Numerical Approximations to Multiscale Solutions in Partial Differential Equations

Thomas Y. Hou*

Applied Mathematics, 217-50, Caltech, Pasadena, CA 91125, USA. Email: hou@ama.caltech.edu.

Abstract. Many problems of fundamental and practical importance have multiple scale solutions. The direct numerical solution of multiple scale problems is difficult to obtain even with modern supercomputers. The major difficulty of direct solutions is the scale of computation. The ratio between the largest scale and the smallest scale could be as large as 10^5 in each space dimension. From an engineering perspective, it is often sufficient to predict the macroscopic properties of the multiple-scale systems, such as the effective conductivity, elastic moduli, permeability, and eddy diffusivity. Therefore, it is desirable to develop a method that captures the small scale effect on the large scales, but does not require resolving all the small scale features. This paper reviews some of the recent advances in developing systematic multiscale methods such as homogenization, numerical samplings, multiscale finite element methods, variational multiscale methods, and wavelets based homogenization. This paper is not intended to be a detailed survey and the discussion is limited by both the taste and expertise of the author.

1 Introduction

Many problems of fundamental and practical importance have multiple scale solutions. Composite materials, porous media, and turbulent transport in high Reynolds number flows are examples of this type. A complete analysis of these problems is extremely difficult. For example, the difficulty in analyzing groundwater transport is mainly caused by the heterogeneity of subsurface formations spanning over many scales. This heterogeneity is often represented by the multiscale fluctuations in the permeability of media. For composite materials, the dispersed phases (particles or fibers), which may be randomly distributed in the matrix, give rise to fluctuations in the thermal or electrical conductivity; moreover, the conductivity is usually discontinuous across the phase boundaries. In turbulent transport problems, the convective velocity field fluctuates randomly and contains many scales depending on the Reynolds number of the flow.

The direct numerical solution of multiple scale problems is difficult even with the advent of supercomputers. The major difficulty of direct solutions is the scale of computation. For groundwater simulations, it is common that

* Research was in part supported by a grant DMS-0073916 from the National Science Foundation

millions of grid blocks are involved, with each block having a dimension of tens of meters, whereas the permeability measured from cores is at a scale of several centimeters. This gives more than 10^5 degrees of freedom per spatial dimension in the computation. Therefore, a tremendous amount of computer memory and CPU time are required, and this can easily exceed the limit of today's computing resources. The situation can be relieved to some degree by parallel computing; however, the size of the discrete problem is *not* reduced. The load is merely shared by more processors with more memory. Whenever one can afford to resolve all the small scale features of a physical problem, direct solutions provide quantitative information of the physical processes at all scales. On the other hand, from an engineering perspective, it is often sufficient to predict the macroscopic properties of the multiscale systems, such as the effective conductivity, elastic moduli, permeability, and eddy diffusivity. Therefore, it is desirable to develop a method that captures the small scale effect on the large scales, but does not require resolving all the small scale features.

The purpose of these lecture notes is to review some recent advances in developing multiscale numerical methods that capture the small scale effect on the large scales, but do not require resolving all the small scale features. The ultimate goal is to develop a general method that works for problems with continuous spectrum of scales. Substantial progress has been made in recent years by combining modern mathematical techniques such as homogenization, numerical samplings, and multiresolution. My lectures can be roughly divided into five parts. In Section 2, I will review some homogenization theory for elliptic and hyperbolic equations as well as for incompressible flows. This homogenization theory provides the critical guideline for designing effective multiscale methods. Section 3 is devoted to numerical homogenization for semilinear hyperbolic systems using particle methods and sampling techniques. For hyperbolic systems, it is important to compute the advection of small scale information accurately and account for the nonlinear interaction properly. We also need to avoid certain resonant sampling of the grid in order to obtain convergence. In Section 4, we focus on some recent developments of numerical homogenization based on the multiscale finite element methods. We also discuss the issue of upscaling one-phase and two-phase flows through heterogeneous porous media. In Sections 5 and 6, I review the main ideas behind the wavelet-based numerical homogenization method and the variational multiscale method. There are many other multiscale methods which we will not cover due to the limited scope of these lectures. The above methods are chosen because they are similar philosophically and the materials complement each other very well. This paper is not intended to be a detailed survey of all available multiscale methods. The discussion is limited by scope of the lectures and expertise of the author.

2 Review of Homogenization Theory

In this section, we will review some classical homogenization theory for elliptic and hyperbolic PDEs. This homogenization theory will play an essential role in designing effective multiscale numerical methods for partial differential equations with multiscale solutions.

2.1 Homogenization Theory for Elliptic Problems

Consider the second order elliptic equation

$$\mathcal{L}(u_\varepsilon) \equiv -\frac{\partial}{\partial x_i}\left(a_{ij}\left(x/\varepsilon\right)\frac{\partial}{\partial x_j}\right)u_\varepsilon + a_0(x/\varepsilon)u_\varepsilon = f,\ u_\varepsilon|_{\partial\Omega} = 0, \tag{2.1}$$

where $a_{ij}(y)$ and $a_0(y)$ are Y-periodic and satisfy $a_{ij}(y)\xi_i\xi_j \geq \alpha\xi_i\xi_i$, with $\alpha > 0$, and $a_0 > \alpha_0 > 0$. Here we have used the Einstein summation notation, i.e. repeated index means summation with respect to that index.

This model equation represents a common difficulty shared by several physical problems. For porous media, it is the pressure equation through Darcy's law, the coefficient a_ε representing the permeability tensor. For composite materials, it is the steady heat conduction equation and the coefficient a_ε represents the thermal conductivity. For steady transport problems, it is a symmetrized form of the governing equation. In this case, the coefficient a_ε is a combination of transport velocity and viscosity tensor.

Homogenization theory is to study the limiting behavior $u_\varepsilon \to u$ as $\varepsilon \to 0$. The main task is to find the homogenized coefficients, a^*_{ij} and a^*_0, and the *homogenized equation* for the limiting solution u

$$-\frac{\partial}{\partial x_i}\left(a^*_{ij}\frac{\partial}{\partial x_j}\right)u + a^*_0 u = f, \quad u|_{\partial\Omega} = 0. \tag{2.2}$$

Define the L^2 and H^1 norms over Ω as follows

$$\|v\|_0^2 = \int_\Omega |v|^2\,dx, \quad \|v\|_1^2 = \|v\|_0^2 + \|\nabla v\|_0^2. \tag{2.3}$$

Further, we define the bilinear form

$$a^\varepsilon(u,v) = \int_\Omega a^\varepsilon_{i,j}(x)\frac{\partial u}{\partial x_j}\frac{\partial v}{\partial x_i}\,dx + \int_\Omega a^\varepsilon_0 uv\,dx. \tag{2.4}$$

It is easy to show that

$$c_1\|u\|_1^2 \leq a^\varepsilon(u,u) \leq c_2\|u\|_1^2, \tag{2.5}$$

with $c_1 = \min(\alpha,\alpha_0)$, $c_2 = \max(\|a_{ij}\|_\infty, \|a_0\|_\infty)$.

The elliptic problem can also be formulated as a variational problem: find $u_\varepsilon \in H^1_0$

$$a^\varepsilon(u_\varepsilon, v) = (f,v), \quad \text{for all} \quad v \in H^1_0(\Omega), \tag{2.6}$$

where (f,v) is the usual L^2 inner product, $\int_\Omega fv\,dx$.

Special Case: One-Dimensional Problem Let $\Omega = (x_0, x_1)$ and take $a_0 = 0$. We have

$$-\frac{d}{dx}\left(a(x/\varepsilon)\frac{du_\varepsilon}{dx}\right) = f, \quad \text{in} \quad \Omega\,, \tag{2.7}$$

where $u_\varepsilon(x_0) = u_\varepsilon(x_1) = 0$, and $a(y) > \alpha_0 > 0$ is y-periodic with period y_0.

By taking $v = u_\varepsilon$ in the bilinear form, we have

$$\|u_\varepsilon\|_1 \leq c.$$

Therefore one can extract a subsequence, still denoted by u_ε, such that

$$u_\varepsilon \rightharpoonup u \quad in \quad H_0^1(\Omega) \quad \text{weakly.} \tag{2.8}$$

On the other hand, we notice that

$$a^\varepsilon \rightharpoonup m(a) = \frac{1}{y_0}\int_0^{y_0} a(y)\,dy \quad in \quad L^\infty(\Omega) \quad \text{weak star.} \tag{2.9}$$

It is tempting to conclude that u satisfies:

$$-\frac{d}{dx}\left(m(a)\frac{du}{dx}\right) = f,$$

where $m(a) = \frac{1}{y_0}\int_0^{y_0} a(y)\,dy$ is the arithmetic mean of a. However, this is *not* true. To derive the correct answer, we introduce

$$\xi^\varepsilon = a^\varepsilon \frac{du^\varepsilon}{dx}.$$

Since a^ε is bounded, and u_x^ε is bounded in $L^2(\Omega)$, so ξ^ε is bounded in $L^2(\Omega)$. Moreover, since $-\frac{d\xi^\varepsilon}{dx} = f$, we have $\xi^\varepsilon \in H^1(\Omega)$. Thus we get

$$\xi^\varepsilon \to \xi \quad \text{in} \quad L^2(\Omega) \quad \text{strongly,}$$

so that

$$\frac{1}{a^\varepsilon}\xi^\varepsilon \to m(1/a)\xi \quad \text{in} \quad L^2(\Omega) \quad weakly.$$

Further, we note that $\frac{1}{a^\varepsilon}\xi^\varepsilon = \frac{du^\varepsilon}{dx}$. Therefore, we arrive at

$$\frac{du}{dx} = m(1/a)\xi.$$

On the other hand, $-\frac{d\xi^\varepsilon}{dx} = f$ implies $-\frac{d\xi}{dx} = f$. This gives

$$-\frac{d}{dx}\left(\frac{1}{m(1/a)}\frac{du}{dx}\right) = f. \tag{2.10}$$

This is the correct homogenized equation for u. Note that $a^* = \frac{1}{m(1/a)}$ is the harmonic average of a^ε. It is in general not equal to the arithmetic average $\overline{a^\varepsilon} = m(a)$.

Multiscale Asymptotic Expansions. The above analysis does not generalize to multi-dimensions. In this subsection, we introduce the multiscale expansion technique in deriving homogenized equations. This technique is very effective and can be used in a number of applications.

We shall look for $u_\varepsilon(x)$ in the form of asymptotic expansion

$$u_\varepsilon(x) = u_0(x, x/\varepsilon) + \varepsilon u_1(x, x/\varepsilon) + \varepsilon^2 u_2(x, x/\varepsilon) + \cdots, \tag{2.11}$$

where the functions $u_j(x, y)$ are Y-periodic in y.

Denote by A^ε the second order elliptic operator

$$A^\varepsilon = -\frac{\partial}{\partial x_i}\left(a_{ij}\left(x/\varepsilon\right)\frac{\partial}{\partial x_j}\right). \tag{2.12}$$

When differentiating a function $\phi(x, x/\varepsilon)$ with respect to x, we have

$$\frac{\partial}{\partial x_j} = \frac{\partial}{\partial x_j} + \frac{1}{\varepsilon}\frac{\partial}{\partial y_j},$$

where y is evaluated at $y = x/\varepsilon$. With this notation, we can expand A^ε as follows

$$A^\varepsilon = \varepsilon^{-2}A_1 + \varepsilon^{-1}A_2 + \varepsilon^0 A_3, \tag{2.13}$$

where

$$A_1 = -\frac{\partial}{\partial y_i}\left(a_{ij}(y)\frac{\partial}{\partial y_j}\right), \tag{2.14}$$

$$A_2 = -\frac{\partial}{\partial y_i}\left(a_{ij}(y)\frac{\partial}{\partial x_j}\right) - \frac{\partial}{\partial x_i}\left(a_{ij}(y)\frac{\partial}{\partial y_j}\right), \tag{2.15}$$

$$A_3 = -\frac{\partial}{\partial x_i}\left(a_{ij}(y)\frac{\partial}{\partial x_j}\right) + a_0\ . \tag{2.16}$$

Substituting the expansions for u_ε and A^ε into $A^\varepsilon u_\varepsilon = f$, and equating the terms of the same power, we get

$$A_1 u_0 = 0, \tag{2.17}$$

$$A_1 u_1 + A_2 u_0 = 0, \tag{2.18}$$

$$A_1 u_2 + A_2 u_1 + A_3 u_0 = f. \tag{2.19}$$

Equation (2.17) can be written as

$$-\frac{\partial}{\partial y_i}\left(a_{ij}(y)\frac{\partial}{\partial y_j}\right)u_0(x, y) = 0, \tag{2.20}$$

where u_0 is periodic in y. The theory of second order elliptic PDEs [35] implies that $u_0(x, y)$ is independent of y, i.e. $u_0(x, y) = u_0(x)$. This simplifies equation (2.18) for u_1,

$$-\frac{\partial}{\partial y_i}\left(a_{ij}(y)\frac{\partial}{\partial y_j}\right)u_1 = \left(\frac{\partial}{\partial y_i}a_{ij}(y)\right)\frac{\partial u}{\partial x_j}(x).$$

Define $\chi^j = \chi^j(y)$ as the solution to the following *cell problem*

$$-\frac{\partial}{\partial y_i}\left(a_{ij}(y)\frac{\partial}{\partial y_j}\right)\chi^j = \frac{\partial}{\partial y_i}a_{ij}(y)\ , \tag{2.21}$$

where χ^j is Y-periodic. The general solution of equation (2.18) for u_1 is then given by

$$u_1(x,y) = -\chi^j(y)\frac{\partial u}{\partial x_j}(x) + \tilde{u}_1(x)\ . \tag{2.22}$$

Finally, we note that the equation for u_2 is given by

$$\frac{\partial}{\partial y_i}\left(a_{ij}(y)\frac{\partial}{\partial y_j}\right)u_2 = A_2u_1 + A_3u_0 - f\ . \tag{2.23}$$

The solvability condition implies that the right hand side of (2.23) must have mean zero in y, i.e.

$$\int_Y (A_2u_1 + A_3u_0 - f)\,dy = 0.$$

This solvability condition for second order elliptic PDEs with periodic boundary condition [35] requires that the right hand side of equation (2.23) have mean zero with respect to the fast variable y. This solvability condition gives rise to the homogenized equation for u:

$$-\frac{\partial}{\partial x_i}\left(a^*_{ij}\frac{\partial}{\partial x_j}\right)u + m(a_0)u = f\ , \tag{2.24}$$

where $m(a_0) = \frac{1}{|Y|}\int_Y a_0(y)\,dy$ and

$$a^*_{ij} = \frac{1}{|Y|}\left(\int_Y (a_{ij} - a_{ik}\frac{\partial\chi^j}{\partial y_k})\,dy\right). \tag{2.25}$$

Justification of formal expansions The above multiscale expansion is based on a formal asymptotic analysis. However, we can justify its convergence rigorously.

Let $z_\varepsilon = u_\varepsilon - (u + \varepsilon u_1 + \varepsilon^2 u_2)$. Applying A^ε to z_ε, we get

$$A^\varepsilon z_\varepsilon = -\varepsilon r_\varepsilon\ ,$$

where $r_\varepsilon = A_2u_2 + A_3u_1 + \varepsilon A_3u_2$. If f is smooth enough, so is u_2. Thus we have $\|r_\varepsilon\|_\infty \le c$.

On the other hand, we have

$$z_\varepsilon|_{\partial\Omega} = -(\varepsilon u_1 + \varepsilon^2 u_2)|_{\partial\Omega}.$$

Thus, we obtain

$$\|z_\varepsilon\|_{L^\infty(\partial\Omega)} \le c\varepsilon.$$

It follows from the maximum principle [35] that

$$\|z_\varepsilon\|_{L^\infty(\Omega)} \leq c\varepsilon$$

and therefore we conclude that

$$\|u_\varepsilon - u\|_{L^\infty(\Omega)} \leq c\varepsilon.$$

Boundary Corrections The above asymptotic expansion does not take into account the boundary condition of the original elliptic PDEs. If we add a boundary correction, we can obtain higher order approximations.

Let $\theta_\varepsilon \in H^1(\Omega)$ denote the solution to

$$\nabla_x \cdot a^\varepsilon \nabla_x \theta_\varepsilon = 0 \text{ in } \Omega, \quad \theta_\varepsilon = u_1(x, x/\varepsilon) \text{ on } \partial\Omega.$$

Then we have

$$(u_\varepsilon - (u + \varepsilon u_1(x, x/\varepsilon) - \varepsilon\theta_\varepsilon))\,|_{\partial\Omega} = 0.$$

Moskow and Vogelius [52] have shown that

$$\|u_\varepsilon - u - \varepsilon u_1(x, x/\varepsilon) + \varepsilon\theta_\varepsilon\|_0 \leq C_\omega \varepsilon^{1+\omega} \|u\|_{2+\omega}, \tag{2.26}$$

$$\|u_\varepsilon - u - \varepsilon u_1(x, x/\varepsilon) + \varepsilon\theta_\varepsilon\|_1 \leq C\varepsilon \|u\|_2, \tag{2.27}$$

where we assume $u \in H^{2+\omega}(\Omega)$ with $0 \leq \omega \leq 1$, and Ω is assumed to be a bounded, convex curvilinear polygon of class C^∞. This improved estimate will be used in the convergence analysis of the multiscale finite element method to be presented in Section 4.

2.2 Homogenization for hyperbolic problems

In this subsection, we will review some homogenization theory for semilinear hyperbolic systems. As we will see below, homogenization for hyperbolic problems is very different from that for elliptic problems. The phenomena are also very rich.

Consider the semilinear Carleman equations [14]:

$$\frac{\partial u_\varepsilon}{\partial t} + \frac{\partial u_\varepsilon}{\partial x} = v_\varepsilon^2 - u_\varepsilon^2,$$

$$\frac{\partial v_\varepsilon}{\partial t} - \frac{\partial v_\varepsilon}{\partial x} = u_\varepsilon^2 - v_\varepsilon^2,$$

with oscillatory initial data, $u_\varepsilon(x, 0) = u_0^\varepsilon(x)$, $v_\varepsilon(x, 0) = v_0^\varepsilon(x)$.

Assume that the initial conditions are positive and bounded. Then it can be shown that there exists a unique *bounded* solution for all times. Thus we can extract a subsequence of u_ε and v_ε such that $u_\varepsilon \rightharpoonup u$ and $v_\varepsilon \rightharpoonup v$ as $\varepsilon \to 0$.

Denote u_m as the weak limit of u_ε^m, and v_m as the weak limit of v_ε^m. By taking the weak limit of both sides of the equations, we get

$$\frac{\partial u_1}{\partial t} + \frac{\partial u_1}{\partial x} = v_2 - u_2,$$
$$\frac{\partial v_1}{\partial t} - \frac{\partial v_1}{\partial x} = u_2 - v_2.$$

By multiplying the Carleman equations by u_ε and v_ε respectively, we get

$$\frac{\partial u_\varepsilon^2}{\partial t} + \frac{\partial u_\varepsilon^2}{\partial x} = 2u_\varepsilon v_\varepsilon^2 - 2u_\varepsilon^3,$$
$$\frac{\partial v_\varepsilon^2}{\partial t} + \frac{\partial v_\varepsilon^2}{\partial x} = 2v_\varepsilon u_\varepsilon^2 - 2v_\varepsilon^3.$$

Thus the weak limit of u_ε^2 depends on the weak limit of u_ε^3 and the weak limit of $u_\varepsilon v_\varepsilon^2$.

Denote by $\overline{w_\varepsilon}$ as the weak limit of w_ε. To obtain a closure, we would like to express $\overline{u_\varepsilon v_\varepsilon^2}$ in terms of the product $\overline{u_\varepsilon}$ and $\overline{v_\varepsilon^2}$. This is *not* possible in general. In this particular case, we can use the Div-Curl Lemma [53,54,58] to obtain a closure.

The Div-Curl Lemma. *Let Ω be an open set of $\mathbb{R}^N$ and u_ε and v_ε be two sequences such that*

$$\mathbf{u}_\varepsilon \rightharpoonup \mathbf{u}, \quad in \quad \big(L^2(\Omega)\big)^N \quad weakly,$$
$$\mathbf{v}_\varepsilon \rightharpoonup \mathbf{v}, \quad in \quad \big(L^2(\Omega)\big)^N \quad weakly.$$

Further, we assume that

$$\operatorname{div} \mathbf{u}_\varepsilon \quad is\ bounded\ in\ L^2(\Omega)(\ or\ compact\ in\ H^{-1}(\Omega))\ ,$$
$$\mathbf{curl}\, \mathbf{v}_\varepsilon \quad is\ bounded\ in \big(L^2(\Omega)\big)^{N^2}\ (or\ compact\ in\ \big(H^{-1}(\Omega)\big)^{N^2}\ ).$$

Let $\langle \cdot, \cdot \rangle$ denote the inner product in $\mathbb{R}^N$, i.e.

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^N u_i v_i.$$

Then we have

$$\langle \mathbf{u}_\varepsilon \cdot \mathbf{v}_\varepsilon \rangle \rightharpoonup \langle \mathbf{u} \cdot \mathbf{v} \rangle \quad \text{weakly.} \tag{2.28}$$

Remark 2.1. We remark that the Div-Curl Lemma is the simplest form of the more general Compensated Compactness Theory developed by Tartar [58] and Murat [53,54].

