

# Multiscale Computations for Flow and Transport in Heterogeneous Media

Yalchin Efendiev and Thomas Y. Hou

Applied Mathematics, 217-50, Caltech, Pasadena, CA 91125, USA. Email: hou@ama.caltech.edu. Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, USA. Email efendieve@math.tamu.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.

The purpose of this lecture note is to review some recent advances in developing multiscale finite element (volume) methods for flow and transport in strongly heterogeneous porous media. Extra effort is made in developing a multiscale computational method that can be potentially used for practical multiscale for problems with a large range of nonseparable scales. Some recent theoretical and computational developments in designing global upscaling methods will be reviewed. The lectures can be roughly divided into 4 parts. In part 1, I will review some homogenization theory for elliptic and hyperbolic equations. This homogenization theory provides the critical guideline for designing effective multiscale methods. In part 2, I will review some recent developments of multiscale finite element (volume) methods. We also discuss the issue of upscaling one-phase, two-phase flows through heterogeneous porous media and the use of limited global information in multiscale finite element (volume) methods. In part 4, we will consider multiscale simulations of two-phase flow immiscible flows using a flow-based adaptive coordinate, and introduce a theoretical framework which enables us to perform global upscaling for heterogeneous media with long range connectivity.

## 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 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. Upscaling

procedures have been commonly applied for this purpose and are effective in many cases. More recently, a number of multiscale techniques have been developed and successfully applied to various areas, e.g., porous media flows. The main idea of upscaling techniques is to form coarse-scale equations with a prescribed analytical form that may differ from the underlying fine-scale equations. In multiscale methods, the fine-scale information is carried throughout the simulation and the coarse-scale equations are generally not expressed analytically, but rather formed and solved numerically.

The purpose of this lecture note is to review some recent advances in developing multiscale finite element (volume) methods for flow and transport in strongly heterogeneous porous media. Extra effort is made in developing a multiscale computational method that can be potentially used for practical multiscale for problems with a large range of nonseparable scales. Substantial progress has been made in recent years by combining modern mathematical techniques such as multiscale analysis, adaptivity and multiresolution. The lectures can be roughly divided into four parts. In Section 2, we 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. In Section 3, we discuss some recent developments of multiscale finite element methods. We also discuss the issue of upscaling one-phase, two-phase flows through heterogeneous porous media and the use of limited global information in multiscale finite element methods. In Section 4, we discuss the generalization of the multiscale finite element methods to nonlinear partial differential equations. In Section 5, we will consider multiscale simulations of two-phase flow immiscible flows using a flow-based adaptive coordinate system. 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

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

where  $a_{ij}(y)$  and  $a_0(y)$  are 1-periodic in both variables of  $y$ , 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  $a_\varepsilon$  representing the permeability tensor. For composite materials, it is the steady heat conduction equation and the coefficient  $a_\varepsilon$  represents the thermal conductivity. For steady transport problems, it is a symmetrized form of the governing equation. In this case, the coefficient  $a_\varepsilon$  is a combination of transport velocity and viscosity tensor.

Homogenization theory is to study the limiting behavior  $u_\varepsilon \rightarrow u$  as  $\varepsilon \rightarrow 0$ . The main task is to find the homogenized coefficients,  $a_{ij}^*$  and  $a_0^*$ , and the *homogenized equation* for the limiting solution  $u$

$$-\frac{\partial}{\partial x_i} \left( a_{ij}^* \frac{\partial}{\partial x_j} \right) u + a_0^* u = f, \quad u|_{\partial\Omega} = 0. \quad (2.2)$$

Define the  $L^2$  and  $H^1$  norms over  $\Omega$  as follows

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

Further, we define the bilinear form

$$a^\varepsilon(u, v) = \int_{\Omega} a_{i,j}^\varepsilon(x) \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} dx + \int_{\Omega} a_0^\varepsilon uv dx. \quad (2.4)$$

It is easy to show that

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

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  $u_\varepsilon \in H_0^1$

$$a^\varepsilon(u_\varepsilon, v) = (f, v), \quad \text{for all } v \in H_0^1(\Omega), \quad (2.6)$$

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

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

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

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  $u_\varepsilon$ , such that

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

On the other hand, we notice that

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

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  $a^\varepsilon$  is bounded, and  $u_x^\varepsilon$  is bounded in  $L^2(\Omega)$ , so  $\xi^\varepsilon$  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

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

so that

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

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

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

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

**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

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

where the functions  $u_j(x, y)$  are double periodic in  $y$  with period 1.

Denote by  $A^\varepsilon$  the second order elliptic operator

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

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  $A^\varepsilon$  as follows

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

where

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

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

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

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

$$A_1 u_0 = 0, \quad (2.17)$$

$$A_1 u_1 + A_2 u_0 = 0, \quad (2.18)$$

$$A_1 u_2 + A_2 u_1 + A_3 u_0 = f. \quad (2.19)$$

Equation (2.17) can be written as

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

where  $u_0$  is periodic in  $y$ . The theory of second order elliptic PDEs [46] 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*

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

where  $\chi^j$  is double periodic in  $y$ . The general solution of equation (2.18) for  $u_1$  is then given by

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

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

$$\frac{\partial}{\partial y_i} \left( a_{ij}(y) \frac{\partial}{\partial y_j} \right) u_2 = A_2 u_1 + A_3 u_0 - f. \quad (2.23)$$

The solvability condition implies that the right hand side of (2.23) must have mean zero in  $y$  over one periodic cell  $Y = [0, 1] \times [0, 1]$ , i.e.

$$\int_Y (A_2 u_1 + A_3 u_0 - f) dy = 0.$$

This solvability condition for second order elliptic PDEs with periodic boundary condition [46] 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  $u$ :

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

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

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

**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  $A^\varepsilon$  to  $z_\varepsilon$ , we get

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

where  $r_\varepsilon = A_2 u_2 + A_3 u_1 + \varepsilon A_3 u_2$ . If  $f$  is smooth enough, so is  $u_2$ . Thus we have  $\|r_\varepsilon\|_\infty \leq c$ .

On the other hand, we have

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

Thus, we obtain

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

It follows from the maximum principle [46] 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 [70] have shown that

$$\begin{aligned} \|u_\varepsilon - u - \varepsilon u_1(x, x/\varepsilon) + \varepsilon \theta_\varepsilon\|_0 &\leq C_\omega \varepsilon^{1+\omega} \|u\|_{2+\omega}, \\ \|u_\varepsilon - u - \varepsilon u_1(x, x/\varepsilon) + \varepsilon \theta_\varepsilon\|_1 &\leq C\varepsilon \|u\|_2, \end{aligned} \tag{2.26}$$

where we assume  $u \in H^{2+\omega}(\Omega)$  with  $0 \leq \omega \leq 1$ , and  $\Omega$  is assumed to be a bounded, convex curvilinear polygon of class  $C^\infty$ . This improved estimate will be used in the convergence analysis of the multiscale finite element method to be presented in Section 3.

## 2.2 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 [69] 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:

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

The pressure field  $p^\varepsilon$  is expanded similarly. From this ansatz, one can show that  $\theta$  is convected by the mean velocity:

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

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 2.1.* One possible way to improve the work of [69] is take into account the oscillation in the Lagrangian characteristics,  $\theta_\varepsilon$ . The oscillatory part of  $\theta_\varepsilon$  in general could have order one contribution to the mean velocity of the incompressible Euler equation. In [53–55], Hou and Yang and co-workers have studied convection of microstructure of the 2-D and 3-D incompressible Euler equations using a new approach. They do not assume that the oscillation is propagated by the mean flow. In fact, they found that it is crucial to include the effect of oscillations in the characteristics on the mean flow. Using this new approach, they 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  $v^\varepsilon$  in powers of  $\varepsilon$

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

The coefficients of  $\varepsilon^{-1}$  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  $\varepsilon^0$  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 [69]

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

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  $u_\varepsilon(x, y, 0) = u_0(x, y)$ . Note that  $y = x_2$  is not a fast variable here.

We assume that  $a_\varepsilon$  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  $u_\varepsilon$ .

Using Laplace Transform and measure theory, Luc Tartar [83] showed that the weak limit  $u$  of  $u_\varepsilon$  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  $\mu_y$  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  $\mu_y$  cannot be expressed explicitly in terms of  $a_\varepsilon$ .

### 3 Numerical Upscaling based on Multiscale Finite Element Methods

In this section, we review the multiscale finite element method (MsFEM) for solving partial differential equations with multiscale solutions, see [50,52,51,37,22,84,3,35]. 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 [52].

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. For example, the high degree of variability and multiscale nature of formation properties in subsurface flows (such as permeability) pose significant challenges for subsurface flow modeling. Geological characterizations that capture these effects are typically developed at scales that are too fine for direct flow simulation, so techniques are required to enable the solution of flow problems in practice. Upscaling procedures have been commonly applied for this purpose and are effective in many cases (see e.g., [60] for reviews and discussion). Our multiscale finite element method can be used for a similar purpose and successfully applied for problems of this type.

As discussed in [60], upscaling methods and multiscale numerical techniques (as applied within the context of subsurface flow modeling) have many similarities and some important differences. Upscaling techniques provide coefficients, which are typically computed in a pre-processing step, for coarse scale equations of prescribed analytical forms. In multiscale methods, the coarse scale equations are formed numerically and fine scale information may be carried throughout the simulation and used at various stages. For example, in multiscale procedures for subsurface flow applications, different grids are often used for flow and transport computations. The advantage of deriving a scale-up equation or performing multiscale computations is that one can perform many useful tests on the coarse 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 coarse-scale equation also allows for larger time steps. This results in additional computational saving.

It should be mentioned that many numerical methods have been developed with goals similar to ours. These include generalized finite element methods [12,10,9], wavelet based numerical homogenization methods [17,27,25,64], methods based on the homogenization theory (cf. [15,31,24]), equation-free computations (e.g., [63]), variational multiscale methods [58,20,59], heterogeneous multiscale methods [33], matrix-dependent multigrid based homogenization [64,25], generalized  $p$ -FEM in homogenization [66,67], and some upscaling methods based on simple physical and/or mathematical motivations (cf. [29,68]). 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 [14,62]. 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. [29,68]). 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.

Most multiscale methods presented to date have applied local calculations for the determination of basis functions. Though effective in many cases, global effects can be important for some problems. The importance of global information has been illustrated within the context of upscaling procedures as well as multiscale computations in recent investigations. These studies have shown that the use of limited global information in



the calculation of the coarse-scale parameters (such as basis functions) can significantly improve the accuracy of the resulting coarse model. In this lecture notes, we describe the use of limited global information in multiscale simulations.

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., [13] and references therein). Babuska et al. applied a similar idea to 1-D problems [12] and to a special class of 2-D problems with the coefficient varying locally in one direction [10]. Most of these methods are based on the special property of one-dimensional properties of the coefficients. 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.

### 3.1 Multiscale Finite Element Methods for Elliptic PDEs.

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

$$L_\varepsilon p := -\nabla \cdot (a(\frac{x}{\varepsilon})\nabla p) = f \text{ in } \Omega, \quad p = 0 \text{ on } \Gamma = \partial\Omega, \quad (3.1)$$

where  $\Omega$  is a convex polygon in  $\mathbb{R}^2$ . The unknown is changed to  $p$ , since it will be used later in porous media flow simulations, where the solution represents the pressure field.  $\varepsilon$  is assumed to be a small parameter, and  $a(x) = (a_{ij}(x/\varepsilon))$  is symmetric and satisfies  $\alpha|\xi|^2 \leq a_{ij}\xi_i\xi_j \leq \beta|\xi|^2$ , 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, which will be discussed later.

Let  $p_0$  be the solution of the homogenized equation

$$L_0 p_0 := -\nabla \cdot (a^* \nabla p_0) = f \text{ in } \Omega, \quad p_0 = 0 \text{ on } \Gamma, \quad (3.2)$$

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

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

It is clear that  $p_0 \in H^2(\Omega)$  since  $\Omega$  is a convex polygon. Denote by  $p_1(x, y) = -\chi^j(y) \frac{\partial p_0(x)}{\partial x_j}$  and let  $\theta_\varepsilon$  be the solution of the problem

$$L_\varepsilon \theta_\varepsilon = 0 \text{ in } \Omega, \quad \theta_\varepsilon(x) = p_1(x, \frac{x}{\varepsilon}) \text{ on } \Gamma. \quad (3.3)$$

Our analysis of the multiscale finite element method relies on the following homogenization result obtained by Moskow and Vogelius [70].

**Lemma 3.1.** *Let  $p_0 \in H^2(\Omega)$  be the solution of (3.2),  $\theta_\varepsilon \in H^1(\Omega)$  be the solution to (3.3) and  $p_1(x) = -\chi^j(x/\varepsilon)\partial p_0(x)/\partial x_j$ . Then there exists a constant  $C$  independent of  $u_0, \varepsilon$  and  $\Omega$  such that*

$$\|p - p_0 - \varepsilon(u_1 - \theta_\varepsilon)\|_{1,\Omega} \leq C\varepsilon(\|p_0\|_{2,\Omega} + \|f\|_{0,\Omega}).$$

Now we are going to introduce the multiscale finite element methods. Let  $\mathcal{T}_h$  be a regular partition of  $\Omega$  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  $\phi^i$  with support in  $S_i$  as follows:

$$L_\varepsilon \phi^i = 0 \text{ in } K, \quad \phi^i = \psi_i \text{ on } \partial K \quad \forall K \in \mathcal{T}_h, K \subset S_i. \quad (3.4)$$

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  $p_h \in V_h$  such that

$$(a(\frac{x}{\varepsilon})\nabla p_h, \nabla v_h) = (f, v_h) \quad \forall v_h \in V_h, \quad (3.5)$$

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  $p$  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  $p$  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 [19] and the variational multiscale method [58,20]. In a recent paper [77], Dr. G. Sangalli derives a multiscale method based on the residual-free bubbles formulation in [77] and compares it with the multiscale finite element method described here. There are many striking similarities between the two approaches.

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.

### 3.2 Error Estimates ( $h < \varepsilon$ )

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

**Lemma 3.2.** *Let  $p$  be the solution of (3.1) and  $p_h$  be the solution of (3.5). Then we have*

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

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

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

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

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

From the definition of the basis function  $\phi^i$  in (3.4) we have

$$L_\varepsilon(I_h p) = 0 \quad \text{in } K, \quad I_h p = \Pi_h p \quad \text{on } \partial K, \quad (3.6)$$

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

**Lemma 3.3.** *Let  $p \in H^2(\Omega)$  be the solution of (3.1). Then there exists a constant  $C$  independent of  $h, \varepsilon$  such that*

$$\|p - I_h p\|_{0,\Omega} + h \|p - I_h p\|_{1,\Omega} \leq Ch^2 (\|p\|_{2,\Omega} + \|f\|_{0,\Omega}). \quad (3.7)$$

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

$$\|p - \Pi_h p\|_{0,\Omega} + h\|p - \Pi_h p\|_{1,\Omega} \leq Ch^2(|p|_{2,\Omega} + \|f\|_{0,\Omega}). \quad (3.8)$$

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

$$\|\Pi_h p - I_h p\|_{0,K} \leq Ch|\Pi_h p - I_h p|_{1,K} \quad \forall K \in \mathcal{T}_h. \quad (3.9)$$

To estimate  $|\Pi_h p - I_h p|_{1,K}$  we multiply the equation in (3.6) by  $I_h p - \Pi_h p \in H_0^1(K)$  to get

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

where  $(\cdot, \cdot)_K$  denotes the  $L^2$  inner product of  $L^2(K)$ . Thus, upon using the equation in (3.1), we get

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

This implies that

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

Hence

$$|I_h p - \Pi_h p|_{1,K} \leq Ch(|p|_{2,K} + \|f\|_{0,K}), \quad (3.10)$$

where we have used (3.9). Now the lemma follows from (3.8)-(3.10).  $\square$

In conclusion, we have the following estimate by using Lemmas 3.2-3.3.

**Theorem 3.1.** *Let  $p \in H^2(\Omega)$  be the solution of (3.1) and  $p_h \in V_h$  be the solution of (3.5). Then we have*

$$\|p - p_h\|_{1,\Omega} \leq Ch(|p|_{2,\Omega} + \|f\|_{0,\Omega}). \quad (3.11)$$

Note that the estimate (3.11) blows up like  $h/\varepsilon$  as  $\varepsilon \rightarrow 0$  since  $|p|_{2,\Omega} = O(1/\varepsilon)$ . This is insufficient for practical applications. In next subsection we derive an error estimate which is uniform as  $\varepsilon \rightarrow 0$ .

### 3.3 Error Estimates ( $h > \varepsilon$ )

In this section, we will show that the multiscale finite element method gives a convergence result uniform in  $\varepsilon$  as  $\varepsilon$  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 3.2.** *Let  $p \in H^2(\Omega)$  be the solution of (3.1) and  $p_h \in V_h$  be the solution of (3.5). Then we have*

$$\|p - p_h\|_{1,\Omega} \leq C(h + \varepsilon)\|f\|_{0,\Omega} + C\left(\frac{\varepsilon}{h}\right)^{1/2} \|p_0\|_{1,\infty,\Omega}, \quad (3.12)$$

where  $p_0 \in H^2(\Omega) \cap W^{1,\infty}(\Omega)$  is the solution of the homogenized equation (3.2).

To prove the theorem, we first denote by

$$p_I(x) = I_h p_0(x) = \sum_{j=1}^J p_0(x_j) \phi^j(x) \in V_h.$$

From (3.6) we know that  $L_\varepsilon p_1 = 0$  in  $K$  and  $p_1 = \Pi_h p_0$  on  $\partial K$  for any  $K \in \mathcal{T}_h$ . The homogenization theory (see (2.26)) implies that

$$\|p_1 - p_{10} - \varepsilon(p_{11} - \theta_{1\varepsilon})\|_{1,K} \leq C\varepsilon(\|f\|_{0,K} + |p_{10}|_{2,K}), \quad (3.13)$$

where  $p_{10}$  is the solution of the homogenized equation on  $K$ :

$$L_0 p_{10} = 0 \text{ in } K, \quad p_{10} = \Pi_h p_0 \text{ on } \partial K, \quad (3.14)$$

$p_{11}$  is given by the relation

$$p_{11}(x, y) = -\chi^j(y) \frac{\partial p_{10}}{\partial x_j} \text{ in } K, \quad (3.15)$$

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

$$L_\varepsilon \theta_{1\varepsilon} = 0 \text{ in } K, \quad \theta_{1\varepsilon}(x) = p_{11}(x, \frac{x}{\varepsilon}) \text{ on } \partial K. \quad (3.16)$$

It is obvious from (3.14) that

$$p_{10} = \Pi_h p_0 \text{ in } K, \quad (3.17)$$

since  $\Pi_h p_0$  is linear on  $K$ . From (3.13) we obtain that

$$\begin{aligned} \|p - p_1\|_{1,\Omega} &\leq \|p_0 - p_{10}\|_{1,\Omega} + \|\varepsilon(p_1 - p_{11})\|_{1,\Omega} \\ &\quad + \|\varepsilon(\theta_\varepsilon - \theta_{1\varepsilon})\|_{1,\Omega} + C\varepsilon\|f\|_{0,\Omega}, \end{aligned} \quad (3.18)$$

where we have used the regularity estimate  $\|p_0\|_{2,\Omega} \leq C\|f\|_{0,\Omega}$ . Now it remains to estimate the terms at the right-hand side of (3.18).

**Lemma 3.4.** *We have*

$$\|p_0 - p_{10}\|_{1,\Omega} \leq Ch\|f\|_{0,\Omega}, \quad (3.19)$$

$$\|\varepsilon(p_1 - p_{11})\|_{1,\Omega} \leq C(h + \varepsilon)\|f\|_{0,\Omega}. \quad (3.20)$$

*Proof.* The estimate (3.19) is a direct consequence of the standard finite element interpolation theory since  $p_{10} = \Pi_h p_0$  by (3.17). Next we note that  $\chi^j(x/\varepsilon)$  satisfies

$$\|\chi^j\|_{0,\infty,\Omega} + \varepsilon\|\nabla\chi^j\|_{0,\infty,\Omega} \leq C \quad (3.21)$$

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

$$\begin{aligned} \|\varepsilon(p_1 - p_{11})\|_{0,K} &\leq C\varepsilon\|\chi^j \frac{\partial}{\partial x_j}(p_0 - \Pi_h p_0)\|_{0,K} \leq Ch\varepsilon|p_0|_{2,K}, \\ \|\varepsilon\nabla(p_1 - p_{11})\|_{0,K} &= \varepsilon\|\nabla(\chi^j \frac{\partial(p_0 - \Pi_h p_0)}{\partial x_j})\|_{0,K} \\ &\leq C\|\nabla(p_0 - \Pi_h p_0)\|_{0,K} + C\varepsilon|p_0|_{2,K} \\ &\leq C(h + \varepsilon)|p_0|_{2,K}. \end{aligned}$$

This completes the proof.  $\square$

**Lemma 3.5.** *We have*

$$\|\varepsilon\theta_\varepsilon\|_{1,\Omega} \leq C\sqrt{\varepsilon}\|p_0\|_{1,\infty,\Omega} + C\varepsilon|p_0|_{2,\Omega}. \quad (3.22)$$

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

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

With this definition, it is clear that  $\theta_\varepsilon - \zeta p_1 = \theta_\varepsilon + \zeta(\chi^j \partial p_0 / \partial x_j) \in H_0^1(\Omega)$ . Multiplying the equation in (3.3) by  $\theta_\varepsilon - \zeta p_1$ , we get

$$(a(\frac{x}{\varepsilon}) \nabla \theta_\varepsilon, \nabla(\theta_\varepsilon + \zeta \chi^j \frac{\partial p_0}{\partial x_j})) = 0,$$

which yields, by using (3.21),

$$\begin{aligned} \|\nabla \theta_\varepsilon\|_{0,\Omega} &\leq C \|\nabla(\zeta \chi^j \partial p_0 / \partial x_j)\|_{0,\Omega} \\ &\leq C \|\nabla \zeta \cdot \chi^j \partial p_0 / \partial x_j\|_{0,\Omega} + C \|\zeta \nabla \chi^j \partial p_0 / \partial x_j\|_{0,\Omega} \\ &\quad + C \|\zeta \chi^j \partial^2 p_0 / \partial^2 x_j\|_{0,\Omega} \\ &\leq C \sqrt{|\partial\Omega|} \cdot \delta \frac{D}{\delta} + C \sqrt{|\partial\Omega|} \cdot \delta \frac{D}{\varepsilon} + C \|p_0\|_{2,\Omega}, \end{aligned} \quad (3.23)$$

where  $D = \|p_0\|_{1,\infty,\Omega}$  and the constant  $C$  is independent of the domain  $\Omega$ . From (3.23) we have

$$\begin{aligned} \|\varepsilon \theta_\varepsilon\|_{0,\Omega} &\leq C \left( \frac{\varepsilon}{\sqrt{\delta}} + \sqrt{\delta} \right) \|p_0\|_{1,\infty,\Omega} + C \varepsilon \|p_0\|_{2,\Omega} \\ &\leq C \sqrt{\varepsilon} \|p_0\|_{1,\infty,\Omega} + C \varepsilon \|p_0\|_{2,\Omega}. \end{aligned} \quad (3.24)$$

Moreover, by applying the maximum principle to (3.3), we get

$$\|\theta_\varepsilon\|_{0,\infty,\Omega} \leq \|\chi^j \partial p_0 / \partial x_j\|_{0,\infty,\partial\Omega} \leq C \|p_0\|_{1,\infty,\Omega}. \quad (3.25)$$

Combining (3.24) and (3.25) completes the proof.  $\square$

**Lemma 3.6.** *We have*

$$\|\varepsilon \theta_{I_\varepsilon}\|_{1,\Omega} \leq C \left( \frac{\varepsilon}{h} \right)^{1/2} \|p_0\|_{1,\infty,\Omega}. \quad (3.26)$$

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

$$L_\varepsilon \theta_{I_\varepsilon} = 0 \text{ in } K, \quad \theta_{I_\varepsilon} = -\chi^j \left( \frac{x}{\varepsilon} \right) \frac{\partial(\Pi_h p_0)}{\partial x_j} \text{ on } \partial K. \quad (3.27)$$

By applying maximum principle and (3.21) we get

$$\|\theta_{I_\varepsilon}\|_{0,\infty,K} \leq \|\chi^j \partial(\Pi_h p_0) / \partial x_j\|_{0,\infty,\partial K} \leq C \|p_0\|_{1,\infty,K}.$$

Thus we have

$$\|\varepsilon \theta_{I_\varepsilon}\|_{0,\Omega} \leq C \varepsilon \|p_0\|_{1,\infty,\Omega}. \quad (3.28)$$

On the other hand, since the constant  $C$  in (3.23) is independent of  $\Omega$ , we can apply the same argument leading to (3.23) to obtain

$$\begin{aligned} \|\varepsilon \nabla \theta_{I_\varepsilon}\|_{0,K} &\leq C \varepsilon \|\Pi_h p_0\|_{1,\infty,K} (\sqrt{|\partial K|} / \sqrt{\delta} + \sqrt{|\partial K|} \cdot \delta / \varepsilon) + C \varepsilon \|\Pi_h p_0\|_{2,K} \\ &\leq C \sqrt{h} \|p_0\|_{1,\infty,K} \left( \frac{\varepsilon}{\sqrt{\delta}} + \sqrt{\delta} \right) \\ &\leq C \sqrt{h \varepsilon} \|p_0\|_{1,\infty,K}, \end{aligned}$$

which implies that

$$\|\varepsilon \nabla \theta_{I_\varepsilon}\|_{0,\Omega} \leq C \left( \frac{\varepsilon}{h} \right)^{1/2} \|p_0\|_{1,\infty,\Omega}.$$

This completes the proof.  $\square$

*Proof (of Theorem 3.2).* The theorem is now a direct consequence of (3.18) and the Lemma 3.4-3.6 and the regularity estimate  $\|p_0\|_{2,\Omega} \leq C\|f\|_{0,\Omega}$ .  $\square$

*Remark 3.1.* As we pointed out earlier, the multiscale FEM indeed gives correct homogenized result as  $\varepsilon$  tends to zero. This is in contrast with the traditional FEM which does not give the correct homogenized result as  $\varepsilon \rightarrow 0$ . 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 ( $\varepsilon$ ) 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.

### 3.4 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 [50]. 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  $\varepsilon$  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  $\psi^j$  be the base functions satisfying the homogeneous elliptic equation in the larger domain  $S \supset K$ . We then form the actual base  $\phi^i$  by linear combination of  $\psi^j$ ,

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

The coefficients  $c_{ij}$  are determined by condition  $\phi^i(\mathbf{x}_j) = \delta_{ij}$ . The corresponding  $\theta_\varepsilon^i$  for  $\phi^i$  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 [37], we perform a careful estimate of the non-conforming 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 [50,51]. 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 (see [48]).

### 3.5 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 [85]). The global linear system on  $\Omega$  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 [50].

**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  $1024 \times 1024$  elements [50]. 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 [50]. It was also found that MsFEM is helpful for improving multigrid convergence when the coefficient  $a_\epsilon$  has very large contrast (i.e., the ratio between the maximum and minimum of  $a_\epsilon$ ).

