}(D)$ , there exists a constant  $C$  such that

$$\sup_{W \in H_0^1(D)} \frac{\hat{A}(U_0; V, W)}{\|W\|_1} \geq C \|V\|_1 \quad \text{for all } V \in H_0^1(D).$$

Using [42, Lemma 2.2], we conclude that there exists a constant  $H_2 > 0$  such that for  $0 < H \leq H_2$ ,

$$(5.14) \quad \sup_{W \in X_H} \frac{\hat{A}(U_0; V, W)}{\|W\|_1} \geq C \|V\|_1 \quad \text{for all } V \in X_H.$$

Therefore there is a unique solution  $P_H U_0 \in X_H$  satisfying (5.12) and

$$(5.15) \quad \|U_0 - P_H U_0\|_{1,\infty} \leq C H^{1-d/p}.$$

Moreover, let  $\hat{G}_H^z$  be the finite element approximation of the regularized Green's function associated with  $\hat{A}(U_0; \cdot, \cdot)$ . Using [42, equation 2.11], or using (5.14), similarly to (2.11), we have

$$(5.16) \quad \|\hat{G}_H^z\|_{1,1} \leq C |\ln H|.$$

Define a nonlinear mapping  $T: X_H \rightarrow X_H$  by

$$\hat{A}(U_0; T(V), W) = \hat{A}(U_0; U_0, W) - R(U_0, V, W) + A(V, W) - A_H(V, W),$$

for any  $W \in X_H$ . Obviously  $T$  is continuous due to (5.14) and (5.6).

Let

$$B := \{ V \in X_H \mid \|V - P_H U_0\|_{1,\infty} \leq e(\text{HMM})^{1/2} + H^{1-d/p} \}.$$

We next prove that there exists a constant  $H_0 > 0$  such that for all  $0 < H \leq H_0$ ,  $T(B) \subset B$ .

Notice that

$$\hat{A}(U_0; T(V) - P_H U_0, W) = -R(U_0, V, W) + A(V, W) - A_H(V, W).$$

Taking  $W = \hat{G}_H^z$  in the above equation, using (5.16) and (5.6), we obtain, for  $\|V\|_{1,\infty} \leq M$ ,

$$\begin{aligned} \|T(V) - P_H U_0\|_{1,\infty} &\leq C(M) \|U_0 - V\|_{1,\infty}^2 |\ln H| + C(e(\text{HMM}) + H) |\ln H| \|V\|_{1,\infty} \\ &\leq C(M) (\|U_0 - P_H U_0\|_{1,\infty}^2 + \|P_H U_0 - V\|_{1,\infty}^2) |\ln H| + CM(e(\text{HMM}) + H) |\ln H| \\ &\leq C(M) (e(\text{HMM}) + H^{2-2d/p}) |\ln H| + CM(e(\text{HMM}) + H) |\ln H|. \end{aligned}$$

Since  $V \in B$  and  $e(\text{HMM})$  is uniformly bounded, e.g.,  $e(\text{HMM}) \leq M_1$ , we have

$$(5.17) \quad \begin{aligned} \|V\|_{1,\infty} &\leq \|V - P_H U_0\|_{1,\infty} + \|P_H U_0\|_{1,\infty} \leq C(U_0) + e(\text{HMM})^{1/2} \\ &\leq C(U_0) + M_1^{1/2} =: M_0. \end{aligned}$$

Combining the above two estimates, we obtain

$$\|T(V) - P_H U_0\|_{1,\infty} \leq C(M_0)(e(\text{HMM}) + H^{2-2d/p} + H)|\ln H|.$$

Define  $\mathcal{M}_1 := 1/C(M_0)$ . Using (5.10), we obtain

$$\|T(V) - P_H U_0\|_{1,\infty} \leq e(\text{HMM})^{1/2} + C(M_0)(H^{2-2d/p} + H)|\ln H|.$$

Therefore there exists a constant  $H_3$  such that for  $0 < H \leq H_3$ , we have

$$\|T(V) - P_H U_0\|_{1,\infty} \leq e(\text{HMM})^{1/2} + H^{1-d/p}.$$

Let  $H_0 := \min(H_2, H_3)$ . Then for  $0 < H \leq H_0$ , we have  $T(B) \subset B$ . An application of Brouwer's fixed point theorem gives the existence of a  $U_H \in B$  such that  $T(U_H) = U_H$ . By definition,  $U_H$  satisfies (5.11)<sub>1</sub>. Together with (5.15) it yields (5.11)<sub>2</sub>.

To prove uniqueness, assume that both  $U_H$  and  $\hat{U}_H$  are solutions of (5.4) satisfying (5.11)<sub>1</sub>. Using (5.14), we obtain

$$\begin{aligned} C\|U_H - \hat{U}_H\|_1 &\leq \sup_{W \in X_H} \frac{\int_0^1 \hat{A}(U_H^t; U_H - \hat{U}_H, W) dt}{\|W\|_1} \\ &\leq \sup_{W \in X_H} \frac{|A(U_H, W) - A(\hat{U}_H, W)|}{\|W\|_1}, \end{aligned}$$

where  $U_H^t = (1-t)\hat{U}_H + tU_H$ . Note that

$$\begin{aligned} A(U_H, W) - A(\hat{U}_H, W) &= (A(U_H, W) - A_H(U_H, W)) \\ &\quad - (A(\hat{U}_H, W) - A_H(\hat{U}_H, W)). \end{aligned}$$

Since both  $U_H$  and  $\hat{U}_H$  sit in the set  $B$ , we can use (5.17) to get  $\|U_H\|_{1,\infty} + \|\hat{U}_H\|_{1,\infty} \leq 2M_0$ . Using (5.13), we obtain

$$\|U_H - \hat{U}_H\|_1 \leq (\eta(2M_0) + C_1 H)\|U_H - \hat{U}_H\|_1.$$

If we choose  $H_1$  such that

$$\eta(2M_0) + C_1 H_1 < 1,$$

then if  $H < H_1$ , we have  $U_H = \hat{U}_H$ . Therefore the HMM solution is locally unique.  $\square$

From here on, when we talk about the HMM solution, we are referring to this particular solution that satisfies the condition in Lemma 5.3.

Based on the above lemma, we prove a nonlinear analog of Theorem 1.1.

**Theorem 5.4.** *Under the assumptions in Lemma 5.3, let  $U_0$  and  $U_H$  be solutions of (5.2) and (5.4), respectively. Assume in addition that  $U_0 \in W^{k+1,\infty}(D)$ . Then there exist constants  $H_0$  and  $\mathcal{M}_1^*$  such that if  $0 < H < H_0$  and  $\mathcal{M}_1 < \mathcal{M}_1^*$ , then*

$$(5.18) \quad \|U_0 - U_H\|_1 \leq C(H^k + e(\text{HMM})),$$

$$(5.19) \quad \|U_0 - U_H\|_{1,\infty} \leq C(H^k + e(\text{HMM}))|\ln H|.$$

*Proof.* Note that  $U_0 \in W^{k+1,\infty}(D)$  and from (5.14), we have

$$(5.20) \quad \|U_0 - P_H U_0\|_1 \leq C H^k, \quad \|U_0 - P_H U_0\|_{1,\infty} \leq C H^k.$$

Using (5.7) with  $V = P_H U_0 - U_H$  and (5.14) and (2.3), we obtain

$$(5.21) \quad \|P_H U_0 - U_H\|_1 \leq C \|U_0 - U_H\|_{1,4}^2 + C(H^k + e(\text{HMM})).$$

Using the interpolation inequality

$$\|U_0 - U_H\|_{1,4}^2 \leq \|U_0 - U_H\|_1 \|U_0 - U_H\|_{1,\infty}$$

together with (5.20) and (5.17) gives

$$\begin{aligned} \|P_H U_0 - U_H\|_1 &\leq C_1 (e(\text{HMM})^{1/2} + H^{1-d/p}) \|P_H U_0 - U_H\|_1 \\ &\quad + C(H^k + e(\text{HMM})). \end{aligned}$$

Let  $V = \hat{G}_H^z$  in (5.7). Using (5.16), we obtain

$$\begin{aligned} \|P_H U_0 - U_H\|_{1,\infty} &\leq C_2 (e(\text{HMM}) + H) |\ln H| \|P_H U_0 - U_H\|_{1,\infty} \\ &\quad + C(\|U_0 - U_H\|_{1,\infty}^2 + e(\text{HMM}) + H^k) |\ln H|. \end{aligned}$$

Since (5.10) holds, using (5.11)<sub>2</sub> and (5.20)<sub>2</sub>, we obtain

$$\begin{aligned} \|P_H U_0 - U_H\|_{1,\infty} &\leq C_2 (e(\text{HMM})^{1/2} + H^{1-d/p}) |\ln H| \|P_H U_0 - U_H\|_{1,\infty} \\ &\quad + C(e(\text{HMM}) + H^k) |\ln H|. \end{aligned}$$

