MODEL REDUCTION TECHNIQUES AND PARAREAL ALGORITHM FOR MULTISCALE PROBLEMS

A Dissertation

by

JIUHUA HU

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Yalchin Efendiev
Co-Chair of Committee,	Tsz Shun Eric Chung
Committee Members,	Eduardo Gildin
	Jianxiin Zhou
Head of Department,	Sarah Witherspoon

August 2021

Major Subject: Mathematics

ABSTRACT

A broad range of scientific and engineering problems involve multiple scales. For example, composite material properties and subsurface propertities can vary over many length scales. Direct numerical methods of multiscale problems is often difficult due to the fact that a very fine mesh of the domain is required to reflect the heterogeneous coefficients. From a computational point of view, the major challenge to solve these problems is the size of the computation, even with the aid of supercomputers. On the other hand, from an engineering perspective, it is often sufficient to predict the macroscopic properties of the multiple scale systems, such as the effective conductivity, permeability, elastic moduli and eddy diffusivity. Therefore, it is desirable to develop fast and effective numerical methods that capture the small scale effect on the large scales, but do not require resolving all the small features.

There has been extensive research effort devoted to developing computational methods for multiscale problems. Among the most popular and developed techniques are homogenization method, multiscale finite element methods and parareal algorithm. The goal of homogenization methods and multiscale finite element methods is to construct numerical solvers on the coarse grid. Their resulting linear systems are typically much smaller than using fine grid. Parareal algorithm facilitates speeding up the numerical solver to time dependent equations on the condition of sufficient processors. Typically, parareal algorithm could result in less wall-clock time than sequentially computing.

In this dissertation, we will design and apply model reduction techniques to time-fractional diffusion equations, parabolic equations and stokes equations in heterogeneous media. Homogenization approach is studied for the time-fractional diffusion equation.

We discuss constraint energy minimizing generalized multiscale finite element method for the incompressible Stokes flow problem in a perforated domain. In this dissertation, we present two methodologies for parabolic problems with heterogeneous coefficients: a novel approach coupling multiscale methods with parareal algorithm and an efficient numerical solver coupling space-time finite element method and Non-local multi-continua technique. The former aims for time-independent permeability field and the latter for time-dependent permeability field.

ACKNOWLEDGMENTS

I would like to express my genuine gratitude to my advisor, Professor Yalchin Efendiev, for his excellent guidance and continuous support throughout my PhD study at Texas A&M University. His guidance helped me in all the time of research and career development. It is my honour to be one of his students. Without his long-term support, this dissertation would not be possible.

I am grateful to my co-advisor, Professor Eric T. Chung, who has always been patient and detailed in explaining his research ideas. I truly appreciate all his contributions of times and ideas to help me complete this dissertation.

I owe a deep sense of gratitude to Professor Guanglian Li. Her timely suggestions with kindness, insightful comments and encouragement has enabled me to complete the dissertation. I also would like to thank Professor Eduardo Gildin and Professor Jianxin Zhou for serving as my committee members, guiding my research and providing me priceless advice in the past several years. My sincere thanks also goes to Professor Raytcho Lazarov and Professor Matthias Maier for their constant encouragement on my research and career development.

My PhD study was much more enriched by my friends from the Department of Mathematics of Texas A&M University. I would also like to thank all my friends for their support and encouragement during my difficult times.

Last but not least, I thank my family for supporting my decision to embark on this path. None of this would have been possible without their support. In particular, I would like to thank my husband Jintao Deng for his encouragement sustained me through graduate school in many ways. Thanks to my sister for taking care of my parents while I am away.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor Yalchin Efendiev and Professor Eric T. Chung of the Department of Mathematics.

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by a fellowship from Texas A&M University.

NOMENCLATURE

Ω	Spatial domain
κ	Permeability
$[0, T]$	Temporal domain
h	Fine spatial mesh size
H	Coarse spatial mesh size
δt	Fine time mesh size
Δt	Coarse time mesh size
$\mathcal{T}^h$	Fine-scale partition of spatial domain
$\mathcal{T}^H$	Coarse-scale partition of spatial domain
$\mathcal{T}^\delta$	Fine-scale partition of time domain
$\mathcal{T}^\Delta$	Coarse-scale partition of time domain
K	Coarse grid element
E	Coarse grid edge
ω	Coarse neighborhood
χ	Partition of unity

TABLE OF CONTENTS

Page

ABSTRACT ii

ACKNOWLEDGMENTS iv

CONTRIBUTORS AND FUNDING SOURCES v

NOMENCLATURE vi

TABLE OF CONTENTS vii

LIST OF FIGURES ix

LIST OF TABLES xii

1. INTRODUCTION 1

2. HOMOGENIZATION OF TIME-FRACTIONAL DIFFUSION EQUATIONS WITH PERIODIC COEFFICIENTS 4

   2.1 Two-scale asymptotic expansion 6
   2.2 First order approximation estimate 11
   2.3 Numerical experiments 21
      2.3.1 Numerical tests with smooth permeability fields 22
      2.3.2 Numerical tests with non-smooth permeability fields 29
      2.3.3 Numerical tests with rough initial data 31

3. CONVERGENCE OF THE CEM-GMSFEM FOR STOKES FLOWS IN HETEROGENEOUS PERFORATED DOMAINS 35

   3.1 Problem setting 36
      3.1.1 Model problem 36
      3.1.2 Function spaces 37
      3.1.3 Variational formulation and fine-grid discretization 38
   3.2 Construction of multiscale spaces 40
      3.2.1 Auxiliary space 41
      3.2.2 Multiscale space 42
      3.2.3 The multiscale method 44

3.3 Convergence analysis 46

4. WAVELET-BASED EDGE MULTISCALE PARAREAL ALGORITHM FOR PARABOLIC EQUATIONS WITH HETEROGENEOUS COEFFICIENTS AND ROUGH INITIAL DATA 61

4.1 Problem setting and the construction of multiscale space 62
4.1.1 Full discretization 63
4.1.2 Multiscale model reduction in the spatial domain D 65
4.1.2.1 Multiscale solver in the spatial domain 65
4.1.2.2 Multiscale space construction 67
4.2 Wavelet-based Edge Multiscale Parareal Algorithm 69
4.3 Convergence study 72
4.3.1 Convergence for Algorithm 1 74
4.3.2 Convergence for Algorithm 2 76
4.4 Numerical results 82
4.4.1 Numerical tests with nonzero source term 84
4.4.2 Numerical tests with a vanishing source term 89

5. SPACE-TIME NON-LOCAL MULTI-CONTINUA UPSCALING FOR PARABOLIC EQUATIONS WITH MOVING CHANNELIZED MEDIA 93

5.1 Problem setting 93
5.1.1 Space-time variational formulation and space-time discretization ... 95
5.1.2 Functional spaces and bilinear forms 96
5.2 Space-time NLMC 99
5.2.1 Global multiscale space 99
5.2.2 Localization of global multiscale basis functions 99
5.3 Convergence analysis 100
5.4 Numerical results 118
5.4.1 Experiment 1: Slow moving permeability 120
5.4.2 Experiment 2: Faster moving permeability 122

6. SUMMARY AND CONCLUSIONS 126

REFERENCES 128

LIST OF FIGURES

FIGURE Page

1.1 Illustration of heterogeneous media. Reprinted from [1]..... 1

2.1 Initial data: $a(x_1, x_2)$ 22

2.2 A smooth periodic permeability field with $\epsilon = \frac{1}{8}$ in a cell and over the domain D: $\kappa(y_1, y_2)$ and $\kappa^{\epsilon}(x_1, x_2)$..... 23

2.3 The fine scale solution $u_h^{\epsilon,k}$ for $k = 11, 51$ and 101 to Problem (2.1) with κ^{ε} in (2.31) and $\epsilon := \frac{1}{8}$. 24

2.4 solutions to cell problem: $\chi_1(y_1, y_2)$ and $\chi_2(y_1, y_2)$ 25

2.5 The homogenized solution $u_h^{0,k}$ for $k = 11, 51$ and 101 to Problem (2.1) with κ in (2.31) and $\epsilon := \frac{1}{8}$..... 26

2.6 The first order approximation solution $U_{1,h}^{\epsilon,k}$ for $k = 11, 51$ and 101 to Problem (2.1) with κ in (2.31) and $\epsilon := \frac{1}{8}$..... 26

2.7 A nonsmooth permeability field of smaller variation: $\kappa_1(x_1, x_2)$ and $\kappa_1^{\epsilon}(x_1, x_2)$ with $\epsilon := \frac{1}{8}$..... 30

2.8 A nonsmooth permeability field of larger variation: $\kappa_2(x_1, x_2)$ and $\kappa_2^{\epsilon}(x_1, x_2)$ with $\epsilon = \frac{1}{8}$. 30

2.9 The initial data $a(x_1, x_2)$ as defined in (2.35). 32

3.1 Illustration of a perforated domain 36

3.2 Illustration of the coarse grid, the fine grid, and the oversampling domain... 40

4.1 Illustration of a coarse neighborhood and coarse element..... 66

4.2 The heterogeneous permeability field κ and the initial data $u_0 = x(1 - x)y(1 - y)$. 83

4.3 Numerical solution U_h^n to (4.4) for $n = 10^3, 3 \times 10^3, 5 \times 10^3$ and 10^4 with $\delta t = 10^{-4}$..... 84

4.4 Numerical solutions U_k^n for $n = 1, 3, 5, 10$ from Algorithm 2 with $\Delta T = 0.1$ and $\delta t = 10^{-3}$, backward Euler scheme: iteration number $k = 0$ (top), $k = 1$ (middle) and $k = 2$ (bottom). 86

4.5 Convergence history of Algorithm 2 in relative $L^2(D)$ error and relative $H^1_\kappa(D)$ error for Experiment 1: backward Euler scheme with $\Delta T = 0.1$ and $\delta t = 10^{-3}$ 86

4.6 Convergence history of Algorithm 2 in relative $L^2(D)$-norm and relative $H^1_\kappa(D)$-norm for Experiment 2: Crank-Nicolson scheme with $\Delta T = 0.1$ and $\delta t = 10^{-3}$ 88

4.7 Convergence history of Algorithm 2 in relative $L^2(D)$ error and relative $H^1_\kappa(D)$ error for Experiment 3: backward Euler scheme with $\Delta T = 10^{-2}$ and $\delta t = 10^{-3}$ 89

4.8 Numerical solutions U_h^n to (4.4) with $f = 0$ for $n = 10, 30, 50$ and 100 with $\delta t = 10^{-3}$ 89

4.9 Numerical solutions U_k^n for $n = 1, 3, 5, 10$ from Algorithm 2 with $\Delta T = 10^{-2}$ and $\delta t = 10^{-3}$, backward Euler scheme: iteration number $k = 0$ (top), $k = 1$ (middle) and $k = 2$ (bottom). 90

4.10 Convergence history of Algorithm 2 in relative $L^2(D)$-norm and relative $H^1_\kappa(D)$-norm for $f = 0$: backward Euler scheme with $\Delta T = 10^{-2}$ and $\delta t = 10^{-3}$ 91

4.11 Convergence history of Algorithm 2 in relative $L^2(D)$-norm and relative $H^1_\kappa(D)$-norm for $f = 0$: Crank-Nicolson scheme with $\Delta T = 10^{-2}$ and $\delta t = 10^{-3}$ 92

5.1 Illustration of oversampling space domain K_1^i. 96

5.2 Permeability field $\kappa(x_1, x_2, 0)$ and $\kappa(x_1, x_2, 0.5)$ for Experiment 1 120

5.3 Snapshot of the reference solutions $\tilde{U}_{h,\delta,t}$ for $t = 0.25, 0.5, 0.75, 1.0$. 121

5.4 Snapshot of the multiscale solutions $\tilde{U}^\ell_{\text{ms},t}$ at $t = 0.25, 0.5, 0.75, 1.0$ with oversampling layer $\ell = 1$ (top), $\ell = 2$ (middle), $\ell = 3$ (bottom). 121

5.5 Permeability field $\kappa(x_1, x_2, 0), \kappa(x_1, x_2, 0.5), \kappa(x_1, x_2, 0.8)$ and $\kappa(x_1, x_2, 1.0)$. 123

5.6 Snapshot of the reference solutions $\tilde{U}_{h,\delta,t}$ for $t = 0.2, 0.5, 0.8, 1.0$. 123

5.7 Snapshot of the multiscale solutions $\tilde{U}^{\ell}_{\mathrm{ms},t}$ at $t = 0.2, 0.5, 0.8, 1.0$ with oversampling layer $\ell = 1$ (top), $\ell = 2$ (middle), $\ell = 3$ (bottom). 124

LIST OF TABLES

TABLE Page

2.1 The convergence history of the first order approximation to Problem (2.1) with κ in (2.31) and $\epsilon := \frac{1}{8}$. 27

2.2 The convergence history of the first order approximation to Problem (2.1) with κ in (2.32) and $\epsilon := \frac{1}{8}$. 28

2.3 The convergence history of the first order approximation to Problem (2.1) with κ in (2.31) and $\epsilon := \frac{1}{16}$. 28

2.4 The convergence history of the first order approximation to Problem (2.1) with κ in (2.31) and $\epsilon := \frac{1}{32}$. 29

2.5 The convergence history of the first order approximation to Problem (2.1) with $\kappa := \kappa_1$ in (2.33) and $\epsilon := \frac{1}{8}$..... 31

2.6 The convergence history of the first order approximation to Problem (2.1) with $\kappa := \kappa_2$ in (2.34) and $\epsilon := \frac{1}{8}$..... 31

2.7 The convergence history of the first order approximation to Problem (2.1) with $a(x)$ defined in (2.35), κ in (2.31) and $\epsilon := \frac{1}{8}$. 33

2.8 The convergence history of the first order approximation to Problem (2.1) with $a(x)$ defined in (2.35), κ in (2.31) and $\epsilon := \frac{1}{16}$. 33

2.9 The convergence history of the first order approximation to Problem (2.1) with $a(x)$ defined in (2.35), κ in (2.31) and $\epsilon := \frac{1}{32}$. 34

5.1 Convergence history of Experiment 1. 122

5.2 Convergence history of Experiment 2. 125

1. INTRODUCTION

A broad range of scientific and engineering problems involve multiple scales. For example, composite material properties and subsurface properties can vary over many length scales. Examples of heterogeneous media are shown in Figure 1.1. Direct numerical methods of multiscale problems is often difficult due to the fact that a very fine mesh of the domain is required to reflect the heterogeneous coefficients. From a computational point of view, the major challenge to solve these problems is the size of the computation, even with the aid of supercomputers. On the other hand, from an engineering perspective, it is often sufficient to predict the macroscopic properties of the multiple scale systems, such as the effiective conductivity, permeability, elastic moduli and eddy diffusivity. Therefore, it is desirable to develop fast and effective numerical methods that capture the small scale effect on the large scales, but do not require resolving all the small features.

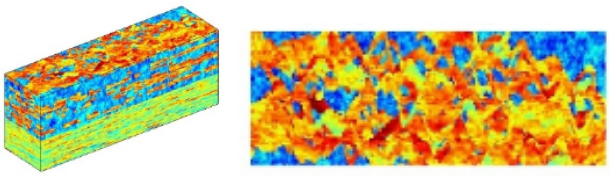

Figure 1.1: Illustration of heterogeneous media. Reprinted from [1].

There have been many existing multiscale model reduction techniques in the literature to deal with multiscale problems. These multiscale approaches include homogenization approaches [4, 5, 6, 8, 33, 39], multiscale finite element methods (MsFEMs) [12, 13, 24, 25, 26, 36], heterogeneous multiscale methods (HMMs) [3, 11, 21, 62], variational multiscale methods [37, 42, 43, 48], flux norm approach [10], generalized multiscale

finite element methods (GMsFEMs) [14, 16, 17, 18, 22, 23] and localized orthogonal decomposition (LOD) [2, 27, 35, 51].