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.

**MsFEM for problems with scale separation** If there is a scale separation in representative volumes smaller than the coarse block, then multiscale finite element basis functions can be computed based on the smaller regions. To demonstrate this, we first consider a periodic case. In this case, the basis functions can be approximated by

$$\phi^j(x) = \phi_0^j(x) + \epsilon \chi^i \nabla_i \phi_0^j.$$

Consequently, the approximation of the basis functions can be carried out in a domain of size  $\epsilon$  via the computation of  $\chi^i$ . This reduces the computational cost. Moreover, the assembly of stiffness matrix can be

also performed in a period, because  $a(x/\epsilon)\nabla\phi^i \cdot \nabla\phi^j$  is a periodic function. The results obtained by this approximation gives the classical numerical homogenization procedure that is based on the computation of effective coefficients based on periodic problems. We would like to note that this approximation procedure is not limited to periodic problems and can be applied to random homogeneous problems with the strong scale separation, i.e., the size of representative volume is much smaller than the coarse mesh size. In general, this holds for problems where homogenization by periodization (see [62]) is true. Random homogeneous case with ergodicity is one of them. We note that a number of methods used in practice employs this strategy (e.g., [76,63,44,33]).

**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 3.1 we present the result for

$$a(\mathbf{x}/\epsilon) = \frac{2 + P \sin(2\pi x/\epsilon)}{2 + P \cos(2\pi y/\epsilon)} + \frac{2 + \sin(2\pi y/\epsilon)}{2 + P \sin(2\pi x/\epsilon)} \quad (P = 1.8), \quad (3.29)$$

$$f(\mathbf{x}) = -1 \quad \text{and} \quad u|_{\partial\Omega} = 0. \quad (3.30)$$

The convergence of three different methods are compared for fixed  $\epsilon/h = 0.64$ , 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 resonance in the

$N$	$\epsilon$	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 3.1.** Convergence for periodic case.

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 [52,50], more extensive convergence tests have been presented.

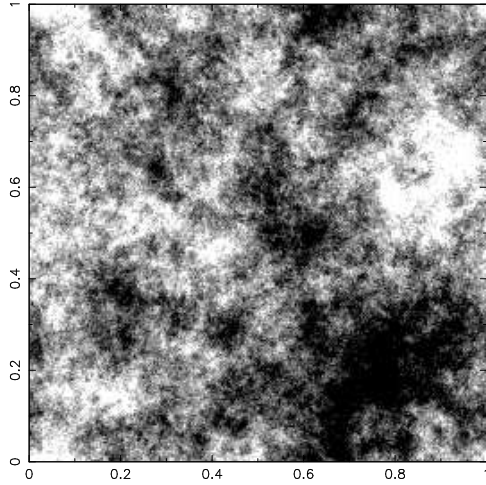
### 3.6 Applications

**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.

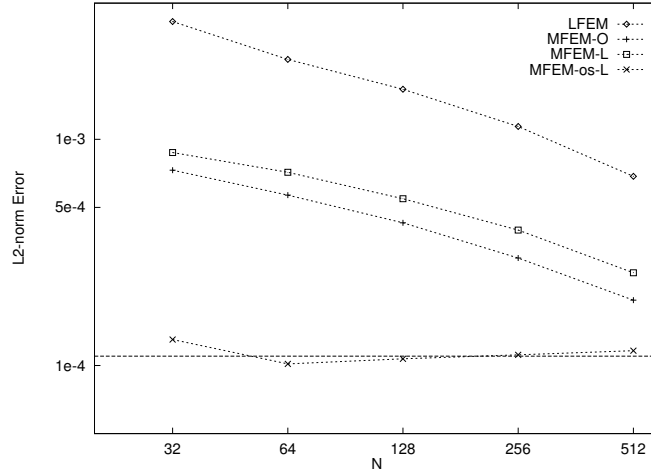
We model the porous media by random distributions of  $a_\epsilon$  generated using a spectral method. In fact,  $a_\epsilon = \alpha 10^{\beta p}$ , where  $p$  is a random field represents porosity, and  $\alpha$  and  $\beta$  are scaling constants to give the desired contrast of  $a_\epsilon$ . In particular, we have tested the method for a porous medium with a statistically fractal porosity field (see Figure 3.1). The fractal dimension is 2.8. This is a model of flow in an oil reservoir or aquifer with uniform injection in the domain and outflow at the boundaries. We note that the problem has a continuous scale because of the fractal distribution.

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





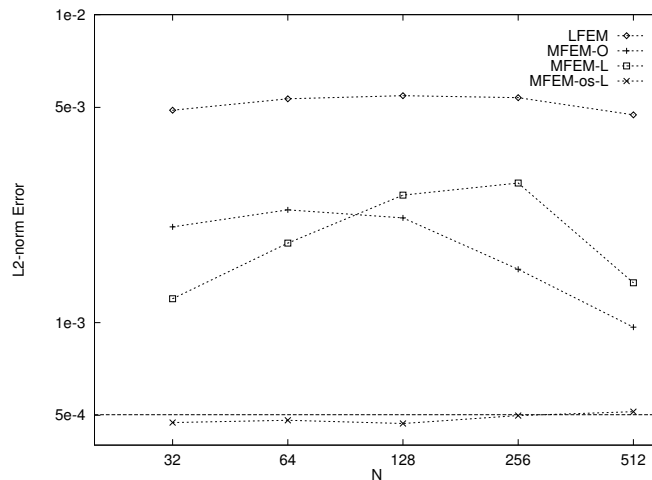
**Fig. 3.1.** Porosity field with fractal dimension of 2.8 generated using the spectral method.



**Fig. 3.2.** The  $l^2$ -norm error of the solutions using various schemes for a fractal distributed permeability field.

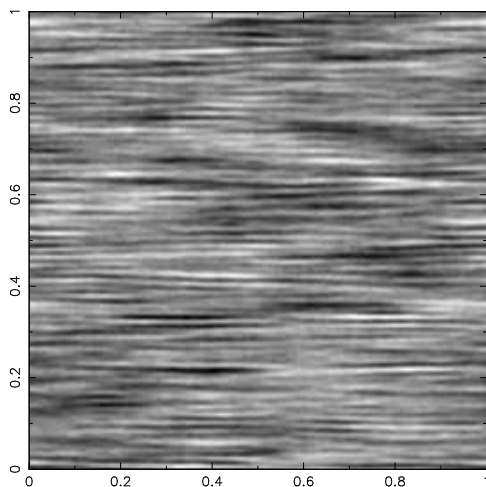
with that of the well-resolved solution obtained using LFEM. However, MsFEM without over-sampling is less accurate. MsFEM-O indicates that oscillatory boundary conditions, obtained from solving some reduced 1-D elliptic equations along  $\partial\mathbf{K}$  (see [50]), are imposed on the base functions. The decay of error in MsFEM is because of the decay of small scales in  $a_\varepsilon$ . The next figure shows the results for a log-normally distributed  $a_\varepsilon$ . In this case, the effect of scale resonance shows clearly for MsFEM-L, i.e., the error increases as  $h$  approaches  $\varepsilon$ . 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. [29]). However, for anisotropic media, especially



**Fig. 3.3.** The  $l^2$ -norm error of the solutions using various schemes for a log-normally distributed permeability field.

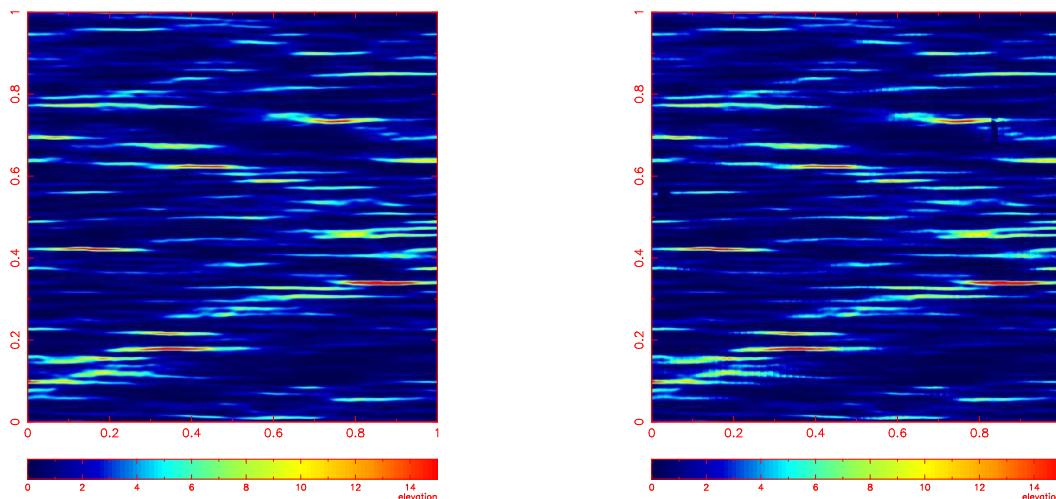
layered ones (Figure 3.4), the velocity in some thin channels can be much higher than the cell average, and these channels often have dominant effects on the transport solutions. In this case, the information about fine scale velocity becomes vitally important. Therefore, an important question for all upscaling methods is how to take those fast-flow channels into account.



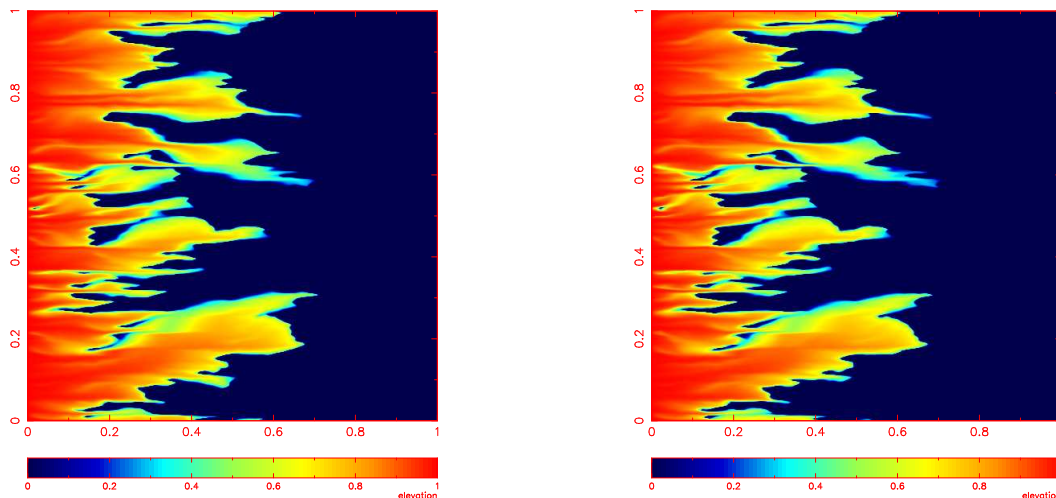
**Fig. 3.4.** A random porosity field with layered structure.

For MsFEM, the fine scale velocity can be easily recovered from the multiscale 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 [37]. 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 3.4: a fluid with red color 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 [30]). 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 3.5a, we plot the horizontal velocity field



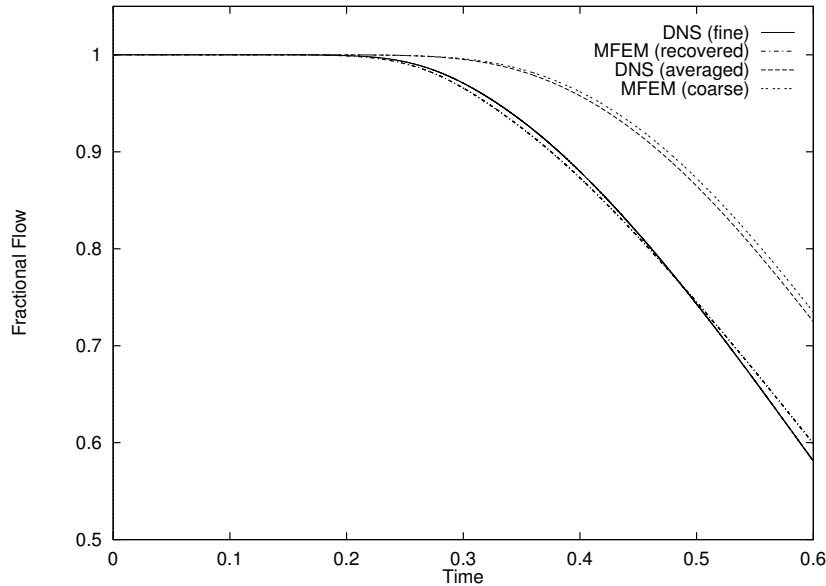
**Fig. 3.5.** (a): Fine grid horizontal velocity field,  $N = 1024$ . (b): Recovered horizontal velocity field from the coarse grid  $N = 64$  calculation using multiscale bases.



**Fig. 3.6.** (a): Fine grid saturation at  $t = 0.06$ ,  $N = 1024$ . (b): Saturation computed using the recovered velocity field from the coarse grid calculation.

obtained by using a fine grid ( $N = 1024$ ) calculation. In Figure 3.5b, 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 3.6a, we plot the saturation at  $t = 0.06$  obtained by the fine grid calculation. Figure 3.6b 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 3.7. 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. 3.7.** 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. [34,41,22]. 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 [22], 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 [65], [86], and [34]

$$\operatorname{div}(k(x)\nabla p) = 0, \quad (3.31)$$

$$\frac{\partial S}{\partial t} + \mathbf{v} \cdot \nabla S = 0, \quad (3.32)$$

where  $p$  is the pressure,  $S$  is the water saturation,  $k(x) = (k_{ij}(x))$  is the relative permeability tensor, and  $\mathbf{v} = -k(x)\nabla p$  is the Darcy velocity. The highly heterogeneous properties of the medium are built into the permeability tensor  $k(x)$  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 [68]. 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 (3.32) to get a complete coarse grid algorithm for the problem (3.31)-(3.32). The numerical upscaling of the saturation equation has been under intensive study in the literature [30,41,65,47,89,87]. Here, we use the upscaling method proposed in [41] and [34] to design an overall coarse grid model for the problem (3.31)-(3.32). The work of [41] 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 (3.32), 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 [41]:

$$\frac{\partial S}{\partial t} + \mathbf{v}_0 \cdot \nabla S = \frac{\partial}{\partial x_i} \left( D_{ij}(x, t) \frac{\partial S}{\partial x_j} \right), \quad (3.33)$$

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

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

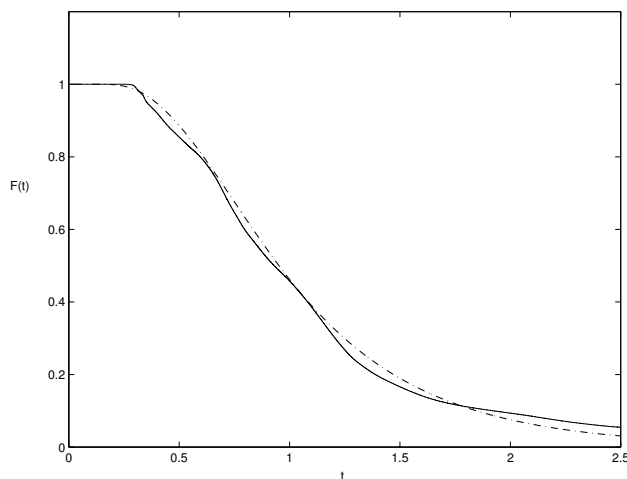
Note that the hyperbolic equation (3.32) is now replaced by a convection-diffusion equation. The convection-dominant parabolic equation (3.33) is solved by the characteristic linear finite element method [28], [75] in our simulation. The flow transport model (3.31)-(3.32) is solved in the coarse grid as follows:

1. Solve the pressure equation (3.31) 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 (3.33) 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.

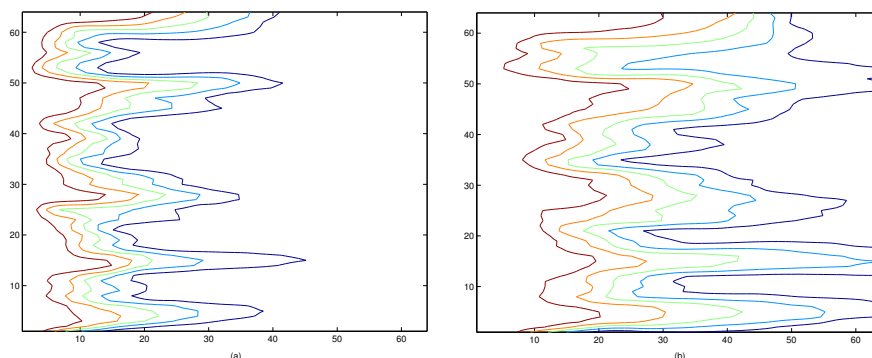
The mixed multiscale finite element method can be readily combined with the above upscaling model for the saturation equation. The local fine grid velocity  $\mathbf{v}'$  will be constructed from the multiscale finite element base functions. The main cost in the above algorithm lies in the computation 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  $64 \times 64$  mesh. The fractional flow curve using the above algorithm is depicted in Figure 3.8. It gives excellent agreement with the ‘‘exact’’ fractional flow curve. The contour plots of the saturation  $S$  on the fine  $1024 \times 1024$  mesh at time  $t = 0.25$  and  $t = 0.5$  computed by using the ‘‘exact’’ velocity field are displayed in Figure 3.10. In Figure 3.9, 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 3.9 approximate the ‘‘exact’’ ones in

Figure 3.10 in certain accuracy but the sharp oil/water interfaces in Figure 3.10 are smeared out. This is due to the parabolic nature of the upscaled equation (3.33). 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.



**Fig. 3.8.** 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.

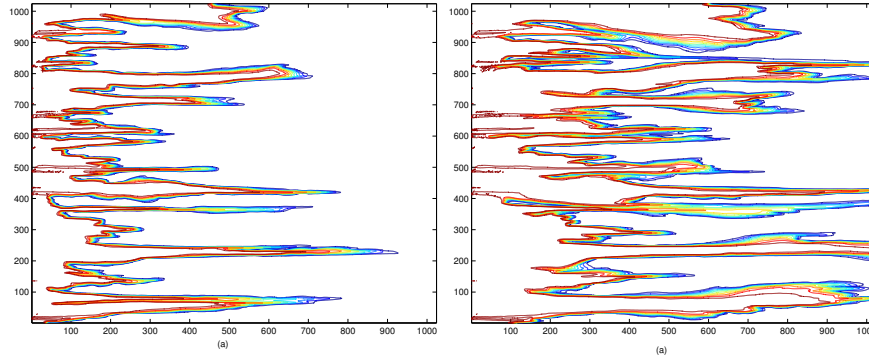


**Fig. 3.9.** The contour plots of the saturation  $S$  computed using the upscaled model on a  $64 \times 64$  mesh at time  $t = 0.25$  (left) and  $t = 0.5$  (right).

Finally, we remark that the upscaling equation (3.33) 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

$$\operatorname{div}(k^*(x)\nabla p^*) = 0$$

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 (3.31) on the fine mesh. From the fine grid computation, one can easily construct the



**Fig. 3.10.** The contour plots of the saturation  $S$  computed on the fine  $1024 \times 1024$  mesh using the “exact” velocity field at time  $t = 0.25$  (left) and  $t = 0.5$  (right).

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 (3.31) 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 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 [30,89]. 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.

### 3.7 Brief overview of mixed finite element and finite volume element methods

**Control volume multiscale finite element method** In this section, we discuss multiscale finite volume element method. Finite volume method is chosen because, by its construction, it satisfies the numerical local conservation which is important in groundwater and reservoir simulations. Let  $\mathcal{K}^h$  denote the collection of coarse elements/rectangles  $K$ . Consider a coarse element  $K$ , and let  $\xi_K$  be its center. The element  $K$  is divided into four rectangles of equal area by connecting  $\xi_K$  to the midpoints of the element’s edges. We denote these quadrilaterals by  $K_\xi$ , where  $\xi \in Z_h(K)$ , are the vertices of  $K$ . Also, we denote  $Z_h = \bigcup_K Z_h(K)$  and  $Z_h^0 \subset Z_h$  the vertices which do not lie on the Dirichlet boundary of  $\Omega$ . The control volume  $V_\xi$  is defined as the union of the quadrilaterals  $K_\xi$  sharing the vertex  $\xi$ .

The key idea of the method is the construction of basis functions on the coarse grids, such that these basis functions capture the small-scale information on each of these coarse grids. As before, the basis functions are constructed from the solution of the leading order homogeneous elliptic equation on each coarse element with some specified boundary conditions. We consider a coarse element  $K$  that has  $d$  vertices, the local basis functions  $\phi_i, i = 1, \dots, d$  are set to satisfy the following elliptic problem:

$$\begin{aligned} -\nabla \cdot (k \cdot \nabla \phi_i) &= 0 \quad \text{in } K \\ \phi_i &= g_i \quad \text{on } \partial K, \end{aligned} \tag{3.34}$$

for some function  $g_i$  defined on the boundary of the coarse element  $K$ . As we discussed earlier, Hou et al. [50] have demonstrated that a careful choice of boundary conditions would improve the accuracy of the method. In previous findings, the function  $g_i$  for each  $i$  is chosen to vary linearly along  $\partial K$  or to be the solution of the local one-dimensional problems [61] or the solution of the problem in a slightly larger domain is chosen to define the boundary conditions. For simplicity, we consider linear boundary conditions and also discuss the boundary conditions obtained from a global solution. We will require  $\phi_i(x_j) = \delta_{ij}$ . Finally, a nodal basis

function associated with the vertex  $x_i$  in the domain  $\Omega$  is constructed from the combination of the local basis functions that share this  $x_i$  and zero elsewhere. We would like to note that one can use an approximate solution of (3.34) when it is possible. For example, in the case of periodic or random homogeneous cases, the basis functions can be approximated using homogenization expansion  $\phi_i = \phi_i^0 + \epsilon \chi_k \nabla_k \phi_i^0$ , where  $\chi_k$  is the solution of the cell problem and  $\phi_i^0$  is standard finite element basis on the coarse mesh (see [36]).

Next, we denote by  $V^h$  the space of our approximate pressure solution, which is spanned by the basis functions  $\{\phi_j\}_{x_j \in Z_h^0}$ . Then we formulate the finite dimensional problem corresponding to finite volume element formulation of pressure equation. A statement of mass conservation on a coarse-control volume  $V_x$  is formed from pressure equation, where the approximate solution is written as a linear combination of the basis functions. Assembly of this conservation statement for all control volumes would give the corresponding linear system of equations that can be solved accordingly. The resulting linear system has incorporated the fine-scale information through the involvement of the nodal basis functions on the approximate solution. To be specific, the problem now is to seek  $p^h \in V^h$  with  $p^h = \sum_{x_j \in Z_h^0} p_j \phi_j$  such that

$$\int_{\partial V_\xi} k \cdot \nabla p^h \cdot n \, dl = 0, \quad (3.35)$$

for every control volume  $V_\xi \subset \Omega$ . Here  $n$  defines the normal vector on the boundary of the control volume,  $\partial V_\xi$  and  $S$  is the fine-scale saturation field at this point. The resulting multiscale method differs from the multiscale finite element method, since it employs the finite volume element method as a global solver, and it is called multiscale finite volume element method (MsFVEM). We would like to note that the coarse-scale velocity field obtained using MsFVEM is conservative in control volume elements  $V_\xi$  (not in  $\mathcal{K}^h$ ).

**Mixed multiscale finite element methods** For simplicity, we assume Neumann boundary conditions. First, we review the mixed multiscale finite element formulation following [22] (see also [6], [1], and [7]). We can rewrite two-phase flow equation as

$$\begin{aligned} k^{-1}u - \nabla p &= 0 & \text{in } \Omega \\ \operatorname{div}(u) &= 0 & \text{in } \Omega \\ k(x)\nabla p \cdot n &= g(x) & \text{on } \partial\Omega. \end{aligned} \quad (3.36)$$

The variational problem associated with (3.36) is to seek  $(u, p) \in H(\operatorname{div}, \Omega) \times L^2(\Omega)/R$  such that  $u \cdot n = g$  on  $\partial\Omega$  and

$$\begin{aligned} (k^{-1}u, v) + (\operatorname{div}v, p) &= 0 \quad \forall v \in H_0(\operatorname{div}, \Omega) \\ (\operatorname{div}u, q) &= 0 \quad \forall q \in L^2(\Omega)/R. \end{aligned} \quad (3.37)$$

where  $H_0(\operatorname{div}, \Omega)$  is  $H(\operatorname{div}, \Omega)$  with homogeneous Neumann boundary conditions. By defining

$$a(u, v) = (k^{-1}u, v), \quad b(v, q) = (\operatorname{div}v, q), \quad (3.38)$$

we can rewrite the weak formulation as

$$\begin{aligned} a(u, v) + b(v, p) &= 0 \quad \forall v \in H_0(\operatorname{div}, \Omega), \\ b(u, q) &= 0 \quad \forall q \in L^2(\Omega)/R. \end{aligned}$$

Let  $V_h \subset H(\operatorname{div}, \Omega)$  and  $Q_h \subset L^2(\Omega)/R$  be finite dimensional spaces and  $V_h^0 = V_h \cap H_0(\operatorname{div}, \Omega)$ . The numerical approximation problem associated with (3.37) is to find  $(u_h, p_h) \in V_h \times Q_h$  such that  $u_h \cdot n = g_h$  on  $\partial\Omega$ , where  $g_h = g_{0,h}n$  on  $\partial\Omega$  and  $g_{0,h} = \sum_{e \in \{\partial K \cap \partial\Omega, K \in \tau_h\}} (\int_e g ds) N_e$ ,  $N_e \in V_h$ , is corresponding basis function to edge  $e$ ,

$$\begin{aligned} (k^{-1}u_h, v_h) + (\operatorname{div}v_h, p_h) &= 0 \quad \forall v_h \in V_h^0 \\ (\operatorname{div}u_h, q_h) &= 0 \quad \forall q_h \in Q_h. \end{aligned} \quad (3.39)$$



One can define a linear operator  $B_h : V_h^0 \rightarrow Q'_h$  by  $b(u_h, q_h) = (B_h u_h, q_h)$ .

Suppose that the following conditions are satisfied

$$a(u_h, u_h) \text{ is } \ker B_h - \text{coercive} \quad (3.40)$$

$$\inf_{q_h \in Q_h} \sup_{v_h \in V_h} \frac{b(v_h, q_h)}{\|v_h\|_{H(\text{div}, \Omega)} \|q_h\|_{L^2(\Omega)}} \geq C. \quad (3.41)$$

Then the following approximation property follows (see e.g., [18]).