Now we choose

$$\mathcal{M}_1^* := \min\left(\frac{|\ln H|}{2C_1}, \frac{1}{2C_2}\right)$$

and  $H_0$  such that  $e(\text{HMM})^{1/2} |\ln H| \leq M_1^*$ , and therefore

$$\begin{aligned} C_1 (e(\text{HMM})^{1/2} + H^{1-d/p}) &\leq \frac{1}{2} + C_1 H_0^{1-d/p} < 1, \\ C_2 (e(\text{HMM})^{1/2} + H^{1-d/p}) |\ln H| &\leq \frac{1}{2} + C_2 H^{1-d/p} |\ln H| < 1. \end{aligned}$$

Thus we obtain

$$\begin{aligned} \|P_H U_0 - U_H\|_1 &\leq C(H^k + e(\text{HMM})), \\ \|P_H U_0 - U_H\|_{1,\infty} &\leq C(H^k + e(\text{HMM})) |\ln H|. \end{aligned}$$

Using (5.20) once again gives (5.18) and (5.19).  $\square$

**5.2. Estimating  $e(\text{HMM})$ .** It remains to estimate  $e(\text{HMM})$  and verify assumptions (5.10) and (5.13). We assume that  $a^\varepsilon(x, u^\varepsilon) = (a_{ij}(x, x/\varepsilon, u^\varepsilon))$ , and for  $1 \leq i, j \leq d$ , the coefficients  $a_{ij}^\varepsilon(x, y, z)$  are smooth in  $x, z$  and periodic in  $y$  with period  $I$ . These types of problems, among others, have been considered in [5, 6, 23]. The homogenized coefficient  $\mathcal{A} = (\mathcal{A}_{ij}(x, p))$  is given for any  $p \in \mathbb{R}$  by

$$\mathcal{A}_{ij}(x, p) = \int_I \left( a_{ij} + a_{ik} \frac{\partial \chi^j}{\partial y_k} \right) (x, y, p) dy,$$

where  $\{\chi^k\}_{k=1}^d$  is defined for any  $p \in \mathbb{R}$  by

$$(5.22) \quad -\frac{\partial}{\partial y_i} \left( a_{ik} \frac{\partial \chi^j}{\partial y_k} \right) (x, y, p) = \frac{\partial}{\partial y_i} a_{ij}(x, y, p),$$

with the periodic boundary condition in  $y$  and  $\int_I \chi^k(x, y, p) dy = 0$ . It is clear that  $\mathcal{A}(x, p)$  is also smooth in  $x$  and  $p$  and satisfies [6, Proposition 3.5]

$$(5.23) \quad \lambda I \leq \mathcal{A} \leq \frac{\Lambda^2}{\lambda} I.$$

Using Lemma 1.9 and Remark 1.10, for the solution of (5.3), we have

$$(5.24) \quad \|\nabla V_\ell\|_{0, I_\delta} \leq \|\nabla v_\ell^\varepsilon\|_{0, I_\delta} \leq \sqrt{\Lambda/\lambda} \|\nabla V_\ell\|_{0, I_\delta}.$$

This gives a bound for  $\mathcal{A}_H$ :

$$\lambda I \leq \mathcal{A}_H \leq \frac{\Lambda^2}{\lambda} I,$$

which together with (5.23) implies

$$(5.25) \quad e(\text{HMM}) \leq 2\Lambda^2/\lambda.$$

This shows that  $e(\text{HMM})$  is uniformly bounded.

To simplify the presentation, we will show how to estimate  $e(\text{HMM})$  when (5.3) is changed slightly to

$$(5.26) \quad \begin{cases} -\operatorname{div}(a^\varepsilon(x, V(x_\ell)) \nabla v_\ell^\varepsilon(x)) = 0 & x \in I_\delta(x_\ell), \\ v_\ell^\varepsilon(x) = V_\ell(x) & x \in \partial I_\delta(x_\ell) \end{cases}$$

and  $A_H(V, W)$  is changed to

$$A_H(V, W) = \sum_{K \in \mathcal{T}_H} |K| \sum_{x_\ell \in K} \omega_\ell \int_{I_\delta(x_\ell)} \nabla w_\ell^\varepsilon(x) \cdot a^\varepsilon(x, V(x_\ell)) \nabla v_\ell^\varepsilon(x) dx.$$

If  $\delta = \varepsilon$ , we replace the Dirichlet boundary condition in (5.26) by the periodic boundary condition, i.e.,  $v_\ell^\varepsilon(x) - V_\ell(x)$  is periodic on  $\partial I_\varepsilon(x_\ell)$ .

**Theorem 5.5.** *If  $(\sqrt{\varepsilon/\delta} + \delta)|\ln H|$  is sufficiently small, then (5.10) and (5.13) hold and*

$$(5.27) \quad e(\text{HMM}) \leq C \left( \left( \frac{\varepsilon}{\delta} \right)^{1/2} + \delta \right).$$

*In the case of  $\delta = \varepsilon$ , if  $\varepsilon|\ln H|$  is sufficiently small, then (5.10) and (5.13) hold and*

$$(5.28) \quad e(\text{HMM}) \leq C\varepsilon.$$

In what follows, we concentrate on the first case. The second case when  $\delta = \varepsilon$  will be commented on.

Let us first fix more notation. Denote by  $\hat{v}_\ell^\varepsilon$  the solutions of (5.26) with the coefficient  $a(x, x/\varepsilon, V(x_\ell))$  replaced by  $a(x_\ell, x/\varepsilon, V(x_\ell))$ . Similarly we define  $w_\ell^\varepsilon$  to be the solution of (5.26) with  $V$  replaced by  $W \in X_H$ . Also,  $\hat{w}_\ell^\varepsilon$  can be defined in the same way and  $w_\ell^\varepsilon$  and  $\hat{w}_\ell^\varepsilon$  can be viewed as the perturbations of  $v_\ell^\varepsilon$  and  $\hat{v}_\ell^\varepsilon$ , respectively. Moreover, we define

$$\begin{aligned} a_V(x_\ell) &= a(x_\ell, \frac{x}{\varepsilon}, V(x_\ell)), & a_W(x_\ell) &= a(x_\ell, \frac{x}{\varepsilon}, W(x_\ell)), \\ a_V(x) &= a(x, \frac{x}{\varepsilon}, V(x_\ell)), & a_W(x) &= a(x, \frac{x}{\varepsilon}, W(x_\ell)). \end{aligned}$$

Observe that  $\hat{v}_\ell^\varepsilon$  and  $\hat{w}_\ell^\varepsilon$  also satisfy (5.24), and using Lemma 1.8, we have

$$(5.29) \quad \|\nabla(v_\ell^\varepsilon - \hat{v}_\ell^\varepsilon)\|_{0, I_\delta} \leq C\delta \|\nabla V_\ell\|_{0, I_\delta}, \quad \|\nabla(w_\ell^\varepsilon - \hat{w}_\ell^\varepsilon)\|_{0, I_\delta} \leq C\delta \|\nabla W_\ell\|_{0, I_\delta}.$$



**Lemma 5.6.** *We have*

$$(5.30) \quad \begin{aligned} \|\nabla(\hat{v}_\ell^\varepsilon - \hat{w}_\ell^\varepsilon)\|_{0,I_\delta} &\leq C(|(V - W)(x_\ell)|(\|\nabla V_\ell\|_{0,I_\delta} + \|\nabla W_\ell\|_{0,I_\delta}) \\ &\quad + \|\nabla(V_\ell - W_\ell)\|_{0,I_\delta}). \end{aligned}$$

*Proof.* Observe that

$$\begin{aligned} -\operatorname{div}(a_V(x_\ell)\nabla(\hat{v}_\ell^\varepsilon - V_\ell)) &= \operatorname{div}(a_V(x_\ell)\nabla V_\ell), \\ -\operatorname{div}(a_W(x_\ell)\nabla(\hat{w}_\ell^\varepsilon - W_\ell)) &= \operatorname{div}(a_W(x_\ell)\nabla W_\ell). \end{aligned}$$

Both  $\hat{v}_\ell^\varepsilon - V_\ell$  and  $\hat{w}_\ell^\varepsilon - W_\ell$  vanish on  $\partial I_\delta(x_\ell)$ . Using Lemma 1.8, we obtain

$$\begin{aligned} \lambda\|\nabla(\hat{v}_\ell^\varepsilon - V_\ell - \hat{w}_\ell^\varepsilon + W_\ell)\|_{0,I_\delta} &\leq \Lambda\|\nabla(V_\ell - W_\ell)\|_{0,I_\delta} \\ &\quad + \max_{x \in I_\delta} |(a_V - a_W)(x_\ell)|(\|\nabla V_\ell\|_{0,I_\delta} + \|\nabla(\hat{w}_\ell^\varepsilon - W_\ell)\|_{0,I_\delta}). \end{aligned}$$

