MULTISCALE SIMULATION AND UNCERTAINTY QUANTIFICATION TECHNIQUES FOR RICHARDS' EQUATION IN HETEROGENEOUS MEDIA

A Dissertation

by

SEUL KI KANG

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2012

Major Subject: Mathematics

MULTISCALE SIMULATION AND UNCERTAINTY QUANTIFICATION TECHNIQUES FOR RICHARDS' EQUATION IN HETEROGENEOUS MEDIA

A Dissertation

by

SEUL KI KANG

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee,	Yalchin Efendiev
	Raytcho Lazarov
Committee Members,	Bani Mallick
	Binayak P. Mohanty
Head of Department,	Emil J. Straube

August 2012

Major Subject: Mathematics

ABSTRACT

Multiscale Simulation and Uncertainty Quantification Techniques for Richards' Equation in Heterogeneous Media. (August 2012)

Seul Ki Kang, B.S., Korea Advanced Institute of Science and Technology

Co-Chairs of Advisory Committee: Dr. Yalchin Efendiev
Dr. Raytcho Lazarov

In this dissertation, we develop multiscale finite element methods and uncertainty quantification technique for Richards' equation, a mathematical model to describe fluid flow in unsaturated porous media. Both coarse-level and fine-level numerical computation techniques are presented. To develop an accurate coarse-scale numerical method, we need to construct an effective multiscale map that is able to capture the multiscale features of the large-scale solution without resolving the small scale details. With a careful choice of the coarse spaces for multiscale finite element methods, we can significantly reduce errors.

We introduce several methods to construct coarse spaces for multiscale finite element methods. A coarse space based on local spectral problems is also presented. The construction of coarse spaces begins with an initial choice of multiscale basis functions supported in coarse regions. These basis functions are complemented using weighted local spectral eigenfunctions. These newly constructed basis functions can capture the small scale features of the solution within a coarse-grid block and give us an accurate coarse-scale solution. However, it is expensive to compute the local basis functions for each parameter value for a nonlinear equation. To overcome this difficulty, local reduced basis method is discussed, which provides smaller dimension spaces with which to compute the basis functions.

Robust solution techniques for Richards' equation at a fine scale are discussed. We construct iterative solvers for Richards' equation, whose number of iterations

is independent of the contrast. We employ two-level domain decomposition preconditioners to solve linear systems arising in approximation of problems with high contrast. We show that, by using the local spectral coarse space for the preconditioners, the number of iterations for these solvers are independent of the physical properties of the media. Several numerical experiments are given to support the theoretical results.

Last, we present numerical methods for uncertainty quantification applications for Richards' equation. Numerical methods combined with stochastic solution techniques are proposed to sample conductivities of porous media given in integrated data. Our proposed algorithm is based on upscaling techniques and the Markov chain Monte Carlo method. Sampling results are presented to prove the efficiency and accuracy of our algorithm.

ACKNOWLEDGMENTS

I would like to present my deep appreciation to Dr. Yalchin Efendiev and Dr. Raytocho Lazarov for giving me wonderful support during my graduate career. Dr. Efendiev gave me many great ideas for my research, showed me invaluable insights, and encouraged me to achieve better results. He also guided me to put my graduate process on the right track. Dr. Lazarov helped me to get mathematical details for my research. He was always patient and he waited until I got my own outcome. In addition, I would like to thank Dr. Binayak P. Mohanty and Dr. Bani Mallick for serving on my commitee and presenting many constructive comments and questions.

I also would like to thank to Dr. Juan Galvis for many great discussions and technical help. I wish to express my gratitude to the Department of Mathematics, especially to Dr. Paulo Lima-Filho and Dr. Peter Howard, who were departmental graduate advisors during my graduate career. I also wish to present my appreciation to Ms. Monique Stewart for her incredible knowledge and ability as a graduate program assistant. I give thanks to all my friends in the Department of Mathematics for their support and friendship.

I would like to acknowledge the financial support that I received during my graduate process. I am grateful to the Department of Mathematics for providing my graduate teaching and research assistantships throughout my study. My research is partially supported by the National Science Foundation under Grant No. 0934837.

Last, I'd like to thank my family for their invaluable support and encouragement throughout my graduate career.

TABLE OF CONTENTS

Page

Page

Page

LIST OF TABLES

TABLE Page

TABLE Page

TABLE Page

TABLE Page

LIST OF FIGURES

FIGURE Page

FIGURE Page

1. INTRODUCTION

Studies on the unsaturated zone have been the subject of extensive research in recent years. Much of this research has been carried out by soil scientists and hydrologists in relation to issues such as crop irrigation and rainfall run-off. In all the studies of unsaturated soil, it is assumed that Richards' equation [62] is modeling subsurface flow. This equation is a mathematical model that is derived by combining Darcy's law with a mass conservation equation of porous media.

A major problem of the flow in natural porous media is that it is largely affected by multiple scales. Data that can be used for constructing flow models can vary over a wide range of length scales. Laboratory studies are performed and soil samples at the micron scale can be identified. Indirect measurements of the soil property data varies on a scale of a foot to thousands of feet. Geophysical data, such as precipitation, can be incorporated into soil properties over scales of several miles. Therefore, it is desirable to develop flow simulations that can capture as many of the scales as possible underlying the available data.

The numerical solution of Richards' equation has received considerable attention in recent years. However, the media properties often vary significantly which introduces an additional level of complexity. A high contrast in the media properties expressed as the ratio between low and high conductivity values brings an additional small scale into the problem. For example, in flow through fractured porous media, it is common to have several orders of magnitude of variations in the conductivity within fractures. Moreover, fracture regions, that are the main carriers of the flow, can have complex heterogeneities.

Numerical discretization of flow problems in heterogeneous media results in very large ill-conditioned systems of linear equations and wherever large disparities in spatial scales are found, the computation to get the numerical solution becomes much

This dissertation follows the style of SIAM Journal of Numerical Analysis.

difficult. To construct a numerical model of coarse scale features for such problems, we need to understand the physical characteristics of the media on different scales. It is important to note that in this case traditional scale separation, such as classical homogenization or upscaling techniques, cannot be straightforwardly applied because they can not capture the high heterogeneity charateristics of the media. Therefore, one needs to take into account the variation in the conductivities when designing efficient coarse-grid simulation techniques. The same problems arise in constructing robust iterative solvers.

In this dissertation, we are interested in modeling the flow of water into a porous medium. Our motivation stems from Richards' equation which describes the infiltration of water into a porous media whose pore spaces are filled with air and water. Richards' equation is given by

$$D_t\theta(u) - div(\kappa(x,u)\nabla(u+x_3)) = f, \qquad x \in \Omega, \tag{1.1}$$

where $\theta(u)$ denotes the volumetric fluid content, u represents the pressure head and $\kappa(x,u) \geq k_0 > 0$ is the relative hydraulic conductivity and k_0 is a constant.

In this study, we will concentrate on solving the Richards' equation (1.1) with some boundary conditions. The objectives of the study are threefold:

(1) Develop a multiscale finite element method (MsFEM) to obtain an accurate solution of Richards' equation on a coarse grid.

(2) Construct robust iterative solvers for Richards' equation which allows the convergence of the numerical solution to be independent of the physical properties of the media.

(3) Design efficient uncertainty quantification techniques in inverse problems for Richards' equation

Before we move on to detailed descriptions of each chapter, we want to motivate the goal of this dissertation and give an overviews of the main results.

Recently, a number of approaches have been introduced where the coupling of small scale information is performed through a numerical formulation of the global problem by incorporating the fine features of the problem into coarse elements. Some approaches involve the solution on a coarse grid (e.g. [16, 18, 20, 21, 36, 42, 53, 79]). In these approaches, coarse-grid properties are computed to represent the media or the solution on the coarse grid.

One of the commonly used approaches is upscaling methods (e.g. [79]) where coarse-grid conductivities are computed, and then the flow equation is solved on this coarse grid. Instead of coarse-grid conductivities, multiscale basis functions can be used to represent the solution on a coarse grid [36, 42]. In the latter, multiscale basis functions are constructed on a coarse grid, and further the approximation for the solution on a fine grid is sought on a finite dimensional space spanned by these basis functions. Multiscale basis functions are typically constructed by solving the local flow equation on a coarse grid subject to some boundary conditions.

A coarse space based on local spectral problems is constructed in [21, 22, 26]. In this dissertation, we discuss the coarse spaces that are used in multiscale finite element methods (MsFEMs) for solving the problem on a coarse grid as well as in two-level preconditioners for iterative solvers. The local spectral basis functions are constructed by incorporating small-scale localizable features of the solution into intial multiscale basis functions and they make us to capture the fine-scale features of the media. Morevover, we prove that when these eigenfunctions are included in the coarse space in the domain decomposition methods, the condition number of the preconditioned matrix is bounded independently of the contrast.

We are also interested in studying efficient techniques for solving flows in highly heterogeneous formations. In many cases, multiscale methods can not provide an accurate approximation of the solution and one needs to solve for the fine-scale solution. The solution of the fine-scale equation is typically prohibitively expensive because of the small scales and high contrast in the conductivity field. For this

reason, to reduce the number of iterations required for solving the fine-scale system of equations, some type of preconditioning is needed.

The design and analysis of preconditioners that converge independently of the contrast are important for many applcations. Domain decomposition techniques use the solutions of local problems and a coarse problem in constructing preconditioners for the fine-scale system. The number of iterations required by domain decomposition preconditioners is largely affected by the high-variabilities of conductivity within the coarse grid [31]. It is known that if high and low conductivity regions can be encompassed within coarse grid blocks such that the variation of the conductivity within each coarse region is bounded, domain decomposition preconditioners result in a system with the condition number independent of the contrast (e.g., [48, 71]). However, in the case of the problems that we are considering, the media has a complex geometry so that it is hard to separate high and low conductivity regions. Therefore, special techniques to handle this situation are required. Constructing a right coarse space, which is used in constructing preconditioners, is a key point in making this happen.

The reduced basis (RB) approach has been introduced to solve large parameter-dependent problems [47, 64]. This technique is proposed to reduce the dimension of the parameter space and make the computation for solving parameter dependent problems much faster. The RB method can be modified and applied to MsFEM for nonlinear equations. The main idea of this combined method is the construction of local spectral basis functions for MsFEM for nonlinear equations, considered as parameter-dependent equatons. This can reduce substantially the overall computation time to solve the equations .

We also consider the application of uncertainty quantification into the inverse problem of Richards' equation. In many engineering applications, one needs to solve inverse problems of fluid dynamics to predict the movement of groundwater flow and soil water content under the ground. These kinds of problems contain uncertainties

which need to be quantified and minimized. We will perform some sampling techniques which include prarameterization of uncertainty and an upscaling process for cheap computation.

We want to outline the dissertation. In Chapter II, we introduce Richards' equation, which is the basic equation describing underground water movement. We present general formulations of the equation and give some popular models that hydrologists use. Numerical solution techniques for the equation is introduced. Several attempt to find a numerical solution for Richards' equation has been introduced in the past few decades, e.g., finite difference methods or finite element methods. In our research, we discretized the equation and define the nonlinear fixed point iteration. Further we show that under some conditions there exists a unique solution of the discrete problem.

Chapter III gives a general overview of MsFEM. There are several ways to construct coarse-scale basis functions and global formulations for MsFEM. We first introduce linear boundary conditions multiscale finite element basis functions for linear elliptic equations and show some examples. It is known that MsFEM can have large errors if the boundary conditions for the basis functions do not reflect the underlying heterogeneities. Therefore, one of the ways to reduce resonance errors when constructing the basis functions, oversampling methods can be introduced. Energy minimizing basis functions for MsFEM also are presented.

We introduce local spectral basis functions. Newly constructed basis functions were first introduced to overcome the difficulties aroused from high contrast problems. The basic idea is that by complementing the standard finite element basis functions with local spectral functions, we can capture the small scale properties of the media with comparably few dimensions. The coarse spaces that are spanned by these basis functions can be used in MsFEM for solving the problem on a coarse grid as well as in two-level preconditioners for iterative solutions. In the last part, we show

its applications to multiscale finite element methods for Richards' equation. Some numerical examples are presented.

Because of the nonlinearity of the problem, multiscale basis functions can be constructed using nonlinear mappings. Given nonlinear equations, which are considered parameter-dependent equations, it can be solved using multiscale basis functions with the help of reduced basis methods. The advantage of this method is that we can save computational time by avoiding whole computations in constructing basis functions for MsFEM. For example, if we compute local spectral basis functions for each parameter value, the computation becomes so expensive. However, reduced basis approach provides a smaller dimension space for which to compute basis functions, and results in cheaper ways to get the desired solution with comparably small errors.

In Chapter IV, we discuss robust solution techniques for separable Richards' equation. We construct iterative solvers for a finite element approximation of Richards' equation, where the number of iterations is independent of the contrast. The proposed iterative procedure involves outer and inner iterations. Outer iterations are designed to handle nonlinearities and inner iterations are employed to solve a linear problem for every outer iteration. For the solution of the linear problem, we employ two-level domain decomposition preconditioners. In particular, we show that the robust iterative techniques designed for a linear system can be re-used for every outer iteration if the nonlinear coefficient function is smooth.

We employ two-level domain decomposition preconditioners. The key point in making inner iterations independent of the contast is constructing a right coarse space for the preconditioners. The local spectral coarse spaces, which are introduced in Chapter III, are used for the preconditioners. By presenting several numerical results, we can show that our developed iterative methods are robust with respect to the contrast values.

Last, in Chapter V, we discuss applications of uncertainty quantification techniques to Richards' equation. In many applications, one needs to solve inverse problems for Richards' equation. Soil moisture prediction is one of the most common applications of these kinds of problems. For predicting moisture content under soil, some types of uncertainty come into this process. Quantifying and reducing this uncertainty with proper techniques allow better predictions.

In this dissertation, we focus on sampling saturated conductivity conditioned on some average flux data. For sampling, we use a Markov chain Monte Carlo (MCMC) method. This method allows to simulate a dependent sequence of random samples from very complex stochastic models. However, to perform MCMC, one needs to compute a realization of the conductivity on a fine grid, which is extremely expensive. Therefore, we employ a multiscale method to coarsen the saturated hyraulic conductivity to get a cheaper computation. Some numerical results support our theoretical view.

2. RICHARDS' EQUATION

Attention to the unsaturated zone has increased in recent years because of growing industrial and agricultural activities that largely affect the quality of the subsurface environment. Fluid flow in unsaturated porous media is often modeled by Richards' equation [62] to describe the relationship among fluid pressures, saturations, and relative permeabilities [5, 73]. This equation is a mathematical model that describes a variably saturated flow and it is derived by combining Darcy's law with a mass conservation equation in porous media [41].

Analytical solutions of Richards' equation have been studied [46, 49, 69]. However, analytic solutions are obtained in very limited cases under some simplifying assumptions. Most of the analytic solutions were obtained using the exponential hydraulic parameter model proposed by Gardner [30] (See also Section 2.3). Such an exponential model allows us to linearize the governing flow equation, which make us possible to find an analytical solution. A detailed review of this approach for unsaturated flow problems was presented by Pullan [56]. Although analytical soltuions may have limited practical applications, they do serve as a means for verifying many numerical models for unsaturated flow. These are especially useful for infiltration in very dry, layered soils where numerical models often suffer from lack of convergence and mass balance problems. In addition, the analytical solutions may enhance our understanding of the infiltration process under a transient state in layered soils.

The analytical solutions mentioned above are restrictive in nature and also limited to one dimensional problems. Consequently, numerical treatments are needed to solve more complicated problems. Numerical methods for solving Richards' equation have been developed in the last few decades. It is desirable to develop appropriate numerical schemes for moisture flow in unsatruated porous media.

A number of finite difference and finite element methods which are the modern tool for solving partial differential equatoins were developed [33, 34, 55, 67]. These

methods often suffer to some degree from mass balance errors as well as from numerical oscillations and dispersion. Additional numerical problems may appear when the gravitational term becomes important. Finite elements are advantageous for several domains in two and three-dimensions. In one dimension finite difference is advantageous because it needs no mass lumping to prevent oscillations [57, 74]. A mass conservative model for solving mixed form of Richards' equation using a finite difference method has been presented in [10]. Other techniques have also been implemented including a mixed finite element [3, 25].

Numerical simulation of ground water flow and transport requires solutions of large, sparse systems of equations. Thus, the complexity and size of problems that can be solved numerically are often constrained by the efficiency of the algorithm used to solve the resulting systems of equations. A number of studies have compared the efficiency of iterative algorithms used to solve a variety of two and three-dimensional, linear and nonlinear ground water flow problems [35,44,50]. These investigations typically compared two or more iterative solvers such as the strongly implicit procedure (SIP), successive over-relaxation (SOR),or conjugate gradient (CG) methods and preconditioned CG methods. Results of these comparisons varied depending on the specific problems, but there is general agreement that preconditioned CG methods provide competitive convergence rates, with an added advantage of not requiring operator-spedified parameters. However, even the preconditioned CG methods for some problems suffer of cost of convergence or robustness that make them impractical, or inapplicable. To overcome this problem, in Chapter 4 we will construct robust iterative solvers using preconditioned CG methods.

2.1 Derivation of Richards' equation using Darcy's law

Richards' equation is a simple consequence of two basic equations. The first one is the fundamental mass conservation law

$$D_t\theta + divV_D = 0, \tag{2.1}$$

where θ is the volumetric water content in the unit volume of soil (the liquid can be water, oil or other fluid), and V_D is the Darcy flux or fluid flow per unit area.

The second equation is a generalization of the empirical Darcy's equation relating the fluid flux, V_D to the total energy potential $H = u - x_3$ (see for example the book [68])

$$V_D = -\kappa(x,u)\nabla H = -\kappa(x,u)\nabla u + \kappa(x,u)\nabla x_3, \tag{2.2}$$

where u represents pressure head and ∇x_3 is the unit vector directed downward, in the positive direction of the axis x_3. The hydraulic conductivity $\kappa(x,u)$ of an unsaturated porous medium can be presented as a the product $\kappa(x,u) = k_s k_r(u,x)$, where $k_s = kg/\nu$ is the saturated hydraulic conductivity in which k is the medium permeability, g is the acceleration due to gravity and ν is the fluid kinematic viscosity. Another variable, $k_r(x,u)$, is the relative hydraulic conductivity of the unsaturated medium, it depends on the volumetric fluid content and satisfies $0 \leq k_r(x,u) \leq 1$. Note that the Darcy law was established originally for water percolation through saturated soils. Later it was generalized to fluid flow through unsaturated porous media with nonconstant hydraulic conductivity $\kappa(u,x)$ depending on the volumetric fluid content. Substituting V_D from Eq. (2.2) into Eq. (2.1), one obtains

$$D_t\theta = div[\kappa(x,u)\nabla(u + x_3)] \tag{2.3}$$

This is commonly known as Richards' equation [62], while Buckingham [6] derived it almost quarter of century earlier.

2.2 Several forms of Richards' equation

Richards' equation is given by

$$D_t\theta(u) - div(\kappa(x,u)\nabla(u+x_3)) = f, \qquad x \in \Omega. \tag{2.4}$$

where $\theta(u)$ denotes the volumetric fluid content, u represents pressure head and $\kappa(x,u) \geq k_0 > 0$ is the relative hydraulic conductivity and k_0 is a constant. The following are assumed [62] for (2.4): (1) the porous media and water are incompressible; (2) the temporal variation of the water saturation is significantly larger than the temporal variation of the water pressure; (3) the air phase is infinitely mobile so that the air pressure remains constant (in this case it is atmospheric pressure which equals zero); and (4) neglect the source/sink terms.

The equation (2.4) is called the coupled form of Richards' equation. This is also called the mixed form of Richards' equation, due to the fact that it involves two variables, namely, the water content θ and the pressure head u. Taking advantage of the differentiability of the soil retention function, one may rewrite (2.4) as follows

$$C(u)D_t u - div(\kappa(x,u)\nabla(u+x_3)) = f, \qquad x \in \Omega, \tag{2.5}$$

where $C(u) = d\theta/du$ is the specific moisture capacity. This version is referred to as the head-form (h-form) of Richards' equation. Another formulation of the Richards' equation is based on the water content θ,

$$D_t\theta - div(D(x,\theta)\nabla\theta) - D_{x_3}\kappa = f, \qquad x \in \Omega, \tag{2.6}$$

where $D(\theta) = \kappa(\theta)/(d\theta/du)$ defines the diffusivity. This form is called the θ- form of Richards' equation.

The three versions of Richards' equation written above have various advantages and disadvantages which in general depend upon the physical situations of the prob-

lems considered. The θ-form, for example, is a conservative form by construction, i.e., it follows the mass conservation law. However, this form only applies to the unsaturated zone, since in a saturated condition the water content becomes constant and D approaches infinity. Furthermore, for multi-layered soils, θ cannot be guaranteed to be continuous across interfaces separating the layers. Thus, this form may be useful only for a homogeneous media.

On the other hand, due to the fact that the pressure head is continuous even for multi-layered soils, the head-form may be advantageous for heterogeneous soil condition. It is also applicable for both unsaturated and saturated media. Nevertheless, as described above, the head-form does not maintain the global conservation of mass. Recently, Rathfelder et al. [60] proposed a method to solve the head-form equation that still maintains the global mass balance. The key to their method is a different way of evaluating the specific moisture capacity C, the so called standard chord slope approximation method.

The coupled-form of Richards' equation is also mass conserved. It is applicable to both saturated and unsaturated porous media. The authors of [10] proposed the so-called modified Picard iteration to solve this equation, and made a comparison with results from the h-form and showed that the coupled-form can maintain the mass conservation throughout the time marching of the simulation. These advantages have caused many researchers and engineers to use this version for various practical problems. Through this disseration, we will also use coupled-form of Richards' equation.

2.3 Constitutive relations

Constitutive relations between θ and u and between κ and u are developted appropriately, which consequently gives nonlinearity behavior in (1.1). Reliable approximations of these relations are in general challenging. Field measurements or laboratory experiments to determine the parameters are relatively expensive, and

furthermore, even if one can come up with such relationships from these works, they will be somehow limited to the particular cases under consideration.

Perhaps the most widely used empirical constitutive relations for moisture content and hydraulic conductivity is due to the work of van Genuchten [73]. He proposed a method of determining the functional relationship of relative hydraulic conductivity to the pressure head by using the field observation knowledge of moisture retention. In turn, the procedure would require curve-fitting to the proposed moisture retention function with the experimental/observational data to establish certain parameters inherent to the resulting hydraulic conductivity model. There are several widely known formulations of the constitutive relationship:

Haverkamp model [33] : This model is devloped by Haverkamp et al. (1977) to characterize the hydraulic properties of a soil. They compared six models, employing different ways of discretization of the non-linear infiltration equation in terms of execution time, accuracy, and programming considerations. The models were tested by comparing water content profiles calculated at given times by each of the models with results obtained from an infiltration experiment carried out in laboratory. All models yielded excellent agreement with water content profiles measured at various times.

The following analytical expressions, obtained by a least square fit thorugh all data points were chosen for characterizing the soil:

$$\theta(u) = \frac{\alpha\,(\theta_s - \theta_r)}{\alpha + |u|^{\beta}} + \theta_r,\ \ \kappa(x,u) = \kappa_s(x)\frac{A}{A + |u|^{\gamma}}. \tag{2.7}$$

where $\kappa_s = 34cm/h$, $A = 1.175 \times 10^6$, $\gamma = 4.74$, and $\theta_s = 0.287, \theta_r = 0.075, \alpha = 1.611 \times 10^6, \beta = 3.96$.

Supscript s refers to saturation, i.e. the value of θ for which $u = 0$, and the subscript r to residual water content.

van Genuchten model [73] : van Genuchten model, motivated by the Mualem [51], was developed for predicting the hydraulic conductivity from knowlege of the soil-water retention curve and the conductivity at satruation. In [73], a closed-form equation of hydarulic conductivity is derived using an equation for the soil-water retention curve which is both continuous and has a continuous slope. The resulting conductivity models generally contain three independent parameters, m, n, and α, which may be obtained by matching the propsed soil-water retention curve to experimental data. In [73], these parameters are estimated as $m = 0.5, n = 2, \alpha = 0.005$.

$$\theta(u) = \frac{\alpha\,(\theta_s - \theta_r)}{[1 + (\alpha|u|)^n]^m} + \theta_r,\ \ \kappa(x,u) = \kappa_s(x)\frac{\left\{1 - (\alpha|u|)^{n-1}\,[1 + (\alpha|u|)^n]^{-m}\right\}^2}{[1 + (\alpha|u|)^n]^{m/2}}. \tag{2.8}$$

Exponential model [77] : Exponential model was introduced in several papers [30, 58] with the aim to find an analytical solution for Richards' equation. For a limited range of values of the u, equation (2.9) can be fitted empirically to match of the capillary conductivity data presently available, but it does not hold well over a wide range of values. Empirical data gives α(units of L^{-1}) having value within the range 0.002 to 0.2 cm^{-1} [58].

$$\theta(u) = \theta_s\, e^{\beta u},\ \ \kappa(x,u) = \kappa_s(x)\, e^{\alpha u}. \tag{2.9}$$

The κ_s in the above model is also known as saturated hydraulic conductivity. It was observed that the hydraulic conductivity has a broad range of values, which together with the functional forms presented above confirm the nonlinear behavior of the process. It can also be seen that the water content and hydraulic conductivity approach zero as the pressure head goes to very large negative values. In other words,

Richards' equation has a tendency to degenerate in a very dry conditions with a large negative pressure.

2.4 Numerical discretization of Richards' equation

Now we consider the numerical discretization of Richards' equation. From now on we consider the steady-state Richards' equation

$$-\text{div}(\kappa(x,u)\nabla(u+x_3)) = f, \qquad x \in \Omega. \tag{2.10}$$

In many practical cases, the heterogeneous portion of the relative conductivity is given by a spatial field that does not depend on u, i.e., $\kappa(x,u) = \kappa(x)\lambda(u)$. We refer to this case as the separable Richards' equation. By denoting, $u + x_3$ as a new variable and assuming λ is smooth, we can write the above equation as

$$-\text{div}(\kappa(x)\lambda(x,u)\nabla u) = f, \qquad x \in \Omega, \tag{2.11}$$

where $\kappa(x)$ is a function with high-variability and $\lambda(x,u)$ is a smooth function that varies moderately in both x and u.

2.4.1 Weak formulation

We multiply the equation (2.11) by a test function $v \in H_0^1(\Omega)$ and integrate over the domain Ω. After applying divergence theorem, we get that the solution u satisfies the following integral identity

$$\int_\Omega k(x)\lambda(x,u)\nabla u \cdot \nabla v dx \;\; = \;\; \int_\Omega f v dx, \quad \text{for all } v \in H_0^1(\Omega).$$

Now we define the space $V = H_0^1(\Omega)$, set of all functions with square integrable generalized derivatives of first order vanishing on the boundary $\partial\Omega$, the form $a(\,\cdot\,,\,\cdot\,;\,\cdot\,)$

$$a(u, v; w) = \int_\Omega k(x)\lambda(x, w)\nabla u \cdot \nabla v dx, \tag{2.12}$$

and the linear functional $F(\,\cdot\,)$

$$F(v) = \int_\Omega f v dx. \tag{2.13}$$

Then the variational form of (2.11) is to find $u \in V$ such that

$$a(u, v; u) = F(v), \qquad \text{for all } v \in V. \tag{2.14}$$

2.4.2 Finite element discretization

Let $\mathcal{T}_h$ be a triangulation of the domain Ω into a finite number of triangular (tetrahedral) elements. We assume that $\mathcal{T}_h$ quasiuniform and regular; see [13]. Let V^h be the finite dimensional subspace of V of piece-wise polynomials with respect to $\mathcal{T}_h$. Let $u_h \in V^h$ be a solution of the following discrete problem.

$$a(u_h, v; u_h) = F(v), \qquad \text{for all } v \in V^h. \tag{2.15}$$

Under suitable conditions, one can ensure the existence of a solution to the above equation. Define the nonlinear map $T_h : V^h \to V^h$ by

$$a(T_h u_h, v; u_h) = F(v), \qquad \text{for all } v \in V^h. \tag{2.16}$$

This is well defined, since $u_h \in V^h$.

2.4.3 A nonlinear fixed point iteration

In this section, we describe a numerical method to approximate the numerical solutions of Richards' equation (2.15). We use a fixed point iteration based on the contractivity of the mapping T_h defined in (2.16). The numerical solution u_h can be approximated to an arbitrary accuracy using Picard iteration.

Starting with an initial guess $u_h^0 \in V^h$, we define the nonlinear fixed point iteration by

$$u_h^{n+1} = T_h u_h^n.$$

That is, given u_h^n, the next approximation u_h^{n+1} is the solution of the linear elliptic equation

$$a(u_h^{n+1}, w; u_h^n) = F(w), \quad \text{for all } w \in V^h. \tag{2.17}$$

In order to define the solution method, we reformulate the problem (2.17) in terms of the linear operator $A^n : V^h \to V^h$ defined for any given $u_h^n \in V^h$ as

$$a(v, w; u_h^n) = (A^n v, w), \quad \text{for all } v, w \in V^h, \tag{2.18}$$

where $(\,\cdot\,,\,\cdot\,)$ is the standard L^2-inner product in V^h. In a similar manner, we present the linear functional $F(w)$ in the form

$$F(w) = (b, w), \quad \text{for all } w \in V^h. \tag{2.19}$$

Obviously, b is the L^2-projection of the right hand side f of (2.11) on V^h. Then the equation (2.17) can be rewritten in the following operator form

$$A^n u_h^{n+1} = b. \tag{2.20}$$

Note that equation (2.17) (and its operator counterpart (2.20)) is an approximation of the linear equation $-\mathrm{div}(\kappa(x)\lambda(x,u_h^n)\nabla u_h^{n+1}) = f$ with u_h^n being the previous iterate.

2.4.4 Existence of the numerical solution

In this section, we show that there exists a solution of the discretization of Richards' equation, (2.15). First, for a given $\mathcal{K} > 0$, we introduce the ball

$$V_h^{\mathcal{K},p} := \{v \in V_h : \|v\|_{W_p^1} \leq \mathcal{K}\}. \tag{2.21}$$

The following three assumptions are used in the proofs of Theorems 1, and 2.

Assumption 1.

(A) $C_0 \leq k(x) \leq M$, *where* C_0, *and* M *is a constant.*

(B) The function $\lambda(x,u)$ *satisfies the following conditions.*

- *(a)* $\lambda(x,u)$ *is Lipschitz continuous with respect to* u, *i.e., there exists a constant* C_1 *such that* $|\lambda(x,u) - \lambda(x,v)| \leq C_1|u-v|$, *for all* $u, v \in V$, $x \in \Omega$,
- *(b)* $\lambda(x,u)$ *is bounded above, i.e. there is a constant* C *such that* $\lambda(x,u) \leq C$ *for all* $x \in \Omega$ *and* $u \in L^\infty(\Omega)$
- *(c)* $\lambda(x,u)$ *is bounded below, i.e. there is a constant* C_2 *such that* $0 < C_2 \leq \lambda(x,u)$ *for all* $x \in \Omega$ *and* $u \in V$.