*Lemma 31* If  $(u, p)$  and  $(u_h, p_h)$  respectively solve the problem (3.37) and (3.39) and the conditions (3.40) and (3.41) hold, then

$$\|u - u_h\|_{H(\text{div}, \Omega)} + \|p - p_h\|_{0, \Omega} \leq \inf_{\substack{v_h \in V_h \\ v_h - g_{0,h} \in V_h^0}} \|u - v_h\|_{H(\text{div}, \Omega)} + \inf_{q_h \in Q_h} \|p - q_h\|_{0, \Omega}. \quad (3.42)$$

Following Chen and Hou [22] (see also [6]), one can construct multiscale basis functions for velocity in each coarse block  $K$

$$\begin{aligned} \text{div}(k(x)\nabla w_i^K) &= \frac{1}{|K|} \text{ in } K \\ k(x)\nabla w_i^K n^K &= \begin{cases} g_i^K & \text{on } e_i^K \\ 0 & \text{else,} \end{cases} \end{aligned} \quad (3.43)$$

where  $g_i^K = \frac{1}{|e_i^K|}$  and  $e_i^K$  are the edges of  $K$ . Then, we can define the finite dimensional space for velocity by

$$\begin{aligned} V_h &= \bigoplus_K \{\Psi_i^K\}, \\ V_h^0 &= V_h \cap H_0(\text{div}, \Omega), \end{aligned}$$

where  $\Psi_i^K = k(x)\nabla w_i^K$ .

### 3.8 MsFEM using limited global information

**Motivation** Multiscale finite element methods and their modifications are used in two-phase flow simulations through heterogeneous porous media. First, we briefly describe the underlying fine-scale equations. We present two-phase flow equations neglecting the effects of gravity, compressibility, capillary pressure and dispersion on the fine scale. Porosity, defined as the volume fraction of the void space, will be taken to be constant and therefore serves only to rescale time. The two phases will be referred to as water and oil and designated by the subscripts  $w$  and  $o$ , respectively. We can then write Darcy's law, with all quantities dimensionless, for each phase  $j$  as follows:

$$\mathbf{v}_j = -\lambda_j(S)k\nabla p, \quad (3.44)$$

where  $\mathbf{v}_j$  is phase velocity,  $S$  is water saturation (volume fraction),  $p$  is pressure,  $\lambda_j = k_{r,j}(S)/\mu_j$  is phase mobility, where  $k_{r,j}$  and  $\mu_j$  are the relative permeability and viscosity of phase  $j$  respectively, and  $\mathbf{k}$  is the permeability tensor, which is here taken to be diagonal,  $\mathbf{k} = kI$ , where  $I$  is the identity matrix,

Combining Darcy's law with conservation of mass,  $\text{div}(\mathbf{v}_w + \mathbf{v}_o) = 0$ , allows us to write the flow equation in the following form

$$\text{div}(\lambda(S)k\nabla p) = f, \quad (3.45)$$

where the total mobility  $\lambda(S)$  is given by  $\lambda(S) = \lambda_w(S) + \lambda_o(S)$  and  $f$  is a source term. The saturation dynamics affects the flow equations. One can derive the equation describing the dynamics of the saturation

$$\frac{\partial S}{\partial t} + \text{div}(\mathbf{F}) = 0, \quad (3.46)$$

where  $\mathbf{F} = \mathbf{v}f_w(S)$ , with  $f_w(S)$ , the fractional flow of water, given by  $f_w = \lambda_w/(\lambda_w + \lambda_o)$ , and the total velocity  $\mathbf{v}$  by:

$$\mathbf{v} = \mathbf{v}_w + \mathbf{v}_o = -\lambda(S)k\nabla p. \quad (3.47)$$

In the presence of capillary effects, an additional diffusion term is present in (3.46).

If  $k_{rw} = S$ ,  $k_{ro} = 1 - S$  and  $\mu_w = \mu_o$ , then the flow equation reduces to

$$\text{div}(k\nabla p^{sp}) = f.$$

This equation, the linear advection pollutant transport equation, will be referred to as the single-phase flow equation associated with (3.45), and  $p^{sp}$  will be referred to as the single-phase flow solution.

As we see from (3.45) and (3.46), the pressure equation is solved many times for different saturation profiles. Thus, computing the basis functions once at time zero is very beneficial and the basis functions are only updated near sharp interfaces. In fact, our numerical results show that only slight improvement can be achieved by updating the basis functions near sharp fronts. However, we have found that for heterogeneous permeability fields with very strong non-local effects, the use of some type of global information can improve multiscale finite element results significantly, which will be discussed next.

We present a representative numerical example for a permeability field generated using two-point geostatistics. To generate this permeability field, we have used GSLIB algorithm [26]. The permeability is log-normally distributed with prescribed variance  $\sigma^2 = 1.5$  ( $\sigma^2$  here refers to the variance of  $\log k$ ) and some correlation structure. The correlation structure is specified in terms of dimensionless correlation lengths in the  $x$  and  $z$ -directions,  $l_x = 0.4$  and  $l_z = 0.04$ , nondimensionalized by the system length. Linear boundary conditions are used for constructing multiscale basis function in (3.34). Spherical variogram is used [26]. In this numerical example, the fine-scale field is  $120 \times 120$ , while the coarse-scale field is  $12 \times 12$  defined in the rectangle with the length 5 and the width 1. For the two-phase flow simulations, the system is considered to initially contain only oil ( $S = 0$ ) and water is injected at inflow boundaries ( $S = 1$  is prescribed), i.e., we specify  $p = 1$ ,  $S = 1$  along the  $x = 0$  edge and  $p = 0$  along the  $x = 5$  edge, and no flow boundary conditions on the lateral boundaries. Relative permeability functions are specified as  $k_{rw} = S^2$ ,  $k_{ro} = (1 - S)^2$ ; water and oil viscosities are set to  $\mu_w = 1$  and  $\mu_o = 5$ . Porosity is constant and serves only to nondimensionalize time. Results are presented in terms of the fraction of oil in the produced fluid (i.e., oil cut, designated  $F$ ) against pore volume injected (PVI). PVI represents dimensionless time and is computed via  $\int Q dt / V_p$  where  $V_p$  is the total pore volume of the system and  $Q$  is the total flow rate.

In our first numerical test, Figure 3.11, we compare the fractional flows. The dashed line corresponds to the calculations performed using a simple saturation upscaling (no subgrid treatment), while dotted line corresponds to the calculations performed by solving the saturation equation on the fine grid using the reconstructed fine-scale velocity field. We observe from this figure that the second approach is very accurate, while the first approach over-predicts the breakthrough time. The saturation snapshots are compared in Figure 3.12. One can observe that there is a very good agreement.

In the next set of numerical results, we consider strongly channelized permeability fields. These permeability fields are proposed in some recent benchmark tests, such as the SPE comparative solution project [23]. In Figure 3.13, one of the layers of this 3-D permeability field is depicted. All the layers have  $220 \times 60$  fine-scale resolution, and we take the coarse grid to be  $22 \times 6$ . As it can be observed, the permeability field contains a high permeability channel, where most flow will occur in our simulation. In Figure 3.14, the fractional flows are compared. The boundary conditions are taken to be  $p = 1$ ,  $S = 1$  along the  $x = 0$  edge and  $p = 0$  along the  $x = 5$  edge, and no flow boundary conditions on the lateral boundaries. Again, the dashed line corresponds to the calculations performed using a simple saturation upscaling (no subgrid treatment), while dotted line corresponds to the calculations performed by solving the saturation equation on the fine grid using the reconstructed fine-scale velocity field. We observe from this figure that the second approach is not very accurate in contrast to the permeability field generated using two-point geostatistics [26]. This is because the local basis functions can not account accurately the global connectivity of the media. Indeed, in the next figure, Figure 3.15, the saturation fields at time PVI = 0.5 are compared. We see that multiscale finite element methods with local basis functions introduce some errors. In the bottom left corner, there is a saturation pocket which is not in the reference solution computed using a fine grid. The reason for this is

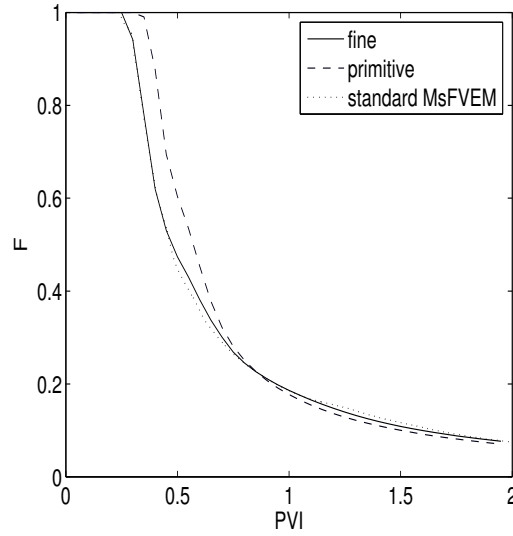


Fig. 3.11. Fractional flow comparison for a permeability field generated using two-point geostatistics.

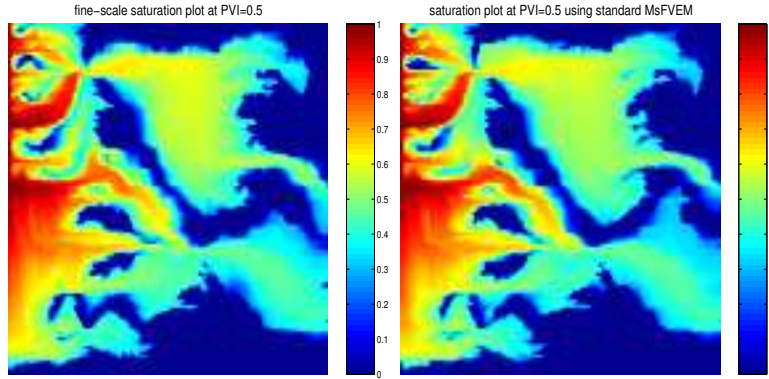
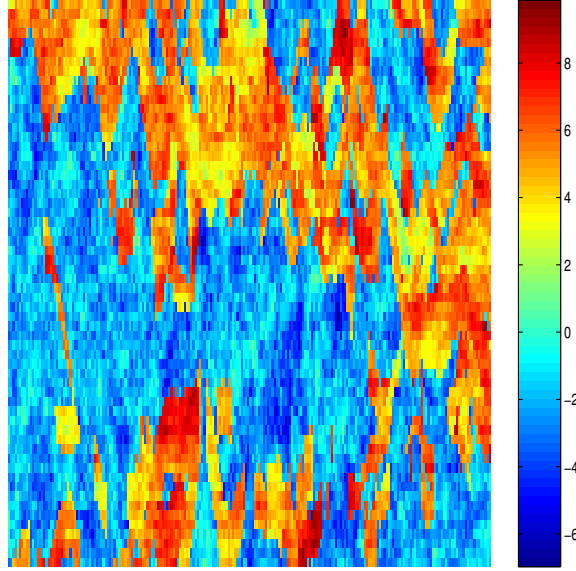


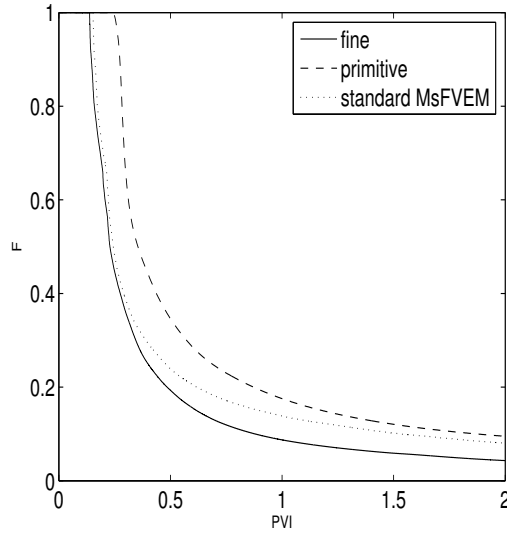
Fig. 3.12. Saturation maps at PVI=0.5 for fine-scale solution (left figure) and standard MsFVEM (right figure)

that the local basis functions in the lower left corner contains high permeability region. However, this high permeability region does not have global connectivity, and the local basis functions can not take this effect into account. Next, we discuss how global information can be incorporated into multiscale basis functions to improve the accuracy of the computations.

**Multiscale finite volume element method** The main idea of the modified multiscale finite volume element method (MsFVEM) is to use the solution of the fine-scale problem at time zero to determine the boundary conditions for the basis functions. This approach is proposed in [35] to handle the permeability fields which are strongly channelized. For this type of permeability fields, some type of global information is needed. Next, we describe the method. We denote the solution of pressure equation at time zero by  $p^{sp}(x)$ . In defining  $p^{sp}(x)$ , we use the actual boundary conditions of the global problem.  $p^{sp}(x)$  depends on global boundary conditions, and, generally, is updated each time when global boundary conditions are changed. The boundary conditions in (3.34) for modified basis functions are defined in the following way. For each rectangular element  $K$  with vertices  $x_i$  ( $i = 1, 2, 3, 4$ ) denote by  $\phi_i(x)$  a restriction of the nodal basis on  $K$ , such that  $\phi_i(x_j) = \delta_{ij}$ . At the edges where  $\phi_i(x) = 0$  at both vertices, we take boundary condition for  $\phi_i(x)$  to be zero. Consequently, the basis functions are localized. We only need to determine the boundary



**Fig. 3.13.** Log-permeability for one of the layers of upper Ness.



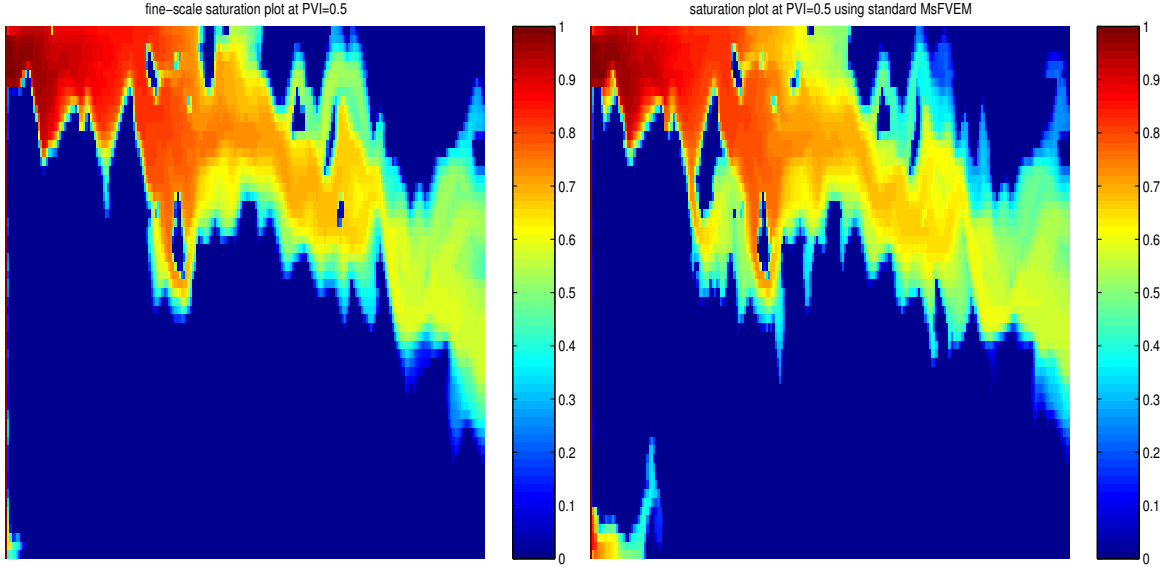
**Fig. 3.14.** Fractional flow comparison for a channelized permeability field.

condition at two edges which have the common vertex  $x_i$  ( $\phi_i(x_i) = 1$ ). Denote these two edges by  $[x_{i-1}, x_i]$  and  $[x_i, x_{i+1}]$  (see Figure 3.16). We only need to describe the boundary condition,  $g_i(x)$ , for the basis function  $\phi_i(x)$ , along the edges  $[x_i, x_{i+1}]$  and  $[x_i, x_{i-1}]$ . If  $p^{sp}(x_i) \neq p^{sp}(x_{i+1})$ , then

$$g_i(x)|_{[x_i, x_{i+1}]} = \frac{p^{sp}(x) - p^{sp}(x_{i+1})}{p^{sp}(x_i) - p^{sp}(x_{i+1})}, \quad g_i(x)|_{[x_i, x_{i-1}]} = \frac{p^{sp}(x) - p^{sp}(x_{i-1})}{p^{sp}(x_i) - p^{sp}(x_{i-1})}.$$

If  $p^{sp}(x_i) = p^{sp}(x_{i+1}) \neq 0$ , then

$$g_i(x)|_{[x_i, x_{i+1}]} = \phi_i^0(x) + \frac{1}{2p^{sp}(x_i)}(p^{sp}(x) - p^{sp}(x_{i+1})),$$



**Fig. 3.15.** Saturation maps at PVI=0.5 for fine-scale solution (left figure) and standard MsFVEM (right figure).

where  $\phi_i^0(x)$  is a linear function on  $[x_i, x_{i+1}]$  such that  $\phi_i^0(x_i) = 1$  and  $\phi_i^0(x_{i+1}) = 0$ . Similarly,

$$g_{i+1}(x)|_{[x_i, x_{i+1}]} = \phi_{i+1}^0(x) + \frac{1}{2p^{sp}(x_{i+1})}(p^{sp}(x) - p^{sp}(x_{i+1})), \quad (3.48)$$

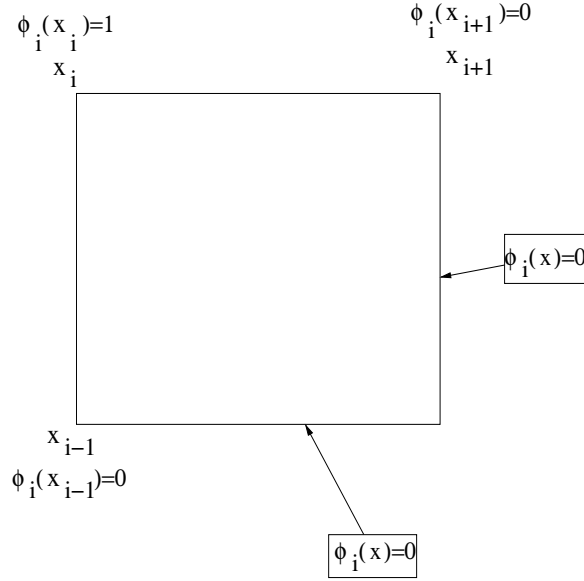
where  $\phi_{i+1}^0(x)$  is a linear function on  $[x_i, x_{i+1}]$  such that  $\phi_{i+1}^0(x_{i+1}) = 1$  and  $\phi_{i+1}^0(x_i) = 0$ . If  $p^{sp}(x_i) = p^{sp}(x_{i+1}) \neq 0$ , then one can also use simply linear boundary conditions. If  $p^{sp}(x_i) = p^{sp}(x_{i+1}) = 0$  then linear boundary conditions are used. In the applications considered in this paper, the initial pressure is always positive. Finally, the basis function  $\phi_i(x)$  is constructed by solving (3.34). The choice of the boundary conditions for the basis functions is motivated by the analysis. In particular, we would like to recover the exact fine-scale solution along each edge if the nodal values of the pressure are equal to the values of exact fine-scale pressure. This is the underlying idea for the choice of boundary conditions. Using this property and Cea's lemma one can show that the pressure obtained from the numerical solution is equal to the underlying fine-scale pressure.

**Mixed multiscale finite element methods** Next, following [1], we present a mixed multiscale finite element method that employs single-phase flow information. Suppose that  $p^{sp}$  solves the single-phase flow equation. We set  $b_i^K = (k \nabla p^{sp}|_{e_i^K}) \cdot n^K$  and assume that  $b_i^K$  is uniformly bounded. Then the new basis functions for velocity is constructed by solving the following local problems (3.43) with  $g_i^K = b_i^K / \beta_i^K$ , where  $\beta_i^K = \int_{e_i^K} k \nabla p^{sp} \cdot n^K ds$ . For further analysis, we assume that  $\beta_i^K \neq 0$ . In general, if  $\beta_i^K = 0$  one can use standard mixed multiscale finite element basis functions. Let  $N_i^K = k(x) \nabla w_i^K$  and the multiscale finite dimensional space  $V_h^0$  for velocity be defined by

$$V_h := \bigoplus_K \{N_i^K\} \subset H(\text{div}, \Omega),$$

$$V_h^0 := V_h \cap H_0(\text{div}, \Omega).$$

First, we will show that the resulting multiscale finite element solution for velocity is exact for single-phase flow (i.e.,  $\lambda(x) = 1$ ). Let  $v_h|_K = \beta_i^K N_i^K$ , then  $\beta_i^K$  is the interpolation value of the fine scale solution.



**Fig. 3.16.** Schematic description of nodal points.

Furthermore, a direct calculation yields  $(v_h|_{e_i^K}) \cdot n^K = k \nabla p^{sp} \cdot n^K$ . Because

$$\text{div} v_h = \beta_i^K \text{div} N_i^K = \frac{1}{|K|} \int_{\partial K} k \nabla p^{sp} \cdot n^K ds = \frac{1}{|K|} \int_K \text{div}(k \nabla p^{sp} d) = 0,$$

the following equation is obtained immediately

$$\text{div} v_h = 0 \quad \text{in } K \tag{3.49}$$

$$v_h \cdot n^K = k \nabla p^{sp} \cdot n^K \quad \text{on } \partial K \tag{3.50}$$

Because  $\text{div}(k \nabla p^{sp}) = 0$ , we get  $v_h = k \nabla p^{sp}$  and the following proposition.

**Proposition** Let  $\beta_i^K = \int_{e_i^K} k \nabla p^{sp} \cdot n^K ds$ , then on each coarse block  $K$

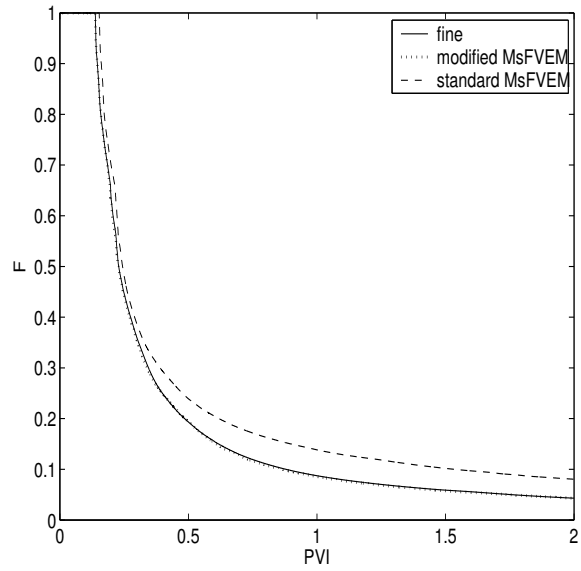
$$k \nabla p^{sp} = \beta_i^K N_i^K. \tag{3.51}$$

*Lemma 32* If  $|\beta_i^K| \geq Ch$  with  $C$  is independent of  $h$ , then

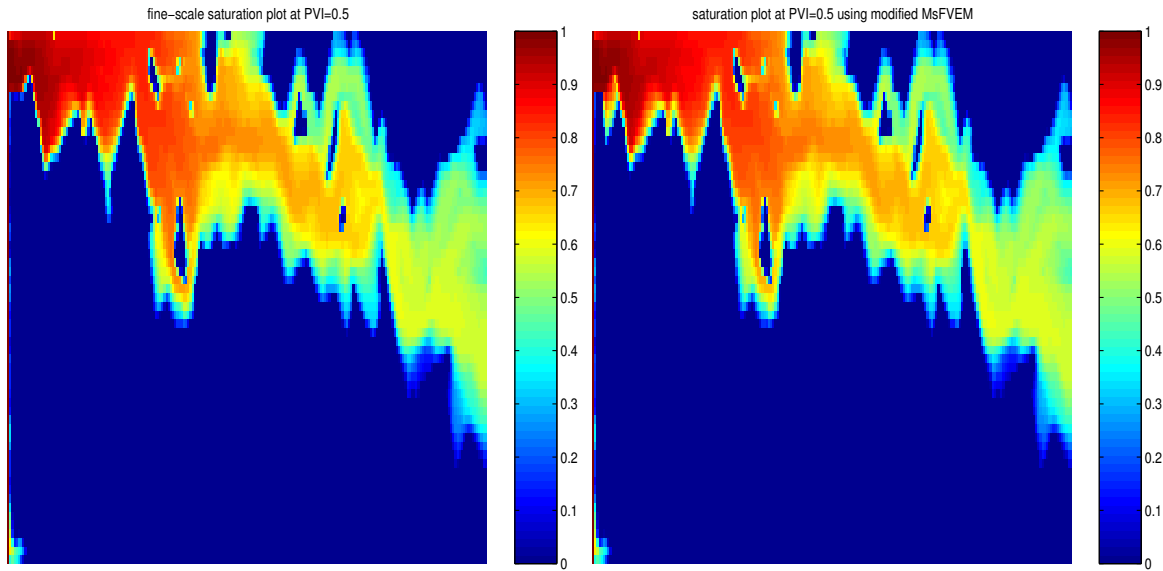
- (1)  $a(u_h, u_h)$  is  $\ker B_h$ -coercive;
- (2)  $\inf_{q_h \in Q_h} \sup_{v_h \in V_h^0} \frac{b(v_h, q_h)}{\|v_h\|_{H(\text{div}, \Omega)} \|q_h\|_{L^2(\Omega)}} \geq C$ .

**Numerical results** Next, we show the numerical results obtained using modified multiscale finite element type methods for the permeability layer depicted in Figure 3.13 and two-phase flow parameters presented earlier. We consider two types of boundary conditions in a rectangular region  $[0, 5] \times [0, 1]$ . For the first type of boundary conditions, we specify  $p = 1, S = 1$  along the  $x = 0$  edge and  $p = 0$  along the  $x = 5$  edge. On the rest of the boundaries, we assume no flow boundary condition. We call this type of the boundary condition as side-to-side. The other type of boundary conditions is obtained by specifying  $p = 1, S = 1$  along the  $x = 0$  edge for  $0.5 \leq z \leq 1$  and  $p = 0$  along the  $x = 5$  edge for  $0 \leq z \leq 0.5$ . On the rest of the boundaries, we assume no flow boundary condition.

In Figure 3.17, the fractional flows are plotted for standard and modified MsFVEM. We observe from this figure that modified MsFVEM is more accurate and provides nearly the same fractional flow response as the direct fine-scale calculations. In Figure 3.18, we compare the saturation fields at PVI=0.5. As we see, the saturation field obtained using modified MsFVEM is very accurate and there is no longer the saturation



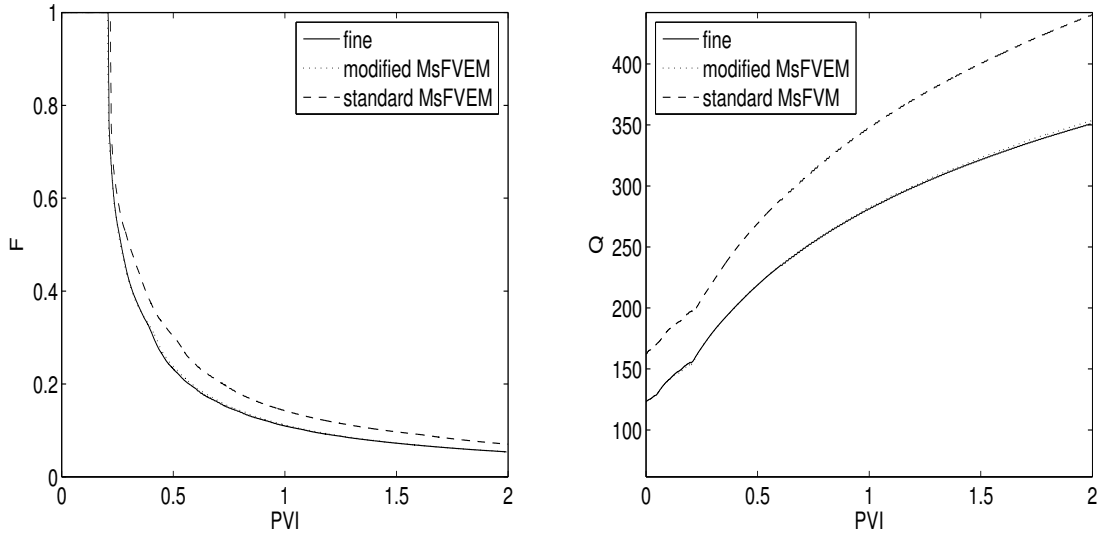
**Fig. 3.17.** Fractional flow comparison for standard MsFVEM and modified MsFVEM for side-to-side flow.