Using (5.24) for  $\hat{w}_\ell^\varepsilon$ , we obtain

$$\|\nabla(\hat{w}_\ell^\varepsilon - W_\ell)\|_{0,I_\delta} \leq \|\nabla\hat{w}_\ell^\varepsilon\|_{0,I_\delta} + \|\nabla W_\ell\|_{0,I_\delta} \leq C(1 + \sqrt{\Lambda/\lambda})\|\nabla W_\ell\|_{0,I_\delta},$$

which together with  $\max_{x \in I_\delta} |(a_V - a_W)(x_\ell)| \leq C|(V - W)(x_\ell)|$  gives (5.30).  $\square$

Next we establish the estimate for  $(v_\ell^\varepsilon - \hat{v}_\ell^\varepsilon) - (w_\ell^\varepsilon - \hat{w}_\ell^\varepsilon)$ . Let  $\psi_\ell^\varepsilon := v_\ell^\varepsilon - \hat{v}_\ell^\varepsilon$  and  $\tilde{\psi}_\ell^\varepsilon := w_\ell^\varepsilon - \hat{w}_\ell^\varepsilon$ . Clearly,  $\psi_\ell^\varepsilon, \tilde{\psi}_\ell^\varepsilon$  vanish on  $\partial I_\delta(x_\ell)$  and satisfy

$$\begin{aligned} -\operatorname{div}(a_V(x)\nabla\psi_\ell^\varepsilon) &= \operatorname{div}((a_V(x) - a_V(x_\ell))\nabla\hat{v}_\ell^\varepsilon) \quad x \in I_\delta(x_\ell), \\ -\operatorname{div}(a_W(x)\nabla\tilde{\psi}_\ell^\varepsilon) &= \operatorname{div}((a_W(x) - a_W(x_\ell))\nabla\hat{w}_\ell^\varepsilon) \quad x \in I_\delta(x_\ell). \end{aligned}$$

**Lemma 5.7.** *We have*

$$(5.31) \quad \begin{aligned} \|\nabla(\psi_\ell^\varepsilon - \tilde{\psi}_\ell^\varepsilon)\|_{0,I_\delta} &\leq C\delta(|(V - W)(x_\ell)|(\|\nabla V_\ell\|_{0,I_\delta} + \|\nabla W_\ell\|_{0,I_\delta}) \\ &\quad + \|\nabla(V_\ell - W_\ell)\|_{0,I_\delta}). \end{aligned}$$

*Proof.* Using Lemma 1.8, we have

$$\begin{aligned} \lambda\|\nabla(\psi_\ell^\varepsilon - \tilde{\psi}_\ell^\varepsilon)\|_{0,I_\delta} &\leq \max_{x \in I_\delta} |a_V(x) - a_V(x_\ell) - a_W(x) + a_W(x_\ell)| \|\nabla\hat{v}_\ell^\varepsilon\|_{0,I_\delta} \\ &\quad + \max_{x \in I_\delta} |a_V(x) - a_W(x)| \|\nabla\tilde{\psi}_\ell^\varepsilon\|_{0,I_\delta} \\ &\quad + \max_{x \in I_\delta} |a_W(x) - a_W(x_\ell)| \|\nabla(\hat{v}_\ell^\varepsilon - \hat{w}_\ell^\varepsilon)\|_{0,I_\delta}. \end{aligned}$$

Using

$$(5.32) \quad \max_{x \in I_\delta} |a_V(x) - a_V(x_\ell) - a_W(x) + a_W(x_\ell)| \leq C\delta|(V - W)(x_\ell)|,$$

from (5.29)<sub>2</sub>, we have  $\|\nabla\tilde{\psi}_\ell^\varepsilon\|_{0,I_\delta} \leq C\delta\|\nabla W_\ell\|_{0,I_\delta}$ . Collecting the above estimates and using (5.30), we obtain (5.31).  $\square$

Define

$$\begin{aligned} \hat{V}_\ell(x) &= V_\ell(x) + \varepsilon\chi_\ell^k\left(x_\ell, \frac{x}{\varepsilon}, V(x_\ell)\right)\frac{\partial V_\ell}{\partial x_k}(x), \\ \hat{W}_\ell(x) &= W_\ell(x) + \varepsilon\chi_\ell^k\left(x_\ell, \frac{x}{\varepsilon}, W(x_\ell)\right)\frac{\partial W_\ell}{\partial x_k}(x), \end{aligned}$$

where  $\{\chi_\ell^k\}_{k=1}^d$  are the solutions of (5.22) with coefficient replaced by  $a_{ij}(x_\ell, y, p)$ .

Denote  $\theta_\ell^\varepsilon = \hat{v}_\ell^\varepsilon - \hat{V}_\ell$  and  $\tilde{\theta}_\ell^\varepsilon = \hat{w}_\ell^\varepsilon - \hat{W}_\ell$ . Observe that

$$\begin{cases} -\operatorname{div}(a_V(x_\ell)\nabla\theta_\ell^\varepsilon(x)) = 0 & x \in I_\delta(x_\ell), \\ \theta_\ell^\varepsilon(x) = -\varepsilon\chi_\ell^k(x_\ell, \frac{x}{\varepsilon}, V(x_\ell))\frac{\partial V_\ell}{\partial x_k} & x \in \partial I_\delta(x_\ell), \end{cases}$$

and

$$\begin{cases} -\operatorname{div}(a_W(x_\ell)\nabla\tilde{\theta}_\ell^\varepsilon(x)) = 0 & x \in I_\delta(x_\ell), \\ \tilde{\theta}_\ell^\varepsilon(x) = -\varepsilon\chi_\ell^k(x_\ell, \frac{x}{\varepsilon}, W(x_\ell))\frac{\partial W_\ell}{\partial x_k} & x \in \partial I_\delta(x_\ell). \end{cases}$$

Similarly to (3.9), we have

$$(5.33) \quad \|\nabla\theta_\ell^\varepsilon\|_{0, I_\delta} \leq C\left(\frac{\varepsilon}{\delta}\right)^{1/2}\|\nabla V_\ell\|_{0, I_\delta}, \quad \|\nabla\tilde{\theta}_\ell^\varepsilon\|_{0, I_\delta} \leq C\left(\frac{\varepsilon}{\delta}\right)^{1/2}\|\nabla W_\ell\|_{0, I_\delta}.$$

Let  $\rho^\varepsilon$  be defined as in (3.14) and define

$$\begin{aligned} \varphi_\ell^\varepsilon &= \theta_\ell^\varepsilon + \varepsilon\chi_\ell^k(x_\ell, \frac{x}{\varepsilon}, V(x_\ell))\frac{\partial V_\ell}{\partial x_k}(1 - \rho^\varepsilon), \\ \tilde{\varphi}_\ell^\varepsilon &= \tilde{\theta}_\ell^\varepsilon + \varepsilon\chi_\ell^k(x_\ell, \frac{x}{\varepsilon}, W(x_\ell))\frac{\partial W_\ell}{\partial x_k}(1 - \rho^\varepsilon). \end{aligned}$$

Observe that  $\varphi_\ell^\varepsilon$  and  $\tilde{\varphi}_\ell^\varepsilon$  vanish on  $\partial I_\delta(x_\ell)$  and satisfy

$$\begin{aligned} -\operatorname{div}(a_V(x_\ell)\nabla\varphi_\ell^\varepsilon(x)) &= \operatorname{div}(a_V(x_\ell)\nabla(\varphi_\ell^\varepsilon - \theta_\ell^\varepsilon)(x)) & x \in I_\delta(x_\ell), \\ -\operatorname{div}(a_W(x_\ell)\nabla\tilde{\varphi}_\ell^\varepsilon(x)) &= \operatorname{div}(a_W(x_\ell)\nabla(\tilde{\varphi}_\ell^\varepsilon - \tilde{\theta}_\ell^\varepsilon)(x)) & x \in I_\delta(x_\ell). \end{aligned}$$

**Lemma 5.8.** *We have*

$$(5.34) \quad \begin{aligned} \|\nabla(\theta_\ell^\varepsilon - \tilde{\theta}_\ell^\varepsilon)\|_{0, I_\delta} &\leq C\left(\frac{\varepsilon}{\delta}\right)^{1/2}\left(\|\nabla(V_\ell - W_\ell)\|_{0, I_\delta} \right. \\ &\quad \left. + |(V - W)(x_\ell)|(\|\nabla V_\ell\|_{0, I_\delta} + \|\nabla W_\ell\|_{0, I_\delta})\right). \end{aligned}$$

