ANALYSIS OF THE HETEROGENEOUS MULTISCALE METHOD FOR PARABOLIC HOMOGENIZATION PROBLEMS

PINGBING MING AND PINGWEN ZHANG

Abstract. The heterogeneous multiscale method (HMM) is applied to various parabolic problems with multiscale coefficients. These problems can be either linear or nonlinear. Optimal estimates are proved for the error between the HMM solution and the homogenized solution.

1. Introduction and Main Results

1.1. **Generality.** Consider the following parabolic problem:

$$\begin{cases} \partial_t u^\varepsilon - \nabla\cdot(a^\varepsilon \nabla u^\varepsilon) = f & \text{in } D\times(0,T) =: \mathcal{Q},\\ u^\varepsilon = 0 & \text{on } \partial D\times(0,T),\\ u^\varepsilon|_{t=0} = u_0. \end{cases} \tag{1.1}$$

Here ε is a small parameter that signifies the multiscale nature of the problem. We let D be a bounded domain in $\mathbb{R}^d$ and T a positive real number. Problem of this type is interesting because of its simplicity, its relevance to several important practical problems such as the flow in porous media and the mechanical properties of composite materials. In contrast to the elliptic problems, there may be oscillations in the temporal direction besides the oscillation in the spatial direction.

On the analytic side, the following fact is known about (1.1). In the sense of parabolic H-convergence (see [25], [8], [12]), introduced with minor modification by Spagnolo and Colombini under the name of G-convergence or PG-convergence (see [11], [22], [23], [24]), that for every $f\in L^2(0,T;H^{-1}(D))$ and $u_0\in L^2(D)$, the sequence $\{u^\varepsilon\}$ the solutions of (1.1) satisfies

$$\begin{aligned} u^\varepsilon &\rightharpoonup U && \text{weakly in } L^2(0,T;H_0^1(D)),\\ a^\varepsilon\nabla u^\varepsilon &\rightharpoonup \mathcal{A}\nabla U && \text{weakly in } L^2(\mathcal{Q};\mathbb{R}^d), \end{aligned}$$

where U is the unique solution of the problem

$$\begin{cases} \partial_t U - \nabla\cdot(\mathcal{A}\nabla U) = f & \text{in } \mathcal{Q},\\ U = 0 & \text{on } \partial D\times(0,T],\\ U|_{t=0} = u_0. \end{cases} \tag{1.2}$$

Date: March 25, 2006.

1991 *Mathematics Subject Classification.* 65N30, 35K05, 65N15.

Key words and phrases. Heterogeneous Multiscale Method, Parabolic Homogenization Problems, Finite Element Methods.

Pingbing Ming is partially supported by the National Natural Science Foundation of China under the grant 10571172 and also supported by the National Basic Research Program under the grant 2005CB321704.

Pingwen Zhang is partially supported by National Natural Science Foundation of China for Distinguished Young Scholars 10225103 and also supported by the National Basic Research Program under the grant 2005CB321704.

We thank the anonymous referees for the improvements of this paper. In particular Corollary A.3 and Corollary A.4.

In general, there is no explicit formulas for the effective matrix $\mathcal{A}$.

Classical numerical methods for this problem are designed to resolve the full details of the fine scale problem (1.1) and without taking into account the special features of the coefficient matrix a^ε. In contrast, the modern multiscale methods are designed specifically for retrieving partial information about u^ε with sublinear cost [16], i.e. the total cost grows sublinearly with the cost of solving the full fine scale problem. To this end, the methods have to take the full advantage of the special features of the problem such as scale separation, self-similarity of the solution. One cannot hope to get an algorithm with sublinear cost for a fully general problem.

The heterogeneous multiscale method introduced in [15] is a general methodology for designing sublinear algorithm by exploiting the scale separation and other special features of the problem. HMM consists of two ingredients: An overall macroscopic scheme for macro variables on a macro grid and estimating the missing macroscopic data from the microscopic model. The efficiency of HMM lies in the ability to extract the missing macroscale data from microscale models with minimum cost, by exploiting scale separation.

For (1.1), the macroscopic solver is chosen to be the standard piecewise linear finite element method [10] over a macroscopic triangulation $\mathcal{T}_H$ with mesh size H as the spatial solver, and the backward Euler scheme as the temporal discretization. Many other conventional discretization methods could be proper candidates as the macroscopic solver. For example, finite difference method and Discontinuous Galerkin method have been employed as the macroscopic solver in [1] and [9], respectively.

We formulate our method as follows. For $1 \le k \le n$, let $t_k = k\Delta t$ with $\Delta t = T/n$. Let $U_H^0 = Q_H u_0$ with Q_H the L^2 projection operator from $H_0^1(D)$ to X_H, where X_H is the macroscopic finite element space. Let $U_H^k \in X_H$ be the solution of the problem

$$(\overline{\partial} U_H^k, V) + A_H(t_k; U_H^k, V) = (f^k, V) \qquad \text{for all } V \in X_H, \tag{1.3}$$

where $\overline{\partial} U_H^k = (U_H^k - U_H^{k-1})/\Delta t$ and $f^k = \Delta t^{-1} \int_{t_k}^{t_k+\Delta t} f(x,s)\,\mathrm{d}s$.

It remains to estimate the stiffness matrix, which amounts to evaluating the effective bilinear form $A_H(t_n; V, W)$ for any $V, W \in X_H$. We write A_H as

$$\begin{aligned} A_H(t_n; V, W) &= \int_D \nabla W \cdot \mathcal{A}_H(x, t_n) \nabla V \,\mathrm{d}x = \sum_{K \in \mathcal{T}_H} \int_K \nabla W \cdot \mathcal{A}_H(x, t_n) \nabla V \,\mathrm{d}x \\ &\simeq \sum_{K \in \mathcal{T}_H} |K| \nabla W \cdot \mathcal{A}_H(x_K, t_n) \nabla V, \end{aligned}$$

where x_K is the barycenter of K. We approximate $\mathcal{A}_H(x_K, t_n)$ by solving the Cauchy-Dirichlet problem:

$$\begin{cases} \partial_t v^\varepsilon - \nabla \cdot \left(a^\varepsilon \nabla v^\varepsilon\right) = 0 & \text{in} \quad (x_K + I_\delta) \times (t_n, t_n + \tau_n), \\ v^\varepsilon = V & \text{on} \quad \partial I_\delta \times (t_n, t_n + \tau_n), \\ v^\varepsilon|_{t=t_n} = V. & \end{cases} \tag{1.4}$$

We then let

$$\nabla W \cdot \mathcal{A}_H(x_K, t_n) \nabla V \simeq \frac{1}{\tau_n |I_\delta|} \int_{t_n}^{t_n+\tau_n} \int_{I_\delta} \nabla w^\varepsilon \cdot a^\varepsilon \nabla v^\varepsilon \,\mathrm{d}x \,\mathrm{d}t,$$

where τ_n denotes the micro simulation time evolves in n-th macro time step, and $I_\delta = \delta Y$ with the unit cell $Y := (-1/2, 1/2)^d$. For simplicity, we denote $I_\delta := x_K + I_\delta, T_n := (t_n, t_n + \tau_n)$, and the cylinder $\mathcal{Q}_n :=I_\delta \times T_n$. We thus rewrite A_H as

$$A_H(t_n; V, W) := \sum_{K\in\mathcal{T}_H} \frac{|K|}{|\mathcal{Q}_n|} \int_{\mathcal{Q}_n} \nabla w^\varepsilon \cdot a^\varepsilon \nabla v^\varepsilon \,\mathrm{d}x\,\mathrm{d}t. \tag{1.5}$$

In (1.4), we use the Dirichlet boundary condition and the Cauchy initial condition. One may also use other boundary conditions and initial conditions. For example, we may use Neumann or periodic boundary condition and periodic initial condition. In the case when $a^\varepsilon = a(x, x/\varepsilon, t)$ and $a(x, y, t)$ is periodic in y, one can take I_δ to be $x_K + \varepsilon Y$ and impose the boundary/initial conditions as $v^\varepsilon - V$ is periodic on the boundary of the cylinder $(x_K + \varepsilon Y) \times (t_n, t_n + \varepsilon^2)$.

So far, the algorithm is quite general. The saving compared with solving the full fine scale problem comes from the fact that we may choose I_δ and $\{\tau_k\}$ much smaller than K and Δt, respectively. The size of the micro cell I_δ and the micro simulation time $\{\tau_k\}$ are mainly determined by the accuracy, the cost and the micro structure of a^ε. The main purpose of the error analysis presented below is to help to assess the performance of the method and give a guidance for the designing of the methods, namely, how we choose δ and $\{\tau_k\}$, or types of cell problem.

Since HMM is based on standard macroscale numerical methods and uses the microscale model only as a supplement, it is possible to analyze its stability and accuracy within the traditional framework of numerical analysis. This has already been illustrated in [14, 15, 17] and will be further elaborated in the present paper. Roughly speaking, we will show that HMM is stable whenever the macroscopic solver is stable. The overall error between the HMM solution and the homogenized solution is controlled by the accuracy of the macroscopic solver and the consistency error emanates from the estimate of the macroscopic data from the microscopic model, which will be denoted by $e(\mathrm{HMM})$. Next we estimate $e(\mathrm{HMM})$ for two cases. One is $a^\varepsilon = a(x, x/\varepsilon, t)$ with $a(x, y, t)$ is periodic in y, and the other is $a^\varepsilon = a(x, x/\varepsilon, t, t/\varepsilon^2)$ with $a(x, y, t, s)$ is periodic in y and s.

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

$$\lambda I \le a^\varepsilon \le \Lambda I$$

for some $\lambda, \Lambda > 0$. We will use $|\cdot|$ to denote the abstract value of a scalar quantity and the volume of a set.

Throughout this paper, the generic constant C is assumed to be independent of the microscale ε, the mesh size H, the time step Δt, the cell size δ and the micro simulation time $\{\tau_k\}_{k=1}^n$. We use the summation convention.

1.2. **Main results.** Define

$$e(\mathrm{HMM}) = \max_{1\le k\le n} e_k(\mathrm{HMM}) \tag{1.6}$$

with

$$e_k(\mathrm{HMM}) = \max_{K\in\mathcal{T}_H} \|(\mathcal{A} - \mathcal{A}_H)(x_K, t_k)\|,$$

where $\|\cdot\|$ denotes the Euclidean norm.

Our main results for the linear problem are as follows.

Theorem 1.1. *Let U and U_H^n be the solutions of* (1.2) *and* (1.3)*, respectively. If U is sufficiently smooth, then there exists a constant C that is independent of $\varepsilon, \delta, \{\tau_k\}_{k=1}^n, H, \Delta t$, such that*

$$\|U_H^n - U(x, t_n)\|_0 + |||U_H^n - U(x, t_n)||| \le C\big(\Delta t + H^2 + e(\mathrm{HMM})\big), \tag{1.7}$$

$$\|U_H^n - U(x, t_n)\|_1 \le C\big(\Delta t + H + e(\mathrm{HMM})\Delta t^{-1/2}\big), \tag{1.8}$$

where $|||\cdot|||$ *is the weighted space-time* H^1 *norm that is defined for every* $V = \{V^k\}_{k=1}^n$ *with* $V^k \in X$ *for* $k = 1, \cdots, n$ *as*

$$|||V|||:= \Big(\sum_{k=1}^{n} \Delta t \|\nabla V^k\|_0^2\Big)^{1/2}.$$

At this stage, no assumption on the form of a^ε is necessary. For U_H^n to converge to $U(x, t_n)$, i.e. $e(\mathrm{HMM}) \to 0$. U must be chosen as the solution of the homogenized equation, which we now assume exists. To obtain qualitative estimate for $e(\mathrm{HMM})$, we must make more assumptions on a^ε.

We estimate $e(\mathrm{HMM})$ for two special cases that depend on the estimate of the homogenized problem (1.1) presented in the Appendix. The extension to other cases [2, 28] is beyond this paper since it depends heavily on the qualitative estimates of the corresponding homogenization problem that seem missing presently.

Theorem 1.2. *For* $a^\varepsilon = a(x, x/\varepsilon, t)$ *with* $a(x, y, t)$ *is periodic in* y *with period* Y*, and the cell problem* (1.4) *is solved with Dirichlet boundary condition and Cauchy initial condition, we have*

$$e(\mathrm{HMM}) \le C\Big[\delta + \frac{\varepsilon}{\delta} + \max_{1\le k\le n}\Big(\tau_k + \frac{\varepsilon^2}{\tau_k}\Big)\Big]. \tag{1.9}$$

Another important case for which the estimate of $e(\mathrm{HMM})$ can be obtained is the so called self-similar case, i.e. $a^\varepsilon = a(x, x/\varepsilon, t, t/\varepsilon^2)$. In this case, we have

Theorem 1.3. *For* $a^\varepsilon = a(x, x/\varepsilon, t, t/\varepsilon^2)$ *with* $a(x, y, t, s)$ *is periodic in* y*and* s *with period* Y *and* 1*, respectively, and the cell problem* (1.4) *is solved with Dirichlet boundary condition and Cauchy initial condition, we have*

$$e(\mathrm{HMM}) \le C\Big[\delta + \Big(\frac{\varepsilon}{\delta}\Big)^{1/2} + \max_{1\le k\le n}\Big(\tau_k + \frac{\varepsilon}{\tau_k^{1/2}}\Big)\Big]. \tag{1.10}$$

Similar results with some modification hold for the nonlinear problems. The details are given in § 4.