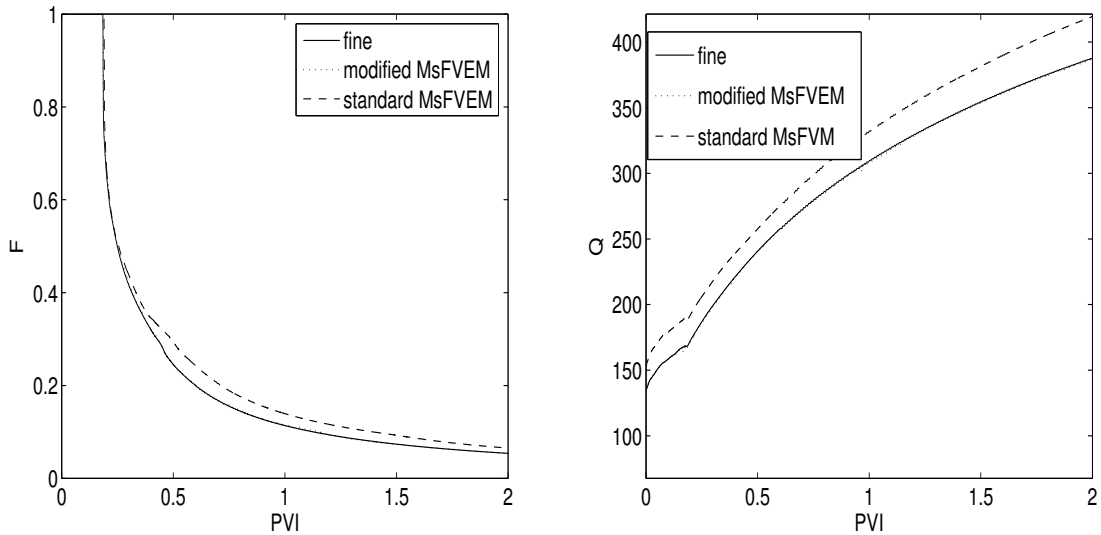
**Fig. 3.18.** Saturation maps at PVI=0.5 for fine-scale solution (left figure) and modified MsFVEM (right figure). Side-to-side boundary condition is used.

pocket at the left bottom corner. Thus, the modified MsFVEM captures the connectivity of the media accurately.

In the next set of numerical results, we test the modified multiscale finite element methods for a different layer (layer 40) of SPE comparative solution project. In Figure 3.19 and 3.20, the fractional flows and total flow rates ( $Q$ ) are compared for two different boundary conditions. One can see clearly that the modified MsFVEM method gives nearly exact results for these integrated responses. The standard MsFVEM tends to over-predict the total flow rate at time zero. This initial error persists at later times. This phenomena is often observed in upscaling of two-phase flows. More numerical results and discussions can be found in [35]. These numerical results demonstrate that modified multiscale finite element methods which use a limited



**Fig. 3.19.** Fractional flow (left figure) and total production (right figure) comparison for standard MsFVEM and modified MsFVEM for side-to-side flow (layer 40).

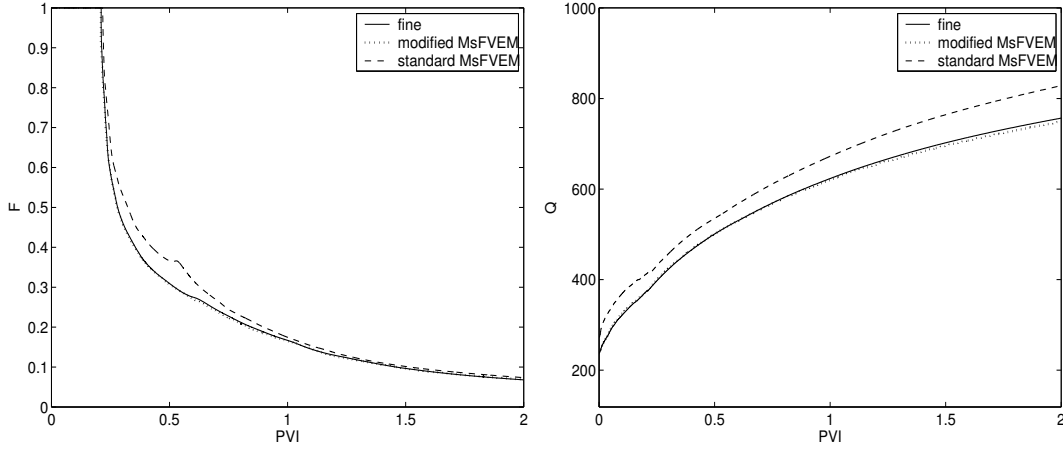


**Fig. 3.20.** Fractional flow (left figure) and total production (right figure) comparison for standard MsFVEM and modified MsFVEM for corner-to-corner (layer 40).

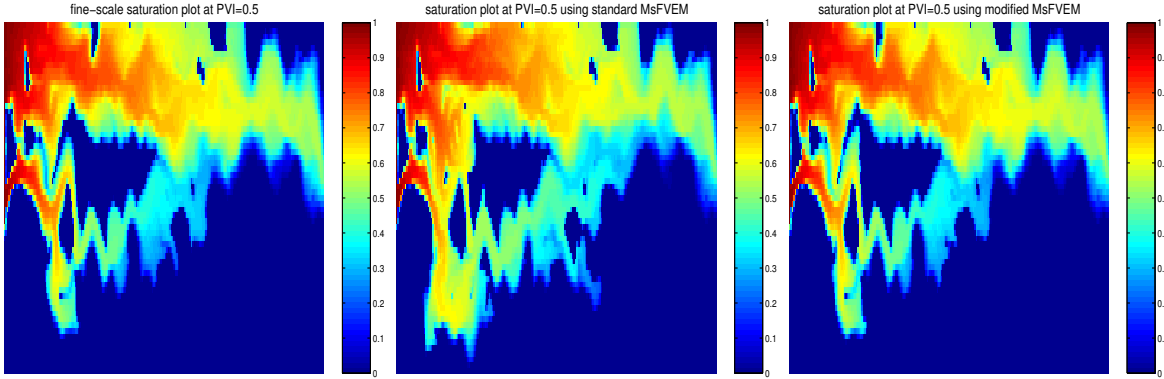
global information are more accurate. Moreover, modified multiscale finite element methods are capable of capturing the long-range flow features accurately for channelized permeability fields.

For the next set of results, we consider another layer of the upper Ness (layer 59). In Figure 3.21, both fractional flow (left figure) and total flow (right figure) are plotted. We observe that the modified MsFVEM gives almost the exact results for these quantities, while the standard MsFVEM overpredicts the total flow rate, and there are deviations in the fractional flow curve around  $PVI \approx 0.6$ . Note that unlike the previous case, fractional flow for standard MsFVEM is nearly exact at later times ( $PVI \approx 2$ ). In Figure 3.22, the saturation maps are plotted at  $PVI = 0.5$ . The left figure represents the fine-scale, the middle figure





**Fig. 3.21.** Fractional flow (left figure) and total production (right figure) comparison for standard MsFVEM and modified MsFVEM for corner-to-corner flow.



**Fig. 3.22.** Saturation maps at  $PVI=0.5$  for fine-scale solution (left figure), standard MsFVEM (middle figure), and modified MsFVEM (right figure). Corner-to-corner boundary condition is used.

represents the results obtained using standard MsFVEM, and the right figure represents the results obtained using the modified MsFVEM. We observe from this figure that the saturation map obtained using standard MsFVEM has some errors. These errors are more evident near the lower left corner. The results of the saturation map obtained using the modified MsFVEM is nearly the same as the fine-scale saturation field. It is evident from these figures that the modified MsFVEM performs better than the standard MsFVEM.

### Analysis

**Galerkin finite element methods with limited global information** We have proposed some analysis for modified multiscale finite element method in [35] and [2]. The main idea is to show that the pressure evolution in two-phase flow simulations is strongly influenced by the initial pressure. To demonstrate this, we consider a channelized permeability field, where the value of the permeability in the channel is large. We assume the permeability has the form  $kI$ , where  $I$  is an identity matrix. In a channelized medium, the dominant flow is within the channels. Our analysis assumes a single channel and restricted to 2-D. Here, we briefly mention the main findings. Denote the initial stream function and pressure by  $\eta = \psi(x, t = 0)$  and  $\zeta = p(x, t = 0)$  ( $\zeta$  is also denoted by  $p^{sp}$  previously). The stream function is defined

$$\partial\psi/\partial x_1 = -v_2, \quad \partial\psi/\partial x_2 = v_1. \tag{3.52}$$

Then the equation for the pressure can be written as

$$\frac{\partial}{\partial \eta} \left( |k|^2 \lambda(S) \frac{\partial p}{\partial \eta} \right) + \frac{\partial}{\partial \zeta} \left( \lambda(S) \frac{\partial p}{\partial \zeta} \right) = 0. \quad (3.53)$$

For simplicity,  $S = 0$  at time zero is assumed. We consider a typical boundary condition that gives high flow within the channel, such that the high flow channel will be mapped into a large slab in  $(\eta, \zeta)$  coordinate system. If the heterogeneities within the channel in  $\eta$  direction is not strong (e.g., narrow channel in Cartesian coordinates), the saturation within the channel will depend on  $\zeta$ . In this case, the leading order pressure will depend only on  $\zeta$ , and it can be shown that

$$p(\eta, \zeta, t) = p_0(\zeta, t) + \text{high order terms}, \quad (3.54)$$

where  $p_0(\zeta, t)$  is the dominant pressure. This asymptotic expansion shows that the time-varying pressure strongly depends on the initial pressure (i.e., the leading order term in the asymptotic expansion is a function of initial pressure and time only). In our analysis, we will assume that  $|p(x, t) - \hat{p}(p^{sp}, t)|_{H^1}$  is small.

Since the analysis of the multiscale finite element methods is carried out only for the pressure equation, we will assume  $t$  (time) is fixed. Then, assuming the function  $\hat{p}$  is sufficiently smooth, one can state the following. There exists  $A_K$  in each  $K$ , such that  $\|\nabla p(x) - A_K \nabla p^{sp}(x)\|_{L^2(\Omega)}$  is small. Note that this assumption indicates that the fine-scale features of pressure solutions of two-phase equations does not change significantly during a simulation (e.g., streamlines do not vary significantly in each coarse block). This phenomena can be observed in numerical simulations of two-phase flows when  $\mu_o/\mu_w > 1$ .

The assumption for the case with scale separation indicates that the coarse-scale features of two-phase flow and single-phase flow are similar (e.g., coarse-scale streamlines do not vary significantly). We will use the following assumption.

*Assumption G.* There exists a sufficiently smooth scalar valued function  $G(\eta)$  ( $G \in C^3$ ), such that

$$|p - G(p^{sp})|_{1,\Omega} \leq C\delta, \quad (3.55)$$

where  $p^{sp}$  is single-phase flow pressure and  $\delta$  is sufficiently small.

We note  $G$  is  $p_0(\zeta, t)$  at fixed  $t$  in (3.54). Moreover, one does not need to know the function  $G$  for computing the multiscale approximation of the solution. It is only necessary that  $G$  has certain smoothness properties, however, it is important that the basis functions span  $p^{sp}$  in each coarse block.

*Theorem 33* Under Assumption G and  $p^{sp} \in W^{1,s}(\Omega)$  ( $s > 2$ ), multiscale finite element method converges with the rate given by

$$|p - p_h|_{1,\Omega} \leq C\delta + Ch^{1-2/s} |p^{sp}|_{W^{1,s}(\Omega)} + Ch^{1-2/s} |p^{sp}|_{1,\Omega} + Ch \|f\|_{0,\Omega} \leq C\delta + Ch^{1-2/s}. \quad (3.56)$$

The proof of this theorem is given in [2]. Note that Theorem(33) shows that MsFEM converges for problems without any scale separation and the proof of this theorem does not use homogenization techniques. Next, we present the proof.

*Proof.* Following standard practice of finite element estimation, we seek  $q_h = c_i \phi_i^K$ , where  $\phi_i^K$  are single-phase flow based multiscale finite element basis functions. Then from Cea's lemma, we have

$$|p - p_h|_{1,\Omega} \leq |p - G(p^{sp})|_{1,\Omega} + |G(p^{sp}) - c_i \phi_i^K|_{1,\Omega}. \quad (3.57)$$

Next, we present an estimate for the second term. We choose  $c_i = G(p^{sp}(x_i))$ , where  $x_i$  are vertices of  $K$ . Furthermore, using Taylor expansion of  $G$  around  $\bar{p}_K$ , which is the average of  $p^{sp}$  over  $K$ ,

$$G(p^{sp}(x_i)) = G(\bar{p}_K) + G'(\bar{p}_K)(p^{sp}(x_i) - \bar{p}_K) + \frac{1}{2} G''(\xi_{x_i})(p^{sp}(x_i) - \bar{p}_K)^2,$$

where  $\xi_{x_i} = \bar{p}_K + \theta(p^{sp}(x_i) - \bar{p}_K)$ ,  $0 < \theta < 1$ , we have in each  $K$

$$\begin{aligned} c_i \phi_i^K &= G(\bar{p}_K) \phi_i^K + G'(\bar{p}_K)(p^{sp}(x_i) - \bar{p}_K) \phi_i^K + \frac{1}{2} G''(\xi_{x_i})(p^{sp}(x_i) - \bar{p}_K)^2 \phi_i^K = \\ &G(\bar{p}_K) + G'(\bar{p}_K)(p^{sp}(x_i) \phi_i^K - \bar{p}_K) + \frac{1}{2} G''(\xi_{x_i})(p^{sp}(x_i) - \bar{p}_K)^2 \phi_i^K. \end{aligned} \quad (3.58)$$

In the last step, we have used  $\sum_i \phi_i^K = 1$ . Similarly, in each  $K$ ,

$$G(p^{sp}(x)) = G(\bar{p}_K) + G'(\bar{p}_K)(p^{sp}(x) - \bar{p}_K) + \frac{1}{2} G''(\xi_x)(p^{sp}(x) - \bar{p}_K)^2, \quad (3.59)$$

where  $\xi_x = \bar{p}_K + \theta(p^{sp}(x) - \bar{p}_K)$ ,  $0 < \theta < 1$ . Using (3.58) and (3.59), we get

$$\begin{aligned} |G(p^{sp}) - c_i \phi_i^K|_{1,K} &\leq |G'(\bar{p}_K)(p^{sp}(x) - p^{sp}(x_i) \phi_i^K)|_{1,K} + \left| \frac{1}{2} G''(\xi_{x_i})(p^{sp}(x_i) - \bar{p}_K)^2 \phi_i^K \right|_{1,K} + \\ &\left| \frac{1}{2} G''(\xi_x)(p^{sp}(x) - \bar{p}_K)^2 \right|_{1,K}. \end{aligned} \quad (3.60)$$

Because of  $|p^{sp}(x) - p^{sp}(x_i)|_{1,K} \leq Ch \|f\|_{0,K}$ , the estimate of the first term is the following

$$|G'(\bar{p}_K)(p^{sp}(x) - p^{sp}(x_i) \phi_i^K)|_{1,K} \leq Ch \|f\|_{0,K}.$$

For the second term on the right hand side of (3.60), assuming  $p^{sp}(x) \in W^{1,s}(\Omega)$ , we have

$$|G''(\xi_{x_i})(p^{sp}(x_i) - \bar{p}_K)^2 \phi_i^K|_{1,K} \leq Ch^{2-4/s} |p^{sp}|_{W^{1,s}(K)}^2 \leq Ch^{1-2/s} |p^{sp}|_{W^{1,s}(K)}.$$

where  $s > 2$ . Here, we have used the inequality (e.g., [4])

$$|u(x) - u(y)| \leq C|x - y|^{1-2/s} |u|_{W^{1,s}},$$

for  $s > 2$ , where  $C$  depends only on  $s$ .

For the third term, since  $G''$  and  $G'''$  are bounded, we have the following estimate:

$$\begin{aligned} |G''(\xi_x)(p^{sp}(x) - \bar{p}_K)^2|_{1,K} &\leq C \|(p^{sp}(x) - \bar{p}_K)^2 \nabla p^{sp}(x)\|_{0,K} + \\ C \|(p^{sp}(x) - \bar{p}_K) \nabla p^{sp}(x)\|_{0,K} &\leq Ch^{2-4/s} |p^{sp}|_{W^{1,s}(K)}^2 |p^{sp}|_{1,K} + \\ Ch^{1-2/s} |p^{sp}|_{1,K} &\leq Ch^{2-4/s} |p^{sp}|_{W^{1,s}(\Omega)}^2 |p^{sp}|_{1,K} + \\ Ch^{1-2/s} |p^{sp}|_{1,K} &\leq Ch^{1-2/s} |p^{sp}|_{1,K}. \end{aligned} \quad (3.61)$$

Combining the above estimates, we have for (3.60),

$$|G(p^{sp}) - c_i \phi_i^K|_{1,K} \leq Ch^{1-2/s} |p^{sp}|_{W^{1,s}(K)} + Ch^{1-2/s} |p^{sp}|_{1,K} + Ch \|f\|_{0,K}. \quad (3.62)$$

Summing (3.62) over all  $K$  and taking into account Assumption G, we have

$$|p - p_h|_{1,\Omega} \leq C\delta + Ch^{1-2/s} |p^{sp}|_{W^{1,s}(\Omega)} + Ch^{1-2/s} |p^{sp}|_{1,\Omega} + Ch \|f\|_{0,\Omega} \leq C\delta + Ch^{1-2/s}. \quad (3.63)$$

Consequently, if  $s > 2$  (see e.g., [8]) single-phase flow based multiscale finite element method converges.

**Extensions of Galerkin finite element methods with limited global information** The multiscale finite element methods considered above employ information from only one single-phase flow solution. In general, depending on the source term, boundary data, and mobility  $\lambda(S)$  (if it contains sharp variations), it might be necessary to use information from multiple global solutions for the computation of accurate two-phase flow solution. The previous multiscale finite element methods can be extended to take into account

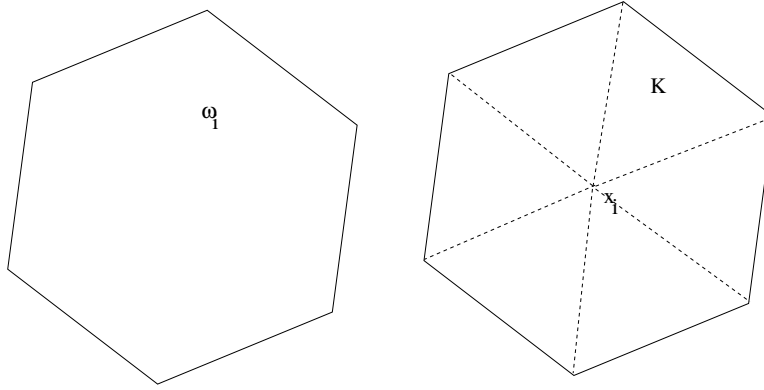
additional global information. Next, we present an extension of the Galerkin multiscale finite element method that uses the partition of unity method [11] (also see e.g., [80], [45], [57]).

Assume that  $u_1, u_2, \dots, u_N$  are the global functions such that  $|p - G(u_1, u_2, \dots, u_N)|_{1,\Omega}$  is sufficiently small. Here,  $u_1, \dots, u_N$  can be possible pressure snapshots for different mobility  $\lambda(S)$  or pressure fields corresponding to different source terms and/or boundary conditions. We would like to note that in a very interesting paper [72], the authors prove under certain conditions on  $f$  (source term) and  $\lambda = 1$  that  $p$  is a smooth function of single-phase flow solutions (elliptic pressure equations) with boundary conditions  $x_1$  and  $x_2$  (it is also extended to multi-dimensional case). In this case,  $u_1$  and  $u_2$  are the solutions of single-phase flow equations with boundary conditions  $u_i = x_i$  ( $i = 1, 2$ ), and it was shown that  $p(u_1, u_2) \in H^2$ . Next, we will formulate the method.

Let  $\omega_i$  be a patch (see Figure 3.23), and define  $\phi_i^0$  to be piecewise linear basis function in patch  $\omega_i$ , such that  $\phi_i^0(x_j) = \delta_{ij}$ . For simplicity of notation, denote  $u_1 = 1$ . Then, the multiscale finite element method for each patch  $\omega_i$  is constructed by

$$\psi_{ij} = \phi_i^0 u_j \quad (3.64)$$

where  $j = 1, \dots, N$  and  $i$  is the index of nodes (see Figure 3.23). First, we note that in each  $K$ ,  $\sum_{i=1}^n \psi_{ij} = u_j$  is the desired single-phase flow solution.



**Fig. 3.23.** Schematic description of patch

We will use the following assumption. *There exists a sufficiently smooth scalar valued function  $G(\eta)$ ,  $\eta \in R^N$  ( $G \in C^3$ ), such that*

$$|p - G(u_1, \dots, u_N)|_{1,\Omega} \leq C\delta, \quad (3.65)$$

where  $\delta$  is sufficiently small.

As before the form of the function  $G$  is not important for the computations, however, it is crucial that the basis functions span  $u_1, \dots, u_N$  in each coarse block. The next theorem shows that MsFEM converges for problems without scale separation in this case.

*Theorem 34* Assume (3.65) and  $u_i \in W^{1,s}(\Omega)$ ,  $s > 2$ ,  $i = 1, \dots, N$ . Then

$$|p - p_h|_{1,\Omega} \leq C\delta + Ch^{1-2/s}.$$

The proof of this theorem is given in [2].

**Mixed finite element methods with limited global information** One can carry out the analysis of mixed multiscale finite element method with limited global information. First, we re-formulate our assumption for the analysis of mixed multiscale finite element methods. From (3.55), it follows that

$$\|\nabla p - G'(p^{sp})\nabla p^{sp}\|_{0,\Omega} \leq C\delta.$$

Using the fact that  $k$  and  $\lambda(x)$  are bounded, we have

$$\|\lambda(x)k\nabla p - G'(p^{sp})\lambda(x)k\nabla p^{sp}\|_{0,\Omega} \leq C\delta.$$

Noting that  $u = \lambda(x)k\nabla p$  and  $u^{sp} = k\nabla p^{sp}$ , it follows that there exists a coarse-scale function scalar  $A(x)$  such that

$$\|u - A(x)u^{sp}\|_{0,\Omega} \leq \delta. \quad (3.66)$$

Since  $A(x)u^{sp}$  approximates  $u$ , we assume that it has small divergence,

$$\left| \int_K \operatorname{div}(A(x)u^{sp}) dx \right| \leq C\delta_1 h^2 \quad (3.67)$$

in each  $K$ , where  $\delta_1$  is a small number. For our analysis, we note that (3.67) gives

$$\left| \int_{\partial K} A(x)u^{sp} n^K ds \right| \leq C\delta_1 h^2. \quad (3.68)$$

We will assume that  $A(x) \in C^\gamma$  ( $0 < \gamma \leq 1$ ). (3.68) can be written as

$$\left| \sum_i A_i \int_{\partial e_i^K} u^{sp} n^K ds \right| \leq C\delta_1 h^2. \quad (3.69)$$

Here  $A_i$ 's are defined as  $A_i = \int_{\partial e_i^K} A(x)u^{sp} n^K ds / \int_{\partial e_i^K} u^{sp} n^K ds$ , since  $\int_{\partial e_i^K} u^{sp} n^K ds = \beta_i^K \neq 0$ . Note that *not* for any  $A(x)$ ,  $A_i$  is necessarily a value of  $A(x)$  along the edge  $e_i^K$  because  $u^{sp} n^K$  can change sign. However, we only need to define  $A(x)$  for each edge by its value  $A_i$  (e.g., the value of  $A(x)$  at the center of edge). Then, for any such  $A(x)$ , (3.66) is satisfied provided  $\delta < h^\gamma$ . This can be directly verified. Thus, our main assumption will be (3.66) and (3.69), where  $A(x)$  is defined, for example, at the center of each edge  $e_i^K$ . We would like to note that from the fact that  $\operatorname{div}(A(x)u^{sp})$  is small in each  $K$ , it follows that  $A(x)$ , for example, can be taken as an approximation of stream function corresponding to  $u^{sp}$ . As before, the form of  $A(x)$  is not important for the computations of multiscale solutions.

The following theorem about the convergence of mixed multiscale finite element methods for problems without scale separation is proven in [2].

*Theorem 35* Assume (3.66) and (3.69) and  $A(x) \in C^\gamma$ ,  $0 < \gamma \leq 1$ . Let  $(u, p)$  and  $(u_h, p_h)$  respectively solve the problem (3.37) and (3.39) with single-phase flow based mixed multiscale finite element, then

$$\|u - u_h\|_{H(\operatorname{div}, \Omega)} + \|p - p_h\|_{0,\Omega} \leq C\delta + C\delta_1 + Ch^\gamma. \quad (3.70)$$

## 4 Multiscale finite element methods for nonlinear partial differential equations

Next, we show that MsFEM can be naturally generalized to nonlinear partial differential equations. The goal of MsFEM is to find a numerical approximation of a homogenized solution *without* solving auxiliary problems (e.g., periodic cell problems) that arise in homogenization. The homogenized solutions are sought on a coarse grid space  $S^h$ , where  $h \gg \epsilon$ . Let  $K^h$  be a partition of  $\Omega$ . We denote by  $S^h$  standard family of finite dimensional space, which possesses approximation properties, e.g., piecewise linear functions over triangular elements,

$$S^h = \{v_h \in C^0(\overline{\Omega}) : \text{the restriction } v_h \text{ is linear for each element } K \text{ and } v_h = 0 \text{ on } \partial\Omega\}. \quad (4.1)$$

In further presentation,  $K$  is a triangular element that belongs to  $K^h$ . To formulate MsFEM for general nonlinear problems, we will need (1) a *multiscale mapping* that gives us the desired approximation containing the small scale information and (2) a *multiscale numerical formulation* of the equation.