*Proof.* Using Lemma 1.8, we obtain

$$\begin{aligned} \lambda\|\nabla(\varphi_\ell^\varepsilon - \tilde{\varphi}_\ell^\varepsilon)\|_{0, I_\delta} &\leq \max_{x \in I_\delta} |a_V(x_\ell) - a_W(x_\ell)|(\|\nabla(\varphi_\ell^\varepsilon - \theta_\ell^\varepsilon)\|_{0, I_\delta} + \|\nabla\tilde{\varphi}_\ell^\varepsilon\|_{0, I_\delta}) \\ &\quad + \max_{x \in I_\delta} |a_W(x_\ell)|\|\nabla(\varphi_\ell^\varepsilon - \theta_\ell^\varepsilon - \tilde{\varphi}_\ell^\varepsilon + \tilde{\theta}_\ell^\varepsilon)\|_{0, I_\delta}. \end{aligned}$$

A direct computation gives that

$$\|\nabla(\varphi_\ell^\varepsilon - \theta_\ell^\varepsilon)\|_{0, I_\delta} \leq C\left(\frac{\varepsilon}{\delta}\right)^{\frac{1}{2}}\|\nabla V_\ell\|_{0, I_\delta} \quad \text{and} \quad \|\nabla(\tilde{\varphi}_\ell^\varepsilon - \tilde{\theta}_\ell^\varepsilon)\|_{0, I_\delta} \leq C\left(\frac{\varepsilon}{\delta}\right)^{\frac{1}{2}}\|\nabla W_\ell\|_{0, I_\delta},$$

which together with (5.33) gives

$$\|\nabla\tilde{\varphi}_\ell^\varepsilon\|_{0, I_\delta} \leq \|\nabla\tilde{\theta}_\ell^\varepsilon\|_{0, I_\delta} + \|\nabla(\tilde{\varphi}_\ell^\varepsilon - \tilde{\theta}_\ell^\varepsilon)\|_{0, I_\delta} \leq C\left(\frac{\varepsilon}{\delta}\right)^{1/2}\|\nabla W_\ell\|_{0, I_\delta}.$$

Note that

$$\begin{aligned} \varphi_\ell^\varepsilon - \theta_\ell^\varepsilon - \tilde{\varphi}_\ell^\varepsilon + \tilde{\theta}_\ell^\varepsilon &= \varepsilon(1 - \rho^\varepsilon) \left( \chi_\ell^k(x_\ell, \frac{x}{\varepsilon}, V(x_\ell)) - \chi_\ell^k(x_\ell, \frac{x}{\varepsilon}, W(x_\ell)) \right) \frac{\partial V_\ell}{\partial x_k} \\ &\quad + \varepsilon(1 - \rho^\varepsilon) \chi_\ell^k(x_\ell, \frac{x}{\varepsilon}, W(x_\ell)) \frac{\partial}{\partial x_k} (V_\ell - W_\ell). \end{aligned}$$

Using the continuity of  $\{\chi_\ell^k\}_{k=1}^d$ , a direct computation gives

$$\|\nabla(\varphi_\ell^\varepsilon - \theta_\ell^\varepsilon - \tilde{\varphi}_\ell^\varepsilon + \tilde{\theta}_\ell^\varepsilon)\|_{0, I_\delta} \leq C \left( \frac{\varepsilon}{\delta} \right)^{1/2} \left( |(V - W)(x_\ell)| \|\nabla V_\ell\|_{0, I_\delta} + \|\nabla(V_\ell - W_\ell)\|_{0, I_\delta} \right).$$

Adding these up, we obtain (5.34).  $\square$

In the next lemma, we shall prove that  $E(V, W)$  has certain continuity with respect to  $V$ .

**Lemma 5.9.** *For any  $V, W, Z \in X_H$  satisfying  $\|V\|_{1, \infty} + \|W\|_{1, \infty} \leq M$ , there exists a constant  $C(M)$  such that (5.13) holds with  $\eta(M) = C(M)(\sqrt{\varepsilon/\delta} + \delta)$ .*

*Proof.* Using the definition of  $v_\ell^\varepsilon$  and noticing that  $v_\ell^\varepsilon = (v_\ell^\varepsilon - \hat{v}_\ell^\varepsilon) + \theta_\ell^\varepsilon + \hat{V}_\ell$ , we obtain

$$\begin{aligned} \int_{I_\delta} \nabla z_\ell^\varepsilon \cdot a_V(x) \nabla v_\ell^\varepsilon \, dx &= \int_{I_\delta} \nabla Z_\ell \cdot a_V(x) \nabla v_\ell^\varepsilon \, dx \\ &= \nabla Z_\ell \cdot \left( \int_{I_\delta} a_V(x) (\nabla(v_\ell - \hat{v}_\ell^\varepsilon) + \nabla\theta_\ell^\varepsilon + \nabla\hat{V}_\ell) \, dx \right), \end{aligned}$$

which together with  $\int_{I_{\kappa\varepsilon}} \nabla Z_\ell \cdot a_V(x_\ell) \nabla \hat{V}_\ell \, dx = \nabla Z(x_\ell) \cdot \mathcal{A}(x_\ell, V(x_\ell)) \nabla V(x_\ell)$  gives

$$E(V, Z) - E(W, Z) = :I_1 + I_2 + I_3,$$

with

$$\begin{aligned} I_1 &= \nabla Z_\ell \cdot \left[ \int_{I_\delta} (a_V(x) \nabla(v_\ell^\varepsilon - \hat{v}_\ell^\varepsilon) - a_W(x) \nabla(w_\ell^\varepsilon - \hat{w}_\ell^\varepsilon)) \, dx \right], \\ I_2 &= \nabla Z_\ell \cdot \left[ \int_{I_\delta} a_V(x) \nabla\theta_\ell^\varepsilon - a_W(x) \nabla\tilde{\theta}_\ell^\varepsilon \, dx \right], \end{aligned}$$

and

$$\begin{aligned} I_3 &= \nabla Z_\ell \cdot \left[ \int_{I_\delta} a_V(x) \nabla\hat{V}_\ell \, dx - \int_{I_{\kappa\varepsilon}} a_V(x_\ell) \nabla\hat{V}_\ell \, dx \right. \\ &\quad \left. - \int_{I_\delta} a_W(x) \nabla\hat{W}_\ell \, dx + \int_{I_{\kappa\varepsilon}} a_W(x_\ell) \nabla\hat{W}_\ell \, dx \right]. \end{aligned}$$

$I_1$  can be decomposed into

$$I_1 = \nabla Z_\ell \cdot \left[ \int_{I_\delta} (a_V(x) - a_W(x)) \nabla(v_\ell^\varepsilon - \hat{v}_\ell^\varepsilon) \, dx + \int_{I_\delta} a_W(x) \nabla(\psi_\ell^\varepsilon - \tilde{\psi}_\ell^\varepsilon) \, dx \right].$$

Using (5.29) and (5.31), we bound  $I_1$  as

$$|I_1| \leq C\delta \left( |\nabla(V_\ell - W_\ell)| + |(V - W)(x_\ell)| (|\nabla V_\ell| + |\nabla W_\ell|) \right) |\nabla Z_\ell|.$$

Similarly, using Lemma 5.8, we bound  $I_2$  as

$$|I_2| \leq C \left( \frac{\varepsilon}{\delta} \right)^{1/2} \left( |\nabla(V_\ell - W_\ell)| + |(V - W)(x_\ell)| (|\nabla V_\ell| + |\nabla W_\ell|) \right) |\nabla Z_\ell|.$$

$I_3$  can be rewritten as

$$\begin{aligned} I_3 &= \nabla Z_\ell \cdot \int_{I_\delta} \left( (a_V(x) - a_V(x_\ell)) \nabla \hat{V}_\ell - (a_W(x) - a_W(x_\ell)) \nabla \hat{W}_\ell \right) dx \\ &\quad - \nabla Z_\ell \cdot \left( \frac{1}{|\kappa\varepsilon|^d} - \frac{1}{\delta^d} \right) \int_{I_\delta} (a_V(x_\ell) \nabla \hat{V}_\ell - a_W(x_\ell) \nabla \hat{W}_\ell) dx \\ &\quad + \frac{1}{|\kappa\varepsilon|^d} \nabla Z_\ell \cdot \int_{I_\delta \setminus I_{\kappa\varepsilon}} (a_V(x_\ell) \nabla \hat{V}_\ell - a_W(x_\ell) \nabla \hat{W}_\ell) dx. \end{aligned}$$