Applying the Div-Curl Lemma to $(u_\varepsilon, u_\varepsilon)$ and $(v_\varepsilon^2, v_\varepsilon^2)$ in the space-time domain, one can show that $\overline{u_\varepsilon v_\varepsilon^2} = \overline{u_\varepsilon}\ \overline{v_\varepsilon^2}$. Similarly, one can show that $\overline{u_\varepsilon^2 v_\varepsilon} =$

$\overline{u_\varepsilon^2}\ \overline{v_\varepsilon}$. Using this fact, Tartar [59] obtained the following infinite hyperbolic system for u_m and v_m [59]:

$$\frac{\partial u_m}{\partial t} + \frac{\partial u_m}{\partial x} = m u_{m-1} v_2 - m u_{m+1},$$
$$\frac{\partial v_m}{\partial t} - \frac{\partial v_m}{\partial x} = m v_{m-1} u_2 - m v_{m+1}.$$

Note that the weak limit of u_ε^m, u_m, depends on the weak limit of u_ε^{m+1}, u_{m+1}. Similarly, v_m depends on v_{m+1}. Thus one cannot obtain a closed system for the weak limits u_ε and v_ε by a finite system. This is a generic phenomenon for nonlinear partial differential equations with microstructure. It is often referred to as the closure problem. On the other hand, for the Carleman equations, Tartar showed that the infinite system is hyperbolic and the system is well-posed.

The situation is very different for a 3×3 system of Broadwell type [13]:

$$\frac{\partial u_\varepsilon}{\partial t} + \frac{\partial u_\varepsilon}{\partial x} = w_\varepsilon^2 - u_\varepsilon v_\varepsilon, \tag{2.29}$$
$$\frac{\partial v_\varepsilon}{\partial t} - \frac{\partial v_\varepsilon}{\partial x} = w_\varepsilon^2 - u_\varepsilon v_\varepsilon, \tag{2.30}$$
$$\frac{\partial w_\varepsilon}{\partial t} + \alpha \frac{\partial w_\varepsilon}{\partial x} = u_\varepsilon v_\varepsilon - w_\varepsilon^2, \tag{2.31}$$

with oscillatory initial data, $u_\varepsilon(x,0) = u_0^\varepsilon(x)$, $v_\varepsilon(x,0) = v_0^\varepsilon(x)$ and $w_\varepsilon(x,0) = w_0^\varepsilon(x)$. When $\alpha = 0$, the above system reduces to the original Broadwell model. We will refer to the above system as the generalized Broadwell model.

Note that in the generalized Broadwell model, the right hand side of the w-equation depends on the product of uv. If we try to obtain an evolution equation for w_ε^2, it will depend on the triple product $u_\varepsilon v_\varepsilon w_\varepsilon$. The Div-Curl Lemma cannot be used here to characterize the weak limit of this triple product in terms of the weak limits of u_ε, v_ε and w_ε.

Assume the initial oscillations are periodic, i.e.

$$u_0^\varepsilon = u_0(x, x/\varepsilon),\ v_0^\varepsilon = v_0(x, x/\varepsilon),\ w_0^\varepsilon = w_0(x, x/\varepsilon).$$

where $u_0(x,y),\ v_0(x,y),\ w_0(x,y)$ are 1-periodic in y.

There are two cases to consider.

Case 1. $\alpha = m/n$ is a rational number. Let $\{U(x,y,t), V(x,y,t), W(x,y,t)\}$ be the homogenized solution which satisfies

$$U_t + U_x = \textstyle\int_0^1 W^2\, dy - U \int_0^1 V\, dy,$$
$$V_t - V_x = \textstyle\int_0^1 W^2\, dy - U \int_0^1 V\, dy,$$
$$W_t + \alpha W_x = -W^2 + \frac{1}{n} \int_0^n U(x, y + (\alpha - 1)z, t) V(x, y + (\alpha + 1)z, t)\, dz,$$

where $U|_{t=0} = u_0(x,y)$, $V|_{t=0} = v_0(x,y)$ and $W|_{t=0} = w_0(x,y)$. Then we have

$$\|u_\varepsilon(x,t) - U(x, \tfrac{x-t}{\varepsilon}, t)\|_{L^\infty} \le C\varepsilon,$$
$$\|v_\varepsilon(x,t) - V(x, \tfrac{x+t}{\varepsilon}, t)\|_{L^\infty} \le C\varepsilon,$$
$$\|w_\varepsilon(x,t) - W(x, \tfrac{x-\alpha t}{\varepsilon}, t)\|_{L^\infty} \le C\varepsilon.$$

Case 2. α is an irrational number. Let $\{U(x,y,t), V(x,y,t), W(x,y,t)\}$ be the homogenized solution which satisfies

$$U_t + U_x = \int_0^1 W^2\,dy - U\int_0^1 V\,dy,$$
$$V_t - V_x = \int_0^1 W^2\,dy - U\int_0^1 V\,dy,$$
$$W_t + \alpha W_x = -W^2 + \left(\int_0^1 U\,dy\right)\left(\int_0^1 V\,dy\right).$$

where $U|_{t=0} = u_0(x,y)$, $V|_{t=0} = v_0(x,y)$ and $W|_{t=0} = w_0(x,y)$. Then we have

$$\|u_\varepsilon(x,t) - U(x, \tfrac{x-t}{\varepsilon}, t)\|_{L^\infty} \le C\varepsilon,$$
$$\|v_\varepsilon(x,t) - V(x, \tfrac{x+t}{\varepsilon}, t)\|_{L^\infty} \le C\varepsilon,$$
$$\|w_\varepsilon(x,t) - W(x, \tfrac{x-\alpha t}{\varepsilon}, t)\|_{L^\infty} \le C\varepsilon.$$

We refer the reader to [37] for the proof of the above results.

Note that when α is a rational number, the interaction of u_ε and v_ε can generate a high frequency contribution to w_ε. This is *not* the case when α is an irrational number. The rational α case corresponds to a resonance interaction.

The derivation and analysis of the above results rely on the following two Lemmas:

Lemma 2.1. *Let $f(x), g(x,y) \in C^1$. Assume that $g(x,y)$ is n-periodic in y, then we have*

$$\int_a^b f(x)g(x, x/\varepsilon)\,dx = \int_a^b f(x)\left(\frac{1}{n}\int_0^n g(x,y)\,dy\right)dx + O(\varepsilon).$$

Lemma 2.2. *Let $f(x,y,z) \in C^1$. Assume that $f(x,y,z)$ is 1-periodic in y and z. If γ_2/γ_1 is an irrational number, then we have*

$$\int_a^b f\left(x, \tfrac{x_1+\gamma_1 x}{\varepsilon}, \tfrac{x_2+\gamma_1 x}{\varepsilon}\right)dx = \int_a^b \left(\int_0^1\int_0^1 f(x,y,z)\,dy\,dz\right)dx + O(\varepsilon).$$

The proof uses some basic ergodic theory. It can be seen easily by expanding in Fourier series in the periodic variables [37]. For the sake of completeness, we present a simple proof of the above homogenization result for the case of $\alpha = 0$ in the next subsection.

Homogenization of the Broadwell Model In this subsection, we give a simple proof of the homogenization result in the special case of $\alpha = 0$. The homogenized equations can be derived by multiscale asymptotic expansions [50].

Consider the Broadwell model

$$\partial_t u + \partial_x u = w^2 - uv \text{ in } \mathbb{R} \times (0,T), \tag{2.32}$$

$$\partial_t v - \partial_x v = w^2 - uv \text{ in } \mathbb{R} \times (0,T), \tag{2.33}$$

$$\partial_t w = uv - w^2 \text{ in } \mathbb{R} \times (0,T), \tag{2.34}$$

with oscillatory initial values

$$u(x,0) = u_0(x, \tfrac{x}{\varepsilon}), \quad v(x,0) = v_0(x, \tfrac{x}{\varepsilon}), \quad w(x,0) = w_0(x, \tfrac{x}{\varepsilon}), \tag{2.35}$$

where $u_0(x,y), v_0(x,y), w_0(x,y)$ are 1-periodic in y. We introduce an extra variable, y, to describe the fast variable, x/ε. Let the solution of the homogenized equation be $\{U(x,y,t), V(x,y,t), W(x,y,t)\}$ which satisfies

$$\partial_t U + \partial_x U + U \int_0^1 V\, dy - \int_0^1 W^2\, dy = 0 \text{ in } \mathbb{R} \times (0,T), \tag{2.36}$$

$$\partial_t V - \partial_x V + V \int_0^1 U\, dy - \int_0^1 W^2\, dy = 0 \text{ in } \mathbb{R} \times (0,T), \tag{2.37}$$

$$\partial_t W + W^2 - \int_0^1 U(x, y-z, t) V(x, y+z, t)\, dz = 0 \text{ in } \mathbb{R} \times (0,T), \tag{2.38}$$

with initial values given by

$$U(x,y,0) = u_0(x,y), \quad V(x,y,0) = v_0(x,y), \quad W(x,y,0) = w_0(x,y). \tag{2.39}$$

Note that $U(x,y,t), V(x,y,t), W(x,y,t)$ are 1-periodic in y and the system (2.36)-(2.39) is a set of partial differential equations in (x,t) with $y \in [0,1]$ as a parameter. The global existence of the systems (2.32)-(2.35) and (2.36)-(2.39) has been established, see the references cited in [32].

Theorem 2.1. *Let (u,v,w) and (U,V,W) be the solutions of the systems (2.32)-(2.35) and (2.36)-(2.39), respectively. Then we have the following error estimate*

$$\max_{0 \le t \le T} E(t) \le \left[5(M(T)^2 + 2TK(T)M(T)) \exp(6M(T)T) \right] \varepsilon := C_1(T)\varepsilon, \tag{2.40}$$

where the error function $E(t)$ is given by

$$E(t) = \max_{x \in \mathbb{R}} \Big\{ \Big| u(x,t) - U(x, \tfrac{x-t}{\varepsilon}, t) \Big| + \Big| v(x,t) - V(x, \tfrac{x+t}{\varepsilon}, t) \Big| + \Big| w(x,t) - W(x, \tfrac{x}{\varepsilon}, t) \Big| \Big\}$$

and the constants $M(T)$ and $K(T)$ are given by

$$M(T) = \max_{(x,y,t)\in\mathbb{R}\times[0,1]\times[0,T]} \Big(|u|, |v|, |w|, |U|, |V|, |W| \Big), \tag{2.41}$$

$$N(T) = \max_{(x,y,t)\in\mathbb{R}\times[0,1]\times[0,T]} \Big(|\partial_x U|, |\partial_t U|, |\partial_x V|, |\partial_t V|, |\partial_x W|, |\partial_t W| \Big). \tag{2.42}$$

This homogenization result was first obtained by McLaughlin, Papanicolaou and Tartar using an L^p norm estimate ($0 < p < \infty$) [50]. Since we need an L^∞ norm estimate in the convergence analysis of our particle method, we give another proof of this result in L^∞ norm. As a first step, we prove the following lemma.

Lemma 2.3. *Let $g(x,y) \in C^1(\mathbb{R}\times[0,1])$ be 1-periodic in y and satisfy the relation $\int_0^1 g(x,y)\,dy = 0$. Then for any $\varepsilon > 0$ and for any constants a and b, the following estimate holds*

$$\Big| \int_a^b g(x, \tfrac{x}{\varepsilon})\,dx \Big| \le B(g)\varepsilon + |b-a| B(\partial_x g)\varepsilon, \tag{2.43}$$

where $B(\zeta) = \max_{(x,y)\in\mathbb{R}\times[0,1]} |\zeta(x,y)|$ for any function ζ defined on $\mathbb{R}\times[0,1]$.

Proof. The estimate (2.43) is a direct consequence of the identity

$$g(x, \tfrac{x}{\varepsilon}) = \frac{d}{dx}\int_a^x g(x, \tfrac{s}{\varepsilon})\,ds - \int_a^x \frac{\partial g}{\partial x}(x, \tfrac{s}{\varepsilon})\,ds$$

and the estimates

$$\Big| \int_a^b g(x, \tfrac{s}{\varepsilon})\,ds \Big| \le B(g)\varepsilon, \quad \Big| \int_a^x \frac{\partial g}{\partial x}(x, \tfrac{s}{\varepsilon})\,ds \Big| \le B(\partial_x g)\varepsilon,$$

which follow from the 1-periodicity of $g(x,y)$ in y and that $\int_0^1 g(x,y)\,dy = 0$. This completes the proof. □

Proof (of Theorem 2.1). Subtracting (2.36) from (2.32) and integrating the resulting equation along the characteristics from 0 to t, we get

$$\begin{aligned}
&u(x,t) - U(x, \tfrac{x-t}{\varepsilon}, t) \\
&= \int_0^t \Big[w(x-t+s,s)^2 - W(x-t+s, \tfrac{x-t+s}{\varepsilon}, s)^2 \Big]\, ds \\
&\quad + \int_0^t \Big[W(x-t+s, \tfrac{x-t+s}{\varepsilon}, s)^2 - \textstyle\int_0^1 W(x-t+s,y,s)^2\, dy \Big]\, ds \\
&\quad - \int_0^t \Big[u(x-t+s,s) v(x-t+s,s) \\
&\qquad\qquad - U(x-t+s, \tfrac{x-t}{\varepsilon}, s) V(x-t+s, \tfrac{x-t+2s}{\varepsilon}, s) \Big]\, ds \\
&\quad - \int_0^t U(x-t+s, \tfrac{x-t}{\varepsilon}, s) \Big[V(x-t+s, \tfrac{x-t+2s}{\varepsilon}, s) \\
&\qquad\qquad - \int_0^1 V(x-t+s,y,s)\, dy \Big]\, ds \\
&:= (\mathrm{I})_1 + \cdots + (\mathrm{I})_4. \qquad (2.44)
\end{aligned}$$

It is clear from the definition of $E(t)$ and $M(T)$ that

$$|(\mathrm{I})_1 + (\mathrm{I})_3| \le 2M(T) \int_0^t E(s)\, ds.$$

To estimate $(\mathrm{I})_2$, we define for fixed $(x,t) \in \mathbb{R} \times [0,T]$,

$$g_{(x,t)}(s,y) = W(x-t+s, \tfrac{x-t}{\varepsilon} + y, s)^2.$$

Since the 1-periodicity of $W(x,y,t)$ in y implies

$$\int_0^1 W(x-t+s,y,s)^2\, dy = \int_0^1 W(x-t+s, \tfrac{x-t}{\varepsilon} + y, s)^2\, dy,$$

we obtain by applying Lemma 2.1 that

$$\begin{aligned}
|(\mathrm{I})_2| &= \Big| \int_0^t \Big[g_{(x,t)}(s, \tfrac{s}{\varepsilon}) - \int_0^1 g_{(x,t)}(s,y)\, dy \Big]\, ds \Big| \\
&\le M(T)^2 \varepsilon + 2M(T)K(T)T\varepsilon.
\end{aligned}$$

Similarly, we have

$$(\mathrm{I})_4 \le M(T)^2 \varepsilon + 2M(T)K(T)T\varepsilon.$$

Substituting these estimates into (2.44) we get

$$\Big| u(x,t) - U(x, \tfrac{x-t}{\varepsilon}, t) \Big| \le 2M(T) \int_0^t E(s)\, ds + 2M(T)^2 \varepsilon + 4M(T)K(T)T\varepsilon. \qquad (2.45)$$

Similarly, we conclude from (2.37)-(2.38) and (2.33)-(2.34) that

$$\left|v(x,t) - V(x, \tfrac{x+t}{\varepsilon}, t)\right| \le 2M(T) \int_0^t E(s)\, ds + 2M(T)^2\varepsilon + 4M(T)K(T)T\varepsilon, \tag{2.46}$$

$$\left|w(x,t) - W(x, \tfrac{x}{\varepsilon}, t)\right| \le 2M(T) \int_0^t E(s)\, ds + M(T)^2\varepsilon + 2M(T)K(T)T\varepsilon. \tag{2.47}$$

Now the desired estimate (2.40) follows from summing (2.45)-(2.47) and using the Gronwall inequality. □

Remark 2.2. The homogenization theory tells us that the initial oscillatory solutions propagate along their characteristics. The nonlinear interaction can generate only low frequency contributions to the u and v components. On the other hand, the nonlinear interaction of $u,\ v$ on w can generate both low and high frequency contribution to w. That is, even if w has no oscillatory component initially, the dynamical interaction of $u,\ v$ and w can generate a high frequency contribution to w at later time. This is not the case for the u and v components. Due to this resonant interaction of $u,\ v$ and w, the weak limit of $u_\varepsilon v_\varepsilon w_\varepsilon$ is not equal to the product of the weak limits of $u_\varepsilon,\ v_\varepsilon,\ w_\varepsilon$. This explains why the Compensated Compactness result does not apply to this 3×3 system [59].

Although it is difficult to characterize the weak limit of the triple product, $u_\varepsilon v_\varepsilon w_\varepsilon$ for arbitrary oscillatory initial data, it is possible to say something about the weak limit of the triple product for oscillatory initial data that have periodic structure, such as the ones studies here. Depending on α being rational or irrational, the limiting behavior is very different. In fact, one can show that $\overline{u_\varepsilon v_\varepsilon w_\varepsilon} = \overline{u_\varepsilon v_\varepsilon}\,\overline{w_\varepsilon}$ when α is equal to an irrational number. This is not true in general when α is a rational number.

2.3 Convection of microstructure

It is most interesting to see if one can apply homogenization technique to obtain an averaged equation for the large scale quantity for incompressible Euler or Navier-Stokes equations. In 1985, McLaughlin, Papanicolaou and Pironneau [51] attempted to obtain a homogenized equation for the 3-D incompressible Euler equations with highly oscillatory velocity field. More specifically, they considered the following initial value problem:

$$u_t + (u \cdot \nabla)u = -\nabla p,$$

with $\nabla \cdot u = 0$ and highly oscillatory initial data

$$u(x,0) = U(x) + W(x, x/\varepsilon).$$

They then constructed multiscale expansions for both the velocity field and the pressure. In doing so, they made an important assumption that the microstructure is convected by the mean flow. Under this assumption, they constructed a multiscale expansion for the velocity field as follows:

$$u^\varepsilon(x,t) = u(x,t) + w(\tfrac{\theta(x,t)}{\varepsilon}, \tfrac{t}{\varepsilon}, x, t) + \varepsilon u_1(\tfrac{\theta(x,t)}{\varepsilon}, \tfrac{t}{\varepsilon}, x, t) + O(\varepsilon^2).$$

The pressure field p^ε is expanded similarly. From this ansatz, one can show that θ is convected by the mean velocity:

$$\theta_t + u \cdot \nabla\theta = 0\,, \qquad \theta(x,0) = x\,.$$

It is a very challenging problems to develop a systematic approach to study the large scale solution in three dimensional Euler and Navier-Stokes equations. The work of McLaughlin, Papanicolaou and Pironneau provided some insightful understanding into how small scales interact with large scale and how to deal with the closure problem. However, the problem is still not completely resolved since the cell problem obtained this way does not have a unique solution. Additional constraints need to be enforced in order to derive a large scale averaged equation. With additional assumptions, they managed to derive a variant of the $k-\varepsilon$ model in turbulence modeling.

Remark 2.3. One possible way to improve the work of [51] is take into account the oscillation in the Lagrangian characteristics, θ_ε. The oscillatory part of θ_ε in general could have order one contribution to the mean velocity of the incompressible Euler equation. With Dr. Danping Yang [41], we have studied convection of microstructure of the 2-D and 3-D incompressible Euler equations using a new approach. We do not assume that the oscillation is propagated by the mean flow. In fact, we found that it is crucial to include the effect of oscillations in the characteristics on the mean flow. Using this new approach, we can derive a well-posed cell problem which can be used to obtain an effective large scale average equation.

More can be said for a passive scalar convection equation.

$$v_t + \frac{1}{\varepsilon}\nabla \cdot (u(x/\varepsilon)v) = \alpha\Delta v,$$

with $v(x,0) = v_0(x)$. Here $u(y)$ is a known incompressible periodic (or stationary random) velocity field with zero mean. Assume that the initial condition is smooth.

Expand the solution v^ε in powers of ε

$$v^\varepsilon = v(t,x) + \varepsilon v_1(t,x,x/\varepsilon) + \varepsilon^2 v_2(t,x,x/\varepsilon) + \cdots\,.$$

The coefficients of ε^{-1} lead to

$$\alpha\Delta_y v_1 - u \cdot \nabla_y v_1 - u \cdot \nabla_x v = 0.$$

Let e_k, $k = 1, 2, 3$ be the unit vectors in the coordinate directions and let $\chi^k(y)$ satisfy the cell problem:

$$\alpha \Delta_y \chi^k - u \cdot \nabla_y \chi^k - u \cdot e_k = 0.$$

Then we have

$$v_1(t, x, y) = \sum_{k=1}^{3} \chi^k(y) \frac{v(t, x)}{\partial x_k}.$$

The coefficients of ε^0 give

$$\alpha \Delta_y v_2 - u \cdot \nabla_y v_2 = u \cdot \nabla_x v_1 - 2\alpha \nabla_x \cdot \nabla_y v_1 - \alpha \Delta_x v + v_t.$$

The solvability condition for v_2 requires that the right hand side has zero mean with respect to y. This gives rise to the equation for homogenized solution v

$$v_t = \alpha \Delta_x v - \overline{u \cdot \nabla_x v_1}.$$

Using the cell problem, McLaughlin, Papanicolaou, and Pironneau obtained [51]

$$v_t = \sum_{i,j=1}^{3} (\alpha \delta_{ij} + \alpha_{T_{ij}}) \frac{\partial^2 v}{\partial x_i \partial x_j},$$

where $\alpha_{T_{ij}} = -\overline{u_i \chi^j}$.