Homogenization is a well-known upscaling method. It constructs homogenized equations whose coefficients depend only on the macroscopic variable. The solutions to the homogenized equations can be solved using coarse mesh and serve as an approximation to the exact solution in the homogenization limit. On the other hand, the main idea of MsFEM and similar methods, like GMsFEM, is to construct multiscale basis functions which capture the small scale information within each coarse grid. The small scale information of the coarse grids is then brought to the large scales. GMsFEM is designed to construct more basis functions for each coarse region. It has been successfully applied in simulating multiscale problems in channelized permeability. This is mainly because the local problems can correctly identify the necessary channels without any geometry interpretation. Constraint Energy Minimizing Generalized Multiscale Finite Element Method (CEM-GMsFEM) shares some ideas of GMsFEM. It constructs multiscale basis functions by solving a minimization problems on oversampling domains. It can be shown that with an appropriate choice of oversampling layer, the convergence of the method is independent of the contrast from the heterogeneities and the error linearly decreases with respect to coarse mesh size. Non-local multi-continua (NLMC) upscaling method follows the general framework of CEM-GMsFEM. It identifies appropriate local problems together with local spectral modes to represent each continuum. These approaches have achieved great success in the efficient and accurate simulation of heterogeneous problems.

With the rapid development of supercomputer, one may utilize parallel computing to obtain multiscale solutions within a much shorter time. Parareal Algorithm is an iterative method which allows parallel computing to solve time-dependent differential equations. Two solvers: inaccurate but cheap coarse solver and accurate but expensive fine solver, are introduced for Parareal Algorithm. Fine solvers could be performed in parallel using

enough processors. The time for the solution to converge is about the same order of time when only using coarse solver. Therefore Parareal Algorithm can substantially reduce wall-clock computation time.

In this dissertation, we will design and apply model reduction techniques to time-fractional diffusion equations, parabolic equations and stokes equations in heterogeneous media. Time-fractional diffusion equations are investigated using homogenization approach in Chapter 2. We consider in Chapter 3 the incompressible Stokes flow problem in a perforated domain and employ the constraint energy minimizing generalized multiscale finite element method to solve this problem. Parabolic equations are investigated in Chapter 4 and Chapter 5. In Chapter 4, we propose the Wavelet-based Edge Multiscale Parareal Algorithm to solve parabolic equations with heterogeneous time-independent coefficients. Finally, we consider the parabolic equations with time-dependent heterogeneous coefficients in Chapter 5.

2. HOMOGENIZATION OF TIME-FRACTIONAL DIFFUSION EQUATIONS WITH PERIODIC COEFFICIENTS *

In this chapter, we study homogenization approach for the initial boundary value problem for the time-fractional diffusion equation with a homogeneous Dirichlet boundary condition and an inhomogeneous initial data in a bounded convex polyhedral domain. Let D be a bounded domain in $\mathbb{R}^d$ $(d = 1, 2, 3)$ with a sufficiently smooth boundary ∂D. We consider a partial differential equation involving a fractional derivative in time t, satisfying:

$$\begin{cases} \partial_t^\alpha u^\epsilon(x,t) = \nabla \cdot (\kappa^\varepsilon(x) \nabla u^\epsilon(x,t)) & \text{in } D, \quad t \in (0,T] \\ u^\epsilon = 0 & \text{on } \partial D, \ t \in (0,T] \\ u^\epsilon(\cdot, 0) = a(x) & \text{in } D. \end{cases} \tag{2.1}$$

Here, $0 < \alpha < 1$ is a given fixed parameter and the matrix $\kappa^\varepsilon(x) := \kappa(x/\epsilon)$ is periodic with period ε. Let $(\kappa_{ij}(x))_{i,j=1}^d$ be symmetric with $\kappa_{ij}(x) \in C^\infty(D)$, $i, j = 1, \ldots, d$. We assume for some constant $\mu > 0$, there holds

$$\kappa_{ij}(x)\xi_i\xi_j \geq \mu|\xi|^2 \text{ for all } \xi \in \mathbb{R}^d \text{ and } x \in D.$$

The initial data $a(x) \in L^2(D)$ is a given macro-scale function and $T > 0$ is a fixed value.

In the model problem (2.1), $\partial_t^\alpha w$ refers to the left-sided Caputo fractional derivative of order α of the function $w(t)$, defined by (see, e.g. [44, p. 91, (2.4.1)] or [56, p. 78])

$$\partial_t^\alpha w(t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{1}{(t-s)^\alpha} w'(s)\, \mathrm{ds}.$$

*Reprinted with permission from "Homogenization of time-fractional diffusion equations with periodic coefficients" by Jiuhua Hu and Guanglian Li, 2020. Journal of Computational Physics, Volume 408, 1 May 2020, 109231, Copyright [2020] by Elsevier.

Fractional diffusion equations were introduced in physics with the aim of describing diffusions in media with fractal geometry [54]. They have been applied to many fields, e.g., in engineering, physics, biology and finance. Their practical applications include electron transport in Xerox photocopier, visco-elastic materials, and protein transport in cell membranes [58, 31, 45]. In this chapter, we are concerned with the time-fractional diffusion problem (2.1) in a heterogeneous periodic medium $\kappa^\epsilon(x)$, which is utilized in many important applications, e.g., porous media and composite material modeling. Most recently, periodic structures are utilized in metamaterials [61] to design novel materials. The main challenge in the classical numerical treatment of these applications is that it becomes prohibitively expensive and even intractable at the microscale as $\epsilon \to 0$.

The goal of this chapter is to construct an efficient numerical solver for (2.1) based on homogenization theory [38]. The main idea of homogenization is to obtain the effective or homogenized problem by solving d cell problems, and then the corresponding homogenized solution u_0 serves as a good approximation to original unknown u^ϵ as the period $\epsilon \to 0$. To the best of our knowledge, there has been no such result for time-fractional diffusion problems so far. However, there are quite a few results on the parabolic equations, i.e., $\alpha = 1$ in (2.1), in the same setting, cf. [55, 65]. Nevertheless, it is nontrivial to generalize the results for the parabolic equations to the time-fractional equations with the same technique, mainly due to the lack of the product rule for the fractional derivative.

We prove in Corollary 2.2.7 the error between the exact solution u^ϵ and its first order approximation u_1^ϵ is of $\mathcal{O}(\epsilon^{\min\{1/2,2/d-1/2\}})$ in the Bochner space $L^p((\theta,T];H^1(D))$ for any $\theta \in (0,T)$ as $\epsilon \to 0$. We also derive in Corollary 2.2.7 the pointwise-in-time error estimate in $L^2(D)$ norm for the first order corrector for $t \in (\theta,T]$ with the same convergence rate in ϵ. Thus, the first order approximation u_1^ϵ in $L^p((\theta,T];H^1(D))$ norm for any $\theta \in (0,T)$ and $p \in [1,\infty)$ norm as $\epsilon \to 0$ achieves optimal convergence rate of $\mathcal{O}(\epsilon^{1/2})$ when $d \le 2$ and a convergence rate of $\mathcal{O}(\epsilon^{1/6})$ when $d = 3$. For the latter case, it is un-

clear whether or not this rate is optimal. The proof of the result relies on the regularity of time-fractional diffusion problems [57] and the introduction of cut-off functions to handle boundary layers and initial data [52, 65]. When the initial data has a better regularity, e.g., $a(x) \in H_0^1(D) \cap H^2(D)$, there is no need to introduce a proper cut-off function for the time domain.

Furthermore, we present in Section 2.3 a number of numerical tests for $d = 2$ to verify Corollary 2.2.7. Specifically, we test the cases when $\kappa(x)$ is smooth and discontinuous, and when $\kappa(x)$ admits large deviations. Our numerical tests demonstrate a higher convergence rate, namely, $\mathcal{O}(\epsilon)$, compared to the theoretical result. We also observe that a larger deviation or discontinuity in the coefficient $\kappa(x)$ results in a larger error.

The remainder of this chapter is organized as follows. In Section 2.1 we present the homogenized equation using two-scale asymptotic expansion for problem (2.1), as well as regularity results on time-fractional diffusion problems. We then introduce auxiliary functions and define the first order approximation in Section 2.2. The main error estimate is derived therein. To verify our theoretical findings, we present in Section 2.3 extensive numerical tests with smooth (or nonsmooth) diffusion coefficient and with smooth (or nonsmooth) initial data $a(x)$.

2.1 Two-scale asymptotic expansion

This section is concerned with the two-scale asymptotic expansion of the solution u^ϵ to (2.1). The approach is standard and can be found, e.g., in [38]. We recall the general procedure for the sake of completeness.

First, analogous to the standard homogenization theory, we denote $y := x/\epsilon$ as the fast variable, and x is referred to as the slow variable. Let $L^2_\#(Y) := \{u \in L^2(Y) : u \text{ is } Y\text{-periodic}\}$ with Y being a unit cell in $\mathbb{R}^d$. Similarly, we can define $H^1_\#(Y)$. Denote $V_\#(Y) := \{v \in H^1_\#(Y) : \langle v \rangle = 0\}$ with $\langle \cdot \rangle := \frac{1}{|Y|}\int_Y \cdot \, \mathrm{d}y$ being the average over the unit

cell Y. The notation $A \lesssim B$ denotes $A \leq CB$ for some constant C independent of the microscale ϵ. Throughout the chapter, we follow the Einstein summation convention.

Under the assumption that the fast variable y and the slow variable x are independent when $\epsilon \to 0$, we seek for an asymptotic expansion of the solution $u^\epsilon(x,t)$ as follows:

$$u^\epsilon(x,t) = u_0(x,y,t) + \epsilon u_1(x,y,t) + \epsilon^2 u_2(x,y,t) + \cdots \tag{2.2}$$

with the functions $u_j(x,y,t) \in H^1_\#(Y)$ being periodic in the fast variable y with period 1. The leading order term $u_0(x,y,t)$ is referred to as the homogenized solution and the following terms $\epsilon^k u_k(x,y,t)$ are the k^{th} order corrector for $u^\epsilon(x,t)$ for $k = 1, 2, \cdots$.

Denote by A^ϵ the second order elliptic operator

$$A^\epsilon = -\frac{\partial}{\partial x_i}\left(\kappa_{ij}\left(x/\epsilon\right)\frac{\partial}{\partial x_j}\right).$$

With this notation, we can expand A^ϵ as follows

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

where

$$\begin{aligned}
A_1 &= -\frac{\partial}{\partial y_i}\left(\kappa_{ij}(y)\frac{\partial}{\partial y_j}\right), \\
A_2 &= -\frac{\partial}{\partial y_i}\left(\kappa_{ij}(y)\frac{\partial}{\partial x_j}\right) - \frac{\partial}{\partial x_i}\left(\kappa_{ij}(y)\frac{\partial}{\partial y_j}\right), \\
A_3 &= -\frac{\partial}{\partial x_i}\left(\kappa_{ij}(y)\frac{\partial}{\partial x_j}\right).
\end{aligned}$$

Substituting the expansions for u^ϵ and A^ϵ into the differential equation (2.1), and equating

the terms with the same power of ϵ, we get

$$A_1 u_0 = 0, \tag{2.3}$$

$$A_1 u_1 + A_2 u_0 = 0, \tag{2.4}$$

$$A_1 u_2 + A_2 u_1 + A_3 u_0 = -\partial_t^\alpha u_0. \tag{2.5}$$

Next we examine these equations one by one. Equation (2.3) is equivalent to seeking $u_0(x, y, t) \in H^1_\#(Y)$, satisfying

$$-\frac{\partial}{\partial y_i}\left(\kappa_{ij}(y)\frac{\partial}{\partial y_j}\right) u_0(x, y, t) = 0.$$

The theory of second order elliptic PDEs implies that $u_0(x, y, t)$ is independent of y. Consequently, we obtain

$$u_0(x, y, t) = u_0(x, t). \tag{2.6}$$

Since there is no micro-scale ϵ in either the boundary or the initial conditions, those conditions are imposed on the leading order term u_0 directly, i.e.,

$$\begin{aligned} u_0 &= 0 \qquad \text{on } \partial D \\ u_0(\cdot, 0) &= a(x) \quad \text{in } D. \end{aligned} \tag{2.7}$$

Thanks to (2.6), the second equation (2.4) can be formulated as seeking for $u_1 \in H^1_\#(Y)$, satisfying

$$-\frac{\partial}{\partial y_i}\left(\kappa_{ij}(y)\frac{\partial}{\partial y_j}\right) u_1 = \left(\frac{\partial}{\partial y_i}\kappa_{ij}(y)\right)\frac{\partial u_0}{\partial x_j}(x, t).$$

Note that u_1 can be defined alternatively independent of the slow variable x. To this end, let $\chi_k \in V_{\#}(Y)$ be the solution to the following *cell problem*:

$$-\frac{\partial}{\partial y_i}\left(\kappa_{ij}(y)\frac{\partial}{\partial y_j}\right)\chi_k = \frac{\partial}{\partial y_i}\kappa_{ik}(y) \text{ in } Y. \tag{2.8}$$

The general solution of equation (2.8) for u_1 then admits the expression

$$u_1(x,y,t) = \chi_j(y)\frac{\partial u_0}{\partial x_j}(x,t) + \tilde{u}_1(x,t)\ . \tag{2.9}$$

For simplicity, we take $\tilde{u}_1(x,t) = 0$.

Finally, we deal with the third term u_2. Because of (2.5), we can seek for $u_2 \in H^1_{\#}(Y)$, such that,

$$\frac{\partial}{\partial y_i}\left(\kappa_{ij}(y)\frac{\partial}{\partial y_j}\right)u_2 = A_2u_1 + A_3u_0 + \partial_t^\alpha u_0\ . \tag{2.10}$$

The solvability condition implies that the right hand side of (2.10) must have mean zero in y over the unit cell $Y = [0,1]^d$, i.e.

$$\int_Y (A_2u_1 + A_3u_0 + \partial_t^\alpha u_0)\,\mathrm{d}y = 0.$$

We note that

$$\int_Y \frac{\partial}{\partial y_i}F(x,y,t)\,\mathrm{d}y = 0$$

for any $F(x,y,t)$ which is periodic with respect to y. This can be easily verified using the Divergence Theorem. After integrating (2.10) over Y, the average over the terms starting with $\frac{\partial}{\partial y_i}$ disappears and we arrive at

$$\partial_t^\alpha u_0 - \frac{\partial}{\partial x_i}\left(\langle\kappa_{ij}(y)\rangle\frac{\partial}{\partial x_j}u_0\right) - \frac{\partial}{\partial x_i}\left(\left\langle\kappa_{ij}(y)\frac{\partial}{\partial y_j}u_1\right\rangle\right) = 0.$$

Substituting the expression for u_1 into this equation, we obtain the homogenized equation:

$$\partial_t^\alpha u_0 - \frac{\partial}{\partial x_i}\left(\kappa_{ij}^* \frac{\partial}{\partial x_j}\right) u_0 = 0\,, \tag{2.11}$$

where

$$\kappa_{ij}^* = \frac{1}{|Y|}\left(\int_Y \kappa_{ij} + \kappa_{ik}\frac{\partial \chi_j}{\partial y_k}\,\mathrm{d}y\right). \tag{2.12}$$

Last, we derive the *a priori* estimate for the homogenized solution u_0, which will be utilized below. We obtain from (2.7) and (2.11) that

$$\begin{cases} \partial_t^\alpha u_0 = \dfrac{\partial}{\partial x_i}\left(\kappa_{ij}^* \dfrac{\partial}{\partial x_j}\right) u_0 & \text{in } D, \quad t \in (0,T] \\ u_0 = 0 & \text{on } \partial D, \ \ t \in (0,T] \\ u_0(\cdot,0) = a(x) & \text{in } D. \end{cases} \tag{2.13}$$

Then by application of [57, Theorem 2.1], we obtain that $u_0 \in C([0,T]; L^2(D)) \cap C((0,T]; H^2(D) \cap H_0^1(D))$. Furthermore, the following estimate holds:

$$t^\alpha \left\|u_0(\cdot,t)\right\|_{H^2(D)} + \left\|u_0(\cdot,t)\right\|_{L^2(D)} + t^\alpha \left\|\partial_t^\alpha u_0(\cdot,t)\right\|_{L^2(D)} \lesssim \left\|a\right\|_{L^2(D)}\,, \quad \text{for all } t \in (0,T]. \tag{2.14}$$

Estimate (2.14) implies that the homogenized solution $u_0(x,t)$ is singular at $t = 0$. The strength of this singularity depends on the regularity of the initial data $a(x)$. This singularity can disappear when the initial data $a(x)$ has certain regularity. Furthermore, the regularity of the solution u_0 from time-fractional diffusion problem has only limited regularity over the space domain D, c.f. [41].