We consider the formulation and analysis of MsFEM for general nonlinear elliptic equations,  $u_\epsilon \in W_0^{1,p}(\Omega)$

$$-div a_\epsilon(x, u_\epsilon, \nabla u_\epsilon) + a_{0,\epsilon}(x, u_\epsilon, \nabla u_\epsilon) = f, \quad (4.2)$$

where  $a_\epsilon(x, \eta, \xi)$  and  $a_{0,\epsilon}(x, \eta, \xi)$ ,  $\eta \in \mathbb{R}$ ,  $\xi \in \mathbb{R}^d$  satisfy the following assumptions:

$$|a_\epsilon(x, \eta, \xi)| + |a_{0,\epsilon}(x, \eta, \xi)| \leq C(1 + |\eta|^{p-1} + |\xi|^{p-1}), \quad (4.3)$$

$$(a_\epsilon(x, \eta, \xi_1) - a_\epsilon(x, \eta, \xi_2), \xi_1 - \xi_2) \geq C|\xi_1 - \xi_2|^p, \quad (4.4)$$

$$(a_\epsilon(x, \eta, \xi), \xi) + a_{0,\epsilon}(x, \eta, \xi) \geq C|\xi|^p. \quad (4.5)$$

Denote

$$H(\eta_1, \xi_1, \eta_2, \xi_2, r) = (1 + |\eta_1|^r + |\eta_2|^r + |\xi_1|^r + |\xi_2|^r), \quad (4.6)$$

for arbitrary  $\eta_1, \eta_2 \in \mathbb{R}$ ,  $\xi_1, \xi_2 \in \mathbb{R}^d$ , and  $r > 0$ . We further assume that

$$\begin{aligned} & |a_\epsilon(x, \eta_1, \xi_1) - a_\epsilon(x, \eta_2, \xi_2)| + |a_{0,\epsilon}(x, \eta_1, \xi_1) - a_{0,\epsilon}(x, \eta_2, \xi_2)| \\ & \leq C H(\eta_1, \xi_1, \eta_2, \xi_2, p-1) \nu(|\eta_1 - \eta_2|) \\ & + C H(\eta_1, \xi_1, \eta_2, \xi_2, p-1-s) |\xi_1 - \xi_2|^s, \end{aligned} \quad (4.7)$$

where  $\eta \in \mathbb{R}$  and  $\xi \in \mathbb{R}^d$ ,  $s > 0$ ,  $p > 1$ ,  $s \in (0, \min(p-1, 1))$  and  $\nu$  is the modulus of continuity, a bounded, concave, and continuous function in  $\mathbb{R}_+$  such that  $\nu(0) = 0$ ,  $\nu(t) = 1$  for  $t \geq 1$  and  $\nu(t) > 0$  for  $t > 0$ . These assumptions guarantee the well-posedness of the nonlinear elliptic problem (4.2). Here  $\Omega \subset \mathbb{R}^d$  is a Lipschitz domain and  $\epsilon$  denotes the small scale of the problem. The homogenization of nonlinear partial differential equations has been studied previously (see, e.g., [73]). It can be shown that a solution  $u_\epsilon$  converges (up to a sub-sequence) to  $u$  in an appropriate norm, where  $u \in W_0^{1,p}(\Omega)$  is a solution of a homogenized equation

$$-div a^*(x, u, Du) + a_0^*(x, u, Du) = f. \quad (4.8)$$

*Multiscale mapping.* Introduce the mapping  $E^{MsFEM} : S^h \rightarrow V_\epsilon^h$  in the following way. For each element  $v_h \in S^h$ ,  $v_{\epsilon,h} = E^{MsFEM} v_h$  is defined as the solution of

$$-div a_\epsilon(x, \eta^{v_h}, \nabla v_{\epsilon,h}) = 0 \text{ in } K, \quad (4.9)$$

$v_{\epsilon,h} = v_h$  on  $\partial K$  and  $\eta^{v_h} = \frac{1}{|K|} \int_K v_h dx$  for each  $K$ . We would like to point out that different boundary conditions can be chosen to obtain more accurate solutions and this will be discussed later. Note that for linear problems,  $E^{MsFEM}$  is a linear operator, where for each  $v_h \in S^h$ ,  $v_{\epsilon,h}$  is the solution of the linear problem. Consequently,  $V_\epsilon^h$  is a linear space that can be obtained by mapping a basis of  $S^h$ . This is precisely the construction presented in [50] for linear elliptic equations.

*Multiscale numerical formulation.* Multiscale finite element formulation of the problem is the following. Find  $u_h \in S^h$  (consequently,  $u_{\epsilon,h} (= E^{MsFEM} u_h) \in V_\epsilon^h$ ) such that

$$\langle A_{\epsilon,h} u_h, v_h \rangle = \int_\Omega f v_h dx \quad \forall v_h \in S^h, \quad (4.10)$$

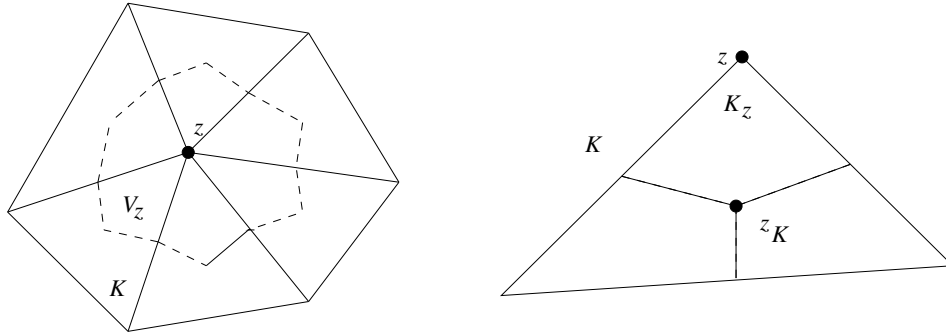
where

$$\langle A_{\epsilon,h} u_h, v_h \rangle = \sum_{K \in K^h} \int_K ((a_\epsilon(x, \eta^{u_h}, \nabla u_{\epsilon,h}), \nabla v_h) + a_{0,\epsilon}(x, \eta^{u_h}, \nabla u_{\epsilon,h}) v_h) dx. \quad (4.11)$$

Note that the above formulation of MsFEM is a generalization of the Petrov-Galerkin MsFEM introduced in [48] for linear problems. MsFEM, introduced above, can be generalized to different kinds of nonlinear problems and this will be discussed later.

#### 4.1 Multiscale finite volume element method (MsFVEM)

The formulation of multiscale finite element (MsFEM) can be extended to a finite volume method. By its construction, the finite volume method has local conservative properties [43] and it is derived from a local relation, namely the balance equation/conservation expression on a number of subdomains which are called control volumes. Finite volume element method can be considered as a Petrov-Galerkin finite element method, where the test functions are constants defined in a dual grid. Consider a triangle  $K$ , and let  $z_K$  be its barycenter. The triangle  $K$  is divided into three quadrilaterals of equal area by connecting  $z_K$  to the midpoints of its three edges. We denote these quadrilaterals by  $K_z$ , where  $z \in Z_h(K)$  are the vertices of  $K$ . Also we denote  $Z_h = \bigcup_K Z_h(K)$ , and  $Z_h^0$  are all vertices that do not lie on  $\Gamma_D$ , where  $\Gamma_D$  is Dirichlet boundaries. The control volume  $V_z$  is defined as the union of the quadrilaterals  $K_z$  sharing the vertex  $z$  (see Figure 4.1). The multiscale finite volume element method (MsFVEM) is to find  $u_h \in S^h$  (consequently,



**Fig. 4.1.** *Left:* Portion of triangulation sharing a common vertex  $z$  and its control volume. *Right:* Partition of a triangle  $K$  into three quadrilaterals

$u_{\epsilon,h} = E^{MsFVEM} u_h$  such that

$$-\int_{\partial V_z} a_\epsilon(x, \eta^{u_h}, \nabla u_{\epsilon,h}) \cdot n \, dS + \int_{V_z} a_{0,\epsilon}(x, \eta^{u_h}, \nabla u_{\epsilon,h}) \, dx = \int_{V_z} f \, dx \quad \forall z \in Z_h^0, \quad (4.12)$$

where  $n$  is the unit normal vector pointing outward on  $\partial V_z$ . Note that the number of control volumes that satisfies (4.12) is the same as the dimension of  $S^h$ . We will present numerical results for both multiscale finite element and multiscale finite volume element methods.

#### 4.2 Examples of $V_\epsilon^h$

**Linear case.** For linear operators,  $V_\epsilon^h$  can be obtained by mapping a basis of  $S^h$ . Define a basis of  $S^h$ ,  $S^h = \text{span}(\phi_0^i)$ , where  $\phi_0^i$  are standard linear basis functions. In each element  $K \in K^h$ , we define a set of nodal basis  $\{\phi_\epsilon^i\}$ ,  $i = 1, \dots, n_d$  with  $n_d (= 3)$  being the number of nodes of the element, satisfying

$$-\text{div} a_\epsilon(x) \nabla \phi_\epsilon^i = 0 \quad \text{in } K \in K^h \quad (4.13)$$

and  $\phi_\epsilon^i = \phi_0^i$  on  $\partial K$ . Thus, we have

$$V_\epsilon^h = \text{span}\{\phi_\epsilon^i; \quad i = 1, \dots, n_d, \quad K \in K^h\} \subset H_0^1(\Omega).$$

Oversampling technique can be used to improve the method [50].

**Special nonlinear case.** For the special case,  $a_\epsilon(x, u_\epsilon, \nabla u_\epsilon) = a_\epsilon(x) b(u_\epsilon) \nabla u_\epsilon$ ,  $V_\epsilon^h$  can be related to the linear case. Indeed, for this case, the local problems associated with the multiscale mapping  $E^{MsFEM}$  (see (4.9)) have the form

$$-\text{div} a_\epsilon(x) b(\eta^{v_h}) \nabla v_{\epsilon,h} = 0 \quad \text{in } K.$$

Because  $\eta^{v_h}$  are constants over  $K$ , the local problems satisfy the linear equations,

$$-div a_\epsilon(x) \nabla \phi_\epsilon^i = 0 \quad \text{in } K,$$

and  $V_\epsilon^h$  can be obtained by mapping a basis of  $S^h$  as it is done for the first example. Thus, for this case one can construct the base functions in the beginning of the computations.

**$V_\epsilon^h$  using subdomain problems.** One can use the solutions of smaller (than  $K \in \mathbf{K}^h$ ) subdomain problems to approximate the solutions of the local problems (4.9). This can be done in various ways based on a homogenization expansion. For example, instead of solving (4.9) we can solve (4.9) in a subdomain  $S$  with boundary conditions  $v_h$  restricted onto the subdomain boundaries,  $\partial S$ . Then the gradient of the solution in a subdomain can be extended periodically to  $K$  to approximate  $\nabla v_{\epsilon,h}$  in (4.11).  $v_{\epsilon,h}$  can be easily reconstructed based on  $\nabla v_{\epsilon,h}$ . When the multiscale coefficient has a periodic structure, the multiscale mapping can be constructed over one periodic cell with a specified average.

### 4.3 Convergence of MsFEM for nonlinear partial differential equations

In [39] it was shown using  $G$ -convergence theory that

$$\lim_{h \rightarrow 0} \lim_{\epsilon \rightarrow 0} \|u_h - u\|_{W_0^{1,p}(\Omega)} = 0, \quad (4.14)$$

(up to a subsequence) where  $u$  is a solution of (4.8) and  $u_h$  is a MsFEM solution given by (4.10). This result can be obtained without any assumption on the nature of the heterogeneities and can not be improved because there could be infinitely many scales,  $\alpha(\epsilon)$ , present such that  $\alpha(\epsilon) \rightarrow 0$  as  $\epsilon \rightarrow 0$ .

For the periodic case, it can be shown that the convergence of MsFEM in the limit as  $\epsilon/h \rightarrow 0$ . To show the convergence for  $\epsilon/h \rightarrow 0$ , we consider  $h = h(\epsilon)$ , such that  $h(\epsilon) \gg \epsilon$  and  $h(\epsilon) \rightarrow 0$  as  $\epsilon \rightarrow 0$ . We would like to note that this limit as well as the proof of the periodic case is different from (4.14), where the double-limit is taken. In contrast to the proof of (4.14), the proof of the periodic case requires the correctors for the solutions of the local problems.

Next we will present the convergence results for MsFEM solutions. For general nonlinear elliptic equations under the assumptions (4.3)-(4.7) the strong convergence of MsFEM solutions can be shown. In the proof of this theorem we show the form of the truncation error (in a weak sense) in terms of the resonance errors between the mesh size and small scale  $\epsilon$ . The resonance errors are derived explicitly. To obtain the convergence rate from the truncation error, one needs some lower bounds. Under the general conditions, such as (4.3)-(4.7), one can prove strong convergence of MsFEM solutions without an explicit convergence rate (cf. [78]). To convert the obtained convergence rates for the truncation errors into the convergence rate of MsFEM solutions, additional assumptions, such as monotonicity, are needed.

Next, we formulate convergence theorems. The proofs can be found in [36].

*Theorem 41* Assume  $a_\epsilon(x, \eta, \xi)$  and  $a_{0,\epsilon}(x, \eta, \xi)$  are periodic functions with respect to  $x$  and let  $u$  be a solution of (4.8) and  $u_h$  is a MsFEM solution given by (4.10). Moreover, we assume that  $\nabla u_h$  is uniformly bounded in  $L^{p+\alpha}(\Omega)$  for some  $\alpha > 0$ . Then

$$\lim_{\epsilon \rightarrow 0} \|u_h - u\|_{W_0^{1,p}(\Omega)} = 0 \quad (4.15)$$

where  $h = h(\epsilon) \gg \epsilon$  and  $h \rightarrow 0$  as  $\epsilon \rightarrow 0$  (up to a subsequence).

*Theorem 42* Let  $u$  and  $u_h$  be the solutions of the homogenized problem (4.8) and MsFEM (4.10), respectively, with the coefficient  $a_\epsilon(x, \eta, \xi) = a(x/\epsilon, \xi)$  and  $a_{0,\epsilon} = 0$ . Then

$$\|u_h - u\|_{W_0^{1,p}(\Omega)}^p \leq c \left(\frac{\epsilon}{h}\right)^{\frac{s}{(p-1)(p-s)}} + c \left(\frac{\epsilon}{h}\right)^{\frac{p}{p-1}} + ch^{\frac{p}{p-1}}. \quad (4.16)$$



#### 4.4 Multiscale finite element methods for nonlinear parabolic equations

We consider

$$\frac{\partial}{\partial t} u_\epsilon - \operatorname{div}(a_\epsilon(x, t, u_\epsilon, \nabla u_\epsilon)) + a_{0,\epsilon}(x, t, u_\epsilon, \nabla u_\epsilon) = f, \quad (4.17)$$

where  $\epsilon$  is a small scale. Our motivation in considering (4.17) mostly stems from the applications of flow in porous media (multi-phase flow in saturated porous media, flow in unsaturated porous media) though many applications of nonlinear parabolic equations of these kinds occur in transport problems. Many problems in subsurface modeling have multiscale nature where the heterogeneities associated with the media is no longer periodic. It was shown that a solution  $u_\epsilon$  converges to  $u$  (up to a subsequence) in an appropriate sense where  $u$  is a solution of

$$\frac{\partial}{\partial t} u - \operatorname{div}(a^*(x, t, u, \nabla u)) + a_0^*(x, t, u, \nabla u) = f. \quad (4.18)$$

In [38] the homogenized fluxes  $a^*$  and  $a_0^*$  are computed under the assumption that the heterogeneities are strictly stationary random fields with respect to both space and time.

The numerical homogenization procedure presented in the previous section can be extended to parabolic equations. To do this we will first formulate MsFEM in a slightly different manner from that presented in [50] for *the linear problem*. Consider a standard finite dimensional  $S^h$  space over a coarse triangulation of  $\Omega$ , (4.1) and define  $E^{MsFEM} : S^h \rightarrow V_\epsilon^h$  in the following way. For each  $u_h \in S^h$  there is a corresponding element  $u_{h,\epsilon}$  in  $V_\epsilon^h$  that is defined by

$$\frac{\partial}{\partial t} u_{h,\epsilon} - \operatorname{div}(a_\epsilon(x, t) \nabla u_{h,\epsilon}) = 0 \text{ in } K \times [t_n, t_{n+1}], \quad (4.19)$$

with boundary condition  $u_{h,\epsilon} = u_h$  on  $\partial K$ , and  $u_{h,\epsilon}(t = t_n) = u_h$ . For the linear equations  $E^{MsFEM}$  is a linear operator and the obtained multiscale space,  $V_\epsilon^h$  is a linear space on  $\Omega \times [t_n, t_{n+1}]$ . Moreover, the basis in the space  $V_\epsilon^h$  can be obtained by mapping the basis functions of  $S^h$ . For *the nonlinear parabolic equations* considered in this paper the operator  $E^{MsFEM}$  is constructed similar to (4.19) using the local problems, i.e., for each  $u_h \in S^h$  there is a corresponding element  $u_{h,\epsilon}$  in  $V_\epsilon^h$  that is defined by

$$\frac{\partial}{\partial t} u_{h,\epsilon} - \operatorname{div}(a_\epsilon(x, t, \eta, \nabla u_{h,\epsilon})) = 0 \text{ in } K \times [t_n, t_{n+1}], \quad (4.20)$$

with boundary condition  $u_{h,\epsilon} = u_h$  on  $\partial K$ , and  $u_{h,\epsilon}(t = t_n) = u_h$ . Here  $\eta = \frac{1}{|K|} \int_K u_h dx$ . Note  $E^{MsFEM}$  is a nonlinear operator and  $V_\epsilon^h$  is no longer a linear space.

The following method that can be derived from general multiscale finite element framework is equivalent to our numerical homogenization procedure. Find  $u_h \in V_\epsilon^h$  such that

$$\int_{t_n}^{t_{n+1}} \int_\Omega \frac{\partial}{\partial t} u_h v_h dx dt + A(u_h, v_h) = \int_{t_n}^{t_{n+1}} \int_\Omega f v_h dx dt, \quad \forall v_h \in S^h,$$

where

$$A(u_h, w_h) = \sum_K \int_{t_n}^{t_{n+1}} \int_K (a_\epsilon(x, t, \eta^{u_h}, \nabla v_\epsilon), \nabla w_h) + a_{0,\epsilon}(x, t, \eta^{u_h}, \nabla v_\epsilon) w_h dx dt,$$

where  $v_\epsilon$  is the solution of the local problem 4.20),  $u_h = l^{u_h}$  in each  $K$ ,  $\eta^{u_h} = \frac{1}{|K|} \int_K l^{u_h} dx$ , and  $u_h$  is known at  $t = t_n$ .

We would like to note that the operator  $E^{MsFEM}$  can be constructed using larger domains as it is done in MsFEM with oversampling [50]. This way one reduces the effects of the boundary conditions and initial conditions. In particular, for the temporal oversampling it is only sufficient to start the computations before  $t_n$  and end them at  $t_{n+1}$ . Consequently, the oversampling domain for  $K \times [t_n, t_{n+1}]$  consists of  $[\tilde{t}_n, t_{n+1}] \times S$ , where  $\tilde{t}_n < t_n$  and  $K \subset S$ . More precise formulation and detail numerical studies of oversampling technique for nonlinear equations are currently under investigation. Further we would like to note that oscillatory initial

conditions can be imposed (without using oversampling techniques) based on the solution of the elliptic part of the local problems (4.20). These initial conditions at  $t = t_n$  are the solutions of

$$-div(a_\epsilon(x, t, \eta, \nabla u_{h,\epsilon})) = 0 \text{ in } K, \quad (4.21)$$

or

$$-div(\bar{a}_\epsilon(x, \eta, \nabla u_{h,\epsilon})) = 0 \text{ in } K, \quad (4.22)$$

where  $\bar{a}_\epsilon(x, \eta, \xi) = \frac{1}{t_{n+1}-t_n} \int_{t_n}^{t_{n+1}} a(T(x/\epsilon^\beta, \tau/\epsilon^\alpha)\omega, \eta, \xi) d\tau$  and  $u_{h,\epsilon} = u_h$  on  $\partial K$ . The latter can become efficient depending on the inter-play between the temporal and spatial scales. This issue is discussed below.

Note that in the case of periodic media the local problems can be solved in a single period in order to construct  $A(u_h, v_h)$ . In general, one can solve the local problems in a domain different from  $K$  (an element) to calculate  $A(u_h, v_h)$ , and our analysis is applicable to these cases. Note that the numerical advantages of our approach over the fine scale simulation is similar to that of MsFEM. In particular, for each Newton's iteration a linear system of equations on a *coarse grid* is solved.

For some special cases the operator  $E^{MsFEM}$  introduced in the previous section can be simplified (see [39]). In general one can avoid solving the local parabolic problems if the ratio between temporal and spatial scales is known, and solve instead a simplified equation. For example, assuming that  $a_\epsilon(x, t, \eta, \xi) = a(x/\epsilon^\beta, t/\epsilon^\alpha, \eta, \xi)$ , we have the following. If  $\alpha < 2\beta$  one can solve instead of (4.20) the local problem  $-div(a_\epsilon(x, t, \eta^{u_h}, \nabla v_\epsilon)) = 0$ , if  $\alpha > 2\beta$  one can solve instead of (4.20) the local problem  $-div(\bar{a}_\epsilon(x, \eta^{u_h}, \nabla v_\epsilon)) = 0$ , where  $\bar{a}_\epsilon(x, \eta, \xi)$  is an average over time of  $a_\epsilon(x, t, \eta, \xi)$ , while if  $\alpha = 2\beta$  we need to solve the parabolic equation in  $K \times [t_n, t_{n+1}]$ , (4.20).

We would like to note that, in general, one can use (4.21) or (4.22) as oscillatory initial conditions and these initial conditions can be efficient for some cases. For example, for  $\alpha > 2\beta$  with initial conditions given by (4.22) the solutions of the local problems (4.20) can be computed easily since they are approximated by (4.22). Moreover, one can expect better accuracy with (4.22) for the case  $\alpha > 2\beta$  because this initial condition is more compatible with the local heterogeneities compare to the artificial linear initial conditions (cf. (4.20)). The comparison of various oscillatory initial conditions including the ones obtained by oversampling method is a subject of future studies.

Finally, we would like to mention that one can prove the following theorem.

*Theorem 4.3*  $u_h = \sum_i \theta_i(t) \phi_i^0(x)$  converges to  $u$ , a solution of the homogenized equation in  $V_0 = L^p(0, T, W_0^{1,p}(\Omega))$  as  $\lim_{h \rightarrow 0} \lim_{\epsilon \rightarrow 0}$  under additional not restrictive assumptions (see [39]).

*Remark 4.1.* The proof of the theorem uses the convergence of the solutions and the fluxes, and consequently it is applicable for the case of general heterogeneities that uses  $G$ -convergence theory. Since the  $G$ -convergence of the operators occurs up to a subsequence the numerical solution converges to a solution of a homogenized equation (up to a subsequence of  $\epsilon$ ).

## 4.5 Numerical results

In this section we present several ingredients pertaining to the implementation of multiscale finite element method for nonlinear elliptic equations. More numerical examples relevant to subsurface applications can be found in [36]. We will present numerical results for both MsFEM and multiscale finite volume element method (MsFVEM). We use an Inexact-Newton algorithm as an iterative technique to tackle the nonlinearity. For the numerical examples below, we use  $a_\epsilon(x, u_\epsilon, \nabla u_\epsilon) = a_\epsilon(x, u_\epsilon) \nabla u_\epsilon$ . Let  $\{\phi_0^i\}_{i=1}^{N_{dof}}$  be the standard piecewise linear basis functions of  $S^h$ . Then MsFEM solution may be written as

$$u_h = \sum_{i=1}^{N_{dof}} \alpha_i \phi_0^i \quad (4.23)$$

for some  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_{N_{dof}})^T$ , where  $\alpha_i$  depends on  $\epsilon$ . Hence, we need to find  $\alpha$  such that

$$F(\alpha) = 0, \quad (4.24)$$

where  $F : \mathbb{R}^{N_{dof}} \rightarrow \mathbb{R}^{N_{dof}}$  is a nonlinear operator such that

$$F_i(\alpha) = \sum_{K \in K^h} \int_K (a_\epsilon(x, \eta^{u_h}) \nabla u_{\epsilon,h}, \nabla \phi_0^i) dx - \int_\Omega f \phi_0^i dx. \quad (4.25)$$

We note that in (4.25)  $\alpha$  is implicitly buried in  $\eta^{u_h}$  and  $u_{\epsilon,h}$ . An inexact-Newton algorithm is a variation of Newton's iteration for nonlinear system of equations, where the Jacobian system is only approximately solved. To be specific, given an initial iterate  $\alpha^0$ , for  $k = 0, 1, 2, \dots$  until convergence do the following:

- Solve  $F'(\alpha^k)\delta^k = -F(\alpha^k)$  by some iterative technique until  $\|F(\alpha^k) + F'(\alpha^k)\delta^k\| \leq \beta_k \|F(\alpha^k)\|$ .
- Update  $\alpha^{k+1} = \alpha^k + \delta^k$ .

In this algorithm  $F'(\alpha^k)$  is the Jacobian matrix evaluated at iteration  $k$ . We note that when  $\beta_k = 0$  then we have recovered the classical Newton iteration. Here we have used

$$\beta_k = 0.001 \left( \frac{\|F(\alpha^k)\|}{\|F(\alpha^{k-1})\|} \right)^2, \quad (4.26)$$

with  $\beta_0 = 0.001$ . Choosing  $\beta_k$  this way, we avoid over-solving the Jacobian system when  $\alpha^k$  is still considerably far from the exact solution.

Next we present the entries of the Jacobian matrix. For this purpose, we use the following notations. Let  $K_i^h = \{K \in K^h : z_i \text{ is a vertex of } K\}$ ,  $I^i = \{j : z_j \text{ is a vertex of } K \in K_i^h\}$ , and  $K_{ij}^h = \{K \in K_i^h : K \text{ shares } \overline{z_i z_j}\}$ . We note that we may write  $F_i(\alpha)$  as follows:

$$F_i(\alpha) = \sum_{K \in K_i^h} \left( \int_K (a_\epsilon(x, \eta^{u_h}) \nabla u_{\epsilon,h}, \nabla \phi_0^i) dx - \int_K f \phi_0^i dx \right), \quad (4.27)$$

with

$$-div a_\epsilon(x, \eta^{u_h}) \nabla u_{\epsilon,h} = 0 \text{ in } K \quad \text{and} \quad u_{\epsilon,h} = \sum_{z_m \in Z_K} \alpha_m \phi_0^m \text{ on } \partial K, \quad (4.28)$$

where  $Z_K$  is all the vertices of element  $K$ . It is apparent that  $F_i(\alpha)$  is not fully dependent on all  $\alpha_1, \alpha_2, \dots, \alpha_d$ . Consequently,  $\frac{\partial F_i(\alpha)}{\partial \alpha_j} = 0$  for  $j \notin I^i$ . To this end, we denote  $\psi_\epsilon^j = \frac{\partial u_{\epsilon,h}}{\partial \alpha_j}$ . By applying chain rule of differentiation to (4.28) we have the following local problem for  $\psi_\epsilon^j$ :

$$-div a_\epsilon(x, \eta^{u_h}) \nabla \psi_\epsilon^j = \frac{1}{3} div \frac{\partial a_\epsilon(x, \eta^{u_h})}{\partial u} \nabla u_{\epsilon,h} \text{ in } K \quad \text{and} \quad \psi_\epsilon^j = \phi_\epsilon^j \text{ on } \partial K. \quad (4.29)$$

The fraction  $1/3$  comes from taking the derivative in the chain rule of differentiation. In the formulation of the local problem, we have replaced the nonlinearity in the coefficient by  $\eta^{u_h}$ , where for each triangle  $K$   $\eta^{u_h} = 1/3 \sum_{i=1}^3 \alpha_i^K$ , which gives  $\partial \eta^{u_h} / \partial \alpha_i = 1/3$ . Moreover, for a rectangular element the fraction  $1/3$  should be replaced by  $1/4$ .