Nonlocal memory effect of homogenization It is interesting to note that for certain degenerate problem, the homogenized equation may have a nonlocal memory effect.

Consider the simple 2-D linear convection equation:

$$\frac{\partial u_\varepsilon(x, y, t)}{\partial t} + a_\varepsilon(y) \frac{\partial u_\varepsilon(x, y, t)}{\partial x} = 0,$$

with initial condition $u_\varepsilon(x, y, 0) = u_0(x, y)$.

We assume that a_ε is bounded and u_0 has compact support. While it is easy to write down the solution explicitly,

$$u_\varepsilon(x, y, t) = u_0(x - a_\varepsilon(y)t, y),$$

it is not an easy task to derive the homogenized equation for the weak limit of u_ε.

Using Laplace Transform and measure theory, Luc Tartar [60] showed that the weak limit u of u_ε satisfies

$$\frac{\partial}{\partial t} u(x, y, t) + A_1(y) \frac{\partial}{\partial x} u(x, y, t) = \int_0^t \int \frac{\partial^2}{\partial x^2} u(x - \lambda(t - s), y, s) d\mu_y(\lambda)\, ds,$$

with $u(x, y, 0) = u_0(x, y)$, where $A_1(y)$ is the weak limit of $a_\varepsilon(y)$, and μ_y is a probability measure of y and has support in $[\min(a_\varepsilon), \max(a_\varepsilon)]$.

As we can see, the degenerate convection induces a nonlocal history dependent diffusion term in the propagating direction (x). The homogenized equation is *not* amenable to computation since the measure μ_y cannot be expressed explicitly in terms of a_ε.

3 Numerical Homogenization Based on Sampling Techniques

Homogenization theory provides a critical guideline for us to design effective numerical methods to compute multiscale problems. Whenever homogenized equations are applicable they are very useful for computational purposes. There are, however, many situations for which we do not have well-posed effective equations or for which the solution contains different frequencies such that effective equations are not practical. In these cases we would like to approximate the original equations directly. In this part of my lectures, we will investigate the possibility of approximating multiscale problems using particle methods together with sampling technique. The classes of equations we consider here include semilinear hyperbolic systems and the incompressible Euler equation with oscillatory solutions.

When we talk about convergence of an approximation to an oscillatory solution, we need to introduce a new definition. The traditional convergence concept is too weak in practice and does not discriminate between solutions which are highly oscillatory and those which are smooth. We need the error to be small essentially independent of the wavelength in the oscillation when the computational grid size is small. On the other hand we cannot expect the approximation to be well behaved pointwise. It is enough if the continuous solution and its discrete approximation have similar local or moving averages.

Definition 3.1 (Engquist [30]). Let v^n be the numerical approximation to u at time $t_n (t_n = n\Delta t)$, ε represents the wave length of oscillation in the solution. The approximation v^n converges to u as $\Delta t \to 0$, essentially independent of ε, if for any $\delta > 0$ and $T > 0$ there exists a set $s(\varepsilon, \Delta t_0) \in (0, \Delta t_0)$ with measure $(s(\varepsilon, \Delta t_0)) \geq (1 - \delta)\Delta t_0$ such that

$$||u(\cdot, t_n) - v^n|| \leq \delta, \qquad 0 \leq t_n \leq T$$

is valid for all $\Delta t \in s(\varepsilon, \Delta t_0)$ and where Δt_0 is independent of ε.

The convergence concept of "essentially independent of ε" is strong enough to mimic the practical case where the high frequency oscillations are not well resolved on the grid. A small set of values of Δt has to be removed in order to avoid resonance between Δt and ε. Compare the almost always convergence for the Monte Carlo methods [55].

It is natural to compare our problem with the numerical approximation of discontinuous solutions of nonlinear conservation laws. Shock capturing methods do not produce the correct shock profiles but the overall solution may still be good. For this the scheme must satisfy certain conditions such as conservation form. We are here interested in analogous conditions on algorithms for oscillatory solutions. These conditions should ideally guarantee that the numerical approximation in some sense is close to the solution of the corresponding effective equation when the wave length of the oscillation tends to zero.

There are three central sources of problems for discrete approximations of highly oscillatory solutions.

(i) The first one is the sampling of the computational mesh points ($x_j = j\Delta x, j = 0, 1, ...$). There is the risk of resonance between the mesh points and the oscillation. For example, if Δx equals the wave length of the periodic oscillation, the discrete initial data may only get values from the peaks of a curve like the upper envelope of the oscillatory solution. We can never expect convergence in that case. Thus Δx cannot be completely independent of the wave length.

(ii) Another problem comes from the approximation of advection. The group velocity for the differential equation and the corresponding discretization are often very different [33]. This means that an oscillatory pulse which is not well resolved is not transported correctly even in average by the approximation. Furthermore, dissipative schemes do not advect oscillations correctly. The oscillations are damped out very fast in time.

(iii) Finally, the nonlinear interaction of different high frequency components in a solution must be modeled correctly. High frequency interactions may produce lower frequencies that influence the averaged solution. We can show that this nonlinear interaction is well approximated by certain particle methods applied to a class of semilinear differential equations. The problem is open for the approximation of more general nonlinear equations.

In [31,32], we studied a particle method approximation to the nonlinear discrete Boltzmann equations in kinetic theory of discrete velocity with multiscale initial data. In such equations, high frequency components can be transformed into lower frequencies through nonlinear interactions, thus affecting the average of solutions. We assume that the initial data are of the form $a(x, x/\varepsilon)$ with $a(x, y)$ 1-periodic in each component of y. As we see from the homogenization theory in the previous section, the behavior of oscillatory solutions for the generalized Broadwell model is very sensitive to the velocity coefficients. It depends on whether a certain ratio among the velocity components is a rational number or an irrational number.

It is interesting to note that the structure of oscillatory solutions for the generalized Broadwell model is quite stable when we perturb the velocity

coefficient α around irrational numbers. In this case, the resonance effect of u and v on w vanishes in the limit of $\varepsilon \to 0$. However, the behavior of oscillatory solutions for the generalized Broadwell model becomes singular when perturbing around integer velocity coefficients. There is a strong interaction between the high frequency components of u and v, and the interaction in the uv term would create an oscillation of order $O(1)$ on the w component. In [59], Tartar showed that for the Carleman model the weak limit of all powers of the initial data will uniquely determine the weak limit of the oscillatory solutions at later times, using the Compensated Compactness Theorem. We found that this is no longer true for the generalized Broadwell model with integer-values velocity coefficients [37].

In [31,32], we showed that this subtle behavior for the generalized Broadwell model with oscillatory initial data can be captured correctly by a particle method even on a coarse grid. The particle method converges to the effective solution essentially independent of ε. For the Broadwell model, the hyperbolic part is solved exactly by the particle method. No averaging is therefore needed in the convergence result. We also analyze a numerical approximation of the Carleman equations with variable coefficients. The scheme is designed such that particle interaction can be accounted for without introducing interpolation. There are errors in the particle method approximation of the linear part of the system. As a result, the convergence can only be proved for moving averages. The convergence proofs for the Carleman and the Broadwell equations have one feature in common. The local truncation errors in both cases are of order $O(\Delta t)$. In order to show convergence, we need to take into account cancellation of the local errors at different time levels. This is very different from the conventional convergence analysis for finite difference methods. This is also the place where numerical sampling becomes crucial in order to obtain error cancellation at different time levels.

In the next two subsections, we present a careful study of the Broadwell model with highly oscillatory initial data in order to demonstrate the basic idea of the numerical homogenization based on sampling techniques.

3.1 Convergence of the Particle Method

Now we consider how to capture this oscillatory solution on a coarse grid using a particle method. Since the discrete velocity coefficients are integers for the Broadwell model, we can express a particle method in the form of a special finite difference method by choosing $\Delta x = \Delta t$. Denote by u_i^n, v_i^n, w_i^n the approximations of $u(x_i, t^n), v(x_i, t^n)$ and $w(x_i, t^n)$ respectively with $x_i = i\Delta x$ and $t^n = n\Delta t$. Our particle scheme is given by

$$u_i^n = u_{i-1}^{n-1} + \Delta t(w^2 - uv)_{i-1}^{n-1}, \tag{3.1}$$

$$v_i^n = v_{i+1}^{n-1} + \Delta t(w^2 - uv)_{i+1}^{n-1}, \tag{3.2}$$

$$w_i^n = w_i^{n-1} - \Delta t(w^2 - uv)_i^{n-1}, \tag{3.3}$$

with the initial conditions given by

$$u_i^0 = u(x_i, 0), \quad v_i^0 = v(x_i, 0), \quad w_i^0 = w(x_i, 0). \tag{3.4}$$

To study the convergence of the particle scheme (3.1)-(3.4) we need the following lemma, which is a discrete analogue of Lemma 2.1.

Lemma 3.1. *Let $g(x,y) \in C^3([0,T] \times [0,1])$ be 1-periodic in y and satisfy the relation $\int_0^1 g(x,y)\,dy = 0$. Let $x_k = kh$ and $r = h/\varepsilon$. If $h \in S(\varepsilon, h_0)$ where*

$$S(\varepsilon, h_0) = \{0 < h \le h_0 : \frac{kh}{\varepsilon} \notin \Big(i - \frac{\tau}{|k|^{3/2}}, i + \frac{\tau}{|k|^{3/2}}\Big),$$
$$\textit{for } i = 1, 2, \cdots, \Big[\frac{kh_0}{\varepsilon}\Big] + 1, 0 \neq k \in Z, 0 < \varepsilon \le 1\},$$

then we have

$$\Big|\sum_{k=0}^{n-1} g(x_k, \frac{x_k}{\varepsilon})h\Big| \le \frac{C_0(1+T)L(g)h}{\tau}, \quad \forall\ n = 1, 2, \cdots, \Big[\frac{T}{h}\Big],$$

where C_0 is a constant independent of h, ε, T, τ and g, and

$$L(g) = \max_{(x,y)\in[0,T]\times[0,1]} \Big(|\partial_y^3 g(x,y)|, |\partial_x \partial_y^3 g(x,y)|\Big).$$

Moreover, it is obvious that

$$|S(\varepsilon, h_0)| \ge h_0\Big(1 - \tau\sum_{k=1}^{\infty} k^{-3/2}\Big) \ge h(1 - 3\tau).$$

Proof. Since g is 1-periodic in y with mean zero, it can be expanded in a Fourier series

$$g(x,y) = \sum_{m\neq 0} a_m(x) \mathrm{e}^{2\pi i m y}, \text{ where } a_m(x) = \int_0^1 g(x,y)\mathrm{e}^{-2\pi i m y}\,dy.$$

Simple integration by parts yields that

$$|a_m(x)| \le \frac{1}{(2\pi|m|)^3} L(g), \quad |a_m'(x)| \le \frac{1}{(2\pi|m|)^3} L(g).$$

Thus we have

$$\begin{aligned}\Big|\sum_{k=0}^{n-1} g(x_k, \frac{x_k}{\varepsilon})h\Big| &= \Big|\sum_{k=0}^{n-1}\sum_{m\neq 0} a_m(x_k)\mathrm{e}^{2\pi i m x_k/\varepsilon} h\Big| \\ &= \Big|\sum_{m\neq 0}\sum_{k=0}^{n-1} a_m(x_k)\mathrm{e}^{2\pi i m k h/\varepsilon} h\Big|.\end{aligned}$$

Summation by parts yields

$$\begin{aligned}\Big|\sum_{k=0}^{n-1} a_m(x_k)\mathrm{e}^{2\pi ikh/\varepsilon}\Big| &\le |a_m(x_{n-1})\sum_{k=0}^{n-1}\mathrm{e}^{2\pi ikh/\varepsilon}\Big| \\ &\quad +\Big|\sum_{k=0}^{n-1}\Big(\sum_{j=1}^{k}\mathrm{e}^{2\pi imjh/\varepsilon}\Big)(a_m(x_k)-a_m(x_{k+1}))\Big| \\ &\le \frac{2(1+T)L(g)}{(2\pi|m|)^3|1-\mathrm{e}^{2\pi imh/\varepsilon}|}.\end{aligned}$$

But for $h \in S(\varepsilon, h_0)$ we have

$$|1-\mathrm{e}^{2\pi imh/\varepsilon}| = 2|\sin(\pi mh/\varepsilon)| \ge \frac{2\pi\tau}{|m|^{3/2}}.$$

Hence, for $h \in S(\varepsilon, h_0)$,

$$\sum_{k=0}^{n-1}\Big|g(x_k, \frac{x_k}{\varepsilon})h\Big| \le \frac{2(1+T)L(g)h}{(2\pi)^4\tau}\sum_{m\neq 0}\frac{1}{|m|^{3/2}} =: \frac{C_0h(1+T)L(g)}{\tau}.$$

This completes the proof. □

Now we are ready to study the approximation property of the particle scheme (3.1)-(3.4). First denote by

$$E^n = \max_i\Big(|u(x_i,t^n)-u_i^n|,\ |v(x_i,t^n)-v_i^n|,\ |w(x_i,t^n)-w_i^n|\Big). \tag{3.5}$$

Integrating (2.29) from 0 to t^n along its characteristics, we get

$$u(x_i,t^n) = u(x_i-t^n,0) + \int_0^{t^n}(w^2-uv)(x_i-t^n+s,s)\,ds. \tag{3.6}$$

¿From (3.1) we know that

$$u_i^n = u_{i-n}^0 + \sum_{k=0}^{n-1}(w^2-uv)_{i-k}^k\Delta t. \tag{3.7}$$

Subtracting (3.7) from (3.6) we obtain that

$$\begin{aligned}&u(x_i,t^n)-u_i^n \\ &= \int_0^{t^n}(w^2-uv)(x_i-t^n+s,s)\,ds - \sum_{k=0}^{n-1}(w^2-uv)(x_i-t^k,t^k)\Delta t \\ &\quad +\sum_{k=0}^{n-1}\Delta t\Big[(w^2-uv)(x_i-t^k,t^k)-(w^2-uv)_{i-k}^k\Big] \\ &:= \text{(II)} + \text{(III)}.\end{aligned} \tag{3.8}$$

Let $M(T)$ be defined as in (2.41) and $N(T)$ be given by

$$N(T) = \max\Big\{|u_i^k|, |v_i^k|, |w_i^k| : i \in Z, 0 \le k \le [T/\Delta t]\Big\}. \tag{3.9}$$

It can be shown that $N(T)$ is bounded for finite time independent of ε, see [32]. Then it is clear that

$$\text{(III)} \le (M(T) + N(T)) \sum_{k=0}^{n-1} \Delta t E^k. \tag{3.10}$$

It remains to estimate (II). For convenience, let $\theta = w^2 - uv$ and

$$\Theta(x,t) = W(x, \tfrac{x}{\varepsilon}, t)^2 - U(x, \tfrac{x-t}{\varepsilon}, t)V(x, \tfrac{x+t}{\varepsilon}, t).$$

Then we have

$$\begin{aligned}
\text{(II)} \;=\; & \int_0^{t^n} \Big[\theta(x_i - t^n + s, s)\, ds - \Theta(x_i - t^n + s, s)\Big]\, ds \\
& + \Big[\int_0^{t^n} \Theta(x_i - t^n + s, s)\, ds - \sum_{k=0}^{n-1} \Delta t \Theta(x_i - t^k, t^k)\Big] \\
& + \sum_{k=0}^{n-1} \Big[\Theta(x_i - t^k, t^k) - \theta(x_i - t^k, t^k)\Big]\Delta t \\
:= \; & \text{(II)}_1 + \cdots + \text{(II)}_3.
\end{aligned} \tag{3.11}$$

By Theorem 2.1 we get

$$|\text{(II)}_1 + \text{(II)}_3| \le 2TM(T)C_1(T)\varepsilon. \tag{3.12}$$

To proceed further, let, for fixed (x_i, t^n),

$$g_i^n(s,y) = W(x_i - t^n + s, \tfrac{x_i - t^n}{\varepsilon} + y, s)^2.$$

It is clear that g_i^n is 1-periodic in y. Now by Lemmas 2.1-2.2 we have

$$\begin{aligned}
& \int_0^{t^n} W(x_i - t^n + s, s)^2\, ds - \sum_{k=0}^{n-1} W(x_i - t^k, t^k)^2 \Delta t \\
= & \int_0^{t^n} \Big[g_i^n(s, \tfrac{s}{\varepsilon}) - \int_0^1 g_i^n(s,y)\, dy\Big]\, ds \\
& + \int_0^{t^n} \int_0^1 g_i^n(s,y)\, dy\, ds - \sum_{k=0}^{n-1} \Delta t \int_0^1 g_i^n(t^{n-k}, y)\, dy \\
& - \sum_{k=0}^{n-1} \Delta t \Big[g(t^{n-k}, \tfrac{t^{n-k}}{\varepsilon}) - \int_0^1 g_i^n(t^{n-k}, y)\, dy\Big] \le C(T)(\varepsilon + \Delta t),
\end{aligned} \tag{3.13}$$

where we have used standard methods to estimate the second term, since the derivative of g_i^n with respect to s is independent of ε. Here and in the remainder of this section, we will always denote by $C(T)$ the various constants which are independent of ε and Δt. Now similar to the reasoning leading to (3.13) we can obtain

$$|(\mathrm{II})_3| \le C(T)(\varepsilon + \Delta t). \tag{3.14}$$

¿From (3.8)-(3.12) and (3.14) we finally get

$$|u(x_i, t^n) - u_i^n| \le C(T)(\varepsilon + \Delta t) + (M(T) + N(T)) \sum_{k=0}^{n-1} \Delta t E^k. \tag{3.15}$$

Similarly, we have

$$|v(x_i, t^n) - v_i^n| \le C(T)(\varepsilon + \Delta t) + (M(T) + N(T)) \sum_{k=0}^{n-1} \Delta t E^k, \tag{3.16}$$

$$|w(x_i, t^n) - w_i^n| \le C(T)(\varepsilon + \Delta t) + (M(T) + N(T)) \sum_{k=0}^{n-1} \Delta t E^k. \tag{3.17}$$

To summarize, we have the following theorem by summing (3.15)-(3.17) and applying the Gronwall inequality.

Theorem 3.1. *Let (u, v, w) be the solution of (2.32)-(2.35) and (u_i^n, v_i^n, w_i^n) be the solution of the particle scheme (3.1)-(3.4). Assume that $\Delta t \in S(\varepsilon, \Delta t_0)$ where $S(\varepsilon, \Delta t_0)$ is defined in Lemma 3.1. Then the following estimate holds*

$$\max_{1 \le n \le [T/\Delta t]} E^n \le C(T)(\varepsilon + \Delta t),$$

where $C(T)$ is independent of ε and Δt, and E^n is defined as in (3.5).

Remark 3.1. It is important that we perform the error analysis globally in time in order to account for cancellation of local truncation errors at different time steps. As we can see from the analysis, the local truncation error is of order Δt in one time step. If we do not take into account the error cancellation in time, we would obtain an error bound of order $O(1)$ which is an overestimate. The error cancellation is closely related to the sampling we choose. This is the place where we can see the difference between a good sampling and a resonant sampling.

Remark 3.2. As we can see from the error analysis, error cancellation along Lagrangian characteristics is essential in obtaining convergence independent of the oscillation. This idea can be generalized to hyperbolic systems with variable coefficient velocity fields. In the special case of the Carleman model with variable coefficients, we have analyzed the convergence of a particle

method in [31]. However, the particle method analyzed in [31] does not generalize to multi-dimensions or 3×3 systems. Together with a Ph.D. student, Razvan Fetecau, we have designed a modified Lagrangian particle method. In this method, each component of the solution is updated along its own characteristic. So there is no fixed grid. When we update one component of the solution, say u, we need values of the other components (say v and w) along the u characteristic. We obtain these values by using some high order interpolation scheme (such as Fourier interpolation or cubic spline). This modified Lagrangian particle method in principle works for any number of families of characteristics and for multi-dimensions. ¿From our preliminary numerical experiments, it produces excellent results for both the Broadwell and Carleman models, even in the oscillatory coefficients case.

Below we describe briefly the results we obtain for the variable coefficient Carleman equations

$$u_t + a(x,t)u_x = v^2 - u^2 \,, \tag{3.18}$$

$$v_t - b(x,t)v_x = u^2 - v^2 \,, \tag{3.19}$$

with initial data $u(x,0) = u_0(x, x/\varepsilon)$, $v(x,0) = v_0(x, x/\varepsilon)$. In Figure 3.1, we illustrate the particle trajectories for the u and v components.

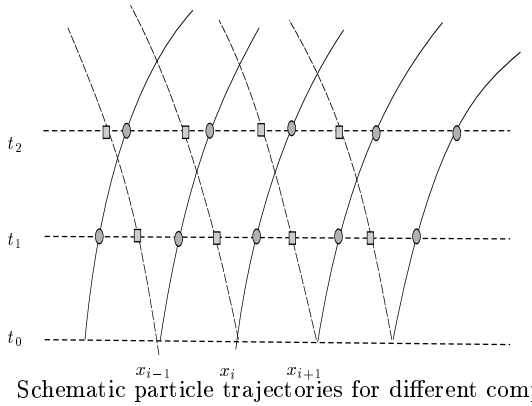

Fig. 3.1. Schematic particle trajectories for different components.