Remark 2.1.1. *Note that [57, Theorem 2.1] still holds when the permeability coefficient κ^ε admits high oscillation. Furthermore, the a priori estimate* (2.14) *is stable with respect*

to the parameter ε.

2.2 First order approximation estimate

We present in this section the first order approximation to u^ε, and then derive its error estimate.

To this end, we will first provide the *a priori* estimate to the solutions of the cell problem (2.8). This result can be found, e.g., in [55, Section 8]. For the completeness, we also present the proof:

Lemma 2.2.1. *Let χ_j be the solution to the cell problem* (2.8) *for all $j = 1, \cdots, d$. Then there holds*

$$\|\nabla \chi_j\|_{L^2(Y)} \leq \mu^{-1} \|\kappa\|_{L^2(Y)}, \tag{2.15}$$

$$\|\chi_j\|_{H^2(Y)} + \|\chi_j\|_{L^\infty(Y)} \lesssim 1. \tag{2.16}$$

Proof. Let e_j be the j^{th} canonical unit vector in $\mathbb{R}^d$. The weak formulation associated to Problem (2.8) is to seek $\chi_j \in V_\#(Y)$ such that

$$\int_Y \kappa(y) \nabla \chi_j \cdot \nabla v \, \mathrm{d}y = -\int_Y \kappa(y) e_j \cdot \nabla v \, \mathrm{d}y, \text{ for all } v \in V_\#(Y).$$

Testing with $v := \chi_j$, we obtain

$$\int_Y \kappa(y) \nabla \chi_j \cdot \nabla \chi_j \, \mathrm{d}y = -\int_Y \kappa(y) e_j \cdot \nabla \chi_j \, \mathrm{d}y.$$

Then the boundedness of $\kappa(y)$ and an application of the Hölder's inequality reveal the assertion (2.15).

Furthermore, by the regularity results of PDEs with periodic boundary conditions that

can be derived from, e.g., [32, Chapter 3] and [46, P69, Chapter 2], we obtain

$$\|\chi_j\|_{H^2(Y)} \lesssim \|\nabla \cdot (\kappa \nabla \chi_j)\|_{L^2(Y)} \lesssim 1. \tag{2.17}$$

This, together with the weak maximum principle, yields the second assertion (2.16). □

Note that both the exact solution u^ϵ and the homogenized solution u_0 satisfy a homogeneous Dirichlet boundary condition. However, the first order corrector $\epsilon\chi_j(x/\epsilon)\frac{\partial u_0}{\partial x_j}(x,t)$ admits ϵ-oscillation over the global boundary ∂D. To account for this fact, we must adapt the corrector near the boundary. We employ an approach developed in [52].

To this end, we introduce the cut-off function ζ^ε corresponding to ∂D. Here $\zeta^\varepsilon = 1$ on ∂D and $\text{supp}(\zeta^\varepsilon) \subset \{x \in \bar{D} : \text{dist}(x, \partial D) \le \varepsilon\}$. With regularity condition $\zeta^\varepsilon \in C^2(\bar{D})$, for $0 \le \ell \le 2$, we have

$$\|D^\ell \zeta^\epsilon\|_{L^\infty(D)} \lesssim \varepsilon^{-\ell}. \tag{2.18}$$

In addition, using (2.18) and the fact that $|\text{supp}(\zeta^\varepsilon)| = O(\varepsilon)$, and also by applying Hölder's inequality, we have the following estimate for derivatives of the cut-off function ζ^ε in $L^q(D)$ given by

$$\left\|D^\ell \zeta^\varepsilon\right\|_{L^q(D)} \le |\text{supp}(\zeta^\varepsilon)|^{1/q} \left\|D^\ell \zeta^\varepsilon\right\|_{L^\infty(D)} \lesssim \varepsilon^{1/q-\ell}. \tag{2.19}$$

To avoid higher regularity condition on the initial data $a(x)$, we employ the trick in [65] and introduce another cutoff function $\eta(t;\theta)$ in the time domain $[0,T]$ for any parameter

$\theta \in (0, T]$. It is defined as follows:

$$\eta(t;\theta) = \begin{cases} 0, & \text{for } t \le \theta/2 \\ 1, & \text{for } t \ge \theta \\ \text{linear} & \text{for } t \in (\theta/2, \theta). \end{cases}$$

Then by definition, we obtain the estimate for the fractional derivative of $\eta(t;\theta)$,

$$\partial_t^\alpha \eta(t;\theta) \lesssim \theta^{-\alpha} \chi_{t>\theta/2}. \tag{2.20}$$

Let the first order approximation to $u^\epsilon(x,t)$, and its modification be defined by

$$\begin{cases} U_1^\epsilon(x,t) := u_0(x,t) + \epsilon \chi_j(x/\epsilon) \dfrac{\partial u_0}{\partial x_j} \\ u_1^\epsilon(x,t) := u_0(x,t) + \epsilon(1-\zeta^\varepsilon)\chi_j(x/\epsilon) \dfrac{\partial u_0}{\partial x_j}. \end{cases} \tag{2.21}$$

Let $R(x,t;\theta) = \eta(t;\theta)(u^\epsilon - u_1^\epsilon)$, for any $\theta \in (0,T]$. Then plugging this expression of u^ϵ into (2.1) yields

$$\partial_t^\alpha \Big(u_1^\epsilon + R(x,t;\theta) \Big) = \nabla \cdot \Big(\kappa(x/\epsilon) \nabla (u_1^\epsilon + R(x,t;\theta)) \Big) \text{ in } D \times (0,T].$$

Collecting the terms and using the homogenized equation (2.11), the boundary and initial conditions for $u_0(x,t)$, we arrive at the following error equation

$$\begin{cases} \partial_t^\alpha R = \nabla \cdot (\kappa^\epsilon \nabla R) + f_1 & \text{in } D, \quad t \in (0,T] \\ R = 0 & \text{on } \partial D, \ t \in (0,T] \\ R(x,0;\theta) = 0 & \text{in } D. \end{cases} \tag{2.22}$$

Here,

$$f_1(x,t;\theta) := \partial_t^\alpha \eta(t;\theta)(u^\epsilon - u_1^\epsilon) + \eta(t;\theta)\Big(- \nabla\cdot(\kappa^\epsilon \nabla u_1^\epsilon) + \nabla\cdot(\kappa^* \nabla u_0)\Big) \\ - \eta(t;\theta)\epsilon(\zeta^\epsilon - 1)\chi_j(\frac{x}{\epsilon})\partial_t^\alpha \frac{\partial u_0}{\partial x_j}.$$

Remark 2.2.2 (Regularity of the modified first order approximation (2.21))**.** *Thanks to the condition $a \in L^2(D)$, we can obtain $u_1^\epsilon(x,t) \in C((0,T];H^1(D))$.*

A fundamental ingredient of the homogenization theory will be a proper *a priori* estimate for the residual $R(x,t;\theta)$. To this end, we first estimate the term f_1, which relies on the following pointwise-in-time error estimate in $L^2(D)$ norm for the modified first order approximation.

Lemma 2.2.3 (Pointwise-in-time error estimate in $L^2(D)$ norm for the modified first order approximation)**.** *For all $\theta \in (0,T)$ and $t \in (\theta/4, T]$, there holds*

$$\|u^\epsilon(\cdot,t) - u_1^\epsilon(\cdot,t)\|_{L^2(D)} \lesssim \Big(\epsilon^{\min\{1/2,2/d-1/2\}} t^{-\alpha} + \epsilon(t-\theta/4)^{-\alpha/2}\Big)\|a\|_{L^2(D)}.$$

Proof. The following identities hold,

$$\begin{cases} \partial_t^\alpha(u^\epsilon - u_1^\epsilon) = \nabla\cdot(\kappa^\epsilon \nabla(u^\epsilon - u_1^\epsilon)) + f_2 & \text{in } D, \quad t \in (\theta/4, T] \\ u^\epsilon - u_1^\epsilon = 0 & \text{on } \partial D,\ t \in (\theta/4, T] \\ (u^\epsilon - u_1^\epsilon)(\cdot,\theta/4) = -\epsilon(1-\zeta^\varepsilon)\chi_j(x/\epsilon)\dfrac{\partial u_0}{\partial x_j}(\cdot,\theta/4) & \text{in } D. \end{cases} \tag{2.23}$$

Here,

$$f_2(x,t) = \nabla\cdot(\kappa^\epsilon \nabla u_1^\epsilon) - \partial_t^\alpha u_1^\epsilon = \nabla\cdot(\kappa^\epsilon \nabla u_1^\epsilon) - \nabla\cdot(\kappa^* \nabla u_0) \\ + \nabla\cdot(\kappa^* \nabla u_0) - \partial_t^\alpha u_1^\epsilon \\ =: T_1(x,t) + T_2(x,t).$$

The definition (2.21), together with the *a priori* estimates (2.14), (2.16) and (2.18), implies

$$\|(u^\epsilon - u_1^\epsilon)(\cdot, \theta/4)\|_{H^{-1}(D)} \lesssim \epsilon \|a\|_{L^2(D)}.$$

We will prove that

$$\|f_2(\cdot,t)\|_{H^{-1}(D)} \lesssim \epsilon^{\min\{1/2, 2/d-1/2\}} t^{-\alpha} \|a\|_{L^2(D)} \text{ for all } t \in (0,T]. \tag{2.24}$$

Then by the linearity of the problem (2.23), we can derive from [40, Theorem 2.3] and [41, Theorem 2.1]

$$\begin{aligned}\|u^\epsilon(\cdot,t) - u_1^\epsilon(\cdot,t)\|_{L^2(D)} &\lesssim (t-\theta/4)^{-\alpha/2} \|(u^\epsilon - u_1^\epsilon)(\cdot,\theta/4)\|_{H^{-1}(D)} \\ &\quad + (t-\theta/4)^{\alpha/2} \|f_2(\cdot,t)\|_{L^\infty((\theta/4,T];H^{-1}(D))} \\ &\lesssim \Big(\epsilon^{\min\{1/2,2/d-1/2\}} t^{-\alpha} + \epsilon (t-\theta/4)^{-\alpha/2}\Big) \|a\|_{L^2(D)}\end{aligned}$$

for all $t \in (\theta/4, T]$. This completes the proof.

In the remaining of the proof, we will prove (2.24). To this end, we only need to estimate $\|T_1(\cdot,t)\|_{H^{-1}(D)}$ and $\|T_2(\cdot,t)\|_{H^{-1}(D)}$ for $t \in (0,T]$. We can further split the first term into

$$T_1(x,t) := \nabla \cdot (\kappa^\epsilon \nabla u_1^\epsilon) - \nabla \cdot (\kappa^* \nabla u_0) := T_{1,1}(x,t) + T_{1,2}(x,t)$$

with

$$\begin{aligned}T_{1,1}(x,t) &:= \nabla \cdot \left(\kappa^\epsilon \nabla \Big(u_0 + \epsilon \chi_j(\frac{x}{\epsilon}) \frac{\partial u_0}{\partial x_j} \Big) \right) - \nabla \cdot (\kappa^* \nabla u_0) \\ T_{1,2}(x,t) &:= -\epsilon \nabla \cdot \left(\kappa^\epsilon \nabla \Big(\zeta^\epsilon \chi_j(\frac{x}{\epsilon}) \frac{\partial u_0}{\partial x_j} \Big) \right).\end{aligned}$$

To estimate $T_{1,1}$, we apply the argument in [38, Section 1.4] and obtain

$$\|T_{1,1}(\cdot,t)\|_{H^{-1}(D)} \lesssim \epsilon t^{-\alpha}\|a\|_{L^2(D)} \text{ for any } t>0. \tag{2.25}$$

Then we estimate $T_{1,2}$. A direct calculation results in

$$\begin{aligned}\|T_{1,2}(\cdot,t)\|_{H^{-1}(D)} :=& \left\|-\epsilon\nabla\cdot\left(\kappa^\epsilon(x)\nabla\Big(\zeta^\epsilon\chi_j(\frac{x}{\epsilon})\frac{\partial u_0}{\partial x_j}(\cdot,t)\Big)\right)\right\|_{H^{-1}(D)}\\ \le&\epsilon\left\|\kappa^\epsilon\nabla\Big(\zeta^\epsilon\chi_j(\frac{x}{\epsilon})\frac{\partial u_0}{\partial x_j}(\cdot,t)\Big)\right\|_{L^2(D)}.\end{aligned}$$

Then by the triangle inequality and the chain rule, we deduce

$$\begin{aligned}\|T_{1,2}(\cdot,t)\|_{H^{-1}(D)} \lesssim \epsilon\|\nabla\zeta^\epsilon\|_{L^\infty(D)}\Big\|\chi_j(\frac{x}{\epsilon})\frac{\partial u_0}{\partial x_j}(\cdot,t)\Big\|_{L^2(D)} &+ \epsilon\Big\|\zeta^\epsilon\chi_j(\frac{x}{\epsilon})\nabla\frac{\partial u_0}{\partial x_j}(\cdot,t)\Big\|_{L^2(D)}\\ &+\Big\|\zeta^\epsilon\nabla\chi_j(\frac{x}{\epsilon})\frac{\partial u_0}{\partial x_j}(\cdot,t)\Big\|_{L^2(D)}.\end{aligned}$$

This, together with the generalized Hölder's inequality and Sobolev embedding, leads to

$$\begin{aligned}\|T_{1,2}(\cdot,t)\|_{H^{-1}(D)} \lesssim \epsilon^{1/2}\|\chi_j\|_{L^\infty(Y)}\left\|\frac{\partial u_0}{\partial x_j}(\cdot,t)\right\|_{L^2(D)} &+ \epsilon\,\|\zeta^\epsilon\|_{L^\infty(D)}\,\|\chi_j\|_{L^\infty(Y)}\,\|u_0(\cdot,t)\|_{H^2(D)}\\ &+\|\zeta^\epsilon\|_{L^q(D)}\,\|\nabla\chi_j\|_{L^r(Y)}\left\|\frac{\partial u_0}{\partial x_j}(\cdot,t)\right\|_{L^r(D)}\end{aligned}$$

where

$$\frac{1}{r}=\begin{cases}0, & \text{when } d=1\\ \delta, & \text{when } d=2\\ \frac{1}{2}-\frac{1}{d}, & \text{when } d=3\end{cases}\qquad\text{and}\qquad \frac{1}{q}=\begin{cases}\frac{1}{2}, & \text{when } d=1\\ \frac{1}{2}-2\delta, & \text{when } d=2\\ \frac{2}{d}-\frac{1}{2}, & \text{when } d=3.\end{cases}$$

Here, the parameter $\delta\in(0,1/2)$ is an arbitrary constant, and d is the dimension of the

domain D.