Thus, provided that  $v_{\epsilon,h}$  has been computed, then we may compute  $\psi_\epsilon^j$  using (4.29). Using the above descriptions we have the expressions for the entries of the Jacobian matrix:

$$\frac{\partial F_i}{\partial \alpha_i} = \sum_{K \in K_i^h} \left( \frac{1}{3} \int_K \left( \frac{\partial a_\epsilon(x, \eta^{u_h})}{\partial u} \nabla u_{\epsilon,h}, \nabla \phi_0^i \right) dx + \int_K (a_\epsilon(x, \eta^{u_h}) \nabla \psi_i, \nabla \phi_0^i) dx \right), \quad (4.30)$$

$$\frac{\partial F_i}{\partial \alpha_j} = \sum_{K \in K_{ij}^h} \left( \frac{1}{3} \int_K \left( \frac{\partial a_\epsilon(x, \eta^{u_h})}{\partial u} \nabla u_{\epsilon,h}, \nabla \phi_\epsilon^i \right) dx + \int_K (a_\epsilon(x, \eta^{u_h}) \nabla \psi_\epsilon^j, \nabla \phi_0^i) dx \right), \quad (4.31)$$

for  $j \neq i, j \in I^i$ .

The implementation of the oversampling technique is similar to the procedure presented earlier, except the local problems in larger domains are used. As in the non-oversampling case, we denote  $\psi_\epsilon^j = \frac{\partial v_{\epsilon,h}}{\partial \alpha_j}$ , such that after applying chain rule of differentiation to the local problem we have:

$$-div a_\epsilon(x, \eta^{u_h}) \nabla \psi_\epsilon^j = \frac{1}{3} div \frac{\partial a_\epsilon(x, \eta^{u_h})}{\partial u} \nabla v_{\epsilon,h} \text{ in } S \quad \text{and} \quad \psi_\epsilon^j = \phi_0^j \text{ on } \partial S, \quad (4.32)$$

where  $\eta^{u_h}$  is computed over the corresponding element  $K$  and  $\phi_0^j$  is understood as the nodal basis functions on oversampled domain  $S$ . Then all the rest of the inexact-Newton algorithms are the same as in the non-oversampling case. Specifically, we also use (4.30) and (4.31) to construct the Jacobian matrix of the system. We note that we will only use  $\psi_\epsilon^j$  from (4.32) pertaining to the element  $K$ .

From the derivation (both for oversampling and non-oversampling) it is obvious that the Jacobian matrix is not symmetric but sparse. Computation of this Jacobian matrix is similar to computing the stiffness matrix resulting from standard finite element, where each entry is formed by accumulation of element by element contribution. Once we have the matrix stored in memory, then its action to a vector is straightforward. Because it is a sparse matrix, devoting some amount of memory for entries storage is inexpensive. The resulting linear system is solved using preconditioned bi-conjugate gradient stabilized method.

We want to solve the following problem:

$$\begin{aligned} -div a(x/\epsilon, u_\epsilon) \nabla u_\epsilon &= -1 & \text{in } \Omega \subset \mathbb{R}^2, \\ u_\epsilon &= 0 & \text{on } \partial\Omega, \end{aligned} \quad (4.33)$$

where  $\Omega = [0, 1] \times [0, 1]$ ,  $a(x/\epsilon, u_\epsilon) = k(x/\epsilon)/(1 + u_\epsilon)^{l(x/\epsilon)}$ , with

$$k(x/\epsilon) = \frac{2 + 1.8 \sin(2\pi x_1/\epsilon)}{2 + 1.8 \cos(2\pi x_2/\epsilon)} + \frac{2 + \sin(2\pi x_2/\epsilon)}{2 + 1.8 \cos(2\pi x_1/\epsilon)} \quad (4.34)$$

and  $l(x/\epsilon)$  is generated from  $k(x/\epsilon)$  such that the average of  $l(x/\epsilon)$  over  $\Omega$  is 2. Here we use  $\epsilon = 0.01$ . Because the exact solution for this problem is not available, we use a well resolved numerical solution using standard finite element method as a reference solution. The resulting nonlinear system is solved using inexact-Newton algorithm. The reference solution is solved on  $512 \times 512$  mesh. Tables 4.2 and 4.4 present the relative errors of the solution with and without oversampling, respectively. In tables 4.3 and 4.5, the relative errors for multiscale finite volume element method are presented. The relative errors are computed as the corresponding error divided by the norm of the solution. In each table, the second, third, and fourth columns list the relative error in  $L^2$ ,  $H^1$ , and  $L^\infty$  norm, respectively. As we can see from these two tables, the oversampling significantly improves the accuracy of the multiscale method.

For our next example, we consider the problem with non-periodic coefficients, where  $a_\epsilon(x, \eta) = k_\epsilon(x)/(1 + \eta)^{\alpha_\epsilon(x)}$ .  $k_\epsilon(x) = \exp(\beta_\epsilon(x))$  is chosen such that  $\beta_\epsilon(x)$  is a realization of a random field with the spherical variogram [26] and with the correlation lengths  $l_x = 0.2$ ,  $l_y = 0.02$  and with the variance  $\sigma = 1$ .  $\alpha_\epsilon(x)$  is chosen such that  $\alpha_\epsilon(x) = k_\epsilon(x) + const$  with the spatial average of 2. As for the boundary conditions we use “left-to-right flow” in  $\Omega = [0, 5] \times [0, 1]$  domain,  $u_\epsilon = 1$  at the inlet ( $x_1 = 0$ ),  $u_\epsilon = 0$  at the outlet ( $x_1 = 5$ ), and no flow boundary conditions on the lateral sides  $x_2 = 0$  and  $x_2 = 1$ . In Table 4.6 we present the relative error for multiscale method with oversampling. Similarly, in Table 4.7 we present the relative error for multiscale finite volume method with oversampling. Clearly, the oversampling method captures the effects induced by the large correlation features. Both  $H^1$  and horizontal flux errors are under five percent. Similar results have been observed for various kinds of non-periodic heterogeneities. In the next set of numerical examples, we test MsFEM for problems with fluxes  $a_\epsilon(x, \eta)$  that are discontinuous in space. The discontinuity in the fluxes is introduced by multiplying the underlying permeability function,  $k_\epsilon(x)$ , by a constant in certain regions, while leaving it unchanged in the rest of the domain. As an underlying permeability field,  $k_\epsilon(x)$ , we choose the random field used for the results in Table 4.6. In the first set of examples, the discontinuities are introduced along the boundaries of the coarse elements. In particular,  $k_\epsilon(x)$  on the left half of the domain is multiplied by a constant  $J$ , where  $J = \exp(1)$ , or  $\exp(2)$ , or  $\exp(4)$ . The results in Tables 4.8-4.10 show that MsFEM converges and the error falls below five percent for relatively large coarsening. For the second set of

examples (Tables 4.11-4.13), the discontinuities are not aligned with the boundaries of the coarse elements. In particular, the discontinuity boundary is given by  $y = x\sqrt{2} + 0.5$ , i.e., the discontinuity line intersects the coarse grid blocks. Similar to the aligned case, various jump magnitudes are considered. These results demonstrate the robustness of our approach for anisotropic fields where  $h$  and  $\epsilon$  are nearly the same, and the fluxes that are discontinuous spatial functions.

As for CPU comparisons, we have observed more than 92 percent CPU savings when using MsFEM without oversampling. With the oversampling approach, the CPU savings depend on the size of the oversampled domain. For example, if the oversampled domain size is two times larger than the target coarse block (half coarse block extension on each side) we have observed 70 percent CPU savings for  $64 \times 64$  and 80 percent CPU savings for  $128 \times 128$  coarse grid. In general, the computational cost will decrease if the oversampled domain size is close to the target coarse block size, and this cost will be close to the cost of MsFEM without oversampling. Conversely, the error decreases if the size of the oversampled domains increases. In the numerical examples studied in our paper, we have observed the same errors for the oversampling methods using either one coarse block extension or half coarse block extensions. The latter indicates that the leading resonance error is eliminated by using a smaller oversampled domain. Oversampled domains with one coarse block extension are previously used in simulations of flow through heterogeneous porous media. As it is indicated in [50], one can use large oversampled domains for simultaneous computations of the several local solutions. Moreover, parallel computations will improve the speed of the method because MsFEM is well suited for parallel computation [50]. For the problems where  $a_\epsilon(x, \eta, \xi) = a_\epsilon(x)b(\eta)\xi$  (see section 4.2 and the next section for applications) our multiscale computations are very fast because the base functions are built in the beginning of the computations. In this case, we have observed more than 95 percent CPU savings.

**Table 4.2.** Relative MsFEM Errors without Oversampling

N	$L^2$ -norm		$H^1$ -norm		$L^\infty$ -norm	
	Error	Rate	Error	Rate	Error	Rate
32	0.029		0.115		0.03	
64	0.053	-0.85	0.156	-0.44	0.0534	-0.94
128	0.10	-0.94	0.234	-0.59	0.10	-0.94

**Table 4.3.** Relative MsFVEM Errors without Oversampling

N	$L^2$ -norm		$H^1$ -norm		$L^\infty$ -norm	
	Error	Rate	Error	Rate	Error	Rate
32	0.03		0.13		0.04	
64	0.05	-0.65	0.19	-0.60	0.05	-0.24
128	0.058	-0.19	0.25	-0.35	0.057	-0.19

Applications of MsFEM to Richards' equation are presented in [36].

#### 4.6 Generalizations of MsFEM and some remarks

Next, we present the framework of MsFEM for general equations. Consider

$$L_\epsilon u_\epsilon = f, \tag{4.35}$$

where  $\epsilon$  is a small scale and  $L_\epsilon : X \rightarrow Y$  is an operator. Moreover, we assume that  $L_\epsilon$   $G$ -converges to  $L^*$  (up to a sub-sequence), where  $u$  is a solution of

$$L^*u = f, \tag{4.36}$$

**Table 4.4.** Relative MsFEM Errors with Oversampling

N	$L^2$ -norm		$H^1$ -norm		$L^\infty$ -norm	
	Error	Rate	Error	Rate	Error	Rate
32	0.0016		0.036		0.0029	
64	0.0012	0.38	0.019	0.93	0.0016	0.92
128	0.0024	-0.96	0.0087	1.14	0.0026	-0.71

**Table 4.5.** Relative MsFVEM Errors with Oversampling

N	$L^2$ -norm		$H^1$ -norm		$L^\infty$ -norm	
	Error	Rate	Error	Rate	Error	Rate
32	0.002		0.038		0.005	
64	0.003	-0.43	0.021	0.87	0.003	0.72
128	0.001	1.10	0.009	1.09	0.001	1.08

**Table 4.6.** Relative MsFEM Errors for random heterogeneities, spherical variogram,  $l_x = 0.20$ ,  $l_z = 0.02$ ,  $\sigma = 1.0$ 

N	$L^2$ -norm		$H^1$ -norm		$L^\infty$ -norm		hor. flux	
	Error	Rate	Error	Rate	Error	Rate	Error	Rate
32	0.0006		0.0505		0.0025		0.025	
64	0.0002	1.58	0.029	0.8	0.001	1.32	0.017	0.57
128	0.0001	1	0.016	0.85	0.0005	1	0.011	0.62

**Table 4.7.** Relative MsFVEM Errors for random heterogeneities, spherical variogram,  $l_x = 0.20$ ,  $l_z = 0.02$ ,  $\sigma = 1.0$ 

N	$L^2$ -norm		$H^1$ -norm		$L^\infty$ -norm		hor. flux	
	Error	Rate	Error	Rate	Error	Rate	Error	Rate
32	0.0006		0.0515		0.0025		0.027	
64	0.0002	1.58	0.029	0.81	0.0013	0.94	0.018	0.58
128	0.0001	1	0.016	0.85	0.0005	1.38	0.012	0.58

**Table 4.8.** Relative MsFEM Errors for random heterogeneities, spherical variogram,  $l_x = 0.20$ ,  $l_z = 0.02$ ,  $\sigma = 1.0$ , aligned discontinuity, jump =  $\exp(1)$ 

N	$L^2$ -norm		$H^1$ -norm		$L^\infty$ -norm		hor. flux	
	Error	Rate	Error	Rate	Error	Rate	Error	Rate
32	0.0006		0.0641		0.0020		0.039	
64	0.0002	1.58	0.0382	0.75	0.0010	1.00	0.027	0.53
128	0.0001	1.00	0.0210	0.86	0.0005	1.00	0.018	0.59

**Table 4.9.** Relative MsFEM Errors for random heterogeneities, spherical variogram,  $l_x = 0.20$ ,  $l_z = 0.02$ ,  $\sigma = 1.0$ , aligned discontinuity, jump =  $\exp(2)$ 

N	$L^2$ -norm		$H^1$ -norm		$L^\infty$ -norm		hor. flux	
	Error	Rate	Error	Rate	Error	Rate	Error	Rate
32	0.0008		0.0817		0.0040		0.061	
64	0.0004	1.00	0.0493	0.73	0.0023	0.80	0.041	0.57
128	0.0002	1.00	0.0256	0.95	0.0011	1.06	0.025	0.71

**Table 4.10.** Relative MsFEM Errors for random heterogeneities, spherical variogram,  $l_x = 0.20$ ,  $l_z = 0.02$ ,  $\sigma = 1.0$ , aligned discontinuity, jump =  $\exp(4)$ 

N	$L^2$ -norm		$H^1$ -norm		$L^\infty$ -norm		hor. flux	
	Error	Rate	Error	Rate	Error	Rate	Error	Rate
32	0.0011		0.1010		0.0068		0.195	
64	0.0006	0.87	0.0638	0.66	0.0045	0.59	0.109	0.84
128	0.0003	1.00	0.0349	0.87	0.0024	0.91	0.063	0.79

**Table 4.11.** Relative MsFEM Errors for random heterogeneities, spherical variogram,  $l_x = 0.20$ ,  $l_z = 0.02$ ,  $\sigma = 1.0$ , nonaligned discontinuity, jump =  $\exp(1)$ 

N	$L^2$ -norm		$H^1$ -norm		$L^\infty$ -norm		hor. flux	
	Error	Rate	Error	Rate	Error	Rate	Error	Rate
32	0.0006		0.0623		0.0023		0.035	
64	0.0002	1.58	0.0366	0.77	0.0014	0.72	0.024	0.54
128	0.0001	1.00	0.0203	0.85	0.0006	1.22	0.016	0.59

**Table 4.12.** Relative MsFEM Errors for random heterogeneities, spherical variogram,  $l_x = 0.20$ ,  $l_z = 0.02$ ,  $\sigma = 1.0$ , nonaligned discontinuity, jump =  $\exp(2)$ 

N	$L^2$ -norm		$H^1$ -norm		$L^\infty$ -norm		hor. flux	
	Error	Rate	Error	Rate	Error	Rate	Error	Rate
32	0.0010		0.0785		0.0088		0.052	
64	0.0003	1.74	0.0440	0.84	0.0052	0.76	0.031	0.75
128	0.0001	1.59	0.0239	0.88	0.0022	1.24	0.017	0.87

**Table 4.13.** Relative MsFEM Errors for random heterogeneities, spherical variogram,  $l_x = 0.20$ ,  $l_z = 0.02$ ,  $\sigma = 1.0$ , nonaligned discontinuity, jump =  $\exp(4)$ 

N	$L^2$ -norm		$H^1$ -norm		$L^\infty$ -norm		hor. flux	
	Error	Rate	Error	Rate	Error	Rate	Error	Rate
32	0.0067		0.1775		0.1000		0.164	
64	0.0016	2.07	0.0758	1.23	0.0288	1.80	0.077	1.09
128	0.0009	0.83	0.0687	0.14	0.0423	-0.55	0.039	0.98

(we refer to [73], page 14 for the definition of  $G$ -convergence for operators). The objective of MsFEM is to approximate  $u$  in  $S^h$ . Denote  $S^h$  a family of finite dimensional space such that it possesses an approximation property (see [90], [74]) as before. Here  $h$  is a scale of computation and  $h \gg \epsilon$ . For (4.35) *multiscale mapping*,  $E^{MsFEM} : S^h \rightarrow V_\epsilon^h$ , will be defined as follows. For each element  $v_h \in S^h$ ,  $v_{\epsilon,h} = E^{MsFEM} v_h$  is defined as

$$L_\epsilon^{map} v_{\epsilon,h} = 0 \text{ in } K, \quad (4.37)$$

where  $L_\epsilon^{map}$  can be, in general, different from  $L_\epsilon$  and allows us to capture the effects of the small scales. Moreover, the domains different from the target coarse block  $K$  can be used in the computations of the local solutions. To solve (4.37) one needs to impose boundary and initial conditions. This issue needs to be resolved on a case by case basis, and the main idea is to interpolate  $v_h$  onto the underlying fine grid. Further, we seek a solution of (4.35) in  $V_\epsilon^h$  as follows. Find  $u_h \in S^h$  (consequently  $u_{\epsilon,h} \in V_\epsilon^h$ ) such that

$$\langle L_\epsilon^{global} u_{\epsilon,h}, v_h \rangle = \langle f, v_h \rangle, \quad \forall v_h \in S^h, \quad (4.38)$$

where  $\langle u, v \rangle$  denotes the duality between  $X$  and  $Y$ , and  $L_\epsilon^{global}$  can be, in general, different from  $L_\epsilon$ . For example, for nonlinear elliptic equations we have  $L_\epsilon u = -\text{div} a_\epsilon(x, u, \nabla u) + a_{0,\epsilon}(x, u, \nabla u)$ ,  $L_\epsilon^{map} u =$

$diva_\epsilon(x, \eta^u, \nabla u)$  in  $K$ , and  $L_\epsilon^{global} = diva_\epsilon(x, \eta^u, \nabla u) + a_{0,\epsilon}(x, \eta^u, \nabla u)$  in  $K$ . The convergence of MsFEM is to show that  $u_h \rightarrow u$  and  $u_{\epsilon,h} \rightarrow u_\epsilon$ , where  $u_{\epsilon,h} = E^{MsFEM} u_h$  in appropriate space. The correct choices of  $L_\epsilon^{map}$  and  $L_\epsilon^{global}$  are the essential part of MsFEM and guarantees the convergence of the method.

In conclusion, we have presented a natural extension of MsFEM to nonlinear problems. This is accomplished by considering a multiscale map instead of the base functions that are considered in linear MsFEM [50]. Our approaches share some common elements with recently introduced HMM [33], where macroscopic and microscopic solvers are also needed. In general, the finding of “correct” macroscopic and microscopic solvers is the main difficulty of the multiscale methods. Our approaches follow MsFEM and, consequently, finite element methods constitute its main ingredient. The resonance errors, that arise in linear problems also arise in nonlinear problems. Note that the resonance errors are the common feature of multiscale methods unless periodic problems are considered and the solutions of the local problems in an exact period are used. To reduce the resonance errors we use oversampling technique and show that the error can be greatly reduced by sampling from the larger domains. The multiscale map for MsFEM uses the solutions of the local problems in the target coarse block. This way one can sample the heterogeneities of the coarse block. If there is a scale separation and, in addition, some kind of periodicity, one can use the solutions of the smaller size problems to approximate the multiscale map. Note that a potential disadvantage of periodicity assumption is that the periodicity can act to disrupt large-scale connectivity features of the flow. For the examples similar to the non-periodic ones considered in this paper, with the use of the smaller size problems for approximating the solutions of the local problems, we have found very large errors (of order 50 percent).

## 5 Multiscale simulations of two-phase immiscible flow in adaptive coordinate system

Previously, we discussed some applications of MsFEM to two-phase flows. In this section, we explore the use of adaptive coordinate system in multiscale simulations of two-phase porous media flows. In particular, we would like to present upscaling of transport equations and its coupling to MsFEM.

As we discussed earlier, the use of global information can improve the multiscale finite element method. In particular, the solution of the pressure equation at initial time is used to construct the boundary conditions for the basis functions. It is interesting to note that the multiscale finite element methods that employ a limited global information reduces to standard multiscale finite element method in flow-based coordinate system. This can be verified directly and the reason behind it is that we have already employed a limited global information in flow-based coordinate system. To achieve high degree of speed-up in two-phase flow computations, we also consider the upscaling of transport equation and its coupling to pressure equation.

We would like to derive an upscaled model for the transport equation. We will assume the velocity is independent of time,  $\lambda(S) = 1$ , and restrict ourselves to the two-dimensional case. Then using the pressure-streamline framework, one obtains

$$\begin{aligned} S_t^\epsilon + v_0^\epsilon f(S^\epsilon)_p &= 0 \\ S(p, \psi, t = 0) &= S_0, \end{aligned} \tag{5.1}$$

where  $\epsilon$  denotes the small scale,  $v_0^\epsilon$  denotes the Jacobian of the transformation and is positive, and  $p$  denotes the initial pressure. For simplicity, we assume  $\mathbf{k}(x) = k(x)\mathbf{I}$  and we have  $\nabla\psi \cdot \nabla p = 0$ . For deriving upscaled equations, we will first homogenize (5.1) along the streamlines, and then to homogenize across the streamlines. The homogenization along the streamlines can be done following Bourgeat and Mikelic [16] or following Hou and Xin [56] and E [32]. The latter uses two-scale convergence theory and we refer to [79] for the results on homogenization of (5.1) using two-scale convergence theory. We note that the homogenization results of Bourgeat and Mikelic is for general heterogeneities without an assumption on periodicity, and thus, is more appropriate for problems considered in the paper. Following [16], the homogenization of (5.1) can be easily derived (see Proposition 3.4 in [16]). Here, we briefly sketch the proof.

For ease of notations, we ignore the  $\psi$  dependence of  $v_0^\epsilon$  and  $S^\epsilon$ , and treat  $\psi$  as a parameter. We consider

$$v_0^\epsilon(p) = v_0(p, \frac{p}{\epsilon}).$$



Moreover, we assume that the domain is a unit interval. Then, for each  $\psi$ , it can be shown that  $S^\epsilon(p, \psi, t) \rightarrow \tilde{S}(p, \psi, t)$  in  $L^1((0, 1) \times (0, T))$ , where  $\tilde{S}$  satisfies

$$\tilde{S}_t + \tilde{v}_0 f(\tilde{S})_p = 0, \quad (5.2)$$

and where  $\tilde{v}_0$  is harmonic average of  $v_0^\epsilon$ , i.e.,

$$\frac{1}{v_0^\epsilon} \rightarrow \frac{1}{\tilde{v}_0} \quad \text{weak}^* \quad \text{in } L^\infty(0, 1).$$

The proof of this fact follows from Proposition 3.4. of [16].

Following [16] and assuming for simplicity  $\int_0^1 \frac{d\eta}{v_0^\epsilon(\eta)} = \int_0^1 \frac{d\eta}{\tilde{v}_0(\eta)} = 1$ , we introduce

$$\frac{dX^\epsilon(p)}{dp} = v_0^\epsilon(X^\epsilon(p)), \quad \frac{dX^0(p)}{dp} = \tilde{v}_0(X^0(p)).$$

Then (Lemma 3.1 of [16]):

$$X^\epsilon \rightarrow X^0 \quad \text{in } C[0, 1] \quad \text{as } \epsilon \rightarrow 0. \quad (5.3)$$

Consequently,

$$\begin{aligned} \int_0^T \int_0^1 |S^\epsilon(p, \tau) - \tilde{S}(p, \tau)| dp d\tau &= \int_0^T \int_0^1 |S^\epsilon(X^\epsilon(p), \tau) - \tilde{S}(X^\epsilon(p), \tau)| v_0^\epsilon(X^\epsilon(p)) dp d\tau \leq \\ &\int_0^T \int_0^1 |S^\epsilon(X^\epsilon(p), \tau) - \tilde{S}(X^0(p), \tau)| v_0^\epsilon(X^\epsilon(p)) dp d\tau + \\ &\int_0^T \int_0^1 |\tilde{S}(X^\epsilon(p), \tau) - \tilde{S}(X^0(p), \tau)| v_0^\epsilon(X^\epsilon(p)) dp d\tau \leq \\ &\int_0^T \int_0^1 |S^\epsilon(X^\epsilon(p), \tau) - \tilde{S}(X^0(p), \tau)| dp d\tau + \int_0^T \int_0^1 |\tilde{S}(X^\epsilon(p), \tau) - \tilde{S}(X^0(p), \tau)| dp d\tau \end{aligned} \quad (5.4)$$

The first term on the right hand side of (5.4) converges to zero because  $S^\epsilon(X^\epsilon(p), \tau)$  and  $\tilde{S}(X^0(p), \tau)$  satisfy the same equation  $u_t + f(u)_p = 0$ , however, with the following initial conditions  $S^\epsilon(X^\epsilon(p), t = 0) = S_0 \circ X^\epsilon$ ,  $\tilde{S}(X^0(p), \tau) = S_0 \circ X^0$ . Because of (5.3) and comparison principle

$$\int_0^T \int_0^1 |S^\epsilon(X^\epsilon(p), \tau) - \tilde{S}(X^0(p), \tau)| dp d\tau \leq C \int_0^1 |S_0 \circ X^\epsilon - S_0 \circ X^0| dp,$$

the first term converges to zero. The convergence of the second term for each  $\psi$  follows from the argument in [16] (page 368) using Lebesgue's dominated convergence theorem.

Next, we provide a convergence rate (see also [79]) of the fine saturation  $S^\epsilon$  to the homogenized limit  $\tilde{S}$  as  $\epsilon \rightarrow 0$ .

*Theorem 51* Assume that  $v_0^\epsilon(p)$  is bounded uniformly

$$C^{-1} \leq v_0^\epsilon(p, \frac{p}{\epsilon}) \leq D.$$

Denote by  $F(t, T)$  the solution to  $S_t + f(S)_T = 0$ . The solution  $\tilde{S}$  of (5.2) converges to  $S^\epsilon$  (assuming initial conditions that don't depend on the fast scale) at a rate given by

$$\|S^\epsilon - \tilde{S}\|_\infty \leq G\epsilon,$$

when  $F$  remains Lipschitz for all time, and

$$\|S^\epsilon - \tilde{S}\|_n \leq G\epsilon^{1/n},$$

when  $F$  develops at most a finite number of discontinuities.