We choose the oscillatory coefficients as follows:

$$a(x,t) = 1 + 0.5\sin\left(\tfrac{xt}{\varepsilon}\right) \text{ and } b(x,t) = 1 + 0.2\cos\left(\tfrac{xt}{\varepsilon}\right).$$

The initial conditions for u and v are chosen as

$$u_0(x, x/\varepsilon) = \begin{cases} 0.5\sin^4(\pi(x-3)/2)(1+\sin(2\pi(x-3)/\varepsilon)), & |x-4| < 1 \\ 0, & |x-4| \geq 1 \end{cases} \tag{3.20}$$

$$v_0(x, x/\varepsilon) = \begin{cases} 0.5\sin^4(\pi(x-4)/2)(1+\sin(2\pi(x-4)/\varepsilon)), & |x-5| < 1 \\ 0, & |x-5| \geq 1 \end{cases} \tag{3.21}$$

In our calculations, we choose $\Delta x = 0.01$, $\Delta t = \frac{\Delta x}{\sqrt{5}}$, and $\varepsilon = \Delta x\sqrt{2} \approx 0.014$. We plot the u-characteristic in Figure 3.2. The coarse grid solution for the u-component is plotted in Figure 3.3a. We can see that it captures very well the high frequency information. In Figure 3.3b, we put the coarse grid solution on top of the corresponding well-resolved solution. The agreement is very good. We also check the accuracy of the moving average [31] of the solution and the average of its second order moments. The results are plotted in Figure 3.4. Again, we observe excellent agreement between the coarse grid calculations and the well-resolved calculations.

We have also performed the same calculations for the 3×3 Broadwell model with rational or irrational coefficient α. The subtle homogenization behavior is captured correctly for both rational α and for irrational α. We do not present the results here.

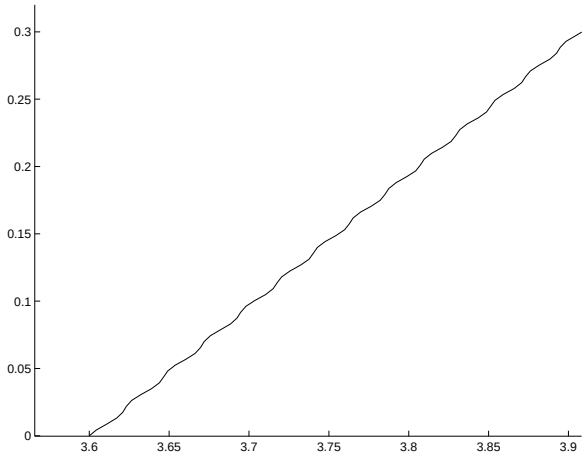

Fig. 3.2. A typical u-characteristic trajectory.

3.2 Vortex methods for incompressible flows

The generalization of the particle method to the incompressible flows is the vortex method. In [26], we have analyzed the convergence of the vortex method for 2-D incompressible Euler equations with oscillatory vorticity field. Our analysis relies on the observation that there are tremendous cancellations among the local errors at different space locations in the velocity approximation. Thus the local errors do not add up to $O(1)$ as predicted by the classical error estimate in the case where the grid size is large compared to the oscillatory wavelength.

Consider the 2-D incompressible Euler equation in vorticity form:

$$\omega_t + (u \cdot \nabla)\omega = 0$$

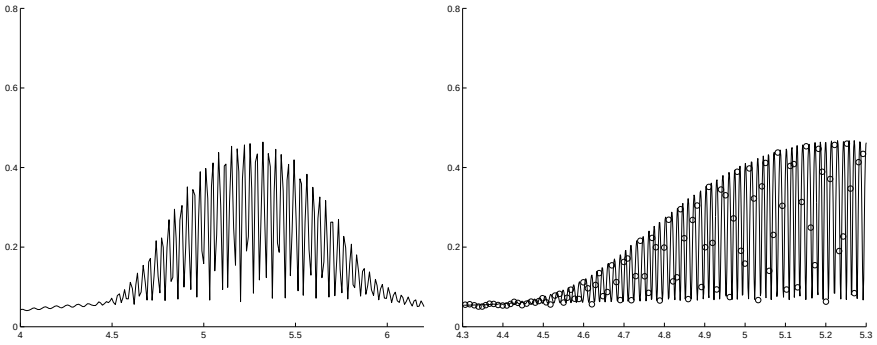

Fig. 3.3. (a): Coarse grid solution u at time $t = 1.28$. (b): Putting the coarse grid solution u on top of a well-resolved computation (solid line).

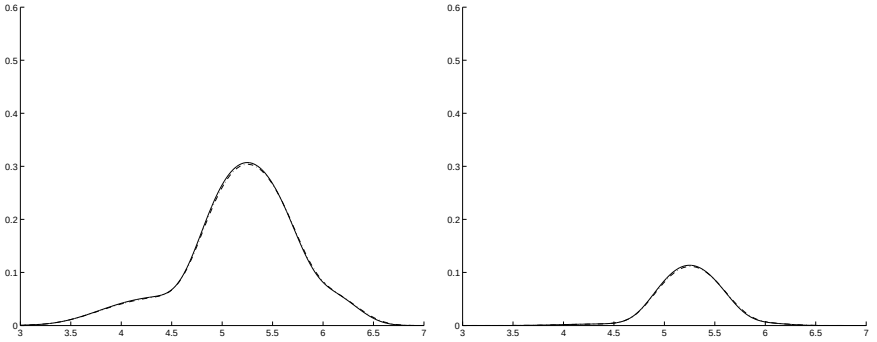

Fig. 3.4. (a): The averaged solution $\overline{u}$ (dashdot line); the solid line represents a very well resolved computation. (b): The averaged second order moment $\overline{u^2}$ (dashdot line); the solid line represents a very well resolved computation.

with oscillatory initial vorticity $\omega(x,0) = \omega_0(x, x/\varepsilon)$.

Define the particle trajectory, denoted as $X(t,\alpha)$,

$$\frac{dX(t,\alpha)}{dt} = u(X(t,\alpha),t), \quad X(0,\alpha) = \alpha.$$

Vorticity is conserved along characteristics:

$$\omega(X(t,\alpha),t) = \omega_0(\alpha).$$

On the other hand, velocity can be expressed in terms of vorticity by the Biot-Savart law:

$$u(X(t,\alpha),t) = \int K(X(t,\alpha) - X(t,\alpha'))\omega_0(\alpha')d\,\alpha'$$

with K given by $K(x) = (-x_2, x_1)/(2\pi|x|^2)$.

The Biot-Savart kernel K has a singularity at the origin. To regularize the kernel, Chorin introduced the vortex blob method (see, e.g. [17], replacing K by $K_\delta = K * \zeta_\delta$,

$$\zeta_\delta = \frac{1}{\delta^2}\zeta\left(\frac{x}{\delta}\right), \quad \delta = h^\sigma, \text{ with } \sigma < 1.$$

ζ is typically chosen as a variant of Gaussian.

The vortex blob method is given by

$$\frac{dX_i^h(t)}{dt} = \sum_j K_\delta(X_i^h(t) - X_j^h(t))\omega_j h^2,$$

where $X_i^h(0) = \alpha_i$, and $w_j = w_0(\alpha_j, \alpha_j/\varepsilon)$.

Together with Weinan E, we have proved that the vortex method converges essentially independent of ε [26].

The case studied in [26] deals with bounded oscillatory vorticity. This assumption leads to strong convergence of the velocity field. It is more physical to consider homogenization for highly oscillatory velocity field. Would the vortex blob method still capture the correct large scale solution with a relatively coarse grid (or small number of particles)? Together with a Ph.D. student, Razvan Fetecau, we have recently derived a modified vortex method for the coarse (or macro) particle system by combining a local subgrid correction with a model reduction technique.

4 Numerical homogenization based on Multiscale FEMs

It is natural to consider the possibility of generalizing the sampling technique to second order elliptic equations with highly oscillatory coefficients. In [3], we showed that finite difference approximations converge essentially independent of the small scale ε for one-dimensional elliptic problems. In several space dimensions we found that only in the case of rapidly oscillating periodic coefficients do the above results generalize, in a weaker form. In the case of almost periodic or random coefficients in several space dimensions we showed, both theoretically and with a simple counterexample, that numerical homogenization by sampling does not work efficiently. New ideas seem to be needed.

In order to overcome the difficulty we mentioned above for the sampling technique, we have introduced a multiscale finite element method (MsFEM) for solving partial differential equations with multiscale solutions, see [38,40,39,28,16,61,1]. The central goal of this approach is to obtain the large scale solutions accurately and efficiently without resolving the small scale details. The main idea is to construct finite element base functions which

capture the small scale information within each element. The small scale information is then brought to the large scales through the coupling of the global stiffness matrix. Thus, the effect of small scales on the large scales is correctly captured. In our method, the base functions are constructed from the leading order homogeneous elliptic equation in each element. As a consequence, the base functions are adapted to the local microstructure of the differential operator. In the case of two-scale periodic structures, we have proved that the multiscale method indeed converges to the correct solution independent of the small scale in the homogenization limit [40].

In practical computations, a large amount of overhead time comes from constructing the base functions. In general, these multiscale base functions are constructed numerically, except for certain special cases. Since the base functions are independent of each other, they can be constructed independently and can be done perfectly in parallel. This greatly reduces the overhead time in constructing these bases. In many applications, it is important to obtain a scale-up equation from the fine grid equation. Our multiscale finite element method can be used for a similar purpose [61]. The advantage of deriving a scale-up equation is that one can perform many useful tests on the scale-up (coarse grid) model with different boundary conditions or source terms. This would be very expensive if we have to perform all these tests on a fine grid. For time dependent problems, the scaled-up equation also allows for larger time steps. This results in additional computational saving. Another advantage of the method is its ability to scale down the size of a large scale computation. This offers a big saving in computer memory.

It should be mentioned that many numerical methods have been developed with goals similar to ours. These include wavelet based numerical homogenization methods [10,21,19,45], methods based on the homogenization theory (cf. [9,25,18,34]), variational multiscale methods [42,12,43], matrix-dependent multigrid based homogenization [45,19], generalized p-FEM in homogenization [47,48], and some upscaling methods based on simple physical and/or mathematical motivations (cf. [23,49]). The methods based on the homogenization theory have been successfully applied to determine the effective conductivity and permeability of certain composite materials and porous media. However, their range of applications is usually limited by restrictive assumptions on the media, such as scale separation and periodicity [7,44]. They are also expensive to use for solving problems with many separate scales since the cost of computation grows exponentially with the number of scales. But for the multiscale method, the number of scales does not increase the overall computational cost exponentially. The upscaling methods are more general and have been applied to problems with random coefficients with partial success (cf. [23,49]). But the design principle is strongly motivated by the homogenization theory for periodic structures. Their application to nonperiodic structures is not always guaranteed to work.

We remark that the idea of using base functions governed by the differential equations has been applied to convection-diffusion equation with boundary layers (see, e.g., [6] and references therein). With a motivation different from ours, Babuska et al. applied a similar idea to 1-D problems [5] and to a special class of 2-D problems with the coefficient varying locally in one direction [4]. However, most of these methods are based on the special property of the harmonic average in one-dimensional elliptic problems. As indicated by our convergence analysis, there is a fundamental difference between one-dimensional problems and genuinely multi-dimensional problems. Special complications such as the resonance between the mesh scale and the physical scale never occur in the corresponding 1-D problems.

4.1 Multiscale Finite Element Methods for Elliptic PDEs.

In this section we consider the multiscale finite element method applied to the following problem

$$L_\varepsilon u := -\nabla \cdot (a(\tfrac{x}{\varepsilon})\nabla u) = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \Gamma = \partial\Omega, \tag{4.1}$$

where Ω is a convex polygon in $\mathbb{R}^2$. ε is assumed to be a small parameter, and $a(x) = (a_{ij}(x/\varepsilon))$ is symmetric and satisfies $\alpha|\xi|^2 \le a_{ij}\xi_i\xi_j \le \beta|\xi|^2$, for all $\xi \in \mathbb{R}^2$ and with $0 < \alpha < \beta$. Furthermore, $a_{ij}(y)$ are smooth periodic function in y in a unit cube Y. We will always assume that $f \in L^2(\Omega)$. In fact, the smoothness assumption on a_{ij} can be relaxed. In [27], Efendiev has proved convergence of the multiscale finite element method in the case where a_{ij} is only piecewise continuous. Efendiev has also obtained convergence of MsFEM in the case where a_{ij} is random [27].

Let u_0 be the solution of the homogenized equation

$$L_0 u_0 := -\nabla \cdot (a^*\nabla u_0) = f \quad \text{in } \Omega, \quad u_0 = 0 \quad \text{on } \Gamma, \tag{4.2}$$

where $\Gamma = \partial\Omega$ and

$$a^*_{ij} = \frac{1}{|Y|}\int_Y a_{ik}(y)(\delta_{kj} - \frac{\partial \chi^j}{\partial y_k})\, dy,$$

and $\chi^j(y)$ is the periodic solution of the cell problem

$$-\nabla_y \cdot (a(y)\nabla_y \chi^j) = \frac{\partial}{\partial y_i} a_{ij}(y) \quad \text{in } Y, \quad \int_Y \chi^j(y)\, dy = 0.$$

It is clear that $u_0 \in H^2(\Omega)$ since Ω is a convex polygon. Denote by $u_1(x,y) = -\chi^j(y)\frac{\partial u_0(x)}{\partial x_j}$ and let θ_ε be the solution of the problem

$$L_\varepsilon \theta_\varepsilon = 0 \quad \text{in } \Omega, \quad \theta_\varepsilon(x) = u_1(x, \tfrac{x}{\varepsilon}) \quad \text{on } \Gamma. \tag{4.3}$$

Our analysis of the multiscale finite element method relies on the following homogenization result obtained by Moskow and Vogelius [52].

Lemma 4.1. *Let $u_0 \in H^2(\Omega)$ be the solution of (4.2), $\theta_\varepsilon \in H^1(\Omega)$ be the solution to (4.3) and $u_1(x) = -\chi^j(x/\varepsilon)\partial u_0(x)/\partial x_j$. Then there exists a constant C independent of u_0, ε and Ω such that*

$$\| u - u_0 - \varepsilon(u_1 - \theta_\varepsilon) \|_{1,\Omega} \leq C\varepsilon(| u_0 |_{2,\Omega} + \| f \|_{0,\Omega}).$$

Now we are going to introduce the multiscale finite element methods. Let $\mathcal{T}_h$ be a regular partition of Ω into triangles. Let $\{x_j\}_{j=1}^J$ be the interior nodes of the mesh $\mathcal{T}_h$ and $\{\psi_j\}_{j=1}^J$ be the nodal basis of the standard linear finite element space $W_h \subset H_0^1(\Omega)$. Denote by $S_i = \text{supp}(\psi_i)$ and define ϕ^i with support in S_i as follows:

$$L_\varepsilon \phi^i = 0 \ \text{ in } K, \quad \phi^i = \psi_i \ \text{ on } \partial K \ \ \forall\, K \in \mathcal{T}_h, K \subset S_i. \tag{4.4}$$

It is obvious that $\phi^i \in H_0^1(S_i) \subset H_0^1(\Omega)$. Finally, let $V_h \subset H_0^1(\Omega)$ be the finite element space spanned by $\{\phi^i\}_{i=1}^J$.

With above notation we can introduce the following discrete problem: find $u_h \in V_h$ such that

$$(a(\tfrac{x}{\varepsilon})\nabla u_h, \nabla v_h) = (f, v_h) \ \ \forall\ v_h \in V_h, \tag{4.5}$$

where and hereafter we denote by $(\cdot,\cdot)$ the L^2 inner product in $L^2(\Omega)$.

As we will see later, the choice of boundary conditions in defining the multiscale bases will play a crucial role in approximating the multiscale solution. Intuitively, the boundary condition for the multiscale base function should reflect the multiscale oscillation of the solution u across the boundary of the coarse grid element. By choosing a linear boundary condition for the base function, we will create a mismatch between the exact solution u and the finite element approximation across the element boundary. In the next section, we will discuss this issue further and introduce an over-sampling technique to alleviate this difficulty. The over-sampling technique plays an important role when we need to reconstruct the local fine grid velocity field from a coarse grid pressure computation for two-phase flows. This technique enables us to remove the artificial numerical boundary layer across the coarse grid boundary element.

We remark that the multiscale finite element method with linear boundary conditions for the multiscale base functions is similar in spirit to the residual-free bubbles finite element method [11] and the variational multiscale method [42,12]. In a recent paper [57], Dr. G. Sangalli derives a multiscale method based on the residual-free bubbles formulation in [11] and compares it with the multiscale finite element method described here. There are many striking similarities between the two approaches. In Section 6, we will discuss the variational multiscale method in some more detail and compare it with the multiscale finite element method.

To gain some insight into the multiscale finite element method, we next perform an error analysis for the multiscale finite element method in the simplest case, i.e. we use linear boundary conditions for the multiscale base functions.

4.2 Error Estimates ($h < \varepsilon$)

The starting point is the well-known Cea's lemma.

Lemma 4.2. *Let u be the solution of (4.1) and u_h be the solution of (4.5). Then we have*

$$\| u - u_h \|_{1,\Omega} \leq C \inf_{v_h \in V_h} \| u - v_h \|_{1,\Omega}.$$

Let $\Pi_h : C(\bar{\Omega}) \to W_h \subset H_0^1(\Omega)$ be the usual Lagrange interpolation operator:

$$\Pi_h u(x) = \sum_{j=1}^{J} u(x_j)\psi_j(x) \quad \forall\ u \in C(\bar{\Omega})$$

and $I_h : C(\bar{\Omega}) \to V_h$ be the corresponding interpolation operator defined through the multiscale base function ϕ

$$I_h u(x) = \sum_{j=1}^{J} u(x_j)\phi^j(x) \quad \forall\ u \in C(\bar{\Omega}).$$

¿From the definition of the basis function ϕ^i in (4.4) we have

$$L_\varepsilon(I_h u) = 0 \ \text{ in } K, \quad I_h u = \Pi_h u \ \text{ on } \partial K, \tag{4.6}$$

for any $K \in \mathcal{T}_h$.

Lemma 4.3. *Let $u \in H^2(\Omega)$ be the solution of (4.1). Then there exists a constant C independent of h, ε such that*

$$\| u - I_h u \|_{0,\Omega} + h\| u - I_h u \|_{1,\Omega} \leq Ch^2(| u |_{2,\Omega} + \| f \|_{0,\Omega}). \tag{4.7}$$

Proof. At first it is known from the standard finite element interpolation theory that

$$\| u - \Pi_h u \|_{0,\Omega} + h\| u - \Pi_h u \|_{1,\Omega} \leq Ch^2(| u |_{2,\Omega} + \| f \|_{0,\Omega}). \tag{4.8}$$

On the other hand, since $\Pi_h u - I_h u = 0$ on ∂K, the standard scaling argument yields

$$\| \Pi_h u - I_h u \|_{0,K} \leq Ch|\Pi_h u - I_h u|_{1,K} \quad \forall\ K \in \mathcal{T}_h. \tag{4.9}$$

To estimate $|\Pi_h u - I_h u|_{1,K}$ we multiply the equation in (4.6) by $I_h u - \Pi_h u \in H_0^1(K)$ to get

$$(a(\tfrac{x}{\varepsilon})\nabla I_h u, \nabla(I_h u - \Pi_h u))_K = 0,$$

where $(\cdot,\cdot)_K$ denotes the L^2 inner product of $L^2(K)$. Thus, upon using the equation in (4.1), we get

$$\begin{aligned}
&(a(\tfrac{x}{\varepsilon})\nabla(I_h u - \Pi_h u), \nabla(I_h u - \Pi_h u))_K \\
&= (a(\tfrac{x}{\varepsilon})\nabla(u - \Pi_h u), \nabla(I_h u - \Pi_h u))_K - (a(\tfrac{x}{\varepsilon})\nabla u, \nabla(I_h u - \Pi_h u))_K \\
&= (a(\tfrac{x}{\varepsilon})\nabla(u - \Pi_h u), \nabla(I_h u - \Pi_h u))_K - (f, I_h u - \Pi_h u)_K.
\end{aligned}$$

This implies that

$$|I_h u - \Pi_h u|_{1,K} \leq Ch| u |_{2,K} + \| I_h u - \Pi_h u \|_{0,K} \| f \|_{0,K}.$$

Hence

$$|I_h u - \Pi_h u|_{1,K} \leq Ch(| u |_{2,K} + \| f \|_{0,K}), \tag{4.10}$$

where we have used (4.9). Now the lemma follows from (4.8)-(4.10). □

In conclusion, we have the following estimate by using Lemmas 4.2-4.3.

Theorem 4.1. *Let $u \in H^2(\Omega)$ be the solution of (4.1) and $u_h \in V_h$ be the solution of (4.5). Then we have*

$$\| u - u_h \|_{1,\Omega} \leq Ch(| u |_{2,\Omega} + \| f \|_{0,\Omega}). \tag{4.11}$$

Note that the estimate (4.11) blows up like h/ε as $\varepsilon \to 0$ since $| u |_{2,\Omega} = O(1/\varepsilon)$. This is insufficient for practical applications. In next subsection we derive an error estimate which is uniform as $\varepsilon \to 0$.

4.3 Error Estimates ($h > \varepsilon$)

In this section, we will show that the multiscale finite element method gives a convergence result uniform in ε as ε tends to zero. This is the main feature of this multiscale finite element method over the traditional finite element method. The main result in this subsection is the following theorem.

Theorem 4.2. *Let $u \in H^2(\Omega)$ be the solution of (4.1) and $u_h \in V_h$ be the solution of (4.5). Then we have*

$$\| u - u_h \|_{1,\Omega} \leq C(h + \varepsilon) \| f \|_{0,\Omega} + C \left(\frac{\varepsilon}{h}\right)^{1/2} \| u_0 \|_{1,\infty,\Omega}, \tag{4.12}$$

where $u_0 \in H^2(\Omega) \cap W^{1,\infty}(\Omega)$ is the solution of the homogenized equation (4.2).