Under these assumptions, we show the following theorems concerning the existence of the solution.

Theorem 1. *Under the Assumption 1, there are constants* $\alpha < \infty$, $h_0 > 0$ *and* $\epsilon > 0$ *such that for all* $0 < h \leq h_0$ *and* $u_h \in V^h$

$$|u_h|_{W_p^1(\Omega)} \leq \alpha \sup_{0 \neq v_h \in V^h} \frac{a(u_h, v_h;\,\cdot\,)}{|v_h|_{W_q^1(\Omega)}}, \quad \textit{with} \quad a(u,v;\,\cdot\,) = \int_\Omega k\nabla u \nabla v \, dx, \tag{2.22}$$

whenever $|2-p| \leq \epsilon$, *q is the dual index to p,* $\frac{1}{p}+\frac{1}{q}=1$ *and* $|\cdot|_{W_q^1(\Omega)}$ *is a semi-norm in* $W_q^1(\Omega)$.

Proof. It was shown in [4] that for $\delta > 0$ there exists $\epsilon > 0$ such that

$$|u_h|_{W_p^1(\Omega)} \leq (1+\delta) \sup_{0 \neq v_h \in V^h} \frac{\langle \nabla u_h, \nabla v_h \rangle}{|v_h|_{W_q^1}(\Omega)}, \quad \text{for all } |2-p| \leq \epsilon, \tag{2.23}$$

where $\frac{1}{p}+\frac{1}{q}=1$ and δ and ϵ are independent of h. Now, we consider a high-contrast case via a perturbation argument.

Define a bilinear form $\mathcal{B} : W_p^1(\Omega) \times W_q^1(\Omega) \to \mathbb{R}$ by

$$\mathcal{B}(u,v) \;\; := \;\; \langle \nabla u, \nabla v \rangle - \frac{1}{M} a(u,v;\,\cdot\,).$$

It follows from Assumption 1 (A) and Hölder's inequality that

$$\mathcal{B}(u,v) \leq \left(1 - \frac{C_0}{M}\right) \int_\Omega |\nabla u(x) \nabla v(x)| dx \leq \left(1 - \frac{C_0}{M}\right) |u|_{W_p^1(\Omega)} |v|_{W_q^1(\Omega)}. \tag{2.24}$$

Note that $C_0/M < 1$. Then the identity $\langle \nabla u, \nabla v \rangle = \mathcal{B}(u,v) + \frac{1}{M} a(u,v;\,\cdot\,)$, together with estimates (2.23) and (2.24) yields

$$M \left(\frac{1}{1+\delta} - \left(1 - \frac{C_0}{M}\right) \right) |u_h|_{W_p^1(\Omega)} \leq \sup_{0 \neq v_h \in V^h} \frac{a(u_h, v_h;\,\cdot)}{|v_h|_{W_q^1(\Omega)}}.$$

Let $\delta = \frac{C_0}{2M - C_0}$, and choose ϵ to be as given in (2.23) for this particular choice of δ. Then, $M\left(\frac{1}{1+\delta} - \left(1 - \frac{C_0}{M}\right)\right) = C_0/2$. Recall that $a(u,v;\,\cdot\,)$ can be very large because of high contrast. Note that ϵ and α depend only on the constants C_0, C^* and M, though the coercivity bound is independent of the contrast M. This completes the proof. □

Theorem 2. *Let the Assumption 1 hold. Then there exists* $\mathcal{K} > 0, p > 2, h_0 > 0$, *and* $\delta > 0$ *such that for all* F *with* $\|F\|_{W_p^{-1}} \leq \delta$, T_h *maps* $V_h^{\mathcal{K},p}$ *into itself for all* $0 <$

$h \le h_0$ and by Browder fixed point Theorem, there exists a solution $\tilde{u}_h$ of equation (2.15) and it satisfies

$$T_h \tilde{u}_h = \tilde{u}_h. \tag{2.25}$$

Proof. For any $u_h \in V_h^{\mathcal{K},p}$, $k(x)\lambda(x,u_h)$ satisfies the conditions of Theorem 1 with a constant M_0 such that

$$M_0 = \sup\{k(x)\lambda(x,s) : \|s\|_{L^\infty} \le c_p |\log h| \mathcal{K}\}, \tag{2.26}$$

where h is the mesh-size of the partition $\mathcal{T}_h$ and c_p is the constant in Sobolev's inequality [4],

$$\|v\|_{L^\infty(\Omega)} \le c_p |\log h| \|v\|_{W_p^1(\Omega)}, \quad \text{for all } v \in W_p^1(\Omega). \tag{2.27}$$

The constant M_0 exists due to Assumption 1. Then $u_h \in V_h^{\mathcal{K},p}$ implies that $\|u_h\|_{L^\infty(\Omega)} \le c_p |\log h| \|u_h\|_{W_p^1(\Omega)} \le c_p |\log h| \mathcal{K}$ and hence $\sup\{k(x)\lambda(x,u_h)\} \le M_0$. For sufficiently small $\mathcal{K}$ (e.g., $\mathcal{K} = C/c_p$) there is a $p > 2$ such that the inequality (2.22) holds. Then

$$\begin{aligned} \|T_h u_h\|_{W_p^1(\Omega)} \le & \quad \alpha \sup_{0 \ne v_h \in V^h} \frac{a(T_h u_h, v_h; u_h)}{|v_h|_{W_q^1(\Omega)}} && \text{(from Theorem 1)} \\ = & \quad \alpha \sup_{0 \ne v_h \in V^h} \frac{F(v_h)}{|v_h|_{W_q^1(\Omega)}} \le \quad C \|F\|_{W_p^{-1}(\Omega)}. \end{aligned}$$

Choose $\|F\|_{W_p^{-1}(\Omega)} \le \mathcal{K}/C$ to get $\|T_h u_h\|_{W_p^1(\Omega)} \le \mathcal{K}$, i.e., T_h maps $V_h^{\mathcal{K},p}$ into itself. By Browder fixed point [15], there exists a solution $\tilde{u}_h$ of equation (2.15) and it satisfies

$$T_h \tilde{u}_h = \tilde{u}_h. \tag{2.28}$$

□

3. MULTISCALE FINITE ELEMENT METHODS FOR SEPARABLE COEFFICIENT RICHARDS' EQUATION

The percolation of water in the soil varies greatly according to the soil structure which has many variations depending on the location or environmental conditions. The difficulty in analyzing groundwater transport is caused mainly by the heterogeneity of subsurface formations which may span many scales. These multiple scales dominate simulation costs and therefore we need to construct solution techniques which resolve all the small-scale effect into the large scales.

In this chaper, we introduce multiscale finite element methods (MsFEM) and coarse spaces that are constructed using selected eigenvectors of a local spectral problem. The main idea of MsFEM is to capture the multiscale structure of the solution via localized basis functions and couple them through a global formulation which will provide a faithful approximation of the solution. MsFEM basis functions are associated with a coarse grid, a partition of the domain into finite elements, and are designed to capture the multiscale features of the solution. Important multiscale features of the solution are incorporated into these localized basis functions which contain information about the scales. The finite element method based on these basis functions provides an accurate approximation of the solution.

The main goal in this chapter is to develop coarse spaces for MsFEM that result in accurate coarse-scale solutions. With careful choice of the coarse spaces, we can significantly reduce the approximation error. It is known that boundary conditions for basis functions need to be carefully chosen in order to obtain accurate coarse-scale approximations of the solution (e.g., [17–19, 29]). In particular, if the boundary conditions for the basis functions do not reflect the underlying heterogeneities, MsFEM can have large errors which come from the resonance between the coarse-grid size and characteristic length scale of the problem. To reduce the resonance error, a number of approaches have been proposed in the past, e.g. oversampling methods [17, 36].

We also consider basis functions obtained using global constraints [20, 53]. Special partitions of unity function using energy minimizing methods [81] was first constructed to facilitate the two requirements of the coarse space in an algebraic multigrid algorithm, stability and approximation. The stability is addressed by minimizing the total energy of the basis functions and the approximation property is maintained by imposing the constraint $\sum_i \phi_i(x) = 1$.

In this chapter, we take the approach in [21], where a coarse space based on local spectral problems was introduced. The construction of coarse spaces starts with an initial choice of multiscale basis functions that are supported in coarse regions sharing a common node. These basis functions are complemented using some weighted local spectral problems. These local spectral problems use the scalar conductivity as a weight function. As a result, we get new sets of basis functions that include the eigenvectors from local spectral problems. These newly constructed basis functions can capture the important features of the solution within a coarse-grid block. Also they can handle the high contrast property which is defined as the ratio between the highest and lowest conductivity values witin a coarse grid.

3.1 Multiscale finite element methods for linear elliptic problems

In this section, we describe multiscale finite element method framework for linear elliptic problems. MsFEM, as a numerical upscaling method, computes the coarse-scale solution by using a multiscale basis functions. Two main aspects of MsFEMs are the global formulation and the construction of basis functions. We can apply this method also to nonlinear problems, such as steady-state Richards' equation.

3.1.1 Global Fomulation

Let $\Omega \subset \mathbb{R}^2$(or $\mathbb{R}^3$) be a polygonal domain and $\mathcal{T}_H = \{K\}$ is a coarse-grid partition, where H denotes the size of the coarse grid. We consider the elliptic equation with heterogeneous coefficients

$$-div(\kappa(x)\nabla u) = f, \tag{3.1}$$

where $\kappa(x)$ is a heterogeneous field with high contrast. In particular, we assume that $\kappa(x) \geq c_0 \geq 0$ (bounded below), while $\kappa(x)$ can have very large variations. We multiply the equation (3.1) by a test function $v \in H_0^1(\Omega)$ and integrate over the domain Ω. By divergence theorem, we get the integral identity

$$\int_\Omega \kappa(x)\nabla u \cdot \nabla v dx = \int_\Omega f v dx, \qquad \text{for all } v \in H_0^1(\Omega).$$

We define the bilinear form a and the linear functional f by

$$a(u,v) = \int_\Omega \kappa(x)\nabla u(x) \cdot \nabla v(x) dx, \qquad \text{for all } u, v \in H_0^1(\Omega), \tag{3.2}$$

and

$$f(v) = \int_\Omega f(x)v(x)dx, \qquad \text{for all } v \in H_0^1(\Omega). \tag{3.3}$$

The variational formulation of this problem is: Find $u \in H_0^1(\Omega)$ such that

$$a(u,v) = f(v) \qquad \text{for all } v \in H_0^1(\Omega). \tag{3.4}$$

Let $\mathcal{T}_h$ be a fine triangulation, which is a refinement of $\mathcal{T}_H$. We denote by $V^h(\Omega)$ the usual finite element discretization of piecewise linear continuous functions with respect to the fine triangulation $\mathcal{T}_h$. Denote also by $V_0^h(\Omega)$ the subset of $V^h(\Omega)$ with vanishing values on $\partial\Omega$.

The Galerkin formulation of (3.4) is; find $u \in V_0^h(\Omega)$ such that

$$a(u, v) = f(v) \qquad \text{for all } v \in V_0^h(\Omega) \tag{3.5}$$

or in matrix form

$$Au = b, \tag{3.6}$$

where for all $u, v \in V^h(\Omega)$ we have

$$u^T A v = \int_\Omega \kappa \nabla u \nabla v \quad \text{and} \quad v^T b = \int_\Omega f v.$$

Here we denote by u, v the vectors and represent $u, v \in V^h(\Omega)$ by its nodal basis functions.

We denote by $\{y_i\}_{i=1}^{N_v}$ the vertices of the coarse mesh $\mathcal{T}_H$ and define the neighborhood of the node y_i by

$$\omega_i = \bigcup \{K_j \in \mathcal{T}_H; \quad y_i \in \overline{K}_j\}. \tag{3.7}$$

and the neighborhood of the coarse element K by

$$\omega_K = \bigcup \{\omega_j \in \mathcal{T}_H; \quad y_i \in \overline{K}\}. \tag{3.8}$$

For given N_c number of linearly independent functions $\{\Phi_i\}_{i=1}^{N_c}$ associated with the coarse mesh $\mathcal{T}_H$, we define a coarse space $V_0 = \text{span}\{\Phi_i\}_{i=1}^{N_c}$. Note that, in MsFEM, the basis functions $\{\Phi_i\}_{i=1}^{N_c}$ are the functions defined on the fine-grid but are supported in ω_i.

MsFEM approximates the solution on a coarse grid as $u_0 = \sum_i c_i \Phi_i$, where c_i are determined from

$$a(u_0, v) = f(v) \qquad \text{for all } v \in V_0.$$

Once c_i's are determined, one can define a finite element approximation of the solution by reconstructing via basis functions, $u_0 = \sum_i c_i \Phi_i$.

One can also views MsFEM in the following way. MsFEM solves the underlying fine-scale equations on the coarse grid. Given coarse-scale basis functions, the coarse matrix is given by

$$A_0 = R_0 A R_0^T \qquad \text{where} \qquad R_0^T = [\Phi_1, \ldots, \Phi_{N_c}]. \tag{3.9}$$

Here Φ_i's are discrete coarse-scale basis functions defined on a fine grid represented by the vector forms. Multiscale finite element solution is the finite element projection of the fine scale solution into the space V_0. More precisely, multiscale solution u_0 is given by

$$A_0 u_0 = f_0 \tag{3.10}$$

where $f_0 = R_0^T b$.

3.1.2 Coarse space construction for MsFEM

In this section, we will discuss some coarse spaces constructed to capture the fine-scale features of the solution. We will introduce three different ways to construct the coarse spaces for MsFEM; linear boundary conditions multiscale coarse space, oversampling techniques, and energy minimzing coarse space. Further, we discuss a coarse space with local spectral information which complement the multiscale coarse spaces. We obseve that using this coarse space, we can get an accurate coarse-scale approximation.

3.1.2.1 Linear boundary conditions

First, let Φ_i^0 be the nodal basis of the standard finite element space W_H with respect to the coarse triangulation $\mathcal{T}_H$. For example, W_H consists of piecewise linear

functions if $\mathcal{T}_H$ consists of triangles or tetrahedra or W_H consists of piecewise bi-linear functions if $\mathcal{T}_H$ consists of rectangles or bricks. Next we define standard multiscale finite element basis functions Φ_i^{ms} that coincide with Φ_i^0 on the boundaries of the coarse mesh. Namely, for each $K \in \omega_i$

$$\text{div}(\kappa \nabla \Phi_i^{ms}) = 0 \ \text{ in } K \in \omega_i, \quad \Phi_i^{ms} = \Phi_i^0 \ \text{ in } \partial K, \ \ \forall \ K \in \omega_i, \tag{3.11}$$

where K is a coarse grid block within ω_i. Now define the linear boundary conditions multiscale coarse spaces

$$V_0^{ms} = \text{span}\{\Phi_i^{ms}\}. \tag{3.12}$$

Note that multiscale basis functions coincide with standard finite element basis functions on the boundaries of coarse grid blocks, and may be oscillatory in the interior of each coarse grid block depending on κ. Even though the choice of Φ_i^0 can be quite arbitrary, our main assumption is that the basis functions satisfy the leading order homogeneous equations. We remark that the MsFEM formulation allows one to take advantage of scale separation. In particular K in (3.11) can be chosen to be a volume smaller than the coarse grid [17].

One-dimensional example: We consider

$$-(\kappa(x)u')' = f, \tag{3.13}$$

$u(x) = u(1) = 0$, where $'$ refers to the spatial derivative. We assume that the interval [0,1] is divided into N segments $0 = x_0 < x_1 < x_2 < \cdots < x_i < x_{i+1} < \cdots < x_N = 1$. The multiscale basis function for the node i is given by

$$(\kappa(x)(\Phi_i^{ms})')' = 0 \tag{3.14}$$

with the support on $[x_{i-1}, x_{i+1}]$. In the interval $[x_{i-1}, x_i]$, the boundary conditions for the basis function Φ_i^{ms} are defined as $\Phi_i^{ms}(x_{i-1}) = 0, \Phi_i^{ms}(x_i) = 1$. In the interval $[x_i, x_{i+1}]$, the boundary conditions for the basis function Φ_i^{ms} are defined as $\Phi_i^{ms}(x_i) = 1, \Phi_i^{ms}(x_{i=1}) = 0$, which are the boundary conditions for standard finite element basis functions. The elements of the stiffness matrix A_0 (see (3.9)) are given by

$$(a_0)_{ij} = \int_{x_{i-1}}^{x_i} \kappa(x)(\Phi_i^{ms})'(\Phi_j^0)'dx + \int_{x_i}^{x_{i+1}} \kappa(x)(\Phi_i^{ms})'(\Phi_j^0)'dx.$$

Note that for the computation of the elements of the stiffness marix, we do not need an explicit expression of Φ_i^{ms} and instead, we simply need to compute $\kappa(x)(\Phi_i^{ms})'$. From (3.14), it is easy to see that $\kappa(x)(\Phi_i^{ms})' = \text{const}$, where the constants are different in $[x_{i-1}, x_i]$ and $[x_i, x_{i+1}]$. This constant can be easily computed by writing $(\Phi_i^{ms})' = \text{const}/\kappa(x)$ and integrating it over $[x_{i-1}, x_i]$. This yields

$$\kappa(x)(\Phi_i^{ms})' = \frac{1}{\int_{x_{i-1}}^{x_i} \frac{dx}{\kappa(x)}}, \qquad x \in [x_{i-1}, x_i],$$

$$\kappa(x)(\Phi_i^{ms})' = -\frac{1}{\int_{x_{i-1}}^{x_i} \frac{dx}{\kappa(x)}}, \qquad x \in [x_i, x_{i+1}].$$

We set,

$$(a_0)_{ij} = \int_{x_{i-1}}^{x_i} \kappa(x)(\Phi_i^{ms})'(\Phi_j^0)'dx + \int_{x_i}^{x_{i+1}} \kappa(x)(\Phi_i^{ms})'(\Phi_j^0)'dx \tag{3.15}$$

$$= \frac{1}{\int_{x_{i-1}}^{x_i} \frac{dx}{\kappa(x)}} \int_{x_{i-1}}^{x_i} (\Phi_j^0)'dx - \frac{1}{\int_{x_i}^{x_{i+1}} \frac{dx}{\kappa(x)}} \int_{x_i}^{x_{i+1}} (\Phi_j^0)'dx. \tag{3.16}$$

Taking into account that $\int_{x_{i-1}}^{x_i} (\Phi_{i-1}^0)'dx = -1, \int_{x_{i-1}}^{x_i} (\Phi_i^0)'dx = 1, \int_{x_i}^{x_{i+1}} (\Phi_i^0)'dx = -1, \int_{x_i}^{x_{i+1}} (\Phi_{i+1}^0)'dx = 1$, we have

$$(a_0)_{i,i-1} = -\frac{1}{\int_{x_{i-1}}^{x_i} \frac{dx}{\kappa(x)}}, (a_0)_{i,i} = \frac{1}{\int_{x_{i-1}}^{x_i} \frac{dx}{\kappa(x)}} + \frac{1}{\int_{x_i}^{x_{i+1}} \frac{dx}{\kappa(x)}}, (a_0)_{i,i+1} = -\frac{1}{\int_{x_i}^{x_{i+1}} \frac{dx}{\kappa(x)}}.$$

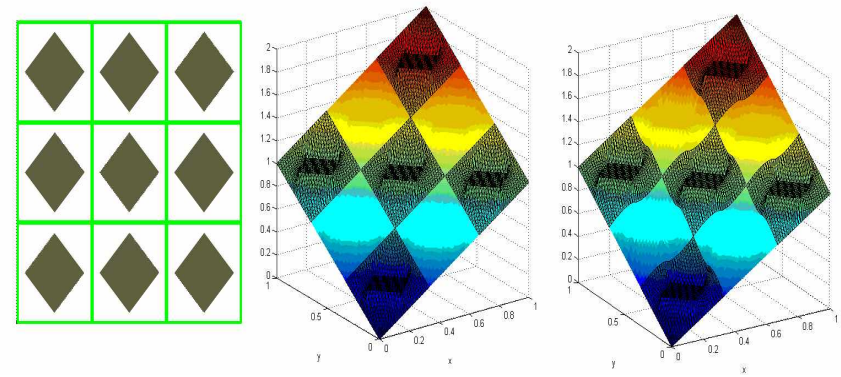

Fig. 3.1. Left: Conductivity field. Middle: Fine-scale solution. Right: Coarse-scale solution with multiscale basis functions that have linear boundary conditions.

3.1.2.2 Oversampling technique

A common difficulty in numerical upscaling methods is that large errors result from the resonance between the grid scale and the scales of the continuous problem. When the coefficient $\kappa(x)$ is a periodic function varying over ϵ-scale ($\kappa(x) = \kappa(x/\epsilon)$), the convergence rate of MsFEM contains a term ϵ/H, which is large when $H \approx \epsilon$, [38]. Recall that H is the coarse mesh size. A previous analysis [38] showed that the error due to the resonance is expressed as a ratio between the wavelength of the small scale oscillation and the grid size. It suggests also that by a judicious choice of boundary conditions for the base function, one can reduce the resonance errors significantly.

This concept is illustrated on Figure 3.1, showing the permeability field (left figure), fine-scale solution (middle figure), and multiscale solution with linear boundary conditions (right figure) computed on a 3×3 coarse grid. We can see that the multiscale solution with linear boundary conditions does not capture the fine-scale features of the solution along the boundaries of the coarse grid. This can lead to large errors.

To overcome the difficulty due to scale resonance, an over-sampling method was proposed [36]. The idea is quite simple. To reduce the errors coming from boundary

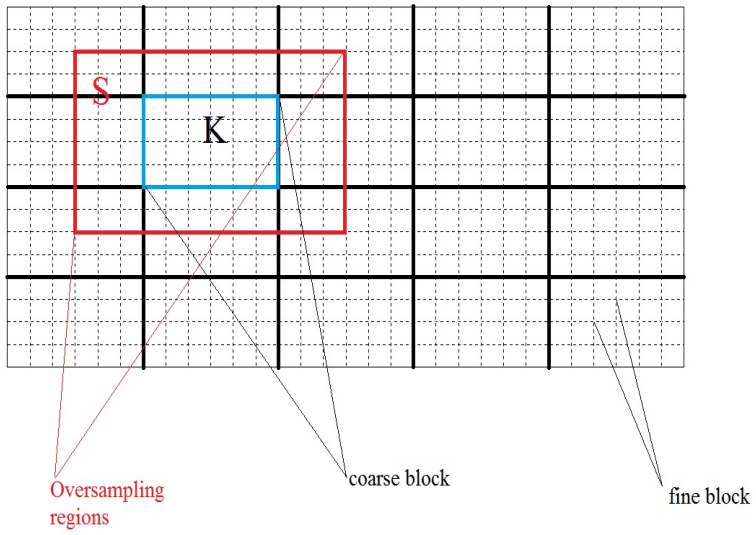

Fig. 3.2. Schematic description of oversampled regions.

layers, we can sample in a domain with a size larger than h and use only the interior sampled information to construct the bases.

Here is the brief description. Let Φ_j^E be the basis functions satisfying the homogeneous elliptic equation in the larger domain $K_E \supset K$(see Figure 3.2). We then form the actual basis Φ_i by linear combination of Φ_j^E,

$$\Phi_i = \sum_j c_{ij} \Phi_j^E.$$

The coefficients c_{ij} are determined by condition $\Phi_i(x_j) = \delta_{ij}$, where x_j are nodal points. Note that this method is nonconforming method, i.e. it could generate a noncoforming error. The analysis [36] show that the nonconforming error is indeed small.

Other boundary conditions have been introduced and analyzed in the literature. For example, reduced boundary conditions are found to be efficient in many porous media applications, see [42].

3.1.2.3 Energy minimizing methods

Energy minimizing has been considered in an algebraic multigrid method for solving large scale algebraic systems. The construction of coarse spaces, a key issue in an algebraic multigrid algorithm, has two major regulations, stability and approximation. In [82], an energy minimizing basis was first constructed to facilitate these two requirements. The stability is addressed by minimizing the total energy of all basis functions

$$\min \sum_i \|\phi_i\|_a^2 \tag{3.17}$$

where $\|\cdot\|_a$ is the energy norm associated with the given elliptic boundary value problem. The approximation property is maintained by imposing the partition of unity constraint

$$\sum_i \phi_i(x) = 1. \tag{3.18}$$

Extensive numerical experiments reported in [76] show that this energy minimizing basis leads to uniformly convergent multigrid methods for many problems of practical interest such as problems with rough coefficients.

As indicated in (3.17), and (3.18), the energy minimizing basis is given in a global optimization problem with a pointwise constaint. A major concern with this approach is the cost for solving this constraint optimization problem. We will follow a procedure similar to that proposed in [81], which provides an easy and efficient method for solving this optimization problem.

To construct the energy minimizing multiscale basis, we solve the following minimization problem

$$\min \sum_i \int_{\omega_i} \kappa |\nabla \Phi_i^{em}|^2 \tag{3.19}$$

subject to $\sum_i \Phi_i^{em} = 1$ with $\text{Supp}(\Phi_i^{em}) \subset \omega_i$, $i = 1, \ldots, N_v$. We can define energy minimizing coarse space V_0^{EM} as

$$V_0^{em} = \text{span}\{\Phi_i^{em}\}. \tag{3.20}$$

We note that to compute these basis functions, one needs to solve a global linear system. This is where expensive compared to the local computation of multiscale finite element basis functions with linear boundary conditions Φ_i^{ms}.

3.1.3 A coarse space with local spectral information

We motivate the choice of the coarse spaces based on the analysis presented in [26–28]. Its motivation stems from flow problems that occur in heterogeneous porous media. In flow problems, the high conductivity regions are the main carrieres of the fluid. Often, there are low permeability layers inside of these high conductivity regions and their complicated connectivities make the flow in such media complicated. Therefore, it is often impossible to separate them into coarse grid blocks without significantly increasing the computational effort.

The coarse spaces discussed in [26–28] complement the original multiscale coarse spaces(e.g. linear boundary multiscale space or energy minizing space) and improve the accuracy of the approximation, especailly in flow problems in heterogeneous porous media. Here is the breif description:

For fixed ω_i consider the eigenvalue problem

$$-\text{div}(\kappa \nabla \psi_\ell^{\omega_i}) = \mu_\ell^{\omega_i} \widetilde{\kappa} \psi_\ell^{\omega_i}, \tag{3.21}$$

where $\mu_\ell^{\omega_i}$ and $\psi_\ell^{\omega_i}$ are eigenvalues and eigenvectors in ω_i and $\widetilde{\kappa}$ is defined by

$$\widetilde{\kappa} = \frac{1}{H^2} \kappa \sum_{j=1}^{N_c} |\nabla \Phi_j^{in}|^2. \tag{3.22}$$

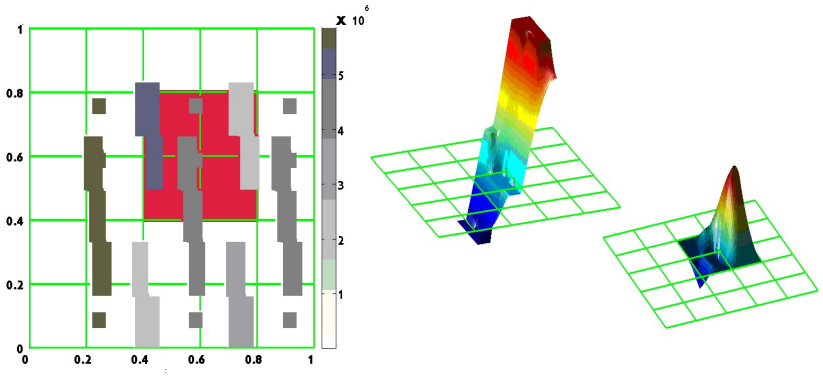

Fig. 3.3. Schematic description of basis function construction. Left: subdomain ω_i. Right-Top: Selected eigenvector ψ_i^ℓ with small eigenvalue. Right-Bottom: product $\Phi_i\psi_i^\ell$ where Φ_i is the initial basis function of node i.

We recall that Φ_j^{in} (simply denoted by Φ_j in further discussions) are the initial multiscale basis functions (either multiscale basis functions with linear boundary conditions or energy minimizing basis functions) and N_c is the number of the coarse nodes. The eigenvalue problem considered above is solved with zero Neumann boundary condition and understood in a discrete setting. Assume eigenvalues are given by

$$\mu_1^{\omega_i} \leq \mu_2^{\omega_i} \leq ....$$

Basis functions are computed by selecting a number of eigenvalues (starting with small ones) and multiplying corresponding eigenvectors by Φ_i. Thus, multiscale space is defined for each i as the span of $\Phi_i\psi_\ell^{\omega_i}$, $\ell = 1, ..., L_i$, where L_i is the number of selected eigenvectors (see Figure 3.3 for an illustration).

We note that $\{\omega_i\}_{y_i \in \mathcal{T}^H}$ is a covering of Ω. Let $\{\Phi_i\}_{i=1}^{N_c}$ be a partition of unity subordinated to the covering $\{\omega_i\}$ such that $\Phi_i \in V_0^h(\omega_i)$ and $|\nabla\Phi_i| \leq \frac{1}{H}$, $i = 1, \ldots, N_c$. Define the set of coarse basis functions

$$\Psi_{i,\ell} = I^h(\Phi_i\psi_\ell^{\omega_i}), \quad \text{for } 1 \leq \ell \leq L_i \text{ and } 1 \leq i \leq N_c, \tag{3.23}$$

where I^h is the fine-grid nodal value interpolation and L_i is an integer number specified for each $i = 1, \ldots, N_c$. Note that in this case, there might be several basis functions per coarse node. The number of basis functions per node is defined via the eigenvalue problem (3.21). Denote by V_0 the *local spectral multiscale* space

$$V_0^{lsm} = \text{span}\{\Psi_{i,\ell} : 1 \leq \ell \leq L_i \text{ and } 1 \leq i \leq N_c\}. \tag{3.24}$$

3.1.3.1 Reduced dimension coarse spaces

We note that the eigenvalues of (3.21) depend on the initial basis functions Φ_i since the weight function $\widetilde{\kappa}$ is determined by the initial basis functions. Basis functions are computed by selecting a number of eigenvalues (starting with small ones) and multiplying corresponding eigenvectors by Φ_i (see Figure 3.3 for the illustration). Thus, multiscale space is defined for each i as the span of $\Phi_i \psi_l^{\omega_i}$, $l = 1, ..., L_i$, where L_i is the number of the selected eigenvectors.