As in  $I_1$  and using (5.32), we bound  $I_3$  as

$$|I_3| \leq C \left( \frac{\varepsilon}{\delta} + \delta \right) \left( |\nabla(V_\ell - W_\ell)| + |(V - W)(x_\ell)| (|\nabla V_\ell| + |\nabla W_\ell|) \right) |\nabla Z_\ell|.$$

Adding these up, we get

$$\begin{aligned} |E(V, Z) - E(W, Z)| &\leq C \left( \frac{\varepsilon}{\delta} + \delta + \left( \frac{\varepsilon}{\delta} \right)^{1/2} \right) \left( |\nabla(V_\ell - W_\ell)| \right. \\ (5.35) \quad &\quad \left. + |(V - W)(x_\ell)| (|\nabla V_\ell| + |\nabla W_\ell|) \right) |\nabla Z_\ell|. \end{aligned}$$

Consequently we obtain

$$\begin{aligned} |K| \sum_{x_\ell \in K} \omega_\ell |E(V, Z) - E(W, Z)| &\leq C \left( \delta + \left( \frac{\varepsilon}{\delta} \right)^{1/2} \right) \|\nabla(V - W)\|_{0,K} \|\nabla Z\|_{0,K} \\ &\quad + C \left( \delta + \left( \frac{\varepsilon}{\delta} \right)^{1/2} \right) \|V - W\|_{L^\infty(K)} (\|\nabla V\|_{0,K} + \|\nabla W\|_{0,K}) \|\nabla Z\|_{0,K}. \end{aligned}$$

Using the inverse inequality (1.21) on each element, we obtain

$$\begin{aligned} \|V - W\|_{L^\infty(K)} \|\nabla V\|_{0,K} &\leq C H_K^{-d/2} \|V - W\|_{0,K} H_K^{d/2} \|\nabla V\|_{L^\infty(K)} \\ &= C \|V - W\|_{0,K} \|\nabla V\|_{L^\infty(K)}. \end{aligned}$$

Similarly  $\|V - W\|_{L^\infty(K)} \|\nabla W\|_{0,K} \leq C \|V - W\|_{0,K} \|\nabla W\|_{L^\infty(K)}$ . A combination of the above estimates gives (5.13) with  $\eta(M) = C(M)(\sqrt{\varepsilon/\delta} + \delta)$ .  $\square$

*Proof of the first case in Theorem 5.5.* Let  $W = 0$  in (5.35) and note that  $V \in W^{1,\infty}(D)$ . We obtain (5.27). Therefore if  $(\sqrt{\varepsilon/\delta} + \delta)|\ln H|$  is sufficiently small, then  $e(\text{HMM})^{1/2}|\ln H|$  can be smaller than any given threshold; this verifies (5.10).

Next let  $\eta(M) = C(M)(\sqrt{\varepsilon/\delta} + \delta)$ . Then if  $(\sqrt{\varepsilon/\delta} + \delta)|\ln H|$  is sufficiently small, we have  $\eta(M) < 1$ ; this verifies (5.13).  $\square$

*Remark 5.10.* Compared to the linear case, the upper bound for  $e(\text{HMM})$  for the case when  $\delta/\varepsilon \notin \mathbb{Z}$  degrades to  $\sqrt{\varepsilon/\delta}$ . This is due to the fact that  $A_H$  is nonsymmetric.

In the case of  $\delta = \varepsilon$ , note that  $\hat{v}_\ell^\varepsilon = \hat{V}_\ell$  and  $\hat{w}_\ell^\varepsilon = \hat{W}_\ell$ . So a direct calculation gives Lemma 5.6 for this case. Lemma 5.7 is also valid with  $\delta$  replaced by  $\varepsilon$ . We also have  $\theta_\ell^\varepsilon = 0$  and  $\tilde{\theta}_\ell^\varepsilon = 0$ . Observing that for any  $V, Z \in X_H$ ,

$$\int_{I_\varepsilon(x_\ell)} \nabla z_\ell^\varepsilon \cdot a_V(x_\ell) \nabla \hat{v}_\ell^\varepsilon dx = \int_K \nabla Z_\ell \cdot \mathcal{A}(x_\ell, V(x_\ell)) \nabla V_\ell dx,$$

we may rewrite  $E(V, Z)$  as

$$E(V, Z) = \int_{I_\varepsilon} \nabla z_\ell^\varepsilon \cdot a_V(x) \nabla (v_\ell^\varepsilon - \hat{v}_\ell^\varepsilon) dx + \int_{I_\varepsilon} \nabla z_\ell^\varepsilon \cdot (a_V(x) - a_V(x_\ell)) \nabla \hat{v}_\ell^\varepsilon dx.$$

Consequently, (5.13) holds with  $\eta(M) = C(M)\varepsilon$ .

*Proof of the second case in Theorem 5.5.* Let  $W = 0$  in (5.35) and note that  $V \in W^{1,\infty}(D)$ . We obtain (5.28). Therefore if  $\varepsilon|\ln H|$  is sufficiently small, then  $e(\text{HMM})^{1/2}|\ln H|$  can be smaller than any given threshold; this verifies (5.10).

Next let  $\eta(M) = C(M)\varepsilon$ . Then if  $\varepsilon|\ln H|$  is sufficiently small, we have  $\eta(M) < 1$ . This proves (5.13).  $\square$

#### APPENDIX A. ESTIMATING $e(\text{HMM})$ FOR PROBLEMS WITH RANDOM COEFFICIENTS

Here we estimate  $e(\text{HMM})$  for the random case. Our basic strategy follows that of [43].

Denote a probability space by  $(\Omega, \mathcal{F}, P)$  and let  $a(y, \omega) = (a_{ij}(y, \omega))$  be a random field whose statistics are invariant under integer shifts and which satisfies the uniform ellipticity condition that there exist constants  $\lambda$  and  $\Lambda$  such that

$$\lambda|\xi|^2 \leq a_{ij}(y, \omega)\xi_i\xi_j \leq \Lambda|\xi|^2,$$

for almost all  $y \in \mathbb{R}^d$  and  $\omega \in \Omega$ . For  $j = 1, \dots, d$ , denote by  $\varphi_j(y, \omega)$  the solutions of the cell problem:

$$(A.1) \quad \mathcal{L}_y \varphi_j := -\operatorname{div}_y (a_{ij}(y, \omega) \nabla_y \varphi_j) = \operatorname{div}_y (a_{ij} \cdot e_j),$$

where  $\{e_j\}_{j=1}^d$  are the standard basis in  $\mathbb{R}^d$ .  $\nabla \varphi_j$  is required to be stationary under integer shift. The existence of  $\varphi_j$  is proved in [28, 36]. In general  $\varphi_j$  is not stationary. The homogenized coefficient  $\mathcal{A}$  [28, 36] is given by

$$\mathcal{A} = \langle a(I + \nabla \varphi) \rangle.$$

Here and in the following, we use the notation

$$\langle f \rangle = \mathbb{E} \int_{[0,1]^d} f(y) dy$$

and

$$[f; m] = \frac{1}{m^d} \int_{[0,m]^d} f(y) dy,$$

where  $\mathbb{E}$  denotes the expectation in the probability space  $(\Omega, \mathcal{F}, P)$ .

As in [43], we will consider the following auxiliary problem:

$$(A.2) \quad \mathcal{L}u + \rho u = f,$$

for any  $\rho > 0$ , where  $f$  is of the form

$$f = \sum_{j=1}^d D_j g_j + h,$$

with  $g_j, h \in \rho_G$  which is defined as

$$\rho_G := \{\psi \mid \langle |\psi|^2 \rangle \leq G^2\},$$

and  $\psi$  is a random field whose statistics are stationary with respect to integer shifts.

The solution of (A.2) can be expressed with the help of a diffusion process  $\eta$  generated by the operator  $\mathcal{L}$ .

For each fixed realization of  $\{a(y, \cdot)\}$ , denote by  $\eta_x$  the diffusion process generated by  $\mathcal{L}_1$  and starting from  $x$  at  $t = 0$ , and denote by  $M_x$  the expectation with respect to  $\eta_x$ . Let

$$\Gamma(\tau) = \int_0^\tau f(\eta(t))e^{-\rho t} dt.$$

Then it is well known [21] that the solution of (A.2) is given by

$$u(x) = M_x \Gamma(\infty).$$

The following results are either standard or proved in [28, 43].