Together with the estimates (2.15), (2.19) and (2.14), by letting $\delta \to 0$, we obtain

$$\|T_{1,2}(\cdot,t)\|_{H^{-1}(D)} \lesssim \varepsilon^{\min\{1/2,2/d-1/2\}} t^{-\alpha} \|a\|_{L^2(D)}. \tag{2.26}$$

Next we estimate the second term T_2. The estimate (2.13) implies

$$\begin{aligned} T_2(x,t) &:= \nabla \cdot \big(\kappa^* \nabla u_0\big) - \partial_t^\alpha u_1^\epsilon = \partial_t^\alpha (u_0 - u_1^\epsilon) \\ &= \epsilon(\zeta^\epsilon - 1)\chi_j(\frac{x}{\epsilon})\partial_t^\alpha \frac{\partial u_0}{\partial x_j}. \end{aligned}$$

By exchanging the fractional derivative with respect to time t and the derivative with respect to the space variable x, we obtain

$$T_2 = \epsilon(\zeta^\epsilon - 1)\chi_j(\frac{x}{\epsilon})\frac{\partial}{\partial x_j}\partial_t^\alpha u_0.$$

Then by application of (2.14), we arrive at

$$\|T_2(\cdot,t)\|_{H^{-1}(D)} \lesssim \epsilon t^{-\alpha} \|a\|_{L^2(D)}.$$

This, together with (2.25), (2.26) and an application of the triangle inequality, proves the desired assertion (2.24). □

Remark 2.2.4. *We cannot derive from Lemma 2.2.3 the pointwise-in-time error estimate in $H^1(D)$ norm for the modified first order approximation due to the insufficient regularity of $a(x)$.*

Lemma 2.2.5. *For all $\theta \in (0,T]$, $f_1 \in L^\infty((0,T];H^{-1}(D))$. Moreover, there holds*

$$\|f_1\|_{L^\infty((0,T);H^{-1}(D))} \lesssim \epsilon^{\min\{1/2,2/d-1/2\}} \theta^{-2\alpha} \|a\|_{L^2(D)}.$$

Proof. Denote

$$f_1(x,t;\theta) \coloneqq \partial_t^\alpha \eta(t;\theta)(u^\epsilon - u_1^\epsilon) + \eta(t;\theta)\Big(- \nabla\cdot(\kappa^\epsilon \nabla u_1^\epsilon) + \nabla\cdot(\kappa^* \nabla u_0)\Big)$$
$$- \eta(t;\theta)\epsilon(\zeta^\epsilon - 1)\chi_j(\frac{x}{\epsilon})\partial_t^\alpha \frac{\partial u_0}{\partial x_j}$$
$$=: T_0(x,t;\theta) + \eta(t;\theta)T_1(x,t) - \eta(t;\theta)T_2(x,t).$$

To estimate $\|f_1(\cdot,t;\theta)\|_{H^{-1}(D)}$, we only need to estimate $\|T_0(\cdot,t;\theta)\|_{H^{-1}(D)}$, $\|T_1(\cdot,t)\|_{H^{-1}(D)}$ and $\|T_2(\cdot,t)\|_{H^{-1}(D)}$ for all $t \in (0,T]$. Firstly, we can obtain from estimates (2.14), (2.20), and Lemma 2.2.3,

$$\begin{aligned}\|T_0(\cdot,t;\theta)\|_{H^{-1}(D)} &\le |\partial_t^\alpha \eta(t;\theta)| \times \|u^\epsilon(\cdot,t) - u_1^\epsilon(\cdot,t)\|_{L^2(D)} \\ &\lesssim \theta^{-\alpha}\chi_{t>\theta/2}\epsilon^{\min\{1/2,2/d-1/2\}}t^{-\alpha}\|a\|_{L^2(D)} \\ &\lesssim \theta^{-2\alpha}\chi_{t>\theta/2}\epsilon^{\min\{1/2,2/d-1/2\}}\|a\|_{L^2(D)}. \end{aligned} \tag{2.27}$$

Next we can obtain from estimates (2.25) and (2.26),

$$\|T_1(\cdot,t)\|_{H^{-1}(D)} \lesssim \varepsilon^{\min\{1/2,2/d-1/2\}}t^{-\alpha}\|a\|_{L^2(D)} \text{ for all } t \in (0,T]. \tag{2.28}$$

Finally, we will estimate the last term T_2. By exchanging the fractional derivative with respect to time t and the derivative with respect to the space variable x, we obtain

$$T_2(x,t) = \epsilon(\zeta^\epsilon - 1)\chi_j(\frac{x}{\epsilon})\frac{\partial}{\partial x_j}\partial_t^\alpha u_0.$$

Then by application of (2.14), we arrive at

$$\|T_2(\cdot,t)\|_{H^{-1}(D)} \lesssim \epsilon t^{-\alpha}\|a\|_{L^2(D)} \text{ for all } t \in (0,T].$$

This, together with (2.27) and (2.28), and an application of the triangle inequality, implies

$$\|f_1(\cdot,t;\theta)\|_{H^{-1}(D)} \lesssim \epsilon^{\min\{1/2,2/d-1/2\}}\theta^{-2\alpha}\|a\|_{L^2(D)} \text{ for all } t \in (\theta/2,T]. \tag{2.29}$$

By noting that $f_1(\cdot,t;\theta) = 0$ for all $t \le \theta/2$, we have proved the desired assertion. □

Finally, we are ready to present the error estimate for the modified first order approximation defined in (2.21):

Theorem 2.2.6 (Error estimate of the modified first order approximation in $L^p((\theta,T);H^1(D))$ norm for $\theta \in (0,T)$ and $p \in [1,\infty)$)**.** *For any $p \in [1,\infty)$ and $\theta \in (0,T)$, let u^ϵ be the solution to Problem* (2.1) *and let the modified first order approximation u_1^ϵ be defined in* (2.21)*, then the following estimates hold*

$$\|u^\epsilon - u_1^\epsilon\|_{L^p((\theta,T];H^1(D))} + \|\partial_t^\alpha(u^\epsilon - u_1^\epsilon)\|_{L^p((\theta,T];H^{-1}(D))} \lesssim \epsilon^{\min\{1/2,2/d-1/2\}}\theta^{-2\alpha}\|a\|_{L^2(D)}. \tag{2.30}$$

Proof. First, we obtain from (2.29),

$$\|f_1(\cdot,t;\theta)\|_{H^{-1}(D)} \lesssim \epsilon^{\min\{1/2,2/d-1/2\}}\theta^{-2\alpha}\|a\|_{L^2(D)} \quad \text{for all } t \in (0,T].$$

Therefore, $f_1 \in L^p((0,T];H^{-1}(D))$ for any $p \ge 1$, and there holds

$$\|f_1\|_{L^p((0,T];H^{-1}(D))} \lesssim \epsilon^{\min\{1/2,2/d-1/2\}}\theta^{-2\alpha}\|a\|_{L^2(D)}.$$

Meanwhile, following the proof of [57, Theorems 2.1 and 2.2] and [40, Theorem 2.3], there exists a unique weak solution $R \in L^p([0,T];H_0^1(D))$ to Problem (2.22) such that

$\partial_t^\alpha R \in L^p([0,T];H^{-1}(D))$, satisfying

$$\|R\|_{L^p((0,T];H^1(D))} + \|\partial_t^\alpha R\|_{L^p((0,T];H^{-1}(D))} \lesssim \|f_1\|_{L^p((0,T];H^{-1}(D))}.$$

Noting that $R(x,t;\theta) = u^\epsilon - u_1^\epsilon$ for $t \geq \theta$, then a combination of the previous two estimates proves the desired assertion. □

We present in the next result the error estimate with the boundary layer effect:

Corollary 2.2.7 (Error estimate for the first order approximation)**.** *Let $a(x) \in L^2(D)$, $p \in [1,\infty)$, $\theta \in (0,T)$ and ϵ be sufficiently small. Let u^ϵ be the solution to Problem* (2.1) *and let the first order approximation U_1^ϵ be defined in* (2.21)*, then the following estimates hold*

$$\|u^\epsilon - U_1^\epsilon\|_{L^p((\theta,T);H^1(D))} + \left\|\partial_t^\alpha\left(u^\epsilon - U_1^\epsilon\right)\right\|_{L^p((\theta,T);H^{-1}(D))} \lesssim \epsilon^{\min\{1/2,2/d-1/2\}}\theta^{-2\alpha}\|a\|_{L^2(D)}$$
$$\|(u^\epsilon - U_1^\epsilon)(\cdot,t)\|_{L^2(D)} \lesssim \epsilon^{\min\{1/2,2/d-1/2\}}\theta^{-\alpha/2}\|a\|_{L^2(D)} \quad \textit{for all } t \in (\theta,T].$$

Proof. Note that we can derive from the definition (2.21) and estimate (2.14),

$$\|u_1^\epsilon - U_1^\epsilon\|_{L^p((\theta,T);H^1(D))} + \left\|\partial_t^\alpha\left(u_1^\epsilon - U_1^\epsilon\right)\right\|_{L^p((\theta,T);H^{-1}(D))} \lesssim \epsilon^{1/2}\theta^{-\alpha+1/p}\|a\|_{L^2(D)}$$
$$\|(u_1^\epsilon - U_1^\epsilon)(\cdot,t)\|_{L^2(D)} \lesssim \epsilon t^{-\alpha/2}\|a\|_{L^2(D)} \quad \text{for all } t \in (0,T].$$

Consequently, the desired assertion follows from Theorem 2.2.6 and Lemma 2.2.3 and an application of the triangle inequality. □

Remark 2.2.8. *When the initial data has a better regularity, e.g., $a(x) \in H_0^1(D) \cap H^2(D)$, there is no need to introduce a proper cut-off-function $\eta(t;\theta)$.*

Remark 2.2.9 (Comparison with homogenization results to parabolic equations with pe-

riodic coefficients). *Due to the boundary layer effect in the bounded domain, we can obtain from [55, 65] that the optimal convergence rate is* $\mathcal{O}(\epsilon^{1/2})$ *in* $L^p((\theta, T); H^1(D))$ *norm for* $p \in (1, \infty)$ *and* $\theta \in (0, T)$. *Corollary 2.2.7 shows that the convergence is* $\mathcal{O}(\epsilon^{\min\{1/2, 2/d-1/2\}})$. *One main restriction to apply similar technique employed in [65] to our current problem of a time-fractional diffusion problem* (2.1) *is the fact that there is no product rule for fractional derivative, but there is for the first derivative. For the same reason, one can not derive the pointwise over the time domain estimate in* $H^1(D)$ *norm for time-fractional diffusion problems given initial data* $a(x) \in L^2(D)$ *with insufficient regularity.*

2.3 Numerical experiments

In this section, we conduct a series of numerical experiments to demonstrate the performance of the first order corrector introduced in Section 2.2. Furthermore, we will validate the convergence result presented in Corollary 2.2.7 corresponding to different permeability fields and fractional order α.

Consider the time-fractional diffusion equation (2.1) in the unit square $D = [0, 1]^2$ with total time $T = 1$ and $\alpha := 0.9$. We will use scalar coefficient $\kappa^\epsilon(x_1, x_2)$ in the following numerical tests. The smooth initial data tested in Sections 2.3.1 and 2.3.2 is

$$a(x_1, x_2) := x_1(1 - x_1)x_2(1 - x_2).$$

We refer to Figure 2.1 for an illustration.

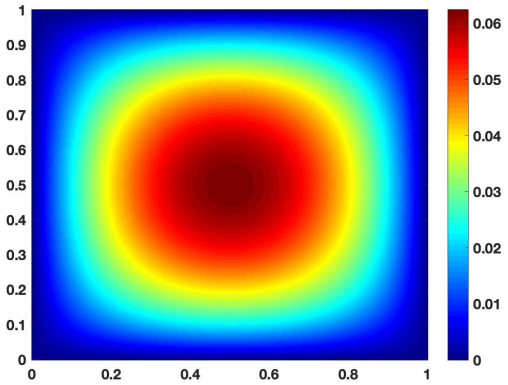

Figure 2.1: Initial data: $a(x_1, x_2)$

In Sections 2.3.1 and 2.3.2, we are concerned with the convergence rate of the first order approximation U_1^ϵ for diffusion coefficients $\kappa^\epsilon(x_1, x_2)$ of different regularities. To this end, we test two kinds of permeability fields, namely, the smooth and nonsmooth permeability fields in these two sections, respectively. Since Corollary 2.2.7 is also valid for rough initial data $a(x) \in L^2(D)$, we present the convergence history of the first order approximation U_1^ϵ with a rough initial data in Section 2.3.3.

2.3.1 Numerical tests with smooth permeability fields

To define the smooth permeability field, we take

$$\kappa(y_1, y_2) := 10 + \sin\Big(2\pi\{y_1\}\{y_2\}\big(1 - \{y_1\}\big)\big(1 - \{y_2\}\big)\Big) \tag{2.31}$$

as the periodic smooth function defined over the unit square Y. Here, $\{\cdot\}$ means taking the fractional part.

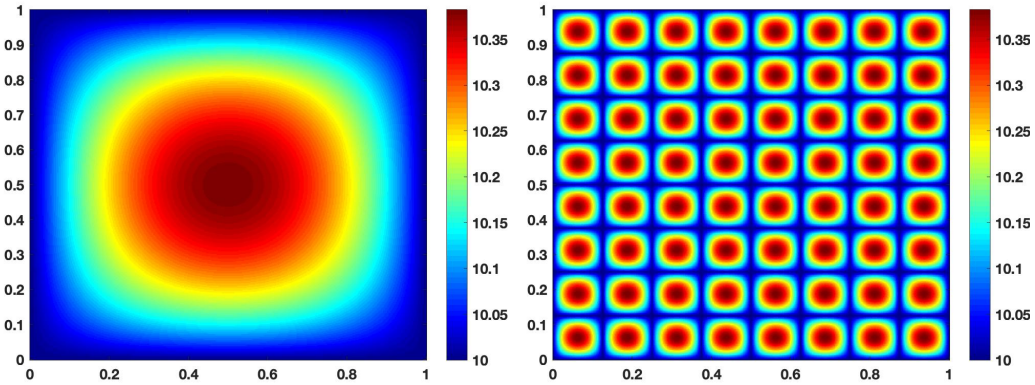

Figure 2.2: A smooth periodic permeability field with $\epsilon = \frac{1}{8}$ in a cell and over the domain D: $\kappa(y_1, y_2)$ and $\kappa^\epsilon(x_1, x_2)$.

Recall that $\kappa^\epsilon(x_1, x_2)$ is a periodic function with period ϵ. Its one cell after stretching over one unit cell is $\kappa(y_1, y_2)$, i.e., $\kappa^\epsilon(x_1, x_2) := \kappa(x_1/\epsilon, x_2/\epsilon)$. See Figure 2.2 for an illustration. Note that the contrast for this permeability is $\frac{11}{9}$. The main aim of this section is to investigate the convergence rate of the first order corrector U_1^ϵ. We will present the absolute and relative errors between U_1^ϵ and u^ϵ in L^2-norm and H^1-norm.