Therefore, the dimension of the coarse space depends on the choice of $\widetilde{\kappa}$. Thus it is important to have a good choice of $\widetilde{\kappa}$ for the local eigenvalue problem (3.21). In this study, we designed $\widetilde{\kappa}$ using initial multiscale basis functions. The special form of the weight function $\widetilde{\kappa}$, (3.22), is motivated by the analysis of the energy stability of the coarse interpolant and the analysis of the stable decomposition in the DD method. From [27], we know that the number of eigenvectors of (3.21) is related to the number of high conductivity inclusions/channels inside of the coarse regions. It is shown that the number of small eigenvalues is the same as the number of high-conductivity regions.

Therefore, we are interested in the partition of unity functions that can elimiate isolated high-conductivity inclusions and thus reduce the dimension of the coarse space. This can be achieved by minimizing the high-conductivity components in $\widetilde{\kappa}$. In particular, by choosing linear boundary multiscale basis functions or energy

minizing basis functions, we can eliminate all isolated high-conductivity inclusions, while preserving the channels.

These newly introduced multiscale techniques in constructing initial basis help to reduce the dimension of the coarse space needed to achieve contrast-independent two-level domain decomposition preconditioners and more accurate coarse-grid solutions. We will show an application of this technique to Richards' equation and show some numerical examples.

3.2 Application of MsFEM to Richards' equation with separable coefficient and numerical examples

As we mentioned earlier in Chapter 2, there are several approaches solving this nonlinear partial differential equation numerically with several different boundary conditions. The finite element, finite volume, and finite difference methods are most commonly used to generate the discretized equation. In our research, we used multi-scale finite element approach to obtain a discretized numerical solution for Richards' equation. In [17], MsFVEM is used to solve Richards' equation on the coarse grid. Basically, in this approach, multiscale bases are constructed and global formulation is developed under finite volume conditions and numerical results show that this methods can be used with success in predicting the solution on the coarse grid.

Here we consider the steady-state equation

$$div(\kappa(x,u)\nabla(u+x_3)) = f, \qquad x \in \Omega, \tag{3.25}$$

Since the equation involves nonlinearity, several assumptions for the coefficient $\kappa(x,u)$ have been made to show the existence of the solution. Together with the ellipticity, a Lipschitz-like continuity condition was assumed for the coefficient [17], and it is shown that there exists a unique solution for this kind of equation.

Here we consider the case when the conductivity $\kappa(x,u)$ has large variations. More specifically we study the case that $\kappa(x,u) = \kappa(x)\lambda(u)$ where $\kappa(x)$ has high-variability while $\lambda(u)$ is a smooth function that varies moderately in both x and u. The computational time to get a numerical solution highly depends on the large variability of the conductivity. Therefore, we need a careful treatment for these kinds of conductivities.

By denoting, $u + x_3$ as a new variable and assuming λ is smooth, we can write the above equation as

$$div(\kappa(x)\lambda(x,u)\nabla u) = f, \qquad x \in \Omega. \tag{3.26}$$

Now we want to find the solution of variational form of (3.26), $a(\,\cdot\,,\,\cdot\,;\,\cdot\,)$, which is given at (2.14).

3.2.1 Coarse-scale Fixed Point Iteration

Define a coarse space V_H by $V_H = \text{span}\{\Phi_i\}_{i=1}^{N_c}$,a span of coarse basis functions $\{\Phi_i\}_{i=1}^{N_c}$ and let $u_H \in V_H$ be a solution of the following discrete problem,

$$a(u_H, v; u_H) = F(v), \qquad \text{for all } v \in V_H, \tag{3.27}$$

We know that under suitable conditions, one can ensure the existence of a solution to the above equation. The solution u_H may be written as

$$u_H = \sum_{i=1}^{N_c} \alpha_i \Phi_i \qquad \text{for some } \alpha = (\alpha_1, \alpha_2, \cdots, \alpha_{N_c}) \in \mathbb{R}^{N_c}. \tag{3.28}$$

Now we want to approximate a numerical solution of the discretized Richards' equation (3.27) using a fixed point iteration technique. We describe a fixed point

iteration based on the contractivity of the nonlinear mapping $T_H : V_H \to V_H$ defined as

$$a(T_H u_H, v; u_H) = F(v), \qquad \text{for all } v \in V_H. \tag{3.29}$$

The fixed point iteration is derived from the linearization of (3.27), i.e., given an initial guess $u_H^0 = \sum_{i=1}^{N_c} \alpha_i^0 \Phi_i \in V_H$, we define the nonlinear fixed point iteration by

$$u_H^{n+1} = T_H u_H^n. \tag{3.30}$$

That is, given u_H^n, the next approximation u_H^{n+1} can be obtained by solving the linear elliptic equation

$$a(u_H^{n+1}, v; u_H^n) = F(v), \qquad \text{for all } v \in V^H. \tag{3.31}$$

Using the linear expansion (3.28),the probelm (3.31) can be written as

$$A_H^n \alpha^{n+1} = b_H, \tag{3.32}$$

where the linear operator $A_H^n : V^H \to V^H$ is defined as

$$(A_H^n)_{ij} = \int_\Omega k(x)\lambda(x, u_H^n)\nabla\Phi_i \cdot \nabla\Phi_j dx, \tag{3.33}$$

and the right hand side b is

$$(b_H)_j = \int_\Omega f\Phi_j dx. \tag{3.34}$$

Here $\alpha^{n+1} = (\alpha_1^{n+1}, \alpha_2^{n+1}, \cdots, \alpha_{N_c}^{n+1})$ is a coefficient of the linear expansion of u_H^{n+1}. The proposed iterative procedure involves outer iterations and inner iterations. Outer iterations are designed to handle nonlinearities by linearizing the equation around the previous state. The simplest is Picard iteration is described by $\mathrm{div}(k(x)\lambda(x, u^n)\nabla u^{n+1}) = f$, where n denotes the outer iteration number. For every outer iteration n, a linear problem needs to be solved. Iterations to solve the linear system are called inner iterations. Since the conductivitiy $k(x, u)$ has complicated heterogenity and there

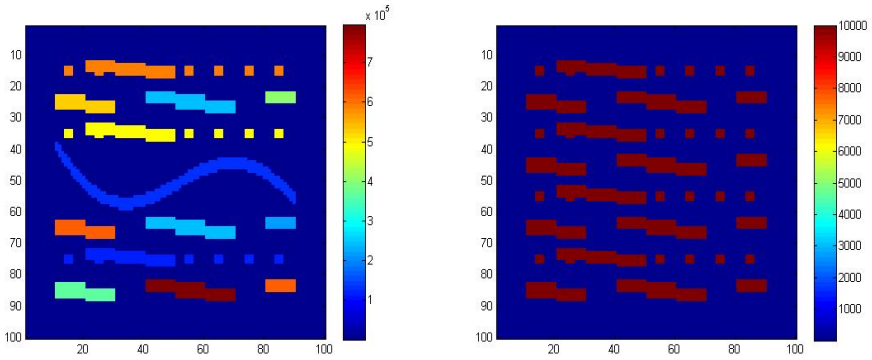

Fig. 3.4. (Left): Conductivity field 1. Blue designates the regions where the coefficient is 1 and other colors designates the regions where the coefficient is a random number between η and $10 * \eta$. (Right): Conductivity field 2. Blue designates the regions where the coefficient is 1 and red designates the regions where the coefficient is η.

are some small scale properties inside of the coarse-grid, big error can be generated during the computation of the coarse-scale discrete solution. Therefore, to reduce the error, we need to construct a suitable coarse space.

3.2.2 Numerical examples

In this section we present some representative numerical examples. We solve the Richards' equation (3.27) in $\Omega = [0,1]\times[0,1]$ with $f(x) = 1$ and homogeneous Dirichlet boundary conditions. We consider several models for the hydraulic conductivity: the Haverkamp, van Genuchten, and Exponential model (see, (2.7),(2.8),(2.9)). The coarse mesh $\mathcal{T}_H$ is obtained by dividing Ω into 10×10 mesh. The fine triangulation is obtained by dividing each coarse-mesh element into 10×10 squares and further dividing each square into two triangles. Thus, the fine-mesh step size is $h = 1/100$.

In all numerical experiments we use the initial approximation for the iterative process u_h^0 that solves

$$a(u_h^0, v; 0) = F(v), \quad \text{for all } v \in V^h. \tag{3.35}$$

We apply the coarse-scale multiscale method to Richards' equation with three different coarse spaces:

1. The coarse space V_0^{ms} described in Section; 3.1.2.1.
2. The coarse space V_0^{em} described in Section; 3.1.2.3.
3. The coarse space V_0^{lsm} described in Section 3.1.3.

We consider different conductivities with complex high-contrast configurations, see Figure 3.4. A number of parameter values in the nonlinearity of the hydraulic conductivity are tested in our simulations. In particular, for each experiment we chose a different set of parameters for the model and a set of contrast values for the hydraulic conductivity. We note that, for each outer iteration we use a PCG iteration. The inner PCG iteration is convergent when the initial residual is reduced by a factor of $tol_{in} = 1e-10$ while the outer tolerance is set to $tol_{out} = 1e-8$. We consider the followings to indicate the performance of coarse-scale solution:

- Coarse space dimension;
- The ratio between the energy norm of the gradient of the error and the gradient of the fine scale solution.
- The ratio between the energy norm of the error and the fine scale solution.

3.2.2.1 Haverkamp model

First, we consider Haverkamp model (see, e.g. [33]) with hydraulic conductivity is given by

$$k(x,u) = k_s(x)\frac{A}{A + (|u|/B)^{\gamma}}. \tag{3.36}$$

The first set of numerical results is presented in Tables 3.1 and 3.2. We use the coarse space V_0^{ms} for MsFEM. We observe that the ratio between the energy norm of the gradient of the error and the gradient of the fine scale solution is quite big with different contrast value η and so is the ratio between the energy norm of the error and the fine scale solution. We observe that the error becomes larger when B and γ (see (4.26)) decrease. This is because the smaller values of B and γ increase the magnitude of the conductivity that comes from its nonlinear component which makes the error larger. Comparing the results in Tables 3.1 and 3.2 that use different conductivity fields, we see that the error in Table 3.2 are smaller than the error in Table 3.1. This is because conductivity field 2 (see Figure 3.4) has simpler heterogeneity structure compared conductivity field 1.

Next, we repeat the above numerical experiments using the coarse space V_0^{em}. Numerical results are presented in Tables 3.3 and 3.4. We observe that, as before, the ratio between the energy norm of the gradient of the error and the gradient of the fine scale solution is comparably big and so is the ratio between the energy norm of the error and the fine scale solution.

Further, we show the numerical experiment using the spectral coarse space V_0^{lsm}. Numerical results are presented in Tables 3.5 and 3.6. We observe in this case energy ratio becomes much smaller than the previous cases. Especially, the ratio between the energy norm of the error and the fine scale solution approaches to almost zero value. These results indicate that when we use loscal spectral coarse space for MsFEM, we can get a very accurate coarse-scale solution.

Table 3.1

Numerical results using V_0^{ms}. Here we use the Haverkamp model $\kappa(x,u) = \kappa(x)\frac{A}{A+(|u|/B)^\gamma}$ with κ depicted in left picture of Figure 3.4. The coarse space dimension is 81.

	$A=1,\ B=1,\ \gamma=1$		$A=1,\ B=0.01,\ \gamma=0.5$	
η	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$
10^3	0.9847	0.9956	0.9952	0.9982
10^4	0.9873	0.9991	0.9965	0.9996
10^5	0.9876	0.9994	0.9966	0.9998

Table 3.2

Numerical results using V_0^{ms}. Here we use the Haverkamp model $\kappa(x,u) = \kappa(x)\frac{A}{A+(|u|/B)^\gamma}$ with κ depicted in right picture of Figure 3.4. The coarse space dimension is 81.

	$A=1,\ B=1,\ \gamma=1$		$A=1,\ B=0.01,\ \gamma=0.5$	
η r	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$
10^3	0.8729	0.8020	0.9065	0.8865
10^4	0.8730	0.8024	0.9066	0.8868
10^5	0.8731	0.8024	0.9066	0.8868

Table 3.3

Numerical results using V_0^{em}. Here we use the Haverkamp model $\kappa(x,u) = \kappa(x)\frac{A}{A+(|u|/B)^\gamma}$ with κ depicted in left picture of Figure 3.4. The coarse space dimension is 81.

	$A=1,\ B=1,\ \gamma=1$		$A=1,\ B=0.01,\ \gamma=0.5$	
η	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$
10^3	0.8162	0.6424	0.8577	0.7544
10^4	0.8188	0.6457	0.8594	0.7569
10^5	0.8191	0.6460	0.8596	0.7571

Table 3.4

Numerical results using V_0^{em}. Here we use the Haverkamp model $\kappa(x,u) = \kappa(x)\frac{A}{A+(|u|/B)^\gamma}$ with κ depicted in right picture of Figure 3.4. The coarse space dimension is 81.

	$A=1,\ B=1,\ \gamma=1$		$A=1,\ B=0.01,\ \gamma=0.5$	
η	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$
10^3	0.8511	0.7502	0.9277	0.8635
10^4	0.8544	0.7575	0.9321	0.8688
10^5	0.8547	0.7583	0.9325	0.8694

Table 3.5

Numerical results using V_0^{lsm}. Here we use the Haverkamp model $\kappa(x,u) = \kappa(x)\frac{A}{A+(|u|/B)^\gamma}$ with κ depicted in left picture of Figure 3.4. The coarse space dimension is 158.

	$A=1,\ B=1,\ \gamma=1$		$A=1,\ B=0.01,\ \gamma=0.5$	
η	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$
10^3	0.1827	0.0113	0.1193	0.0201
10^4	0.1829	0.0113	0.1192	0.0201
10^5	0.1829	0.0113	0.1192	0.0201

Table 3.6

Numerical results using V_0^{lsm}. Here we use the Haverkamp model $\kappa(x,u) = \kappa(x)\frac{A}{A+(|u|/B)^\gamma}$ with κ depicted in right picture of Figure 3.4. The coarse space dimension is 158.

	$A=1,\ B=1,\ \gamma=1$		$A=1,\ B=0.01,\ \gamma=0.5$	
η	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$
10^3	0.0575	0.0011	0.4624	0.4343
10^4	0.0575	0.0011	0.0644	0.0023
10^5	0.0575	0.0011	0.0644	0.0023

3.2.2.2 Exponential model

We present numerical results for exponential model where hydraulic conductivity depend exponentially on the pressure head u;

$$k(x,u) = k_s(x)e^{\alpha u/B}. \tag{3.37}$$

We present the first set of numerical results in Tables 3.7 and 3.8. First, we use the the coarse space V_0^{ms} for MsFEM. We observe the ratio between the energy norm of the gradient of the error and the gradient of the fine scale solution is relatively large and so is the ratio between the energy norm of the error and the fine scale solution.

We can see from Tables 3.7 and 3.8 (these use different conductivity fields) that the errors in Table 3.8 are smaller than the corresponding errors in Table 3.7. This is because conductivity field 2 has simpler subgrid structure, i.e. it has less complicated channel structure and smaller contrast value, compared to conductivity field 1.

Next, we repeat the numerical experiment using the coarse space V_0^{em} and the coarse space V_0^{lsm}. Numerical results for the coarse space V_0^{em} are presented in Tables 3.9 and 3.10 while the results for V_0^{lsm} are presented in Tables 3.11 and 3.12. For space V_0^{em}, the errors are comparably big. On the other hand, the errors are quite small when V_0^{lsm} is used as a coarse space. In conclusion, V_0^{lsm} provides an accurate coarse-scale solution.

3.2.2.3 VanGenutchen model

Next, we consider Van Genuchten model (see [73]) with hydraulic conductivity which is given by

$$k(x,u) = k_s(x)\frac{\{1-(\alpha|u|/B)^{n-1}[1+(\alpha|u|)^n]^{-m}\}^2}{[1+(\alpha|u|)^n]^{\frac{m}{2}}}. \tag{3.38}$$

Table 3.7

Numerical results using V_0^{ms}. Here we use the Exponential model $\kappa(x,u) = \kappa(x)e^{\alpha(u/B)}$ with κ depicted in left picture of Figure 3.4. The coarse space dimension is 81.

	$\alpha = 1,\ B = 1$		$\alpha = 2,\ B = 1$	
η	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$
10^3	0.9841	0.9955	0.9838	0.9955
10^4	0.9868	0.9990	0.9865	0.9990
10^5	0.9871	0.9994	0.9868	0.9994

Table 3.8

Numerical results using V_0^{ms}. Here we use the Exponential model $\kappa(x,u) = \kappa(x)e^{\alpha(u/B)}$ with κ depicted in right picture of Figure 3.4. The coarse space dimension is 81.

	$\alpha = 1,\ B = 1$		$\alpha = 2,\ B = 1$	
η	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$ r	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$
10^3	0.8720	0.7962	0.8713	0.7933
10^4	0.8721	0.7966	0.8718	0.7937
10^5	0.8721	0.7966	0.8718	0.7938

Table 3.9

Numerical results using V_0^{em}. Here we use the Exponential model $\kappa(x,u) = \kappa(x)e^{\alpha(u/B)}$ with κ depicted in left picture of Figure 3.4. The coarse space dimension is 81.

	$\alpha = 1,\ B = 1$		$\alpha = 2,\ B = 1$	
η	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$
10^3	0.8154	0.6373	0.8151	0.6348
10^4	0.8180	0.6406	0.8177	0.6382
10^5	0.8183	0.6410	0.8179	0.6385

Table 3.10

Numerical results using V_0^{em}. Here we use the Exponential model $\kappa(x,u) = \kappa(x)e^{\alpha(u/B)}$ with κ depicted in right picture of Figure 3.4. The coarse space dimension is 81.

	$\alpha = 1,\ B = 1$		$\alpha = 2,\ B = 1$	
η	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$
10^3	0.8472	0.7429	0.8450	0.7393
10^4	0.8490	0.7504	0.8481	0.7469
10^5	0.8507	0.7512	0.8484	0.7477

Table 3.11

Numerical results using V_0^{lsm}. Here we use the Exponential model $\kappa(x,u) = \kappa(x)e^{\alpha(u/B)}$ with κ depicted in left picture of Figure 3.4. The coarse space dimension is 158.

	$\alpha = 1,\ B = 1$		$\alpha = 2,\ B = 1$	
η	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$
10^3	0.1872	0.0108	0.1894	0.0105
10^4	0.1873	0.0108	0.1896	0.0105
10^5	0.1873	0.0108	0.1896	0.0105

Table 3.12

Numerical results using V_0^{lsm}. Here we use the Exponential model $\kappa(x,u) = \kappa(x)e^{\alpha(u/B)}$ with κ depicted in right picture of Figure 3.4. The coarse space dimension is 158.

	$\alpha = 1,\ B = 1$		$\alpha = 2,\ B = 1$	
η	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$
10^3	0.0605	0.0011	0.0622	0.0010
10^4	0.0605	0.0011	0.0622	0.0010
10^5	0.0605	0.0011	0.0622	0.0010

As before, we present numerical results for all three coarse spaces. First, in Tables 3.13 and 3.14 we present the numerical results for the coarse space V_0^{ms}. We observe the ratio between the energy norm of the gradient of the error and the gradient of the fine scale solution is comparably big and so is the ratio between the energy norm of the error and the fine scale solution. Now we compare Table 3.13 and Table 3.14 for two different conductivity fields shown in Figure 3.4. We observe that the errors presented in Table 3.14 is smaller than those presented in Table 3.13 which is consistent with our previous observations.

Numerical results for the coarse spaceV_0^{em} are presented in Tables 3.15 and 3.16, while numerical results for the coarse space V_0^{lsm} are presented in Tables 3.17 and 3.18. We observe that the errors bewteen the coars-scale solution and the fine-scale solution is quite small which make us to conclude that spectral coarse space gives an accurate MsFEM coarse scale solution.

Table 3.13

Numerical results using V_0^{ms}. Here we use the van Genuchten model $\kappa(x,u) = \kappa(x)\frac{\{1-(\alpha(|u|/B))^{n-1}[1+(\alpha(|u|/B))^n]^{-m}\}^2}{[1+(\alpha(|u|/B))^n]^{m/2}}$ with κ depicted in left picture of Figure 3.4. The coarse space dimension is 81.

	$\alpha = 0.005$, $B = 1$, $n = 2$, $m = 0.5$	
η	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$
10^3	0.9840	0.9956
10^4	0.9864	0.9991
10^5	0.9867	0.9994

Table 3.14

Numerical results using V_0^{ms}. Here we use the van Genuchten model $\kappa(x,u) = \kappa(x)\frac{\{1-(\alpha(|u|/B))^{n-1}[1+(\alpha(|u|/B))^n]^{-m}\}^2}{[1+(\alpha(|u|/B))^n]^{m/2}}$ with κ depicted in right picture of Figure 3.4. The coarse space dimension is 81.

	$\alpha = 0.005$, $B = 1$, $n = 2$, $m = 0.5$	
η	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$
10^3	0.8719	0.7991
10^4	0.8720	0.7995
10^5	0.8720	0.7995

Table 3.15

Numerical results using V_0^{em}. Here we use the van Genuchten model $\kappa(x,u) = \kappa(x)\frac{\{1-(\alpha(|u|/B))^{n-1}[1+(\alpha(|u|/B))^n]^{-m}\}^2}{[1+(\alpha(|u|/B))^n]^{m/2}}$ with κ depicted in left picture of Figure 3.4. The coarse space dimension is 81.

	$\alpha = 0.005$, $B = 1$, $n = 2$, $m = 0.5$	
η	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$
10^3	0.8169	0.6398
10^4	0.8195	0.6432
10^5	0.8194	0.6435

Table 3.16

Numerical results using V_0^{em}. Here we use the van Genuchten model $\kappa(x,u) = \kappa(x)\frac{\{1-(\alpha(|u|/B))^{n-1}[1+(\alpha(|u|/B))^n]^{-m}\}^2}{[1+(\alpha(|u|/B))^n]^{m/2}}$ with κ depicted in right picture of Figure 3.4. The coarse space dimension is 81.

	$\alpha = 0.005$, $B = 1$, $n = 2$, $m = 0.5$	
η	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$
10^3	0.8475	0.7466
10^4	0.8507	0.7540
10^5	0.8510	0.7548

Table 3.17

Numerical results using V_0^{lsm}. Here we use the van Genuchten model $\kappa(x,u) = \kappa(x)\frac{\{1-(\alpha(|u|/B))^{n-1}[1+(\alpha(|u|/B))^n]^{-m}\}^2}{[1+(\alpha(|u|/B))^n]^{m/2}}$ with κ depicted in left picture of Figure 3.4. The coarse space dimension is 158.

	$\alpha = 0.005$, $B = 1$, $n = 2$, $m = 0.5$	
η	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$
10^3	0.1849	0.0110
10^4	0.1851	0.0110
10^5	0.1851	0.0110

Table 3.18

Numerical results using V_0^{lsm}. Here we use the van Genuchten model $\kappa(x,u) = \kappa(x)\frac{\{1-(\alpha(|u|/B))^{n-1}[1+(\alpha(|u|/B))^n]^{-m}\}^2}{[1+(\alpha(|u|/B))^n]^{m/2}}$ with κ depicted in right picture of Figure 3.4. The coarse space dimension is 158.

	$\alpha = 0.005$, $B = 1$, $n = 2$, $m = 0.5$	
η	$\frac{\int \kappa\|\nabla e\|^2}{\int \kappa\|\nabla u_f\|_2}$	$\frac{\int \kappa\|e\|^2}{\int \kappa\|u_f\|^2}$
10^3	0.1753	0.0110
10^4	0.1754	0.0110
10^5	0.1754	0.0110

3.3 MsFEM for Richards' equation with non-separable coefficient

In this section, we consider the numerical discretization of Richards' equation with non-separable coefficient, i.e., we will try to find the numerical solution of steady-state Richards' equation

$$-div(\kappa(x,u)\nabla u) = 0 \qquad \text{in } \Omega \tag{3.39}$$

with nonlinear coefficient $\kappa(x,u)$ which has a high-conductivity charateristic.

3.3.1 MsFEMs for nonlinear equations

MsFEM for nonlinear problems, like as MsFEM for linear problems, has two main ingredients, a global formulation and localized multiscale basis functions. However, unlike MsFEM for linear problems, basis functions for nonlinear problems need to be defined via nonlinear maps that map coarse-scale functions into fine-scale functions.

3.3.1.1 Multiscale basis construction

Recall that multiscale basis functions for linear elliptic equation are found by solving

$$div(\kappa(x)\nabla\Phi_i^{ms}) = 0. \tag{3.40}$$

In the case of nonlinear equations, the basis functions construction is a little bit different.

Let's W_H be the standard finite element space with respect to the coarse triangulation $\mathcal{T}_H$ and define the multiscale space as V_h. We can define the nonlinear mapping $T^{ms} : W_H \to V_h$ in the following way. For each coarse-scale function $u_H \in W_H$, we denote by $u_{r,h}$ the corresponding fine-scale response, $u_{r,h} = T^{ms}u_H$ that satisfies the following equation solved approximately on the fine grid

$$div(\kappa(x, \eta^{u_H})\nabla u_{r,h}) = 0 \qquad \text{in } K \tag{3.41}$$

where $u_{r,h} = u_H$ on ∂K and $\eta^{u_H} = (1/|K|)\int_K u_H dx$ for each K. The map T^{ms} is nonlinear; however, for a fixed u_H, this map is linear. One can represent $u_{r,h}$ using multiscale basis functions as $u_{r,h} = \sum_i \alpha_i \Phi_i^{u_H}$, where $\alpha_i = u_H(x_i)$, x_i being nodal points) and $\Phi_i^{u_H}$ are multiscale basis functions defined by

$$div(\kappa(x, \eta^{u_H})\nabla\Phi_i^{u_H}) = 0 \qquad \text{in } K, \quad \Phi_i^{u_H} = \Phi_i^0 \text{ on } \partial K. \tag{3.42}$$

where Φ^0 is a standard finite element basis.

3.3.1.2 Numerical formulation

One can use various global formulations for MsFEM. In this section, we present the following global numerical formulation: Find $u_H \in W_H$ (consequently, $u_{r,h} = T^{ms}u_H$) such that

$$< \kappa_{r,h} u_H, v_H >= \int_\Omega f v_H dx, \qquad \forall v_H \in W_H, \tag{3.43}$$

where

$$< \kappa_{r,h} u_H, v_H >= \int_\Omega \kappa(x, \eta^{u_H}) \nabla u_{r,h} \cdot \nabla v_H dx \tag{3.44}$$

Take $v_H = \Phi_i^0$, then the equation (3.43) can be written as a nonlinear system of equatitons,

$$A_{r,h}(u_H) = f_{r,h} \tag{3.45}$$

such that

$$A_{r,h}(u_H) = \int_\Omega \kappa(x, \eta^{u_H}) \nabla u_{r,h} \cdot \nabla \Phi_i^0 dx \tag{3.46}$$

and

$$f_{r,h} = \int_\Omega f \Phi_i^0 dx. \tag{3.47}$$

One-dimensional example: We consider a simple one-dimensional case

$$-(\kappa(x,u)u')' = f, \tag{3.48}$$

$u(0) = u(1) = 0$. We assume that the interval $[0, 1]$ is divided into N segments $0 = x_0 < x_1 < x_2 < \cdots < x_i < x_{i+1} < \cdots < x_N = 1$. For a given $u_H \in W_H$, $u_{r,h}$ is the solution of

$$(\kappa(x, \eta^{u_H})u'_{r,h})' = 0, \tag{3.49}$$

where $u_{r,h}(x_i) = u_H(x_i)$ for every interior node x_i. In the interval $[x_{i-1}, x_i]$, the equation (3.49) can be solved. To compute (3.44), we only need to evaluate $\kappa(x, \eta^{u_H})u'_{r,h}$. Noting that this quantity is constant, $\kappa(x, \eta^{u_H})u'_{r,h} = c(x_{i-2}, x_i)$, we can easily find that

$$u'_{r,h} = c(x_{i-2}, x_i)/\kappa(x, \eta^{u_H}), \tag{3.50}$$

where $\eta^{u_H} = \frac{1}{2}(u_H(x_{i-1}) + u_H(x_i))$. Taking the integral of (3.50) over $[x_{i-1}, x_i]$, we have

$$u_H(x_i) - u_H(x_{i-1}) = c(x_{i-2}, x_i) \int_{x_{i-1}}^{x_i} \frac{1}{\kappa(x, \eta^{u_H})} dx.$$

Consequently,

$$c(x_{i-2}, x_i) = \kappa(x, \eta^{u_H})u'_{r,h} = \frac{u_H(x_i) - u_H(x_{i-1})}{\int_{x_{i-1}}^{x_i} \frac{1}{\kappa(x, \eta^{u_H})} dx}.$$

To evaluate (3.45), we have

$$\begin{aligned} A_{r,h}(u_H) &= \int_{x_{i-1}}^{x_i} c(x_{i-1}, x_i)(\Phi_i^0)' dx + \int_{x_i}^{x_{i+1}} c(x_i, x_{i+1})(\Phi_i^0)' dx \\ &= \frac{u_H(x_i) - u_H(x_{i-1})}{\int_{x_{i-1}}^{x_i} \frac{1}{\kappa(x, \eta^{u_H})} dx} \int_{x_{i-1}}^{x_i} (\Phi_i^0)' dx + \frac{u_H(x_{i+1}) - u_H(x_i)}{\int_{x_i}^{x_{i+1}} \frac{1}{\kappa(x, \eta^{u_H})} dx} \int_{x_i}^{x_{i+1}} (\Phi_i^0)' dx \end{aligned}$$

Taking into account that $\int_{x_{i-1}}^{x_i} (\Phi_i^0)' dx = 1, \int_{x_i}^{x_{i+1}} (\Phi_i^0)' dx = -1$, we have

$$A_{r,h}(u_H) = \frac{u_H(x_i) - u_H(x_{i-1})}{\int_{x_{i-1}}^{x_i} \frac{1}{\kappa(x, \eta^{u_H})} dx} - \frac{u_H(x_{i+1}) - u_H(x_i)}{\int_{x_i}^{x_{i+1}} \frac{1}{\kappa(x, \eta^{u_H})} dx}. \tag{3.51}$$

3.3.2 Multiscale methods for Richards' equation. Parameter dependent flows

3.3.2.1 Motivation

For designing a numerical method for Richards' equation with nonseparable coefficient, we will use the fixed point iteration, as we have done for the separable coefficient case (see (3.31)),

$$div(\kappa(x, u^n)\nabla u^{n+1}) = 0, \tag{3.52}$$

i.e., given u^n, we need to find the next iterative solution u^{n+1}.