1.3. **Parameter choices.** In this part, we analyze the sources of each terms appear in the upper bound of $e(\mathrm{HMM})$. It is clear that the term ε/δ comes from the boundary condition, while the term ε^2/τ_k comes from the initial condition. It is clear to see the corresponding terms vanishes if we let $\delta/\varepsilon, \tau_k/\varepsilon^2 \in \mathbb{N}$ and $v^\varepsilon - V$ be periodic on $\partial\mathfrak{Q}_n$.

For $a^\varepsilon = a(x, x/\varepsilon, t)$, we may choose $\delta = M_1\varepsilon \simeq \varepsilon^{1/2}$, and $\tau_k = M_2\varepsilon^2 \simeq \varepsilon$ for $k = 1, \cdots, n$. With such choice of parameters, we get

$$e(\mathrm{HMM}) \le C\varepsilon^{1/2}. \tag{1.11}$$

For $a^\varepsilon = a(x, x/\varepsilon, t, t/\varepsilon^2)$, we may choose $\delta = M_1\varepsilon \simeq \varepsilon^{1/3}$, and $\tau_k = M_2\varepsilon^2 \simeq \varepsilon^{2/3}$ for $k = 1, \cdots, n$. With such choice of parameters, we have the overall estimate for $e(\mathrm{HMM})$ as

$$e(\mathrm{HMM}) \le C\varepsilon^{1/3}. \tag{1.12}$$

Actually, classical homogenization result suggests that there is no oscillation in the temporal direction when $a^\varepsilon = a(x, x/\varepsilon, t)$. Therefore, we may replace (1.4) by an elliptic cell problem:

$$\begin{cases} -\nabla\cdot\big(a^\varepsilon(\cdot, t_n)\nabla v^\varepsilon\big) = 0 & \text{in} \quad I_\delta, \\ v^\varepsilon = V & \text{on} \quad \partial I_\delta, \end{cases} \tag{1.13}$$

Define w^ε in the same way and $\mathcal{A}_H$ is defined as

$$\nabla W\cdot\mathcal{A}_H(x_K,t_n)\nabla V=\frac{1}{|I_\delta|}\int\limits_{I_\delta}\nabla w^\varepsilon\cdot a^\varepsilon(\cdot,t_n)\nabla v^\varepsilon\,\mathrm{d}x.$$

Corollary 1.4. *For $a^\varepsilon=a(x,x/\varepsilon,t)$ with $a(x,y,t)$ is periodic in y with period Y, if we use the cell problem* (1.13)*, then*

$$e(\mathrm{HMM})\le C\Big(\delta+\frac{\varepsilon}{\delta}\Big). \tag{1.14}$$

The proof of (1.14) is essentially the same as the elliptic case as we have done in [17]. Actually, it may also follow the proof of Theorem 1.2 literally, we omit the proof.

2. Analysis of the Method

2.1. **Preliminaries and notations.** We introduce some notations. Denote by $L^2(D),H^m(D)$ and $H_0^m(D),m\in\mathbb{Z}$ the usual Lebesgue space and Sobolev spaces. $(\cdot,\cdot)_D$ and $\|\cdot\|_{m,D}$ will be denoted as the L^2 inner-product and norms, respectively, the subscript will be omitted if no confusion can occur. $\fint_D u\,\mathrm{d}x$ is defined as the mean value of u over D. For any Banach space U with norm $\|\cdot\|_U$, the space $L^2(0,T;U)$ consists of all measurable functions $u:[0,T]\to U$ with

$$\|u\|_{L^2(0,T;U)}:=\Big(\int_0^T\|u(t)\|_U^2\,\mathrm{d}t\Big)^{1/2}.$$

The space $H^m(0,T;U)$ comprises of all functions $d^ku/dt^k\in L^2(0,T;U)$ for $0\le k\le m$, which is equipped with the norm

$$\|u\|_{H^m(0,T;U)}:=\Big(\int_0^T\sum_{0\le k\le m}\|d^ku/dt^k\|_U^2\,\mathrm{d}t\Big)^{1/2}.$$

The space $\mathcal{C}([0,T];U)$ comprises of all continuous functions $u:[0,T]\to U$ with

$$\|u\|_{\mathcal{C}([0,T];U)}=\max_{0\le t\le T}\|u(t)\|_U.$$

For vectors $\boldsymbol{x}=(x_1,x_2)$ and $\boldsymbol{y}=(y_1,y_2)\in\mathbb{R}^2$, $\boldsymbol{x}\otimes\boldsymbol{y}$ is a 2×2 matrix with elements $(\boldsymbol{x}\otimes\boldsymbol{y})_{ij}:=x_iy_j$. A matrix product is defined by $A:B=\mathrm{tr}(A^TB)$, where $\mathrm{tr}(A)$ is the trace of a 2×2 matrix A.

The following simple result underlines the stability of HMM for problem (1.1). A similar one for the elliptic problem can be found in [17, Lemma 1.9].

Lemma 2.1. *Given a domain $\Omega\in\mathbb{R}^d,T>0$ and a linear function V, let φ be the solution of*

$$\begin{cases}\partial_t\varphi-\nabla\cdot\big(a\nabla\varphi\big)=0 & \textit{in}\quad \Omega\times(0,T],\\ \varphi=V & \textit{on}\quad \partial\Omega\times(0,T],\\ \varphi|_{t=0}=V,\end{cases} \tag{2.1}$$

where $a=\big(a_{ij}\big)$ satisfies

$$\lambda I\le a\le\Lambda I\qquad a.e.\ (x,t)\in\Omega\times(0,T].$$

Then for any $t>0$, we have

$$\|\nabla V\|_{0,\Omega}\le\|\nabla\varphi(x,t)\|_{0,\Omega}\quad\textit{and}\quad\Big(\int\limits_0^t\int\limits_\Omega\nabla\varphi\cdot a\nabla\varphi\Big)^{1/2}\le\Big(\int\limits_0^t\int\limits_\Omega\nabla V\cdot a\nabla V\Big)^{1/2}. \tag{2.2}$$

Proof. Notice that $\varphi = V$ on the boundary of Ω, using the fact that ∇V is a constant in Ω, and integration by parts leads to

$$\int_\Omega \nabla(\varphi - V)(x,t)\nabla V(x)\,\mathrm{d}x = 0 \qquad \text{for any } t > 0,$$

which implies

$$\int_\Omega |\nabla\varphi(x,t)|^2\,\mathrm{d}x = \int_\Omega |\nabla V(x)|^2\,\mathrm{d}x + \int_\Omega |\nabla(\varphi - V)(x,t)|^2\,\mathrm{d}x.$$

This gives the first result of (2.2). Multiplying the first equation of (2.1) by $\varphi - V$ and integrating by parts, we obtain

$$\begin{aligned}\frac{1}{2}\int_\Omega |\varphi(x,t) - V|^2\,\mathrm{d}x &+ \int_0^t \int_\Omega \nabla\varphi(x,s)\cdot a(x,s)\nabla\varphi(x,s)\,\mathrm{d}x\,\mathrm{d}s \\ &= \int_0^t \int_\Omega \nabla V(x)\cdot a(x,s)\nabla\varphi(x,s)\,\mathrm{d}x\,\mathrm{d}s. \end{aligned}\tag{2.3}$$

By Cauchy-Schwartz inequality,

$$\begin{aligned}\int_\Omega \nabla V(x)\cdot a(x,s)\nabla\varphi(x,s)\,\mathrm{d}x\,\mathrm{d}s &\le \Big(\int_0^t \int_\Omega \nabla\varphi(x,s)\cdot a(x,s)\nabla\varphi(x,s)\,\mathrm{d}x\,\mathrm{d}s\Big)^{1/2} \\ &\quad \times \Big(\int_0^t \int_\Omega \nabla V(x)\cdot a(x,s)\nabla V(x)\,\mathrm{d}x\,\mathrm{d}s\Big)^{1/2}.\end{aligned}$$

A combination of the above two gives the second part of (2.2). $\square$

Remark 2.2. *For this result, the coefficient $a = (a_{ij})$ may depend on the solution, i.e.* (2.1) *may be nonlinear.*

2.2. **Generality.** Using (2.2) with $\Omega = I_\delta$, for any $V \in X_H$ and $1 \le k \le n$, we have

$$\begin{aligned}A_H(t_k;V,V) &= \sum_{K\in\mathcal{T}_H} |K| \fint_{\Omega_k} \nabla v^\varepsilon \cdot a^\varepsilon \nabla v^\varepsilon \ge \lambda \sum_{K\in\mathcal{T}_H} |K| \fint_{\Omega_k} |\nabla v^\varepsilon|^2 \\ &\ge \lambda \sum_{K\in\mathcal{T}_H} |K| \fint_{\Omega_k} |\nabla V|^2 = \lambda \sum_{K\in\mathcal{T}_H} \int_K |\nabla V|^2 \\ &= \lambda \|\nabla V\|_0^2. \end{aligned}\tag{2.4}$$

Similarly, we get

$$\begin{aligned}A_H(t_k;V,W) &\le \sum_{K\in\mathcal{T}_H} |K| \Big(\fint_{\Omega_k} \nabla v^\varepsilon \cdot a^\varepsilon \nabla v^\varepsilon\Big)^{1/2} \Big(\fint_{\Omega_k} \nabla w^\varepsilon \cdot a^\varepsilon \nabla w^\varepsilon\Big)^{1/2} \\ &\le \sum_{K\in\mathcal{T}_H} |K| \Big(\fint_{\Omega_k} \nabla V \cdot a^\varepsilon \nabla V\Big)^{1/2} \Big(\fint_{\Omega_k} \nabla W \cdot a^\varepsilon \nabla W\Big)^{1/2} \\ &\le \Lambda \sum_{K\in\mathcal{T}_H} |K|\,|\nabla V|\,|\nabla W| = \Lambda \sum_{K\in\mathcal{T}_H} \Big(\int_K |\nabla V|^2\Big)^{1/2} \Big(\int_K |\nabla W|^2\Big)^{1/2} \\ &\le \Lambda \|\nabla V\|_0 \|\nabla W\|_0. \end{aligned}\tag{2.5}$$

The stability of the method is included in the following lemma. The proof is standard by (2.4) and (2.5), we refer to [26] for details.

Lemma 2.3. *There exists a constant C such that*

$$\|U_H^n\|_0 + |||U_H^n||| \le C\Big(\|u_0\|_0 + \Big(\sum_{k=1}^n \Delta t \|f^k\|_{-1,h}^2\Big)^{1/2}\Big), \tag{2.6}$$

$$\|\nabla U_H^n\|_0 \le C\Big(\|u_0\|_1 + \Big(\sum_{k=1}^n \|f^k\|_{-1,h}^2\Big)^{1/2}\Big), \tag{2.7}$$

where $\|\cdot\|_{-1,h}$ is defined for any $G \in L^2(D)$ as

$$\|G\|_{-1,h} = \sup_{V\in X_H} \frac{(G,V)}{\|\nabla V\|_0}.$$

To prove Theorem 1.1, we define an auxiliary function $\widetilde{U}_H^n \in X_H$ as: Let $\widetilde{U}_H^0 = Q_H u_0$, and for $1 \le k \le n$, $\widetilde{U}_H^k \in X_H$ satisfies

$$(\overline{\partial}\widetilde{U}_H^k, V) + A(t_k; \widetilde{U}_H^k, V) = (f^k, V) \qquad \text{for all } V \in X_H, \tag{2.8}$$

where A is defined as $A(t_k; V, W) = \sum_{K\in\mathcal{T}_H} |K| \nabla W \cdot \mathcal{A}(x_K, t_k)\nabla V$ for all $V, W \in X_H$.

The error estimate for the above problem is well-known [26]:

$$\|\widetilde{U}_H^n - U(x,t_n)\|_0 + |||\widetilde{U}_H^n - U(x,t_n)||| \le C(\Delta t + H^2), \quad \|\widetilde{U}_H^n - U(x,t_n)\|_1 \le C(\Delta t + H). \tag{2.9}$$

Proof for Theorem 1.1 For $1 \le k \le n$, define $E^k \colon = U_H^k - \widetilde{U}_H^k$. For any $V \in X_H$, it is clear that

$$(\overline{\partial}E^k, V) + A(t_k; E^k, V) = (F^k, V), \tag{2.10}$$

where $(F^k, V) \colon = A(t_k; U_H^k, V) - A_H(t_k; U_H^k, V)$. By definition,

$$\|F^k\|_{-1,h} \le e_k(\text{HMM}) \|\nabla U_H^k\|_0.$$

By (2.6) we have, since $E^0 = 0$,

$$\|E^n\|_0 + |||E^n||| \le Ce(\text{HMM}) |||U_H^n||| \le Ce(\text{HMM}). \tag{2.11}$$

Combining the above inequality and the first part of (2.9), we obtain (1.7).