*Proof.* First, we note that the velocity bound implies that  $\tilde{C}^{-1} \leq \tilde{v}_0(p) \leq \tilde{D}$ , uniformly in  $\psi, \zeta$ . We transform the equations for  $S^\epsilon$  (5.1) and  $\tilde{S}$  (5.2) to the time of flight variable defined by

$$\frac{dT^\epsilon}{dp} = \frac{1}{v_0^\epsilon(p, \psi)} \text{ for } S^\epsilon \text{ and } \frac{d\tilde{T}}{dp} = \frac{1}{\tilde{v}(p, \psi, \frac{\psi}{\epsilon})} \text{ for } \tilde{S}.$$

$$T^\epsilon(0) = 0 \quad \tilde{T}(0) = 0$$

Both equations reduce to

$$S_t + f(S)_T = 0.$$

The solution to this equation is  $F(t, T)$ . Since the initial condition does not depend on  $\epsilon$  neither does  $F$ . Then  $S = F(t, T^\epsilon(P, \Psi))$ ,  $\tilde{S} = F(t, \tilde{T}(P, \Psi))$ . Using these expressions for the saturation we can obtain the desired estimates by following the same steps as in the linear case. When  $F$  remains Lipschitz for all times we can easily obtain a pointwise estimate in terms of the Lipschitz constant  $M$   $\|S^\epsilon - \tilde{S}\|_\infty = \|F(t, T^\epsilon) - F(t, \tilde{T})\|_\infty \leq M\|T^\epsilon - \tilde{T}\|_\infty \leq G\epsilon$ . Otherwise we will need the time of flight bound that we derived for the linear flux that reduces here to

$$|T^\epsilon(P) - \tilde{T}(P)| \leq 2C\epsilon. \quad (5.5)$$

We will divide the domain in regions where  $F$  is Lipschitz with constant  $M$  in the second variable, denoted by  $A_2$ , and shock regions, denoted by  $A_1$ , and estimate the difference of  $S^\epsilon$  and  $\tilde{S}$  in each region separately. To fix the notation, let that there be  $n$  discontinuities in  $F(t, \cdot)$  of magnitude less than  $\Delta F$ , which does not have to be small, at  $\{T = T_i\}_{i=1, \dots, n}$ . We will denote the thin strips of width  $2C\epsilon$  around the discontinuities with  $A_1$

$$A_1 = \{T \text{ such that } |T - T_i| \leq 2C\epsilon, \text{ for some } i = 1, \dots, n\}$$

and with  $A_2$  its complement. We selected the width of the strip based on (5.5), so that for any point  $P$ , if  $T^\epsilon(P) \notin A_1$ , then  $T^\epsilon(P)$  and  $\tilde{T}(P)$  are on the same side of any jump  $T_i$ . When  $T^\epsilon(P) \in A_2$ ,  $F$  is Lipschitz in the region between  $T^\epsilon$  and  $\tilde{T}$ , and we can show

$$\int_{A_2} (S^\epsilon - \tilde{S})^2 dp d\psi = \int_{A_2} (F(t, T^\epsilon) - F(t, \tilde{T}))^2 dp d\psi \leq M^2 \|T^\epsilon - \tilde{T}\|_\infty^2 |T^\epsilon(A_2)^{-1}| \leq N^2 \epsilon^2 |T^\epsilon(A_2)^{-1}|,$$

where we used the time of flight bound (5.5). By  $|T^\epsilon(A_2)^{-1}|$  we denoted the image of  $A_2$  under the inverse of  $T^\epsilon(P)$ . Inside the strip  $A_1$ , even though  $S^\epsilon$  and  $\tilde{S}$  differ by an  $O(1)$  quantity we can use the smallness of the area of the strip to make the  $L_2$  norm of their difference small

$$\int_{A_1} (S^\epsilon - \tilde{S})^2 dp d\psi = \int_{A_2} (F(t, T^\epsilon) - F(t, \tilde{T}))^2 dp d\psi \leq (\Delta S + N\epsilon)^2 |T^\epsilon(A_1)^{-1}| \leq (\Delta S + N\epsilon)^2 4CDn\epsilon.$$

We estimated the area  $|T^\epsilon(A_1)^{-1}|$  by using the definition of  $A_1$  and the fact that the Jacobian of the transformation  $T^\epsilon(P)^{-1}$  is  $v_0^\epsilon$  and is bounded uniformly in  $p, \psi$ . Putting together the two estimates for regions  $A_1$  and  $A_2$  we obtain  $\|S^\epsilon - \tilde{S}\|_2 \leq G\epsilon^{1/2}$ . Estimates in terms of the other  $L_p$  norms follow similarly.

The homogenized operator given by (5.2) still contains variation of order  $\epsilon$  through the fast variable  $\frac{\psi}{\epsilon}$ , however there it does not contain any derivatives in that variable. Its dependence on  $\frac{\psi}{\epsilon}$  is only parametric. We can homogenize the dependence of the partially homogenized operator on  $\frac{\psi}{\epsilon}$  and arrive at a homogenized operator that is independent of the small scale. In the latter case, we will only obtain weak convergence of the partially homogenized solution. When we homogenized along the streamlines, the resulting equation was of hyperbolic type like the original equation. In a seminal and celebrated paper, Tartar [83] showed that homogenization across streamlines leads to transport with the average velocity plus a time-dependent diffusion term, referred to as macrodispersion, a physical phenomenon that was not present in the original fine equation. In particular, if the velocity field does not depend on  $p$  inside the cells, that is,  $\tilde{v}(\psi, \frac{\psi}{\epsilon})$ , then the homogenized solution,  $\bar{S}$ , (weak\* limit of  $\tilde{S}$ , which will be denoted by  $\bar{S}$ ), satisfies

$$\bar{S}_t + \bar{v}_0 \bar{S}_p = \int_0^t \int \bar{S}_{pp}(p - \lambda(t - \tau), \psi, \tau) d\mu_{\frac{\psi}{\epsilon}}(\lambda) d\tau. \quad (5.6)$$

Here,  $d\nu_{\frac{\psi}{\epsilon}}$  the Young measure associated with the sequence  $\tilde{v}_0(\psi, \cdot)$  and  $d\mu_{\frac{\psi}{\epsilon}}$  is a Young measure that satisfies

$$\left( \int \frac{d\nu_{\frac{\psi}{\epsilon}}(\lambda)}{\frac{s}{2\pi i q} + \lambda} \right)^{-1} = \frac{s}{2\pi i q} + \bar{v}_0 - \int \frac{d\mu_{\frac{\psi}{\epsilon}}(\lambda)}{\frac{s}{2\pi i q} + \lambda}.$$

We have denoted by  $\bar{v}_0$  the weak limit of the velocity. This equation has no dependence on the small scale and we consider it to be the full homogenization of the fine saturation equation. Efendiev and Popov [42] have extended this method for the Riemann problem in the case of nonlinear flux. Note that the homogenization across streamlines provides a weak limit of partially homogenized solution. Because the original solution  $S^\epsilon$  strongly converges to partially homogenized solution for each  $\psi$ , it can be easily shown that  $S^\epsilon \rightarrow \bar{S}$  weakly. We omit this proof here.

In numerical simulations, it is difficult to use (5.6) as a homogenized operator, and often a second order approximation of this equation is used. These approximate equations can be also derived using perturbation analysis. In particular, using the higher moments of the saturation and the velocity, one can model the macrodispersion. In the context of two-phase flow this idea was introduced by Efendiev, Durlafsky, and Lee [41], [40], Chen and Hou [22] and Hou et al., [49]. In our case, the computation of the macrodispersion is much simpler because the transport equations have been already averaged along the streamlines, and thus we will be applying perturbation technique to one dimensional problem.

We expand  $\tilde{S}$ ,  $\tilde{v}_0$  (following [41]) as an average over the cells in the pressure-streamline frame and the corresponding fluctuations

$$\begin{aligned} \tilde{S} &= \bar{S}(p, \psi, t) + S'(p, \psi, \zeta, t) \\ \tilde{v}_0 &= \bar{v}_0(p, \psi, t) + \tilde{v}'_0(p, \psi, \zeta, t). \end{aligned} \quad (5.7)$$

We will derive the homogenized equation for  $f(S) = S$ . Averaging equations (5.2) with respect to  $\psi$  we find an equation for the mean of the saturation

$$\bar{S}_t + \bar{v}_0 \bar{S}_p + \overline{\tilde{v}'_0 S'_p} = 0.$$

An equation for the fluctuations is obtained by subtracting the above equation from (5.2)

$$S'_t + (\tilde{v}_0 - \bar{v}_0) \bar{S}_p + \tilde{v}_0 S'_p - \overline{\tilde{v}'_0 S'_p} = 0.$$

Together, the equations for the saturation are

$$\begin{aligned} \bar{S}_t + \bar{v}_0 \bar{S}_p + \overline{\tilde{v}'_0 S'_p} &= 0 \\ S'_t + \tilde{v}'_0 \bar{S}_p + \tilde{v}_0 S'_p - \overline{\tilde{v}'_0 S'_p} &= 0. \end{aligned} \quad (5.8)$$

We can consider the second equation to be the auxiliary (cell) problem and the first equation to be the upscaled equation. We note that the cell problem for a hyperbolic equation is  $O(1)$  whereas for an elliptic it is  $O(\epsilon)$ . We can obtain an approximate numerical method by solving the cell problem only near the shock region in space time, where the macrodispersion term is largest. In that case it is best to diagonalize these equations by adding the first to the second one

$$\begin{aligned} \bar{S}_t + \bar{v}_0 \bar{S}_p &= -\overline{\tilde{v}'_0 (\tilde{S}_p - \bar{S}_p)} \\ \tilde{S}_t + \tilde{v}_0 \tilde{S}_p &= 0. \end{aligned}$$

Compared to (5.8), it has fewer forcing terms and no cross fluxes, which leads to a numerical method with less numerical diffusion that is easier to implement .

## 5.1 Numerical Averaging across Streamlines

The derivation in the previous sections contained no approximation. In this section, we follow the same idea as in the derivation to solve the equation for the fluctuations along the characteristics, but with the

purpose of deriving an equation on the coarse grid. To achieve this, we will not perform analytical upscaling in the sense of deriving a continuous upscaled equation as in the previous section. We will first discretize the equation with a finite volume method in space and then upscale the resulting equation. Our upscaled equation will therefore be dependent on the numerical scheme.

We use the same definition for the average saturation and the fluctuations as in (5.7) and follow the same steps until equation (5.8). We discretize the macrodispersion term in the equation for the average saturation

$$\overline{\tilde{v}'_0 S'_p} = \frac{\overline{\tilde{v}'_0 S'^{i+1}} - \overline{\tilde{v}'_0 S'^i}}{\Delta p} + O(\Delta p).$$

A superscript  $\cdot^i$  refers to a discrete quantity defined at the center of the conservation cell. Instead of solving the equation for the fluctuations on the fine characteristics as before, which would lead to a fine grid algorithm, we solve it on the coarse characteristics defined by

$$\frac{dP}{dt} = \overline{\tilde{v}_0}, \text{ with } P(p, 0) = p.$$

Compared to the equation that we obtained in the previous section for  $S'$ , this equation for  $S'$  has an extra term, which appears second

$$S' = - \int_0^t \left( \tilde{v}'_0(P(p, \tau), \psi) \overline{S}_p(P(p, \tau), \psi, \tau) + \tilde{v}'_0(P(p, \tau), \psi) S'_p(P(p, \tau), \psi, \tau) + \overline{\tilde{v}'_0 S'_p} \right) d\tau.$$

The second term is second-order in fluctuating quantities, and we expect it to be smaller than the first term so we neglect it. As before, we multiply by  $\tilde{v}'_0$  and average over  $\psi$  to find

$$\overline{\tilde{v}'_0 S'} = - \int_0^t \overline{\tilde{v}'_0 \tilde{v}_0(P(p, \tau), \psi) \overline{S}_p(P(p, \tau), \psi, \tau)} d\tau.$$

In this form at time  $t$  it is necessary to know information about the past saturation in  $(0, t)$  to compute the future saturation. Following [41], it can be easily shown that  $\overline{S}_p(P(p, \tau))$  depends weakly on time, in the sense that the difference between  $\overline{S}_p(P(p, \tau))$  and  $\overline{S}_p(P(p, t))$  is of third-order in fluctuating quantities. Therefore we can take  $\overline{S}_p(P(p, \tau))$  out of the time integral to find

$$\overline{\tilde{v}'_0 S'} = - \int_0^t \overline{\tilde{v}'_0 \tilde{v}_0(P(p, \tau), \psi)} d\tau \overline{S}_p.$$

The term inside the time integral is the covariance of the velocity field along each streamline. The macrodispersion in this form can be computed independent of the past saturation.

The nonlinearity of the flux function introduces an extra source of error in the approximation. We expand  $f(\tilde{S})$  near  $\overline{S}$  (cf. [40]) and keep only the first term

$$\begin{aligned} \tilde{S} &= \overline{S}(p, \psi, t) + S'(p, \psi, \zeta, t) \\ \tilde{v}_0 &= \overline{v}_0(p, \psi, t) + \tilde{v}'_0(p, \psi, \eta, t) \\ \overline{f}(\tilde{S}) &= f(\overline{S}) + f_S(\overline{S}) S' + O(S'^2) \\ f(S)_p &= f_S(\overline{S}) \overline{S}_p + f(\overline{S}) S' + \dots \end{aligned} \tag{5.9}$$

This approximation is not accurate near the shock because  $S'$  is not small near sharp fronts. The region near the shock is important because the macrodispersion is large. Due to the dependence of the jump in the saturation on the mobility we expect this approximation to be better for lower mobilities. Nevertheless this approximation works well in practice. For more accuracy, it is also possible to retain more terms in the Taylor expansion. We will show that in realistic examples these higher-order terms are not important in our setting.

Using these definitions we derive the following equations for the average saturation and the fluctuations (see [79] for more details)

$$\begin{aligned}\overline{S}_t + \overline{v_0 f(\overline{S})}_p + \overline{v'_0(f_S(\overline{S})S')}_p &= 0 \\ S'_t + \overline{v'_0 f_S(\overline{S})\overline{S}}_p + \overline{v_0 f_S(\overline{S})S'}_p - \overline{v'_0 S'}_p &= 0.\end{aligned}\tag{5.10}$$

The macrodispersion is discretized as

$$\overline{v'_0(f_S(\overline{S})S')}_p = \frac{\overline{v'_0 f_S(\overline{S})S'}^{i+1} - \overline{v'_0 f_S(\overline{S})S'}^i}{\Delta p} + O(\Delta p).$$

We solve the second equation on the coarse characteristics defined by

$$\frac{dP}{dt} = \overline{v_0 f_S(\overline{S})}, \text{ with } P(p, 0) = p$$

and form the terms that appear in the macrodispersion

$$\overline{v'_0 f_S(\overline{S})S'} = - \int_0^t \overline{v'_0 f_S(\overline{S})v'_0(P(p, \tau), \psi) f_S(\overline{S}(P(p, \tau), \psi, \tau))\overline{S}_p(P(p, \tau), \psi, \tau)} d\tau.$$

As before we have dropped terms that are second-order in fluctuating quantities. It can be shown (see [79]) that  $f_S(\overline{S}(P(p, \tau), \psi, \tau))\overline{S}_p(P(p, \tau), \psi, \tau)$  does not vary significantly along the streamlines and it can be taken out of the integration in time:

$$\overline{v'_0 f_S(\overline{S})S'} = - \int_0^t \overline{v'_0 v'_0(P(p, \tau), \psi) d\tau f_S(\overline{S})^2 \overline{S}_p}.\tag{5.11}$$

This expression is similar to the one obtained in the linear case, however the macrodispersion depends on the past saturation through the equation for the coarse characteristics.

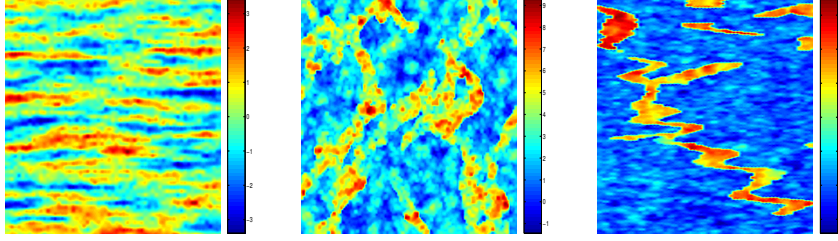
Even though the macrodispersion depends on the past saturation it is possible to compute it incrementally as it is done in [40]. Given its value  $D(t)$  at time  $t$  we compute the values at  $t + \Delta t$  using the macrodispersion at the previous time

$$D(t + \Delta t) = \int_0^{t+\Delta t} \dots d\tau = \int_0^t \dots d\tau + \int_t^{t+\Delta t} \dots d\tau.$$

This is possible because in the derivation for the approximate expression for the macrodispersion we took the terms that depend on  $S(\tau)$  outside the time integration. The integrand, the average covariance of the velocity field along the streamlines, needs to be computed only once at the beginning. Then updating the macrodispersion takes  $O(n^2)$  computations, as many as it takes to update  $\overline{S}$ .

## 5.2 Numerical Results

In this section, we first show representative simulation results for  $\lambda(S) = 1$  for flux functions  $f(S) = S$  and nonlinear  $f(S)$  with viscosity ratio  $\mu_o/\mu_w = 5$ . For such setting, the pressure and saturation equations are decoupled and we can investigate the accuracy of saturation upscaling independently from the pressure upscaling. At the end of the section we will present numerical results for two-phase flow. We consider two type of permeability fields. The first type includes a permeability field generated using two-point geostatistics with correlation lengths  $l_x = 0.3$ ,  $l_z = 0.03$  and  $\sigma^2 = 1.5$  (see Figure 5.1, left). The second type of permeability fields correspond to a channelized system, and we consider two examples. The first example (middle figure of Figure 5.1) is a synthetic channelized reservoir generated using both multi-point geostatistics (for the channels) and two-point geostatistics (for permeability distribution within each facies). The second channelized system is one of the layers of the benchmark test (representing the North Sea reservoir), the SPE comparative project [23] (upper Ness layers). These permeability fields are highly heterogeneous, channelized, and



**Fig. 5.1.** Permeability fields used in the simulations. Left - permeability field with exponential variogram, middle - synthetic channelized permeability field, right - layer 36 of SPE comparative project [23]

difficult to upscale. Because the permeability fields are highly heterogeneous, they are refined to  $400 \times 400$  in order to obtain accurate comparisons.

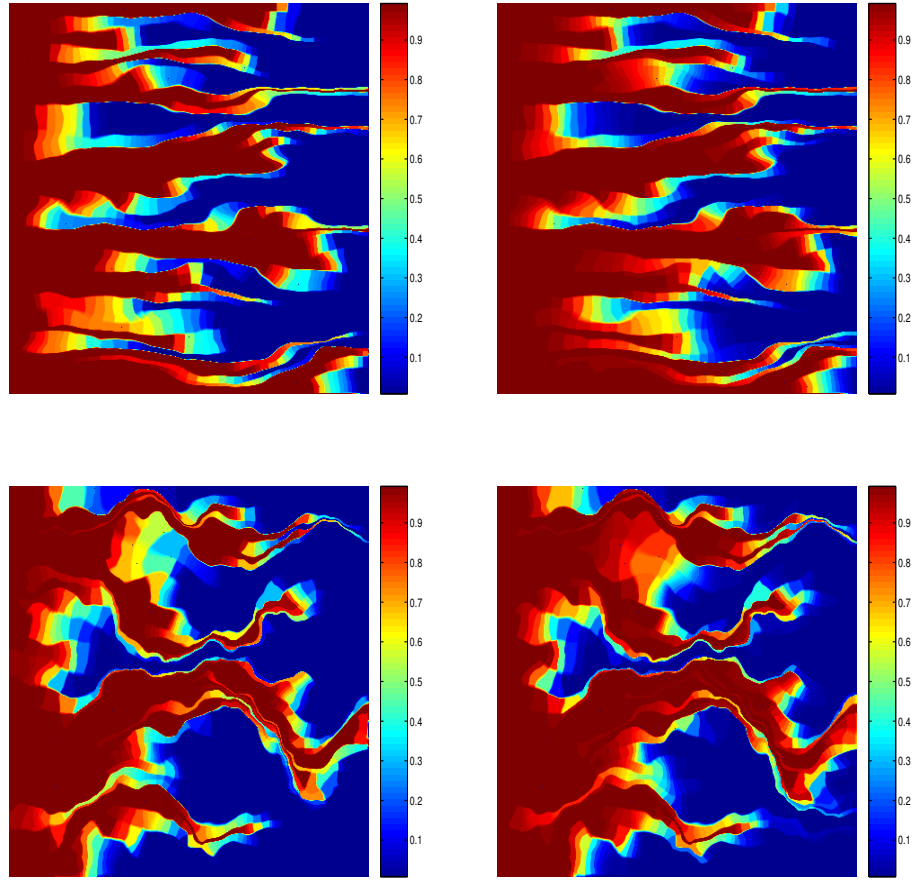
Simulation results will be presented for saturation snapshots as well as the oil cut as a function of pore volume injected (PVI). Note that the oil cut is also referred to as the fractional flow of oil. The oil cut (or fractional flow) is defined as the fraction of oil in the produced fluid and is given by  $q_o/q_t$ , where  $q_t = q_o + q_w$ , with  $q_o$  and  $q_w$  being the flow rates of oil and water at the production edge of the model. In particular,  $q_w = \int_{\partial\Omega^{out}} f(S)\mathbf{v} \cdot \mathbf{n}dl$ ,  $q_t = \int_{\partial\Omega^{out}} \mathbf{v} \cdot \mathbf{n}dl$ , and  $q_o = q_t - q_w$ , where  $\partial\Omega^{out}$  is the outer flow boundary. We will use the notation  $\Omega$  for total flow  $q_t$  and  $F$  for fractional flow  $q_o/q_t$  in numerical results. Pore volume injected, defined as  $PVI = \frac{1}{V_p} \int_0^t q_t(\tau)d\tau$ , with  $V_p$  being the total pore volume of the system, provide a dimensionless time for the displacement.

When using multiscale finite element methods for two-phase flow, one can update the basis functions near the sharp fronts. Indeed, sharp fronts modify the local heterogeneities and this can be taken into account by re-solving the local equations, (3.34), for basis functions. If the saturation is smooth in the coarse block, it can be approximated by its average in (3.34), and consequently, the basis functions do not need to be updated. It can be shown that this approximation yields first-order errors (in terms of coarse mesh size). In our simulations, we have found only a slight improvement when the basis functions are updated, thus the numerical results for the MsFVEM presented in this paper do not include the basis function update near the sharp fronts. Since a pressure-streamline coordinate system is used the boundary conditions are given by  $P = 1$ ,  $S = 1$  along the  $p = 1$  edge and  $P = 0$  along the  $p = 0$  edge, and no flow boundary condition on the rest of the boundaries.

For the upscaled saturation equation, which is a convection-diffusion equation, we need to observe an extra CFL-like condition to obtain a stable numerical scheme  $\Delta t \leq \frac{\Delta p^2}{2\nu}$ , where  $\nu$  is the diffusivity. In our case the diffusivity is  $\int_{cell} \int_0^t \tilde{v}'_0(p(\tau), \psi) \tilde{v}'_0(p, \psi) d\tau d\psi$ . If the macrodispersion is large this can be a very restrictive condition. To remedy this, we used an implicit discretization for the macrodispersion. This is straightforward since the problem is one-dimensional. The resulting system was solved by a tridiagonal solver very fast. Since the order of the highest derivative in the equation has increased, we require extra boundary conditions. For the computation of the macrodispersion term, we impose no flux on both boundaries of the domain.

In the upscaled algorithm, a moving mesh is used to concentrate the points of computation near the sharp front. Since the saturation equation is one dimensional in the pressure-streamline coordinates, the implementation of the moving mesh is straightforward and efficient. For the details we refer to [79]. We compare the saturation right before the breakthrough time so that the shock front is largest. For this comparison we also average the fine saturation over the coarse blocks, since the upscaled model is defined on a coarser grid. In Figures 5.2 5.3, we plot the saturation for linear and nonlinear (with  $\mu_o/\mu_w = 5$ )  $f(S)$ . As we see in both cases, we have very accurate representation of the saturation profile.

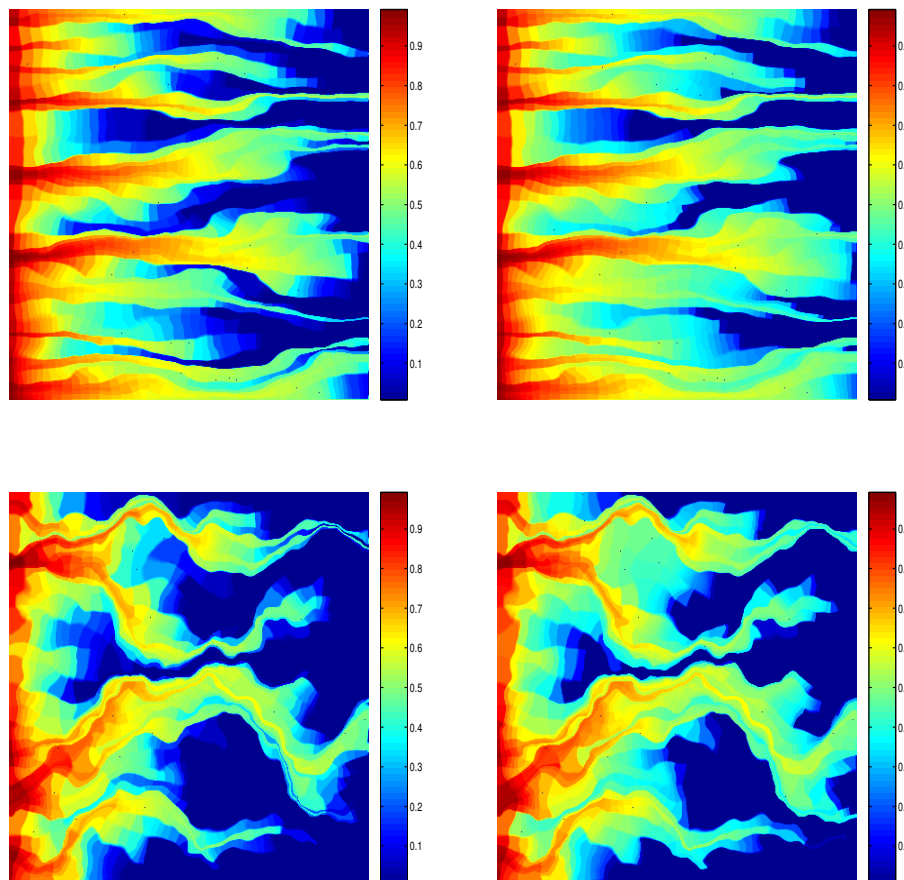
We proceed with a quantitative description of the error. We will distinguish between two sources of errors. We will refer to the difference between the upscaled and the exact equation as the upscaling or modeling error and to the difference between the solution of continuous upscaled equations and the solution to the numerical scheme as the discretization error. We will refer to the difference between the solutions of the



**Fig. 5.2.** Saturation snapshots for variogram based permeability field (top) and synthetic channelized permeability field (bottom). Linear flux is used. Left figures represent the upscaled saturation plots and the right figures represent the fine-scale saturation plots.

continuous fine equations and the numerical scheme of the upscaled equations as the total error. To separate the upscaling error from the total error we will solve the upscaled equations on the fine grid, which is the grid on which we solve to the fine equation. We will also solve them on the coarse grid to compute the total error. The errors are computed in the  $p, \psi$  frame and are relative errors. We display the upscaling error against the number of coarse cells for the computations of the previous section in Tables 5.14, 5.15, 5.16. As we see from this table that upscaling using macrodispersion decreases the upscaling errors. We also see that the effects of macrodispersion are more significant in the case of linear flux when the jump discontinuity in the saturation profile is larger.