We can see this problem in the following form

$$div(\kappa(x, \mu)\nabla u) = 0, \tag{3.53}$$

thus, we consider a parameter-dependent elliptic equation. We assume that the coefficients have both small scales and high contrast. In a similar fashion to solving the separable coefficient case, we will construct local basis functions that encode the local features to approximate the solution of the parameter-dependent flow equation. To construct local basis functions, we first find initial multiscale basis functions and construct local spectral problems for complementing the initial coarse space.

However, for parameter dependent problems, solving the local eigenvalue problem for each parameter can be expensive, especially in the cases of large size coarse blocks. Therefore we need a special techniques to compute basis functions inexpensively. Here, we will borrow some ideas from Reduced Basis (RB) approach.

RB approach is proposed in [47,64] to solve many parameter-dependent problems and reduce the computational effort needed to obtain a solution. The main idea of RB method is that, instead of applying a original discretization method, a surrogate method that will allow to approximate the behavior of the solution is used. This is done by offline computations which involves the solution of global problems that

are used to compute new basis functions. Note that online procedures, a process to find the solutions, should be largely appreciated by offline computations because lots of computation time can be saved due to these offline steps. Here, we will use RB techniques for solving local problems across the parameter space.

In many parameter-dependent problems, the method has been successfully appied and proved to be efficient. This method is valid in case when the set $U = \{u(\mu), \mu \in \Lambda\}$ has a simple structure and the solution $u(\mu)$ is regular with respect to μ. The complexity of the problem resulting from this approach should lead to a simplification of the original problem.

3.3.2.2 Problem Setting

We consider the following parameter-dependent equation

$$-div(k(x,\mu)\nabla u) = f, \qquad x \in \Omega, \quad \mu \in \Lambda \subset \mathbb{R}^p \tag{3.54}$$

with homogeneous Dirichlet boundary conditions. This equation represents Richards' equation if μ is replaced by u. The weak formulation of the problem (3.54) is: find $u \in H_0^1(\Omega)$ such that

$$a(u, v; \mu) = F(v), \qquad \text{for all } v \in H_0^1(\Omega), \tag{3.55}$$

with

$$a(u, v; \mu) = \int_\Omega k(x,\mu)\nabla u \nabla v dx, \qquad \text{for all } u, v \in H_0^1(\Omega), \tag{3.56}$$

and

$$F(v) = \int_\Omega f v dx, \qquad \text{for all } v \in H_0^1(\Omega). \tag{3.57}$$

Our objective is to construct a coarse-grid reduced order model, so that for any values of the parameter μ and any value of the right hand side f are able to compute, in a inexpensive way, the solution of (3.54).

In the reduced model, we compute a set of multiscale basis functions $\Psi_i^j (= \Psi_i^j(x;\mu))$ which strongly depend on μ. It is often expensive to compute multiscale basis functions as one may need to solve local spectral problems to find the appropriate number of basis functions. Moreover, in order the MsFEM to be accurate, we need to define a new basis for each single value of the parameter μ. For this purpose, we will construct a reduced order problem to compute the basis functions. The construction of the basis functions involves the approximation of local parametric eigenvalue problems. We will apply reduced basis techniques to these local parametric eigenvalue problems posed on coarse regions.

In [52], RB methods and associated a posteriori error estimation procedures have been developed for parameter dependent ellliptic PDEs with affine parameter dependence, in particular, problems which accept an affine decomposition hypothesis: for some finite (preferably small) integer Q, a may be expressed as

$$a(u,v;\mu) = \sum_{q=1}^{Q} \Theta^q(\mu) a^q(u,v), \qquad \forall \omega, v \in H_0^1(\Omega), \forall \mu \in \Lambda \tag{3.58}$$

$$= \sum_{q=1}^{Q} \Theta^q(\mu) \int_\Omega \kappa_q(x) \nabla u \cdot \nabla v dx, \tag{3.59}$$

where $\Theta^q : \Lambda \to \mathbb{R}$ and $\kappa_q(x)$ is the heterogeneous spatial field with multiple scales and high contrast.

3.3.2.3 Algorithm

We shall describe the overall procedure [24] which consists of two parts: offline process and online process. In offline process, we get a reduced dimension of the eigenvalue problem and in the online process, we solve those reduced dimension eigenvalue problem to get local spectral basis. These further are used for solving the coarse-grid system of Richards' equation.

Offline Process We assume here that a fine grid, on which κ is resolved, is given. We also assume that we know the parameter space $\mu \in \Lambda$, a threshold τ and a number N_{rb}, that corresponds to the number of selected parameter value. We then proceed as follows:

1. Define a coarse mesh $\mathcal{T}_H$ and the standard finite element model basis functions, Φ_i^0 so that $W_H = \text{span}\{\Phi_i^0\}$.
2. Define Λ_{trial} as a discrete subset of Λ.
3. Choose a partition of unity $\{\Phi_i\}_{1\leq i\leq N_c}$. Here we choose the linear basis function (Φ_i^0), but other choices can be made, as well.
4. For each coarse region ω_i,
5. Compute on the fine grid the following stiffness and mass matrices for all $1 \leq q \leq Q$

$$v^T A_q^{\omega_i} u := \int_{\omega_i} \kappa_q(x) \nabla u \nabla v dx \tag{3.60}$$

$$v^T M_q^{\omega_i} u := \int_{\omega_i} \kappa_q(x) \sum_{k=1}^{N_c} |\nabla \Phi_k|^2 uv dx. \tag{3.61}$$

6. Define the sequence $\{\mu_m\}_{1\leq m\leq N_{rb}}$, using a Greedy procedure (see Algoritm 3.3.2.4), and solve

$$\sum_{q=1}^{Q} \Theta_q(\mu_j)(A_q^{\omega_i} - \lambda_\ell^{\omega_i}(\mu_m) M_q^{\omega_i}) \psi_\ell^{\omega_i}(\mu_m) = 0, \quad \lambda_\ell^{\omega_i}(\mu_m) \leq \tau. \tag{3.62}$$

7. Construct the matrix

$$R_{\omega_i}^T := [\psi_\ell^{\omega_i}(\mu_m),\ \lambda_\ell^{\omega_i}(\mu_m) \leq \tau,\ 1 \leq m \leq N_{rb}]. \tag{3.63}$$

8. end for

9. Outputs of the offline stage are $R_{\omega_i}, A_q^{\omega_i}, M_q^{\omega_i}$ and $R_0^T := [\Phi_i, 1 \leq i \leq N_c]$

Online process The purpose of this setp is to compute the solution for a given value of μ and a give f:

1. For each coarse region ω_i

2. Solve the reduced order eigenvalue problem and keep only some eigenvectors that correspond to eigenvalues below a certain thereshold $\lambda_\ell^{\omega_i,N_{rb}}(\mu) \leq \tau$

$$\sum_{q=1}^{Q} \Theta_q(\mu)(R_{\omega_i}^T A_q^{\omega_i} R_{\omega_i} - \lambda_\ell^{\omega_i,N_{rb}}(\mu) R_{\omega_i}^T M_q^{\omega_i} R_{\omega_i})\psi_\ell^{\omega_i,N_{rb}}(\mu) = 0 \tag{3.64}$$

3. Compute the multiscale basis functions

$$\Psi_i^j := \Phi_i \psi_j^{\omega_i,N_{rb}} \tag{3.65}$$

4. end for

5. Solve the coarse-grid system.

3.3.2.4 Reduced Basis for eigenvalue problems

As we mentioned eariler, RB method is used for solving the eigenvalue problem with much smaller dimension so that we can get local spectral basis functions with a much reduced workload. In the RB method, we first need to get a space $\Sigma_{N_{rb}}$ that

can approximate the manifold $\{\psi_\ell^{\omega_i}(x;\mu)$, with $\psi_\ell^{\omega_i}$ solution to (3.62), $\mu \in \Lambda_{trial}\}$. i.e.,

$$\Sigma_{N_{rb}} := \text{Span}\{\psi_\ell^{\omega_i}(\mu_m), 1 \leq m \leq N_{rb}, \lambda_\ell^{\omega_i}(\mu_m) \leq \tau\}. \tag{3.66}$$

It is spanned by N_{rb} global basis functions $\zeta_n, 1 \leq n \leq N_{rb}$ where N_{rb} is small. Typically, these basis functions are constructed from a 'snapshot set' which consists of solutions of the underlying eigenvalue problem at seleted parameter points.

If $\Sigma_{N_{rb}}$ is constructed, we solve the reduced eigenvalue problem within this space, namely,

$$\int_{\omega_i} \kappa(x,\mu)\nabla\psi_\ell^{\omega_i,N_{rb}}\nabla v dx = \lambda_\ell^{\omega_i,N_{rb}} \int_{\omega_i} \tilde{\kappa}(x,\mu)\psi_\ell^{\omega_i,N_{rb}} v dx \qquad \forall v \in \Sigma_{N_{rb}}, \tag{3.67}$$

where $\psi_\ell^{\omega_i,N_{rb}} \in \Sigma_{N_{rb}}, \lambda_\ell^{\omega_i,N_{rb}} \in \mathbb{R}$.

A Greedy algorithm is used to select iteratively the snapshots. Here is a breif description. We want to choose the set of parameter $\{\mu_m\}_{1\leq m\leq N_{rb}}$ which makes the error between the solution and the reduced basis approximation minimal. Let's denote the error function

$$e(\mu) = \|\psi_\ell^{\omega_i}(\mu) - \psi_\ell^{\omega_i,N_{rb}}(\mu)\|_{H^1(\omega_i)} \tag{3.68}$$

Then the judicious choice of μ would be

$$\{\mu_m\}_{1\leq m\leq N_{rb}} = \inf_{\mu\in\Lambda^{N_{rb}}}(\sup_{\mu\in\Lambda} e(\mu)) \tag{3.69}$$

However, choosing μ from the set Λ which can be a continuous set is difficult and expensive procedure. Also finding the eigenfunctions $\psi_\ell^{\omega_i}$ for given value μ might be computationaly very expensive. Therefore, we discretize the set Λ and try to get the parameter sets $\{\mu_m\}_{1\leq m\leq N_{rb}}$ solving the following infsup problem,

$$\{\mu_m\}_{1\leq m\leq N_{rb}} = \inf_{\mu\in\Lambda_{trial}^{N_{rb}}}(\sup_{\mu\in\Lambda_{trial}} \Delta_{N_{rb}}(\mu)) \tag{3.70}$$

where Λ_{trial} is a discret subset of Λ and $\Delta_{N_{rb}}(\mu)$ is an estimator of the error $e(\mu)$. In [24], the judicious choice of the estimator $\Delta_{N_{rb}}(\mu)$ is described. We use this estimator in our numerical simulation.

Now we choose the parameter set $\{\mu_m\}_{1\leq m\leq N_{rb}}$ iteratively using the following process, called Greedy procedure.

1: Choose $\mu_1 \in \Lambda_{trial}$
2: Solve (3.62) with $\mu = \mu_1$ to define $\Sigma_1 := \text{span}(\psi_\ell^{\omega_i}(\mu_1), \lambda_\ell^{\omega_i}(\mu_1) \leq \tau)$,
3: **for** $m = 2\text{to}N_{rb}$ do **do**
4: Choose $\mu_m := \max_{\mu\in\Lambda_{trial}}\Delta_{m-1}(\mu)$,
5: Solve (3.62) with $\mu = \mu_m$ to define $\Sigma_m := \text{span}(\psi_\ell^{\omega_i}(\mu_j), \lambda_\ell^{\omega_i}(\mu_j) \leq \tau, 1 \leq j \leq m)$,
6: **end for**

Algorithm 1
Greedy Procedure

3.3.2.5 Computing Spectral Multiscale basis using RB technique

In this section, we will use RB techniques to compute cheap online approximations for the eigenvalue problem and get a parameter dependent spectral multiscale basis. Recall the construction of spectal multiscale space from the Section 3.1.3. In this problematic settings, we will have parameter dependent coarse space for the the coarse-scale solution.

First we choose the initial coarse basis function Φ_i. Note that a possible choice for the inital basis function is to select a reasonable value of μ_0 and for each element $K_i \in \mathcal{T}_H$ compute the following multiscale partiton of unity functions:

$$-div(\kappa(x,\mu_0)\nabla\Phi_i(x)) = 0 \text{ in } K_i, \quad \Phi_i = \Phi_i^0 \text{ on } \partial K_i, \tag{3.71}$$

where Φ_i^0 is a standard finite element basis function. We need to complement this coarse basis if more accurate coarse-scale solutions are sought and this completion

is accomplished via local spectral problems. For the case of parameter-dependent elliptic problems, this completion is done at the online stage.

After finding out the eigenpairs of the eigenvalue problem (3.62) with cheap computation using RB tehcniques, we can construct multiscale basis functions. The multiscale basis functions are constructed as

$$\Psi_i^\ell = \Phi_i^\ell(x;\mu) = \Phi_i \psi_\ell^{\omega_i, N_{rb}}, \quad \text{for } 1 \leq \ell \leq L_i \text{ and } 1 \leq i \leq N_c. \tag{3.72}$$

Then we define the coarse-grid spaces as

$$V_0 = V_0(\mu) = \text{span}(\Psi_i^\ell). \tag{3.73}$$

and use it for solving the global problem.

3.3.3 Numerical examples for Richards' equation with nonseparable heterogeneities and nonlinearities

In this section, we present numerical results for

$$-div(\kappa(x,u)\nabla u) = f(x),$$

with $u = 0$ on $\partial\Omega$, $f(x) = 1$ and

$$\kappa(x,u) = \lambda_0(u)\left(\kappa_1(x) + (e^{\alpha u} - 1)\kappa_2(x)\right), \tag{3.74}$$

where $\kappa_1(x)$ and $\kappa_2(x)$ are defined as in the previous example.(See Figure 3.4. Left picture represents $\kappa_1(x)$ and Right picture represents $\kappa_2(x)$.) We take $\lambda_0(u) = e^{2u}$ and $\alpha = 1$ in this numerical experiment. In this example, we will take the average of u, $\bar{u}$, being a parameter in each coarse-grid block as discussed in Section 3.3.1.1.

Table 3.19

Relative errors in energy norm and the coarse space dimension in the last iteration. LSM+n indicates that the coarse spaces include eigenvectors corresponding to small, asymptotically vanishing eigenvalues, and additional n eigenvectors corresponding to the next n eigenvalues. Here, $h = 0.001$ and $\eta = 10^6$.

$H = 0.1$	$N_{rb} = 1$	$N_{rb} = 2$	$N_{rb} = 3$
LSM+0	$8.13e-005(104)$	$3.44e-005(141)$	$9.88e-006(156)$
LSM+1	$9.42e-005(226)$	$3.55e-005(218)$	$3.83e-006(237)$
LSM+2	$9.34e-005(308)$	$2.29e-005(302)$	$3.68e-006(318)$

Recall the bilinear form $a(\,\cdot\,,\,\cdot\,;\,\cdot\,)$ and the linear functional $F(\,\cdot\,)$ from (3.56) and (3.57).

The numerical solution u_h can be approximated to an arbitrary accuracy using a Picard iteration. Starting with an initial guess $u_h^0 \in V^h$, we define the nonlinear fixed point iteration as follows. Given u_h^n, the next approximation u_h^{n+1} is the solution of the linear elliptic equation

$$a(u_h^{n+1}, w; \overline{u_h^n}) = F(w), \quad \text{for all } w \in V^h. \tag{3.75}$$

where $\overline{u_h^n} = \frac{1}{|K_i|}\int_{K_i} u_h^n dx$, $K_i \in \mathcal{T}_H$. This is an approximation of the linear equation $-\mathrm{div}(k(x)\lambda(x, \overline{u_h^n})\nabla u_h^{n+1}) = f$ with u_h^n being the previous iterate.

We reformulate the iteration (3.75) in matrix form. Define A^n by

$$a(v, w; \overline{u_h^n}) = w^T A^n v, \quad \text{for all } v, w \in V^h, \tag{3.76}$$

and define the vector F by

$$F(w) = w^T b, \quad \text{for all } w \in V^h. \tag{3.77}$$

Then the equation (3.75) can be rewritten in the following matrix form

$$A^n u_h^{n+1} = b. \tag{3.78}$$

Furthermore, we solve (3.78) using reduced basis MsFEM and we show our numerical results next.

In Table 3.19, we present numerical results for the accuracy of our proposed MsFEM. Relative errors in energy norm and the coarse space dimension in the last iteration is displayed. We consider two cases for N_{rb}, a number of basis functions for online computations, $N_{rb} = 1$, $N_{rb} = 2$ and $N_{rb} = 3$.

We observe that, in all three cases, the relative error is comparably small when the multiscale basis functions and local spectral problems are solved with reduced basis. However, the computation cost at the online stage is much lower. In particular, when $N_{rb} = 3$, we notice that the errors are small and decrease as we increase the dimension of the local spectral coarse spaces. It indicates that the more number of basis functions are used in the proposed procedure, the more fine scale features can be captured. The case $N_{rb} = 3$ still gives a good approximation with reduced computational cost.

4. ROBUST SOLUTION TECHNIQUE FOR SEPARABLE COEFFICIENT RICHARDS' EQUATION

4.1 Introduction

In this section, we present robust preconditioners for the finite element system resulting from the discretization of nonlinear equations when $\kappa(x,u)$ is heterogeneous with respect to space. We consider the steady-state separable Richards' equation

$$-div(\kappa(x)\lambda(x,u)\nabla u) = f, \qquad x \in \Omega, \tag{4.1}$$

where $\kappa(x)$ is a have high-variability, while $\lambda(x,u)$ is a smooth function that varies moderately in both x and u.

Various iterative methods for solving nonlinear equations have been proposed and studied in the past, e.g. [7, 8, 14, 43, 70, 80]. For example, in [7, 70], a nonlinear iterative procedure has been proposed and its optimality has been established, in [43], multilevel iterative methods have been studied for Richards' equation, in [14], two-level domain decomposition methods have been proposed and analyzed. To the best of our knowledge, the techniques developed in the previous works have not considered robustness with respect to the contrast in the case of highly heterogeneous conductivity fields, which is the main objective of this paper.

The proposed iterative procedure involves outer iterations and inner iterations, a technique that is commonly used in the literature. Outer iterations are designed to handle nonlinearities by linearizing the equation around the previous state. The simplest is Picard iteration that is described by $-div(\kappa(x)\lambda(x,u^n)\nabla u^{n+1}) = f$, where n denotes the outer iteration number. For every outer iteration n, a linear problem needs to be solved. For the solution of the linear problem, we employ two-level domain decomposition preconditioners within conjugate gradient (CG) iterative technique. Both inner and outer iteration can, in general, depend on the contrast and

small scales. One of our main goals is to construct iterative process that converges independently of both, the small scales and the contrast. In particular, we show that the robust iterative techniques designed for a linear system can be re-used for every outer iteration if λ is a smooth function. Therefore, it is important to use efficient preconditioners for solving the linear system arising in approximation of a problem with highly heterogeneous coefficients. Such preconditioners, designed in the earlier works [26–28], are discussed below.

For every outer iteration, the resulting linear system on the fine scale is solved using two-level domain decomposition preconditioner (e.g., [48, 71]), which involves local (subdomain) and global (coarse) problems. The number of iterations required by domain decomposition preconditioners is typically affected by the contrast in the media properties (e.g., [48,71]) that are within each coarse grid block. Because of the complex geometry of fine-scale features, it is often impossible to separate low and high conductivity regions into different coarse grid blocks. Consequently, without proper preconditioner, the number of iterations can be very large, which substantially reduces the efficiency of the iterative method, particularly for nonlinear flows.

In this section, for every outer iteration we use the preconditioners designed in [26, 27]. The main idea of these preconditioners consists of augmenting the coarse space in the domain decomposition methods. In particular, a coarse space based on local spectral problems using multiscale functions is constructed. We prove that when the coarse space in the domain decomposition methods includes these eigenfunctions, the condition number of the preconditioned matrix is bounded independently of the contrast. The choice of multiscale space is important to achieve small dimensional coarse spaces. By incorporating small-scale localizable features of the solution into initial multiscale basis functions, we have shown that one can achieve coarse spaces of lower dimension without sacrificing the convergence properties of the preconditioners. Initial multiscale spaces can employ constructions proposed for multiscale finite element methods in [17, 19, 36, 42].

We show that both, the number of outer iterations and the number of inner iterations, are bounded independently of physical parameters, such as the contrast and small spatial scales. We first prove that under some assumptions the number of outer iterations depends on the contraction constant that is independent of the contrast in the conductivity field. Our reasoning takes into account the high variations of the contrast in the conductivity field and follows the standard for such nonlinear problems technique, e.g., [4]. As for inner iterations, we use two-level preconditioners developed in [26, 27] that provide independent of the contrast condition number for every outer iteration. We use the same preconditioner for every outer iterations repeatedly without sacrificing the convergence of the overall method.

4.2 Problem setting

4.2.1 Finite Element Discretization

We follow the settings from Section 3.1.1, i.e., we consider the variational form of (4.1), $a(\,\cdot\,,\,\cdot\,;\,\cdot\,)$, defined on (3.4). Let $\mathcal{T}_h$ be a triangulation of the domain Ω into a finite number of triangular (tetrahedral) elements. We assume that $\mathcal{T}_h$ quasiuniform and regular; see [13]. Let V^h be the finite dimensional subspace of V of piece-wise polynomials with respect to $\mathcal{T}_h$. Let $u_h \in V^h$ be a solution of the following discrete problem.

$$a(u_h, v; u_h) = F(v), \qquad \text{for all } v \in V^h. \tag{4.2}$$

We know that under suitable conditions, one can ensure the existence of a solution to the above equation. Define the nonlinear map $T_h : V^h \to V^h$ by

$$a(T_h u_h, v; u_h) = F(v), \qquad \text{for all } v \in V^h. \tag{4.3}$$

This is well defined, since $u_h \in V^h$.

4.2.2 A nonlinear fixed point iteration

In this section we describe a robust numerical method to approximate the numerical solutions of the Richards' equation (4.2). We use a fixed point iteration based on the contractivity of the mapping T_h defined in (4.3). The numerical solution u_h can be approximated to an arbitrary accuracy via using Picard iteration.

Starting with an initial guess $u_h^0 \in V^h$, we define the nonlinear fixed point iteration by

$$u_h^{n+1} = T_h u_h^n.$$

That is, given u_h^n, the next approximation u_h^{n+1} is the solution of the linear elliptic equation

$$a(u_h^{n+1}, w; u_h^n) = F(w), \quad \text{for all } w \in V^h. \tag{4.4}$$

In order to define the solution method, we reformulate the problem (4.4) in terms of the linear operator $A^n : V^h \to V^h$ defined for any given $u_h^n \in V^h$ as

$$a(v, w; u_h^n) = (A^n v, w), \quad \text{for all } v, w \in V^h, \tag{4.5}$$

where $(\,\cdot\,,\,\cdot\,)$ is the standard L^2-inner product in V^h. In a similar manner, we present the linear functional $F(w)$ in the form

$$F(w) = (b, w), \quad \text{for all } w \in V^h. \tag{4.6}$$

Obviously, b is the L^2-projection of the right hand side f of (4.1) on V^h. Then the equation (4.4) can be rewritten in the following operator form

$$A^n u_h^{n+1} = b. \tag{4.7}$$

Note that equation (4.4) (and its operator counterpart (4.7)) is an approximation of the linear equation $-\mathrm{div}(\kappa(x)\lambda(x, u_h^n)\nabla u_h^{n+1}) = f$ with u_h^n being the previous

iterate. It is known that the presence of the high-contrast coefficient $\kappa(x)$ makes computationally difficult to construct appropriated robust linear solvers for computing u_h^{n+1}. Moreover, taking into account the contractivity of the operator T_h, in order to get a robust method to compute the solution of the Richards' equation (4.3), we only need a robust method for solving the linear problem (4.4).

The hydraulic conductivity $\kappa(x)$ has small scale features and high-contrast. Because of the small scales and high contrast in the conductivity field, the solution of this system (of size proportional the fine grid points) is prohibitively expensive. Therefore, an adequate robust iterative method is needed.

The construction of robust solvers for high-contrast linear elliptic equation has been considered by many authors. We will use as a preconditioner a two-level domain decomposition method proposed in [26–28], which involves solutions of appropriate local spectral problems. If B^{-1} is the preconditioner, our goal is to have the condition number of $B^{-1}A^n$ bounded independent of the contrast and n (i.e, independent of u_h^n). Now we describe a construction of such preconditioner for (4.7), which will give a robust with respect to the contrast method for Richards' equation.

4.3 Finite element discretization and two level domain decomposition preconditioner

4.3.1 Finite element approximation and local spaces

First, we provide an overview of the use of domain decomposition techniques for constructing preconditioners for multiscale finite element approximations of high-contrast elliptic equations (cf., [26–28, 31, 32]). For an extension to multilevel methods, we refer to [22]. Next, we briefly describe two-level domain decomposition setting that we use and introduce the local spaces and the coarse space.

Let $\mathcal{T}_H$ and $\mathcal{T}_h$ be coarse and fine partitions of Ω into finite elements K (or nonoverlapping subdomains) that consists of triangles, quadrilaterals, etc.. We as-

sume that the coarse elements of $\mathcal{T}_H$ consist of a number of fine elements from $\mathcal{T}_h$. Practically, we first introduce the coarse-grid $\mathcal{T}_H$ and then obtain the fine grid $\mathcal{T}_h$ by partitioning each coarse element into a number of smaller ones. Let Φ_i be the nodal basis of the standard finite element space with respect to the coarse triangulation $\mathcal{T}_H$. We denote by N_c the number of coarse nodes, by $\{y_i\}_{i=1}^{N_c}$ the vertices of the coarse mesh $\mathcal{T}_H$, and define a neighborhood of each node y_i by

$$\omega_i = \bigcup \{K_j \in \mathcal{T}_H; \quad y_i \in \overline{K}_j\}. \tag{4.8}$$

Let $V_0^h(\omega_i) \subset V^h$ be the set of finite element functions with support in ω_i and $R_i^T : V_0^h(\omega_i) \to V^h$ denote the extension by zero operator.

We define, for later use, the one level additive preconditioner (e.g. [48, 71])

$$B_1^{-1} = \sum_{i=1}^{N_v} R_i^T (A_i^0)^{-1} R_i, \tag{4.9}$$

where the operators $A_i^0 : V_0^h(\omega_i) \to V_0^h(\omega_i)$ are defined by

$$(A_i^0 v, w) = a(v, w; u_h^0), \quad \text{for all } v, w \in V_0^h(\omega_i), \quad i = 1, \ldots, N_v. \tag{4.10}$$

The application of the preconditioner B_1^{-1} involves $(A_i^0)^{-1}$ which means solving local problems subdomain-wise in each iteration. The operator A_i^0, defined by the bilinear form $a(\,\cdot\,,\,\cdot\,; u_h^0)$ restricted to $V_0^h(\omega_i)$, is local and invertible.

4.3.2 Coarse space construction

For given M_c number of linearly independent functions $\{\Phi_i\}_{i=1}^{N_c}$ associated with the coarse mesh $\mathcal{T}_H$ (these will be introduced later), we define a coarse space V_0 by

$$V_0 = \text{span}\{\Phi_i\}_{i=1}^{N_c}. \tag{4.11}$$

Below we shall give three choices of sets $\{\Phi_i\}_{i=1}^{N_c}$, that have been already used in the construction of a robust preconditioner for A^n. These are: (1) multiscale coarse space (see, e.g. [17] and the references therein), (2) energy minimizing coarse space (see, e.g. [81]), and (3) a coarse space with local spectral information, (see, e.g. [26–28]). On an abstract level, the main assumption is that $\Phi_i \in V^h$, but the support of each Φ_i is related to the coarse mesh $\mathcal{T}_H$ so that $M_c << \dim V^h$. Below we refer to the Φ_i's as coarse-scale basis functions. The coarse space V_0 defines an operator

$$A_c : V_0 \to V_0, \quad (A_c v, w; u_h^0) = a(v, w; u_h^0), \quad \forall v, w \in V_0.$$

Note that if $R_c^T : V_0 \to V^h$ is the natural interpolation operator then we have

$$A_c = R_c A^0 R_c^T \quad \text{with } A^0 \text{ defined by (4.5) for } n = 0. \tag{4.12}$$

Note that the operator A_c uses the initial guess $u_h^0 \in V^h$ and is constructed only once at the beginning of the fixed point nonlinear iteration. Likewise, the coarse basis functions $\{\Phi_i\}_{i=1}^{N_c}$ are related to the form $a(\,\cdot\,,\,\cdot\,; u_h^0)$ and are constructed only one time. These can be regarded as a preprocessing step. Once the coarse space V_0 is constructed and the coarse-scale operator A_c is defined, we can use the two level additive preconditioner of the form

$$B^{-1} = R_c^T A_c^{-1} R_c + \sum_{i=1}^{N_v} R_i^T (A_i^0)^{-1} R_i = R_c^T A_c^{-1} R_c + B_1^{-1}. \tag{4.13}$$

The preconditioner B^{-1} involves solving one coarse-scale system and N_v local problems in each overlapping subdomain ω_i, $i = 1, \ldots, N_v$. The goal is to reduce the number of iterations in the iterative procedure, e.g., a preconditioned conjugate gradient. An appropriate construction of the coarse space V_0 plays a key role in obtaining robust iterative domain decomposition method. In the next Section 4.3.3 we present

examples of such coarse space constructions. We summarize the fixed point iteration in Algorithm 2.

1: Initialize: Choose $u_h^0 \in V^h$ and compute the residual $r^0 = b - A^0 u_h^0$.
2: Construct the coarse basis $\{\Phi_j\}$, the coarse space V_0 in (4.11), and the coarse operator A_c in (4.12) .
3: **for** $n = 1, 2, \ldots$ until convergence **do**
4: Set the linear system $A^n u_h^{n+1} = b$ (see (4.7)).
5: Using PCG with preconditioner B^{-1} in (4.13) solve the linear system in 4: to get u_h^{n+1}.
6: Compute the residual $r^{n+1} = b - A^{n+1} u_h^{n+1}$.
7: **end for**

Algorithm 2
Fixed point iteration

Remark 1. *In the abstract domain decomposition method setting the overlapping subdomains $\{\omega_i\}$ could be chosen independently of the coarse triangulation $\mathcal{T}^H$. However, in this dissertation, we will only consider the partition introduced above.*

4.3.3 Some multiscale coarse spaces

In this subsection we review several possibilities for construction of coarse basis functions that have been used in design of two level preconditioners that are robust with respect to the contrast.