Repeating the above steps, using (2.7) and (2.6), we obtain

$$\|\nabla E^n\|_0 \le Ce(\text{HMM})\Delta t^{-1/2} |||U_H^n||| \le Ce(\text{HMM})\Delta t^{-1/2}.$$

The estimate (1.8) follows from the above estimate and the second part of (2.9). $\square$

Remark 2.4. *Note that $E^n \in X_H$ for any n, using* (2.11) *and the* inverse estimate [10], *we get*

$$\|E^n\|_1 \le (C/H)\|E^n\|_0 \le Ce(\text{HMM})/H,$$

which together with the second part of (2.9) *leads to*

$$\|U_H^n - U(x,t_n)\|_1 \le C(H + \Delta t + e(\text{HMM})/H). \tag{2.12}$$

3. Estimating e(HMM)

In this section, we estimate e(HMM) for two cases, one is $a^\varepsilon = a(x, x/\varepsilon, t)$ and the other is $a^\varepsilon = a(x, x/\varepsilon, t, t/\varepsilon^2)$. In both cases, the cell problem (1.4) is solved with Dirichlet boundary condition and Cauchy initial condition. We will use $a^\varepsilon_{K,n} = a(x_K, x/\varepsilon, t_n)$ or $a^\varepsilon_{K,n} = a(x_K, x/\varepsilon, t_n, t/\varepsilon^2)$ and $\boldsymbol{\chi}_{K,n} = \boldsymbol{\chi}(x_K, x/\varepsilon, t_n)$ or $\boldsymbol{\chi}_{K,n} = \boldsymbol{\chi}(x_K, x/\varepsilon, t_n, t/\varepsilon^2)$ for simplicity, where $\boldsymbol{\chi}$ is the solution of certain cell problems (cf. (3.4) and (3.15)).

Estimating e(HMM) consists of two steps. First, we estimate $\|\widetilde{\mathcal{A}} - \mathcal{A}\|$. The auxiliary operator $\widetilde{\mathcal{A}}$ is defined by

$$\nabla W \cdot \widetilde{\mathcal{A}}(x_K, t_n)\nabla V = \fint_{\mathfrak{Q}_n} \nabla \widehat{W}^\varepsilon \cdot a^\varepsilon_{K,n} \nabla \widehat{V}^\varepsilon \qquad \text{for any } W, V \in X_H, \tag{3.1}$$

where

$$\widehat{V}^\varepsilon = V + \varepsilon \boldsymbol{\chi}_{K,n} \cdot \nabla V \qquad \text{and} \qquad \widehat{W}^\varepsilon = W + \varepsilon \boldsymbol{\chi}_{K,n} \cdot \nabla W.$$

Next we estimate $\|\widetilde{\mathcal{A}} - \mathcal{A}_H\|$. This is achieved by

$$\begin{aligned}\nabla W \cdot (\widetilde{\mathcal{A}} - \mathcal{A}_H)(x_K, t_n)\nabla V = &\fint_{\mathfrak{Q}_n} [\nabla \widehat{W}^\varepsilon \cdot a^\varepsilon_{K,n} \nabla(\widehat{V}^\varepsilon - v^\varepsilon) + \nabla \widehat{V}^\varepsilon \cdot a^\varepsilon_{K,n} \nabla(\widehat{W}^\varepsilon - w^\varepsilon)] \\ &- \fint_{\mathfrak{Q}_n} [\nabla w^\varepsilon \cdot (a^\varepsilon - a^\varepsilon_{K,n})\nabla v^\varepsilon + \nabla(w^\varepsilon - \widehat{W}^\varepsilon) \cdot a^\varepsilon_{K,n} \nabla(v^\varepsilon - \widehat{V}^\varepsilon)].\end{aligned} \tag{3.2}$$

Finally, estimating e(HMM) follows from the triangle inequality.

3.1. Estimating e(HMM) for the case when $a^\varepsilon = a(x, x/\varepsilon, t)$.

Denote by $\hat{v}^\varepsilon$ the solution of (1.4) with a^ε replaced by $a^\varepsilon_{K,n}$. By standard a priori estimate and (2.2), we have

$$\|\nabla(v^\varepsilon - \hat{v}^\varepsilon)\|_{L^2(\mathfrak{Q}_n)} \le C(\delta + \tau_n)\|\nabla v^\varepsilon\|_{L^2(\mathfrak{Q}_n)} \le C(\delta + \tau_n)\|\nabla V\|_{L^2(\mathfrak{Q}_n)}. \tag{3.3}$$

For $j = 1, \cdots, d$, $\boldsymbol{\chi} = \{\chi^j\}_{j=1}^d$ is periodic in y with period Y and satisfies

$$\frac{\partial}{\partial y_i}\Big(a_{ik}\frac{\partial \chi^j}{\partial y_k}\Big)(x, y, t) = -\Big(\frac{\partial}{\partial y_i} a_{ij}\Big)(x, y, t) \quad \text{in } Y, \qquad \int_Y \chi^j(x, y, t)\,\mathrm{d}y = 0. \tag{3.4}$$

This problem is solvable and there exists a constant C such that for $j = 1, \cdots, d$,

$$|\nabla_y \chi^j(x, y, t)| \le C \quad \text{for all } (x, t) \in \mathfrak{Q} \text{ and } y \in Y. \tag{3.5}$$

The effective matrix is given by

$$\mathcal{A}_{ij}(x, t) = \fint_Y \Big(a_{ij} + a_{ik}\frac{\partial \chi^j}{\partial y_k}\Big)(x, y, t)\,\mathrm{d}y \quad i, j = 1, \cdots, d. \tag{3.6}$$

A straightforward calculation gives

$$\nabla \cdot \big(a^\varepsilon_{K,n} \nabla \widehat{V}^\varepsilon\big) = 0 \qquad \text{and} \qquad \nabla \cdot \big(a^\varepsilon_{K,n} \nabla \widehat{W}^\varepsilon\big) = 0. \tag{3.7}$$

Define $\theta^\varepsilon = \hat{v}^\varepsilon - \widehat{V}^\varepsilon$, which obviously satisfies

$$\begin{cases} \partial_t \theta^\varepsilon - \nabla \cdot \big(a^\varepsilon_{K,n} \nabla \theta^\varepsilon\big) = 0 & \text{in} \quad \mathfrak{Q}_n, \\ \theta^\varepsilon = -\varepsilon \boldsymbol{\chi}_{K,n} \cdot \nabla V & \text{on} \quad \partial I_\delta \times T_n, \\ \theta^\varepsilon|_{t=t_n} = -\varepsilon \boldsymbol{\chi}_{K,n} \cdot \nabla V. & \end{cases} \tag{3.8}$$

Lemma 3.1. *Let θ^ε be solution of* (3.8)*. There exists a constant independent of ε, δ and τ_n such that*

$$\|\nabla \theta^\varepsilon\|_{L^2(\mathfrak{Q}_n)} \le C\Big(\frac{\varepsilon}{\tau_n^{1/2}} + \Big(\frac{\varepsilon}{\delta}\Big)^{1/2}\Big)\|\nabla V\|_{L^2(\mathfrak{Q}_n)}. \tag{3.9}$$

Proof. Multiplying both sides of $(3.8)_1$ by $\theta_1^\varepsilon := \theta^\varepsilon + (\widehat{V}^\varepsilon - V)(1-\rho^\varepsilon)$ and integrating over I_δ, we obtain

$$\frac{1}{2}\frac{\partial}{\partial t}\int_{I_\delta}|\theta_1^\varepsilon|^2 + \int_{I_\delta}\nabla\theta_1^\varepsilon\cdot a_{K,n}^\varepsilon\nabla\theta_1^\varepsilon = \int_{I_\delta}\nabla(\theta_1^\varepsilon-\theta^\varepsilon)\cdot a_{K,n}^\varepsilon\nabla\theta_1^\varepsilon, \tag{3.10}$$

where the cut-off function $\rho^\varepsilon \in C_0^\infty(I_\delta), |\nabla\rho^\varepsilon| \le C/\varepsilon$ and

$$\rho^\varepsilon = \begin{cases} 1 & \text{if } \operatorname{dist}(x,\partial I_\delta)\ge 2\varepsilon,\\ 0 & \text{if } \operatorname{dist}(x,\partial I_\delta)\le \varepsilon.\end{cases}$$

It is clear to see

$$|\int_{I_\delta}\nabla(\theta_1^\varepsilon-\theta^\varepsilon)\cdot a_{K,n}^\varepsilon\nabla\theta_1^\varepsilon| \le \Big(\int_{I_\delta}\nabla(\theta_1^\varepsilon-\theta^\varepsilon)\cdot a_{K,n}^\varepsilon\nabla(\theta_1^\varepsilon-\theta^\varepsilon)\Big)^{1/2}\Big(\int_{I_\delta}\nabla\theta_1^\varepsilon\cdot a_{K,n}^\varepsilon\nabla\theta_1^\varepsilon\Big)^{1/2}.$$

Substituting the above inequality into (3.10), we obtain

$$\frac{\partial}{\partial t}\int_{I_\delta}|\theta_1^\varepsilon|^2 + \int_{I_\delta}\nabla\theta_1^\varepsilon\cdot a_{K,n}^\varepsilon\nabla\theta_1^\varepsilon \le \int_{I_\delta}\nabla(\theta_1^\varepsilon-\theta^\varepsilon)\cdot a_{K,n}^\varepsilon\nabla(\theta_1^\varepsilon-\theta^\varepsilon).$$

Integrating the above inequality over T_n, we get

$$\lambda\|\nabla\theta_1^\varepsilon\|_{L^2(\mathcal{Q}_n)}^2 \le \|\theta_1^\varepsilon(x,t_n)\|_{L^2(I_\delta)}^2 + \Lambda\|\nabla(\theta_1^\varepsilon-\theta^\varepsilon)\|_{L^2(\mathcal{Q}_n)}^2,$$

which implies

$$\|\nabla\theta^\varepsilon\|_{L^2(\mathcal{Q}_n)} \le \lambda^{-1/2}\|\theta_1^\varepsilon(x,t_n)\|_{L^2(I_\delta)} + \big(1+(\Lambda/\lambda)^{1/2}\big)\|\nabla(\theta_1^\varepsilon-\theta^\varepsilon)\|_{L^2(\mathcal{Q}_n)}.$$

A direct calculation gives

$$\begin{aligned}\|\nabla(\theta_1^\varepsilon-\theta^\varepsilon)\|_{L^2(\mathcal{Q}_n)} &\le C\Big(\frac{\varepsilon}{\delta}\Big)^{1/2}\|\nabla V\|_{L^2(\mathcal{Q}_n)},\\ \|\theta_1^\varepsilon(x,t_n)\|_{L^2(I_\delta)} &= \varepsilon\|\rho^\varepsilon(\widehat{V}^\varepsilon - V)\|_{L^2(I_\delta)} \le C\varepsilon\|\nabla V\|_{L^2(I_\delta)}.\end{aligned}$$

A combination of the above three inequalities leads to (3.9). □

Next lemma concerns estimating $\|\widetilde{\mathcal{A}}-\mathcal{A}\|$.

Lemma 3.2. *There exists a constant C such that*

$$\|(\widetilde{\mathcal{A}}-\mathcal{A})(x_K,t_n)\| \le C\frac{\varepsilon}{\delta}. \tag{3.11}$$

Proof. Denote by $I_{\kappa\varepsilon} = \kappa Y$, where κ is the integer part of δ/ε, i.e. $\kappa = \lfloor\delta/\varepsilon\rfloor$, integrating by parts and using (3.7), we get

$$⨍_{I_{\kappa\varepsilon}}\nabla(\widehat{W}^\varepsilon - W)\cdot a_{K,n}^\varepsilon\nabla\widehat{V}^\varepsilon = 0.$$

Using the expression of $\widehat{V}^\varepsilon$ and (3.6), we obtain

$$⨍_{I_{\kappa\varepsilon}}\nabla W\cdot a_{K,n}^\varepsilon\nabla\widehat{V}^\varepsilon = \nabla W\cdot\mathcal{A}(x_K,t_n)\nabla V.$$

It follows from the above two equations that

$$⨍_{I_{\kappa\varepsilon}}\nabla\widehat{W}^\varepsilon\cdot a_{K,n}^\varepsilon\nabla\widehat{V}^\varepsilon = \nabla W\cdot\mathcal{A}(x_K,t_n)\nabla V.$$

Since $\widehat{V}^\varepsilon, \widehat{W}^\varepsilon$ and $a_{K,n}^\varepsilon$ are independent of t, we write $\widetilde{\mathcal{A}}$ as

$$\nabla W\cdot\widetilde{\mathcal{A}}(x_K,t_n)\nabla V = ⨍_{I_\delta}\nabla\widehat{W}^\varepsilon\cdot a_{K,n}^\varepsilon\nabla\widehat{V}^\varepsilon \qquad \text{for any } W,V\in X_H,$$