To prove the theorem, we first denote by

$$u_{\rm I}(x) = I_h u_0(x) = \sum_{j=1}^{J} u_0(x_j) \phi^j(x) \in V_h.$$

¿From (4.6) we know that $L_\varepsilon u_{\rm I} = 0$ in K and $u_{\rm I} = \Pi_h u_0$ on ∂K for any $K \in \mathcal{T}_h$. The homogenization theory in Lemma 3.1 implies that

$$\| u_{\rm I} - u_{\rm I0} - \varepsilon (u_{\rm I1} - \theta_{{\rm I}\varepsilon}) \|_{1,K} \leq C\varepsilon (\| f \|_{0,K} + | u_{\rm I0} |_{2,K}), \tag{4.13}$$

where u_{I0} is the solution of the homogenized equation on K:

$$L_0 u_{I0} = 0 \ \text{ in } K, \quad u_{I0} = \Pi_h u_0 \ \text{ on } \partial K, \tag{4.14}$$

u_{I1} is given by the relation

$$u_{I1}(x,y) = -\chi^j(y)\frac{\partial u_{I0}}{\partial x_j} \ \text{ in } K, \tag{4.15}$$

and $\theta_{I\varepsilon} \in H^1(K)$ is the solution of the problem:

$$L_\varepsilon \theta_{I\varepsilon} = 0 \ \text{ in } K, \quad \theta_{I\varepsilon}(x) = u_{I1}(x, \tfrac{x}{\varepsilon}) \ \text{ on } \partial K. \tag{4.16}$$

It is obvious from (4.14) that

$$u_{I0} = \Pi_h u_0 \ \text{ in } K, \tag{4.17}$$

since $\Pi_h u_0$ is linear on K. From Lemma 3.1 and (4.13) we obtain that

$$\begin{aligned} \| u - u_I \|_{1,\Omega} \le \| u_0 - u_{I0} \|_{1,\Omega} + \| \varepsilon(u_1 - u_{I1}) \|_{1,\Omega} \\ + \| \varepsilon(\theta_\varepsilon - \theta_{I\varepsilon}) \|_{1,\Omega} + C\varepsilon \| f \|_{0,\Omega}, \end{aligned} \tag{4.18}$$

where we have used the regularity estimate $\| u_0 \|_{2,\Omega} \le C \| f \|_{0,\Omega}$. Now it remains to estimate the terms at the right-hand side of (4.18).

Lemma 4.4. *We have*

$$\| u_0 - u_{I0} \|_{1,\Omega} \le Ch \| f \|_{0,\Omega}, \tag{4.19}$$

$$\| \varepsilon(u_1 - u_{I1}) \|_{1,\Omega} \le C(h+\varepsilon) \| f \|_{0,\Omega}. \tag{4.20}$$

Proof. The estimate (4.19) is a direct consequence of the standard finite element interpolation theory since $u_{I0} = \Pi_h u_0$ by (4.17). Next we note that $\chi^j(x/\varepsilon)$ satisfies

$$\| \chi^j \|_{0,\infty,\Omega} + \varepsilon \| \nabla \chi^j \|_{0,\infty,\Omega} \le C \tag{4.21}$$

for some constant C independent of h and ε. Thus we have, for any $K \in \mathcal{T}_h$,

$$\| \varepsilon(u_1 - u_{I1}) \|_{0,K} \le C\varepsilon \| \chi^j \frac{\partial}{\partial x_j}(u_0 - \Pi_h u_0) \|_{0,K} \le Ch\varepsilon | u_0 |_{2,K},$$

$$\begin{aligned} \| \varepsilon \nabla (u_1 - u_{I1}) \|_{0,K} &= \varepsilon \| \nabla (\chi^j \frac{\partial (u_0 - \Pi_h u_0)}{\partial x_j}) \|_{0,K} \\ &\le C \| \nabla (u_0 - \Pi_h u_0) \|_{0,K} + C\varepsilon | u_0 |_{2,K} \\ &\le C(h+\varepsilon) | u_0 |_{2,K}. \end{aligned}$$

This completes the proof. □

Lemma 4.5. *We have*

$$\| \varepsilon\theta_\varepsilon \|_{1,\Omega} \le C\sqrt{\varepsilon}\| u_0 \|_{1,\infty,\Omega} + C\varepsilon| u_0 |_{2,\Omega}. \tag{4.22}$$

Proof. Let $\zeta \in C_0^\infty(\mathbb{R}^2)$ be the cut-off function which satisfies $\zeta \equiv 1$ in $\Omega\backslash\Omega_{\delta/2}$, $\zeta \equiv 0$ in Ω_δ, $0 \le \zeta \le 1$ in $\mathbb{R}^2$, and $|\nabla\zeta| \le C/\delta$ in Ω, where for any $\delta > 0$ sufficiently small, we denote by Ω_δ as

$$\Omega_\delta = \{x \in \Omega : \mathrm{dist}(x, \partial\Omega) \ge \delta\}.$$

With this definition, it is clear that $\theta_\varepsilon - \zeta u_1 = \theta_\varepsilon + \zeta(\chi^j \partial u_0/\partial x_j) \in H_0^1(\Omega)$. Multiplying the equation in (4.3) by $\theta_\varepsilon - \zeta u_1$, we get

$$(a(\tfrac{x}{\varepsilon})\nabla\theta_\varepsilon, \nabla(\theta_\varepsilon + \zeta\chi^j \frac{\partial u_0}{\partial x_j})) = 0,$$

which yields, by using (4.21),

$$\begin{aligned}
\| \nabla\theta_\varepsilon \|_{0,\Omega} &\le C\| \nabla(\zeta\chi^j \partial u_0/\partial x_j) \|_{0,\Omega} \\
&\le C\| \nabla\zeta \cdot \chi^j \partial u_0/\partial x_j \|_{0,\Omega} + C\| \zeta\nabla\chi^j \partial u_0/\partial x_j \|_{0,\Omega} \\
&\quad + C\| \zeta\chi^j \partial^2 u_0/\partial^2 x_j \|_{0,\Omega} \\
&\le C\sqrt{|\partial\Omega|\cdot\delta}\frac{D}{\delta} + C\sqrt{|\partial\Omega|\cdot\delta}\frac{D}{\varepsilon} + C| u_0 |_{2,\Omega},
\end{aligned} \tag{4.23}$$

where $D = \| u_0 \|_{1,\infty,\Omega}$ and the constant C is independent of the domain Ω. From (4.23) we have

$$\begin{aligned}
\| \varepsilon\theta_\varepsilon \|_{0,\Omega} &\le C(\frac{\varepsilon}{\sqrt{\delta}} + \sqrt{\delta})\| u_0 \|_{1,\infty,\Omega} + C\varepsilon| u_0 |_{2,\Omega} \\
&\le C\sqrt{\varepsilon}\| u_0 \|_{1,\infty,\Omega} + C\varepsilon| u_0 |_{2,\Omega}.
\end{aligned} \tag{4.24}$$

Moreover, by applying the maximum principle to (4.3), we get

$$\| \theta_\varepsilon \|_{0,\infty,\Omega} \le \| \chi^j \partial u_0/\partial x_j \|_{0,\infty,\partial\Omega} \le C\| u_0 \|_{1,\infty,\Omega}. \tag{4.25}$$

Combining (4.24) and (4.25) completes the proof. □

Lemma 4.6. *We have*

$$\| \varepsilon\theta_{\mathrm{I}\varepsilon} \|_{1,\Omega} \le C\Big(\frac{\varepsilon}{h}\Big)^{1/2}\| u_0 \|_{1,\infty,\Omega}. \tag{4.26}$$

Proof. First we remember that for any $K \in \mathcal{T}_h$, $\theta_{\mathrm{I}\varepsilon} \in H^1(K)$ satisfies

$$L_\varepsilon\theta_{\mathrm{I}\varepsilon} = 0 \ \text{ in } K, \quad \theta_{\mathrm{I}\varepsilon} = -\chi^j(\frac{x}{\varepsilon})\frac{\partial(\Pi_h u_0)}{\partial x_j} \ \text{ on } \partial K. \tag{4.27}$$

By applying maximum principle and (4.21) we get

$$\| \theta_{\mathrm{I}\varepsilon} \|_{0,\infty,K} \le \| \chi^j \partial(\Pi_h u_0)/\partial x_j \|_{0,\infty,\partial K} \le C\| u_0 \|_{1,\infty,K}.$$

Thus we have

$$\| \varepsilon \theta_{I\varepsilon} \|_{0,\Omega} \leq C\varepsilon \| u_0 \|_{1,\infty,\Omega}. \tag{4.28}$$

On the other hand, since the constant C in (4.23) is independent of Ω, we can apply the same argument leading to (4.23) to obtain

$$\begin{aligned} \| \varepsilon \nabla \theta_{I\varepsilon} \|_{0,K} &\leq C\varepsilon \| \Pi_h u_0 \|_{1,\infty,K} (\sqrt{|\partial K|}/\sqrt{\delta} + \sqrt{|\partial K| \cdot \delta}/\varepsilon) + C\varepsilon | \Pi_h u_0 |_{2,K} \\ &\leq C\sqrt{h} \| u_0 \|_{1,\infty,K} (\frac{\varepsilon}{\sqrt{\delta}} + \sqrt{\delta}) \\ &\leq C\sqrt{h\varepsilon} \| u_0 \|_{1,\infty,K}, \end{aligned}$$

which implies that

$$\| \varepsilon \nabla \theta_{I\varepsilon} \|_{0,\Omega} \leq C \left(\frac{\varepsilon}{h}\right)^{1/2} \| u_0 \|_{1,\infty,\Omega}.$$

This completes the proof. □

Proof (of Theorem 3.2.). The theorem is now a direct consequence of (4.18) and the Lemmas 4.4-4.6 and the regularity estimate $\| u_0 \|_{2,\Omega} \leq C \| f \|_{0,\Omega}$. □

Remark 4.1. As we pointed out earlier, the multiscale FEM indeed gives correct homogenized result as ε tends to zero. This is in contrast with the traditional FEM which does not give the correct homogenized result as $\varepsilon \to 0$. The error would grow like $O(h^2/\varepsilon^2)$. On the other hand, we also observe that when $h \sim \varepsilon$, the multiscale method attains large error in both H^1 and L^2 norms. This is what we call the *resonance* effect between the grid scale (h) and the small scale (ε) of the problem. This estimate reflects the intrinsic scale interaction between the two scales in the *discrete* problem. Our extensive numerical experiments confirm that this estimate is indeed generic and sharp. From the viewpoint of practical applications, it is important to reduce or completely remove the resonance error for problems with many scales since the chance of hitting a resonance sampling is high. In the next subsection, we propose an over-sampling method to overcome this difficulty.

4.4 The Over-Sampling Technique

As illustrated by our error analysis, large errors result from the "resonance" between the grid scale and the scales of the continuous problem. For the two-scale problem, the error due to the resonance manifests as a ratio between the wavelength of the small scale oscillation and the grid size; the error becomes large when the two scales are close. A deeper analysis shows that the boundary layer in the first order corrector seems to be the main source of the resonance effect. By a judicious choice of boundary conditions for the base

function, we can eliminate the boundary layer in the first order corrector. This would give a nice conservative difference structure in the discretization, which in turn leads to *cancellation of resonance errors* and gives an improved rate of convergence.

Motivated by our convergence analysis, we propose an *over-sampling* method to overcome the difficulty due to scale resonance [38]. The idea is quite simple and easy to implement. Since the boundary layer in the first order corrector is thin, $O(\varepsilon)$, we can sample in a domain with size larger than $h + \varepsilon$ and use only the interior sampled information to construct the bases; here, h is the mesh size and ε is the small scale in the solution. By doing this, we can reduce the influence of the boundary layer in the larger sample domain on the base functions significantly. As a consequence, we obtain an improved rate of convergence.

Specifically, let ψ^j be the base functions satisfying the homogeneous elliptic equation in the larger domain $S \supset K$. We then form the actual base ϕ^i by linear combination of ψ^j,

$$\phi^i = \sum_{j=1}^{d} c_{ij} \psi^j \ .$$

The coefficients c_{ij} are determined by condition $\phi^i(\mathbf{x}_j) = \delta_{ij}$. The corresponding θ^i_ε for ϕ^i are now free of boundary layers. Our extensive numerical experiments have demonstrated that the over-sampling technique does improve the numerical error substantially in many applications. On the other hand, the over-sampling technique results in a *non-conforming* MsFEM method. In [28], we perform a careful estimate of the nonconforming errors in both H^1 norm and the L^2 norm. The analysis shows that the non-conforming error is indeed small, consistent with our numerical results [38,39]. Our analysis also reveals another source of resonance, which is the mismatch between the mesh size and the "perfect" sample size. In case of a periodic structure, the "perfect" sample size is the length of an integer multiple of the period. We call the new resonance the "cell resonance". In the error expansion, this resonance effect appears as a *higher* order correction. In numerical computations, we found that the cell resonance error is generically small, and is rarely observed in practice. Nonetheless, it is possible to completely eliminate this cell resonance error by using the over-sampling technique to construct the base functions but using piecewise linear functions as test functions. This reduces the nonconforming error and eliminates the resonance error completely.

4.5 Performance and Implementation Issues

The multiscale method given in the previous section is fairly straightforward to implement. Here, we outline the implementation and define some notations that are used in the discussion below. We consider solving problems in a unit

square domain. Let N be the number of elements in the x and y directions. The mesh size is thus $h = 1/N$. To compute the base functions, each element is discretized into $M \times M$ subcell elements with mesh size $h_s = h/M$. To implement the over-sampling method, we partition the domain into sampling domains and each of them contains many elements. From the analysis and numerical tests, the size of the sampling domains can be chosen freely as long as the boundary layer is avoided. In practice, though, one wants to maximize the efficiency of over-sampling by choosing the largest possible sample size which reduces the redundant computation of overlapping domains to a minimum.

In general, the multiscale (sampling) base functions are constructed numerically, except for certain special cases. They are solved in each K or S using standard FEM. The linear systems are solved using a robust multigrid method with matrix dependent prolongation and ILLU smoothing (MG-ILLU, see [62]). The global linear system on Ω is solved using the same method. Numerical tests show that the accuracy of the final solution is insensitive to the accuracy of base functions.

Since the base functions are independent of each other, their construction can be carried out in parallel perfectly. In our parallel implementation of over-sampling, the sample domains are chosen such that they can be handled within each processor without communication. The multigrid solver is also modified to better suit the parallelization. In particular, the ILLU smoothing is replaced by Gauss-Seidel iterations. More implementation details can be found in [38].

Cost and Performance In practical computations, a large amount of overhead time comes from constructing the base functions. On a sequential machine, the operation count of our method is about twice that of a conventional FEM for a 2-D problem. However, due to good parallel efficiency, this difference is reduced significantly on a massively parallel computer. For example, using 256 processors on an Intel Paragon, our method with $N = 32$ and $M = 32$ only spends 9% more CPU time than the conventional linear FEM method using 1024×1024 elements [38]. Note that this comparison is made for a single solve of the problem. In practice, multiple solves are often required, then the overhead of base construction is negligible. A detailed study of MsFEM's parallel efficiency has been conducted in [38]. It was also found that MsFEM is helpful for improving multigrid convergence when the coefficient a_ε has very large contrast (i.e., the ratio between the maximum and minimum of a_ε).

Significant computational savings can be obtained for time dependent problems (such as two-phase flows) by constructing the multiscale bases adaptively. Multiscale base functions are updated only for those coarse grid elements where the saturation changes significantly. In practice, the number of such coarse grid elements are small. They are concentrated near the interface

separating oil and water. Also, the cost of solving a base function in a small cell is more efficient than solving the fine grid problem globally because the condition number for solving the local base function in each coarse grid element is much smaller than that of the corresponding global fine grid pressure system. Thus, updating a small number of multiscale base functions dynamically is much cheaper than updating the fine grid pressure field globally.

Another advantage of the multiscale finite element method is its ability to scale down the size of a large scale problem. This offers a big saving in computer memory. For example, let N be the number of elements in each spatial direction, and M be the number of subcell elements in each direction for solving the base functions. Then there are total $(MN)^n$ (n is dimension) elements at the fine grid level. For a traditional FEM, the computer memory needed for solving the problem on the fine grid is $O(M^n N^n)$. In contrast, MsFEM requires only $O(M^n + N^n)$ amount of memory. For a typical value of $M = 32$ in a 2-D problem, the traditional FEM needs about 1000 times more memory than MsFEM.

Convergence and Accuracy Since we need to use an additional grid to compute the base function numerically, it makes sense to compare our MsFEM with a traditional FEM at the subcell grid, $h_s = h/M$. Note that MsFEM only captures the solution at the coarse grid h, while FEM tries to resolve the solution at the fine grid h_s. Our extensive numerical experiments demonstrate that the accuracy of MsFEM on the coarse grid h is comparable to that of FEM on the fine grid. In some cases, MsFEM is even more accurate than the FEM (see below and the next section).

As an example, in Table 4.1 we present the result for

$$a(\mathbf{x}/\varepsilon) = \frac{2 + P\sin(2\pi x/\varepsilon)}{2 + P\cos(2\pi y/\varepsilon)} + \frac{2 + \sin(2\pi y/\varepsilon)}{2 + P\sin(2\pi x/\varepsilon)} \quad (P = 1.8), \tag{4.29}$$

$$f(\mathbf{x}) = -1 \qquad \text{and} \qquad u|_{\partial\Omega} = 0. \tag{4.30}$$

The convergence of three different methods are compared for fixed $\varepsilon/h = 0.64$, where "-L" indicates that linear boundary condition is imposed on the multiscale base functions, "os" indicates the use of over-sampling, and LFEM stands for standard FEM with linear base functions. We see clearly the scale

N	ε	MsFEM-L		MsFEM-os-L		LFEM	
		$\|\|E\|\|_{l^2}$	rate	$\|\|E\|\|_{l^2}$	rate	MN	$\|\|E\|\|_{l^2}$
16	0.04	3.54e-4		7.78e-5		256	1.34e-4
32	0.02	3.90e-4	-0.14	3.83e-5	1.02	512	1.34e-4
64	0.01	4.04e-4	-0.05	1.97e-5	0.96	1024	1.34e-4
128	0.005	4.10e-4	-0.02	1.03e-5	0.94	2048	1.34e-4

Table 4.1. Convergence for periodic case.

resonance in the results of MsFEM-L and the (almost) first order convergence

(i.e., no resonance) in MsFEM-os-L. Evident also is the error of MsFEM-os-L being smaller than those of LFEM obtained on the fine grid. In [40,38], more extensive convergence tests have been presented.

4.6 Applications

Flow in Porous Media One of the main application of our multiscale method is the flow and transport through porous media. This is a fundamental problem in hydrology and petroleum engineering. Here, we apply MsFEM to solve the single phase flow, which is a good test problem in practice.

We model the porous media by random distributions of a_ε generated using a spectral method. In fact, $a_\varepsilon = \alpha 10^{\beta p}$, where p is a random field represents porosity, and α and β are scaling constants to give the desired contrast of a_ε. In particular, we have tested the method for a porous medium with a statistically fractal porosity field (see Figure 4.1). The fractal dimension is 2.8. Such a model is widely used as the areal field in the oil industry. We note that the problem has a continuous scale because of the fractal distribution.

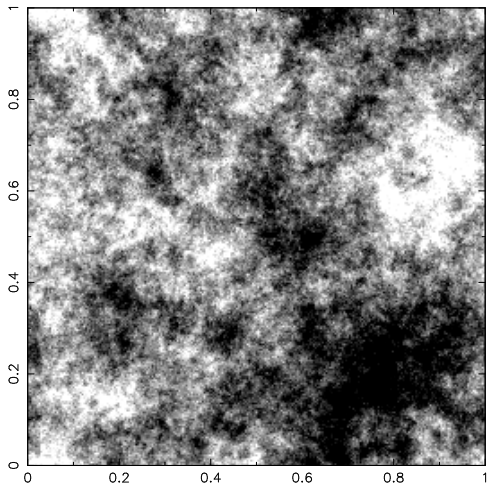

Fig. 4.1. Porosity field with fractal dimension of 2.8 generated using the spectral method.