4.3.3.1 Linear boundary conditions multiscale coarse spaces

Recall the linear boundary multiscale finite element basis functions Φ_i^{ms} from 3.1.2.1. Then we can define linear boundary conditions multiscale coarse spaces V_0^{ms} as following:

$$V_0^{ms} = \text{span}\{\Phi_i^{ms}\}. \tag{4.14}$$

Note that these multiscale basis functions coincide with standard finite element basis functions on the boundaries of coarse grid blocks, while are oscillatory in the interior of each coarse grid block.

4.3.3.2 Energy minimizing coarse spaces

Coarse basis functions can be obtained by minimizing the energy of the basis functions subject to a global constraint (see, [81]). More precise construction of energy minimizing functions Φ_i^{em} are described in 3.1.2.3. Now define energy minimizing coarse space V_0^{em} as

$$V_0^{em} = \text{span}\{\Phi_i^{em}\}. \tag{4.15}$$

We note that the computation of these basis functions requires the solution of a global linear system, a procedure more expensive compared to the local computation of multiscale finite element basis functions with linear boundary conditions Φ_i^{ms}.

4.3.3.3 A coarse space with local spectral information

A coarse space with local spectral information is described in 3.1.3. Local spectral basis functions $\Psi_{i,\ell}$ are computed by selecting a number of eigenvalues from the eigenvalue problem (3.21) and multiplying corresponding eigenvectors by initial multiscale basis functions Φ_i, i.e. $\Psi_{i,\ell} = \Phi_i \psi_\ell^{\omega_i}$. Thus, multiscale space is defined for each i as the span of $\Psi_{i,\ell}$, $\ell = 1, ..., L_i$, where L_i is the number of selected eigenvectors. Denote by V_0^{lsm} the *local spectral multiscale* space

$$V_0^{lsm} = \text{span}\{\Psi_{i,\ell} : 1 \leq \ell \leq L_i \text{ and } 1 \leq i \leq N_v\}. \tag{4.16}$$

4.3.4 Condition number estimates

In this section, we present a theoretical result which shows that the number of outer iterations is independent of the contrast. Recall the bounded ball $V_h^{\mathcal{K},p}$ for a given $\mathcal{K} > 0$ from (2.21). The following three assumptions are used in the proofs of Theorems 3 and 4.

Assumption 2.

(A) See Assumption 1

(B) See (4.21).

Under these assumptions, we show the following theorems concerning the boundedness of the contraction constant.

Theorem 3. *Let the Assumption 2 holds. Then The map $T_h : V_h^{\mathcal{K},p} \to V_h^{\mathcal{K},p}$ is a contraction and the contraction constant is independent of the contrast.*

Proof. We shall show the mapping T_h is contraction and also that the contractivity constant is independent of the contrast.

Suppose $u_h, v_h \in V_h^{\mathcal{K},p}$ satisfy $a(T_h u_h, w; u_h) = F(w)$ and $a(T_h v_h, w; v_h) = F(w)$. Thus,

$$a(T_h u_h, w; u_h) - a(T_h v_h, w; v_h) \quad = \quad 0. \tag{4.17}$$

Since $a(\,\cdot\,,\,\cdot\,,\,\cdot\,)$ is a bilinear form, from equation (4.17) we get

$$a(T_h u_h - T_h v_h, w; u_h) \quad = \quad a(T_h v_h, w; v_h) - a(T_h v_h, w; u_h). \tag{4.18}$$

Now using the definition of $a(\,\cdot\,,\,\cdot\,,\,\cdot\,)$, the right hand side of the equation (4.18) can be written as

$$\begin{aligned}
&\int_\Omega k(x)(\lambda(x,v_h)-\lambda(x,u_h))\nabla T_h v_h \nabla w dx \\
&\quad\le \left(\int_\Omega k(x)(\nabla T_h v_h)^2|\lambda(x,v_h)-\lambda(x,u_h)|^2 dx\right)^{\frac{1}{2}} \left(\int_\Omega k(x)(\nabla w)^2 dx\right)^{\frac{1}{2}} \\
&\quad\le \left(\int_\Omega |k(x)|^q|\nabla T_h v_h|^{2q} dx\right)^{\frac{1}{2q}} \text{(By Hőlder's inequality, } \frac{1}{q}+\frac{1}{q'}=1) \\
&\qquad\cdot \left(\int_\Omega |\lambda(x,v_h)-\lambda(x,u_h)|^{2q'} dx\right)^{\frac{1}{2q'}} \left(\int_\Omega k(x)(\nabla w)^2 dx\right)^{\frac{1}{2}} := A
\end{aligned} \tag{4.19}$$

Then using Lipschitz continuity of λ,

$$\begin{aligned}
A = \quad &\le \left(\int_\Omega |k(x)|^q|\nabla T_h v_h|^{2q} dx\right)^{\frac{1}{2q}} \left(C_1\int_\Omega |v_h-u_h|^{2q'} dx\right)^{\frac{1}{2q'}} \\
&\qquad\cdot \left(\int_\Omega k(x)(\nabla w)^2 dx\right)^{\frac{1}{2}} \\
&\le \left(\int_\Omega |k(x)|^q|\nabla T_h v_h|^{2q} dx\right)^{\frac{1}{2q}} \left(C_1 C_{2q'}\int_\Omega (\nabla(v_h-u_h))^2 dx\right)^{\frac{1}{2}} \\
&\qquad\cdot \left(\int_\Omega k(x)(\nabla w)^2 dx\right)^{\frac{1}{2}}, \quad \text{(by Sobolev inequality),}
\end{aligned} \tag{4.20}$$

where we have used the Sobolev inequality $\|u\|_{L^{2q'}(\Omega)} \le C_{2q'}\|Du\|_{L^2(\Omega)}$ with $2q' \in [1,\infty]$ for function u with bounded mean oscillation. Next, we want to bound $\left(\int_\Omega |k(x)|^q|\nabla T_h v_h|^{2q} dx\right)^{\frac{1}{2q}}$ with some constant which is independent of the contrast, i.e., the constant doesn't depend on $k(\,\cdot\,)$.

Now we make the following assumption.

Assumption 3. *Given the equation $a(T_h v_h, T_h v_h, v_h) = F(T_h v_h)$ (see (4.3)), we assume that*

$$\int_\Omega (k(x)|\nabla T_h v_h|^2)^{q/2} dx \le C_F^q, \tag{4.21}$$

where $C_F^q \to 0$ as $\|F\|_{W_q^{-1}(\Omega)} \to 0$ for some $q > 2$.

We note that when $F = 0$ then $C_F^2 = 0$, thus, $T_h v_h$ is zero almost everywhere. Moreover, if $\|F\|_{W_2^{-1}(\Omega)}$ is small, then C_F^2 is small and C_F^2 converges to zero as $\|F\|_{W_2^{-1}(\Omega)}$ goes to zero. The inequality (4.21) assumes that we have continuity of C_F^q with respect to $\|F\|_{W_2^{-1}(\Omega)}$ for any $q > 2$ that is sufficiently close to 2. We note that $\|T_h v_h\|_{W_q^1(\Omega)}$ is bounded by $\|F\|_{W_q^{-1}(\Omega)}$ as shown above. This is typically used to show the contractivity of the map T_h.

Now, we can conclude that

$$\begin{aligned}\int_\Omega k(x)(\lambda(x,v_h)-\lambda(x,u_h))\nabla T_h v_h \nabla w dx \\ \leq C\left(\int_\Omega (\nabla(v_h-u_h))^2 dx\right)^{\frac{1}{2}}\left(\int_\Omega k(x)(\nabla w)^2 dx\right)^{\frac{1}{2}},\end{aligned} \tag{4.22}$$

where the constant C depends on Lipschitz constant C_1.

Now put $w = T_h u_h - T_h v_h$, then left hand side of (4.18) is bounded below,

$$\begin{aligned}a(T_h u_h - T_h v_h, T_h u_h - T_h v_h, u_h) &= \int_\Omega (k(x)\lambda(x,u_h)(\nabla(T_h u_h - T_h v_h))^2 dx \\ &\geq C_2 \int_\Omega k(x)(\nabla(T_h u_h - T_h v_h))^2 dx.\end{aligned} \tag{4.23}$$

Combine equations (4.22) and (4.23), then we get

$$\begin{aligned}\int_\Omega k(x)(\nabla(T_h u_h - T_h v_h))^2 dx &\leq C_2^{-1} C\left(\int_\Omega (\nabla(v_h-u_h))^2 dx\right)^{\frac{1}{2}} \\ &\quad \cdot \left(\int_\Omega k(x)(\nabla(T_h u_h - T_h v_h))^2 dx\right)^{\frac{1}{2}}.\end{aligned}$$

Then using the Assumption 2 (a), we get

$$C_0^{\frac{1}{2}}\left(\int_\Omega (\nabla(T_h u_h - T_h v_h)^2 dx\right)^{\frac{1}{2}} \leq C_2^{-1} C\left(\int_\Omega (\nabla(v_h-u_h))^2 dx\right)^{\frac{1}{2}}.$$

So we can deduce that

$$|T_h u_h - T_h v_h|_{W_2^1} \leq C_0^{-\frac{1}{2}} C_2^{-1} C |u_h - v_h|_{W_2^1}, \tag{4.24}$$

i.e., the mapping T_h is a contraction if C is chosen sufficiently small (see Assumption 3). □

Theorem 4. *Under the assumptions of Theorem 1, we have* $cond(B^{-1}A^n) \leq C$, *where* C *is independent of the contrast.*

Proof. From Lemma 1 and Lemma 10 of [26] we have that there is a *stable decomposition*, that is, there exists $v_0 \in V_0^{lsm}$, $v_i \in V_0^h(\omega_i)$, $i = 1, \ldots, N_v$, such that

$$\int_D \kappa |\nabla v_0| + \sum_{i=1}^{N_v} \int_{\omega_i} \kappa |\nabla v_i|^2 \leq C_0 \left(1 + \frac{1}{H^2 \mu_{L+1}}\right) \int_D \kappa |\nabla v|^2,$$

for some positive constant independent of the contrast and $\mu_{L+1} = \min_i \mu_{L_i+1}$. Here we select the first L_i smallest eigenvalues of (3.21). Then, for a fixed w we have stable decomposition,

$$\begin{aligned} &\int_D \lambda(x,w)\kappa(x)|\nabla v_0(x)| + \sum_{i=1}^{N_v} \int_{\omega_i} \lambda(x,w)\kappa(x)|\nabla v_i|^2 \\ &\quad \leq \quad C_0 \frac{\max_{x\in D} \lambda(x,w)}{\min_{x\in D} \lambda(x,\omega)} \left(1 + \frac{1}{H^2 \mu_{L+1}}\right) \int_D \lambda(x,w) k(x) |\nabla v|^2. \end{aligned}$$

According to the abstract theory of domain decomposition, see [48, 71], we conclude that the condition number of the preconditioned matrix is of order

$$cond(B^{-1}A) \leq C_0 \frac{\max_{x\in D} \lambda(x,w)}{\min_{x\in D} \lambda(x,\omega)} \left(1 + \frac{1}{H^2 \mu_{L+1}}\right).$$

Further noting that the number of nonlinear outer iterations is bounded (see Theorem 2), we conclude that the proposed iterative procedure converges independent of the contrast. □

Remark 2 (Degenerate case)**.** *In some practical cases the diffusion coefficient of the Richards' equation can approach to zero. For the analysis of mixed finite element approximations of such problems we refer to e.g. [1, 61]. In general the proposed in this paper methods are not directly applicable. Nevertheless, in some cases, our algorithms will yield robust results. These cases includes coefficients where the regions of degeneracy can be resolved by the coarse grid, or when those regions are union of isolated inclusions. However, when the region of degeneracy is complex and cannot be resolved by the coarse grid, additional basis functions or some refinement may be required in order to get contrast-independent preconditioners. And since we use a fixed point iteration, our outer loop will allow us to determine approximately the region of degeneracy. We note that this requires additional analysis and is a subject of future research.*

4.4 Numerical examples

In this section we present some representative numerical examples. We solve the Richards' equation (4.2) in $\Omega = [0,1]\times[0,1]$ with $f(x) = 1$ and homogeneous Dirichlet boundary conditions. We consider several models for the hydraulic conductivity: the Haverkamp, van Genuchten, and Exponential model (see, e.g. [11, 33, 59, 73]), as introduced below. The coarse mesh $\mathcal{T}_H$ is obtained by dividing Ω into 10×10 mesh. The fine triangulation is obtained by dividing each coarse-mesh element into 10×10 squares and further dividing each square into two triangles. Thus, the fine-mesh step

size is $h = 1/100$. In all the numerical experiments we use the initial approximation for the iterative process u_h^0 that solves

$$a(u_h^0, v; 0) = F(v), \quad \text{for all } v \in V^h. \tag{4.25}$$

We apply the Algorithm (2). As stated in Algorithm (2) we use the preconditioner B^{-1} in (4.13) with three different coarse spaces:

1. The coarse space V_0^{ms} described in Section 4.3.3.1. In this case B^{-1} is denoted by B_{ms}^{-1};

2. The coarse space V_0^{em} described in Section 4.3.3.2. In this case B^{-1} is denoted by B_{em}^{-1};

3. The coarse space V_0^{lsm} described in Section 4.3.3.3. In this case B^{-1} is denoted by B_{lsm}^{-1}.

We study the performance of Algorithm 2 with initial guess u_h^0 and preconditioners B_{ms}^{-1}, B_{em}^{-1}, and B_{lsm}^{-1}. We consider different permeabilities with complex high-contrast configurations, see Figure 3.4. A number of parameter values in the nonlinearity of the hydraulic conductivity are tested in our simulations. In particular, for each experiment we chose a different set of parameters for the model and a set of contrast values for the hydraulic conductivity. We note that, for each outer iteration in Algorithm 2 we have a PCG iteration. The inner PCG iteration is convergent when the initial residual is reduced by a factor of $tol_{in} = 1e-10$ while the outer tolerance is set to $tol_{out} = 1e-8$.

We consider the following indicators for the performance of the preconditioners:

- Coarse space dimension;

- The number of outer iterations of the nonlinear fixed point iteration (R-iter);

- The maximum and minimum number of inner PCG iterations over all outer iterations (CG-iter) and the estimated maximum condition numbers (Cond).

We also verify numerically our main assumption in the proof of Theorem 1. That is, for every outer iteration update we compute $EN := \|\sqrt{\kappa}|\nabla u|\|_p^p = \int_D (\sqrt{\kappa}|\nabla u|)^p \, dx$, $p = 1, 2, 3, \ldots, 10$. We observe that this quantity remains bounded in all experiments.

4.4.1 Haverkamp model

First, we will study Haverkamp model. In the Haverkamp model, (see, e.g. [33]), the hydraulic conductivity is given by

$$k(x, u) = k_s(x) \frac{A}{A + (|u|/B)^\gamma}. \tag{4.26}$$

We present the first set of numerical results in Tables 4.1 and 4.2. We use the preconditioner B_{ms}^{-1} based on the coarse space V_0^{ms}. We observe from these tables that the numbers of outer iterations do not change when the contrast value η increases. However, the condition number of the preconditioned system grows as η increases. We also observe that the quantity $\|\sqrt{\kappa}|\nabla u|\|_p^p, p = 1, 2, 3, \ldots, 10$, that is related to the number of outer iterations, is bounded. We observe that the number of outer iterations is larger when B and γ (see (4.26)) decrease. This is because the smaller values of B and γ increase the magnitude of the conductivity that comes from its nonlinear component. Comparing Tables 4.1 and 4.2 that use different conductivity fields, we see that the condition numbers in Table 4.2 are smaller than the condition numbers in Table 4.1. This is because conductivity field 2 (see Figure 3.4) has simpler heterogeneity structure compared conductivity field 1.

Next, we repeat the above numerical experiments using the preconditioner B_{em}^{-1} based on the coarse space V_0^{em}. Numerical results are presented in Tables 4.3 and 4.4. We observe that, as before, the number of outer iterations is fixed with increasing η. On the other hand, the condition number of the PCG iteration grows as the contrast

increases. This condition number is much larger compared to the case when spectral basis functions are used as presented in the next tables.

Further, we show the numerical experiment using the preconditioner B_{lsm}^{-1} based on the spectral coarse space V_0^{lsm}. Numerical results are presented in Tables 4.5 and 4.6. As before, we observe that the number of outer iterations is independent of the contrast. We observe that the condition number is also independent of the contrast. Note that the condition number is substantially smaller than the one of the preconditioned system using B_{ms}^{-1} or B_{em}^{-1}. In general, the number of inner PCG iterations is much smaller compared to those when other coarse spaces are used.

Table 4.1
Numerical results for preconditioner B_{ms}^{-1}. Here we use the Haverkamp model $\kappa(x,u) = \kappa(x)\frac{A}{A+(|u|/B)^{\gamma}}$ with κ depicted in left picture of Figure 3.4. The coarse space dimension is 81.

	$A=1,\ B=1,\ \gamma=1$				$A=1,\ B=0.01,\ \gamma=0.5$			
η	R-iter	CG-iter	Max Cond	EN	R-iter	CG-iter	Max Cond	EN
10^3	4	$119,124$	$1.1e+3$	0.12	11	$131,139$	$1.4e+3$	0.12
10^4	4	$166,178$	$1.1e+4$	0.12	11	$179,199$	$1.4e+4$	0.12
10^5	4	$224,224$	$1.1e+5$	0.12	11	$224,224$	$1.4e+5$	0.12
10^6	4	$278,278$	$1.1e+6$	0.12	11	$278,278$	$1.4e+6$	0.12

4.4.2 van Genuchten Model

Next, we consider *Van Genuchten model* (see [73]) that is one of widely used empirical constitutive relations. In this model, the hydraulic conductivity is given by

$$k(x,u) = k_s(x)\frac{\{1-(\alpha|u|/B)^{n-1}[1+(\alpha|u|)^n]^{-m}\}^2}{[1+(\alpha|u|)^n]^{\frac{m}{2}}}. \tag{4.27}$$

Table 4.2

Numerical results for preconditioner B_{ms}^{-1}. Here we use the Haverkamp model $\kappa(x,u) = \kappa(x)\frac{A}{A+(|u|/B)^{\gamma}}$ with κ depicted in right picture of Figure 3.4. The coarse space dimension is 81.

	$A=1,\ B=1,\ \gamma=1$				$A=1,\ B=0.01,\ \gamma=0.5$			
η	R-iter	CG-iter	Max Cond	EN	R-iter	CG-iter	Max Cond	EN
10^3	4	$113, 113$	$2.6e+2$	0.13	11	$107, 123$	$3.9e+2$	0.13
10^4	4	$163, 171$	$2.5e+3$	0.13	11	$180, 193$	$3.6e+3$	0.13
10^5	4	$224, 232$	$2.5e+4$	0.13	11	$238, 255$	$3.6e+4$	0.13
10^6	4	$288, 295$	$2.5e+5$	0.13	11	$308, 324$	$3.6e+5$	0.13

Table 4.3

Numerical results for preconditioner B_{em}^{-1}. Here we use the Haverkamp model $\kappa(x,u) = \kappa(x)\frac{A}{A+(|u|/B)^{\gamma}}$ with κ depicted in left picture of Figure 3.4. The coarse space dimension is 81.

	$A=1,\ B=1,\ \gamma=1$				$A=1,\ B=0.01,\ \gamma=0.5$			
η	R-iter	CG-iter	Max Cond	EN	R-iter	CG-iter	Max Cond	EN
10^3	3	$83, 83$	$1.3e+2$	0.12	8	$90, 102$	$1.9e+2$	0.12
10^4	3	$88, 88$	$2.5e+2$	0.12	8	$95, 109$	$3.9e+2$	0.12
10^5	3	$89, 90$	$3.0e+2$	0.12	8	$97, 113$	$4.6e+2$	0.12
10^6	3	$95, 103$	$3.1e+2$	0.12	8	$103, 115$	$4.7e+2$	0.12

Table 4.4

Numerical results for preconditioner B_{em}^{-1}. Here we use the Haverkamp model $\kappa(x,u) = \kappa(x)\frac{A}{A+(|u|/B)^{\gamma}}$ with κ depicted in right picture of Figure 3.4. The coarse space dimension is 81.

	$A=1,\ B=1,\ \gamma=1$				$A=1,\ B=0.01,\ \gamma=0.5$			
η	R-iter	CG-iter	Max Cond	EN	R-iter	CG-iter	Max Cond	EN
10^3	3	$90, 90$	$1.6e+2$	0.13	8	$84, 98$	$2.7e+2$	0.13
10^4	3	$94, 94$	$3.7e+2$	0.13	8	$88, 102$	$6.2e+2$	0.13
10^5	3	$95, 95$	$4.2e+2$	0.13	8	$89, 103$	$7.1e+2$	0.13
10^6	3	$96, 96$	$4.3e+2$	0.13	8	$91, 104$	$7.2e+2$	0.13

Table 4.5

Numerical results for preconditioner B_{lsm}^{-1}. Here we use the Haverkamp model $\kappa(x,u) = \kappa(x)\frac{A}{A+(|u|/B)^{\gamma}}$ with κ depicted in left picture of Figure 3.4. The coarse space dimension is 166.

	$A=1,\ B=1,\ \gamma=1$				$A=1,\ B=0.01,\ \gamma=0.5$			
η	R-iter	CG-iter	Max Cond	EN	R-iter	CG-iter	Max Cond	EN
10^3	4	$34, 34$	6.9	0.13	8	$37, 39$	9.6	0.13
10^4	4	$35, 35$	7.0	0.13	8	$39, 41$	9.7	0.13
10^5	4	$35, 37$	7.0	0.13	8	$40, 42$	9.7	0.13
10^6	4	$36, 36$	7.0	0.13	8	$41, 44$	9.7	0.13

Table 4.6

Numerical results for preconditioner B_{lsm}^{-1}. Here we use the Haverkamp model $\kappa(x,u) = \kappa(x)\frac{A}{A+(|u|/B)^{\gamma}}$ with κ depicted in right picture of Figure 3.4. The coarse space dimension is 184.

	$A=1,\ B=1,\ \gamma=1$				$A=1,\ B=0.01,\ \gamma=0.5$			
η	R-iter	CG-iter	Max Cond	EN	R-iter	CG-iter	Max Cond	EN
10^3	3	$31, 31$	6.2	0.13	8	$35, 37$	8.1	0.13
10^4	3	$33, 33$	6.3	0.13	8	$36, 37$	8.0	0.13
10^5	3	$33, 33$	6.3	0.13	8	$38, 43$	8.0	0.13
10^6	3	$34, 34$	6.3	0.13	8	$38, 41$	8.0	0.13

As before, we will present numerical results for all three coarse spaces. First, in Tables 4.7 and 4.8 we present the numerical results for the preconditioner B_{ms}^{-1}. We observe that the number of outer iterations is smaller compared to the other two models. The number of outer iterations stays the same while increasing η. On the other hand, the condition number of the linearized system increases as η increases. We observe that the value $\|\sqrt{\kappa}|\nabla u|\|_p^p, p = 1, 2, 3, \ldots, 10$ is bounded independent of the contrast. Now we compare Table 4.7 and Table 4.8 for two different conductivity fields depicted in Figure 3.4. We observe that the condition numbers presented in

Table 4.8 is smaller than those presented in Table 4.7 which is consistent with our previous observations.

Numerical results for the preconditioner B_{em}^{-1} are presented in Tables 4.9 and 4.10, while numerical results for the preconditioner B_{lsm}^{-1} are presented in Tables 4.11 and 4.12. As before, we observe that the number of outer iteration does not change with η increasing. However, the condition number of the inner iteration is increasing for B_{em}^{-1}, while the condition number of the inner iteration does not change (and is much smaller) for B_{lsm}^{-1}.

Table 4.7

Numerical results for preconditioner B_{ms}^{-1}. Here we use the van Genuchten model $\kappa(x,u) = \kappa(x)\frac{\{1-(\alpha(|u|/B))^{n-1}[1+(\alpha(|u|/B))^n]^{-m}\}^2}{[1+(\alpha(|u|/B))^n]^{m/2}}$ with κ depicted in left picture of Figure 3.4. The coarse space dimension is 81.

	$\alpha = 0.005$, $B = 1$, $n = 2$, $m = 0.5$				$\alpha = 0.01$, $B = 1$, $n = 4$, $m = 0.75$			
η	R-iter	CG-iter	Max Cond	EN	R-iter	CG-iter	Max Cond	EN
10^3	2	116, 116	$1.1e+3$	0.13	2	115, 116	$1.1e+3$	0.13
10^4	2	168, 168	$1.1e+4$	0.13	2	174, 174	$1.1e+4$	0.13
10^5	2	219, 219	$1.1e+5$	0.13	2	219, 219	$1.1e+5$	0.13
10^6	2	273, 290	$1.1e+6$	0.13	2	267, 272	$1.1e+6$	0.13

4.4.3 Exponential Model

Finally, we present numerical results for exponential model. Here the hydraulic conductivity depend exponentially on the pressure head u, that is,

$$k(x,u) = k_s(x)e^{\alpha u/B}. \tag{4.28}$$

Table 4.8

Numerical results for preconditioner B_{ms}^{-1}. Here we use the van Genuchten model $\kappa(x,u) = \kappa(x)\frac{\{1-(\alpha(|u|/B))^{n-1}[1+(\alpha(|u|/B))^n]^{-m}\}^2}{[1+(\alpha(|u|/B))^n]^{m/2}}$ with κ depicted in right picture of Figure 3.4. The coarse space dimension is 81.

	$\alpha = 0.005$, $B = 1$, $n = 2$, $m = 0.5$				$\alpha = 0.01$, $B = 1$, $n = 4$, $m = 0.75$			
η	R-iter	CG-iter	Max Cond	EN	R-iter	CG-iter	Max Cond	EN
10^3	2	$98, 99$	$2.5e+2$	0.13	2	$99, 99$	$2.5e+2$	0.13
10^4	2	$134, 134$	$2.5e+3$	0.13	2	$160, 160$	$2.5e+3$	0.13
10^5	2	$183, 184$	$2.5e+4$	0.13	2	$219, 223$	$2.5e+4$	0.13
10^6	2	$222, 225$	$2.5e+5$	0.13	2	$286, 287$	$2.5e+5$	0.13

Table 4.9

Numerical results for preconditioner B_{em}^{-1}. Here we use the van Genuchten model $\kappa(x,u) = \kappa(x)\frac{\{1-(\alpha(|u|/B))^{n-1}[1+(\alpha(|u|/B))^n]^{-m}\}^2}{[1+(\alpha(|u|/B))^n]^{m/2}}$ with κ depicted in left picture of Figure 3.4. The coarse space dimension is 81.

	$\alpha = 0.005$, $B = 1$, $n = 2$, $m = 0.5$				$\alpha = 0.01$, $B = 1$, $n = 4$, $m = 0.75$			
η	R-iter	CG-iter	Max Cond	EN	R-iter	CG-iter	Max Cond	EN
10^3	2	$82, 82$	$1.3e+2$	0.13	1	81	$1.3e+2$	0.13
10^4	2	$85, 85$	$2.5e+2$	0.13	1	84	$2.5e+2$	0.13
10^5	2	$88, 88$	$3.0e+2$	0.13	1	87	$3.0e+2$	0.13
10^6	2	$93, 101$	$3.1e+2$	0.13	1	95	$3.1e+2$	0.13

This nonlinear equation can also be derived by homogenizing Stokes equation in porous media when the fluid viscosity exponentially depends on the pressure [59].

We present the first set of numerical results in Tables 4.13 and 4.14. First, we use the preconditioner B_{ms}^{-1} based on the coarse space V_0^{ms}. We observe that the number of the outer iterations does not change when the contrast η increases. However, the condition number of the preconditioned system increases proportional to η. We also observe that the quantity $\|\sqrt{\kappa}|\nabla u|\|_p^p, p = 1, 2, 3, \ldots, 10$ is bounded independent of contrast η. We see that the number of outer iterations stays the same for both set

Table 4.10

Numerical results for preconditioner B_{em}^{-1}. Here we use the van Genuchten model $\kappa(x,u) = \kappa(x)\frac{\{1-(\alpha(|u|/B))^{n-1}[1+(\alpha(|u|/B))^n]^{-m}\}^2}{[1+(\alpha(|u|/B))^n]^{m/2}}$ with κ depicted in right picture of Figure 3.4. The coarse space dimension is 81.

	$\alpha = 0.005$, $B = 1$, $n = 2$, $m = 0.5$				$\alpha = 0.01$, $B = 1$, $n = 4$, $m = 0.75$			
η	R-iter	CG-iter	Max Cond	EN	R-iter	CG-iter	Max Cond	EN
10^3	2	76, 76	$1.6e+2$	0.13	1	88	$1.6e+2$	0.13
10^4	2	79, 79	$3.6e+2$	0.13	1	90	$3.6e+2$	0.13
10^5	2	79, 79	$4.2e+2$	0.13	1	87	$4.1e+2$	0.13
10^6	2	80, 81	$4.2e+2$	0.13	1	90	$4.2e+2$	0.13

Table 4.11

Numerical results for preconditioner B_{lsm}^{-1}. Here we use the van Genuchten model $\kappa(x,u) = \kappa(x)\frac{\{1-(\alpha(|u|/B))^{n-1}[1+(\alpha(|u|/B))^n]^{-m}\}^2}{[1+(\alpha(|u|/B))^n]^{m/2}}$ with κ depicted in left picture of Figure 3.4. The coarse space dimension is 166.

	$\alpha = 0.005$, $B = 1$, $n = 2$, $m = 0.5$				$\alpha = 0.01$, $B = 1$, $n = 4$, $m = 0.75$			
η	R-iter	CG-iter	Max Cond	EN	R-iter	CG-iter	Max Cond	EN
10^3	2	33, 33	6.8	0.13	1	33	6.8	0.13
10^4	2	34, 34	6.8	0.13	1	34	6.8	0.13
10^5	2	35, 35	6.8	0.13	1	35	6.8	0.13
10^6	2	36, 36	6.8	0.13	1	36	6.8	0.13

of parameters for nonlinearities which means larger α values do not affect the outer iterations. We observe from Tables 4.13 and 4.14 (these use different conductivity fields) that the condition numbers in Table 4.14 are smaller than the corresponding condition numbers in Table 4.13. This is because conductivity field 2 has simpler subgrid structure compared to conductivity field 1.

Next, we repeat the numerical experiment using the preconditioner B_{em}^{-1} based on the coarse space V_0^{em} and B_{lsm}^{-1} with coarse space V_0^{lsm}. Numerical results for the coarse space B_{em}^{-1} are presented in Tables 4.15 and 4.16 while the results for V_0^{lsm}

Table 4.12

Numerical results for preconditioner B_{lsm}^{-1}. Here we use the van Genuchten model $\kappa(x,u) = \kappa(x)\frac{\{1-(\alpha(|u|/B))^{n-1}[1+(\alpha(|u|/B))^n]^{-m}\}^2}{[1+(\alpha(|u|/B))^n]^{m/2}}$ with κ depicted in right picture of Figure 3.4. The coarse space dimension is 184.