In follows from the above equation and (3.5) that

$$
\begin{aligned}
|\nabla W\cdot(\mathcal{A}-\widetilde{\mathcal{A}})(x_K,t_n)\nabla V| &\le \Big(1-\frac{|I_{\kappa\varepsilon}|}{|I_\delta|}\Big)\fint_{I_{\kappa\varepsilon}}|\nabla W\cdot a^\varepsilon_{K,n}\nabla\widehat{V}^\varepsilon|+|I_\delta|^{-1}\int_{I_\delta\setminus I_{\kappa\varepsilon}}|\nabla W\cdot a^\varepsilon_{K,n}\nabla\widehat{V}^\varepsilon|\\
&\le C\frac{\varepsilon}{\delta}|\nabla W|\,|\nabla V|, \qquad (3.12)
\end{aligned}
$$

which in turn implies (3.11). □

Proof of (1.9) Using the first part of (3.7) and noting $[\widehat{W}^\varepsilon\rho^\varepsilon-w^\varepsilon+W(1-\rho^\varepsilon)](x,t)=0$ for $(x,t)\in\partial I_\delta\times T_n$, integrating by parts, we have

$$
\fint_{\Omega_n}\nabla\widehat{V}^\varepsilon\cdot a^\varepsilon_{K,n}\nabla(\widehat{W}^\varepsilon\rho^\varepsilon-w^\varepsilon+W(1-\rho^\varepsilon))=0.
$$

Therefore, we get

$$
\begin{aligned}
\fint_{\Omega_n}\nabla\widehat{V}^\varepsilon\cdot a^\varepsilon_{K,n}\nabla(\widehat{W}^\varepsilon-w^\varepsilon)&=\fint_{\Omega_n}\nabla\widehat{V}^\varepsilon\cdot a^\varepsilon_{K,n}\nabla[(\widehat{W}^\varepsilon-W)(1-\rho^\varepsilon)]\\
&=\fint_{I_\delta}\nabla\widehat{V}^\varepsilon\cdot a^\varepsilon_{K,n}\nabla[(\widehat{W}^\varepsilon-W)(1-\rho^\varepsilon)].
\end{aligned}
$$

Symmetrically, using the second part of (3.7), we have

$$
\fint_{\Omega_n}\nabla\widehat{W}^\varepsilon\cdot a^\varepsilon_{K,n}\nabla(\widehat{V}^\varepsilon-v^\varepsilon)=\fint_{I_\delta}\nabla\widehat{W}^\varepsilon\cdot a^\varepsilon_{K,n}\nabla[(\widehat{V}^\varepsilon-V)(1-\rho^\varepsilon)]. \qquad (3.13)
$$

Using the above two identities, we rewrite (3.2) as

$$
\begin{aligned}
&\nabla W\cdot(\widetilde{\mathcal{A}}-\mathcal{A}_H)(x_K,t_n)\nabla V\\
&=\fint_{I_\delta}\big[\nabla\widehat{W}^\varepsilon\cdot a^\varepsilon_{K,n}\nabla[(\widehat{V}^\varepsilon-V)(1-\rho^\varepsilon)]+\nabla\widehat{V}^\varepsilon\cdot a^\varepsilon_{K,n}\nabla[(\widehat{W}^\varepsilon-W)(1-\rho^\varepsilon)]\big]\\
&\quad-\fint_{\Omega_n}[\nabla w^\varepsilon\cdot(a^\varepsilon-a^\varepsilon_{K,n})\nabla v^\varepsilon+\nabla(w^\varepsilon-\widehat{W}^\varepsilon)\cdot a^\varepsilon_{K,n}\nabla(v^\varepsilon-\widehat{V}^\varepsilon)]=:I_1+I_2. \qquad (3.14)
\end{aligned}
$$

A direct calculation gives

$$
|I_1|\le C\frac{\varepsilon}{\delta}|\nabla W|\,|\nabla V|.
$$

It follows from (3.3) and (3.9) that

$$
\|\nabla(v^\varepsilon-\widehat{V}^\varepsilon)\|_{L^2(\Omega_n)}\le\|\nabla(v^\varepsilon-\hat{v}^\varepsilon)\|_{L^2(\Omega_n)}+\|\nabla\theta^\varepsilon\|_{L^2(\Omega_n)}\le C\Big(\delta+\tau_n+\Big(\frac{\varepsilon}{\delta}\Big)^{1/2}+\frac{\varepsilon}{\tau_n^{1/2}}\Big)\|\nabla V\|_{L^2(\Omega_n)}.
$$

Similarly, we have

$$
\|\nabla(w^\varepsilon-\widehat{W}^\varepsilon)\|_{L^2(\Omega_n)}\le C\Big(\delta+\tau_n+\Big(\frac{\varepsilon}{\delta}\Big)^{1/2}+\frac{\varepsilon}{\tau_n^{1/2}}\Big)\|\nabla W\|_{L^2(\Omega_n)}.
$$

Using the above two inequalities, we obtain

$$
\begin{aligned}
|I_2|&\le C\frac{\delta+\tau_n}{|\Omega_n|}\|\nabla w^\varepsilon\|_{L^2(\Omega_n)}\|\nabla v^\varepsilon\|_{L^2(\Omega_n)}+\frac{\Lambda}{|\Omega_n|}\|\nabla(w^\varepsilon-\widehat{W}^\varepsilon)\|_{L^2(\Omega_n)}\|\nabla(v^\varepsilon-\widehat{V}^\varepsilon)\|_{L^2(\Omega_n)}\\
&\le C|\Omega_n|^{-1}\Big(\delta+\tau_n+\frac{\varepsilon}{\delta}+\frac{\varepsilon^2}{\tau_n}\Big)\|\nabla W\|_{L^2(\Omega_n)}\|\nabla V\|_{L^2(\Omega_n)}\\
&=C\Big(\delta+\tau_n+\frac{\varepsilon}{\delta}+\frac{\varepsilon^2}{\tau_n}\Big)|\nabla W|\,|\nabla V|.
\end{aligned}
$$

Summing up the estimates for I_1 and I_2, we obtain

$$\|(\widetilde{\mathcal{A}} - \mathcal{A}_H)(x_K, t_n)\| \le C\Big(\delta + \tau_n + \frac{\varepsilon}{\delta} + \frac{\varepsilon^2}{\tau_n}\Big),$$

which together with (3.11) gives (1.9). $\square$

3.2. **Estimating e(HMM) for the case when $a^\varepsilon = a(x, x/\varepsilon, t, t/\varepsilon^2)$.** Next we estimate e(HMM) for the case $a^\varepsilon = a(x, x/\varepsilon, t, t/\varepsilon^2)$ when $a(x, y, t, s)$ is periodic in y and s with period Y and 1, respectively. We assume that (1.4) is solved with Dirichlet boundary condition and Cauchy initial condition. For $j = 1, \cdots, d$, $\boldsymbol{\chi}(x, y, t, s) = \{\chi^j\}_{j=1}^d$ is periodic in y and s with periods Y and 1, respectively, and satisfies

$$\partial_s \chi^j - \partial_{y_i}\Big(a_{ik}\frac{\partial \chi^j}{\partial y_k}\Big)(x, y, t, s) = (\partial_{y_i} a_{ij})(x, y, t, s) \quad \text{and} \quad \int_0^1\int_Y \chi^j(x, y, t, s)\,\mathrm{d}y\,\mathrm{d}s = 0. \tag{3.15}$$

The existence of χ^j is classical since

$$\int_0^1\int_Y (\partial_{y_i} a_{ij})(x, y, t, s)\,\mathrm{d}y\,\mathrm{d}s = 0.$$

By [20], there exists a constant C such that for $j = 1, \cdots, d$,

$$|\chi^j(x, y, s, t)| + |\nabla_y \chi^j(x, y, s, t)| \le C \quad \text{for all } (x, t) \in \mathcal{Q}, y \in Y \text{ and } s \in (0, 1). \tag{3.16}$$

Denote by $\hat{v}^\varepsilon$ the solution of (1.4) with a^ε replaced by $a^\varepsilon_{K,n}$. Using the standard a priori estimate and Lemma 2.1, we have

$$\|\nabla(v^\varepsilon - \hat{v}^\varepsilon)\|_{L^2(\mathcal{Q}_n)} \le C(\delta + \tau_n)\|\nabla V\|_{L^2(\mathcal{Q}_n)}. \tag{3.17}$$

It is easy to verify that

$$\partial_t \widehat{V}^\varepsilon - \nabla\cdot\big(a^\varepsilon_{K,n}\nabla\widehat{V}^\varepsilon\big) = 0 \quad \text{and} \quad \partial_t \widehat{W}^\varepsilon - \nabla\cdot\big(a^\varepsilon_{K,n}\nabla\widehat{W}^\varepsilon\big) = 0, \tag{3.18}$$

and

$$\begin{cases} \partial_t\theta^\varepsilon - \nabla\cdot\big(a^\varepsilon_{K,n}\nabla\theta^\varepsilon\big) = 0 & \text{in} \quad \mathcal{Q}_n, \\ \qquad\qquad \theta^\varepsilon = -\varepsilon\boldsymbol{\chi}_{K,n}\cdot\nabla V & \text{on} \quad \partial I_\delta \times T_n, \\ \qquad \theta^\varepsilon|_{t=t_n} = -\varepsilon(\boldsymbol{\chi}_{K,n}\cdot\nabla V)|_{t=t_n}. & \end{cases} \tag{3.19}$$

For the correction θ^ε, we have the following estimate (cf. (3.9)).

Lemma 3.3. *There exists a constant C independent of ε, δ and τ_n such that*

$$\|\nabla\theta^\varepsilon\|_{L^2(\mathcal{Q}_n)} \le C\Big(\Big(\frac{\varepsilon}{\delta}\Big)^{1/2} + \frac{\varepsilon}{\tau_n^{1/2}}\Big)\|\nabla V\|_{L^2(\mathcal{Q}_n)}. \tag{3.20}$$

The proof of (3.20) is essentially the same as Lemma 3.1. The difference lies in the second term in the right-hand side of the equation below.

Proof. Multiplying both sides of $(3.19)_1$ by $\theta_1^\varepsilon := \theta^\varepsilon + (\widehat{V}^\varepsilon - V)(1 - \rho^\varepsilon)$ and integrating by parts, we get

$$\frac{1}{2}\frac{\partial}{\partial t}\int_{I_\delta}|\theta_1^\varepsilon|^2 + \int_{I_\delta}\nabla\theta_1^\varepsilon\cdot a^\varepsilon_{K,n}\nabla\theta_1^\varepsilon = \int_{I_\delta}\nabla\theta_1^\varepsilon\cdot a^\varepsilon\nabla(\theta_1^\varepsilon - \theta^\varepsilon) + \frac{1}{2}\int_{I_\delta}\theta_1^\varepsilon\partial_t(\theta_1^\varepsilon - \theta^\varepsilon). \tag{3.21}$$

It follows from (3.15) that

$$\begin{aligned}\int_{I_\delta}\theta_1^\varepsilon\partial_t(\theta_1^\varepsilon-\theta^\varepsilon)&=\varepsilon^{-1}\int_{I_\delta}\partial_s\boldsymbol{\chi}_{K,n}\cdot\nabla V(1-\rho^\varepsilon)\theta_1^\varepsilon\\&=\varepsilon^{-1}\int_{I_\delta}\nabla_y\cdot(a_{K,n}^\varepsilon(I+\nabla_y\boldsymbol{\chi}_{K,n}))\nabla V(1-\rho^\varepsilon)\theta_1^\varepsilon\\&=\int_{I_\delta}\nabla\cdot(a_{K,n}^\varepsilon(I+\nabla_y\boldsymbol{\chi}_{K,n}))\nabla V(1-\rho^\varepsilon)\theta_1^\varepsilon.\end{aligned}$$

Integrating by parts, we obtain

$$\begin{aligned}\int_{I_\delta}\theta_1^\varepsilon\partial_t(\theta_1^\varepsilon-\theta^\varepsilon)&=-\int_{I_\delta}\nabla(\theta_1^\varepsilon(1-\rho^\varepsilon)\nabla V):a_{K,n}^\varepsilon(I+\nabla_y\boldsymbol{\chi}_{K,n})\\&=-\int_{I_\delta}(1-\rho^\varepsilon)[\nabla\theta_1^\varepsilon\otimes\nabla V]:a_{K,n}^\varepsilon(I+\nabla_y\boldsymbol{\chi}_{K,n})\\&\quad+\int_{I_\delta}\theta_1^\varepsilon\nabla\rho^\varepsilon\cdot a_{K,n}^\varepsilon(I+\nabla_y\boldsymbol{\chi}_{K,n})\nabla V.\end{aligned}\tag{3.22}$$

Using (3.16), we bound the first term in the right-hand side of the above equation as

$$\begin{aligned}&|\int_{I_\delta}(1-\rho^\varepsilon)[\nabla\theta_1^\varepsilon\otimes\nabla V]:a_{K,n}^\varepsilon(I+\nabla_y\boldsymbol{\chi}_{K,n})|\\&\le\Lambda\max_{(x,t)\in\Omega_n}\|I+\nabla_y\boldsymbol{\chi}_{K,n}\|\,\|\nabla\theta_1^\varepsilon\|_{L^2(I_\delta)}\|\nabla V\|_{L^2(I_\delta\setminus I_{(\kappa-2)\varepsilon})}\\&\le C\Big(\frac{\varepsilon}{\delta}\Big)^{1/2}\|\nabla\theta_1^\varepsilon\|_{L^2(I_\delta)}\|\nabla V\|_{L^2(I_\delta)}.\end{aligned}$$

By maximum principle [20], we have

$$\max_{(x,t)\in\Omega_n}|\theta^\varepsilon(x,t)|\le\varepsilon\max_{(x,t)\in\Omega_n}|\boldsymbol{\chi}_{K,n}(x,t)|\,|\nabla V|.\tag{3.23}$$