The pressure field due to uniform injection is solved and the error is shown in Figure 3.2. The horizontal dash line indicates the error of the LFEM solution with $N = 2048$. The coarse-grid solutions are obtained with different number of elements, N, but fixed $NM = 2048$. We note that error of MsFEM-os-L almost coincide with that of the well-resolved solution obtained using LFEM. However, MsFEM without over-sampling is less accurate. MsFEM-O indicates that oscillatory boundary conditions, obtained from solving some reduced 1-D elliptic equations along $\partial\mathbf{K}$ (see [38]), are imposed on the base

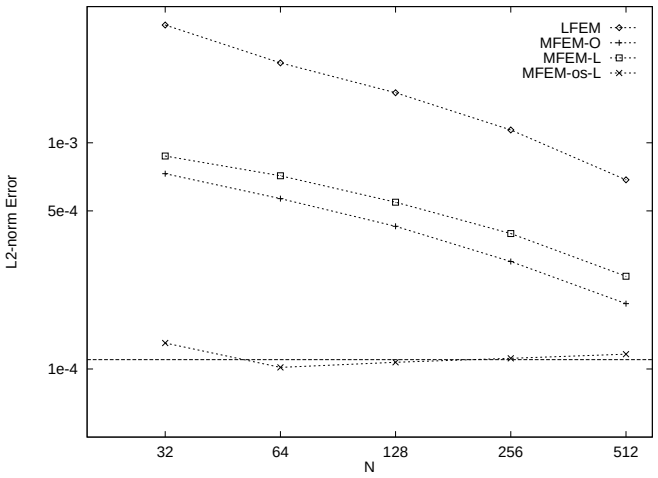

Fig. 4.2. The l^2-norm error of the solutions using various schemes for a fractal distributed permeability field.

functions. The decay of error in MsFEM is because of the decay of small scales in a_ε. The next figure shows the results for a log-normally distributed a_ε. In this case, the effect of scale resonance shows clearly for MsFEM-L, i.e., the error increases as h approaches ε. Here $\varepsilon \sim 0.004$ roughly equals the correlation length. Using the oscillatory boundary conditions (MsFEM-O) gives better results, but it does not completely eliminate resonance. On the other hand, the multiscale method with over-sampling agrees extremely well with the well-resolved calculation. One may wonder why the errors do not decrease as the number of coarse grid elements increase. This is because we use the same subgrid mesh size, which is the same as the well-resolved grid size, to construct the base functions for various coarse grid sizes ($N = 32, 64, 128$, etc). In some special cases, one can construct multiscale base functions analytically. In this case, the errors for the coarse grid computations will indeed decrease as the number of coarse grid elements increase.

Fine Scale Recovery To solve transport problems in the subsurface formations, as in oil reservoir simulations, one needs to compute the velocity field from the elliptic equation for pressure, i.e $\mathbf{v} = -a_\varepsilon \nabla u$, here u is pressure. In some applications involving isotropic media, the cell-averaged velocity is sufficient, as shown by some computations using the local upscaling methods (cf. [23]). However, for anisotropic media, especially layered ones (Figure 4.4), the velocity in some thin channels can be much higher than the cell average, and these channels often have dominant effects on the transport solutions. In this case, the information about fine scale velocity becomes vitally important. Therefore, an important question for all upscaling methods is how to take those fast-flow channels into account.

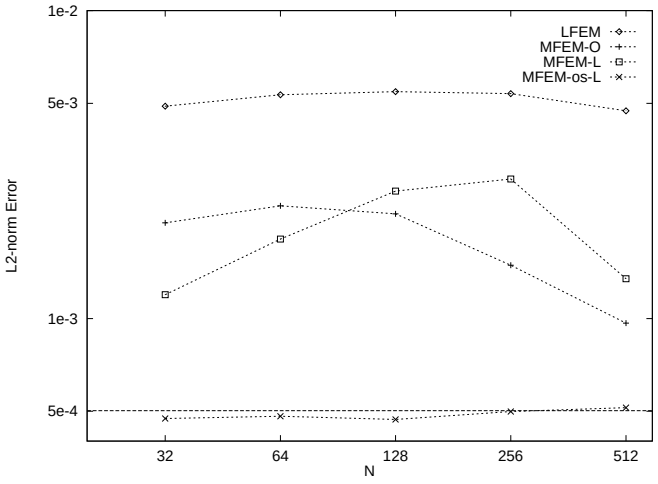

Fig. 4.3. The l^2-norm error of the solutions using various schemes for a log-normally distributed permeability field.

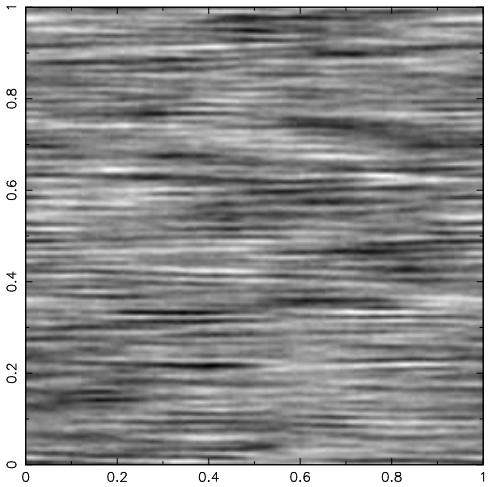

Fig. 4.4. A random porosity field with layered structure.

For MsFEM, the fine scale velocity can be easily recovered from the multi-scale base functions, noting that they provide interpolations from the coarse h-grid to the fine h_s-grid. Using the over-sampling technique, the error in velocity is $O(\varepsilon/h)$, as proved in [28]. We remark that the resonance effect seems unavoidable in the velocity. On the other hand, our numerical tests indicate that the error is small when $\varepsilon \approx h$. The cell-averaged velocity can also be obtained and its error is even smaller.

To demonstrate the accuracy of the recovered velocity and effect of small-scale velocity on the transport problem, we show the fractional flow result of a "tracer" test using the layered medium in Figure 4.4: a fluid with red color

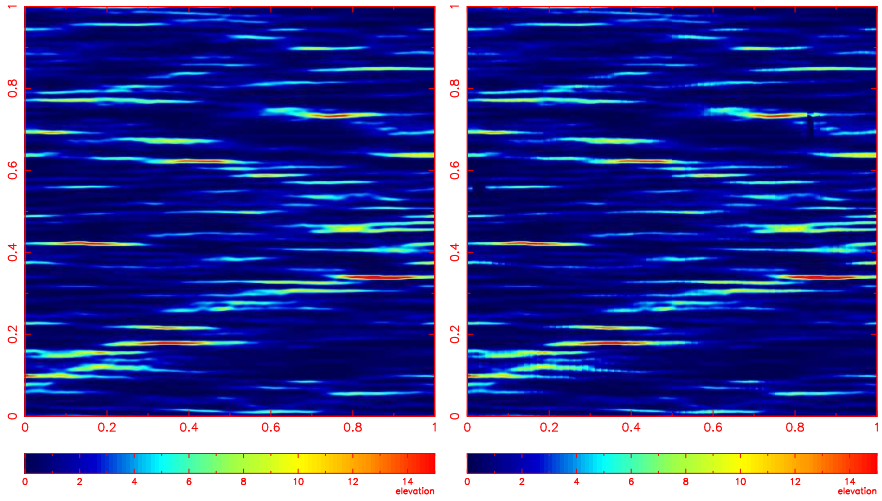

Fig. 4.5. (a): Fine grid horizontal velocity field, $N = 1024$. (b): Recovered horizontal velocity field from the coarse grid $N = 64$ calculation using multiscale bases.

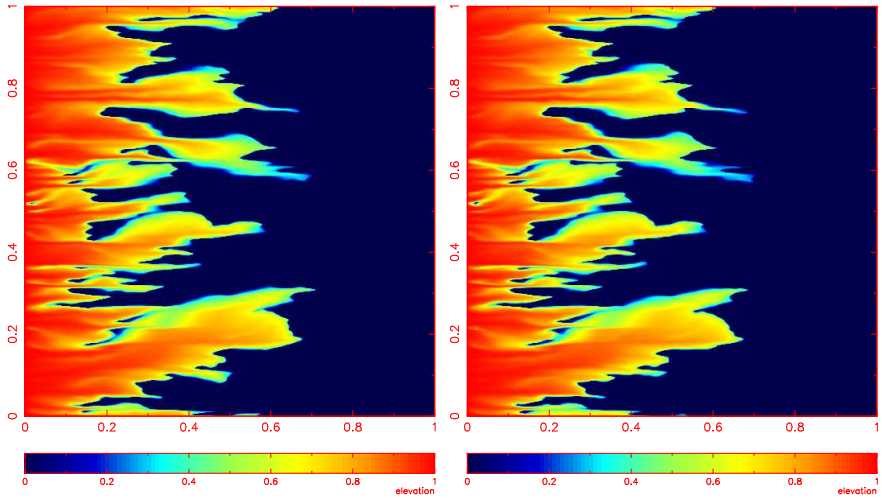

Fig. 4.6. (a): Fine grid saturation at $t = 0.06$, $N = 1024$. (b): Saturation computed using the recovered velocity field from the coarse grid calculation.

originally saturating the medium is displaced by the same fluid with blue color injected by flow in the medium at the left boundary, where the flow is created by a unit horizontal pressure drop. The linear convection equation is solved to compute the saturation of the red fluid (for details, see [24]). To demonstrate that we can recover the fine grid velocity field from the coarse grid pressure calculation, we plot the horizontal velocity fields obtained by two methods. In Figure 4.5a, we plot the horizontal velocity field obtained by using a fine grid ($N = 1024$) calculation. In Figure 4.5b, we plot the same

horizontal velocity field obtained by using the coarse grid pressure calculation with $N = 64$ and using the multiscale finite element bases to interpolate the fine grid velocity field. We can see that the recovered velocity field captures very well the layer structure in the fine grid velocity field. Further, we use the recovered fine grid velocity field to compute the saturation in time. In Figure 4.6a, we plot the saturation at $t = 0.06$ obtained by the fine grid calculation. Figure 4.6b shows the corresponding saturation obtained using the recovered velocity field from the coarse grid calculation. The agreement is striking.

We also check the fractional flow curves obtained by the two calculations. The fractional flow of the red fluid, defined as $F = \int S_{red} v_x \, dy / \int v_x \, dy$ (S being the saturation), at the right boundary is shown in Figure 4.7. The top pair of curves are the solutions of the transport problem using the cell-averaged velocity obtained from a well-resolved solution and from MsFEM; the bottom pair are solutions using well-resolved fine scale velocity and the recovered fine scale velocity from the MsFEM calculation. Two conclusions can be made from the comparisons. First, the cell-averaged velocity may lead to a large error in the solution of the transport equation. Second, both recovered fine scale velocity and the cell-averaged velocity obtained from MsFEM give faithful reproductions of respective direct numerical solutions.

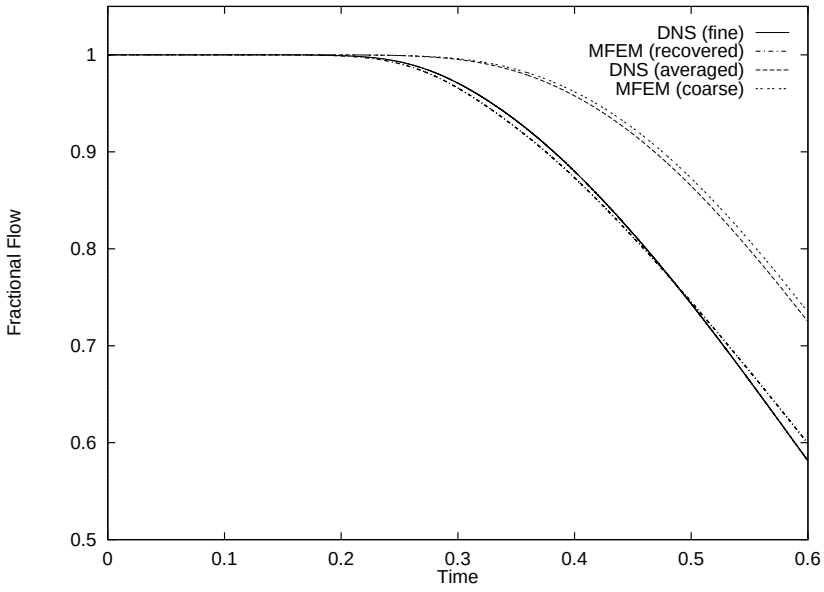

Fig. 4.7. Variation of fractional flow with time. DNS: well-resolved direct numerical solution using LFEM ($N = 512$). MsFEM: over-sampling is used ($N = 64$, $M = 8$).

Scale-up of one-phase flows The multiscale finite element method has been used in conjunction with some moment closure models to obtain an

upscaled method for one-phase flows, see, e.g. [27,29,16]. Note that the multiscale finite element method presented above does not conserve mass. For long time integration, it may lead to significant loss of mass. This is an undesirable feature of the method. In a recent work with Zhiming Chen [16], we have designed and analyzed a mixed multiscale finite element method, and we have applied this mixed method to study the scale up of one-phase flows and found that mass is conserved very well even for long time integration. Below we describe our results in some detail.

In its simplest form, neglecting the effect of gravity, compressibility, capillary pressure, and considering constant porosity and unit mobility, the governing equations for the flow transport in highly heterogeneous porous media can be described by the following partial differential equations [46], [63], and [27]

$$\mathrm{div}(\mathcal{K}(x)\nabla p) = 0, \tag{4.31}$$

$$\frac{\partial S}{\partial t} + \mathbf{v}\cdot\nabla S = 0, \tag{4.32}$$

where p is the pressure, S is the water saturation, $\mathcal{K}(x) = (\mathcal{K}_{ij}(x))$ is the relative permeability tensor, and $\mathbf{v} = -\mathcal{K}(x)\nabla p$ is the Darcy velocity. The highly heterogeneous properties of the medium are built into the permeability tensor $\mathcal{K}(x)$ which is generated through the use of sophisticated geological and geostatistical modeling tools. The detailed structure of the permeability coefficients makes the direct simulation of the above model infeasible. For example, it is common in real simulations to use millions of grid blocks, with each block having a dimension of tens of meters, whereas the permeability measured from cores is at a scale of centimeters [49]. This gives more than 10^5 degrees of freedom per spatial dimension in the computation. This makes a direct simulation to resolve all small scales prohibitive even with today's most powerful supercomputers. On the other hand, from an engineering perspective, it is often sufficient to predict the macroscopic properties of the solutions. Thus it is highly desirable to derive effective coarse grid models to capture the correct large scale solution without resolving the small scale features. Numerical upscaling is one of the commonly used approaches in practice.

Now we describe how the (mixed) multiscale finite element can be combined with the existing upscaling technique for the saturation equation (4.32) to get a complete coarse grid algorithm for the problem (4.31)-(4.32). The numerical upscaling of the saturation equation has been under intensive study in the literature [24,29,46,36,66,64]. Here, we use the upscaling method proposed in [29] and [27] to design an overall coarse grid model for the problem (4.31)-(4.32). The work of [29] for upscaling the saturation equation involves a moment closure argument. The velocity and the saturation are separated into a local mean quantity and a small scale perturbation with zero mean. For example, the Darcy velocity is expressed as $\mathbf{v} = \mathbf{v}_0 + \mathbf{v}'$ in (4.32), where

$\mathbf{v}_0$ is the average of velocity $\mathbf{v}$ over each coarse element, $\mathbf{v}' = (\mathbf{v}'_1, \mathbf{v}'_2)$ is the deviation of the fine scale velocity from its coarse scale average. After some manipulations, an average equation for the saturation S can be derived as follows [29]:

$$\frac{\partial S}{\partial t} + \mathbf{v}_0 \cdot \nabla S = \frac{\partial}{\partial x_i}\Big(D_{ij}(x,t)\frac{\partial S}{\partial x_j}\Big), \tag{4.33}$$

where the diffusion coefficients $D_{ij}(x,t)$ are defined by

$$D_{ii}(x,t) = \langle|\mathbf{v}'_i(x)|\rangle L_i^0(x,t), \quad D_{ij}(x,t) = 0, \quad \text{for } i \neq j,$$

$\langle|\mathbf{v}'_i(x)|\rangle$ stands for the average of $|\mathbf{v}'_i(x)|$ over each coarse element. $L_i^0(x,t)$ is the length of the coarse grid streamline in the x_i direction which starts at time t at point x, i.e.

$$L_i^0(x,t) = \int_0^t y_i(s)\, ds,$$

where $y(s)$ is the solution of the following system of ODEs

$$\frac{dy(s)}{ds} = \mathbf{v}_0(y(s)), \quad y(t) = x.$$

Note that the hyperbolic equation (4.32) is now replaced by a convection-diffusion equation. The convection-dominant parabolic equation (4.33) is solved by the characteristic linear finite element method [22], [56] in our simulation. The flow transport model (4.31)-(4.32) is solved in the coarse grid as follows:

1. Solve the pressure equation (4.31) by the over-sampling mixed multiscale finite element method and obtain the fine scale velocity field using the multiscale basis functions.
2. Compute the coarse grid average $\mathbf{v}_0$ and the fine scale deviation $\langle|\mathbf{v}'_i(x)|\rangle$ on the coarse grid.
3. At each time step, solve the convection-diffusion equation (4.33) by the characteristic linear finite element method on the coarse grid in which the lengths $L_i^0(x,t)$ of the streamline are computed for the center of each coarse grid element.

The mixed multiscale finite element method can be readily combined with the above upscaling model for the saturation equation. The local fine grid velocity $\mathbf{v}'$ will be constructed from the multiscale finite element base functions. The main cost in the above algorithm lies in the computation of multiscale bases which can be done a priori and completely in parallel. This algorithm is particularly attractive when multiple simulations must be carried out due to the change of boundary and source distribution as it is often the case in engineering applications. In such a situation, the cost of computing the multiscale

base functions is just an over-head. Moreover, once these base functions are computed, they can be used for subsequent time integration of the saturation. Because the evolution equation is now solved on a coarse grid, a larger time step can be used. This also offers additional computational saving. For many oil recovery problems, due to the excessively large fine grid data, upscaling is a necessary step before performing many simulations and realizations on the upscaled coarse grid model. If one can coarsen the fine grid by a factor of 10 in each dimension, the computational saving of the coarse grid model over the original fine model could be as large as a factor 10,000 (three space dimensions plus time).

We perform a coarse grid computation of the above algorithm on the coarse 64×64 mesh. The fractional flow curve using the above algorithm is depicted in Figure 4.8. It gives excellent agreement with the "exact" fractional flow curve. The contour plots of the saturation S on the fine 1024×1024 mesh at time $t = 0.25$ and $t = 0.5$ computed by using the "exact" velocity field are displayed in Figure 4.10. In Figure 4.9, we show the contour plots of the saturation obtained using the recovered velocity field from the coarse grid pressure calculation $N = 64$. We can see that the the contour plots in Figure 4.9 approximate the "exact" ones in Figure 4.10 in certain accuracy but the sharp oil/water interfaces in Figure 4.10 are smeared out. This is due to the parabolic nature of the upscaled equation (4.33). We have also performed many other numerical experiments to test the robustness of this combined coarse grid model. We found that for permeability fields with strong layered structure, the above coarse grid model is very robust. The agreement with the fine grid calculations is very good. We are currently working toward some qualitative and quantitative understanding of this upscaling model.

Finally, we remark that the upscaling equation (4.33) uses small scale information $\mathbf{v}'$ of the velocity field to define the diffusion coefficients. This information can be constructed *locally* through the mixed multiscale basis functions. This is an important property of our multiscale finite element method. It is clear that solving directly the homogenized pressure equation

$$\mathrm{div}(\mathcal{K}^*(x)\nabla p^*) = 0$$

will not provide such small scale information. On the other hand, whenever one can afford to resolve all the small scale feature by a fine grid, one can use fast linear solvers, such as multigrid methods, to solve the pressure equation (4.31) on the fine mesh. ¿From the fine grid computation, one can easily construct the average velocity $\mathbf{v}_0$ and its deviation $\mathbf{v}'$. However, when multiple simulations must be carried out due to the change of boundary conditions, the pressure equation (4.31) will then have to be solved again on the fine mesh. The multiscale finite element method only solves the pressure equation once on a coarse mesh, and the fine grid velocity can be constructed locally through the finite element bases. This is the main advantage of our mixed multiscale finite element method. This process becomes more difficult for the

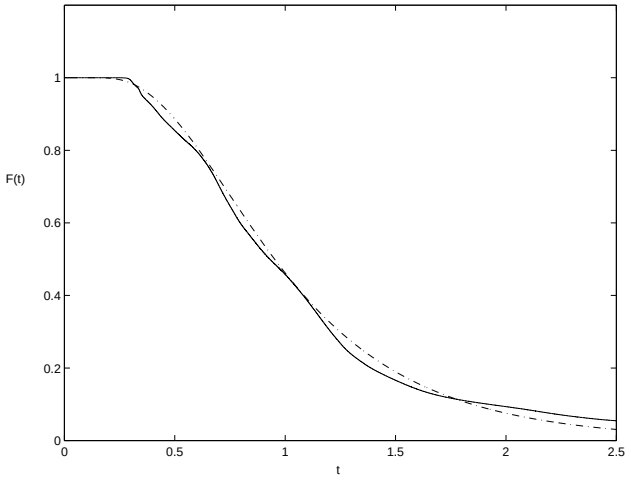

Fig. 4.8. The accuracy of the coarse grid algorithm. Solid line is the "exact" fractional flow curve using mixed finite element method solving the pressure equation. The slash-dotted line is the fractional flow curve using above coarse grid algorithm.

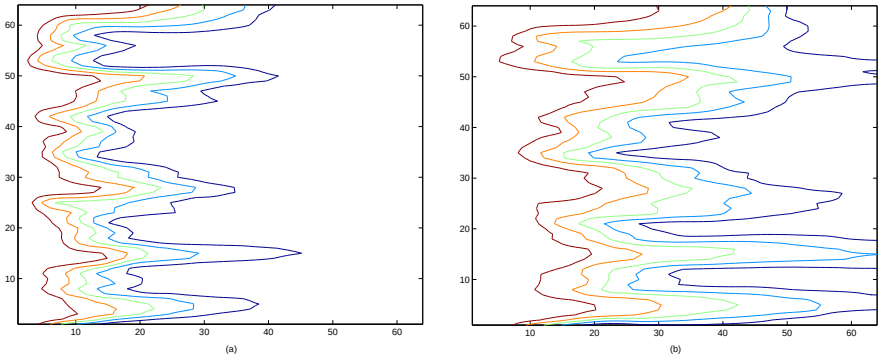

Fig. 4.9. The contour plots of the saturation S computed using the upscaled model on a 64×64 mesh at time $t = 0.25$ (left) and $t = 0.5$ (right).