**Lemma A.1.** *If  $u$  is the solution of (A.2), then there exists a constant  $C$  independent of  $\rho$  such that*

$$(A.3) \quad \langle |\nabla u|^2 \rangle + \rho \langle u^2 \rangle \leq C \left( \langle g^2 \rangle + \frac{1}{\rho} \langle h^2 \rangle \right),$$

$$(A.4) \quad \langle (M_x \Gamma(\infty))^2 \rangle^{1/2} \leq C \frac{G^2}{\rho},$$

$$(A.5) \quad \langle M_x (\Gamma(\infty) - \Gamma(s))^2 \rangle \leq C \frac{G^2}{\rho} e^{-2s\rho}.$$

Because of the lowest order term  $\rho u$ , the Green's function associated with the operator  $\mathcal{L}_1 + \rho I$  decays exponentially with rate  $\mathcal{O}(\sqrt{\rho})$ . To make this statement precise, we define the norm  $\|x\| := \max_i |x_i|$ , and

$$Q_\rho := \left\{ x \in \mathbb{R}^d \mid \|x\| \leq \left( \frac{1}{\rho} \right)^{\frac{1}{2} [\ln(1/\rho)]^{1/2}} \right\}.$$

Let  $\tau$  be the first exit time of  $Q_\rho$  starting at  $x \in Q_\rho$ . Let  $\hat{\varphi}_\rho(x) = M_x \Gamma(\tau)$ .

**Lemma A.2** ([43]). *If  $\rho$  is sufficiently small, then*

$$(A.6) \quad \mathbb{E} \int_{\|x\| \leq 10} |\varphi_\rho(x) - \hat{\varphi}_\rho(x)|^2 dx \leq C G^2 e^{-C |\ln(1/\rho)|^2},$$

$$(A.7) \quad \mathbb{E} \int_{\|x\| \leq 1} |\nabla \varphi_\rho(x) - \nabla \hat{\varphi}_\rho(x)|^2 dx \leq C G^2 e^{-C |\ln(1/\rho)|^2}.$$

To prepare for the discussion on the consequence of the mixing condition, we mention

**Lemma A.3** ([43]). *Let  $\{a_{ij}, g_j\}$  and  $\{\tilde{a}_{ij}, \tilde{g}_j\}$  be two sets of data such that*

$$\{a_{ij}(y), g_j(y)\} = \{\tilde{a}_{ij}(y), \tilde{g}_j(y)\}$$

for  $y \notin B$ , where  $B$  is a domain in  $\mathbb{R}^d$ , and let  $\varphi_\rho$  and  $\hat{\varphi}_\rho$  be the solutions of (A.2) associated with  $\{a_{ij}, g_j\}$  and  $\{\tilde{a}_{ij}, \tilde{g}_j\}$ , respectively (with  $h = 0$ ). Then

$$(A.8) \quad \int_{\mathbb{R}^d} |\varphi_\rho(x) - \tilde{\varphi}_\rho(x)|^2 dx \leq \frac{C}{\rho} \int_{\mathbb{R}^d} (G^2 + |\nabla \varphi_\rho(x)|^2) I_B(x) dx,$$

where  $I_B$  is the indicator function of the domain  $B$ .

Now we introduce the crucial mixing condition [22]. Let  $B$  be a domain in  $\mathbb{R}^d$ . Denote by  $\mathcal{F}(B)$  the  $\sigma$ -algebra generated by  $\{a_{ij}(y, \omega), y \in B\}$ . Let  $\xi, \eta$  be two random variables that are measurable with respect to  $\mathcal{F}(B_1)$  and  $\mathcal{F}(B_2)$ , respectively. Then

$$(A) \quad \frac{|\mathbb{E}\xi\eta - \mathbb{E}\xi\mathbb{E}\eta|}{(\mathbb{E}\xi^2)^{1/2}(\mathbb{E}\eta^2)^{1/2}} \leq e^{-\lambda q},$$

where  $q = \text{dist}(B_1, B_2)$ ,  $\lambda > 0$  is a fixed constant.

**Lemma A.4.** *Under condition (A), we have*

$$(A.9) \quad \mathbb{E}[\varphi_\rho; m]^2 \leq C \left( \frac{G^2}{\rho} \left( \frac{|\ln(1/\rho)|^2}{\rho^{1/2}m} \right)^d + e^{-c(\ln(1/\rho))^2} \right).$$

*Proof.* For  $\ell = (\ell_1, \dots, \ell_d) \in \mathbb{Z}^d \cap [0, m]^d := \mathbb{Z}_m^d$ , denote by  $I^\ell$  the cube of size 1 centered at  $\ell + \frac{1}{2} = (\ell_1 + \frac{1}{2}, \dots, \ell_d + \frac{1}{2})$ , and let  $\varphi^\ell = \int_{I^\ell} \varphi_\rho(x) dx$ . Then

$$[\varphi_\rho; m] = \frac{1}{m^d} \sum_{\ell \in \mathbb{Z}_m^d} \varphi^\ell.$$

We first estimate  $\mathbb{E}(\varphi^\ell \varphi^k)$ . If  $|\ell - k| \leq C\rho^{-1/2}|\ln(1/\rho)|^2$ , then

$$\mathbb{E}(\varphi^\ell \varphi^k) \leq \langle \varphi_\rho^2(x) \rangle^{1/2} \langle \varphi_\rho^2(x) \rangle^{1/2} \leq C \frac{G^2}{\rho}.$$

If  $|\ell - k| \geq C\rho^{-1/2}|\ln(1/\rho)|^2$ , then let  $B_1 := \ell + B(\rho^{-1/2}|\ln(1/\rho)|^2)$  and  $B_2 := k + B(\rho^{-1/2}|\ln(1/\rho)|^2)$ , where  $B(s)$  is a ball of size  $s$  in the norm  $\|\cdot\|$ . Denote by  $\tilde{\varphi}_1(x)$  the solution of (A.2) in which the coefficient  $(a_{ij}(y, \omega))$  is modified in  $\mathbb{R}^d \setminus B_1$  such that it is measurable with respect to  $\mathcal{F}(B_1)$ , and similarly denote by  $\tilde{\varphi}_2(x)$  the solution of (A.2) in which the coefficient  $(a_{ij}(y, \omega))$  is modified in  $\mathbb{R}^d \setminus B_2$  such that it is measurable with respect to  $\mathcal{F}(B_2)$ . The modified coefficients  $(\tilde{a}_{ij}(y, \omega))$  should still satisfy the condition on  $a_{ij}$  listed in the beginning of this subsection. From  $\tilde{\varphi}_1$  and  $\tilde{\varphi}_2$ , we can similarly define  $\tilde{\varphi}_1^\ell$  and  $\tilde{\varphi}_2^k$ . Using (A.6), we have

$$\begin{aligned} \mathbb{E}(\tilde{\varphi}_1^\ell - \tilde{\varphi}^\ell)^2 &\leq CG^2 e^{-C|\ln(1/\rho)|^2}, \\ \mathbb{E}(\tilde{\varphi}_1^k - \tilde{\varphi}^k)^2 &\leq CG^2 e^{-C|\ln(1/\rho)|^2}. \end{aligned}$$

Since

$$\begin{aligned} \mathbb{E}(\varphi^\ell \varphi^k) &= \mathbb{E}(\tilde{\varphi}^\ell \tilde{\varphi}^k) + \mathbb{E}(\varphi^\ell - \tilde{\varphi}_1^\ell) \tilde{\varphi}_2^k + \mathbb{E} \tilde{\varphi}_1^\ell (\varphi^k - \tilde{\varphi}_2^k) \\ &\quad + \mathbb{E}(\varphi^\ell - \tilde{\varphi}_1^\ell) (\varphi^k - \tilde{\varphi}_2^k) \end{aligned}$$

and

$$\begin{aligned} |\mathbb{E} \tilde{\varphi}_1^\ell|, |\mathbb{E} \tilde{\varphi}_2^k| &\leq CGe^{-C|\ln(1/\rho)|^2}, \\ |\mathbb{E} \tilde{\varphi}_1^\ell \tilde{\varphi}_2^k| &\leq C \left( \frac{G^2}{\rho} e^{-C|\ell-k|} + G^2 e^{-C|\ln(1/\rho)|^2} \right), \end{aligned}$$

we thus have

$$|\mathbb{E} \varphi^\ell \varphi^k| \leq C \left( \frac{G^2}{\rho} e^{-C|\ell-k|} + G^2 e^{-C|\ln(1/\rho)|^2} \right).$$