We thus get

$$\max_{(x,t)\in\Omega_n}|\theta_1^\varepsilon(x,t)|\le\max_{(x,t)\in\Omega_n}\big(|\theta^\varepsilon(x,t)|+\varepsilon|\boldsymbol{\chi}_{K,n}(x,t)||\nabla V|\big)\le2\varepsilon\max_{(x,t)\in\Omega_n}|\boldsymbol{\chi}_{K,n}(x,t)||\nabla V|.$$

Therefore, we bound the second term in the right-hand side of (3.22) as

$$\begin{aligned}|\int_{I_\delta}\theta_1^\varepsilon\nabla\rho^\varepsilon\cdot a_{K,n}^\varepsilon(I+\nabla_y\boldsymbol{\chi}_{K,n})\nabla V|&\le2\Lambda\max_{(x,t)\in\Omega_n}\|I+\nabla_y\boldsymbol{\chi}_{K,n}(x,t)\|\int_{I_\delta}|\nabla V|^2|\varepsilon\nabla\rho^\varepsilon|\\&\le C\frac{\varepsilon}{\delta}\|\nabla V\|_{L^2(I_\delta)}^2.\end{aligned}$$

Substituting the above two estimates into (3.21), we obtain

$$\begin{aligned}\frac12\frac{\partial}{\partial t}\int_{I_\delta}|\theta_1^\varepsilon|^2+\int_{I_\delta}\nabla\theta_1^\varepsilon\cdot a_{K,n}^\varepsilon\nabla\theta_1^\varepsilon&\le\frac12\int_{I_\delta}\nabla\theta_1^\varepsilon\cdot a_{K,n}^\varepsilon\nabla\theta_1^\varepsilon+\int_{I_\delta}\nabla(\theta^\varepsilon-\theta_1^\varepsilon)\cdot a_{K,n}^\varepsilon\nabla(\theta^\varepsilon-\theta_1^\varepsilon)\\&\quad+C\frac{\varepsilon}{\delta}\|\nabla V\|_{L^2(I_\delta)}^2.\end{aligned}$$

Therefore, integrating the above inequality over T_n, we obtain

$$\|\nabla\theta_1^\varepsilon\|_{L^2(\Omega_n)}\le C\Big(\|\theta_1^\varepsilon(x,t_n)\|_{L^2(I_\delta)}+\|\nabla(\theta^\varepsilon-\theta_1^\varepsilon)\|_{L^2(\Omega_n)}+\Big(\frac{\varepsilon}{\delta}\Big)^{1/2}\|\nabla V\|_{L^2(\Omega_n)}\Big),$$

which in turn implies

$$\|\nabla\theta^\varepsilon\|_{L^2(\Omega_n)}\le C\Big(\|\theta_1^\varepsilon(x,t_n)\|_{L^2(I_\delta)}+C\|\nabla(\theta^\varepsilon-\theta_1^\varepsilon)\|_{L^2(\Omega_n)}+C\Big(\frac{\varepsilon}{\delta}\Big)^{1/2}\|\nabla V\|_{L^2(\Omega_n)}\Big).$$

A direct calculation gives

$$\|\theta_1^\varepsilon(x,t_n)\|_{L^2(I_\delta)} \le C\varepsilon\|\nabla V\|_{L^2(I_\delta)},$$
$$\|\nabla(\theta^\varepsilon-\theta_1^\varepsilon)\|_{L^2(\mathcal{Q}_n)} \le C\Big(\frac{\varepsilon}{\delta}\Big)^{1/2}\|\nabla V\|_{L^2(\mathcal{Q}_n)}.$$

A combination of the above three inequalities leads to (3.20). □

Similar to Lemma 3.2, we have

Lemma 3.4. *There exists a constant C such that*

$$\|(\mathcal{A}-\widetilde{\mathcal{A}})(x_K,t_n)\| \le C\Big(\frac{\varepsilon}{\delta}+\frac{\varepsilon^2}{\tau_n}\Big). \tag{3.24}$$

Proof. Let $\ell\colon=\lfloor\tau_n/\varepsilon^2\rfloor$, and $\widetilde{\mathcal{Q}}_n\colon=I_{\kappa\varepsilon}\times(t_n,t_n+\ell\varepsilon^2)$. The key to the proof is the following observation: For any $V,W\in X_H$, we have

$$\nabla W\cdot\mathcal{A}(x_K,t_n)\nabla V\colon=\fint_{\widetilde{\mathcal{Q}}_n}\nabla\widehat{W}^\varepsilon\cdot a_{K,n}^\varepsilon\nabla\widehat{V}^\varepsilon. \tag{3.25}$$

Integration by parts and using the first part of (3.18), we obtain

$$\begin{aligned}\fint_{\widetilde{\mathcal{Q}}_n}\nabla(\widehat{W}^\varepsilon-W)\cdot a_{K,n}^\varepsilon\nabla\widehat{V}^\varepsilon&=-\fint_{\widetilde{\mathcal{Q}}_n}(\widehat{W}^\varepsilon-W)\nabla\cdot(a_{K,n}^\varepsilon\nabla\widehat{V}^\varepsilon)=-\fint_{\widetilde{\mathcal{Q}}_n}(\widehat{W}^\varepsilon-W)\partial_t\widehat{V}^\varepsilon\\&=-\fint_{\widetilde{\mathcal{Q}}_n}(\widehat{W}^\varepsilon-W)\partial_t(\widehat{V}^\varepsilon-V).\end{aligned}$$

A direct calculation leads to

$$\fint_{\widetilde{\mathcal{Q}}_n}\nabla W\cdot a_{K,n}^\varepsilon\nabla\widehat{V}^\varepsilon=\nabla W\cdot\mathcal{A}(x_K,t_n)\nabla V.$$

Adding up the above two equations, we obtain

$$\nabla W\cdot\mathcal{A}(x_K,t_n)\nabla V-\fint_{\widetilde{\mathcal{Q}}_n}\nabla\widehat{W}^\varepsilon\cdot a_{K,n}^\varepsilon\nabla\widehat{V}^\varepsilon=\fint_{\widetilde{\mathcal{Q}}_n}(\widehat{W}^\varepsilon-W)\partial_t(\widehat{V}^\varepsilon-V)$$

Exchanging W and V and notice that a^ε and $\mathcal{A}$ are symmetric, we get

$$\nabla W\cdot\mathcal{A}(x_K,t_n)\nabla V-\fint_{\widetilde{\mathcal{Q}}_n}\nabla\widehat{W}^\varepsilon\cdot a_{K,n}^\varepsilon\nabla\widehat{V}^\varepsilon=\fint_{\widetilde{\mathcal{Q}}_n}(\widehat{V}^\varepsilon-V)\partial_t(\widehat{W}^\varepsilon-W).$$

Adding up the above two equations and using the explicit expressions of $\widehat{V}^\varepsilon$ and $\widehat{W}^\varepsilon$, we get

$$\nabla W\cdot\mathcal{A}(x_K,t_n)\nabla V-\fint_{\widetilde{\mathcal{Q}}_n}\nabla\widehat{W}^\varepsilon\cdot a_{K,n}^\varepsilon\nabla\widehat{V}^\varepsilon=\frac{1}{2}\fint_{\widetilde{\mathcal{Q}}_n}\partial_t[(\widehat{V}^\varepsilon-V)(\widehat{W}^\varepsilon-W)]=0,$$

which gives (3.25).

By (3.25), proceeding as that in (3.12) and using (3.16), we get (3.24). □

Proof for (1.10) It follows from (3.2), (3.17) and Lemma 3.3 that

$$
\begin{aligned}
|\nabla W\cdot(\widetilde{\mathcal{A}}-\mathcal{A}_H)(x_K,t_n)\nabla V| \le{}& C\Big(\delta+\tau_n+\Big(\frac{\varepsilon}{\delta}\Big)^{1/2}+\frac{\varepsilon}{\tau_n^{1/2}}\Big)|\nabla W|\,|\nabla V|\\
&+C\frac{\delta+\tau_n}{|\mathcal{Q}_n|}\|\nabla w^\varepsilon\|_{L^2(\mathcal{Q}_n)}\|\nabla v^\varepsilon\|_{L^2(\mathcal{Q}_n)}\\
&+\frac{\Lambda}{|\mathcal{Q}_n|}\|\nabla(w^\varepsilon-\widehat{W}^\varepsilon)\|_{L^2(\mathcal{Q}_n)}\|\nabla(v^\varepsilon-\widehat{V}^\varepsilon)\|_{L^2(\mathcal{Q}_n)}\\
\le{}& C\Big(\delta+\tau_n+\Big(\frac{\varepsilon}{\delta}\Big)^{1/2}+\frac{\varepsilon}{\tau_n^{1/2}}\Big)|\nabla W||\nabla V|,
\end{aligned}
$$

which implies

$$
\|(\widetilde{\mathcal{A}}-\mathcal{A}_H)(x_K,t_n)\|\le C\Big(\delta+\tau_n+\Big(\frac{\varepsilon}{\delta}\Big)^{1/2}+\frac{\varepsilon}{\tau_n^{1/2}}\Big). \tag{3.26}
$$

This estimate together with (3.24) leads to (1.10). □

Remark 3.5. *One may wonder whether the estimate* (1.10) *can be improved to* (1.9)*. This is actually not the case due to* (3.26)*.*

4. Nonlinear Problem

We consider the following nonlinear problem

$$
\begin{cases}
\partial_t u^\varepsilon-\nabla\cdot\big(a^\varepsilon\big(x,t,u^\varepsilon\big)\nabla u^\varepsilon\big)=f & \text{in}\quad \mathcal{Q},\\
u^\varepsilon=0 & \text{on}\quad \partial D\times(0,T),\\
u^\varepsilon|_{t=0}=u_0.
\end{cases}\tag{4.1}
$$

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

$$
\lambda|\xi|^2\le a^\varepsilon_{ij}(x,t,z)\xi_i\xi_j\le\Lambda|\xi|^2\qquad\text{for all }\xi\in\mathbb{R}^d\ \text{ and for all }(x,t)\in\mathcal{Q}\text{ and }z\in\mathbb{R}
$$

with $0<\lambda\le\Lambda$. Moreover, we assume that $a^\varepsilon(x,t,z)$ is Lipschitz continuous in z uniformly with respect to x and t. The existence of u^ε is classic. Similar problem in the elliptic case has been discussed in [7], the extension to (4.1) is straightforward. We refer to [19] for more general nonlinear problems. The homogenized problem, if it exists, is of the following form:

$$
\begin{cases}
\partial_t U-\nabla\cdot\big(\mathcal{A}\big(x,t,U\big)\nabla U\big)=f & \text{in}\quad \mathcal{Q},\\
U=0 & \text{on}\quad \partial D\times(0,T),\\
U|_{t=0}=u_0.
\end{cases}\tag{4.2}
$$

To formulate HMM, for any $V\in X_H$, define v^ε to be the solution of

$$
\begin{cases}
\partial_t v^\varepsilon-\nabla\cdot\big(a^\varepsilon\big(x,t,v^\varepsilon\big)\nabla v^\varepsilon\big)=0 & \text{in}\quad \mathcal{Q}_n,\\
v^\varepsilon=V & \text{on}\quad \partial I_\delta\times T_n,\\
v^\varepsilon|_{t=t_n}=V.
\end{cases}\tag{4.3}
$$

We can define w^ε similarly.

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

$$
\nabla W\cdot\mathcal{A}_H(x_K,t_n,V)\nabla V:=\fint_{\mathcal{Q}_n}\nabla w^\varepsilon\cdot a^\varepsilon(x,t,v^\varepsilon)\nabla v^\varepsilon,
$$

and $A_H(t_n;V,W)=\sum_{K\in\mathcal{T}_H}|K|\nabla W\cdot\mathcal{A}_H(x_K,t_n,V)\nabla V$.

The HMM solution is given by the problem:

Problem 4.1. *Let $U_H^0 = Q_H u_0$, for $k = 1, \cdots, n$, find $U_H^k \in X_H$ such that*

$$(\overline{\partial} U_H^k, V) + A_H(t_k; U_H^k, V) = (f^k, V) \qquad \text{for all } V \in X_H. \tag{4.4}$$

Remark 4.2. *We only consider a special nonlinear problem, the algorithm applies to much more general nonlinear problem, cf.* [19]*, which together with realistic application will be dealt in a forthcoming paper.*

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

$$E_k(V, W) := \nabla W \cdot (\mathcal{A}_H - \mathcal{A})(x_K, t_k, V)\nabla V,$$

and

$$e(\text{HMM}) = \max_{\substack{K\in\mathcal{T}_H, V\in X_H,\\ 1\le k\le n}} \frac{E_k(V, W)}{|\nabla W||\nabla V|}.$$

Proceeding along the same line of Lemma 2.1, we get the same estimate for v^ε. Notice that a^ε in the second part of (2.2) depends on the solution v^ε. Obviously, for any $V \in X_H$, we have

$$A_H(t_k; V, V) \ge \lambda \|\nabla V\|_0^2. \tag{4.5}$$

By (4.5), it is easy to derive a stability result that is similar to (2.6) and (2.7).