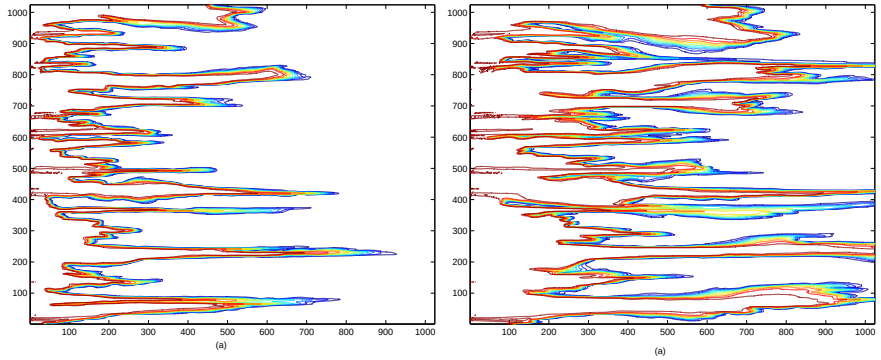

Fig. 4.10. The contour plots of the saturation S computed on the fine 1024×1024 mesh using the "exact" velocity field at time $t = 0.25$ (left) and $t = 0.5$ (right).

nonlinear two-phase flow due to the dynamic coupling between the pressure and the saturation. We are now investigating the possibility of upscaling the two-phase flow by using multiscale finite element base functions that are constructed from the one-phase flow (time independent). In this case, we need to provide corrections to the pressure equation to account for the scale interaction near the oil/water interface.

It should be noted that some adaptive scale-up strategies have also been developed [24,66]. The idea is to refine the mesh around the fast-flow channels in order to capture their effect directly. The approach seems to work well when the channels are isolated. For MsFEM, it is also possible to adjust the coarse mesh adaptively based on the recovered velocity. In particular, one does not need to use the fine recovered velocity in the regions with no fast-flow channels; in those regions, the coarse mesh and cell-averaged velocity are sufficient. On the other hand, one can simply keep the fine mesh when the channels are too many. How to develop a consistent upscaling equation for the saturation equation is still open when the capillary pressure effect is neglected, which is the common practice in oil reservoir simulations. One approach is to combine grid adaptivity with multiscale modeling. We use a dynamic adaptive coarse grid [15] to capture the isolated small scale features, such as the flow channels and use the multiscale finite element method to capture the small scale feature within each adaptive coarse grid block. By doing this, we take into account the local flow orientation and anisotropy in upscaling the saturation equation. We are also investigating the possibility to develop a consistent upscaling model for the saturation equation by combining multiscale finite element methods and systematic multiscale modeling for the saturation equation.

5 Wavelet-based Homogenization (WBH)

The material in this section is based on the work of Borobantu and Engquist in [21] and the lectures given by Engquist in the Morningside Summer School on Multiscale Analysis and Computation in Beijing in August, 1997. Please also consult with the related work of Brewster and Beylkin in [10] for integral differential equations.

We begin by considering a simple, elliptic equation. Let u^ε be the variable in question defined on the domain Ω. We study the boundary-value problem

$$\begin{aligned} -\nabla \cdot a_\varepsilon(x)\nabla u^\varepsilon(x) &= f(x), && x \in \Omega \\ \text{B.C.} \quad \text{for} \quad u^\varepsilon, & && x \in \partial\Omega \end{aligned}$$

where $a_\varepsilon(x) = a(x_1/\varepsilon, x_2/\varepsilon)$, and $a(y_1, y_2)$ is 1-periodic in y_1, y_2.

As we discussed in Section 1, one can derive a homogenized equation using the multiscale expansion technique. Roughly speaking, homogenization is a mathematical method that allows us to upscale differential equations.

This method not only offers formulae for upscaling but also provides tools for producing rigorous mathematical convergence proofs.

The homogenization problem can be stated in various formulations. A classical formulation, see e.g. Bensoussan *et al.* [7], is the following: Consider a family of operators L_ε indexed by the small parameter ε. For any function f, let u^ε solve the problem

$$L_\varepsilon u^\varepsilon = f. \tag{5.1}$$

Assume that $u^\varepsilon \rightharpoonup \overline{u}$, as $\varepsilon \to 0$. The homogenization problem is to find an operator $\overline{L}$ such that

$$\overline{L}\overline{u} = f. \tag{5.2}$$

The homogenized equation does not contain the ε-scale and it can be solved on a coarse grid. In some cases $\overline{L}$ has a closed form analytic expression [7]. If that is not the case, but the general structure of $\overline{L}$ is known, a numerical approximation is still possible [23].

The following diagram describes two common approaches to a coarse grid discretization of the multiscale problem, $L_\varepsilon u^\varepsilon = f$.

$$\begin{array}{ccc} & L_\varepsilon u^\varepsilon = f & \\ \clubsuit \swarrow & & \diamondsuit \searrow \\ \overline{L}\overline{u} = \overline{f} & & L_{\varepsilon,h} u_h^\varepsilon = f_{\varepsilon,h} \\ \diamondsuit \searrow & & \heartsuit \swarrow \\ & L_H \overline{u}_H = f_H & \end{array}$$

where $\clubsuit$ stands for homogenization, $\diamondsuit$ stands for discretization (for example, finite element methods, etc.); $\heartsuit$ stands for numerical upscaling (sampling, wavelet based homogenization, etc, $\cdots$).

Homogenization is a model reduction. In the traditional analysis, reduction is often based on physical conservation (derivation from first principles). For example, in fluid mechanics, fundamental particle interaction $\longrightarrow$ kinetic model (Boltzmann equation) $\longrightarrow$ Navier-Stokes equation (Hydrodynamic approximation). ¿From the Navier-Stokes equations, one can further derive turbulence models, the Euler equations, and potential equations.

In contrast, one can derive reduced models using numerical techniques starting from the discretized problem. Wavelet-Based-Homogenization, Sampling, Multigrid are some of examples in discrete model reduction.

The framework for the wavelet based method is discrete. The operator L is a finite-dimensional approximation of the original differential operator. It can be written in the form

$$L = P(\Delta, A, h),$$

where Δ is a collection of difference operators, A is a discretized variable coefficient, typically diagonal matrices, and h represents the grid size. We

seek a homogenized, discrete, operator $\overline{L}$ that can be written in a form similar to L,

$$\overline{L} = P(\Delta, H, \overline{h}),$$

but with $\overline{h} \gg h$ and the structure of H close to the structure of A, typically diagonal dominant and sparse. We interpret H as the subgrid model of A. If A corresponds to a material coefficient, H can be seen as the effective material coefficient.

5.1 Wavelets

For multiscale problems, the first important concept is to define scales, what we mean by large scales and what we mean by small scales. The classical way is to use the Fourier series or Fourier transform as a systematic way of defining scales

$$u(x) = \sum_j a_j \mathrm{e}^{ijx}.$$

Note that the Fourier bases, e^{ijx}, are global. This introduces some difficulty in characterizing local small scales. Wavelets introduce orthonormal bases which are localized in space or time. This offers some very attractive features which have important applications in signal and image processing.

When solving differential equations, we often express the solution as a linear combination of some pre-determined basis

$$u(x) = \sum_k a_k \varphi_k(x).$$

For example, $\varphi_k(x) = \varphi(x - x_k)$ in finite element method. It is a translation operator, whereas in the spectral method, we have

$$u(x) = \sum_j a_j \varphi_j(x)$$

where $\varphi_j(x) = \mathrm{e}^{ijx}$, which is a dilation operator.

Using both translation and dilation, the functions are expressed by wavelets

$$u(x) = \sum_{j,k} a_{jk} \psi_{jk},$$

where $\psi_{jk}(x) = 2^{j/2}\psi(2^j x - k)$.

Example: A well-known example of wavelet basis is the Haar wavelet

$$H_{jk}(x) = 2^{j/2} H(2^j x - k)\ ,$$

where

$$H(x) = \begin{cases} 1, & \text{if} \quad 0 \le x \le 1/2, \\ -1, & \text{if} \quad 1/2 < x \le 1, \\ 0, & \text{otherwise.} \end{cases}$$

Note that the mean of $H(x)$ is zero, $\int_{\mathbf{R}} H(x)\,dx = 0$.

We now describe a subgrid scale model within the framework of wavelet based multiresolution analysis (MRA). A wavelet representation lends itself naturally to analyzing the fine and coarse scales as well as the localization properties of a function. For a detailed description of MRA we refer the reader to the book by Daubechies [20]. For fast wavelet algorithms, we refer the reader to [8].

We consider a ladder of spaces $V_J \subset V_{J+1}$ which are spanned by the dilates and integer translates of one shape-function $\phi \in V_0$:

$$V_J = \text{span}\{\phi_{J,k}(x) = 2^{J/2}\phi(2^J x - k)\}.$$

The functions $\phi_{J,k}$ form an L^2-orthonormal basis. The orthogonal complement of V_J in V_{J+1} is denoted by W_J and it is generated by another orthonormal basis $\psi_{J,k}(x) = 2^{J/2}\psi(2^J x - k)$, where ψ is called the mother wavelet. The transformation

$$\mathcal{W}_J : V_{J+1} \to W_J \oplus V_J$$

that maps the basis $\{\phi_{J+1,k}\}$ into $\{\psi_{J+1,k}, \phi_{J+1,k}\}$ is an orthogonal operator and we denote its inverse by $\mathcal{W}^T$. The product $\mathcal{W}_J\mathcal{W}_{J-1}\cdots\mathcal{W}_0$ maps V_{J+1} into $V_0 \oplus \sum_{0\le j\le J} W_j$ is called the wavelet transform and it can be optimally implemented using the fast wavelet transform. We denote by P_j and Q_j the L^2-projections onto V_j and W_j.

5.2 Introduction to Wavelet-Based Homogenization (WBH)

Given the fully discrete solution operator on a fine grid, the underlying idea of wavelet-based homogenization is to find an operator of lower dimension that extracts only the coarse scales of the solution. The numerical homogenization can be described as the following process:

$$L_\varepsilon u^\varepsilon = f \to L_{\varepsilon,h} u_h^\varepsilon = f_h \to L_H u_H = f_H\ .$$

The wavelet-based homogenization is described as follows:

1. Start from the discretized differential equation;
2. Represent the operator and the solution in wavelet basis;
3. Project to the (coarse and small) scale spaces;
4. Compress the operator after transformation.

We illustrate the main idea by considering the following differential equation

$$-\frac{d}{dx}\left(a_\varepsilon(x)\frac{d}{dx}u_\varepsilon(x)\right) = f(x)\ ,$$

where $a_\varepsilon(x) = a(\frac{x}{\varepsilon})$. Discretizing the equation on a fine mesh h gives $L_{\varepsilon,h}u_h^\varepsilon = f_h$. We denote by L_{J+1} the discrete operator on the finest level. We have

$$L_{J+1}U = F.$$

If we use a centered difference approximation to the derivative operator, we have

$$L_{J+1} = h^{-2}\Delta_+ \operatorname{diag}(a_\varepsilon)\Delta_-$$

where $\Delta_+ u_k = u_{k+1} - u_k$ is the forward difference operator, and $\Delta_- u_k = u_k - u_{k-1}$ is the backward difference operator, which satisfies $\Delta_- = \Delta_+^T$. Here $\operatorname{diag}(a)$ is the diagonal matrix with diagonal entry given by $a(x_i)$.

We first identify the pointwise value with coefficient in the space of V_{J+1}. A linear operator L_{J+1} acting on the space V_{J+1} can be decomposed into four operators $L_{J+1} = A_J + B_J + C_J + L_J$ acting on the subspace W_J and V_J, where

$$\begin{aligned} A_J &= Q_J L_{J+1} Q_J : \qquad W_J \to W_J \\ B_J &= Q_J L_{J+1} P_J : \qquad V_J \to W_J \\ C_J &= P_J L_{J+1} Q_J : \qquad W_J \to V_J \\ L_J &= P_J L_{J+1} P_J : \qquad V_J \to V_J. \end{aligned}$$

Applying the transformation $\mathcal{W}_J$ on L_{J+1}, we have for $U \in V_{J+1}$,

$$\mathcal{W}_J L_{J+1} \mathcal{W}_J^T (\mathcal{W}_J U) = \begin{bmatrix} A_J & B_J \\ C_J & L_J \end{bmatrix} \begin{bmatrix} Q_J U \\ P_J U \end{bmatrix}, \tag{5.3}$$

or simply

$$\mathcal{W}_J L_{J+1} \mathcal{W}_J^T = \begin{bmatrix} A_J & B_J \\ C_J & L_J \end{bmatrix}. \tag{5.4}$$

Let us now consider

$$L_{J+1} U = F, \qquad U, F \in V_{J+1}.$$

This equation may originate from a finite difference, finite element or finite volume discretization of a given equation. We identify U as a piecewise constant approximation of $u(x)$, the solution to the continuous problem. After the same wavelet transformation as in (5.3), we have

$$\begin{pmatrix} A_J & B_J \\ C_J & L_J \end{pmatrix} \begin{pmatrix} U_h \\ U_l \end{pmatrix} = \begin{pmatrix} F_h \\ F_l \end{pmatrix}, \qquad U_h, F_h \in W_J, \qquad U_l, F_l \in V_J,$$

where $U_h = Q_J U$ and $U_l = P_J U$ and similarly for F. For the Haar basis this means that U_h is essentially the high frequency part and U_l is the low frequency part of U. The first equation is

$$U_h = A_J^{-1}(F_h - B_J U_l).$$

Eliminating U_h yields the equation for U_l

$$(L_J - C_J A_J^{-1} B_J) U_l = F_l - C_J A_J^{-1} F_h.$$

Our new "coarse grid operator" is the Schur complement

$$\overline{L}_J = L_J - C_J A_J^{-1} B_J, \tag{5.5}$$

which includes subgrid phenomena via $C_J A_J^{-1} B_J$. We also get the homogenized right hand side,

$$\bar{F}_J = F_l - C_J A_J^{-1} F_h.$$

Note that this is in fact a block Gaussian elimination procedure. Further note that the above procedure can be repeated on $\bar{L}_J$ to get $\bar{L}_{J-1}$ and so on. To make this efficient in real applications it is necessary to be able to approximate $\bar{L}_J$ with a sparse matrix. This sparse matrix can be seen as a discretization of a local differential operator.

There is a striking relation between the Schur complement $\bar{L}_j$ in (5.5) and the analytically homogenized operator (2.24)-(2.25) in Section 1, repeated here for convenience,

$$\overline{L\overline{u}} = -\frac{\partial}{\partial x_i}\left(a_{ij}^* \frac{\partial}{\partial x_i \partial x_j}\right)\overline{u} + \left(\frac{1}{|Y|}\int_Y a_0(y)\,dy\right)\overline{u}\,, \tag{5.6}$$

where

$$a_{ij}^* = \frac{1}{|Y|}\int_Y (a_{ij} - a_{ik}\frac{\partial \chi^j(y)}{\partial y_k})\,dy. \tag{5.7}$$

The first terms in (5.5) and (5.7) both represent averaged operators, L_J in a discrete sense and

$$-\left(\sum_{ij}\frac{1}{|Y|}\int_Y a_{ij}\,dy\right)\frac{\partial^2}{\partial x_i \partial x_j} \tag{5.8}$$

in an integral sense. In both formulations a correction term is subtracted from the average. Furthermore, in the correction term χ is the solution of an elliptic equation and A_j^{-1} is a discrete positive definite operator.

The above discussion on one-dimensional problems can be generalized to two-dimensional problems. In the two-dimensional case, the maps

$$\mathcal{W}_J : \mathbf{V}_{J+1} \to \mathbf{W}_J \oplus \mathbf{V}_J \tag{5.9}$$

can be written as a tensor product of one-dimensional transforms,

$$\mathcal{W}_J^{2d} = \mathcal{W}_J \otimes \mathcal{W}_J.$$

A linear operator L_{J+1} that acts on the space $\mathbf{V}_{J+1}$ can be decomposed in a way similar to the one-dimensional case. To get a convenient matrix representation we use $\widetilde{\mathcal{W}} = P\mathcal{W}$, instead of $\mathcal{W}$. The matrix P is a suitable permutation. The equation

$$L_{J+1}U = F, \qquad U, F \in \mathbf{V}_{J+1},$$

can then be transformed to

$$\begin{pmatrix} A_J & B_J \\ C_J & L_J \end{pmatrix} \begin{pmatrix} U_h \\ U_l \end{pmatrix} = \begin{pmatrix} F_h \\ F_l \end{pmatrix}, \qquad U_h, F_h \in \mathbf{W}_J, \qquad U_l, F_l \in \mathbf{V}_J,$$

and the coarse grid operator is again the Schur complement,

$$\bar{L}_J = L_J - C_J A_J^{-1} B_J.$$

Note that the high frequency part of U can be decomposed as

$$U_h = \begin{bmatrix} U_{hh} \\ U_{lh} \\ U_{hl} \end{bmatrix}, \quad U_{hh} \in W_J \otimes W_J, \quad U_{lh} \in V_J \otimes W_J, \quad U_{hl} \in W_J \otimes V_J.$$

Similar tensor product extensions can be made also for high dimensions.

Unlike the homogenized operator $\overline{L}$ in the continuous case, the discrete "homogenized" operator, $\overline{L}_J$, is a nonlocal dense operator, since A_J^1 is dense. For elliptic operators, A_J is diagonally dominant. The compression property of wavelets makes it possible to approximate A_J^{-1} by a sparse matrix. This is an essential property that makes this numerical homogenization procedure efficient. This discrete homogenization procedure can be applied recursively to yield a coarse grid operator, $\overline{L}_I$ at a desired coarse level.

It turns out that it is more effective to write $\overline{L}_J$ in a conservative form. In the 1-D case, this amounts to expressing $\overline{L}_J$ as follows:

$$\overline{L}_J = \frac{1}{(2h)^2} \Delta_+ H \Delta_- \ ,$$

where $\Delta_+ u_j = u_{j+1} - u_j$ and $\Delta_- u_j = u_j - u_{j-1}$, and H is a strongly diagonal dominant matrix.

We look at the extreme case when $a(x) = \overline{a} + \tilde{a}(x)$ is the sum of a constant and the highest frequency represented on the grid, i.e., $a(x_m) = \overline{a} + |\tilde{a}|(-1)^m$. We have that $\overline{a}$ and $\tilde{a}$ are represented as constant vectors in the bases of V_J and W_J. The fact that $a(x) > 0$ implies $|\tilde{a}| < |\overline{a}|$.

The following theorem shows that the wavelet homogenized operator $\frac{1}{h^2}\Delta_+ H \Delta_-$ equals the discrete form $\alpha \frac{1}{h^2}\Delta_+ \Delta_-$ of the classically homogenized equation, apart from a second order error term of order h^2.

Theorem 5.1. *[21] Let $a(x) = \overline{a} + \tilde{a} \in V_{J+1}$ be such that $\overline{a} \in V_0$ is a constant and the oscillatory part $\tilde{a} \in W_J$ has constant amplitude and satisfies the condition $|\tilde{a}| < \overline{a}$. Let $L_{J+1} = \frac{1}{(h/2)^2}\Delta_+ a \Delta_-$ and α be the harmonic average*

$$\alpha = \left(\frac{1}{2h} \int_0^{2h} \frac{1}{a(x)}\, dx \right)^{-1}.$$

Then there exists a function $v(x)$ with a continuous and bounded fourth derivative such that

$$\|\overline{L}_J v - \alpha \frac{1}{h^2} \Delta_+ \Delta_- v\|_\infty \le C h^2 \|v^{(4)}\|_{L^\infty}.$$

In practice, we want to approximate the homogenized operator $\bar{L}_J$ by a sparse approximation. Due to the decay in the off-diagonal entries, we can approximate $\overline{L}_J$ by a band-diagonal matrix $\overline{L}_{J,\nu}$ where ν is the band-width. Let us consider the operator band defined by

$$\text{band}(M,\nu)_{i,j} = \begin{cases} M_{i,j}, & \text{if} \quad 2|i-j| \le \nu - 1 \ , \\ 0, & \text{otherwise.} \end{cases}$$

We have in fact two obvious strategies available for producing $\overline{L}_{J,\nu}$: We can set directly $\overline{L}_{J,\nu} = \text{band}(\overline{L}_J, \nu)$ or use the homogenized coefficient form and build $\overline{L}_{J,\nu} = \frac{1}{h^2}\Delta_+ \,\text{band}(H, \nu-2)\Delta_-$. Both approaches produce small perturbations of $\overline{L}_J$. However, important properties, such as divergence form, are lost in the first approach and numerical experiments show that ν needs to be rather large to compensate for this. The second approach produces $\overline{L}_{J,\nu}$ in divergence form. Moreover, the approximation error can be estimated, as in the following result:

Theorem 5.2. *[21] If the conditions of Theorem 5.1 are valid, then we have*

$$\|H - \text{band}(H,\nu)\| \le C\rho^{\nu}, \qquad \rho = \frac{2(\overline{a} + |\tilde{a}|)}{6\overline{a} - 2|\tilde{a}|} < 1.$$

If v is the discretization of a smooth function $v(x)$, then

$$\|(\overline{L}_J - \overline{L}_{J,\nu} v\|_{\infty} \le C\rho^{\nu} \|v''\|_{L^{\infty}}.$$

In the current approach, the numerical homogenization starts from a fine grid, and the operator is global. It would be nice to derive a local procedure to implement this idea. Further, it remains to study the decay rate of H_{ij} away from diagonal for more general a_ε and more general wavelet bases. One also needs to find an efficient way of computing $\overline{L}_J$ (incomplete LU decomposition of A). Compare with multigrid method and preconditional conjugate gradient method.