	$\alpha = 0.005$, $B = 1$, $n = 2$, $m = 0.5$				$\alpha = 0.01$, $B = 1$, $n = 4$, $m = 0.75$			
η	R-iter	CG-iter	Max Cond	EN	R-iter	CG-iter	Max Cond	EN
10^3	2	32, 32	6.5	0.14	1	31	6.2	0.13
10^4	2	33, 33	6.6	0.13	1	32	6.3	0.13
10^5	2	33, 33	6.6	0.13	1	33	6.3	0.13
10^6	2	35, 35	6.6	0.13	1	34	6.3	0.13

are presented in Tables 4.17 and 4.18. As before, we observe that the number of outer iterations is independent of the contrast. However, for space V_0^{em} the condition number increases as we increase the contrast. On the other hand, the condition number is independent of contrast when V_0^{lsm} is used as a coarse space. Moreover, we observe that the condition number produced by V_0^{lsm}, is only 6 while the condition number for V_0^{em} is about 400 for $\eta = 10^6$. In conclusion, B_{lsm}^{-1} provides truly independent-of-contrast solver.

Table 4.13

Numerical results for preconditioner B_{ms}^{-1}. Here we use the Exponential model $\kappa(x,u) = \kappa(x)e^{\alpha(u/B)}$ with κ depicted in left picture of Figure 3.4. The coarse space dimension is 81.

	$\alpha = 1$, $B = 1$				$\alpha = 2$, $B = 1$			
η	R-iter	CG-iter	Max Cond	EN	R-iter	CG-iter	Max Cond	EN
10^3	4	119, 120	$1.0e+3$	0.13	4	120, 122	$1.1e+3$	0.13
10^4	4	166, 178	$1.1e+4$	0.13	4	173, 181	$1.1e+4$	0.13
10^5	4	224, 224	$1.1e+5$	0.13	4	226, 227	$1.1e+5$	0.13
10^6	4	274, 284	$1.1e+6$	0.13	4	277, 287	$1.1e+6$	0.13

Table 4.14

Numerical results for preconditioner B_{ms}^{-1}. Here we use the Exponential model $\kappa(x,u) = \kappa(x)e^{\alpha(u/B)}$ with κ depicted in right picture of Figure 3.4. The coarse space dimension is 81.

	$\alpha = 1,\ B = 1$				$\alpha = 2,\ B = 1$			
η	R-iter	CG-iter	Max Cond	EN	R-iter	CG-iter	Max Cond	EN
10^3	4	$113, 113$	$2.5e+2$	0.13	4	$114, 115$	$2.5e+2$	0.13
10^4	4	$164, 164$	$2.5e+3$	0.13	4	$164, 164$	$2.5e+3$	0.13
10^5	4	$223, 232$	$2.5e+4$	0.13	4	$227, 231$	$2.5e+4$	0.13
10^6	4	$290, 294$	$2.5e+5$	0.13	4	$289, 302$	$2.5e+5$	0.13

Table 4.15

Numerical results for preconditioner B_{em}^{-1}. Here we use the Exponential model $\kappa(x,u) = \kappa(x)e^{\alpha(u/B)}$ with κ depicted in left picture of Figure 3.4. The coarse space dimension is 81.

	$\alpha = 1,\ B = 1$				$\alpha = 2,\ B = 1$			
η	R-iter	CG-iter	Max Cond	EN	R-iter	CG-iter	Max Cond	EN
10^3	3	$83, 84$	$1.3e+2$	0.13	3	$84, 84$	$1.3e+2$	0.13
10^4	3	$88, 88$	$2.5e+2$	0.13	3	$89, 90$	$2.6e+2$	0.13
10^5	3	$90, 91$	$3.0e+2$	0.13	3	$92, 92$	$3.1e+2$	0.13
10^6	3	$96, 97$	$3.1e+2$	0.13	3	$97, 98$	$3.1e+2$	0.13

Table 4.16

Numerical results for preconditioner B_{em}^{-1}. Here we use the Exponential model $\kappa(x,u) = \kappa(x)e^{\alpha(u/B)}$ with κ depicted in right picture of Figure 3.4. The coarse space dimension is 81.

	$\alpha = 1,\ B = 1$				$\alpha = 2,\ B = 1$			
η	R-iter	CG-iter	Max Cond	EN	R-iter	CG-iter	Max Cond	EN
10^3	3	$91, 91$	$1.6e+2$	0.13	3	$91, 92$	$1.6e+2$	0.13
10^4	3	$95, 95$	$3.6e+2$	0.13	3	$95, 96$	$3.7e+2$	0.13
10^5	3	$95, 95$	$4.2e+2$	0.13	3	$98, 99$	$4.2e+2$	0.13
10^6	3	$98, 98$	$4.2e+2$	0.13	3	$99, 99$	$4.3e+2$	0.13

Table 4.17
Numerical results for preconditioner B_{lsm}^{-1}. Here we use the Exponential model $\kappa(x,u) = \kappa(x)e^{\alpha(u/B)}$ with κ depicted in left picture of Figure 3.4. The coarse space dimension is 166.

	$\alpha = 1,\ B = 1$				$\alpha = 2,\ B = 1$			
η	R-iter	CG-iter	Max Cond	EN	R-iter	CG-iter	Max Cond	EN
10^3	3	33, 33	6.8	0.13	3	34, 34	6.8	0.13
10^4	3	35, 35	6.8	0.13	3	35, 35	6.8	0.13
10^5	3	36, 36	6.8	0.13	3	36, 36	6.9	0.13
10^6	3	37, 37	6.8	0.13	3	37, 37	6.9	0.13

Table 4.18
Numerical results for preconditioner B_{lsm}^{-1}. Here we use the Exponential model $\kappa(x,u) = \kappa(x)e^{\alpha(u/B)}$ with κ depicted in right picture of Figure 3.4. The coarse space dimension is 184.

	$\alpha = 1,\ B = 1$				$\alpha = 2,\ B = 1$			
η	R-iter	CG-iter	Max Cond	EN	R-iter	CG-iter	Max Cond	EN
10^3	3	32, 32	6.4	0.13	3	32, 32	6.6	0.13
10^4	3	34, 34	6.8	0.13	3	34, 34	6.7	0.13
10^5	3	34, 34	6.5	0.13	3	35, 35	6.7	0.13
10^6	3	36, 36	6.8	0.13	3	35, 36	6.7	0.13

4.5 Reduced Basis approach for nonseparable coefficient Richards' equation

In this section, we apply the approach in Section 3.3 to construct robust iterative methods of the domain decomposition type. Recall that this approach helps to construct a reduced dimensional local approximation that allows quick computation of the local spectral problem. This process is also necessary for constructing a robust iterative method of the fine-scale system of nonseparable coefficient Richards' equation.

For every outer iteration, we need to solve inner iterations using a preconditioner which is constructed based on the local spectral space. However, when the equation is

parameter-dependent, then constructing a preconditioner for each parameter value is extremely computationaly expensive. Therefore, we need the RB technique to reduce this computational time in multiscale basis calculations.

In this approach, via offline computations, one reduces the costs associated with computing local solutions. We speed-up basis computations reducing these parameter-dependent problems to much smaller dimensional ones at the online stage. Moreover we have lower computational costs compared to standard RB approaches that use global solutions.

4.5.1 Problem setting

Recall the nonseparable coefficient Richards' equation (3.39). To get a fine-scale solution of the nonlinear elliptic equation, we use the fixed point iteration, (3.52), and this can be considered as parameter-dependent elliptic equation, like (3.53).

We will use coarse spaces constructed via the RB techniques (see Section 3.3.2.5) in two-level additive Schwarz preconditioners. We brefly describe the two-level domain decomposition setting. Let $\mathcal{T}_H$ and $\mathcal{T}_h$ be coarse and fine partitions of Ω into finite elements K (or nonoverlapping subdomains) that consists of triangles, quadrilaterals, etc.. We assume that the coarse elements of $\mathcal{T}_H$ consist of a number of fine elements from $\mathcal{T}_h$. Define V^h be the finite dimensional subspace of H_0^1 of piece-wise polynomials with respect to $\mathcal{T}_h$. Let Φ_i be the nodal basis of the standard finite element space with respect to the coarse triangulation $\mathcal{T}_H$. We denote by N_v the number of coarse nodes, by $\{y_i\}_{i=1}^{N_v}$ the vertices of the coarse mesh $\mathcal{T}_H$, and define a neighborhood of each node y_i by ω_i which is defined on (4.8). Let $V_0^h(\omega_i) \subset V^h$ be the set of finite element functions with support in ω_i and $R_i^T : V_0^h(\omega_i) \to V^h$ denote the extension by zero operator.

We consider iterative methods to find the solution of the fine-grid finite element problem

$$a(u, v; \mu) = f(v), \qquad \forall v \in V^h, \tag{4.29}$$

where the bilinear form a defined in (3.56). The matrix of this linear systems is written as

$$A(\mu)u(\mu) = b. \tag{4.30}$$

We note that we are representing finite element functions and their vector of coordinates by the same symbols. We can solve the fine-scale linear system iteratively with the preconditioned conjugate gradient(PCG) method. Any other suitable iterative scheme can be used as well. We introduce the two level additive preconditioner of the form

$$B^{-1}(\mu) = R_0^T(\mu)A_0^{-1}(\mu)R_0(\mu) + \sum_{i=1}^{N_v} R_i(\mu)^T A_i^{-1}(\mu)R_i(\mu), \tag{4.31}$$

where the local matrices are defined by

$$uA_i(\mu)v = a(u, v; \mu), \qquad \text{for all} u, v \in V_0^i(\omega_i). \tag{4.32}$$

The coarse projection matrix R_0^T is defined by $R_0^T = R_0^T(\mu) = [\Phi_1, \ldots, \Phi_{N_c}]$ and $A_0(\mu) = R_0(\mu)A(\mu)R_0^T(\mu)$. The columns Φ_i's are fine-grid coordinate vectors corresponding to the basis functions$\{\Phi_i^j\}$. We will use RB procedures to construct the coarse space basis, i.e., $R_0(\mu)$, for any given value of the parameter μ.

4.5.2 Constructing preconditioner using RB techniques

In Section 4.3, we observed that constructing a right coarse space $V_0 = \text{span}\{\Phi_i\}_{i=1}^{N_c}$ plays a key role in obtaining robust iterative domain decomposition method. This fact is also applicapable to the parameter-dependent fine-scale problem. We are going to use local spectral space for our preconditioner. The problem is, when you solve the local spectral problem on each subdomain to get multiscale bases, it can be expensive to compute realization-by-realization, i.e. computing multiscale bases for each realization could be very high-priced.

Therefore, we use earlier proposed technique to get a reduced-dimensional eigenvalue problem and we can rapidly solve the local spectral problem to constrct multiscale basis functions. This approach is especially can be effectively used when the parameter μ is a coarse-grid spatial function, i.e. μ varies on a coarse grid.

We compute the coarse space using the method described in Section 3.3.2. Using these coarse spaces we can design preconditioners for our fine-scale systems of equation. Two-level additive Schwarz preconditioners are constructed

$$B^{-1}(\mu) = (R_0^{RB}(\mu))^T (A_0^{RB}(\mu))^{-1} R_0^{RB}(\mu) + \sum_{i=1}^{M} R_i(\mu)^T A_i^{-1}(\mu) R_i(\mu), \tag{4.33}$$

where R^{RB} and A^{RB} are computed by the basis functions which is constructed in Section 3.3.2. The local matrices $A_i(\mu)$ correspond to the exact local solver for each μ as in the standard two-level additive Schwarz. We will show that these preconditioners are optimal with respect to the contrast, i.e. $cond(B^{-1}(\mu)A(\mu))$ is bounded independent of the contrast.We are currently implementing this.

5. APPLICATION OF UNCERTAINTY QUANTIFICATION TECHNIQUES FOR RICHARDS' EQUATION

5.1 Introduction

In large applications of soil moisture problems, one needs to solve inverse problems for Richards' equation. These problems are motivated by applications of soil moisture predictions. To predict future moisture content under the soil or water flow through the subsurface, we need to conjecture the hydraulic conductivity conditioned to some average dynamic data(e.g. average flux). There is uncertainty under this process. To make a better decision in water flow management, one requires a method for quantifying and reducing the uncertainty.

Since predicting water flow in porous media is nonlinear, it is generally difficult to calculate directly the posterior probability distribution of soil moisture. Instead, we estimate this probability distribution from the outcomes of flow predictions for a large number of realizations of the media. Therefore, it is essential that the conductivity realizations used in the flow simulation adequately reflect the uncertainty in the porous media.

The prediction of conductivity fields based on dynamic data is a challenging problem because conductivity fields are typically defined on a large number of grid blocks. Moreover, flow dynamic data is typically measured at different scales varying from point scale to remote sensing scale. Therefore, we need some judicious methods to link these two values.

In this chapter, we want to predict saturated hydraulic conductivity fields conditoined to some average dynamic data. In a probability context, this problem can be regarded as conditioning the saturated conductivity K_s, to the dynamic data F with measurement error. We would like to sample the saturated conductivity K_s from the conditional distribution $P(K_s|F)$. We used Markov chain Monte Carlo (MCMC) methods to perform this samplng process. MCMC methods permit a practitioner to

simulate a dependent sequence of random draws from very complicated stochastic models. The good thing about an MCMC algorithm is that it can generate a Markov chain $\{K_s\}$ whose stationary distribution is its target distribution $P(K_s|F)$.

In our sampling process, to compute a realization of the conductivity field, one needs a computation on a scale fine enough to determine the dynamic data fileds accurately. However, the need for fine-scale computations at each step of the MCMC algorithm makes the problems computationaly very expensive. Therefore, we need some coarsening process. In our research, we employed an upscaling method to coarsen the saturated hydraulic conductivity field using flow-based techniques. As for the nonlinear parameters of the conductivity term, we used simple volume averaging. Based upon this coarsening process, we proposed the two-stage MCMC algorithm which makes the entire computation much faster. Using this methodology, we can screen out bad proposals for the sampling by just doing the coarse-scale computations.

In many practical applications one wishes to express the saturated conductivity K_s as an expression of some parameters, rather than simply a function in physical space. In this dissertation, we discuss a particular type of expansion, known as the Karhunen-Loéve expansion, or KLE. Using the KLE, the high-dimensional conductivity field can be represented by a small number of parameters. Log-conductivity values which are used in sampling can be written as a linear comination of eigenvectors of the covariance matrix times the square root of the eigenvalues.

5.2 Fine-scale model

In this section we introduce the fine-scale models for Richards' equation. Recall that the formula describing Richards' equation under some assumptions is given by

$$D_t\theta(u) - \operatorname{div}(\kappa(x,u)D_x(u+x_3)) = 0 \quad \text{in}\,\Omega, \tag{5.1}$$

where $\theta(u)$ is volumetric water content and u is the pressure head. Constitutive relations between θ and u and between κ and u are developed appropriately, which consequently gives nonlinearity behavior in (5.1). See Section 2.3

There are several widely known formulations of the constitutive relations e.g. Haverkamp model, Exponential model (See (2.7), (2.9)) but in this paper, we use van Genuchten model, whose $\theta(u)$ and $\kappa(x, u)$ are described in (2.8)

5.2.1 Parametrization

In this section, we describe the model parametrization. It is known that model parameters vary depending on the soil properties. We consider heterogeneous soil environment where the soil consists of sand, silt, and clay and each component has heterogeneous spatial distribution. Depending on the fraction of sand, silt, and clay, the model parameters, $\theta_r(x)$, $\theta_s(x)$, $\alpha(x)$, $n(x)$, $K_s(x)$ (see (2.8)) will vary spatially. Next, we will describe the spatial parametrization. For simplicity of further discussions, we denote by Ψ, $\Psi(x) = \{\theta_r(x), \theta_s(x), \alpha(x), n(x), K_s(x)\}$,where θ_r, θ_s, α, and n are the parameter values for van Genuchten model and K_s is saturated conductivity.

For the distribution of each component sand, silt, and clay, we assume that each has a log-normal Gaussian distribution and independent. In our future work, we plan to investigate the correlation among these parameters. Let $Y_i(x) = \log[X_i(x)]$ where X_i is the distribution of each sand,silt, and clay, where Y_i ,$(i = 1, 2, 3)$, has a log-normal Gaussian distribution. Further, we denote by p_i, $(i = 1, 2, 3)$, the percentage of ith component of the soil and assume that it is described by two-point correlation function (that will be introduced in the next section). The percentages p_i can be represented by

$$p_i = \frac{\exp(Y_i)}{\exp(Y_1) + \exp(Y_2) + \exp(Y_3)} \times 100, \qquad i = 1, 2, 3. \tag{5.2}$$

Equipped with these percentages, ROSETTA database [66] can be used to get the material parameters θ_r, θ_s, α, and n, which are the parameter values for van Genuchten model, and K_s, saturated conductivity of the field. Material properties are computed using the ROSETTA software for soil hydraulic parameters [66]. The ROSETTA program is an artificial neural networks-based pedotransfer function that allows for the estimation of the van Genuchten soil hydraulic parameters using different levels of input data such as the soil textural class, soil textural composition, and/or more such details as bulk density, and water contents at specific pressure heads. In this study, we use ROSETTA to estimate the soil hydraulic parameters from the soil textural composition (percentages of sand, silt, and clay), denoted as the p_i 's above.

As for the forward solve, we use SWMS-3D [39] to solve Richards' equation using these material parameters. The SWMS-3D package is designed to simulate water and solute transport through the porous media (soils) by numerically solving the Richards' equations for water flow and the convection-dispersion equation (CDE) for solute transport. As the title signifies, the program handles the flow and transport phenomena in three dimensions. The flow equation is solved using a Galerkin-type linear finite element model. Further details about the structure and development of the program can be found in the work of Simunek et al [39].

As for boundary conditions, the top surface of the study domain is assigned an atmospheric boundary condition which allows for the flow of water into and out of the domain based on the environmental conditions (precipitation, evaporation, etc.). The bottom boundary of the domain was assigned a variable flux condition with a water table at 1m below the surface. The vertical sides of the domain were held impermeable, with no flow going into or out of the domain from these boundaries.

5.2.2 Karhunen-Loève expansion

Next, we describe the parametrization based on two-point statistics. We assume that for each component the spatial covariance description is prescribed by

$C^i_{\ell m} = E[Y_i(x_\ell)Y_i(x_m)]$, $(\ell, m = 1, 2, \ldots, N_f)$, where x_ℓ and x_m is the mesh points corresponding to the fine-scale model and E is the expectation operator. For simplicity, we assume that $E[Y_i] = 0$. We are interested in an orthonormal basis $\{\phi^i_k\}$ which makes $Y_i(x_\ell)$ and $Y_i(x_m), \ell \neq m$ uncorrelated. The basis functions $\{\phi^i_k\}$ are found by solving following eigenvalue problem,

$$C^i_{\ell m}\phi^i_k(x_\ell) = \lambda^i_k\phi^i_k(x_\ell), \qquad k = 1, 2, \ldots, N_f. \tag{5.3}$$

Using this basis, we can get the Karhunen-Loève expansion (KLE) of Y_i,

$$Y_i(x_\ell, \omega) = \sum_{k=1}^{N_f} \alpha^i_k(\omega)\phi^i_k(x_\ell), \qquad \ell = 1, 2, \ldots, N_f.$$

Denote $z^i_k = \alpha^i_k/\sqrt{\lambda^i_k}$, then z^i_k satisfy $E(z^i_k) = 0$ and $E(z^i_k z^i_l) = \delta_{kl}$. It follows that

$$Y_i(x_l, \omega) = \sum_{k=1}^{N_f} \sqrt{\lambda^i_k} z^i_k(\omega)\phi^i_k(x_l), \qquad \ell = 1, 2, \ldots, N_f. \tag{5.4}$$

Generally, we only need to keep the leading order terms (quantified by the magnitude of λ^i_k) and still capture most of the energy of the stochastic process Y_i. For an N-term KLE approximation $(N < N_f)$

$$Y^N_i = \sum_{k=1}^{N} \sqrt{\lambda^i_k} z^i_k \phi^i_k, \tag{5.5}$$

define the energy ratio of the approximation as

$$e_i(N) := \frac{E\|Y^N_i\|^2}{E\|Y_i\|^2} = \frac{\sum_{k=1}^{N} \lambda^i_k}{\sum_{k=1}^{\infty} \lambda^j_k}.$$

Since the eigenvalues can decay fast, the truncated KLE approximates the stochastic process Y_i fairly well in the L^2 sense. Therefore, we can sample Y_i from the truncated KLE (5.5) by generating the random variables z^i_k.

The advantages of the truncated KLE is that, first, we can use uncorrelated random variable to sample a realization and, second, it allows dimension reduction as we can remove small eigenvalues. Once Y_i^N's are sampled, we compute the percentages p_i based on (5.2) and then Ψ.

5.3 Bayesian approach

5.3.1 Markov chain Monte Carlo

In this section, we will briefly introduce Markov chain Monte Carlo algorithm. We will use Ψ to denote the parameters as defined earlier,

$$\Psi(x) = \{\theta_r(x), \theta_s(x), \alpha(x), n(x), K_s(x)\},$$

where θ_r, θ_s, α, and n are the parameter values for van Genuchten model and K_s is saturated conductivity. Note that Ψ is determined by Y_i's that represent the percentages of silt, sand, and clay. The problem under consideration consists of sampling the Ψ, with given measured data.

Measured data is taken as the soil moisture on the surface. While we use a synthetically generated case as the ground conditions for this study, data such as soil moisture are widely available, or easily measured, using ground/in-situ moisture sensors, or air-borne or satellite-based remote sensors. Measured soil moisture values are used as state variables and inputs for hydrologic modeling, and using that as the known data is thus appropriate. We denote the measured data by F. We would like to note that our method is not limited to this type of measurement and various measurements can be used in a straightforward way within the proposed techniques. Since the data are provided as an integrated response, the map from the data to the super-parameter value is not one-to-one. Hence this problem is ill-posed in the sense that there exist many different realizations for the given data.

From the probabilistic point of view, our problem can be regarded as conditioning Ψ to the measured data with associated measurement errors. Consequently, our goal is to sample from the conditional distribution $P(\Psi|F)$.

Using the Bayes formula we can write

$$P(\Psi|F) \propto P(F|\Psi)P(\Psi), \tag{5.6}$$

where $P(\Psi)$ is the prior distribution. In practice, the measured fluxes contain measurement errors. In this paper, we assume that the measurement error satisfies a Gaussian distribution; thus, the likelihood function $P(F|\Psi)$ takes the form

$$P(F|\Psi) \propto \exp\Big(-\frac{\|F_{obs} - F_\Psi\|^2}{\sigma_f^2}\Big), \tag{5.7}$$

where F_{obs} is observed data, F_Ψ is the response calculated using the parameter Ψ, and σ_f is the measurement precision.

Since both F_{obs} and F_Ψ are functions of time and space (denoted by x and t), the norm $\|F_{obs} - F_\Psi\|^2$ is defined as the L_2 norm

$$\|F_{obs} - F_\Psi\|^2 = \int_0^X \int_0^T \left[F_{obs}(x,t) - F_\Psi(x,t)\right]^2 dtdx.$$

Denote the sampling target distribution as

$$\pi(\Psi) = P(\Psi|F) \propto \exp\Big(-\frac{\|F_{obs} - F_\Psi\|^2}{\sigma_f^2}\Big) P(\Psi). \tag{5.8}$$

Since different spatial fields Ψ can give the same average flux response, the distribution $\pi(\Psi)$ is a function of Ψ with multiple local maxima. Sampling from the distribution $\pi(\Psi)$ can be accomplished by the Markov chain Monte Carlo (MCMC) method. For a given proposal distribution $q(y|x)$, the Metropolis-Hastings MCMC algorithm (see, e.g., [65]) consists of the following steps.

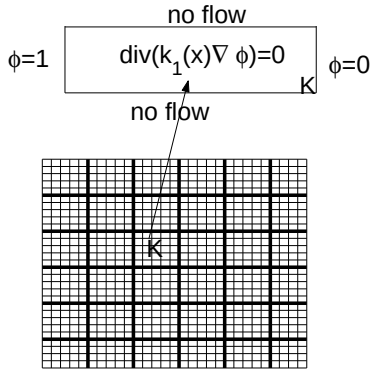

Fig. 5.1. Schematic description of the upscaling. Bold lines illustrate a coarse-scale partitioning, while thin lines show a fine-scale partitioning within coarse-grid cells.

Algorithm I (Metropolis-Hastings MCMC, Robert and Casella [65])

- Step 1. At Ψ^n generate Y from $q(Y|\Psi^n)$.
- Step 2. Accept Y as a sample with probability

$$p(\Psi^n, Y) = \min\left(1, \frac{q(\Psi^n|Y)\pi(Y)}{q(Y|\Psi^n)\pi(\Psi^n)}\right), \tag{5.9}$$

  i.e. take $\Psi^{n+1} = Y$ with probability $p(\Psi^n, Y)$, and $\Psi^{n+1} = \Psi^n$ with probability $1 - p(\Psi^n, Y)$.

The MCMC algorithm generates a Markov chain $\{\Psi^n\}$ whose stationary distribution is $\pi(\Psi)$.

5.3.2 A multiscale approach for surrogate modeling

The link between the coarse and the fine-scale conductivity fields is usually nontrivial because one needs to take into account the effects of all the scales present at the fine level. In this paper, we seek an inexpensive upscaling/multiscale procedure

that can screen out the proposals and avoid unnecessary fine-scale simulations. In our coarse-scale model, we will only upscale saturated hydraulic conductivity field using flow-based techniques [79]. As for nonlinear parameters, we will use simple volume averaging. This corresponds to a multiscale procedure where multiscale basis functions are used for pressure equation [17]. One can use nonlinear homogenization techniques for upscaling these parameters; however, this will make the upscaling CPU demanding. Later, we will discuss the effects of upscaling errors on our sampling.

First, we briefly describe flow-based upscaling techniques [79] which is one of the commonly used multiscale methods in several researches. In these techniques, coarse-grid conductivities are computed rather than multiscale basis functions are constructed on a coarse grid. The flow equation is solved on the coarse grid with these computed coarse-grid conductivities. Here is the brief description.

Consider the fine-scale hydraulic conductivity that is defined in the domain with underlying fine grid as shown in Figure 5.1. On the same graph we illustrate a coarse-scale partition of the domain. To calculate the coarse-scale saturated hydraulic conductivity field at this level we need to determine it for each coarse block, K. The coarse block conductivity can be defined both using the solutions of local or global problems. The main idea of the calculation of a coarse-scale conductivity is that it delivers the same average response as that of the underlying fine-scale problem locally. The calculation of the coarse-scale conductivity based on local solutions is schematically depicted in Figure 5.1. For each coarse domain K we solve the local problems

$$div(K_s(x)\nabla\phi_j) = 0, \tag{5.10}$$

with some coarse-scale boundary conditions. One of such boundary conditions is given by $\phi_j = 1$ and $\phi_j = 0$ on the opposite sides along the direction e_j and no flow

boundary conditions on all other sides. For these boundary conditions the coarse-scale conductivity is given by

$$(K_s^*(x)e_j, e_l) = \frac{1}{|K|}\int_K (K_s(x)\nabla\phi_j(x), e_l)dx, \tag{5.11}$$

where ϕ_j is the solution of (5.10) with prescribed boundary conditions. Various boundary conditions can have some influence on the accuracy of the calculations, including periodic, Dirichlet and etc. These issues have been discussed in [79]. In particular, for determining the coarse-scale conductivity field one can choose the local domains that are larger than target coarse block, K, for (5.10). Further (5.11) is used in the domain K, where ϕ_m are computed in the larger domains with correct scaling (see [79]). This way one reduces the effects of the artificial boundary conditions imposed on K. These issues have been also discussed, for example, in [79].

As for parameters representing nonlinearities, we use volume averaging. More precisely, we define

$$\theta_r^* = \frac{1}{|K|}\int_K \theta_r.$$

Similarly, we average the other quantities θ_s, α, n and get the volume averaging value $\theta_s^*, \alpha^*, n^*$. Furthermore, these coarse-grid quantities are used to solve Richards' equation on a coarse grid. This computation is inexpensive as it involves fewer coarse domains. For example, elapsed time to solve Richards' equation on the coarse grid for one sample is 6.7 seconds. On the other hand, it takes 115.6 seconds to solve it on the fine grid. Next, we briefly demonstrate some numerical results that compare coarse- and fine-grid simulations results. In Figures 5.2 and 5.3, we depict the soil moisture response on the top surface (measured data) for fine and coarse models. The figures on the left represent the soil moisture distribution for the fine-scale model. The figures in the middle represent the average fine-scale model. The figures on the right represent the response from coarse-grid models. As we see that our coarse model captures the main features of the fine-scale models.

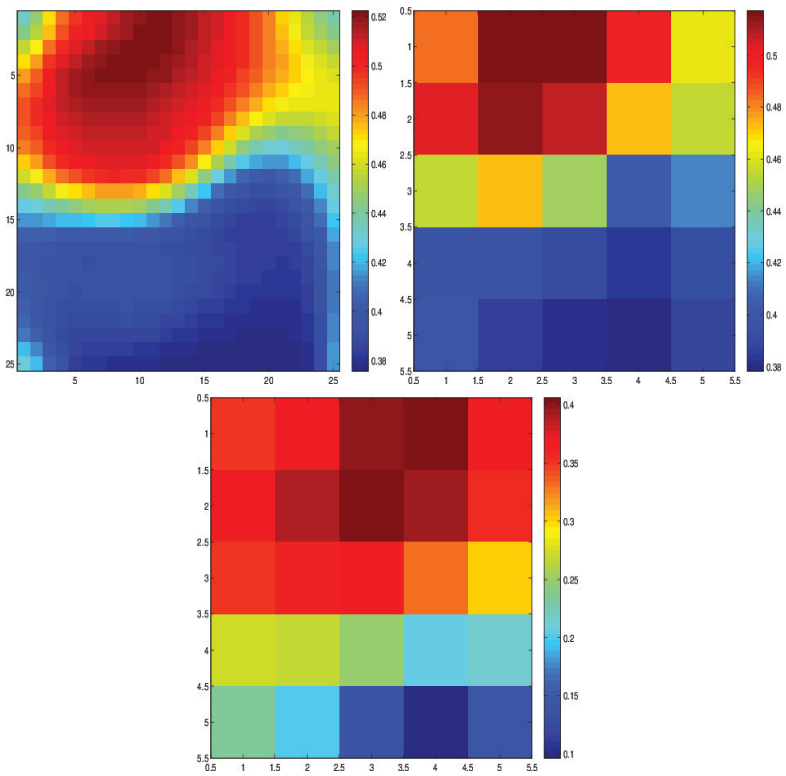

Fig. 5.2. (Left): Soil moisture on the fine grid. (Middle): Fine-scale soil moisture averaged on a coarse grid. (Right): Coarse-scale soil moisture data obtained using upscaled model.