Similar to the second part of (2.2), for any $W \in X_H$, we have

$$\Big(\int_0^t \int_\Omega \nabla w^\varepsilon \cdot a^\varepsilon(x, t, w^\varepsilon)\nabla w^\varepsilon\Big)^{1/2} \le \Big(\int_0^t \int_\Omega \nabla W \cdot a^\varepsilon(x, t, w^\varepsilon)\nabla W\Big)^{1/2}.$$

Using the above inequality, we get

$$\begin{aligned}
A_H(t_k; V, W) &\le \sum_{K\in\mathcal{T}_H} |K| \Big(\frac{\Lambda}{\lambda}\Big)^{1/2} \Big(\fint_{\mathrm{Q}_k} \nabla v^\varepsilon \cdot a^\varepsilon(x, t, v^\varepsilon)\nabla v^\varepsilon\Big)^{1/2} \Big(\fint_{\mathrm{Q}_k} \nabla w^\varepsilon \cdot a^\varepsilon(x, t, w^\varepsilon)\nabla w^\varepsilon\Big)^{1/2}\\
&\le \sum_{K\in\mathcal{T}_H} |K| \Big(\frac{\Lambda}{\lambda}\Big)^{1/2} \Big(\fint_{\mathrm{Q}_k} \nabla V \cdot a^\varepsilon(x, t, v^\varepsilon)\nabla V\Big)^{1/2} \Big(\fint_{\mathrm{Q}_k} \nabla W \cdot a^\varepsilon(x, t, w^\varepsilon)\nabla W\Big)^{1/2}\\
&\le \Lambda \Big(\frac{\Lambda}{\lambda}\Big)^{1/2} \sum_{K\in\mathcal{T}_H} |K|\,|\nabla V|\,|\nabla W| = \Lambda \Big(\frac{\Lambda}{\lambda}\Big)^{1/2} \sum_{K\in\mathcal{T}_H} \Big(\int_K |\nabla V|^2\Big)^{1/2} \Big(\int_K |\nabla W|^2\Big)^{1/2}\\
&\le \Lambda (\Lambda/\lambda)^{1/2} \|\nabla V\|_0 \|\nabla W\|_0. \qquad (4.6)
\end{aligned}$$

The existence of the solution easily follows from the standard approach in [13] by (4.5) and (4.6), while the uniqueness is more involved, which together with the error estimate will be addressed in Theorem 4.3.

The error estimate for Problem 4.1 is essentially the same as the linear case. Define $\widetilde{U}_H^n$ as: Let $\widetilde{U}_H^0 = Q_H u_0$, for $k = 1, \cdots, n$, $\widetilde{U}_H^k \in X_H$ satisfies

$$(\overline{\partial} \widetilde{U}_H^k, V) + A(t_k; \widetilde{U}_H^k, V) = (f^k, V) \qquad \text{for all } V \in X_H,$$

where

$$A(t_k; \widetilde{U}_H^k, V) = \sum_{K\in\mathcal{T}_H} |K| \nabla V \cdot \mathcal{A}(x_K, t_k, \widetilde{U}_H^k)\nabla \widetilde{U}_H^k.$$

For notation simplicity, we associate A with an operator $\hat{\mathcal{A}}$ as

$$(\hat{\mathcal{A}}(x, t_k, V)\nabla V, \nabla W) = A(t_k; V, W) \qquad \text{for all } V, W \in X_H.$$

By [7, Theorem 3.1], the effective matrix $\mathcal{A}$ satisfies

$$\lambda I \le \mathcal{A} \le (\Lambda^2/\lambda) I.$$

Moreover, by [7, Proposition 3.5], $\mathcal{A}(x,t,z)$ (so does $\hat{\mathcal{A}}$) is Lipschitz continuous in z uniformly with respect to all $(x,t) \in \Omega$, and the Lipschitz constant is denoted by L. By [26],

$$\|\widetilde{U}_H^n - U(x,t_n)\|_0 \le C(\Delta t + H^2), \tag{4.7}$$

and there exists a constant $K_1 := C_*(\Delta t^{1/2} + H + \Delta t/H)$ such that

$$\Delta t^{1/2} \|\nabla \widetilde{U}_H^n\|_{L^\infty} \le K_1, \tag{4.8}$$

where C_* depends on U.

Theorem 4.3. *Let U and U_H^n be solutions of (4.2) and (4.4), respectively. Then, under the appropriate regularity assumption on U, we have, for small Δt,*

$$\|U_H^n - U(x,t_n)\|_0 \le C\big(H^2 + \Delta t + e(\mathrm{HMM})\big). \tag{4.9}$$

Moreover, for $M = K_1 + CH^{-1}e(\mathrm{HMM})$ with C a generic constant independent of $\varepsilon, \delta, H, \tau_n, X, Z$ and V, if M satisfies

$$L^2 M^2 < \lambda, \tag{4.10}$$

and there exists a constant $\eta(M)$ with $0 < \eta(M) < \lambda/2$ such that

$$\int_D |E_k(X,V) - E_k(Z,V)| \,\mathrm{d}x \le \eta(M) \|X - Z\|_1 \|\nabla V\|_0 \tag{4.11}$$

for all $X, Z \in X_H \cap W^{1,\infty}(D)$ and $V \in X_H$ satisfying $\|X\|_{1,\infty}, \|Z\|_{1,\infty} \le M$, then the HMM solution is locally unique.

Proof. Define $E^n = U_H^n - \widetilde{U}_H^n$, we have for any $V \in X_H$,

$$\begin{aligned}(\overline{\partial} E^k, V) + (\hat{\mathcal{A}}(x,t_k,U_H^k)\nabla E^k, \nabla V) &= (A - A_H)(t_k; U_H^k, V) \\ &\quad + \big((\hat{\mathcal{A}}(x,t_k,\widetilde{U}_H^k) - \hat{\mathcal{A}}(x,t_k,U_H^k))\nabla \widetilde{U}_H^k, \nabla V\big).\end{aligned}$$

Taking $V = E^k$ in the above equation and using (4.5), we get

$$\frac{1}{2\Delta t}\Big(\|E^k\|_0^2 - \|E^{k-1}\|_0^2\Big) + \lambda \|\nabla E^k\|_0^2 \le e(\mathrm{HMM})\|\nabla U_H^k\|_0 \|\nabla E^k\|_0 + C\|\nabla \widetilde{U}_H^k\|_{L^\infty} \|E^k\|_0 \|\nabla E^k\|_0.$$

Using (4.8) and a kickback of $\|\nabla E^k\|$, we get

$$\frac{1}{2\Delta t}\Big(\|E^k\|_0^2 - \|E^{k-1}\|_0^2\Big) + \frac{\lambda}{2}\|\nabla E^k\|_0^2 \le (e^2(\mathrm{HMM})/\lambda)\|\nabla U_H^k\|_0^2 + C\|E^k\|_0^2. \tag{4.12}$$

There exists a constant M_1 such that for $\Delta t < M_1$, there holds

$$\|E^k\|_0^2 \le (1 + C\Delta t)\|E^{k-1}\|_0^2 + C\Delta t\, e^2(\mathrm{HMM})\|\nabla U_H^k\|_0^2.$$

Hence, by recursive application of the above inequality and noting that $E^0 = 0$, we obtain

$$\|E^n\|_0^2 \le Ce^2(\mathrm{HMM})\Delta t \sum_{k=1}^n (1 + C\Delta t)^{n-k} \|\nabla U_H^k\|_0^2 \le Ce^2(\mathrm{HMM}) |||U_H|||^2. \tag{4.13}$$

This together with (4.7) gives (4.9).

Let $U_H^n = X$ and $U_H^n = Z$ be solutions of Problem 4.1 with U_H^{n-1} given, then by substraction, we get for all $V \in X_H$,

$$(X - Z, V) + \Delta t\, A_H(t_n; X, V) = \Delta t\, A_H(t_n; Z, V),$$

which can be rewrite as

$$(X-Z,V)+\Delta t(\hat{\mathcal{A}}(x,t_n,X)\nabla(X-Z),\nabla V)=\Delta t(A_H-A)(t_n;Z,V)-\Delta t(A_H-A)(t_n;X,V)\\ +\Delta t\big([\hat{\mathcal{A}}(x,t_n,Z)-\hat{\mathcal{A}}(x,t_n,X)]\nabla Z,\nabla V\big).$$

Taking $V=X-Z$ in the above equation and using (4.11), we get

$$\|X-Z\|_0^2+\lambda\Delta t\|\nabla(X-Z)\|_0^2\le\eta(M)\Delta t\|\nabla(X-Z)\|_0^2+L\Delta t\|\nabla Z\|_{L^\infty}\|X-Z\|_0\|\nabla(X-Z)\|_0.$$

After a kickback of $\|\nabla(X-Z)\|_0$, we obtain

$$\|X-Z\|_0^2+(\lambda/2)\Delta t\|\nabla(X-Z)\|_0^2\le\eta(M)\Delta t\|\nabla(X-Z)\|_0^2+\frac{L^2\Delta t}{2\lambda}\|\nabla Z\|_{L^\infty}^2\|X-Z\|_0^2. \tag{4.14}$$

It follows from (4.12) and (4.13) that

$$\Delta t\|\nabla E^n\|_0^2\le C\big(\Delta t\|E^n\|_0^2+\|E^{n-1}\|_0^2+\Delta te^2(\mathrm{HMM})\big)\le Ce^2(\mathrm{HMM}).$$

This together with (4.8) and the inverse inequality give

$$\Delta t^{1/2}\|\nabla Z\|_{L^\infty}\le K_1+CH^{-1}\Delta t^{1/2}\|\nabla E^n\|_0\le K_1+CH^{-1}e(\mathrm{HMM}).$$

Substituting the above inequality into (4.14), we get

$$\|X-Z\|_0^2+(\lambda/2)\Delta t\|\nabla(X-Z)\|_0^2\le\eta(M)\Delta t\|\nabla(X-Z)\|_0^2+(L^2M^2/\lambda)\|X-Z\|_0^2$$

with $M=K_1+CH^{-1}e(\mathrm{HMM})$. Using (4.10) and (4.11), we get $X=Z$, i.e. the HMM solution is locally unique. $\square$

Remark 4.4. *Conditions* (4.10) *and* (4.11) *show that the HMM solution may not be unique if the estimating data procedure is not accurate enough. This is indeed the case even if the homogenized solution U is unique. We refer to* [3] *for related discussion on the approximation of the quasilinear elliptic problems.*

To simplify the presentation, we will show how to estimate $e(\mathrm{HMM})$ when (4.3) is changed slightly to

$$\begin{cases}\partial_t v^\varepsilon-\nabla\cdot\big(a^\varepsilon\big(x,t,V(x_K)\big)\nabla v^\varepsilon\big)=0 & \text{in}\quad \mathcal{Q}_n,\\ v^\varepsilon=V & \text{on}\quad \partial I_\delta\times T_n,\\ v^\varepsilon|_{t=t_n}=V,\end{cases} \tag{4.15}$$

and A_H is changed to

$$A_H(t_n;V,W)=\sum_{K\in\mathcal{T}_H}|K|\fint_{\mathcal{Q}_n}\nabla w^\varepsilon\cdot a^\varepsilon(x,t,V(x_K))\nabla v^\varepsilon.$$

Estimating $e(\mathrm{HMM})$ with cell problem (4.3) is more involved and we will address it in a forthcoming paper.

Theorem 4.5. *If we assume that $a^\varepsilon(x,t,u^\varepsilon)=a(x,x/\varepsilon,t,u^\varepsilon)$ with $a(x,y,t,p)$ periodic in y with period Y, and the cell problem* (4.15) *is employed, then*

$$e(\mathrm{HMM})\le C\Big(\delta+\Big(\frac{\varepsilon}{\delta}\Big)^{1/2}+\max_{1\le k\le n}\Big(\tau_k+\frac{\varepsilon}{\tau_k^{1/2}}\Big)\Big). \tag{4.16}$$

If $\big(\delta+(\varepsilon/\delta)^{1/2}+\tau_n+\varepsilon/\tau_n^{1/2}\big)/\Delta t^{1/2}$, $\big(\delta+(\varepsilon/\delta)^{1/2}+\tau_n+\varepsilon/\tau_n^{1/2}\big)/H$ and $\Delta t/H$ are sufficiently small, then (4.10) *and* (4.11) *hold.*

Proof. By the homogenization result in [4] and proceeding along the same line of (1.9), we may get (4.16). The only modification lies in the fact that $\mathcal{A}_H$ is not symmetric, therefore, the identity (3.13) is invalid, which actually accounts for the accuracy loss in (4.16).

To verify the validity of (4.10) and (4.11), we proceed in the same fashion of [17, Theorem 5.5]. Define

$$\mathcal{K}_1 = \delta + \Big(\frac{\varepsilon}{\delta}\Big)^{1/2} + \tau_n + \frac{\varepsilon}{\tau_n^{1/2}}.$$

It follows from (4.16) that

$$L^2 M^2 \le 2L^2(K_1^2 + CH^{-2}\mathcal{K}_1^2) = 2L^2 C_*^2(\Delta t + H^2 + (\Delta t/H)^2) + CL^2H^{-2}\mathcal{K}_1^2).$$

Therefore, there exists $\rho_0 > 0$ and $\rho_1 > 0$ such that if $\Delta t, H, \Delta t/H < \rho_0$ and $\mathcal{K}_1/H < \rho_1$, we get (4.10).