Hence

$$\begin{aligned}
\mathbb{E}[\varphi_\rho; m]^2 &= \frac{1}{m^{2d}} \sum_{\ell, k \in \mathbb{Z}_m^d} \mathbb{E} \varphi^\ell \varphi^k \\
&= \frac{1}{m^{2d}} \left( \sum_{|\ell-k| \leq \rho^{-1/2} |\ln(1/\rho)|^2} \mathbb{E} \varphi^\ell \varphi^k + \sum_{|\ell-k| > \rho^{-1/2} |\ln(1/\rho)|^2} \mathbb{E} \varphi^\ell \varphi^k \right) \\
&\leq \frac{1}{m^{2d}} \left( \frac{G^2}{\rho} m^d \left( \rho^{-1/2} |\ln(1/\rho)|^2 \right)^d + m^{2d} G^2 e^{-C |\ln(1/\rho)|^2} \right) \\
&\quad \times \frac{G^2}{\rho} \sum_{|\ell-k| \geq \rho^{-1/2} |\ln(1/\rho)|^2} e^{-C |\ell-k|} \\
&\leq C \frac{G^2}{\rho} \left( \frac{\ln(1/\rho)}{\rho^{1/2} m} \right)^d + e^{-C |\ln(1/\rho)|^2}.
\end{aligned}$$

□

Proceeding along the same line as in [43, Theorem 2.1], using condition **(A)**, we have

**Lemma A.5.** *For any  $0 < \gamma < 1/2$ , under condition **(A)**, there exists a constant  $C$  such that*

$$(A.10) \quad |\mathcal{A} - \langle a(I + \nabla \chi_\rho) \rangle| \leq C \rho^{\frac{d-2-2\gamma}{4+d}},$$

where  $\chi_\rho = \{\chi_{k,\rho}\}_{k=1}^d$ , and  $\chi_{k,\rho}$  is the solution of (A.1) with  $g = (a_{k1}, \dots, a_{kd})$ .

Now we are ready to estimate  $e(\text{HMM})$ . Define  $m = \frac{\delta}{2\varepsilon}$  and denote by  $\varphi_j^m$  the solution of (A.1) on  $I_m = [0, m]^d$  with the boundary condition that  $\varphi_j^m(y) = 0$  on  $\partial I_m$ , and let  $\varphi^m = (\varphi_1^m, \dots, \varphi_d^m)$ . Then

$$e(\text{HMM}) = |\mathcal{A} - [a(I + \nabla \varphi^m)]_m|.$$

Define  $\varphi_\rho, \varphi_\rho^m$  similarly. We have

$$e(\text{HMM}) \leq E_1 + E_2 + E_3,$$

where

$$\begin{aligned}
E_1 &= |\mathcal{A} - [a(I + \nabla \varphi_\rho)]_m|, \\
E_2 &= |[a(I + \nabla \varphi_\rho) - a(I + \nabla \varphi_\rho^m)]_m|, \\
E_3 &= |[a \nabla (\varphi_\rho^m - \varphi^m)]_m|.
\end{aligned}$$

Obviously,

$$E_1 = |\mathcal{A} - \langle a(I + \nabla \varphi_\rho) \rangle + [\tilde{\psi}]_m|,$$

with  $\tilde{\psi} = \langle a(I + \nabla \varphi_\rho) \rangle - a(I + \nabla \varphi_\rho)$ . It was proved in [43, Lemma 2.5] that

$$\mathbb{E} |[\tilde{\psi}]_m| \leq (\mathbb{E} [\tilde{\psi}_m^2])^{1/2} \leq C \left( \frac{|\ln(1/\rho)|^2}{\rho^{1/2} m} \right)^{d/2}.$$

The above inequality together with Lemma A.5 gives

$$\mathbb{E} E_1 \leq C \left( G \rho^{\frac{d-2-2\gamma}{4+d}} + \left( \frac{|\ln(1/\rho)|^2}{\rho^{1/2} m} \right)^{d/2} \right).$$



To estimate  $E_2$ , denote by  $\tau_m$  the first exit time for the domain  $I_{2m}$ . Then  $\varphi_\rho^{2m} = M_x \Gamma(\tau_m)$ . For any  $s > 0$ ,

$$\begin{aligned} |\varphi_\rho - \varphi_\rho^{2m}| &= |M_x(\Gamma(\infty) - \Gamma(\tau_m))| \\ &\leq M_x\{|\Gamma(\infty)| + |\Gamma(\tau_m)|; \tau_m \leq s\} \\ &\quad + M_x\{e^{-s\rho} M_{\rho(s)} |\Gamma(\infty) - \Gamma(\tau_m)|; \tau_m > s\} \\ &\leq C \left( M_x((\Gamma(\infty))^2 + \Gamma(\tau_m)^2) \right)^{1/2} \{P_x(\tau_m \leq s)\}^{1/2} + e^{-s\rho}. \end{aligned}$$

Since

$$P_x\{\tau_m \leq s\} \leq e^{-Cm^2/s},$$

we get

$$\mathbb{E}[|\varphi_\rho - \varphi_\rho^{2m}|^2]_{2m} \leq C \frac{G^2}{\rho} \left( e^{-Cm^2/s} + e^{-s\rho} \right)^2.$$

Optimizing in  $s$ , we get

$$\mathbb{E}[|\varphi_\rho - \varphi_\rho^{2m}|^2]_{2m} \leq C \frac{G^4}{\rho^2} e^{-Cm\rho^{1/2}}.$$

Using standard interior estimates, we have

$$\mathbb{E} E_2 \leq \left( \mathbb{E}[|\nabla(\varphi_\rho - \varphi_\rho^m)|^2]_m \right)^{1/2} \leq \frac{C}{m} \left( \mathbb{E}[|\varphi_\rho - \varphi_\rho^{2m}|^2]_{2m} \right)^{1/2} \leq \frac{CG^2}{m\rho} e^{-Cm\rho^{1/2}}.$$

As for  $E_3$ , proceeding along the same line as in the estimate of  $E_1$ , we get

$$\mathbb{E} E_3 \leq C \left( G^2 \rho^{\frac{d-2-2\gamma}{4+d}} + \left( \frac{|\ln(1/\rho)|^2}{\rho^{1/2}m} \right)^{d/2} \right).$$

To sum up, we have

$$\mathbb{E} e(\text{HMM}) \leq C \left( G^2 \rho^{\frac{d-2-2\gamma}{4+d}} + \left( \frac{|\ln(1/\rho)|^2}{\rho^{1/2}m} \right)^{d/2} + \frac{CG^2}{m\rho} e^{-Cm\rho^{1/2}} \right).$$

Optimizing in  $\rho$  with respect to the first two terms, we get  $\rho_0 = m^{-\frac{2d}{d+4\alpha}}$  with  $\alpha = (d-2-2\gamma)/(d+4)$ . Hence

$$(A.11) \quad \mathbb{E} e(\text{HMM}) \leq C \left( \frac{|\ln m|^d}{m^\kappa} + \frac{CG^2}{m\rho_0} e^{-Cm\rho_0^{1/2}} \right) \leq C \frac{|\ln m|^d}{m^\kappa},$$

with

$$\kappa = \frac{d/2}{1 + \frac{d(d+4)/4}{d-2-2\gamma}}, \quad 0 < \gamma < \frac{1}{2}.$$

Obviously, the  $|\ln m|^d$  factor in (A.11) can be absorbed into the factor  $m^{-\kappa}$ . This proves Theorem 1.3 for  $d = 3$ .

*Remark A.6.* The estimate (A.11) is unlikely to be optimal. If  $d = 1$ , a direct calculation gives

$$\mathbb{E} e(\text{HMM}) \leq C \left( \frac{\varepsilon}{\delta} \right)^{1/2},$$

whereas the estimate (A.11) does not apply for  $d = 1$  and  $d = 2$ . We may use the techniques in [13] to derive improved bounds for  $e(\text{HMM})$  if the magnitude of the oscillation in the coefficients  $(a_{ij}^\varepsilon)$  is sufficiently small.