Let $\mathcal{T}_h$ be a decomposition of the domain D into non-overlapping shape-regular rectangular elements with maximal mesh size $h := 2^{-9}$. Let V_h be the conforming piecewise affine finite element space associated with the partition $\mathcal{T}_h$:

$$V_h := \{v \in C^0(D) : v|_T \in \mathcal{Q}_1(T) \text{ for all } T \in \mathcal{T}_h\} \cap H^1(D),$$

where $\mathcal{Q}_1(T)$ denotes the space of affine polynomials on each element $T \in \mathcal{T}_h$. Let $V_h^0 := V_h \cap H_0^1(D)$. We discretize the time interval $[0, 1]$ with a time step $\Delta t := 1/100$. Let $t_n = (n-1)\Delta t$ for $n = 1, 2, \cdots, N$ with $N := \Delta t^{-1} + 1$.

We adopt one popular scheme [49] to discretize the time variable t in (2.1) and (2.13), and apply the conforming Galerkin method to approximate the exact solution u^ϵ. We

denote its approximation at $t = t_{k+1}$ by $u_h^{\epsilon,k+1}$ for $k = 0, 1, \cdots, N-1$.

To this end, we seek for $u_h^{\epsilon,k+1} \in V_h^0$ for $k = 0, 1, \cdots, N-1$, satisfying

$$\forall\, v_h \in V_h^0 : \int_D u_h^{\epsilon,k+1} v_h \,\mathrm{d}x + \Gamma(2-\alpha)\Delta t^\alpha \int_D \kappa(x/\epsilon)\nabla u_h^{\epsilon,k+1} \cdot \nabla v_h \,\mathrm{d}x = \sum_{j=0}^{k-1}(b_j - b_{j+1}) \int_D u_h^{\epsilon,k-j} v_h \,\mathrm{d}x + b_k \int_D a(x) v_h \,\mathrm{d}x.$$

Here, the parameters b_j are given by

$$b_j := (j+1)^{1-\alpha} - j^{1-\alpha} \text{ for all } j = 0, 1, \cdots, N.$$

The numerical solutions $u_h^{\epsilon,k}$ for $k = 11, 51$ and 101 are depicted in Figure 2.3.

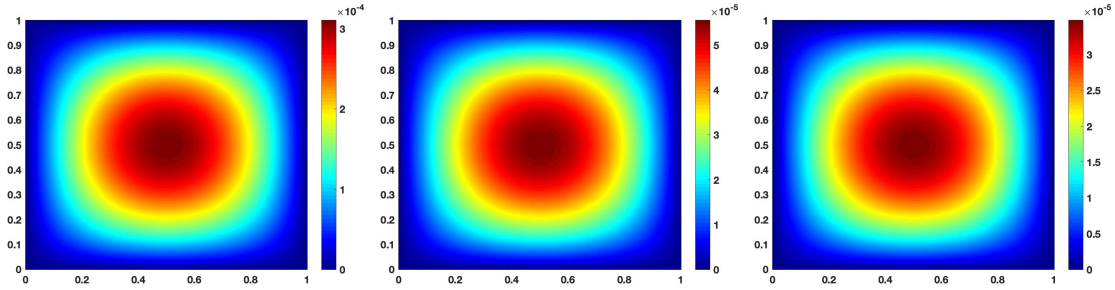

Figure 2.3: The fine scale solution $u_h^{\epsilon,k}$ for $k = 11, 51$ and 101 to Problem (2.1) with κ^ε in (2.31) and $\epsilon := \frac{1}{8}$.

We denote $U_{1,h}^{\epsilon,k}$ (or $u_{0,h}^k$) for $k = 1, 2\cdots, N$ as the numerical approximation to $U_1^\epsilon(t)$ (or $u_0(t)$) for $t = 0, \Delta t, \cdots, 1$. In order to obtain the first order approximation $U_{1,h}^{\epsilon,k}$, we first solve the cell problem (2.8). To this end, we first divide the computational domain Y into non-overlapping shape-regular rectangular elements with a maximal mesh size $h := 2^{-6}$. Then we solve the cell problem (2.8) with continuous piece-wise bilinear Lagrange

Finite Element Method using the conforming Galerkin formulation. We plot the two cell solutions χ_1 and χ_2 in Figure 2.4 with the smooth permeability field in Figure 2.2.

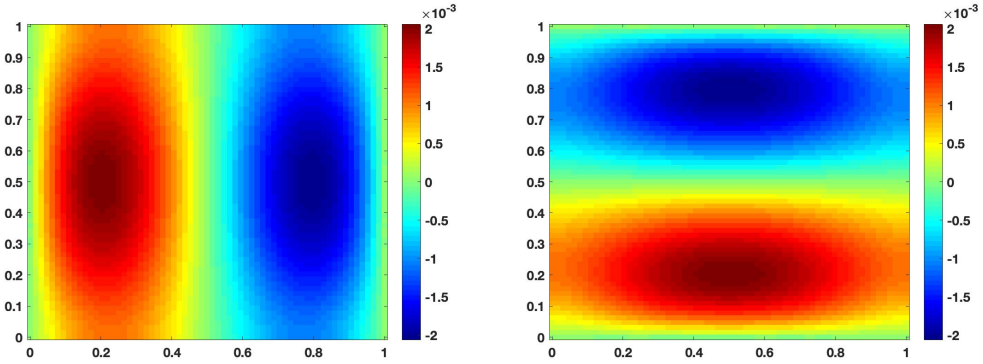

Figure 2.4: solutions to cell problem: $\chi_1(y_1, y_2)$ and $\chi_2(y_1, y_2)$

Utilizing the solutions to the cell problem (2.8), i.e., χ_1 and χ_2, we can obtain the effective coefficient κ^* from (2.12), and then solve for u_0 by (2.13). Note that there is no microscale oscillation in the effective coefficient κ^*, thus the solution u_0 can be solved in a much coarser mesh compared to the mesh $\mathcal{T}_h$ associated to the original problem (2.1). To simplify our notations, we adopt the same mesh $\mathcal{T}_h$ and finite element space V_h^0 as before, and utilize the conforming Galerkin formulation to solve for $u_{0,h}^{k+1}$ for $k = 0, 1, \cdots, N-1$.

We present the fine scale approximate solutions $u_{0,h}^k$ for $k = 11, 51$ and 101 in Figure 2.5.

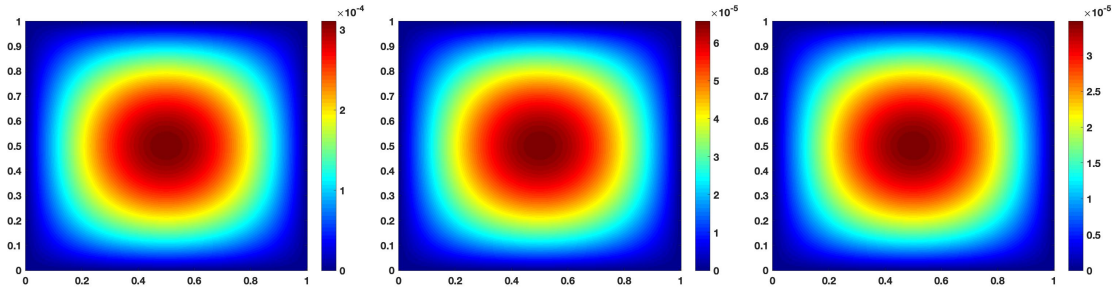

Figure 2.5: The homogenized solution $u_h^{0,k}$ for $k = 11, 51$ and 101 to Problem (2.1) with κ in (2.31) and $\epsilon := \frac{1}{8}$.

Finally, the first order approximation U_1^ϵ can be estimated using formula (2.21). We present the graphs of the approximate solutions $U_{1,h}^{\epsilon,k}$ for $k = 11, 51$ and 101 in Figure 2.6.

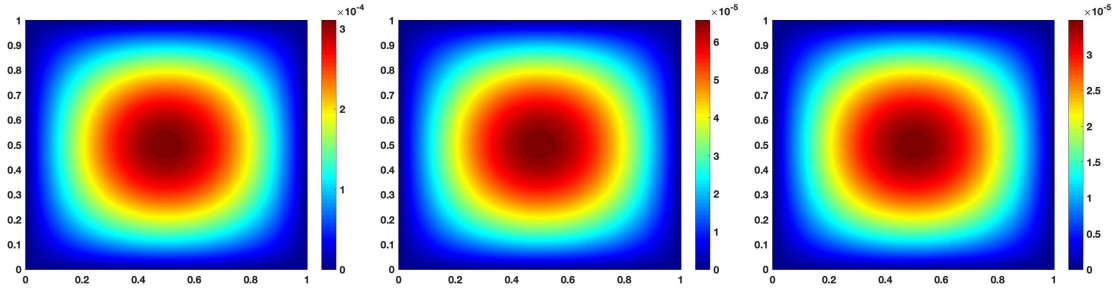

Figure 2.6: The first order approximation solution $U_{1,h}^{\epsilon,k}$ for $k = 11, 51$ and 101 to Problem (2.1) with κ in (2.31) and $\epsilon := \frac{1}{8}$.

We present the absolute error and relative error in the L^2 norm and H^1 norm in Table 2.1. The first column displays the discrete time steps at which we calculate the error. The next two columns display the absolute error and relative error in the L^2 norm between the fine-scale solution u_h^ϵ and the first order approximation $U_{1,h}^\epsilon$. The last two columns display the absolute error and relative error in the H^1 norm between the fine-scale solution u_h^ϵ and the first order approximation $U_{1,h}^\epsilon$.

t	$\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{L^2(D)}$	$\frac{\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{L^2(D)}}{\left\|u_h^\epsilon\right\|_{L^2(D)}}$	$\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{H^1(D)}$	$\frac{\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{H^1(D)}}{\left\|u_h^\epsilon\right\|_{H^1(D)}}$
0.1	2.1235e-9	1.3488e-5	1.4463e-7	2.0673e-4
0.5	4.5471e-10	1.3637e-5	3.0813e-8	2.0794e-4
1	2.411e-10	1.3656e-5	1.6325e-8	2.0809e-4

Table 2.1: The convergence history of the first order approximation to Problem (2.1) with κ in (2.31) and $\epsilon := \frac{1}{8}$.

Furthermore, one numerical experiment is conducted with larger variation in the coefficient κ compared to (2.31) to see the influence of the variation on the accuracy of homogenization. In this experiment, we take the same initial data $a(x_1, x_2)$ and the period $\epsilon := \frac{1}{8}$. We set

$$\kappa(y_1, y_2) := 10 + 9\sin\Big(2\pi\{y_1\}\{y_2\}\big(1 - \{y_1\}\big)\big(1 - \{y_2\}\big)\Big). \tag{2.32}$$

Note that the variation in this coefficient κ^ε is much larger than that defined in (2.31). The convergence history of the first order approximation with $\kappa(y_1, y_2)$ in (2.32) is presented in Table 2.2. Compared with the results in Table 2.1, a larger variation in the diffusion coefficient $\kappa(y_1, y_2)$ results in a larger error in the first order approximation.

t	$\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{L^2(D)}$	$\frac{\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{L^2(D)}}{\left\|u_h^\epsilon\right\|_{L^2(D)}}$	$\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{H^1(D)}$	$\frac{\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{H^1(D)}}{\left\|u_h^\epsilon\right\|_{H^1(D)}}$
0.1	1.6061e-8	1.16596e-4	1.0832e-6	1.7659e-3
0.5	3.4581e-9	1.1753e-4	2.3239e-7	1.7737e-3
1	1.8342e-9	1.1765e-4	1.2321e-7	1.7747 e-3

Table 2.2: The convergence history of the first order approximation to Problem (2.1) with κ in (2.32) and $\epsilon := \frac{1}{8}$.

In order to verify the convergence rate of the first approximation U_1^ϵ, we test two different values of the parameter ϵ being $\frac{1}{16}$ and $\frac{1}{32}$, respectively. Their corresponding convergence histories are shown in Tables 2.3 and 2.4.

t	$\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{L^2(D)}$	$\frac{\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{L^2(D)}}{\left\|u_h^\epsilon\right\|_{L^2(D)}}$	$\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{H^1(D)}$	$\frac{\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{H^1(D)}}{\left\|u_h^\epsilon\right\|_{H^1(D)}}$
0.1	5.2208e-10	3.3162e-6	7.1850e-8	1.0270e-4
0.5	1.1117e-10	3.3341e-6	1.5262e-8	1.0300e-4
1	5.8895e-11	3.3363e-6	8.0828e-9	1.0303e-4

Table 2.3: The convergence history of the first order approximation to Problem (2.1) with κ in (2.31) and $\epsilon := \frac{1}{16}$.

t	$\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{L^2(D)}$	$\frac{\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{L^2(D)}}{\left\|u_h^\epsilon\right\|_{L^2(D)}}$	$\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{H^1(D)}$	$\frac{\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{H^1(D)}}{\left\|u_h^\epsilon\right\|_{H^1(D)}}$
0.1	1.4377e-10	9.1322e-7	3.8067e-8	5.4412e-5
0.5	3.0547e-11	9.1614e-7	8.0778e-9	5.4515e-5
1	1.6179e-11	9.1649e-7	4.2776e-9	5.4527e-5

Table 2.4: The convergence history of the first order approximation to Problem (2.1) with κ in (2.31) and $\epsilon := \frac{1}{32}$.

One can calculate directly from Tables 2.1, 2.3 and 2.4 that the first order approximation maintains a convergence rate of $\mathcal{O}(\epsilon^{0.9623})$, $\mathcal{O}(\epsilon^{0.9657})$ and $\mathcal{O}(\epsilon^{0.9661})$ for $t = 0.1$, $t = 0.5$ and $t = 1$, respectively.

We also test different values of the parameter $\alpha \in (0, 1)$ and they all exhibit a similar convergence rate, as proved in Corollary 2.2.7. For the brevity of presentation, we will not present these results.

2.3.2 Numerical tests with non-smooth permeability fields

Even though the theoretical result presented in Corollary 2.2.7 is proved under the assumption that the diffusion coefficient $\kappa(y_1, y_2)$ is sufficiently smooth, we investigate in this section how well the first order approximation $U_1^\epsilon(x, t)$ performs when the coefficient $\kappa(y_1, y_2)$ is rough and admits large variation.

Firstly, we define two rough permeability fields $\kappa(x_1, x_2)$ with different variations. Let $U := [\frac{1}{5}, \frac{4}{5}]^2$ be a rectangle. Recall that $Y = [0, 1]^2$ is a unit square. We set the variation in the first permeability field to be 1.1, which is defined by

$$\kappa_1(x_1, x_2) := \begin{cases} 11, & \text{if } (x_1, x_2) \in U \\ 10, & \text{if } (x_1, x_2) \in Y \setminus U. \end{cases} \tag{2.33}$$

See Figure 2.7 for an illustration.

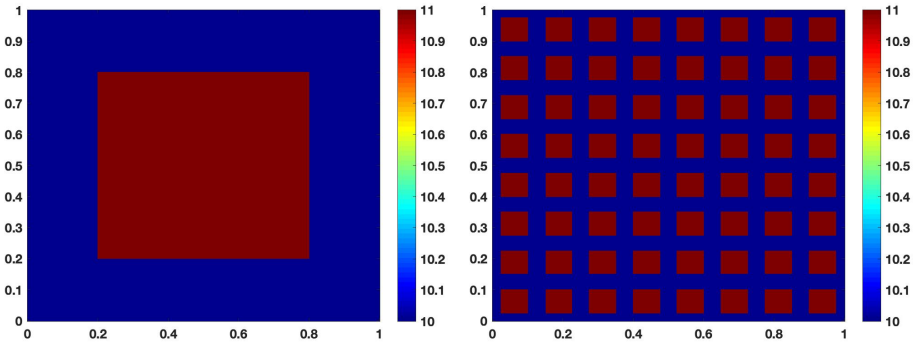

Figure 2.7: A nonsmooth permeability field of smaller variation: $\kappa_1(x_1, x_2)$ and $\kappa_1^\epsilon(x_1, x_2)$ with $\epsilon := \frac{1}{8}$.