Next, proceeding in the same fashion of [17, Lemma 5.9], we may take $\eta(M) = C(1 + M\Delta t^{-1/2})\mathcal{K}_1$. Invoking (4.16) once again, we obtain

$$\begin{aligned}\eta(M) &\le C(1 + K_1\Delta t^{-1/2})\mathcal{K}_1 + CH^{-1}\Delta t^{-1/2}\mathcal{K}_1^2\\ &\le C(1 + C_*)\mathcal{K}_1 + C_*(H/\Delta t^{1/2} + \Delta t^{1/2}/H)\mathcal{K}_1 + CH^{-1}\Delta t^{-1/2}\mathcal{K}_1^2.\end{aligned}$$

Therefore, there exists a constant ρ_2 such that if $\mathcal{K}_1/\Delta t^{1/2} < \rho_2$, we have $\eta(M) < \lambda/2$. Finally, let $\rho = \min(\rho_1, \rho_2)$, if $\mathcal{K}_1/\Delta t^{1/2}, \mathcal{K}_1/H < \rho$ and $\Delta t, H, \Delta t/H < \rho_0$, then (4.10) and (4.11) hold true. □

Remark 4.6. *For the case when $a^\varepsilon = a(x, x/\varepsilon, t, u^\varepsilon)$, a formal asymptotic expansion shows that there is no oscillation in the temporal direction and u^ε plays a role of a parameter in the cell problem. Taking into account these special features of the problem, we may employ the following cell problem: For any $s \in \mathbb{R}$, let v_s^ε be solution of*

$$\begin{cases} -\nabla\cdot\big(a(x, t_n, s)\nabla v^\varepsilon\big) = 0 & \text{in} \quad I_\delta,\\ v^\varepsilon = V & \text{on} \quad \partial I_\delta, \end{cases} \tag{4.17}$$

Define w_s^ε similarly. For any $s \in \mathbb{R}$, we define $\mathcal{A}_H(x_K, t_n, s)$ as

$$\nabla W \cdot \mathcal{A}_H(x_K, t_n, s)\nabla V = \fint_{I_\delta} \nabla w_s^\varepsilon \cdot a^\varepsilon(x, t_n, s)\nabla v_s^\varepsilon \,\mathrm{d}x \qquad \text{for } W, V \in X_H.$$

Given $\mathcal{A}_H$, we may get the following estimate for $e(\mathrm{HMM})$ as:

$$e(\mathrm{HMM}) \le C\Big(\delta + \frac{\varepsilon}{\delta}\Big).$$

The details will be given elsewhere.

Appendix A. Error estimates for the locally periodic parabolic homogenization problems

The homogenization procedure for the parabolic problem is by now well-understood, see [5, 6, 29] and the references therein. However, there are very few results concerning the error estimate for the difference between u^ε and the homogenization solution U, or the difference between u^ε and the first order approximation u_1^ε and the second order approximation u_2^ε (see (A.2) and (A.6) for the definitions). In this appendix, we shall prove such error estimates for the locally periodic parabolic homogenization problem [6, 8].

As to the locally periodic parabolic homogenization problem, the homogenization matrix $\mathcal{A}$ is given by (3.6). We have the following regularity estimate for the solution of (1.2) (see-[18]):

$$\begin{aligned}\|\nabla U\|_{L^2(\Omega)} + \|D^2 U\|_{L^2(\Omega)} &\le C(\|f\|_{L^2(\Omega)} + \|u_0\|_1),\\ \|\nabla\partial_t U\|_{L^2(\Omega)} &\le C(\|\partial_t f\|_{L^2(\Omega)} + \|u_0\|_2).\end{aligned} \tag{A.1}$$

Set

$$u_1^\varepsilon := U + \varepsilon\boldsymbol{\chi}\cdot\nabla U. \tag{A.2}$$

A direct calculation yields

$$\begin{aligned}\Big(a_{ij}\frac{\partial u_1^\varepsilon}{\partial x_j}\Big)(x,x/\varepsilon,t) &= \Big(\mathcal{A}_{ij}\frac{\partial U}{\partial x_j}\Big)(x,t) + \mathcal{G}(x,x/\varepsilon,t)\nabla U\\ &\quad + \varepsilon\Big(a_{ij}\frac{\partial \chi^k}{\partial x_j}\Big)(x,x/\varepsilon,t)\frac{\partial U}{\partial x_k} + \varepsilon(a_{ij}\chi^k)(x,x/\varepsilon,t)\frac{\partial^2 U}{\partial x_k\partial x_j},\end{aligned} \tag{A.3}$$

where $\mathcal{G} = \{g_i^j\}_{i,j=1}^d$ is defined as

$$g_i^j(x,y,t) := \Big(a_{ij} + a_{ik}\frac{\partial\chi^j}{\partial y_k}\Big)(x,y,t) - \mathcal{A}_{ij}(x,t).$$

Obviously,

$$\int_Y g_i^j(x,y,t)\,\mathrm{d}y = 0 \quad\text{and}\quad g_i^j(x,y,t)\ \text{ is periodic in}\, y.$$

Notice that $\partial_{y_i} g_i^j(x,y,t) = 0$ for $j = 1,\cdots,d$, therefore, there exists a skew-symmetric matrix $\boldsymbol{\alpha}(x,y,t) = \{\alpha_{ij}^k(x,y,t)\}_{i,j,k=1}^d$ such that

$$g_i^j(x,y,t) = \frac{\partial}{\partial y_k}\alpha_{ik}^j(x,y,t), \qquad \int_Y \alpha_{ik}^j(x,y,t)\,\mathrm{d}y = 0.$$

Thus, we obtain

$$\begin{aligned}g_i^j(x,x/\varepsilon,t)\frac{\partial U}{\partial x_j} &= \varepsilon\frac{\partial}{\partial x_k}\Big(\alpha_{ik}^j(x,x/\varepsilon,t)\frac{\partial U}{\partial x_j}\Big) - \varepsilon\alpha_{ik}^j(x,x/\varepsilon,t)\frac{\partial^2 U}{\partial x_k\partial x_j}\\ &\quad - \varepsilon\frac{\partial\alpha_{ik}^j}{\partial x_j}(x,x/\varepsilon,t)\frac{\partial U}{\partial x_j}.\end{aligned} \tag{A.4}$$

Let the corrector θ^ε be the solution of

$$\begin{cases}\partial_t\theta^\varepsilon - \nabla\cdot\big(a(x,x/\varepsilon,t)\nabla\theta^\varepsilon\big) = 0 & \text{in } \Omega,\\ \theta^\varepsilon = -\varepsilon\boldsymbol{\chi}\cdot\nabla U & \text{on } \partial D\times(0,T),\\ \theta^\varepsilon|_{t=0} = -\varepsilon\boldsymbol{\chi}|_{t=0}\cdot\nabla u_0 & \text{in } D.\end{cases} \tag{A.5}$$

Define

$$u_2^\varepsilon := u_1^\varepsilon + \theta^\varepsilon. \tag{A.6}$$

We estimate $u^\varepsilon - u_2^\varepsilon$ in the following theorem.

Theorem A.1. *Assume that $u_0 \in H^2(D)$ and $f \in H^1(0,T;L^2(D))$, then*

$$\begin{aligned}&\sup_{0<t\le T}\|(u^\varepsilon - u_2^\varepsilon)(t)\|_0 + \|\nabla(u^\varepsilon - u_2^\varepsilon)\|_{L^2(\Omega)}\\ &\qquad\le C\varepsilon(\|u_0\|_2 + \|f\|_{L^2(\Omega)} + \|\partial_t f\|_{L^2(\Omega)}).\end{aligned} \tag{A.7}$$

Proof. For any $\phi \in \mathcal{C}(0,T;L^2(D)) \cap L^2(0,T;H_0^1(D))$ with $\phi(x,0)=0$, we write the weak form of (1.2) and (A.5) as

$$\int_D \big(\phi\partial_s U + \nabla\phi\cdot\mathcal{A}\nabla U\big)\,\mathrm{d}x = \int_D f\phi\,\mathrm{d}x \quad \text{and} \quad \int_D \big(\phi\partial_s\theta^\varepsilon + \nabla\phi\cdot a^\varepsilon\nabla\theta^\varepsilon\big)\,\mathrm{d}x = 0.$$

Invoking (A.3) and the above equations, we obtain

$$\begin{aligned}&\int_D \phi\partial_s(u^\varepsilon - u_2^\varepsilon)\,\mathrm{d}x + \int_D \nabla\phi\cdot a^\varepsilon\nabla(u^\varepsilon - u_2^\varepsilon)\,\mathrm{d}x\\ &= -\varepsilon\int_D \partial_s\big(\boldsymbol{\chi}\cdot\nabla U\big)\phi\,\mathrm{d}x - \int_D \nabla\phi\cdot\mathcal{G}\nabla U\,\mathrm{d}x - \varepsilon\int_D \nabla\phi\cdot a^\varepsilon\nabla(\boldsymbol{\chi}\cdot\nabla U)\,\mathrm{d}x.\end{aligned}\tag{A.8}$$

In view of (A.4) and the fact that $\boldsymbol{\alpha}$ is a skew-symmetric matrix, we get

$$\int_D \nabla\phi\cdot\mathcal{G}\nabla U\,\mathrm{d}x = -\varepsilon\int_D \big(\nabla\phi\cdot\boldsymbol{\alpha} : D^2U + \nabla\phi\cdot(\nabla\cdot\boldsymbol{\alpha})\nabla U\big)\,\mathrm{d}x.$$

Substituting the above identity into (A.8), we get

$$\begin{aligned}&\int_D \partial_s(u^\varepsilon - u_2^\varepsilon)\phi\,\mathrm{d}x + \int_D \nabla\phi\cdot a^\varepsilon\nabla(u^\varepsilon - u_2^\varepsilon)\,\mathrm{d}x\\ &= -\varepsilon\int_D \partial_s(\boldsymbol{\chi}\cdot\nabla U)\phi\,\mathrm{d}x - \varepsilon\int_D \nabla\phi\cdot a^\varepsilon\nabla(\boldsymbol{\chi}\cdot\nabla U)\,\mathrm{d}x\\ &\quad + \varepsilon\int_D \big(\nabla\phi\cdot\boldsymbol{\alpha} : D^2U + \nabla\phi\cdot(\nabla\cdot\boldsymbol{\alpha})\nabla U\big)\,\mathrm{d}x.\end{aligned}\tag{A.9}$$

Taking $\phi = u^\varepsilon - u_2^\varepsilon$ in the above identity since $(u^\varepsilon - u_2^\varepsilon) \in H_0^1(D)$ and $(u^\varepsilon - u_2^\varepsilon)|_{t=0} = 0$, integrating from 0 to t for any $0 < t \le T$, we obtain

$$\|(u^\varepsilon - u_2^\varepsilon)(x,t)\|_0 + \Big(\int_0^t \|\nabla(u^\varepsilon - u_2^\varepsilon)\|_0^2\,\mathrm{d}s\Big)^{1/2} \le C\varepsilon\Big(\int_0^t (\|\partial_s U\|_1^2 + \|U\|_2^2)\,\mathrm{d}s\Big)^{1/2}.$$

A combination of the above inequality and the regularity estimate (A.1) gives (A.7). $\square$

In what follows, we turn to the estimates for the corrector and the first order approximation u_1^ε. No error estimates for the correctors are available to the best of the author's knowledge.

Theorem A.2. *Assume that $u_0 \in H^2(D)$ and $f, \partial_t f \in L^2(\mathcal{Q})$, then*

$$\begin{aligned}\sup_{0<t\le T} \|(u^\varepsilon - u_1^\varepsilon)(t)\|_0 &+ \|\nabla(u^\varepsilon - u_1^\varepsilon)\|_{L^2(\mathcal{Q})}\\ &\le C\sqrt{\varepsilon}(\|u_0\|_2 + \|f\|_{L^2(\mathcal{Q})} + \|\partial_t f\|_{L^2(\mathcal{Q})}),\end{aligned}\tag{A.10}$$

and

$$\sup_{0<t\le T} \|(u^\varepsilon - U)(t)\|_0 \le C\sqrt{\varepsilon}(\|u_0\|_2 + \|f\|_{L^2(\mathcal{Q})} + \|\partial_t f\|_{L^2(\mathcal{Q})})).\tag{A.11}$$

Proof. Define $\psi^\varepsilon \in \mathcal{C}_0^\infty(D)$, which equals 1 in $D/D_{2\varepsilon}$ and equals 0 in D_ε, where

$$D_\varepsilon{:} = \{\, x \in D \mid \mathrm{dist}(x, \partial D) \le \varepsilon \,\}.$$

Obviously, $|\nabla\psi^\varepsilon| \le C/\varepsilon$.