6 Variational Multiscale Method

In this section, we will briefly review the main idea of the variational multiscale method introduced by Hughes and Brezzi et al in [42,12,43].

Consider an abstract variational problem: Find $u \in V$ such that

$$a(u,v) = F(v), \quad \text{for all} \quad v \in V, \tag{6.1}$$

where V is a Hilbert space, $a(\cdot,\cdot)$ is a continuous and coercive bilinear form on V, and $F(\cdot)$ is a continuous linear form on V.

A typical example is the elliptic problem (2.1) in Section 1. In this case, we have

$$a(u,v) = \int_\Omega a_\varepsilon(x)\nabla u \cdot \nabla v \, dx + \int_\Omega a_0(x) uv \, dx, \quad F(v) = \int_\Omega f(x)v(x)\, dx.$$

If we choose $V = H_0^1(\Omega)$, then the above variational problem is equivalent to the elliptic equation (2.1) in Section 1.

The classical Galerkin approximation of (6.1) consists of taking a finite dimensional subspace V_h of V and solving (6.1) in V_h, i.e. find $u_h \in V_h$ such that

$$a(u_h, v_h) = F(v_h), \qquad \text{for all} \quad v_h \in V_h. \tag{6.2}$$

Let $\mathcal{T}_h = \{K\}$ be a triangulation of Ω, $h_K = \text{diam}\{K\}$, $h = \max_K h_K$. Typically V_h consists of continuous functions which are polynomials of some degrees on a triangular element K.

To be specific, we consider piecewise linear elements. We set

$$V_R^h = \left\{ v_R \in H_0^1(\Omega),\ v_R|_K \text{ is linear in each } \quad K \right\}.$$

The variational problem (6.2) can be written as follows: Find $u_R \in V_R^h$ such that

$$a(u_R, v_R) = F(v_R), \quad \text{for all } \ v_R \in V_R^h. \tag{6.3}$$

Here, u_R represents the resolvable part of the solution.

Let V_U^b be a closed subspace of $H_0^1(\Omega)$ such that $V_R^h \cap V_U = \{0\}$.

Further, we define

$$V_h = V_R^h \oplus V_U^b. \tag{6.4}$$

We can consider V_h as the augmented space of V_R^h.

Using the decomposition $V_h = V_R^h \oplus V_U^b$, we can express any $v_h \in V_h$ as the sum of a *resolvable part*, $v_R \in V_R^h$, and an *unresolvable part*, $v_U \in V_U^b$ in a unique way:

$$v_h = v_R + v_U \in V_R^h \oplus V_U^b \,.$$

In turn, the variational problem (6.2) can be expressed as follows: Find $u_h = u_R + u_U \in V_R^h \oplus V_U^b$ such that

$$a(u_R + u_U, v_R) = F(v_R) \,, \quad \text{for all } v_R \in V_R^h \tag{6.5}$$

$$a(u_R + u_U, v_U) = F(v_U) \,, \quad \text{for all } v_U \in V_U^b \,. \tag{6.6}$$

Using the bilinearity of $a(\cdot,\cdot)$, equation (6.6) can be written as

$$a(u_U, v_U) = -\left(a(u_R, \cdot) - F(\cdot)\right)(v_U) \,, \quad \text{for all } v_U \in V_U^b \,. \tag{6.7}$$

Problem (6.7) can be "solved" for any $u_R \in V_R^h$, and the solution can be formally written as

$$u_U = M(\mathcal{L}u_R - f), \tag{6.8}$$

where the operator $\mathcal{L}$ is defined as in (2.1), M is a linear solution operator from $H^{-1}(\Omega)$ to $H_0^1(\Omega)$. One can also view M as the fine grid solution operator or the discrete Green function operator acting on the unresolvable scales.

Substituting the unresolvable part of the solution, $u_U = M(\mathcal{L}u_R - f)$ into equation (6.5) for the resolvable part, we get

$$a(u_R, v_R) + \underbrace{a(M(\mathcal{L}u_R - f), v_R)}_{\text{effect of the space } V_U^h} = (f, v_R), \quad \text{for all } v_R \in V_R^h \ . \quad (6.9)$$

The term $a(M(\mathcal{L}u_R - f), v_R)$ represents the contribution of small scales to large scales, which resembles the so-called "Reynolds stress" term in turbulence modeling.

Solving u_U exactly would be as expensive as solving the fine grid solution globally. In order to *localize* the computation of u_U, the authors in [42,12,43] made the following assumption:

$$V_U = \oplus_K H_0^1(K).$$

In other words, they take into account only those unresolvable scales that *vanish* on the boundaries of the coarse grid elements. In some sense, the multiscale finite element method with linear boundary conditions for the multiscale base functions is very similar to the variational multiscale method described here. As we see from the analysis of the multiscale finite element method in [38,28], by forcing the unresolvable bases to vanish on the boundaries of the coarse grid elements, the resulting multiscale method may introduce $O(1)$ errors when the physical small scale is of the same order as the coarse grid size.

Using the assumption $u_U|_K = 0$, u_U can be uniquely decomposed among each element, K:

$$u_U = \sum_K u_{U,K} \ , \quad u_{U,K} \in H_0^1(K).$$

The variational problem now becomes: Find $u_h = u_R + u_U = u_R + \sum_K u_{U,K} \in V_R^h \oplus V_U$ such that

$$a(u_R + u_U, v_R) = F(v_R) \ , \quad \text{for all } v_R \in V_R^h \quad (6.10)$$

$$a(u_R + u_{U,K}, v_{U,K})_K = F(v_{U,K})_K \ , \quad \text{for all } v_{U,K} \in H_0^1(K), \ \forall \ K, \quad (6.11)$$

where $a(u_R + u_{U,K}, v_{U,K})_K$ and $F(v_{U,K})_K$ are the restrictions of $a(u_R + u_{U,K}, v_{U,K})$ and $F(v_{U,K})$ on K respectively. Again, we obtain an equation for u_R as

$$a(u_R, v_R) + \sum_K a(u_{U,K}, v_R)_K = F(v_R) \ , \quad \text{for all } v_R \in V_R^h, \ \forall \ K. \quad (6.12)$$

The "local equation" for $u_{U,R}$ becomes

$$a(u_{U,K}, v_{U,K})_K = -[a(u_R, \cdot) - F(\cdot)]|_K(v_{U,K}), \quad \text{for } v_{U,K} \in H_0^1(K).$$

Equivalently, we have

$$\mathcal{L}u_{U,K} = -\underbrace{(\mathcal{L}u_R - f)}_{\text{residual}}, \quad \text{in } K, \quad \text{with } u_{U,K} = 0 \quad \text{on } \partial K.$$

In some sense, the un-resolvable bases, $\{u_{U,K}\}$, play the same role as the residual-free bubbles in the residual-free bubbles finite element method introduced by Brezzi and Russo in [11]. Let g_y^K be the Green function on K for operator $\mathcal{L}$, i.e.

$$\mathcal{L}g_y^K(x) = \delta_y(x), \quad \text{for} \quad x \in K, \quad g_y^K = 0 \quad \text{on } \partial K.$$

Then we can write $u_{U,K}$ formally as

$$u_{U,K} = -\int_K g_y^K(x)(\mathcal{L}u_R - f)(y)\, dy.$$

This is not very practical since it is expensive to construct the Green function numerically in each element. In [42,12,43], various approximations to the discrete Green function are proposed to study the stabilizing residual-free bubble method. Hughes and his co-workers have also applied this idea to a number of interesting applications [43]. Recently, Todd Arbogast has introduced a subgrid upscaling method for two-phase flow in porous media using a similar approach [2]. From the analytical view point, the variational multiscale method or the residual-free bubble approach provides a good framework to design multiscale methods in a systematic way.

References

1. J. Aarnes and T. Y. Hou *An Efficient Domain Decomposition Preconditioner for Multiscale Elliptic Problems with High Aspect Ratios,* Acta Mathematicae Applicatae Sinica, **18** (2002), 63-76.
2. T. Arbogast, *Numerical Subgrid Upscaling of Two-Phase Flow in Porous Media,* in Numerical treatment of multiphase flows in porous media, Z. Chen et al., eds., Lecture Notes in Physics 552, Springer, Berlin, 2000, pp. 35-49.
3. M. Avellaneda, T. Y. Hou and G. Papanicolaou, *Finite Difference Approximations for Partial Differential Equations with Rapidly Oscillating Coefficients,* Mathematical Modelling and Numerical Analysis, **25** (1991), 693-710.
4. I. Babuska, G. Caloz, and E. Osborn, *Special Finite Element Methods for a Class of Second Order Elliptic Problems with Rough Coefficients,* SIAM J. Numer. Anal., **31** (1994), 945-981.
5. I. Babuska and E. Osborn, *Generalized Finite Element Methods: Their Performance and Their Relation to Mixed Methods,* SIAM J. Numer. Anal., **20** (1983), 510-536.
6. I. Babuska and W. G. Szymczak, *An Error Analysis for the Finite Element Method Applied to Convection-Diffusion Problems,* Comput. Methods Appl. Math. Engrg, **31** (1982), 19-42.

7. A. Bensoussan, J. L. Lions, and G. Papanicolaou, *Asymptotic Analysis for Periodic Structures,* Volume 5 of Studies in Mathematics and Its Applications, North-Holland Publ., 1978.
8. G. Beylkin, R. Coifman, and V. Rokhlin *Fast Wavelet Transforms and Numerical Algorithm I* Comm. Pure Appl. Math., **44** (1991), 141-183.
9. A. Bourgeat, *Homogenized Behavior of Two-Phase Flows in Naturally Fractured Reservoirs with Uniform Fractures Distribution,* Comp. Meth. Appl. Mech. Engrg, **47** (1984), 205-216.
10. M. Brewster and G. Beylkin, *A Multiresolution Strategy for Numerical Homogenization,* ACHA, **2**(1995), 327-349.
11. F. Brezzi and A. Russo, *Choosing Bubbles for Advection-Diffusion Problems,* Math. Models Methods Appl. Sci, **4** (1994), 571-587.
12. F. Brezzi, L. P. Franca, T. J. R. Hughes and A. Russo, $b = \int g$, Comput. Methods in Appl. Mech. and Engrg., **145** (1997), 329-339.
13. J. E. Broadwell, *Shock Structure in a Simple Discrete Velocity Gas,* Phys. Fluids, **7** (1964), 1243-1247.
14. T. Carleman, *Problèms Mathématiques dans la Théorie Cinétique de Gaz,* Publ. Sc. Inst. Mittag-Leffler, Uppsala, 1957.
15. H. Ceniceros and T. Y. Hou, *An Efficient Dynamically Adaptive Mesh for Potentially Singular Solutions.* J. Comput. Phys., **172** (2001), 609-639.
16. Z. Chen and T. Y. Hou, *A Mixed Finite Element Method for Elliptic Problems with Rapidly Oscillating Coefficients,* to appear in Math. Comput..
17. A. J. Chorin, *Vortex Models and Boundary Layer Instabilities,* SIAM J. Sci. Statist. Comput., **1** (1980), 1-21.
18. M. E. Cruz and A. Petera, *A Parallel Monte-Carlo Finite Element Procedure for the Analysis of Multicomponent Random Media,* Int. J. Numer. Methods Engrg, **38** (1995), 1087-1121.
19. J. E. Dendy, J. M. Hyman, and J. D. Moulton, *The Black Box Multigrid Numerical Homogenization Algorithm,* J. Comput. Phys., **142** (1998), 80-108.
20. I. Daubechies *Ten Lectures on Wavelets.* SIAM Publications, 1991.
21. M. Dorobantu and B. Engquist, *Wavelet-based Numerical Homogenization,* SIAM J.Numer. Anal., **35** (1998), 540-559.
22. J. Douglas, Jr. and T.F. Russell, *Numerical Methods for Convection-dominated Diffusion Problem Based on Combining the Method of Characteristics with Finite Element or Finite Difference Procedures,* SIAM J. Numer. Anal. **19** (1982), 871–885.
23. L. J. Durlofsky, *Numerical Calculation of Equivalent Grid Block Permeability Tensors for Heterogeneous Porous Media,* Water Resour. Res., **27** (1991), 699-708.
24. L.J. Durlofsky, R.C. Jones, and W.J. Milliken, *A Nonuniform Coarsening Approach for the Scale-up of Displacement Processes in Heterogeneous Porous Media,* Adv. Water Resources, **20** (1997), 335–347.
25. B. B. Dykaar and P. K. Kitanidis, *Determination of the Effective Hydraulic Conductivity for Heterogeneous Porous Media Using a Numerical Spectral Approach: 1. Method,* Water Resour. Res., **28** (1992), 1155-1166.
26. W. E and T. Y. Hou, *Homogenization and Convergence of the Vortex Method for 2-D Euler Equations with Oscillatory Vorticity Fields,* Comm. Pure and Appl. Math., **43** (1990), 821-855.
27. Y. R. Efendiev, *Multiscale Finite Element Method (MsFEM) and its Applications*, Ph. D. Thesis, Applied Mathematics, Caltech, 1999.

28. Y. R. Efendiev, T. Y. Hou, and X. H. Wu, *Convergence of A Nonconforming Multiscale Finite Element Method,* SIAM J. Numer. Anal., **37** (2000), 888-910.
29. Y. R. Efendiev, L. J. Durlofsky, S. H. Lee, *Modeling of Subgrid Effects in Coarse-scale Simulations of Transport in Heterogeneous Porous Media,* WATER RESOUR RES, **36** (2000), 2031-2041.
30. B. Engquist, *Computation of Oscillatory Solutions to Partial Differential Equations,* in Proc. Conference on Hyperbolic Partial Differential Equations, Carasso, Raviart, and Serre, eds, Lecture Notes in Mathematics No. 1270, Springer-Verlag, 10-22, 1987.
31. B. Engquist and T. Y. Hou, *Particle Method Approximation of Oscillatory Solutions to Hyperbolic Differential Equations,* SIAM J. Numer. Anal., **26** (1989), 289-319.
32. B. Engquist and T. Y. Hou, *Computation of Oscillatory Solutions to Hyperbolic Equations Using Particle Methods,* Lecture Notes in Mathematics No. 1360, Anderson and Greengard eds., Springer-Verlag, 68-82, 1988.
33. B. Engquist and H. O. Kreiss, *Difference and Finite Element Methods for Hyperbolic Differential Equations,* Comput. Methods Appl. Mech. Engrg., **17/18** (1979), 581-596.
34. B. Engquist and E.D. Luo, *Convergence of a Multigrid Method for Elliptic Equations with Highly Oscillatory Coefficients,* SIAM J. Numer. Anal., **34** (1997), 2254-2273.
35. D. Gilbarg and N. S. Trudinger, Elliptic Partial Differential Equations of Second Order. Springer, Berlin, New York, 2001.
36. J. Glimm, H. Kim, D. Sharp, and T. Wallstrom *A Stochastic Analysis of the Scale Up Problem for Flow in Porous Media,* Comput. Appl. Math., **17** (1998), 67-79.
37. T. Y. Hou, *Homogenization for Semilinear Hyperbolic Systems with Oscillatory Data,* Comm. Pure and Appl. Math., **41** (1988), 471-495.
38. T. Y. Hou and X. H. Wu, *A Multiscale Finite Element Method for Elliptic Problems in Composite Materials and Porous Media,* J. Comput. Phys., **134** (1997), 169-189.
39. T. Y. Hou and X. H. Wu, *A Multiscale Finite Element Method for PDEs with Oscillatory Coefficients,* Proceedings of 13th GAMM-Seminar Kiel on Numerical Treatment of Multi-Scale Problems, Jan 24-26, 1997, Notes on Numerical Fluid Mechanics, Vol. 70, ed. by W. Hackbusch and G. Wittum, Vieweg-Verlag, 58-69, 1999.
40. T. Y. Hou, X. H. Wu, and Z. Cai, *Convergence of a Multiscale Finite Element Method for Elliptic Problems With Rapidly Oscillating Coefficients,* Math. Comput., **68** (1999), 913-943.
41. T. Y. Hou and D.-P. Yang, *Convection of Microstructure in Two and Three Dimensional Incompressible Euler Equations,* in preparation, 2002.
42. T. J. R. Hughes, *Multiscale Phenomena: Green's Functions, the Dirichlet-to-Neumann Formulation, Subgrid Scale Models, Bubbles and the Origins of Stabilized Methods,* Comput. Methods Appl. Mech Engrg., **127** (1995), 387-401.
43. T. J. R. Hughes, G. R. Feijóo, L. Mazzei, J.-B. Quincy, *The Variational Multiscale Method – A Paradigm for Computational Mechanics,* Comput. Methods Appl. Mech Engrg., **166**(1998), 3-24.
44. V. V. Jikov, S. M. Kozlov, and O. A. Oleinik, *Homogenization of Differential Operators and Integral Functionals,* Springer-Verlag, 1994, Translated from Russian.

45. S. Knapek, *Matrix-Dependent Multigrid-Homogenization for Diffusion Problems,* in the Proceedings of the Copper Mountain Conference on Iterative Methods, edited by T. Manteuffal and S. McCormick, volume I, SIAM Special Interest Group on Linear Algebra, Cray Research , 1996.
46. P. Langlo and M.S. Espedal, *Macrodispersion for Two-phase, Immiscible Flow in Porous Media,* Adv. Water Resources **17** (1994), 297–316.
47. A. M. Matache, I. Babuska, and C. Schwab, *Generalized p-FEM in Homogenization,* Numer. Math. **86**(2000), 319-375.
48. A. M. Matache and C. Schwab, *Homogenization via p-FEM for Problems with Microstructure,* Appl. Numer. Math. **33** (2000), 43-59.
49. J. F. McCarthy, *Comparison of Fast Algorithms for Estimating Large-Scale Permeabilities of Heterogeneous Media,* Transport in Porous Media, **19** (1995), 123-137.
50. D. W. McLaughlin, G. C. Papanicolaou, and L. Tartar, *Weak Limits of Semilinear Hyperbolic Systems with Oscillating Data,* Lecture Notes in Physics **230** (1985), 277-289, Springer-Verlag, Berlin, New York.
51. D. W. McLaughlin, G. C. Papanicolaou, and O. Pironneau, *Convection of Microstructure and Related Problems,* SIAM J. Applied Math, **45** (1985), 780-797.
52. S. Moskow and M. Vogelius, *First Order Corrections to the Homogenized Eigenvalues of a Periodic Composite Medium: A Convergence Proof,* Proc. Roy. Soc. Edinburgh, A, **127** (1997), 1263-1299.
53. F. Murat, *Compacité par Compensation,* Ann. Scuola Norm. Sup. Pisa, **5** (1978), 489-507.
54. F. Murat, *Compacité par compensation II,*, Proceedings of the International Meeting on Recent Methods in Nonlinear Analysis, Rome, May 8-12, 1978, ed. by E. De Giorgi, E. Magenes and U. Mosco, Pitagora Editrice, Bologna, 245-256, 1979.
55. H. Neiderriter, *Quasi-Monte Carlo Methods and Pseudo-Random Numbers,* Bull. Amer. Math. Soc., **84** (1978), 957-1041.
56. O. Pironneau, *On the Transport-diffusion Algorithm and its Application to the Navier-Stokes Equations*, Numer. Math. **38** (1982), 309–332.
57. G. Sangalli, *Capturing Small Scales in Elliptic Problems Using a Residual-Free Bubbles Finite Element Method,* to appear in Multiscale Modeling and Simulation.
58. L. Tartar, *Compensated Compactness and Applications to P.D.E.,* Nonlinear Analysis and Mechanics, Heriot-Watt Symposium, Vol. IV, ed. by R. J. Knops, Research Notes in Mathematics **39**, Pitman, Boston, 136-212, 1979.
59. L. Tartar, *Solutions oscillantes des équations de Carleman,* Seminaire Goulaouic-Meyer-Schwartz (1980-1981), exp. XII. Ecole Polytechnique (Palaiseau), 1981.
60. L. Tartar, *Nonlocal Effects Induced by Homogenization,* in PDE and Calculus of Variations, ed by F. Culumbini, et al, Birkhäuser, Boston, 925-938, 1989.
61. X.H. Wu, Y. Efendiev, and T. Y. Hou, *Analysis of Upscaling Absolute Permeability,* Discrete and Continuous Dynamical Systems, Series B, **2** (2002), 185-204.
62. P. M. De Zeeuw, *Matrix-dependent Prolongation and Restrictions in a Blackbox Multigrid Solver,* J. Comput. Applied Math, **33**(1990), 1-27.
63. S. Verdiere and M.H. Vignal, *Numerical and Theoretical Study of a Dual Mesh Method Using Finite Volume Schemes for Two-phase Flow Problems in Porous Media,* Numer. Math. **80** (1998), 601–639.

64. T. C. Wallstrom, M. A. Christie, L. J. Durlofsky, and D. H. Sharp, *Effective Flux Boundary Conditions for Upscaling Porous Media Equations,* Transport in Porous Media, **46** (2002), 139-153.
65. T. C. Wallstrom, M. A. Christie, L. J. Durlofsky, and D. H. Sharp, *Application of Effective Flux Boundary Conditions to Two-phase Upscaling in Porous Media,* Transport in Porous Media, **46** (2002), 155-178.
66. T. C. Wallstrom, S. L. Hou, M. A. Christie, L. J. Durlofsky, and D. H. Sharp, *Accurate Scale Up of Two Phase Flow Using Renormalization and Nonuniform Coarsening,* Comput. Geosci, **3** (1999), 69-87.