We set the variation of the second rough permeability field to 2 and define

$$\kappa_2(x_1, x_2) := \begin{cases} 20, & \text{if } (x_1, x_2) \in U \\ 10, & \text{if } (x_1, x_2) \in Y \setminus U. \end{cases} \tag{2.34}$$

We present the graphs of such $\kappa(y_1, y_2)$ and $\kappa^\epsilon(x_1, x_2)$ in Figure 2.8.

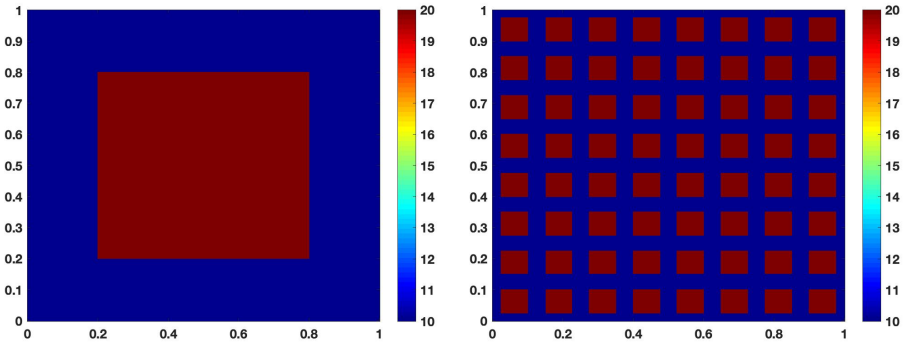

Figure 2.8: A nonsmooth permeability field of larger variation: $\kappa_2(x_1, x_2)$ and $\kappa_2^\epsilon(x_1, x_2)$ with $\epsilon = \frac{1}{8}$.

Let the parameter $\epsilon := 1/8$. We present the convergence histories of the first order approximation $U_1^\epsilon(x,t)$ with those two nonsmooth permeability fields in (2.33) and (2.34) in Tables 2.5 and 2.6. One can observe that the former outperforms the latter.

t	$\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{L^2(D)}$	$\frac{\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{L^2(D)}}{\left\|u_h^\epsilon\right\|_{L^2(D)}}$	$\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{H^1(D)}$	$\frac{\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{H^1(D)}}{\left\|u_h^\epsilon\right\|_{H^1(D)}}$
0.1	8.3877e-10	2.5129e-2	5.6728e-9	3.8002e-2
0.5	1.8973e-10	2.5130e-2	1.2832e-9	3.8003e-2
1	1.0121e-10	2.5130e-2	6.8450e-10	3.8003e-2

Table 2.5: The convergence history of the first order approximation to Problem (2.1) with $\kappa := \kappa_1$ in (2.33) and $\epsilon := \frac{1}{8}$.

t	$\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{L^2(D)}$	$\frac{\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{L^2(D)}}{\left\|u_h^\epsilon\right\|_{L^2(D)}}$	$\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{H^1(D)}$	$\frac{\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{H^1(D)}}{\left\|u_h^\epsilon\right\|_{H^1(D)}}$
0.1	5.5466e-9	1.6618e-1	3.6977e-8	2.4771e-1
0.5	1.2546e-9	1.6618e-1	8.3640e-9	2.4772e-1
1	6.6929e-10	1.6618e-1	4.4618e-9	2.4772e-1

Table 2.6: The convergence history of the first order approximation to Problem (2.1) with $\kappa := \kappa_2$ in (2.34) and $\epsilon := \frac{1}{8}$.

2.3.3 Numerical tests with rough initial data

Now we study the convergence rate of the first order approximation $U_1^\epsilon(x,t)$ when the initial data $a(x)$ is nonsmooth. To this end, we take a rough initial data $a(x_1,x_2)$ defined

by

$$a(x_1, x_2) := \begin{cases} 1, & \text{if } (x_1, x_2) \in (0.5, 1)^2 \\ 0, & \text{if } (x_1, x_2) \in (0, 1)^2 \setminus (0.5, 1)^2, \end{cases} \tag{2.35}$$

which is depicted in Figure 2.9.

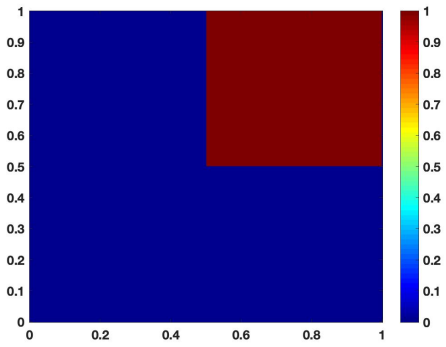

Figure 2.9: The initial data $a(x_1, x_2)$ as defined in (2.35).

To emphasize the effect of a rough initial data on the convergence rate of the first order approximation, we take the smooth permeability field $\kappa(x_1, x_2)$ as defined in (2.31). We adopt the same numerical scheme to calculate the numerical solutions as in Section 2.3.1. Like before, we test $\epsilon = \frac{1}{8}, \frac{1}{16}$ and $\frac{1}{32}$ to validate the convergence rate proved in Corollary 2.2.7. Their corresponding convergence histories of the first order approximation to Problem (2.1) are presented in Tables 2.7, 2.8 and 2.9.

t	$\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{L^2(D)}$	$\frac{\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{L^2(D)}}{\left\|u_h^\epsilon\right\|_{L^2(D)}}$	$\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{H^1(D)}$	$\frac{\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{H^1(D)}}{\left\|u_h^\epsilon\right\|_{H^1(D)}}$
0.1	2.0111e-7	1.8092e-4	5.1832e-6	8.3829e-4
0.5	4.4356e-8	1.8618e-4	1.1545e-6	8.6128e-4
1	2.3592e-8	1.8678e-4	6.1475e-7	8.6391e-4

Table 2.7: The convergence history of the first order approximation to Problem (2.1) with $a(x)$ defined in (2.35), κ in (2.31) and $\epsilon := \frac{1}{8}$.

Comparing the results presented in Table 2.7 with that in Table 2.1, one can observe that the latter admits a first order approximation with a much smaller error as expected, since a rough initial data produces singularity in the solution when the time t is small.

t	$\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{L^2(D)}$	$\frac{\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{L^2(D)}}{\left\|u_h^\epsilon\right\|_{L^2(D)}}$	$\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{H^1(D)}$	$\frac{\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{H^1(D)}}{\left\|u_h^\epsilon\right\|_{H^1(D)}}$
0.1	5.3061e-8	4.7739e-5	2.7039e-6	4.3744e-4
0.5	1.1742e-8	4.9292e-5	6.0373e-7	4.5053e-4
1	6.24783e-9	4.9471e-5	3.2156e-7	4.5204e-4

Table 2.8: The convergence history of the first order approximation to Problem (2.1) with $a(x)$ defined in (2.35), κ in (2.31) and $\epsilon := \frac{1}{16}$.

t	$\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{L^2(D)}$	$\frac{\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{L^2(D)}}{\left\|u_h^\epsilon\right\|_{L^2(D)}}$	$\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{H^1(D)}$	$\frac{\left\|u_h^\epsilon - U_{1,h}^\epsilon\right\|_{H^1(D)}}{\left\|u_h^\epsilon\right\|_{H^1(D)}}$
0.1	1.3154e-8	1.1835e-5	1.3839e-6	2.2390e-4
0.5	2.9144e-9	1.2235e-5	3.0918e-7	2.3074e-4
1	1.5510e-9	1.2281e-5	1.6469e-7	2.3153e-4

Table 2.9: The convergence history of the first order approximation to Problem (2.1) with $a(x)$ defined in (2.35), κ in (2.31) and $\epsilon := \frac{1}{32}$.

One can calculate from Tables 2.7, 2.8 and 2.9 that the first order approximation has a convergence rate of $\mathcal{O}(\epsilon^{0.9523})$, $\mathcal{O}(\epsilon^{0.9502})$ and $\mathcal{O}(\epsilon^{0.9499})$ for $t = 0.1$, $t = 0.5$ and $t = 1$, respectively.

3. CONVERGENCE OF THE CEM-GMSFEM FOR STOKES FLOWS IN HETEROGENEOUS PERFORATED DOMAINS *

In this chapter, we will develop and analyze a novel multiscale method for incompressible Stokes flows in perforated domains. Our idea is motivated by CEM-GMsFEM. It is based on the framework of GMsFEM to design multiscale basis functions such that the convergence of the method is independent of the contrast from the heterogeneities; and the error linearly decreases with respect to coarse mesh size if oversampling parameter is appropriately chosen. Our approach of solving velocity has two ingredients. Firstly, we construct auxiliary multiscale basis functions by solving a local eigenvalue problem on each coarse block. The global auxiliary space is formed by extending these auxiliary basis and the auxiliary space contains the information related to the pores. Secondly, the multiscale basis is sought in a weakly divergence free space by solving a minimization problem in an oversampling domain. The impose of weakly divergence free condition on the multiscale basis enables us solving velocity solitarily. We prove in Lemma 3.3.8 that the multiscale bases decay exponentially outside the corresponding local oversampling regions. This exponential decay property plays a vital role in the convergence analysis of the proposed method and justifies the use of local multiscale basis functions.

We organize this chapter as follows. In Section 3.1, we present the model problem and its variational formulation. In Section 3.2, we introduce auxiliary space and the construction of multiscale basis functions for pressure using relaxed constraint energy minimization. The multiscale basis functions are constructed by solving a class of local spectral problems and constrained minimization problems. We analyze convergence results in

*Reprinted with permission from "Convergence of the CEM-GMsFEM for Stokes flows in heterogeneous perforated domains" by Eric Chung, Jiuhua Hu and Sai-Mang Pun, 2020. Journal of Computational and Applied Mathematics, Volume 389, June 2021, 113327, Copyright [2020] by Elsevier.

Section 3.3.

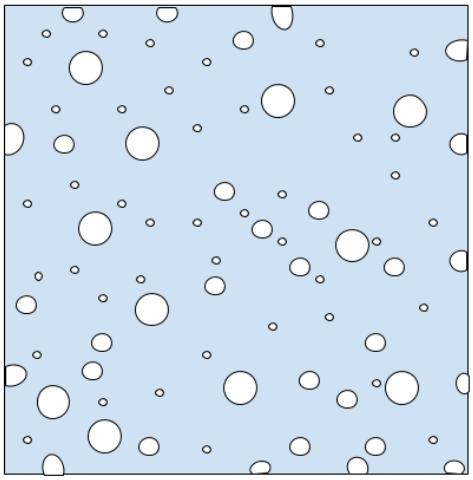

Figure 3.1: Illustration of a perforated domain

3.1 Problem setting

In this section, we start with stating the Stokes flow in heterogenous perforated domains. Some notations and function spaces are introduced. We also introduce its corresponding variational formulation.

3.1.1 Model problem

Let $\Omega \subseteq \mathbb{R}^d$ be a bounded domain and B_ϵ be a set of perforations within this domain. The set of perforations B_ϵ is assumed to be a union of connected circular disks. Each of these disks is of diameter of order $0 < \epsilon \ll \text{diam}(\Omega)$ and might be separated by a distance also of order ϵ. Here, the quantity $\text{diam}(S)$ denotes the diameter of a generic set S. In this work, we do not assume the set B_ϵ has periodic structure (see Figure 3.1 for an illustration). We denote $\Omega^\epsilon := \Omega \setminus \overline{B_\epsilon}$ the perforated domain. Then, we consider the basic linear model for incompressible fluid mechanics, i.e, Stokes equations. It consists of finding a vector

function $\mathbf{u}: \Omega^\epsilon \to \mathbb{R}^d$ and a scalar function $p: \Omega^\epsilon \to \mathbb{R}$ satisfying

$$\begin{aligned} -\mu\Delta\mathbf{u} + \nabla p &= \mathbf{f} && \text{in } \Omega^\epsilon, \\ \nabla \cdot \mathbf{u} &= 0 && \text{in } \Omega^\epsilon, \\ \mathbf{u} &= \mathbf{g} && \text{on } \partial\Omega^\epsilon \cap \partial\Omega, \\ \mathbf{u} &= \mathbf{0} && \text{on } \partial\Omega^\epsilon \cap B_\epsilon, \end{aligned} \tag{3.1}$$

where the vector filed $\mathbf{f}: \Omega^\epsilon \to \mathbb{R}^d$ is the body force acting on the fluid, $\mathbf{u}$ can be interpreted as the velocity of an incompressible fluid motion, p is the associated pressure, and the constant $\mu > 0$ is the viscosity coefficient of the fluid.

The boundary condition $\mathbf{g}$ defined on the outer boundary $\partial\Omega$ should satisfy the so-called compatibility condition $\int_{\partial\Omega^\epsilon \cap \partial\Omega} \mathbf{g} \cdot \mathbf{n}_{\partial\Omega^\epsilon \cap \partial\Omega}\, dS = 0$, where $\mathbf{n}_{\partial\Omega^\epsilon \cap \partial\Omega}$ is the unit outward normal vector field of the boundary $\partial\Omega^\epsilon \cap \partial\Omega$. In this work, we consider only the homogeneous Dirichlet boundary condition on ∂B_ϵ, where we assume that the perforation is impenetrable (see [53]). For the sake of simplicity, we consider homogeneous Dirichlet boundary for the velocity, i.e., $\mathbf{g} = \mathbf{0}$. The viscosity constant is assumed to be $\mu = 1$. The extension to the general viscosity constant is straightforward. See the remark in Section 3.1.3 for more details about inhomogeneous Dirichlet boundary condition. Other kinds of boundary conditions on the perforation need a completely new effort, which is out of the scope of this work. Since the pressure p is uniquely defined up to a constant, we assume that $\int_{\Omega^\epsilon} p\, dx = 0$ so that the problem has a unique solution. In this model, the primary source of the heterogeneity comes from the perforations in the computational domain; model reduction is necessary for practical simulation in this case.

3.1.2 Function spaces

In this subsection, we clarify the notations used throughout the article. We write $(\cdot, \cdot)$ to denote the inner product in $L^2(\Omega^\epsilon)$ and $\|\cdot\|$ for the corresponding norm. We denote

$L_0^2(\Omega^\epsilon)$ the subspace of $L^2(\Omega^\epsilon)$ containing functions with zero mean. Let $H^1(\Omega^\epsilon)$ be the classical Sobolev space with the norm $\|v\|_1 := \left(\|v\|^2 + \|\nabla v\|^2\right)^{1/2}$ for any $v \in H^1(\Omega^\epsilon)$ and $H_0^1(\Omega^\epsilon)$ the subspace of functions having a vanishing trace. For vector-valued functions, we denote $\mathbf{L}^2(\Omega^\epsilon) := (L^2(\Omega^\epsilon))^d$ and $\mathbf{H}_0^1(\Omega^\epsilon) := (H_0^1(\Omega^\epsilon))^d$. We write $\langle\cdot,\cdot\rangle$ to denote the inner product in $\mathbf{L}^2(\Omega^\epsilon)$. We also denote $\|\cdot\|$ the norm induced by the inner product $\langle\cdot,\cdot\rangle$. To shorten notations, we define the spaces for the velocity field $\mathbf{u}$ and the pressure p by

$$\mathbf{V}_0 := \mathbf{H}_0^1(\Omega^\epsilon) \quad \text{and} \quad Q_0 := L_0^2(\Omega^\epsilon).$$

3.1.3 Variational formulation and fine-grid discretization

In this subsection, we provide the variational formulation corresponding to the system (3.1). We multiply the first equation and the second one with test functions from $\mathbf{V}_0$ and Q_0, respectively. Then, applying Green's formula and making use of the boundary condition, the associated variational formulation of Stokes equation reads: Find $(\mathbf{u}, p) \in \mathbf{V}_0 \times Q_0$ such that

$$\begin{aligned} a(\mathbf{u},\mathbf{v}) - b(\mathbf{v},p) &= \langle \mathbf{f},\mathbf{v}\rangle \qquad \text{for all } \mathbf{v} \in \mathbf{V}_0, \\ b(\mathbf{u},q) &= 0 \qquad\qquad \text{for all } q \in Q_0, \end{aligned} \tag{3.2}$$

where

$$a(\mathbf{u},\mathbf{v}) := \int_{\Omega^\epsilon} \nabla\mathbf{u} : \nabla\mathbf{v}\, dx, \quad \text{and} \quad b(\mathbf{u},q) := \int_{\Omega^\epsilon} q\nabla\cdot\mathbf{u}\, dx.$$

The well-posedness of (3.2) can be proved (see, for example [28, Chapter 4]). Throughout this work, we denote $\|\cdot\|_a := \sqrt{a(\cdot,\cdot)}$ the energy norm.