## REFERENCES

1. R.A. Adams and J.J. F. Fournier, *Sobolev Spaces*, second edition, Academic Press, New York, 2003.
2. G. Allaire, *Homogenization and two-scale convergence*, SIAM J. Math. Anal., **23**(1992), 1482–1518. MR1185639 (93k:35022)
3. I. Babuska, *Homogenization and its applications, mathematical and computational problems*, Numerical Solutions of Partial Differential Equations-III (SYNSPADE 1975, College Park, MD, May 1975) (B. Hubbard ed.), Academic Press, New York, 1976, pp. 89–116. MR0502025 (58:19215)
4. I. Babuska, *Solution of interface by homogenization*. I, II, III, SIAM J. Math. Anal., **7** (1976), 603–634, 635–645, **8** (1977), 923–937. MR0509273 (58:23013a); MR0509277 (58:23013b); MR0509282 (58:23014)
5. A. Bensoussan, J.L. Lions and G.C. Papanicolaou, *Asymptotic Analysis for Periodic Structures*, North-Holland, Amsterdam, 1978. MR0503330 (82h:35001)
6. L. Boccardo and F. Murat, *Homogénéisation de problèmes quasi-linéaires*, Publ. IRMA, Lille., **3** (1981), no. 7, 13–51.
7. J.F. Bourgat, *Numerical experiments of the homogenization method for operators with periodic coefficients*, Lecture Notes in Mathematics, Vol. 707. Springer-Verlag, 1977, pp. 330–356. MR0540121 (80e:65108)
8. A. Brandt, *Multi-level adaptive solutions to boundary-value problems*, Math. Comp., **31** (1977), 333–390. MR0431719 (55:4714)
9. S.C. Brenner and L.R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, New York, 1994. MR1278258 (95f:65001)
10. P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978. MR0520174 (58:25001)
11. P.G. Ciarlet and P.-A. Raviart, *The combined effect of curved boundaries and numerical integration in isoparametric finite element methods*, The Mathematics Foundations of the Finite Element Method with Applications to Partial Differential Equations (A. K. Aziz ed.), Academic Press, Inc., 1972, pp. 409–474. MR0421108 (54:9113)
12. P. Clément, *Approximation by finite element functions using local regularization*, RAIRO., **9** (1975), 77–84. MR0400739 (53:4569)
13. J.G. Conlon and A. Naddaf, *On homogenization of elliptic equations with random coefficients*, Electron J. Probab., **5** (2000), 1–58. MR1768843 (2002j:35328)
14. L.J. Durlofsky, *Numerical calculation of equivalent grid block permeability tensors for heterogeneous poros-media*, Water. Resour. Res., **28** (1992), 699–708.
15. W. E, *Homogenization of linear and nonlinear transport equations*, Comm. Pure Appl. Math., **45** (1992), 301–326. MR1151269 (92k:35026)
16. W. E and B. Engquist, *The heterogeneous multiscale methods*, Comm. Math. Sci., **1**(2003), 87–132. MR1979846 (2004b:35019)
17. W. E and B. Engquist, *The heterogeneous multiscale method for homogenization problems*, submitted to MMS, 2002.
18. W. E and B. Engquist, *Multiscale modeling and computation*, Notice Amer. Math. Soc., **50** (2003), 1062–1070. MR2002752
19. W. E and X.Y. Yue, *Heterogeneous multiscale method for locally self-similar problems*, Comm. Math. Sci., **2** (2004), 137–144.
20. Y.R. Efendiev, T. Hou and X. Wu, *Convergence of a nonconforming multiscale finite element method*, SIAM J. Numer. Anal., **37** (2000), 888–910. MR1740386 (2002a:65176)
21. M.I. Freidlin, *Functional Integration and Partial Differential Equations*, Princeton, NJ, 1985. MR0833742 (87g:60066)

22. M.I. Freidlin and A.D. Wentzell, *Random Perturbations of Dynamical Systems*, second edition, Springer-Verlag, 1998. MR1652127 (99h:60128)
23. N. Fusco and G. MoscarIELlo, *On the homogenization of quasilinear divergence structure operators*, Ann. Mat. Pura. Appl., **146** (1987), 1–13. MR0916685 (89a:35025)
24. D. Gilbarg and N.S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, second edition, Springer-Verlag, New York, 1983. MR0737190 (86c:35035)
25. T. Hou and X. Wu, *A multiscale finite element method for elliptic problems in composite materials and porous media*, J. Comput. Phys., **134** (1997), 169–189. MR1455261 (98e:73132)
26. I.G. Kevrekidis, C.W. Gear, J.M. Hyman, P.G. Kevrekidis, O. Runborg and C. Theodoropoulos, *Equation-free, coarse-grained multiscale computation: enabling microscopic simulators to perform system-level analysis*, Comm. Math. Sci., **1** (2003), 715–762. MR2041455
27. J. Knap and M. Ortiz, *An analysis of the quasicontinuum method*, J. Mech. Phys. Solids., **49** (2001), 1899–1923.
28. S.M. Kozlov, *Homogenization of random operators*, Math. USSR Sb., **37** (1980), 167–180. MR0542557 (81m:35142)
29. A.M. Matache, I. Babuska and C. Schwab, *Generalized  $p$ -FEM in homogenization*, Numer. Math., **86** (2000), 319–375. MR1777492 (2001f:65139)
30. G.W. Milton, *The Theory of Composites*, Cambridge University Press, 2002. MR1899805 (2003d:74077)
31. P.B. Ming and X.Y. Yue, *Numerical methods for multiscale elliptic problems*, preprint, 2003.
32. S. Moskow and M. Vogelius, *First order corrections to the homogenized eigenvalues of a periodic composite medium*, Proc. Royal Soc. Edinburgh. Sec. A., **127** (1997), 1263–1299. MR1489436 (99g:35018)
33. F. Murat and L. Tartar, *H-Convergence*, Cours Peccot, 1977. Reprinted in Topics in the Mathematical Modeling of Composite Materials (A. Cherkaev and R. Kohn eds.), Birkhäuser, Boston, 1997, pp. 21–43. MR1493039
34. G. Nguetseng, *A general convergence result for a functional related to the theory of homogenization*, SIAM J. Math. Anal., **20** (1989), 608–623. MR0990867 (90j:35030)
35. J.T. Oden and K.S. Vemaganti, *Estimation of local modeling error and goal-oriented adaptive modeling of heterogeneous materials. I. Error estimates and adaptive algorithms*, J. Comput. Phys., **164** (2000), 22–47. MR1786241 (2001e:74079)
36. G.C. Papanicolaou and S.R. S. Varadhan, *Boundary value problems with rapidly oscillating random coefficients*, at Proceedings of the conference on Random Fields, Esztergom, Hungary, 1979, published in Seria Colloquia Mathematica Societatis Janos Bolyai, North-Holland, **27** (1981), pp. 835–873. MR0712714 (84k:58233)
37. R. Rannacher and R. Scott, *Some optimal error estimates for piecewise linear finite element approximations*, Math. Comp., **38** (1982), 437–445. MR0645661 (83e:65180)
38. C. Schwab and A.-M. Matache, *Generalized FEM for homogenization problems*, Multiscale and Multiresolution Methods: Theory and Applications, Lecture Notes in Computational Sciences and Engineering, Vol. 20 (T. J. Barth, T. Chan and R. Heimes eds.), Springer-Verlag, Heidelberg, 2002, pp. 197–237. MR1928567 (2003i:65115)
39. R. Scott, *Optimal  $L^\infty$  estimates for the finite element method on irregular meshes*, Math. Comp., **30** (1976), 681–697. MR0436617 (55:9560)
40. S. Spagnolo, *Convergence in energy for elliptic operators*, Numerical Solutions of Partial Differential Equations-III (SYNSPADE 1975, College Park, MD, May 1975) (B. Hubbard ed.), Academic Press, New York, 1976, pp. 469–499. MR0477444 (57:16971)
41. L. Tartar, *An introduction to the homogenization method in optimal design*, Optimal shape Design (A. Cellina and A. Ornelas eds.), Lecture Notes in Mathematics, Vol. 1740. Springer-Verlag, 2000. pp. 47–156. MR1804685
42. J. Xu, *Two-grid discretization techniques for linear and nonlinear PDE*, SIAM J. Numer. Anal., **33** (1996), 1759–1777. MR1411848 (97i:65169)
43. V.V. Yurinskii, *Averaging of symmetric diffusion in random media*, Sibirsk. Mat. Zh. **23** (1982), no. 2, 176–188; English transl. in Siberian Math J., **27** (1986), 603–613. MR0867870 (88e:35190)
44. V.V. Zhikov, *On an extension of the method of two-scale convergence and its application*, Sbornik: Mathematics, **191** (2000), 973–1014. MR1809928 (2001k:35026)
45. V.V. Zhikov, S.M. Kozlov and O.A. Oleinik, *Homogenization of Differential Operators and Integral Functionals*, Springer-Verlag, Heidelberg, 1994. MR1329546 (96h:35003b)

DEPARTMENT OF MATHEMATICS AND PACM, PRINCETON UNIVERSITY, PRINCETON, NEW JERSEY 08544 AND SCHOOL OF MATHEMATICAL SCIENCES, PEKING UNIVERSITY, BEIJING 100871, PEOPLE'S REPUBLIC OF CHINA

*E-mail address:* `weinan@Princeton.EDU`

NO. 55, ZHONG-GUAN-CUN EAST ROAD, INSTITUTE OF COMPUTATIONAL MATHEMATICS, ACADEMY OF MATHEMATICS AND SYSTEMS SCIENCE, CHINESE ACADEMY OF SCIENCES, BEIJING 100080, PEOPLE'S REPUBLIC OF CHINA

*E-mail address:* `mpb@lsec.cc.ac.cn`

SCHOOL OF MATHEMATICAL SCIENCES, PEKING UNIVERSITY, BEIJING 100871, PEOPLE'S REPUBLIC OF CHINA

*E-mail address:* `pzhang@pku.edu.cn`