Define $w^\varepsilon := U + \varepsilon\psi^\varepsilon \boldsymbol{\chi} \cdot \nabla U$, obviously, $w^\varepsilon(x,t) \in H_0^1(D)$ for $a.e., t \in (0,T]$. A direct calculation gives

$$\sup_{0<t\le T} \|(u_1^\varepsilon - w^\varepsilon)(t)\|_{L^2(D)} + \|\nabla(u_1^\varepsilon - w^\varepsilon)\|_{L^2(\Omega)} \\ \le C\sqrt{\varepsilon}(\|u_0\|_1 + \|\nabla U\|_{L^2(\Omega)} + \|\nabla\partial_t U\|_{L^2(\Omega)} + \|D^2U\|_{L^2(\Omega)}). \tag{A.12}$$

It remains to bound $u^\varepsilon - w^\varepsilon$, as that in the proof of (A.7), we have for any $\phi \in \mathcal{C}(0,T;L^2(D)) \cap L^2(0,T;H_0^1(D))$,

$$\begin{aligned}
&\int_D \partial_s(u^\varepsilon - w^\varepsilon)\phi \,\mathrm{d}x + \int_D \nabla\phi \cdot a^\varepsilon \nabla(u^\varepsilon - w^\varepsilon)\,\mathrm{d}x \\
&= -\varepsilon \int_D \partial_s(\boldsymbol{\chi}\cdot\nabla U)\phi\psi^\varepsilon \,\mathrm{d}x + \varepsilon\int_D \nabla\phi\cdot a^\varepsilon\nabla(\boldsymbol{\chi}\cdot\nabla U)\,\mathrm{d}x \\
&\quad - \varepsilon\int_D \nabla\phi\cdot a^\varepsilon(\boldsymbol{\chi}\cdot D^2U)\psi^\varepsilon\,\mathrm{d}x - \int_D \boldsymbol{r}^\varepsilon\cdot\nabla\phi\,\mathrm{d}x,
\end{aligned}$$

where $\boldsymbol{r}^\varepsilon$ is defined by

$$\boldsymbol{r}^\varepsilon := \nabla U\cdot a^\varepsilon\nabla_y\boldsymbol{\chi}(\psi^\varepsilon - 1) + \varepsilon\nabla\psi^\varepsilon\cdot a^\varepsilon(\boldsymbol{\chi}\nabla U).$$

The terms except the last one in the right-hand side of of the above expansion can be easily bounded by $C\varepsilon(|\partial_s U|_1 + |U|_1 + |U|_2)\|\nabla\phi\|_0$.

By virtue of [21, Lemma 2.5], we get

$$|U|_{1,D_{2\varepsilon}} \le C\sqrt{\varepsilon}(|U|_1 + |U|_2).$$

We thus bound $\boldsymbol{r}^\varepsilon$ as

$$\|\boldsymbol{r}^\varepsilon\|_0 \le C|U|_{1,D_{2\varepsilon}} \le C\sqrt{\varepsilon}(|U|_1 + |U|_2).$$

Therefore, we get

$$\int_D \partial_s(u^\varepsilon - w^\varepsilon)\phi\,\mathrm{d}x + \int_D \nabla\phi\cdot a^\varepsilon\nabla(u^\varepsilon - w^\varepsilon)\,\mathrm{d}x \le C\sqrt{\varepsilon}(|\partial_s U|_1 + |U|_1 + |U|_2)\|\nabla\phi\|_0,$$

let $\phi = u^\varepsilon - w^\varepsilon$, integrating the above inequality from 0 to t, we obtain

$$\begin{aligned}
\|(u^\varepsilon - w^\varepsilon)(t)\|_0^2 + \lambda\int_0^t \|\nabla(u^\varepsilon - w^\varepsilon)\|_0^2 &\le \|(u^\varepsilon - w^\varepsilon)(x,0)\|_0^2 \\
&\quad + C\varepsilon\int_0^t (\|\nabla\partial_s U\|_0^2 + \|\nabla U\|_0^2 + \|D^2U\|_0^2)\,\mathrm{d}s.
\end{aligned}$$

Using $\|(u^\varepsilon - w^\varepsilon)(x,0)\|_0 \le C\varepsilon\|u_0\|_1$, we get

$$\max_{0<t\le T} \|(u^\varepsilon - w^\varepsilon)(t)\|_0 + \|\nabla(u^\varepsilon - w^\varepsilon)\|_{L^2(\Omega)} \\ \le C\sqrt{\varepsilon}(\|u_0\|_1 + \|\nabla\partial_t U\|_{L^2(\Omega)} + \|D^2U\|_{L^2(\Omega)}).$$

This inequality together with (A.12) and the regularity estimate (A.1) give the desired estimate (A.10). The estimate (A.11) follows from (A.7) and (A.10). □

If U is smoother, then we may improve (A.11) to $\mathcal{O}(\varepsilon)$.

Corollary A.3. *If $\nabla U \in L^\infty(\Omega)$, then we have*

$$\sup_{0<t\le T}\|(u^\varepsilon - U)(t)\|_0 \le C\varepsilon(\|u_0\|_2 + \|f\|_{L^2(\Omega)} + \|\partial_t f\|_{L^2(\Omega)} + \|\nabla U\|_{L^\infty(\Omega)}). \tag{A.13}$$

Proof. By maximum principle [20], we have

$$\max_{(x,t)\in\Omega}|\theta^\varepsilon(x,t)| \le C\varepsilon \max_{(x,t)\in\Omega}|\nabla U(x,t)|, \tag{A.14}$$

which together with (A.7) gives

$$\begin{aligned}
\sup_{0<t\le T}\|(u^\varepsilon - U)(t)\|_0 &\le \sup_{0<t\le T}\|(u^\varepsilon - u_2^\varepsilon)(t)\|_0 + \varepsilon\sup_{0<t\le T}\|(\boldsymbol{\chi}\cdot\nabla U)(t)\|_0 + \sup_{0<t\le T}\|\theta^\varepsilon(\cdot,t)\|_0\\
&\le C\varepsilon(\|u_0\|_2 + \|f\|_{L^2(\Omega)} + \|\partial_t f\|_{L^2(\Omega)}) + C\varepsilon\max_{0<t\le T}\|\nabla U(\cdot,t)\|_0\\
&\quad + C\varepsilon\|\nabla U\|_{L^\infty(\Omega)}\\
&\le C\varepsilon(\|u_0\|_2 + \|f\|_{L^2(\Omega)} + \|\partial_t f\|_{L^2(\Omega)} + \|\nabla U\|_{L^\infty(\Omega)}).
\end{aligned}$$

This gives (A.13). □

Notice that (A.14) also holds true for the case when $a^\varepsilon = a(x, x/\varepsilon, t, t/\varepsilon^2)$. Therefore, we may proceed as that in Lemma 3.3 to obtain the following estimate (A.15) for the corrector. But we cannot obtain (A.10) since we cannot obtain (A.7) by the method herein.

Corollary A.4. *For $a^\varepsilon = a(x, x/\varepsilon, t, t/\varepsilon^2)$ with $a(\cdot, y, \cdot, s)$ is periodic in y and s respectively with periods Y and 1, if $\nabla U \in L^\infty(\Omega)$, then we have*

$$\|\nabla\theta^\varepsilon\|_{L^2(\Omega)} \le C\sqrt{\varepsilon}(\|u_0\|_2 + \|f\|_{L^2(\Omega)} + \|\partial_t f\|_{L^2(\Omega)} + \|\nabla U\|_{L^\infty(\Omega)}). \tag{A.15}$$

Remark A.5. *In case of one-dimensional problem, the following error estimates are stated in* [5, pp. 43, Theorem 1].

$$\|\nabla(u^\varepsilon - u_2^\varepsilon)\|_{L^2(\Omega)} \le C(T)\varepsilon, \qquad \|\nabla(u^\varepsilon - u_1^\varepsilon)\|_{L^2(\Omega)} \le C(T)\varepsilon.$$

It is not surprising that the error estimate for the first order approximation is $\mathcal{O}(\varepsilon)$, since there is no boundary layer for one-dimensional problem.

References

1. A. Abdulle and W. E, *Finite difference heterogeneous multi-scale method for homogenization problems*, J. Comput. Phys. **191** (2003), 18-39.
2. G. Alexopoulos, *An application of homogenization theory to harmonic analysis: Harnack inequalities and Reisz transforms on Lie groups of polynomial growth*, Canad. J. Math. **44** (1992), 691-727.
3. N. André and M. Chipot, *Uniqueness and nonuniqueness for the approximation of quasiliear elliptic equations*, SIAM J. Numer. Anal. **33** (1996), 1981-1994.
4. M. Artola and G. Duvaut, *Un résultat d'homogénéisation pour une classe de problémes de diffusion non stationnaires*, Ann. Fac. Sci. Toulouse. (5) **IV** (1982), 1-28.
5. N. Bakhvalov and G. Panasenko, *Homogenisation: Averaging Processes in Periodic Media, Mathematical Problems in the Mechanics of Materials*, Kluwer Academic Publishers, 1989.
6. A. Bensoussan, J. L. Lions and G. C. Papanicolaou, *Asymptotic Analysis for Periodic Structures*, North-Holland, Amsterdam, 1978.
7. L. Boccardo and F. Murat, *Homogénéisation de problémes quasi-linéaires*, Publ. IRMA, Lille. **3** (1981), no. 7, 13-51.
8. S. Brahim-Otsmane, G. A. Francfort and F. Murat, *Correctors for the homogenization of the wave and heat equations*, J. Math. Pures Appl. **71** (1992), 197-231.
9. S. Q. Chen, W. E and C-W. Shu, *The heterogeneous multiscale method based on the discontinuous Galerkin method for hyperbolic and parabolic problems*, Multiscale Model. Simul. **3** (2005), 871-894.
10. P. G. Ciarlet, *The Finite Element Method for the Elliptic Problems*, North-Holland, Amsterdam, 1978.

11. F. Colombini and S. Spagnolo, *Sur la convergence de solutions d'équations paraboliques*, J. Math. Pures Appl. **56** (1977), 263-305.
12. A. Dall'aglio and F. Murat, *A corrector result for H-converging parabolic problems with time-dependent coefficients*, Ann. Scuola Norm. Sup. Pisa C1. Sci. **XXV** (1997), 329-373.
13. J. Douglas and T. Dupont, *Galerkin methods for parabolic equations*, SIAM J. Numer. Anal. **7** (1970), 575-626.
14. W. E, *Analysis of the heterogeneous multiscale method for ordinary differential equations*, Commun. Math. Sci. **1** (2003), 423-436.
15. W. E and B. Engquist, *The heterogeneous multi-scale methods*, Commun. Math. Sci. **1** (2003), 87-132.
16. W. E and B. Engquist, *Multiscale modeling and computation*, Notices Amer. Math. Soc. **50** (2003), 1062-1070.
17. W. E, P. B. Ming and P. W. Zhang, *Analysis of the heterogeneous multiscale method for elliptic homogenization problems*, J. Amer. Math. Soc. **18** (2005), 121-156.
18. L. C. Evans, *Partial Differential Equations*, American Mathmatics Society, Providence, Rhode Island: AMS, 1998.
19. J. Garcıa-Azorero, C. E. Gutierrez and I. Peral, *Homogenization of quasilinear parabolic equations in periodic media*, Comm. Partial Differential Equations. **28** (2003), 1887-1910.
20. O. A. Ladyzhenskaya, V. A. Solonnikov and N. N. Ural'tseva, *Linear and Quasilinear Equations of Parabolic Type*, Translations of Mathematics Monographys, Vol 25, Providence, Rhode Island: AMS, 1968.
21. O. A. Oleinik, A. S. Shamaev and G. A. Yosifian, *Mathematical Problems in Elasticity and Homogenization*, North-Holland, Amsterdam, 1992.
22. S. Spagnolo, *Sul limite delle soluzioni di problemi di Cauchy relativi all'equazione del calore*, Ann. Scuola Norm. Sup. Pisa (3), **21** (1967), 657-699.
23. S. Spagnolo, *Sulla convergenza di soluzioni di equationi paraboliche ed ellittiche*, Ann. Scuola Norm. Sup. Pisa (3), **22** (1968), 571-597.
24. S. Spagnolo, *Convergence of parabolic equations*, Boll. Un. Mat. Ital. B (5), **14** (1977), 547-568.
25. L. Tartar, *H-Convergence*, Course Peccot, Collége de France, March 1977. Partially written by F. Murat. Séminaire d'Analyse Fonctionnelle et Numérique de l'Université d'Alger, 1977–1978. English Translation: F. Murat and L. Tartar: *H-Convergence*, in *Topics in the Mathematical Modeling of Composite Materials*, A. Cherkaev and R. Kohn, eds, Birkhäuser, Boston, 1997, pp. 21-43.
26. V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Springer-Verlag Berlin Heidelberg, 1997.
27. V. V. Zhikov, S. M. Kozlov and O. A. Oleinik, *G-convergence of parabolic operators*, Uspekhi Mat. Nauk **36: 1** (1981), 11-58. English Translation: Russ. Math. Surv. **36** (1981), 9-60.
28. V. V. Zhikov, S. M. Kozlov and O. A. Oleinik, *Averaging of parabolic operators*, Trudy Mosk. Mat. O.-va **45** (1982), 182-236. English Translation: Trans. Moscow Math. Soc. **45** (1984), 189-241.
29. V. V. Zhikov, S. M. Kozlov and O. A. Oleinik, *Homogenization of Differential Operators and Integral Functionals*, Springer-verlag Berlin Heidelberg, 1994.

Institute of Computational Mathematics and Scientific/Engineering Computing, AMSS, Chinese Academy of Sciences, No. 55 Zhong-Guan-Cun East Road, Beijing, 100080, China

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

LMAM and School of Mathematical Sciences, Peking University, Beijing, 100871, China

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