Remark 3.1.1. *For the case of inhomogeneous Dirichlet boundary condition* $\mathbf{g} \neq \mathbf{0}$*, we*

may write

$$\mathbf{u} = \mathbf{u}_0 + \tilde{\mathbf{g}},$$

where $\tilde{\mathbf{g}} \in \mathbf{H}^1(\Omega^\epsilon)$ *satisfies* $\tilde{\mathbf{g}}|_{\partial\Omega^\epsilon \cap B_\epsilon} = \mathbf{0}$, $\tilde{\mathbf{g}}|_{\partial\Omega^\epsilon \cap \partial\Omega} = \mathbf{g}$ *and* $b(\tilde{\mathbf{g}}, q) = 0$ *for any* $q \in Q_0$. *Then, the pair of functions* $(\mathbf{u}_0, p)$ *satisfies*

$$\begin{aligned} a(\mathbf{u}_0, \mathbf{v}) - b(\mathbf{v}, p) &= \langle \mathbf{f}, \mathbf{v} \rangle - a(\tilde{\mathbf{g}}, \mathbf{v}) + \int_{\partial\Omega^\epsilon \cap \partial\Omega} (\nabla \mathbf{g}) \mathbf{n}_{\partial\Omega^\epsilon \cap \partial\Omega} \cdot \mathbf{v} \, dS \qquad \text{for all } \mathbf{v} \in \mathbf{V}_0, \\ b(\mathbf{u}_0, q) &= 0 \qquad \text{for all } q \in Q_0, \end{aligned}$$

Thus, we only consider the case with $\mathbf{g} = 0$ *in this work. For a wide class of incompressible or compressible flow models, similar reduction techniques can be applied (see [34, Section 2] for more details) to make the problem homogeneous in terms of Dirichlet boundary condition.*

To discretize the variational problem (3.2), let $\mathcal{T}^h$ be a conforming partition for the computational domain Ω^ϵ with grid size $h := \max_{\tau \in \mathcal{T}^h} h_\tau$ with $h_\tau := \text{diam}(\tau)$ for any fine-grid element $\tau \in \mathcal{T}^h$. We remark that $\mathcal{T}^h$ is referred to as the *fine grid*. Next, let $\mathbf{V}_h$ and Q_h be any conforming stable pair of finite element spaces with respect to the fine grid $\mathcal{T}^h$. For the coupling numerical scheme, one may use continuous Galerkin (CG) formulation: Find $(\mathbf{u}_h, p_h) \in \mathbf{V}_h \times Q_h$ such that

$$\begin{aligned} a(\mathbf{u}_h, \mathbf{v}_h) - b(\mathbf{v}_h, p_h) &= \langle \mathbf{f}, \mathbf{v}_h \rangle \qquad \text{for all } \mathbf{v}_h \in \mathbf{V}_h, \\ b(\mathbf{u}_h, q_h) &= 0 \qquad \text{for all } q_h \in Q_h. \end{aligned} \tag{3.3}$$

We remark that this classical approach will serve as a reference solution. The aim of this research is to construct a reduced system based on (3.3). To this end, we introduce finite-dimensional multiscale spaces $\mathbf{V}_{\text{ms}} \subseteq \mathbf{V}_0$ and $Q_{\text{ms}} \subseteq Q_0$, whose dimensions are much smaller, for approximating the solution on some feasible coarse grid.

3.2 Construction of multiscale spaces

In this section, we construct multiscale spaces on a coarse grid. Let $\mathcal{T}^H$ be a conforming partition of the computational domain Ω^ϵ such that $\mathcal{T}^h$ is a refinement of $\mathcal{T}^H$. We call $\mathcal{T}^H$ the *coarse grid* and each element of $\mathcal{T}^H$ a coarse block. We denote $H := \max_{K\in\mathcal{T}^H}\text{diam}(K)$ the coarse grid size. Let N_c be the total number of (interior) vertices of $\mathcal{T}^H$ and N be the total number of coarse elements. We remark that the coarse element $K \in \mathcal{T}^H$ is a closed subset (of the domain Ω^ϵ) with nonempty interior and piecewise smooth boundary. Let $\{x_i\}_{i=1}^{N_c}$ be the set of nodes in $\mathcal{T}^H$. Figure 3.2 illustrates the fine grid and a coarse element K_i.

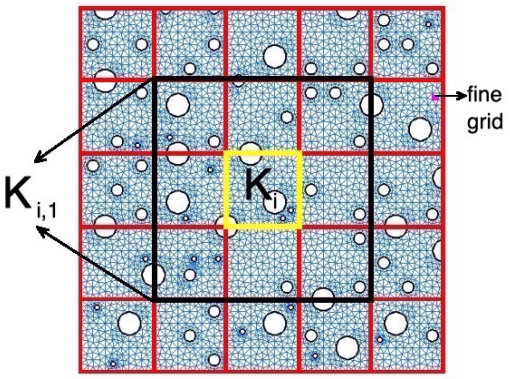

Figure 3.2: Illustration of the coarse grid, the fine grid, and the oversampling domain.

The construction of the multiscale spaces consists of two steps. The first step is to construct auxiliary multiscale spaces using the concept of GMsFEM. Based on the auxiliary spaces, we can then construct multiscale spaces containing basis functions whose energy

are minimized in some subregions of the domain. These energy-minimized basis functions will be shown to decay exponentially outside the oversampling domain, and can be used to construct a multiscale solution.

3.2.1 Auxiliary space

In this section, we begin with the construction of the auxiliary multiscale basis functions. Let $\mathbf{V}(S)$ be the restriction of $\mathbf{V}_0$ on $S \subset \Omega^\epsilon$ and $\mathbf{V}_0(S)$ be the subspace of $\mathbf{V}(S)$, whose element is of zero trace on ∂S. We also define $Q_0(S) := L_0^2(S)$. Consider the following local spectral problem: Find $(\phi_j^i, \lambda_j^i) \in \mathbf{V}(K_i) \times \mathbb{R}$ such that

$$a_i(\phi_j^i, \mathbf{v}) = \lambda_j^i s_i(\phi_j^i, \mathbf{v}) \quad \text{for all } \mathbf{v} \in \mathbf{V}(K_i), \tag{3.4}$$

where $a_i(\cdot,\cdot)$ and $s_i(\cdot,\cdot)$ are defined as follows:

$$a_i(\mathbf{u}, \mathbf{v}) := \int_{K_i} \nabla \mathbf{u} : \nabla \mathbf{v} \, dx \quad \text{and} \quad s_i(\mathbf{u}, \mathbf{v}) := \int_{K_i} \tilde{\kappa} \mathbf{u} \cdot \mathbf{v} \, dx \tag{3.5}$$

for any $\mathbf{u}, \mathbf{v} \in \mathbf{V}(K_i)$. Here, we define $\tilde{\kappa} := \sum_{j=1}^{N_c} |\nabla \chi_j^{\text{ms}}|^2$, where $\{\chi_j^{\text{ms}}\}_{j=1}^{N_c}$ is a set of neighborhood-wise defined partition of unity functions [7] on the coarse grid. In particular, the function χ_j^{ms} satisfies $H|\nabla \chi_j^{\text{ms}}| = O(1)$ and $0 \le \chi_j^{\text{ms}} \le 1$.

Assume that the eigenvalues are arranged in ascending order such that

$$0 \le \lambda_1^i \le \cdots \le \lambda_{\ell_i}^i \le \cdots$$

for each $i \in \{1, \cdots, N\}$. Also, we assume that the eigenfunctions satisfy the normalization condition $s_i(\phi_j^i, \phi_j^i) = 1$. Then, we choose the first $\ell_i \in \mathbb{N}^+$ eigenfunctions and define $V_{\text{aux}}^i := \text{span}\{\phi_j^i : j = 1, \cdots, \ell_i\}$. Based on these local spaces, the global auxiliary space V_{aux}

is defined to be

$$V_{\text{aux}} := \bigoplus_{i=1}^{N} V_{\text{aux}}^{i} \quad \text{with inner product} \quad s(\mathbf{u},\mathbf{v}) := \sum_{i=1}^{N} s_i(\mathbf{u},\mathbf{v})$$

for any $\mathbf{u},\mathbf{v} \in V_{\text{aux}}$. Further, we define an orthogonal projection $\pi : \mathbf{V}_0 \to V_{\text{aux}}$ such that

$$\pi(\mathbf{v}) := \sum_{i=1}^{N} \pi_i(\mathbf{v}), \quad \text{where } \pi_i(\mathbf{v}) := \sum_{j=1}^{\ell_i} s_i(\mathbf{v},\phi_j^i)\phi_j^i$$

for all $\mathbf{v} \in \mathbf{V}_0$.

3.2.2 Multiscale space

In this section, we construct multiscale basis functions based on constraint energy minimization. For each coarse element K_i, we define the oversampled region $K_{i,k_i} \subseteq \Omega^\epsilon$ by enlarging K_i by $k_i \in \mathbb{N}$ layer(s), i.e.,

$$K_{i,0} := K_i, \quad K_{i,k_i} := \bigcup \{K \in \mathcal{T}^H : K \cap K_{i,k_i-1} \neq \varnothing\} \quad \text{for } k_i = 1,2,\cdots.$$

We call k_i a parameter of oversampling related to the coarse element K_i. See Figure 3.2 for an illustration of $K_{i,1}$. For simplicity, we denote K_i^+ a generic oversampling region related to the coarse element K_i with a specific oversampling parameter k_i. Next, we define multiscale basis function possessing the property of constraint energy minimization [15]. In particular, for each auxiliary function $\phi_j^i \in V_{\text{aux}}$, we solve the following minimization problem: Find $\psi_{j,\text{ms}}^i \in \mathbf{V}_0(K_i^+)$ such that

$$\psi_{j,\text{ms}}^i := \operatorname{argmin}\left\{a(\psi,\psi) + s\left(\pi(\psi)-\phi_j^i, \pi(\psi)-\phi_j^i\right) : \psi \in \mathbf{V}_0(K_i^+) \text{ and } \nabla\cdot\psi = 0\right\}. \quad (3.6)$$

Note that problem (3.6) is equivalent to the local problem: Find $(\psi^i_{j,\text{ms}}, \xi^i_{j,\text{ms}}) \in \mathbf{V}_0(K_i^+) \times Q_0(K_i^+)$ such that

$$\begin{aligned} a(\psi^i_{j,\text{ms}}, v) + s\left(\pi(\psi^i_{j,\text{ms}}), \pi(v)\right) + b(v, \xi^i_{j,\text{ms}}) &= s\left(\phi^i_j, \pi(v)\right) \qquad \text{for all } v \in \mathbf{V}_0(K_i^+), \\ b(\psi^i_{j,\text{ms}}, q) &= 0 \qquad \text{for all } q \in Q_0(K_i^+). \end{aligned} \tag{3.7}$$

Finally, for fixed parameters k_i and ℓ_i, the multiscale space V_{ms} is defined by

$$V_{\text{ms}} := \text{span}\left\{\psi^i_{j,\text{ms}} : 1 \le j \le \ell_i,\ 1 \le i \le N\right\}.$$

The multiscale basis functions can be interpreted as approximations to global multiscale basis functions $\psi^i_j \in \mathbf{V}_0$ defined by

$$\psi^i_j := \text{argmin}\left\{a(\psi,\psi) + s\left(\pi(\psi) - \phi^i_j, \pi(\psi) - \phi^i_j\right) :\ \psi \in \mathbf{V}_0 \text{ and } \nabla\cdot\psi = 0\right\},$$

which is equivalent to the following variational formulation: Find $(\psi^i_j, \xi^i_j) \in \mathbf{V}_0 \times Q_0$ such that

$$\begin{aligned} a(\psi^i_j, v) + s\left(\pi(\psi^i_j), \pi(v)\right) + b(v, \xi^i_j) &= s\left(\phi^i_j, \pi(v)\right) \qquad \text{for all } v \in \mathbf{V}_0, \\ b(\psi^i_j, q) &= 0 \qquad \text{for all } q \in Q_0. \end{aligned} \tag{3.8}$$

These basis functions have global support in the domain Ω^ϵ, but, as shown in Lemma 3.3.8, decay exponentially outside some local (oversampled) region. This property plays a vital role in the convergence analysis of the proposed method and justifies the use of local basis functions in V_{ms}. Furthermore, we define $V_{\text{glo}} := \text{span}\left\{\psi^i_j : 1 \le j \le \ell_i, 1 \le i \le N\right\}$ and $\tilde{V} := \{\mathbf{v} \in \mathbf{V}_0^{\text{div}} : \pi(\mathbf{v}) = 0\}$, where $\mathbf{V}_0^{\text{div}}$ is the closed subspace of $\mathbf{V}_0$ containing divergence-free vector fields. Then, one can show that $\mathbf{V}_0^{\text{div}} = V_{\text{glo}} \bigoplus_a \tilde{V}$.

Remark 3.2.1. *Suppose that $S \subset \Omega^\epsilon$ is any non-empty connected union of coarse elements*

$K_i \in \mathcal{T}^H$. *Denote* $\mathcal{D}_S : \mathbf{H}_0^1(S) \to L_0^2(S)$ *the divergence operator corresponding to the set* S. *We have the following auxiliary result from functional analysis.*

Lemma 3.2.2 (cf. Theorem 6.14-1 in [20])**.** *Suppose that* S *is any non-empty connected union of coarse elements. Restricting the domain of* $\mathcal{D}_S$ *on the orthogonal complement (with respect to standard* L^2 *inner product) of its kernel, the divergence operator* $\mathcal{D}_S$ *is injective and surjective. Moreover, it has a continuous inverse and there is a generic constant* $\beta_S > 0$ *such that*

$$\beta_S \left\| \mathcal{D}_S^{-1} \mu \right\|_a \leq \|\mu\| \quad \textit{for any } \mu \in L_0^2(S).$$

Using the result of Lemma 3.2.2, one can show that (3.7) and (3.8) are well-posed. Let S be the whole domain Ω^ϵ or an oversampled region K_i^+. Then, for any non-zero element $v \in \mathbf{H}_0^1(S)$, we have

$$\sup_{v \in \mathbf{H}_0^1(S), v \neq 0} \frac{|b(v,\mu)|}{\|v\|_a} \geq \frac{|b(\mathcal{D}_S^{-1}\mu, \mu)|}{\|\mathcal{D}_S^{-1}\mu\|_a} = \frac{\|\mu\|^2}{\|\mathcal{D}_S^{-1}\mu\|_a} \geq \beta_S \|\mu\| \tag{3.9}$$

for any $\mu \in L_0^2(S)$, which shows that the inf-sup condition holds for (3.8). Similarly, we can prove the inf-sup condition holds for (3.7).