In Tables 5.17, 5.18, 5.19, we show the total error, that is, the modeling and discretization error when we use a moving mesh to solve the saturation equation. It is interesting that the convergence of  $\tilde{S}$  to  $S$  is observed even though the upscaling error is larger than the numerical error of the fine solution, which is 0.02 for the linear flux and 0.002 for the nonlinear flux in the  $L_1$  norm, as mentioned before. The reason is that the location of the moving mesh points was selected so that the points are as dense near the shock as the fine solution using the parameter  $h_{min}$ . This was done to observe the upscaling error clearly and also to have similar CFL constraints on the time step, which allows a clean comparison of computational times. We compare the require CPU times in Table 5.20. We note that it took 26 units of time to interpolate one quantity from the Cartesian to the pressure-streamline frame. The upscaled solutions were computed on a



**Fig. 5.3.** Saturation snapshots for variogram based permeability field (top) and synthetic channelized permeability field (bottom). Nonlinear flux is used. Left figures represent the upscaled saturation plots and the right figures represent the fine-scale saturation plots.

**Table 5.14.** Upscaling error for permeability generated using two-point geostatistics

LINEAR FLUX	25x25	50x50	100x100	200x200
$L_1$ error of $\bar{S}$	0.0021	$6.57 \times 10^{-4}$	$2.15 \times 10^{-4}$	$8.75 \times 10^{-5}$
$L_1$ error of $\bar{S}$ with macrodispersion	0.115	0.0696	0.0364	0.0135
$L_1$ error of $\bar{S}$ fine without macrodispersion	0.1843	0.0997	0.0505	0.0191
NONLINEAR FLUX	25x25	50x50	100x100	200x200
$L_1$ error of $\bar{S}$	0.0023	$8.05 \times 10^{-4}$	$2.89 \times 10^{-4}$	$1.29 \times 10^{-4}$
$L_1$ error of $\bar{S}$ with macrodispersion	0.116	0.0665	0.0433	0.0177
$L_1$ error of $\bar{S}$ fine without macrodispersion	0.151	0.0805	0.0432	0.0186



**Table 5.15.** Upscaling error for for synthetic channelized permeability field

LINEAR FLUX	25x25	50x50	100x100	200x200
$L_1$ error of $\bar{S}$	0.0222	0.0171	0.0122	0.0053
$L_1$ error of $\bar{S}$ with macrodispersion	0.0819	0.0534	0.0333	0.0178
$L_1$ error of $\bar{S}$ fine without macrodispersion	0.123	0.0834	0.0486	0.0209

NONLINEAR FLUX	25x25	50x50	100x100	200x200
$L_1$ error of $\bar{S}$	0.0147	0.0105	0.0075	0.0040
$L_1$ error of $\bar{S}$ with macrodispersion	0.0842	0.0658	0.0371	0.0207
$L_1$ error of $\bar{S}$ fine without macrodispersion	0.119	0.0744	0.0424	0.0214

**Table 5.16.** Upscaling error for SPE 10, layer 36

LINEAR FLUX	25x25	50x50	100x100	200x200
$L_1$ error of $\bar{S}$	0.0128	0.0093	0.0072	0.0042
$L_1$ error of $\bar{S}$ with macrodispersion	0.0554	0.0435	0.0307	0.0176
$L_1$ error of $\bar{S}$ fine without macrodispersion	0.123	0.0798	0.0484	0.0258

NONLINEAR FLUX	25x25	50x50	100x100	200x200
$L_1$ error of $\bar{S}$	0.0089	0.0064	0.0054	0.0033
$L_1$ error of $\bar{S}$ with macrodispersion	0.0743	0.0538	0.0348	0.0189
$L_1$ error of $\bar{S}$ fine without macrodispersion	0.0924	0.0602	0.0395	0.0202

$25 \times 25$  grid and the fine solution was computed on a  $400 \times 400$  grid so we expect the  $\bar{S}$  computations to be 256 times or more faster. The extra gain comes from a less restrictive CFL condition since we use an averaged velocity. The computations in the Cartesian frame are much slower.

**Table 5.17.** Total error for permeability field generated using two-point geostatistics

LINEAR FLUX	25x25	50x50	100x100	200x200
$L_1$ upscaling error of $\bar{S}$	0.0021	$6.57 \times 10^{-4}$	$2.15 \times 10^{-4}$	$8.75 \times 10^{-5}$
$L_1$ error of $\bar{S}$ computed on coarse grid	0.0185	0.0062	0.0019	0.0015
$L_1$ upscaling error of $\bar{S}$	0.115	0.0696	0.0364	0.0135
$L_1$ error of $\bar{S}$ computed on coarse grid	0.139	0.0779	0.0390	0.0144

NONLINEAR FLUX	25x25	50x50	100x100	200x200
$L_1$ upscaling error of $\bar{S}$	0.0023	$8.05 \times 10^{-4}$	$2.89 \times 10^{-4}$	$1.29 \times 10^{-4}$
$L_1$ error of $\bar{S}$ computed on coarse grid	0.0268	0.0099	0.0027	$9.38 \times 10^{-4}$
$L_1$ upscaling error of $\bar{S}$	0.116	0.0665	0.0433	0.0177
$L_1$ error of $\bar{S}$ computed on coarse grid	0.146	0.0797	0.0461	0.0184

The application of the proposed method to two-phase immiscible flow can be performed using the implicit pressure and explicit saturation (IMPES) framework. This procedure consists of computing the velocity and then using the velocity field in updating the saturation field. When updating the saturation field, we consider the velocity field to be time independent and we can use our upscaling procedure at each IMPES time step. First, we note that in the proposed method, the mapping is done between the current pressure-streamline and initial pressure-streamline. This mapping is nearly the identity for the cases when  $\mu_o > \mu_w$ . In Figure 5.4, we plot the level sets of the pressure and streamfunction at time  $t = 0.4$  in a Cartesian coordinate system (left plot) and in the coordinate system of the initial pressure and streamline (right plot). Clearly, the level sets are much smoother in initial pressure-streamline frame compared to Cartesian frame. This also

**Table 5.18.** Total error for synthetic channelized permeability field

LINEAR FLUX	25x25	50x50	100x100	200x200
$L_1$ upscaling error of $\tilde{S}$	0.0222	0.0171	0.0122	0.0053
$L_1$ error of $\tilde{S}$ computed on coarse grid	0.0326	0.0161	0.0107	0.0113
$L_1$ upscaling error of $\tilde{S}$	0.0819	0.0534	0.0333	0.0178
$L_1$ error of $\tilde{S}$ computed on coarse grid	0.135	0.0849	0.0477	0.0274

NONLINEAR FLUX	25x25	50x50	100x100	200x200
$L_1$ upscaling error of $\tilde{S}$	0.0147	0.0105	0.0075	0.0040
$L_1$ error of $\tilde{S}$ computed on coarse grid	0.0494	0.0295	0.0150	0.0130
$L_1$ upscaling error of $\tilde{S}$	0.0842	0.0658	0.0371	0.0207
$L_1$ error of $\tilde{S}$ computed on coarse grid	0.17	0.11	0.0541	0.0303

**Table 5.19.** Total error for SPE10 layer 36

LINEAR FLUX	25x25	50x50	100x100	200x200
$L_1$ upscaling error of $\tilde{S}$	0.0128	0.0093	0.0072	0.0042
$L_1$ error of $\tilde{S}$ computed on coarse grid	0.023	0.0095	0.0069	0.0052
$L_1$ upscaling error of $\tilde{S}$	0.0554	0.0435	0.0307	0.0176
$L_1$ error of $\tilde{S}$ computed on coarse grid	0.0683	0.052	0.0361	0.0205

NONLINEAR FLUX	25x25	50x50	100x100	200x200
$L_1$ upscaling error of $\tilde{S}$	0.0089	0.0064	0.0054	0.0033
$L_1$ error of $\tilde{S}$ computed on coarse grid	0.0338	0.0148	0.0074	0.0037
$L_1$ upscaling error of $\tilde{S}$	0.0743	0.0538	0.0348	0.0189
$L_1$ error of $\tilde{S}$ computed on coarse grid	0.115	0.0720	0.0406	0.0204

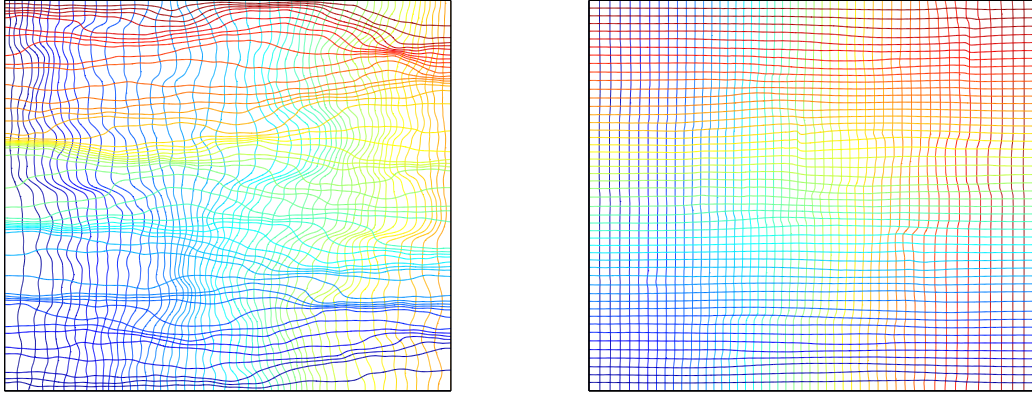
**Table 5.20.** Computational cost

	fine $x,y$	fine $p, \psi$	$\tilde{S}$	$\tilde{S}$
layered, linear flux	5648	257	9	1
layered, nonlinear flux	14543	945	28	4
percolation, linear flux	8812	552	12	1
percolation, nonlinear flux	23466	579	12	1
SPE10 36, linear flux	40586	1835	34	2
SPE10 36, nonlinear flux	118364	7644	25	2

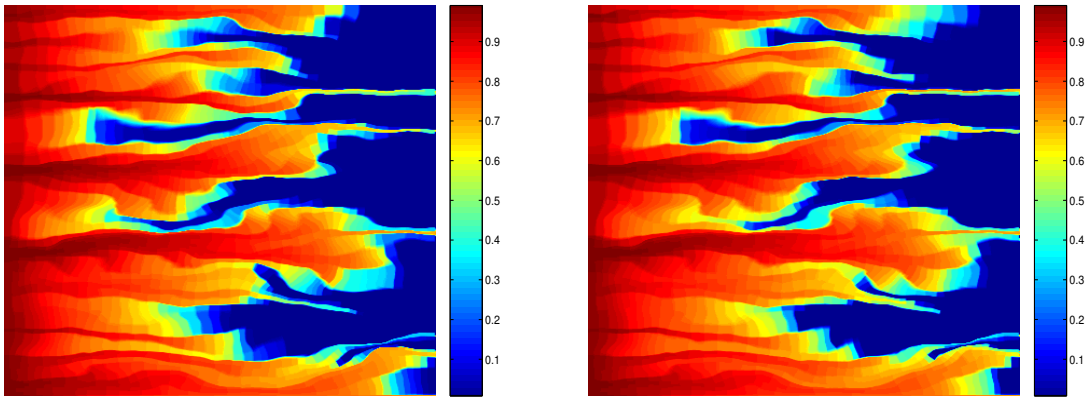
explains the observed convergence of upscaling methods as we refined the coarse grid. In Figure 5.5, we plot the saturation snapshots right before the breakthrough. In Figure 5.6, the fractional flow is plotted. Again, the moving mesh algorithm is used to track the front separately. The convergence table is presented in Table 5.21. We see from this table that the errors decrease as first order which indicates that the pressure and saturation is smooth functions of initial pressure and streamline.

## 6 Conclusions

In these lecture notes, we reviewed some of the recent advances in developing systematic multiscale methods with particular emphasis on multiscale finite element methods and their applications to fluid flows in heterogeneous porous media. In particular, the local approaches and their convergence properties for various flow problems are discussed. Moreover, improved subgrid capturing techniques through a judicious choice of local boundary conditions or through oversampling techniques or through the use of limited global information are reviewed. Other topics, such as homogenization, the sampling techniques in numerical homogenization,



**Fig. 5.4.** Left: Pressure and streamline function at time  $t = 0.4$  in Cartesian frame. Right: pressure and streamline function at time  $t = 0.4$  in initial pressure-streamline frame.

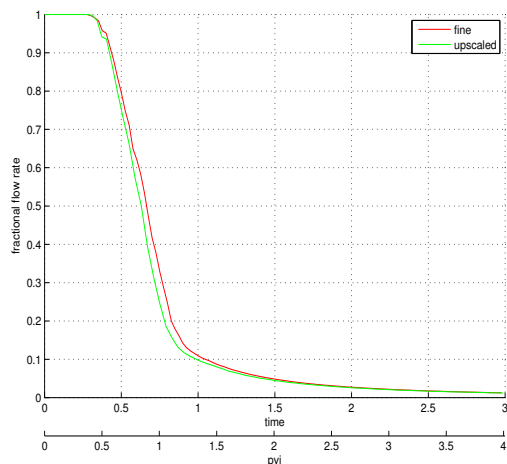


**Fig. 5.5.** Left: Saturation plot obtained using coarse-scale model. Right: The fine-scale saturation plot. Both plots are on coarse grid. Variogram based permeability field is used.  $\mu_o/\mu_w = 5$ .

**Table 5.21.** Convergence of the upscaling method for two-phase flow for variogram based permeability field

with $\bar{S}$	50x50	100x100	200x200
$L_2$ pressure error at $t = \frac{3T_{final}}{4}$	0.0014	0.007	0.004
$L_2$ velocity error at $t = \frac{3T_{final}}{4}$	0.0235	0.0137	0.0072
$L_1$ saturation error $t = T_{final}$	0.0105	0.0052	0.0027

with $\bar{S}$	50x50	100x100	200x200
$L_2$ pressure error at $t = \frac{3T_{final}}{4}$	0.0046	0.0021	0.0008
$L_2$ velocity error at $t = \frac{3T_{final}}{4}$	0.0530	0.0335	0.0246
$L_1$ saturation error $t = T_{final}$	0.0546	0.0294	0.0134



**Fig. 5.6.** Comparison of fractional flow for coarse- and fine-scale models. Variogram based permeability field is used.  $\mu_o/\mu_w = 5$ .

and multiscale simulations of two-phase flows in heterogeneous porous media are also presented. Although the results presented in this paper are encouraging, there is scope for further exploration. These include the development and mathematical analysis of efficient numerical homogenization techniques for nonlinear convection-diffusion equations with various Peclet numbers (e.g., convection dominated), inexpensive approximations of multiscale basis functions, further exploration of accurate boundary conditions based on local multiscale solutions, the use of limited global information for nonlinear problems, development of adaptive criteria for multiscale basis functions (selection of coarse grid), applications of MsFEM to more multi-phase/multi-component porous media flows, and etc.

## References

1. J. Aarnes, *On the use of a mixed multiscale finite element method for greater flexibility and increased speed or improved accuracy in reservoir simulation*, SIAM MMS, 2 (2004), pp. 421–439.
2. J. Aarnes, Y. R. Efendiev, and L. Jiang, *Analysis of multiscale finite element methods using global information for two-phase flow simulations*. submitted.
3. 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.
4. R. A. ADAMS, *Sobolev spaces*, Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Pure and Applied Mathematics, Vol. 65.
5. 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.
6. T. Arbogast, *Implementation of a locally conservative numerical subgrid upscaling scheme for two-phase Darcy flow*, Comput. Geosci., 6 (2002), pp. 453–481. Locally conservative numerical methods for flow in porous media.
7. T. Arbogast and K. Boyd, *Subgrid upscaling and mixed multiscale finite elements*. to appear in SIAM Num. Anal.
8. M. AVELLANEDA AND F.-H. LIN, *Compactness methods in the theory of homogenization*, Comm. Pure Appl. Math., 40 (1987), pp. 803–847.
9. I. Babuska, U. Banerjee, and J. E. Osborn, *Survey of meshless and generalized finite element methods: A unified approach*, Acta Numerica, 2003, pp. 1-125.
10. 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.
11. I. Babuška and J. M. Melenk, *The partition of unity method*, Internat. J. Numer. Methods Engrg., 40 (1997), pp. 727–758.
12. 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.

13. 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.
14. 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.
15. 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.
16. A. Bourgeat and A. Mikelić, *Homogenization of two-phase immiscible flows in a one-dimensional porous medium*, *Asymptotic Anal.*, 9 (1994), pp. 359-380.
17. M. Brewster and G. Beylkin, *A Multiresolution Strategy for Numerical Homogenization*, *ACHA*, **2**(1995), 327-349.
18. F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Springer-Verlag, Berlin – Heidelberg – New-York, 1991.
19. F. Brezzi and A. Russo, *Choosing Bubbles for Advection-Diffusion Problems*, *Math. Models Methods Appl. Sci*, **4** (1994), 571-587.
20. 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.
21. J. E. Broadwell, *Shock Structure in a Simple Discrete Velocity Gas*, *Phys. Fluids*, **7** (1964), 1243-1247.
22. Z. Chen and T. Y. Hou, *A mixed multiscale finite element method for elliptic problems with oscillating coefficients*, *Math. Comp.*, 72 (2002), pp. 541-576 (electronic).
23. M. Christie and M. Blunt, *Tenth SPE comparative solution project: A comparison of upscaling techniques*, *SPE Reser. Eval. Eng.*, 4 (2001), pp. 308-317.
24. 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.
25. J. E. Dendy, J. M. Hyman, and J. D. Moulton, *The Black Box Multigrid Numerical Homogenization Algorithm*, *J. Comput. Phys.*, **142** (1998), 80-108.
26. C. V. DEUTSCH AND A. G. JOURNAL, *GSLIB: Geostatistical software library and user's guide, 2nd edition*, Oxford University Press, New York, 1998.
27. M. Dorobantu and B. Engquist, *Wavelet-based Numerical Homogenization*, *SIAM J. Numer. Anal.*, **35** (1998), 540-559.
28. 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.
29. L. J. Durlofsky, *Numerical Calculation of Equivalent Grid Block Permeability Tensors for Heterogeneous Porous Media*, *Water Resour. Res.*, **27** (1991), 699-708.
30. 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.
31. 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.
32. W. E, *Homogenization of linear and nonlinear transport equations*, *Comm. Pure Appl. Math.*, XLV (1992), pp. 301-326.
33. W. E and B. Engquist, *The heterogeneous multi-scale methods*, *Comm. Math. Sci.*, 1(1) (2003), pp. 87-133.
34. Y. R. Efendiev, *Multiscale Finite Element Method (MsFEM) and its Applications*, Ph. D. Thesis, Applied Mathematics, Caltech, 1999.
35. Y. Efendiev, V. Ginting, T. Hou, and R. Ewing, *Accurate multiscale finite element methods for two-phase flow simulations*. to appear in *Comp. Physics*.
36. Y. Efendiev, T. Hou, and V. Ginting, *Multiscale finite element methods for nonlinear problems and their applications*, *Comm. Math. Sci.*, 2 (2004), pp. 553-589.
37. 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.
38. Y. Efendiev and A. Pankov, *Homogenization of nonlinear random parabolic operators*, *Advances in Differential Equations*, vol. 10, Number 11, 2005, pp., 1235-1260
39. Y. Efendiev and A. Pankov, *Numerical homogenization of nonlinear random parabolic operators*, *SIAM Multiscale Modeling and Simulation*, 2(2) (2004), pp. 237-268.
40. Y. R. Efendiev and L. J. Durlofsky, *Numerical modeling of subgrid heterogeneity in two phase flow simulations*, *Water Resour. Res.*, 38(8) (2002), p. 1128.
41. 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.

42. Y. R. Efendiev and B. Popov, *On homogenization of nonlinear hyperbolic equations*, Communications on Pure and Applied Analysis, 4(2) (2005), pp. 295–309.
43. R. Eymard, T. Gallouët, and R. Herbin, *Finite volume methods*, in Handbook of numerical analysis, Vol. VII, Handb. Numer. Anal., VII, North-Holland, Amsterdam, 2000, 713–1020.
44. J. Fish and K.L. Shek, *Multiscale Analysis for Composite Materials and Structures*, Composites Science and Technology: An International Journal, **60** (2000), 2547-2556.
45. J. Fish and Z. Yuan, *Multiscale enrichment based on the partition of unity*, International Journal for Numerical Methods in Engineering, **62**, (2005), 1341–1359.
46. D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*. Springer, Berlin, New York, 2001.
47. 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.
48. T. Hou, X. Wu, and Y. Zhang, *Removing the cell resonance error in the multiscale finite element method via a petrov-galerkin formulation*, Communications in Mathematical Sciences, **2**(2) (2004), 185–205.
49. T. Y. Hou, A. Westhead, and D. P. Yang, *A framework for modeling subgrid effects for two-phase flows in porous media*. to appear in SIAM Multiscale Modeling and Simulation.
50. 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.
51. 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.
52. 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.
53. T. Y. Hou, D.-P. Yang, and K. Wang, *Homogenization of Incompressible Euler Equation*. J. Comput. Math., **22** (2004), 220-229.
54. T. Y. Hou, D. P. Yang, and H. Ran, *Multiscale Analysis in the Lagrangian Formulation for the 2-D Incompressible Euler Equation*, Discrete and Continuous Dynamical Systems, **13** (2005), 1153-1186.
55. T. Y. Hou, D.-P. Yang, and H. Ran, *Multiscale Computation of Isotropic Homogeneous Turbulent Flow*, in Inverse Problems, Multi-Scale Analysis and Effective Medium Theory, Contemporary Mathematics, Vol. 408, pp. 111-135, 2006, ed. H. Ammari and H. Kang.
56. T. Y. Hou and X. Xin, *Homogenization of linear transport equations with oscillatory vector fields*, SIAM J. Appl. Math., 52 (1992), pp. 34–45.
57. Y. Huang and J. Xu, *A partition-of-unity finite element method for elliptic problems with highly oscillating coefficients*.
58. 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.
59. 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.
60. M. GERRITSEN AND L. J. DURLOFSKY, *Modeling of fluid flow in oil reservoirs*, Annual Reviews in Fluid Mechanics, 37 (2005), pp. 211–238.
61. P. Jenny, S. H. Lee, and H. Tchelepi, *Adaptive multi-scale finite volume method for multi-phase flow and transport in porous media*, Multiscale Modeling and Simulation, 3 (2005), pp. 30–64.
62. V. V. Jikov, S. M. Kozlov, and O. A. Oleinik, *Homogenization of Differential Operators and Integral Functionals*, Springer-Verlag, 1994, Translated from Russian.
63. Ioannis G. Kevrekidis, C. William Gear, James M. Hyman, Panagiotis G. Kevrekidis, Olof Runborg, and Constantinos Theodoropoulos, *Equation-free, coarse-grained multiscale computation: enabling microscopic simulators to perform system-level analysis*, Commun. Math. Sci. **1** (2003), no. 4, 715–762.
64. S. Knapek, *Matrix-Dependent Multigrid-Homogenization for Diffusion Problems*, in the Proceedings of the Copper Mountain Conference on Iterative Methods, edited by T. Manteuffel and S. McCormick, volume I, SIAM Special Interest Group on Linear Algebra, Cray Research , 1996.
65. P. Langlo and M.S. Espedal, *Macrodispersion for Two-phase, Immiscible Flow in Porous Media*, Adv. Water Resources **17** (1994), 297–316.
66. A. M. Matache, I. Babuska, and C. Schwab, *Generalized p-FEM in Homogenization*, Numer. Math. **86**(2000), 319-375.
67. A. M. Matache and C. Schwab, *Homogenization via p-FEM for Problems with Microstructure*, Appl. Numer. Math. **33** (2000), 43-59.

68. J. F. McCarthy, *Comparison of Fast Algorithms for Estimating Large-Scale Permeabilities of Heterogeneous Media*, *Transport in Porous Media*, **19** (1995), 123-137.
69. D. W. McLaughlin, G. C. Papanicolaou, and O. Pironneau, *Convection of Microstructure and Related Problems*, *SIAM J. Applied Math*, **45** (1985), 780-797.
70. 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.
71. H. Owhadi and L. Zhang, *Homogenization of parabolic equations with a continuum of space and time scales*. submitted.
72. ———, *Metric based up-scaling*. to appear in *Comm. Pure and Applied Math*.
73. A. Pankov, *G-convergence and homogenization of nonlinear partial differential operators*, Kluwer Academic Publishers, Dordrecht, 1997.
74. W. V. Petryshyn, *On the approximation-solvability of equations involving A-proper and pseudo-A-proper mappings*, *Bull. Amer. Math. Soc.*, **81** (1975), pp. 223-312.
75. O. Pironneau, *On the Transport-diffusion Algorithm and its Application to the Navier-Stokes Equations*, *Numer. Math.* **38** (1982), 309-332.
76. R.E. Rudd and J.Q. Broughton, *Coarse-grained molecular dynamics and the atomic limit of finite elements*, *Phys. Rev. B* **58**, R5893 (1998).
77. G. Sangalli, *Capturing Small Scales in Elliptic Problems Using a Residual-Free Bubbles Finite Element Method*, *Multiscale Modeling and Simulation*, **1** (2003), no. 3, 485-503
78. I. V. Skrypnik, *Methods for analysis of nonlinear elliptic boundary value problems*, vol. 139 of *Translations of Mathematical Monographs*, American Mathematical Society, Providence, RI, 1994. Translated from the 1990 Russian original by Dan D. Pascali.
79. T. Srinopoulos, *Upscaling of immiscible two-phase flows in an adaptive frame*, PhD thesis, California Institute of Technology, Pasadena, 2005.
80. T. Strouboulis, I. Babuška, and K. Copps, *The design and analysis of the generalized finite element method*, *Comput. Methods Appl. Mech. Engrg.*, **181** (2000), pp. 43-69.
81. 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.
82. L. Tartar, *Solutions oscillantes des équations de Carleman*, *Seminaire Goulaouic-Meyer-Schwartz (1980-1981)*, exp. XII. Ecole Polytechnique (Palaiseau), 1981.
83. L. Tartar, *Nonlocal Effects Induced by Homogenization*, in *PDE and Calculus of Variations*, ed by F. Cullumbini, et al, Birkhäuser, Boston, 925-938, 1989.
84. 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.
85. P. M. De Zeeuw, *Matrix-dependent Prolongation and Restrictions in a Blackbox Multigrid Solver*, *J. Comput. Applied Math*, **33**(1990), 1-27.
86. S. Verdier 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.
87. 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.
88. 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.
89. 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.
90. E. Zeidler, *Nonlinear functional analysis and its applications. II/B*, Springer-Verlag, New York, 1990. *Nonlinear monotone operators*, Translated from the German by the author and Leo F. Boron.