Note that the upscaled model will be used as a preconditioner in the sampling process. For the efficiency of the algorithm, we need to be able to predict the results of the fine-scale simulations based on the results of upscaled models. One can use a different upscaling strategy, e.g., average Y_i and X_i. We did not compare various upscaled methods in this paper and would like to use flow-based upscaling technique to speed-up the sampling procedure.

u

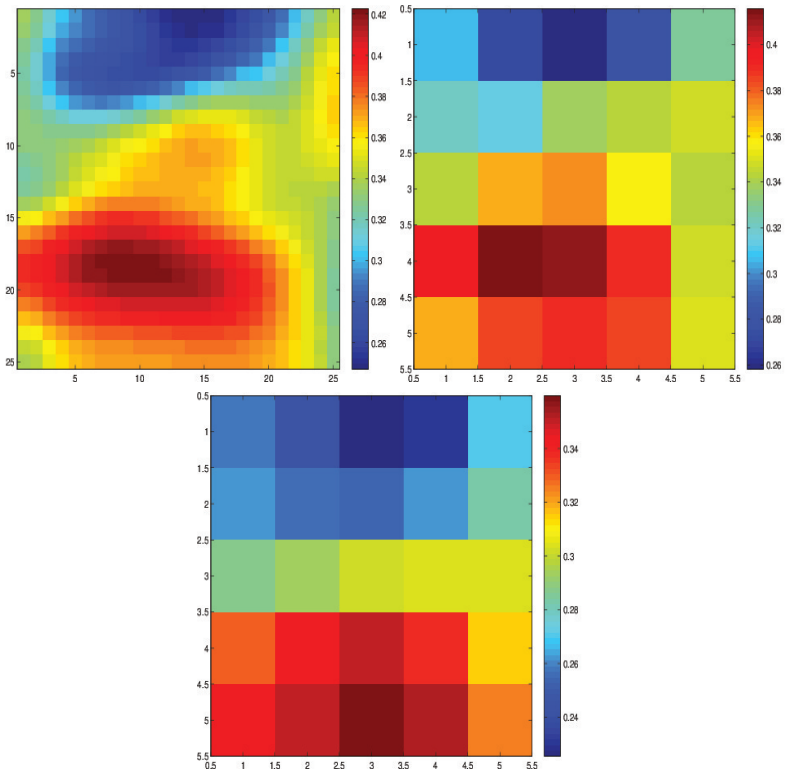

Fig. 5.3. (Left): Soil moisture on the fine grid. (Middle): Fine-scale soil moisture averaged on a coarse grid. (Right): Coarse-scale soil moisture data obtained using upscaled model.

5.3.3 Two-stage algorithm

To introduce two-stage algorithm, we consider the posterior probability distribution that corresponds to the coarse-grid model. It is given by

$$\pi^*(\Psi) \propto \exp\left(\frac{-\int_0^{T_{final}}(F^*_\Psi(t) - F_{obs}(t))^2 dt}{\sigma_c^2}\right) P(\Psi), \tag{5.12}$$

where $F^*_\Psi(t)$ is an average flux value calculated with the parameter Ψ on the coarse grid, F_{obs} is the measured flux data as defined earlier, and σ_f is modified error measurement for the coarse-grid response. This quantity can be computed based

on correlation between fine-scale and coarse-scale results performed based on a few simulations. Also consider the posterior probability distribution that corresponds to the fine-grid model given by

$$\pi(\Psi) \propto \exp\left(\frac{-\int_0^{T_{final}}(F_\Psi(t) - F_{obs}(t))^2 dt}{\sigma_f^2}\right) P(\Psi),$$

where $F_\Psi(t)$ is an average flux value calculated with the parameter Ψ on the fine grid. For a given proposal distribution $q(y|x)$, the two stage Metropolis-Hasting MCMC method can be defined as following.

Algorithm II (Two stage Metropolis-Hasting MCMC)

- Step1. Given parameter Ψ^n generate Y from $q(Y|\Psi^n)$.
- Step2. Accept Y with probability

$$Pr_1 = min\left(1, \frac{q(\Psi^n|Y)\pi^*(Y)}{q(Y|\Psi^n)\pi^*(\Psi^n)}\right),$$

- Step3. If Y is accepted at the step2, accept Y as a sample with probability

$$Pr_2 = min\left(1, \frac{\pi(Y)\pi^*(\Psi^n)}{\pi(\Psi^n)\pi^*(Y)}\right),$$

The advantage of the two-stage MCMC is that we can screen out bad proposals based on the upscaled models that are cheap to implement.

5.4 Numerical Experiments

5.4.1 Numerical implementation

In this section we discuss the numerical implementation for the two stage MCMC algorithm.

The domain we take for the implementation is the box with the dimensions $100m \times 100m \times 3.31m$. We discretize the domain with a rectangular mesh. We take a 15×15 and 25×25 fine scale model, and a 5×5 coarse scale model. We assume that the heterogeneities are only in the horizontal direction aerially, and not vertical (layering). We thus perform only areal coarsening. However, for the solution of the finite element mesh, 44 layers are introduced in the domain, with finer discretization near the top and bottom boundaries, and coarser in the center of the domain.

In our implementation,we assume that the distribution of each component, sand, silt, and clay, has a log-normal Gaussian distribution., which means that $Y_i(x, \omega)$ is a Gaussian process with the covariance kernel. We take a Gaussian covariance kernel for the KLE for each material (sand, silt, and clay),

$$C^i(x, y) = \sigma^2 \exp\Big(-\frac{|x_1 - y_1|^2}{2L_{1i}^2} - \frac{|x_2 - y_2|^2}{2L_{2i}^2}\Big), \tag{5.13}$$

with correlation lengths $L_{11} = 0.3, L_{21} = 0.5, L_{12} = 0.6, L_{22} = 0.3, L_{13} = 0.3$, and $L_{23} = 0.3$ respectively for sand, silt, and clay, and a constant $\sigma^2 = E(Y_i^2) = 2$. Note that one can choose different correlation lengths and also assume the correlation lengths to be random within the proposed framework.

We take z_k^i (see (5.4)) to be independent standard Gaussian random variables. Now given initial value z^0, we get the next proposal

$$z^{n+1} = z^n + \delta N(0, 1), \tag{5.14}$$

where δ is a constant and $N(0, 1)$ is a standard normal distribution. We keep 25 terms for the truncated KLE for each constituent. We impose hard constraints on the distributions of sand, silt, and clay and assume that their values are given at some fixed locations. This is done by solving a linear system resulting in a linear subspace of our parameter space (a hyperplane). For example, if we have truncated

KLE which consists of 25 terms, the expansion for Y_i can be represented by following matrix multiplication.

$$\Phi^i z^n = [\phi_1^i(x_\ell), \phi_2^i(x_\ell), \ldots, \phi_{25}^i(x_\ell)]z^n,$$

where ϕ_j^i, $(j = 1, \ldots 25)$ are eigenfunctions from the eigenvalue problem (5.3), x_ℓ, $(\ell = 1, \ldots N_f)$ are grid points on the fine-scale and z^n is a random vector proposal defined on (5.14). However, if Y_i values are known at 6 locations, one needs to choose 19 of elements of z^n randomly and solve the 6 equations with 6 unknowns, i.e. we need to solve $(\Phi^i)'x = b^i$ where $(\Phi^i)'$ is a truncated matrix from the matrix Φ^i and b^i is the known Y_i values at the specific locations. In our simulations, we search 6×6 matrices in 6×25 matrix with the lowest condition number. This can be easily done using QR decomposition. Now using these expansion parameters, we get KL expansion of Y_i , $i = 1, 2, 3$ for each material sand, silt, and clay.

After obtaining the KLE for each material, we compute the percentage of each and get the material parameters using ROSETTA program equipped with these percentages. We run SWMS-3D to solve Richards' equation on the coarse grid. We have the volume averaging material parameter values and upscaled conductivity field for the input of SWMS-3D. We get nodal values of the water content for a output of SWMS-3D. Now let us define

$$G^*(z^n) = \int_0^T (F_{z^n}^*(t) - F_{obs}(t))^2 dt.$$

where $F_{z^n}^*$ is an average flux value calculated on the coarse grid with the proposal z^n. We accept the trial z^{n+1} with probability

$$Pr = \exp(-\frac{G^*(z^{n+1}) - G^*(z^n)}{\sigma_c^2}).$$

Now if this sample is accepted, we run the SWMS-3D on the fine grid and get the water content value for the fine grid vertices. Then we compute

$$G(z^n) = \int_0^T (F_{z^n}(t) - F_{obs}(t))^2 dt.$$

where F_{z^n} is an average flux value calculated on the fine grid with the proposal z^n. We accept this sample with probability

$$Pr = exp\left((-\frac{G(z^{n+1})}{\sigma_f^2}) - (-\frac{G(z^n)}{\sigma_f^2}) + (-\frac{G^*(z^{n+1})}{\sigma_c^2}) - (-\frac{G^*(z^n)}{\sigma_c^2})\right).$$

In our implementation, we take $\sigma_c = \sigma_f = \sigma$ though one can choose a different σ_c depending on some initial runs. We test 10,000 samples in the MCMC for every simulation.

5.4.2 Numerical results

We present numerical results for four cases:

- 15×15: $\sigma = 0.005$ and $\delta = 0.1$ (Figure 5.4)
- 25×25: $\sigma = 0.01$ and $\delta = 0.1$ (Figure 5.5, 5.6, 5.7, 5.8)
- 25×25: $\sigma = 0.02$ and $\delta = 0.1$ (Figure 5.10, 5.11, 5.12, 5.13)
- 25×25: $\sigma = 0.02$ and $\delta = 0.2$ (Figure 5.15, 5.16, 5.17, 5.18)

Here 15×15 means that we used 15×15 discretized domain as our fine-scale model. In every case, we use 5×5 disctretized domain as our coarse-domain model. σ is a measurement precision and δ is a variance of the proposal distribution.

The acceptance rate for each case is presented as: which can be calculated by the ratio of the number of the accepted trials at the first stage to the number of the

Table 5.1
Accepted rate

Mesh size	σ	δ	acceptance rate
15×15	0.01	0.1	0.1010
25×25	0.01	0.1	0.0317
25×25	0.02	0.1	0.0156
25×25	0.02	0.2	0.0038

accepted trials at the second stage. Note that using single-stage MCMC based on fine-scale runs, these computations are difficult to perform due to long CPU times.

In the Table 5.1, we can observe that the acceptance ratio is higher when the fine grid is taken to be 15×15 instead of 25×25. This is because the coarse-scale model gives more accurate results for 15×15 as we use the less coarsening. The table also shows that when we use smaller δ, the acceptance rate is higher. This is because at each proposal step, the new proposal value is closer to the previous proposal value when you use the smaller δ. We also observe that larger σ gives less acceptance rate. This is because using larger measurement precision means we need more accurate sampled data to the real observed data. So our proposal needs to be closer to the reference response that means we accept only accurate samples. This is the reason why the acceptance rate is smaller when we use the larger precision value σ.

For each case we present the plots for the average errors $G(z^n)$ and some saturated conductivity realizations. Figures 5.4, 5.5, 5.10 and 5.15 show that the preceding algorithm gives us convergences to steady states. We can observe that all the average errors with different σ and δ show the convergence. Moreover, we can see that the the results with smaller δ value show more rapid convergence. This is because smaller δ makes the next trial be close from the previous trial and it makes the difference between the global errors of previous and next steps smaller. For example, if we get a good sample from some trial, then next trial would be chosen closely from the good sample if we use a small δ value. When we compare Figures 5.5 and 5.10,

we see that larger σ value give less fluctuational error movement. We notice that larger σ gives smaller accepted probability which results in more accurate samples for each iterative step. Therefore, the error between the previous trial and next trial is smaller with larger σ value.

In Figures 5.6, 5.11, 5.16, the reference and three accepted distributions of the saturated conductivities are presented. Also, in Figures 5.7, 5.8, 5.9, 5.12, 5.13, 5.14 and 5.17, 5.18, 5.19, the reference configuration is shown in the upper left, and three accepted distributions from the preceding algorithm are presented in the rest of plots for sand, silt, and clay, respectively. We can observe good sampling results especially with smaller σ and δ values. Let's compare the cases, Figures 5.7-5.9 and Figure 5.12-5.14. As we mentioned above, larger σ value allow us to accept only accurate samples. Therefore, the case, Figures 5.7-5.9, only accepts 130 samples out of 10,000 trials but the case, Figures 5.12-5.14, accepts 160 samples. This is why smaller σ gives a bit better sampling results out of limited number of trials. Also we observe from Figures 5.12-5.14 and 5.17-5.19 that larger δ gives worse sampling results. This is because larger δ makes the next trial be far way from the previous trial and it makes the error big which results in poor sampling.

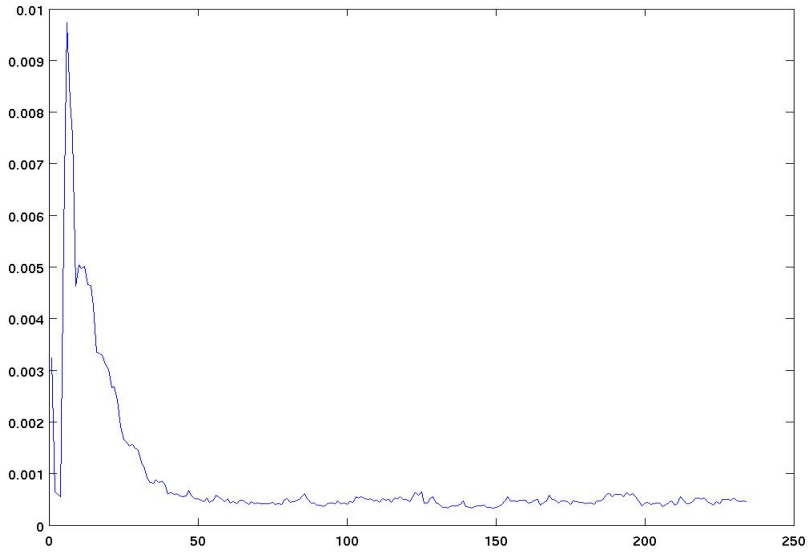

Fig. 5.4. Average errors for the case of 15×15: $\sigma = 0.005$ and $\delta = 0.1$

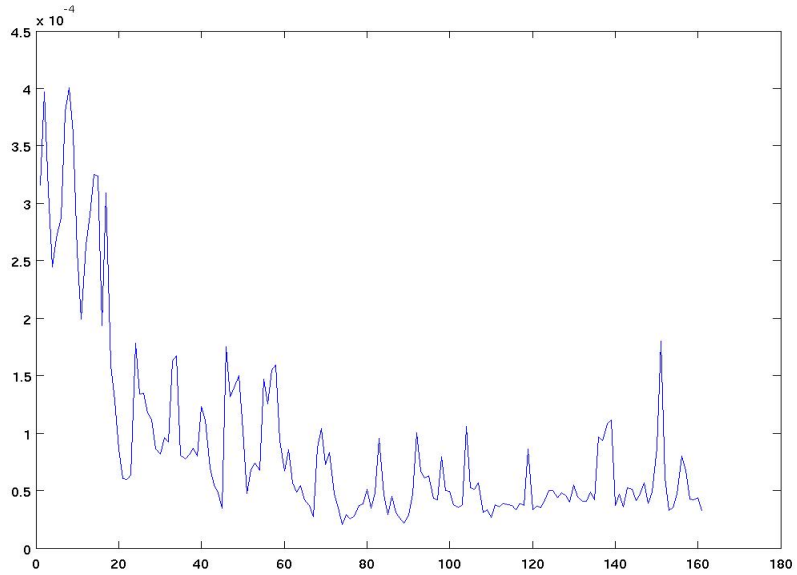

Fig. 5.5. Average errors for the case of 25×25: $\sigma = 0.01$ and $\delta = 0.1$

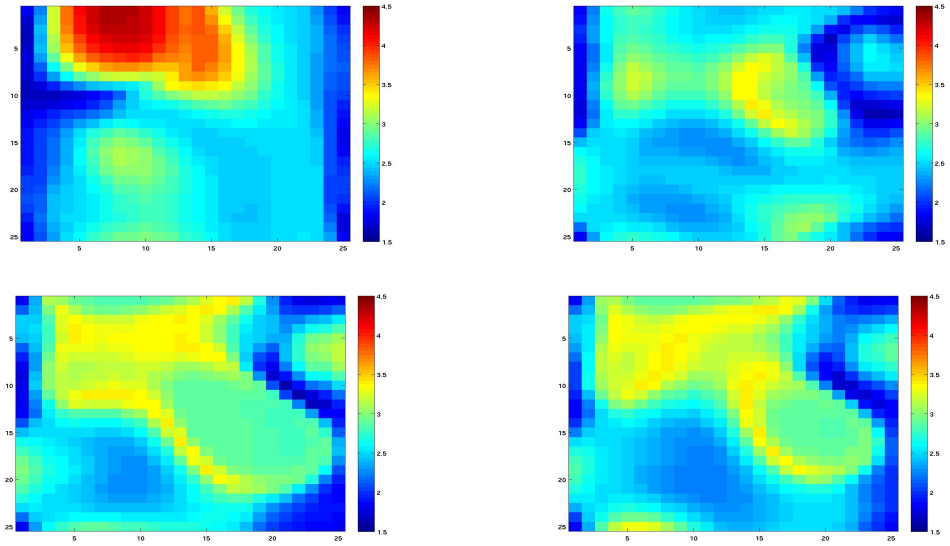

Fig. 5.6. Reference (upper left) and three different accepted k_s distribution for the case of 25×25: $\sigma = 0.01$ and $\delta = 0.1$

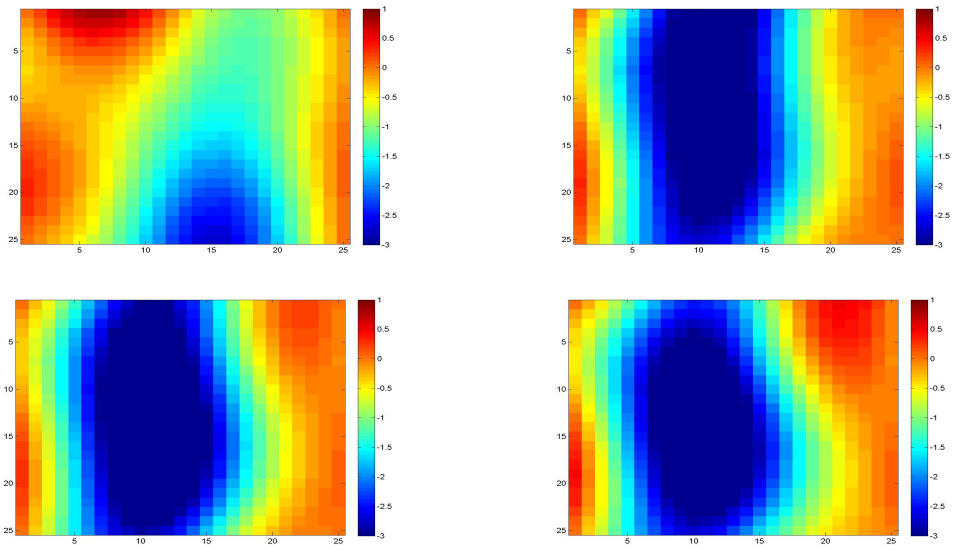

Fig. 5.7. Reference (upper left) and three different accepted sand distribution for the case of 25×25: $\sigma = 0.01$ and $\delta = 0.1$

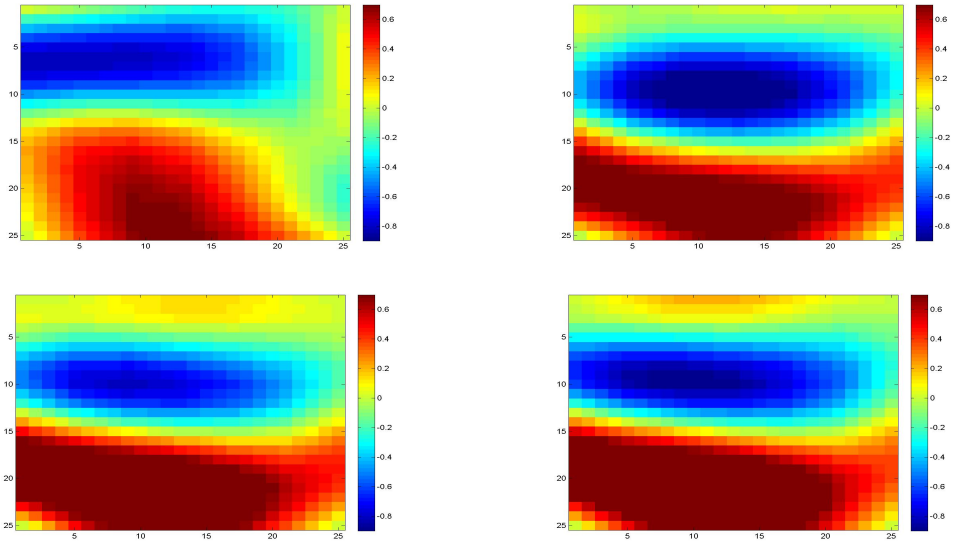

Fig. 5.8. Reference (upper left) and three different accepted silt distribution for the case of 25×25: $\sigma = 0.01$ and $\delta = 0.1$

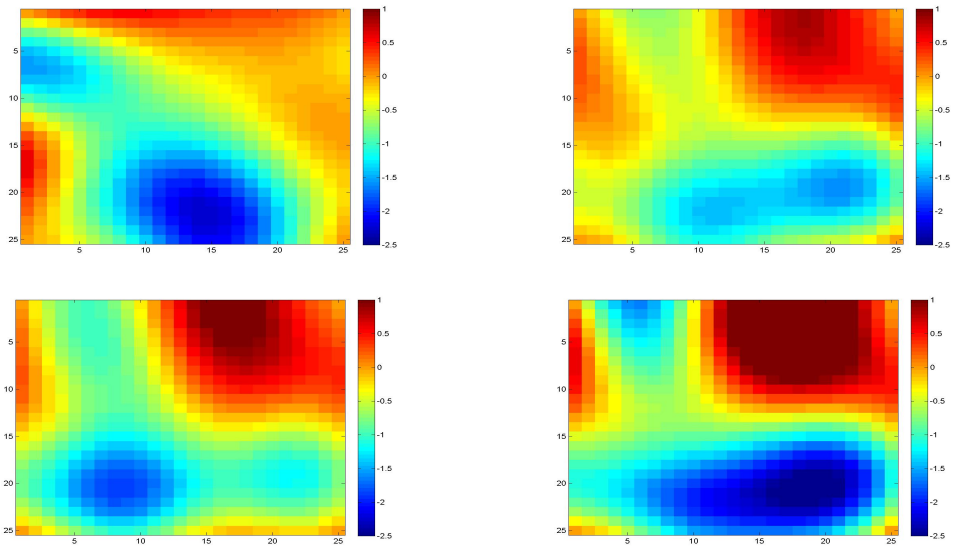

Fig. 5.9. Reference (upper left) and three different accepted clay distribution for the case of 25×25: $\sigma = 0.01$ and $\delta = 0.1$

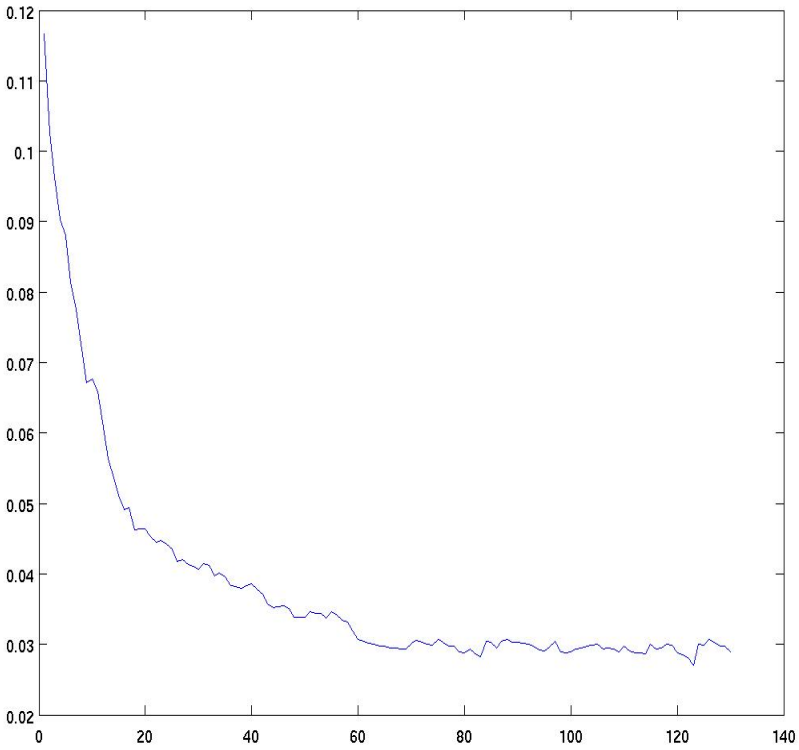

Fig. 5.10. Average errors for the case of 25×25: $\sigma = 0.02$ and $\delta = 0.1$

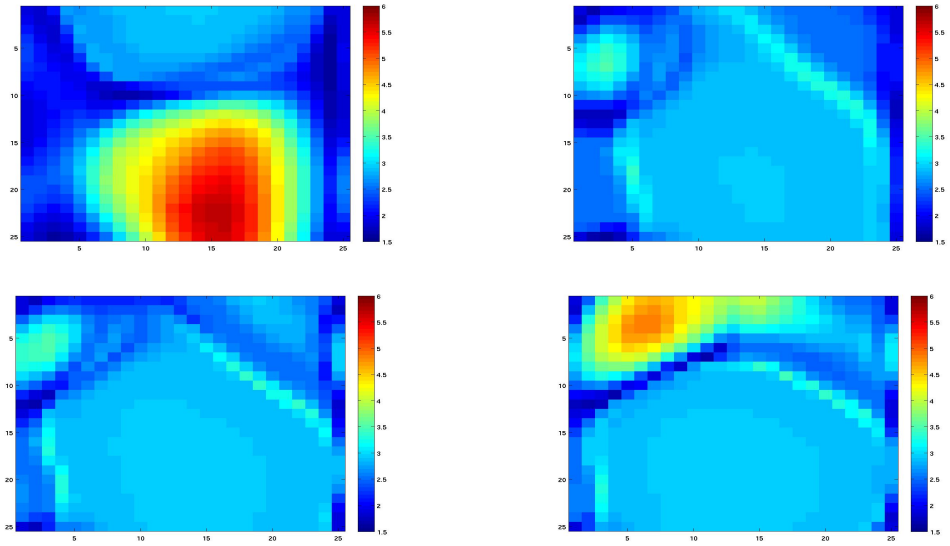

Fig. 5.11. Reference (upper left) and three different accepted k_s distribution for the case of 25×25: $\sigma = 0.02$ and $\delta = 0.1$

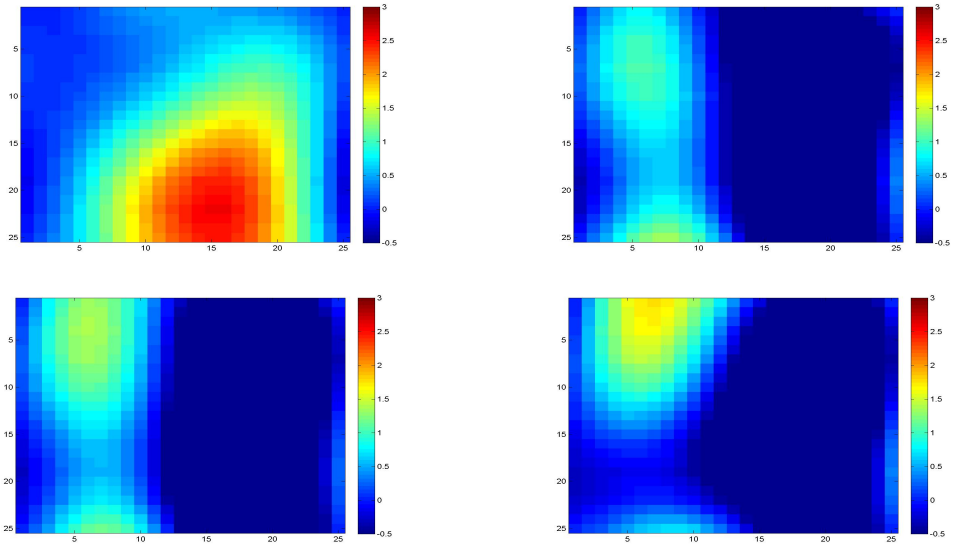

Fig. 5.12. Reference (upper left) and three different accepted sand distribution for the case of 25×25: $\sigma = 0.02$ and $\delta = 0.1$

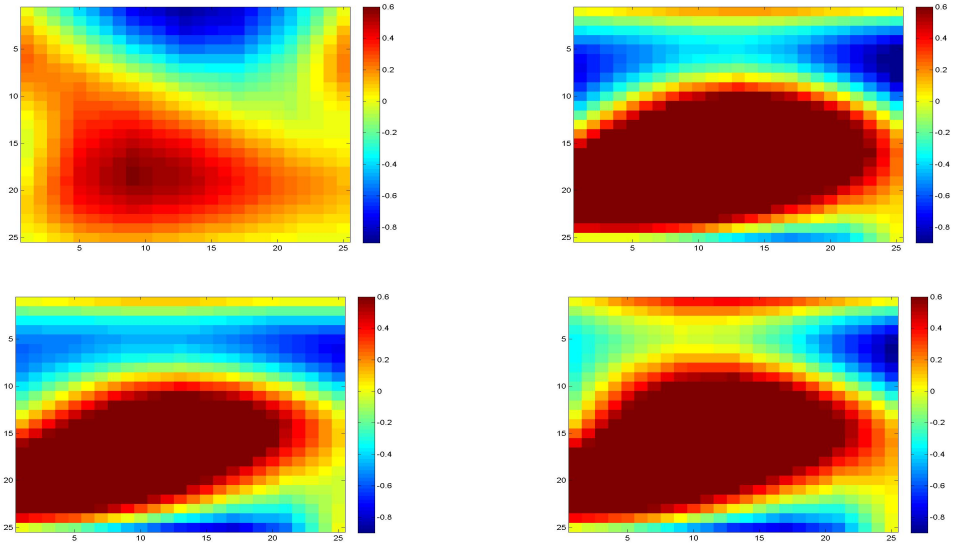

Fig. 5.13. Reference (upper left) and three different accepted silt distribution for the case of 25×25: $\sigma = 0.02$ and $\delta = 0.1$

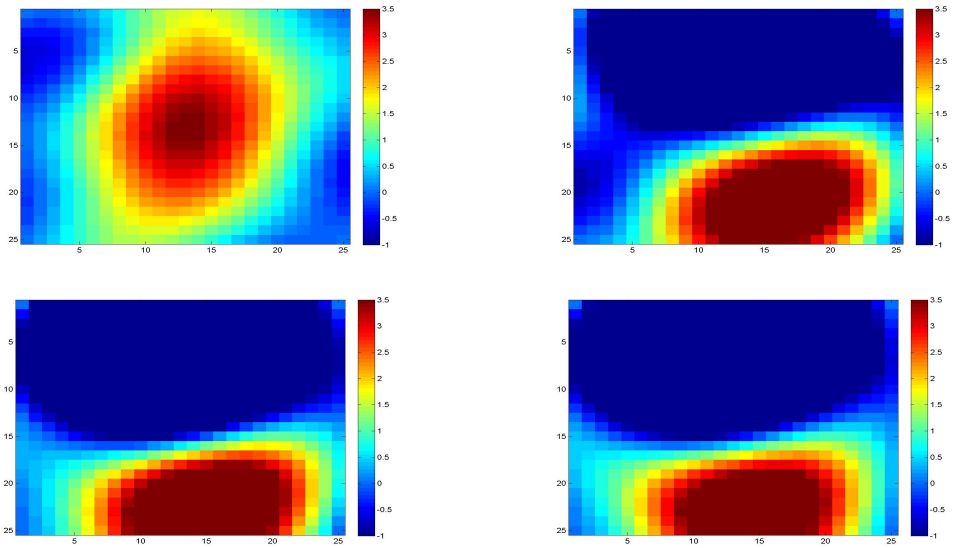

Fig. 5.14. Reference (upper left) and three different accepted clay distribution for the case of 25×25: $\sigma = 0.02$ and $\delta = 0.1$

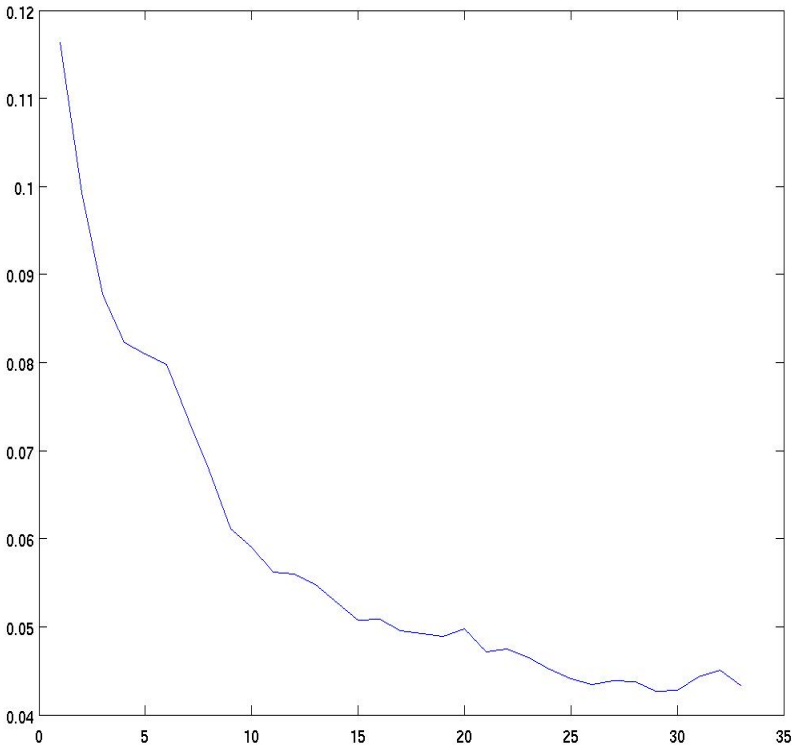

Fig. 5.15. Average errors for the case of 25×25: $\sigma = 0.02$ and $\delta = 0.2$

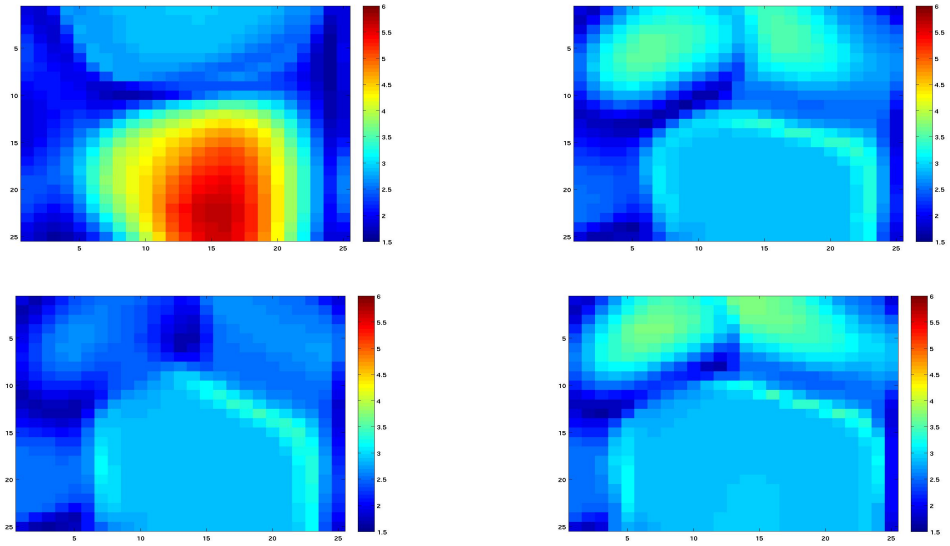

Fig. 5.16. Reference (upper left) and three different accepted k_s distribution for the case of 25×25: $\sigma = 0.02$ and $\delta = 0.2$

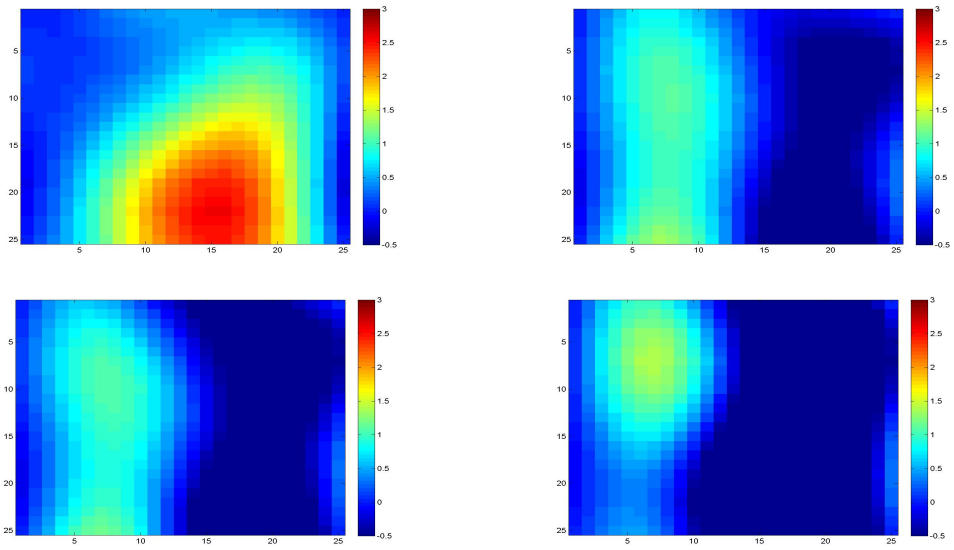

Fig. 5.17. Reference (upper left) and three different accepted sand distribution for the case of 25×25: $\sigma = 0.02$ and $\delta = 0.2$

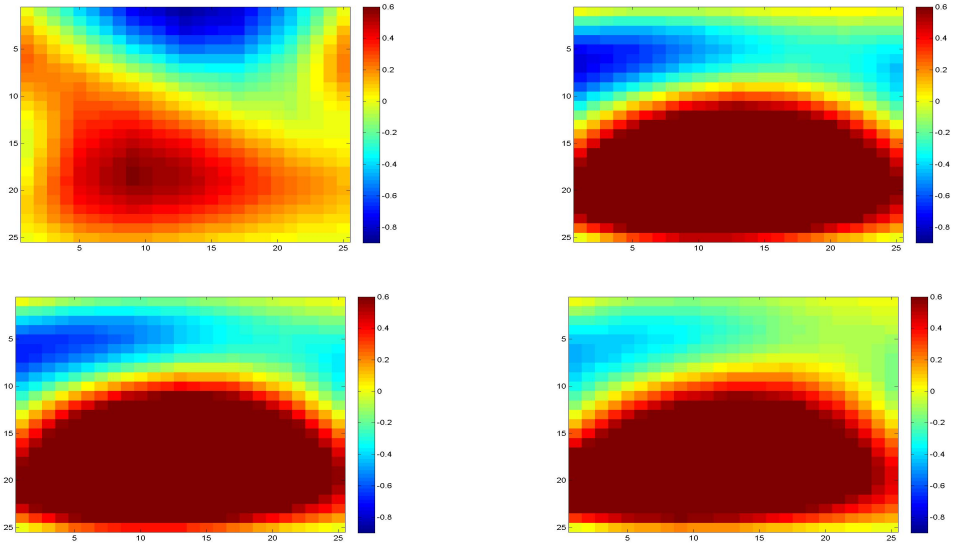

Fig. 5.18. Reference (upper left) and three different accepted silt distribution for the case of 25×25: $\sigma = 0.02$ and $\delta = 0.2$

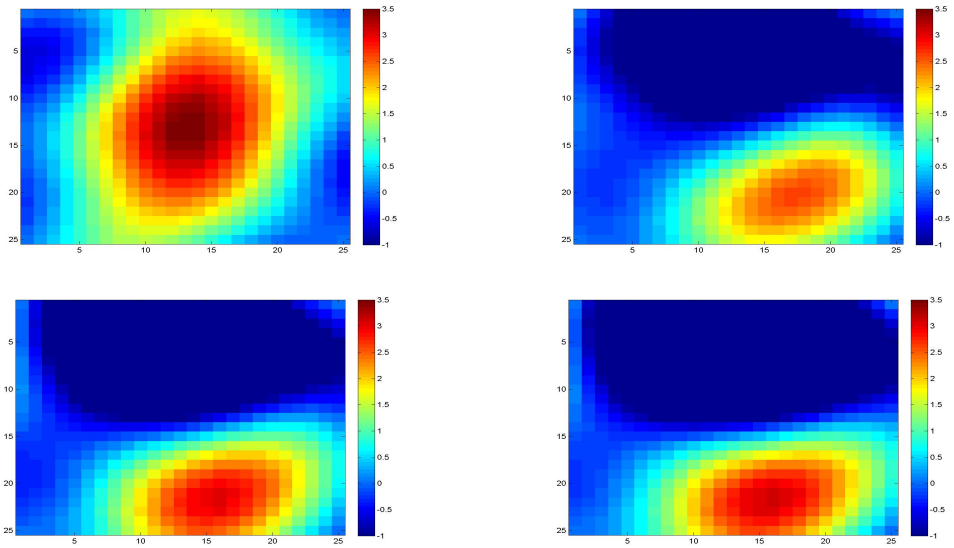

Fig. 5.19. Reference (upper left) and three different accepted clay distribution for the case of 25×25: $\sigma = 0.02$ and $\delta = 0.2$

6. CONCLUSION

6.1 Summary

This dissertation concentrates on designing multiscale finite element methods and uncertainty quantification techniques for Richards' equation especially in the highly heterogeneous porous media. In Chapter II, we presented Richards' equation, a mathematical model that describes fluid flow in unsaturated porous media. There are several formulations of the Richards' equation, such as the head-form, θ-form, and the coupled-form and through this dissertation we used coupled-form of Richards' equation due to its charateristics of mass conservancy. Also there are several constitutive relations for the moisture content and hydraulic conductivity which are the main components to model Richards' equation. Haverkamp model, exponential model, and van Genuchten model are the most widely used empiricaly construced models that describes the functional relation of relative hydraulic conductivity to the pressure head. For our numerical simulations in this dissertation, we used these three models for Richards' equation. In this chapter, we also described the finite element based numerical discretization of Richards' equation. We have shown that under suitable conditions, there exists a unique solution for this discretized equation.

In Chapter III, MsFEM for Richards' equation has been presented. MsFEM consists of two parts; constructing basis functions that capture the multiscale structure of the solution and developing a global formulation which couples the basis functions to provide a reasonable approximation of the solution. Many coarse spaces have been introduced and constructed to capture the fine-scale features and give an accurate coarse-scale solution. We introduced three differerent coarse spaces for MsFEM, linear boundary conditions coarse space, coarse space from oversampling techniques and energy minimizing coarse space. Further, we presented coarse spaces with local spectral information which gives faithful coarse-scale approximaion by complementing the multiscale coarse space. We applied MsFEM to Richards' equation with

separable coefficient and displayed some numerical examples using introduced coarse spaces. We observed from the numerical results that using local spectral spaces for MsFEM, we can get an accurate coarse-scale solution for Richards' equation.

We used MsFEM for parameter-dependent problems for numerical simulation of Richards' equation. In this simulation, nonseparable Richards' equation has been considered and we regarded it as a parameter-dependent elliptic equation. One of the main difficulties to solve this equation using MsFEM is expensive computational time to construct the basis functions. To overcome this problem, we employed techniques from RB which gives an efficient way to solve local spectral problems for each parameter value by constructing the solution of the eigenvalue problem using eigenvectors for parameter values computed in the offline stage.

In Chapter IV, we study robust iterative solvers for finite element discretizations of steady-state Richards' equation. We assume that the nonlinear conductivity field can be written as a product of a nonlinear function and a heterogeneous spatial function that has high contrast. Due to spatial heterogeneities, the number of iterations in an iterative method, in general, will depend on the contrast. To alleviate this problem, we design and investigate iterative solvers that converge independent of the physical parameters (small spatial scales and large contrast). The proposed iterative solvers consist of outer and inner iterations, as it is commonly done in the literature. Outer iterations, designed to handle nonlinearities, linearize the equation around the previous solution state. We show that this linearization yields contrast independent iterative procedure. For inner iterations, we constructed domain decomposition preconditioners to solve this problem. Especially we used the local spectral space for our preconditioner and show that it makes the inner iterations of our iterative solver be independent of the contrast. Since the same preconditioner was used for every outer iteration, this makes the overall solution process quite efficient. Numerical results are presented to confirm the theoretical findings.

Lastly, uncertainty quantification techniques for Richards' equation is introduced in Chapter V. We are interested in sampling hydraulic conductivities conditioned on given flux data. We performed MCMC algorithm to sample the conductivities. For each step of MCMC, we need to solve Richards' equation with given realization of the conductivity field, which results in huge computational time especially on a fine grid. Therefore, we need some coarsening process. We performed an upscaling method, one of the commonly used multiscale methods, to coarsen the conductivity field and suggeted a two-stage MCMC algorithm. Compared to one-stage MCMC, two-stage MCMC can screen out all the bad proposals by just computing solution on a coarse-grid and gives much faster sampling process. Several sampling results with different parameter values are presented and showed that our sampling technique is efficient.

REFERENCES

[1] T. Arbogast, M.F. Wheeler, and N.Y. Zhang, *A nonlinear mixed finite element method for a degenerate parabolic equationarising in flow in porous media,* SIAM J. Numer. Anal., 33 (1996), pp. 1669-1687.

[2] A. Abdulle, and G. Vilmart, *A priori error estimates for finite element methods with numerical quadrature for nonmonotone non linear elliptic problems*, Numerische Mathematik, 121 (2012) , pp. 397-431. Submitted for publication.

[3] M. Bause, and P. Knabner, *Computation of variably saturated subsurface flow by adaptive mixed hybrid finite element methods* Advances in Water Resources, 27 (2004), pp. 565-581.

[4] S.C. Brenner, and L.R. Scott, *The Mathemethical Theory of Finite Element Methods*, Springer-Verlag, Berlin,1994.

[5] R.H. Brooks, and A.T. Corey, *Properties of porous media affecting fluid flow,* Journal of the Irrigation and Drainage Division, Vol. 92(1966), pp. 61-90.

[6] E. Buckingham E. *Studies on the movement of soil moisture*, U.S. Dept. of Agriculture Bureau of Soils Bulletin, 38 (1907), pp. 28-61.

[7] X.C. Cai, and D.E. Keyes, *Nonlinearly preconditioned inexact Newton algorithms*, SIAM J. Sci. Comput., 24 (2002), pp. 183-200.

[8] X.C. Cai, L. Marcinkowski, and P. Vassilevski, *An element agglomeration nonlinear additive Schwarz preconditioned Newton method for unstructured finite element problems*, Appl. Math., 50 (2005), pp. 247–275.

[9] V. Calo, Y. Efendiev, and J. Galvis, *A note on vatiational multiscale methods for high-contrast heterogeneous flows with rough source terms*, Advances in Water Resources, 34 (2011), pp. 1177-1185.

[10] M.A. Celia, E.T. Bouloutas, and R.L. Zarba, *A general mass-conservative numerical solution for the unsaturated flow equation*, Water Resources Research, 26 (1990), pp. 1483-1496.

[11] G. Chavent. and J. Jaffré, h *Mathematical Models and Finite Elements for Reservoir Simulation; Single Phase, Multiphase and Multicomponent Flows through Porous Media*, Elsevier Science Publishers, New York, 1986.

[12] Z. Chen, W.B. Deng, and H. Ye, *Upscaling of a class of nonlinear parabolic equations for the flow transport in heterogeneous porous media*, Communications in Mathematical Sciences 3 (2005), pp. 493-515.

[13] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, SIAM, Philadelphia, 2002.

[14] M. Dryja, and W. Hackbusch, *On the nonlinear domain decomposition method*, BIT Numerical Mathematics, Volume 37 (1997), pp. 296-311.

[15] J. Dugundgi, *Topology*, Allyn and Bacon, Boston, 1966.

[16] L.J. Durlofsky, *Coarse scale models of two-phase flow in heterogeneous reservoirs: Volume averaged equations and their relation to existing upscaling techniques*, Comp. Geosciences 2 (1998), pp. 73-92.

[17] Y .Efendiev, and T. Hou, *Multiscale Finite Element Methods, Theory and Applications*, Springer, New York, 2009.

[18] Y. Efendiev, T. Hou, and V. Ginting, *Mutlscale finite element methods for nonlinear problems and their applications*, Comm. Math. Sci. 2 (2004), pp. 553-589.

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

[20] Y. Efendiev, V. Ginting, T. Hou, and R. Ewing, *Accurate multiscale finite element methods for two-phase flow simulations*, J. Comp. Physics, 220 (2006), pp. 155-174.

[21] Y. Efendiev, J. Galvis, and X. H. Wu, *Multiscale finite element methods for high-contrast problems using local spectral basis functions*, Journal of Computational Physics, 230 (2011), pp. 937-955.

[22] Y. Efendiev, J. Galvis, and P. Vassielvski, *Spectral Element Agglomerate Algebraic Multigrid Methods for Elliptic Problems with High-Contrast Coefficients*, Domain Decomposition Methods in Science and Engineering XIX, 78 (2011), pp. 407-414.

[23] Y. Efendiev, and J. Galvis, *A domain decomposition preconditioner for multiscale high-contrast problems*, in Domain Decomposition Methods in Science and Engineering XIX 78 (2011), pp. 189-196.

[24] Y. Efendiev, J. Galvis and F. Thomines, *A systematic coarse-scale model reduction technique for parameter-dependent flows in highly heterogeneous media and its applications*, accepted, 2012.

[25] M.W. Farthing, C.E. Kees, and C.T. Miller, *Mixed finite element methods and higher order temporal approximations for variably saturated groundwater flow*, Advances in Water Resources, 26 (2003), pp. 373-394.

[26] J. Galvis and Y. Efendiev, *Domain decomposition preconditioners for multiscale flows in high contrast media*, SIAM MMS, 8 (2010), pp. 1461-1483.

[27] J. Galvis, and Y. Efendiev, *Domain decomposition preconditioners for multiscale flows in high-contrast media: Reduced dimension coarse spaces*, SIAM MMS, Volume 8 (2010), pp. 1621-1644 .

[28] J. Galvis, and Y. Efendiev, *Domain decomposition preconditioners for multiscale flows in high contrast media. Reduced dimension coarse spaces*, Multiscale Model. Simul. 8 (2010), pp. 1621-1644.

[29] J. Galvis, and Y. Efendiev, *Domain decomposition preconditioners for multiscale flows in high contrast media*, Multiscale Model. Simul. 8 (2010), pp. 1461-1483.

[30] W.R. Gardner, *Some steady-state solutions of the unsaturated mositure flow equation with application to evaporation from a water table*, Soil Sci. 85 (1958), pp. 228-232.

[31] I.G. Graham, P.O. Lechner, and R. Scheichl, *Domain decomposition for multiscale PDEs*, Numer. Math., 106 (2007), pp. 589-626.

[32] I.G. Graham, and R. Scheichl, *Robust domain decomposition algorithms for multiscale PDEs*, Numer. Methods Partial Differential Equations, 23 (2007), pp. 859-878.

[33] R. Haverkamp, M. Vauclin, J. Touma, P. Weirenga, and G. Vachaud, *Comparison of numerical simulation models for one-dimensional infiltration*, Soil Sci. Soc. Am. J., 41 (1977), pp. 285-294.

[34] R. Haverkamp, and M. Vauclin, *A note on estimating finite difference interblock hydraulic conductivity values for transient unsaturated flow problems*, Water Resources Research, 15 (1979), pp. 181-187.

[35] M.C. Hill, *Solving groundwater flow problems by conjugate gradient methods and the strongly implicit procedure*, Water Resources Research 26 (1990), pp. 1961-1969.

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

[37] T.Y. Hou, X.H. Wu, and Y. Zhang, *Removing the cell resonance error in the multiscale finite element method via a Petrov-Galerkin formulation*, Communications in Mathematical Sciences (2004), pp. 185-205.

[38] 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), pp. 913-943.

[39] J. K. Huang, and M. Th. van Genuchten, *The SWMS-3D Code for Simulating Water Flow and Solute Transport in Three-Dimensional Variably-Saturated Media, Version 1.0*, Research Report No. 139, U.S. Salinity Laboratory, USDA-ARS, Riverside, California, 1995.

[40] K. Huang, R. Zhang, and M.T. Van Genuchten, *An Eulerian-Lagrangian approach with an adaptively corrected method of characteristics to simulate variably saturated water flow*, Water Resources Research, 30 (1994): pp. 499-507.

[41] P.S. Huyakorn, and G.F. Pinder, *Computational methods in subsurface flow*, Academic Press,San Diego, 1985.

[42] P. Jenny, S.H. Lee, and H. Tchelepi, *Multi-scale finite volume method for elliptic problems in subsurface flow simulation*, J. Comput. Phys., 187 (2003), pp. 47-67.

[43] T. Kim, J. Pasciak, and P. Vassilevski, *Mesh-independent convergence of the modified inexact Newton method for a second order non-linear problem*, Numer. Linear Algebra Appl., 13 (2006), pp. 23-47.

[44] L.K. Kuiper, *A comparison of iterative methods as applied to the solution of the nonlinear three-dimensional groundwater flow equation.*, SIAM Journal of Scientific and Statistical Computing 8 (1987), pp. 521-528.

[45] M.Loéve, *Probability Theory,fourth ed.*,Springer, Berlin, 1977.

[46] Z. Lu, and D. Zhang, *Analytical solution to steady state unsaturated flow in layered randomly heterogeneous soils via Kirchhoff transformation*, Advances in Water Resources, 27 (2004), pp. 775-784.

[47] Y. Maday, *Reduced-basis method for the rapid and reliable solution of partial differential equations*, In: Proceedings of international conference of mathematicians, Madrid. European Mathematical Society, Zurich, 2006.

[48] T.P.A. Mathew, *Domain decomposition methods for the numerical solution of partial differential equations*, volume 61 of Lecture Notes in Computational Science and Engineering, Springer-Verlag, Berlin, 2008.

[49] M. Menziani, S. Pugnaghi, and S. Vincenzi, *Analytical solution of the linearized Richards equation for discrete arbitrary initial and boundary condition*, Journal of Hydrology, 332 (2007), pp. 214-225.

[50] P.D. Meyer, A.J. Valocchi, S.F. Ashby, and P.E. Saylar, *A numerical investigation of the conjugate gradient method as applied to three dimensional groundwater flow problems in randomly heterogeneous porous media*, Water Resources Research 25 (1989), pp. 1440-1446.

[51] Y. Mualem, *A new model for predicting the hydraulic conductivity of unsaturated porous media.* Water Resour. Res. 12 (1976), pp. 513-522.

[52] N.C Nguyen, *Multiscale reduced-basis method for parameterized elliptic partial differential equations with multiple scales*, J.Comp. Physics, 227 (208), 9807-9822.

[53] H. Owhadi, and L. Zhang, *Metric based up-scaling*, Comm. Pure and Applied Math., (2007), 675-723.

[54] M. Panfilov, *Macroscale models of flow through highly heterogeneous porous media*, Springer, Netherlands, 2000.

[55] C. Paniconi, S Ferraris, M. Putti, G Pini and G. Gambolati, *Three-dimensional numerical codes for simulating groundwater contamination: FLOW3D, flow in saturated and unsaturated porous media*, Computer Techniques in Environmental Studies V, Vol. 1: Pollution Modeling, (1994), pp. 149–156.

[56] A.J. Pullan, *The quasi linear approximation for unsaturated porous media flow*, Water Resour. Res., 26 (1990), pp. 1219-1234.

[57] L. Pan, A.W. Warrick, and P.J. Wierenga, *Finite elements methods for modeling water flow in variably saturated porous media: Numerical oscillation and mass distributed schemes*, Water Resources Research, 32 (1996), 1883-1889.

[58] j.R. Philip, *Theory of infiltration*, Advan. Hydrosci.5 (1969), pp. 215-296.

[59] S.K. Rajagopal, *On a hierarchy of approximate models for flows in incompressible fluids through porous solids*, Math. Models Methods Appl. Sci., 17 (2007), pp. 215-252.

[60] K. Rathfelde,r and L. M. Abriola, *Mass conservative numerical solutions of the head-based Richards' equation*, Water Resour. Res., (1994), pp. 2579-2586.

[61] F.A. Radu, I.S. Pop, and P. Knabner, *Error estimates for a mixed finite element discretization of some degenerate parabolic equations,* Numer. Math., 109 (2008), pp. 285-311.

[62] L.A. Richards, *Capillary conuction of liquids through porous mediums.*,Physics, 1 (1931), pp. 318-333.

[63] C. Robert and G. Casella, *Monte Carlo Statistical Methods*, Springer-Verlag, New-York, 1999.

[64] G. Rozza, D.B.P Huynh, and A.T. Patera, *Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations. Application to trasport and continuum mechanics.*, Arch Comput Methods Eng 15 (2008), pp. 229-275.

[65] C. Robert and G. Casella, *Monte Carlo Statistical Methods*, Springer-Verlag, New-York, 1999.

[66] M.G. Schaap, F. J. Leij, and M. Th. van Genuchten,*ROSETTA: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions*, J. Hydrol., 251 (2001), pp. 163-176.

[67] J. Simunek, M. Sejna, and M.T. van Genuchten, *HYDRUS-2D: simulating water flow and solute transport in two-dimensional variably saturated media* Tech. Rep., IGWMC, Golden, CO, USA, 1999.

[68] R.E. Smith, *Infiltration Theory for Hydrologic Applications*, AGU, Washington, 2002.

[69] R. Srivastava, and T.C.J. Yeh, *Analytical solutions for one-dimensional transient infiltration toward the water table in homogeneous and layered soils*, Water Resources Research, 27 (1991), pp. 753-762.

[70] X.C. Tai, and M. Espedal, *Applications of a space decomposition method to linear and nonlinear elliptic problems*, Numer. Methods Partial Differential Equations 14, no. 6 (1998), pp. 717-737.

[71] A. Toselli, and O. Widlund. *Domain decomposition methods—algorithms and theory*, volume 34 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 2005.

[72] J. Van lent, R. Scheichl, and I.G. Graham, *Energy-minimizing coarse spaces for two-level Schwarz methods for multiscale PDEs.* Numer. Linear Algebra Appl. 16 (2009), pp. 775-799.

[73] M.T. van Genuchten, *A closed-form equation for predicting the hydraulic onductivity of unsaturated soils*, Soil.Sci.Soc.Am.J.,44 (1980), pp. 892-898.

[74] M.T. Van Genuchten, *A comparison of numerical solution of the one-dimensional unsaturated-saturated flow and mass transport equations*, Advances in Water Resources, 5 (1982), pp. 47-55.

[75] A.W. Warrick, *Time-dependent linearized infiltration: III. strip and disc sources*,Soil.Sci.Soc.Am.J., 40 (1976), pp. 639-643.

[76] W. L. Wan, T. F. Chan, and B. Smith. *An energy-minimizing interpolation for robust multigrid methods*, SIAM J. Sci. Comput., 21 (2000), pp. 1632–1649.

[77] A.W. Warrick, *Time-dependent linearized infiltration: III. strip and disc sources*,Soil.Sci.Soc.Am.J.,40 (1976), pp. 639-643.

[78] E. Wong, *Stochastic Processes in Information and Dynamical Systems*, MCGraw-Hill, New York, 1971.

[79] X.H. Wu, Y. Efendiev, and T.Y. Hou, *Analysis of Upscaling Absolute Permeability,* Discrete and Continuous Dynamical Systems, Series B, (2002), pp. 185-204.

[80] J. Xu, *Two-grid discretization techniques for linear and nonlinear PDEs*, SIAM Journal on Numerical Analysis 33 (1996), pp. 1759-1777

[81] J. Xu, and L. Zikatanov, *On an energy minimizing basis for algebraic multigrid methods*, Comput. Visual Sci., 7 (2004), pp. 121-127.

[82] J. Xu and L. Zikatanov. *The method of subspace corrections and the method of alternating projections in Hilbert space*, J. of Amer. Math. Soc., 15 (2002), pp. 573–597.

VITA

Seul Ki Kang was born in Seoul, South Korea. She received her B.S. degree in mathematics from Korea Advanced Institute of Science and Technology, South Korea in February 2007. She then began to study at Texas A&M University and received her Ph.D. from the Department of Mathematics in August 2012. Seul ki Kang can be reached by writing to: 105-1801, SamsungRaemian APT, Munjungdong, Songphagu, Seoul, South Korea, or to the email address semodol@gmail.com.