3.2.3 The multiscale method

From the above, we have the multiscale space V_{ms} for the approximation of velocity field. The multiscale solution $\mathbf{u}_{\text{ms}} \in V_{\text{ms}}$ is obtained by solving the following equation:

$$a(\mathbf{u}_{\text{ms}}, \mathbf{v}) = \langle \mathbf{f}, \mathbf{v} \rangle \quad \text{for all } \mathbf{v} \in V_{\text{ms}}. \tag{3.10}$$

To approximate the pressure based on coarse grid, we will construct a specific solution space of finite dimension.

Let $W(K_i) := \{v \in H^1(K_i) : \int_{K_i} v\ dx = 0,\ b(\mathbf{w}, v) = 0 \text{ for all } \mathbf{w} \in (I - \pi)\mathbf{V}_0\}$. We consider the following spectral problem: Find $(q_j^i, \zeta_j^i) \in W(K_i) \times \mathbb{R}$ such that

$$\mathcal{A}_i(q_j^i, v) = \zeta_j^i \mathcal{S}_i(q_j^i, v) \quad \text{for all } v \in W(K_i), \tag{3.11}$$

where $\mathcal{A}_i(\cdot,\cdot)$ and $\mathcal{S}_i(\cdot,\cdot)$ are defined as follows:

$$\mathcal{A}_i(u, v) := \int_{K_i} \nabla u \cdot \nabla v\ dx \quad \text{and} \quad \mathcal{S}_i(u, v) := \int_{K_i} \tilde{\kappa} u v\ dx \tag{3.12}$$

for any $u, v \in H^1(K_i)$. Assume that for each $i \in \{1, \cdots, N\}$ the eigenvalues ζ_j^i are arranged in ascending order such that $0 \leq \zeta_1^i \leq \zeta_2^i \leq \cdots$. We then define a finite dimensional solution space Q_H as follows:

$$Q_H := \text{span}\left\{q_j^i :\ i = 1, \cdots, N,\ j = 1, \cdots, \ell_i\right\}.$$

Then, we solve the following variational problem over the domain Ω^ϵ: Find $p_{\text{ms}} \in Q_H$ such that

$$b(\mathbf{v}, p_{\text{ms}}) = a(\mathbf{u}_{\text{ms}}, \mathbf{v}) - \langle \mathbf{f}, \mathbf{v} \rangle \quad \text{for all } \mathbf{v} \in V_{\text{aux}}. \tag{3.13}$$

Note that $\dim(Q_H) = \dim(V_{\text{aux}})$. To prove the well-posedness of (3.13), it suffices to verify inf-sup condition for the bilinear form $b(\cdot,\cdot)$ over V_{aux} and Q_H. Recall that the variational formulation (3.2) is well-posed and inf-sup condition holds for $b(\cdot,\cdot)$ under spaces $\mathbf{V}_0$ and Q_0. Hence, for any $q \in Q_H$, there exists $\mathbf{w} \in \mathbf{V}_0$ such that $b(\mathbf{w}, q) \geq C \|\mathbf{w}\|_a \|q\|$ for some constant $C > 0$. Choosing $\mathbf{v} := \pi\mathbf{w}$, we have $\mathbf{v} \in V_{\text{aux}}$ and

$$b(\mathbf{v}, q) = b(\pi\mathbf{w}, q) = b(\mathbf{w}, q) \geq C \|\mathbf{w}\|_a \|q\| \geq C \|\mathbf{v}\|_a \|q\|.$$

Therefore, the problem (3.13) is well-posed. Note that the pressure p solves the following equation:

$$b(\mathbf{v}, p) = a(\mathbf{u}, \mathbf{v}) - \langle \mathbf{f}, \mathbf{v} \rangle \qquad \text{for all } \mathbf{v} \in \mathbf{V}_0.$$

Then, we have

$$b(\mathbf{v}, p - p_{\text{ms}}) = a(\mathbf{u} - \mathbf{u}_{\text{ms}}, \mathbf{v}) \leq \|\mathbf{u} - \mathbf{u}_{\text{ms}}\|_a \|\mathbf{v}\|_a ,$$

for all $\mathbf{v} \in V_{\text{aux}}$. It implies that

$$\sup_{\mathbf{v} \in V_{\text{aux}}} \frac{b(\mathbf{v}, p - p_{\text{ms}})}{\|\mathbf{v}\|_a} \leq \|\mathbf{u} - \mathbf{u}_{\text{ms}}\|_a .$$

The multiscale solution p_{ms} serves as an approximation of the solution p and $\|p - p_{\text{ms}}\| \lesssim \|\mathbf{u} - \mathbf{u}_{\text{ms}}\|_a$.

3.3 Convergence analysis

In this section, we analyze the proposed method. We denote $\|\cdot\|_s := \sqrt{s(\cdot, \cdot)}$ the s-norm. In particular, $\|\mathbf{v}\|_s^2 = \sum_{i=1}^d \|v_i\|_s^2$ for any $\mathbf{v} = (v_1, \cdots, v_d)^T$. We also denote $\text{spt}(v)$ the support of a given function or vector field. We write $a \lesssim b$ if there exists a generic constant $C > 0$ such that $a \leq Cb$. Define $\Lambda := \min_{1 \leq i \leq N} \lambda^i_{\ell_i+1}$ and $\Gamma := \max_{1 \leq i \leq N} \lambda^i_{\ell_i}$. For a given subregion $S \subset \Omega^\epsilon$, we define local norms $\|\mathbf{v}\|_{a(S)} := \left(\int_S |\nabla \mathbf{v}|^2 \, dx\right)^{1/2}$ and $\|\mathbf{v}\|_{s(S)} := \left(\int_S \tilde{\kappa} |\mathbf{v}|^2 \, dx\right)^{1/2}$ for any $\mathbf{v} \in \mathbf{V}_0$.

Before estimating the error between global and local multiscale basis functions, we introduce some notions that will be used in the analysis. First, we introduce cutoff function with respect to oversampling region. Given a coarse block $K_i \in \mathcal{T}^H$ and a parameter of oversampling $m \in \mathbb{N}$, we recall that $K_{i,m} \subset \Omega^\epsilon$ is an m-layer oversampling region corresponding to K_i.

Definition 3.3.1. *For two positive integers M and m with $M > m \geq 1$, we define cutoff function $\chi_i^{M,m} \in span\{\chi_j^{ms}\}_{j=1}^{N_c}$ such that $0 \leq \chi_i^{M,m} \leq 1$ and*

$$\chi_i^{M,m} = \begin{cases} 1 & \text{in } K_{i,m}, \\ 0 & \text{in } \Omega^\epsilon \setminus K_{i,M}. \end{cases}$$

Note that, we have $K_{i,m} \subset K_{i,M} \subset \Omega^\epsilon$ and $spt(\chi_i^{M,m}) \subset K_{i,M}$.

First, we establish the following auxiliary results for later use in the analysis.

Lemma 3.3.2. *Let $\mathbf{v} \in \mathbf{V}_0$ and $k \geq 2$ be an integer. Then, the following inequalities hold:*

(i) $\|\mathbf{v}\|_a \leq \Gamma^{1/2} \|\mathbf{v}\|_s$ *if* $\mathbf{v} \in V_{aux}$;

(ii) $\|\mathbf{v}\|_s \leq \Lambda^{-1/2} \|\mathbf{v}\|_a$ *if* $\mathbf{v} \notin V_{aux}$;

(iii) $\|\mathbf{v}\|_s^2 \lesssim \Lambda^{-1} \|(I-\pi)\mathbf{v}\|_a^2 + \|\pi\mathbf{v}\|_s^2$;

(iv) $\left\|(1-\chi_i^{k,k-1})\mathbf{v}\right\|_a^2 \leq 2(1+\Lambda^{-1}) \|\mathbf{v}\|_{a(\Omega^\epsilon \setminus K_{i,k-1})}^2 + 2\|\pi\mathbf{v}\|_{s(\Omega^\epsilon \setminus K_{i,k-1})}^2$

(v) $\left\|(1-\chi_i^{k,k-1})\mathbf{v}\right\|_s^2 \leq \Lambda^{-1} \|\mathbf{v}\|_{a(\Omega^\epsilon \setminus K_{i,k-1})}^2 + \|\pi\mathbf{v}\|_{s(\Omega^\epsilon \setminus K_{i,k-1})}^2$.

Proof. Note that one can write $\mathbf{v} = \sum_{i=1}^N \sum_{j\geq 1} \alpha_j^i \phi_j^i$ with $\alpha_j^i \in \mathbb{R}$ for any $\mathbf{v} \in \mathbf{V}_0$.

(i) Since $\mathbf{v} \in V_{\text{aux}}$, then $\alpha_j^i = 0$ for $j \geq \ell_i + 1$. Using the local spectral problem (3.4), we obtain

$$\|\mathbf{v}\|_a^2 = \sum_{i=1}^N \sum_{j=1}^{\ell_i} \alpha_j^i a(\phi_j^i, \mathbf{v}) = \sum_{i=1}^N \sum_{j=1}^{\ell_i} \alpha_j^i \lambda_j^i s(\phi_j^i, \mathbf{v}) \leq \Gamma \sum_{i=1}^N \sum_{j=1}^{\ell_i} \alpha_j^i s(\phi_j^i, \mathbf{v}) = \Gamma \|\mathbf{v}\|_s^2.$$

(ii) For any $\mathbf{v} \notin V_{\text{aux}}$, one can write $\mathbf{v} = \sum_{i=1}^N \sum_{j\geq \ell_i+1} \alpha_j^i \phi_j^i$. Then, we have

$$\|\mathbf{v}\|_a^2 = \sum_{i=1}^N \sum_{j\geq \ell_i+1} \alpha_j^i a(\phi_j^i, \mathbf{v}) = \sum_{i=1}^N \sum_{j\geq \ell_i+1} \alpha_j^i \lambda_j^i s(\phi_j^i, \mathbf{v}) \geq \Lambda \sum_{i=1}^N \sum_{j\geq \ell_i+1} \alpha_j^i s(\phi_j^i, \mathbf{v}) = \Lambda \|\mathbf{v}\|_s^2.$$

(iii) The result follows from (i), (ii), and the triangle inequality.

(iv) By using the property of cutoff function $\chi_i^{k,k-1}$ and (iii), we have

$$\begin{aligned}\left\|(1-\chi_i^{k,k-1})\mathbf{v}\right\|_a^2 &\leq 2\left(\int_{\Omega^\epsilon\setminus K_{i,k-1}}(1-\chi_i^{k,k-1})^2|\nabla\mathbf{v}|^2+|\mathbf{v}\nabla\chi_i^{k,k-1}|^2dx\right)\\ &\leq 2\left(\|\mathbf{v}\|^2_{a(\Omega^\epsilon\setminus K_{i,k-1})}+\|\mathbf{v}\|^2_{s(\Omega^\epsilon\setminus K_{i,k-1})}\right)\\ &\leq 2(1+\Lambda^{-1})\|\mathbf{v}\|^2_{a(\Omega^\epsilon\setminus K_{i,k-1})}+2\|\pi\mathbf{v}\|^2_{s(\Omega^\epsilon\setminus K_{i,k-1})}.\end{aligned}$$

(v) For any $k\geq 2$, we have

$$\begin{aligned}\left\|(1-\chi_i^{k,k-1})\mathbf{v}\right\|_s^2 &\leq \|\mathbf{v}\|^2_{s(\Omega^\epsilon\setminus K_{i,k-1})}\\ &\leq \Lambda^{-1}\|\mathbf{v}\|^2_{a(\Omega^\epsilon\setminus K_{i,k-1})}+\|\pi\mathbf{v}\|^2_{s(\Omega^\epsilon\setminus K_{i,k-1})}.\end{aligned}$$

This completes the proof. □

First, we present the convergence of using global basis functions constructed in (3.8). We define $\mathbf{u}_{\text{glo}}\in V_{\text{glo}}$ as the global multiscale solution satisfying

$$a(\mathbf{u}_{\text{glo}},\mathbf{v})=\langle\mathbf{f},\mathbf{v}\rangle\quad\text{for all }\mathbf{v}\in V_{\text{glo}}. \tag{3.14}$$

Theorem 3.3.3. *Let* $\mathbf{u}$ *be the solution of* (3.2) *and* $\mathbf{u}_{glo}$ *be the solution of* (3.14)*. We have*

$$\|\mathbf{u}-\mathbf{u}_{glo}\|_a\lesssim\Lambda^{-1/2}\left\|\tilde{\kappa}^{-1/2}\mathbf{f}\right\|.$$

Moreover, if $\{\chi_j^{ms}\}_{j=1}^{N_c}$ *is a set of bilinear partition of unity, we have*

$$\|\mathbf{u}-\mathbf{u}_{glo}\|_a\lesssim H\Lambda^{-1}\|\mathbf{f}\|.$$

Proof. Note that the multiscale basis functions for velocity are divergence free and thus

the problem is elliptic. The proof mainly follows the lines of proof of [15, Lemma 1] and is stated for being self-contained. By the definition of $\mathbf{u}$ and $\mathbf{u}_{\text{glo}}$, we have

$$a(\mathbf{u} - \mathbf{u}_{\text{glo}}, \mathbf{v}) = 0 \quad \text{for all } \mathbf{v} \in V_{\text{glo}}. \tag{3.15}$$

Hence, we have $\mathbf{u} - \mathbf{u}_{\text{glo}} \in \tilde{V}$ and

$$a(\mathbf{u} - \mathbf{u}_{\text{glo}}, \mathbf{u} - \mathbf{u}_{\text{glo}}) = a(\mathbf{u} - \mathbf{u}_{\text{glo}}, \mathbf{u}) = (\mathbf{f}, \mathbf{u} - \mathbf{u}_{\text{glo}}) \leq \left\| \tilde{\kappa}^{-\frac{1}{2}} \mathbf{f} \right\| \left\| \mathbf{u} - \mathbf{u}_{\text{glo}} \right\|_s. \tag{3.16}$$

Since $\mathbf{u} - \mathbf{u}_{\text{glo}} \in \mathbf{V}_0 - V_{\text{aux}}$, it follows from Lemma 3.3.2 (ii) that

$$\left\| \mathbf{u} - \mathbf{u}_{\text{glo}} \right\|_s^2 \leq \Lambda^{-1} \left\| \mathbf{u} - \mathbf{u}_{\text{glo}} \right\|_a^2. \tag{3.17}$$

The result follows by combining (3.16) and (3.17). The second part follows from the fact that $|\nabla \chi_j^{\text{ms}}| = O(H^{-1})$ when $\{\chi_j^{\text{ms}}\}_{j=1}^{N_c}$ is a set of bilinear partition of unity functions. □

Next we analyze the convergence of the proposed multiscale method. We first recall Projection Theorem, which can be found in many functional analysis literature, e.g., [20, Section 4.3].

Theorem 3.3.4 (Projection Theorem)**.** *Let $\mathcal{V}$ be a closed subspace of the Hilbert space $\mathcal{H}$ equipped with an inner product $(\cdot,\cdot)_{\mathcal{H}}$. Then, for any given element $f \in \mathcal{H}$, there exists a unique element $p \in \mathcal{V}$ such that*

$$\|f - p\| = \min_{v \in \mathcal{V}} \|f - v\|.$$

Here, $\|\cdot\|$ is the norm induced by the inner product $(\cdot,\cdot)_{\mathcal{H}}$. Moreover, the mapping $\mathbb{P} : f \mapsto p$ is linear and satisfies the inequality $\|\mathbb{P} f\| \leq \|f\|$ for any $f \in \mathcal{H}$.

In the following lemma, we show the existence of a projection from $\mathbf{V}_0(D)$ to $\mathbf{V}_0^{\text{div}}(D)$ using the Projection Theorem.

Lemma 3.3.5. *Let $\mathcal{D} \subseteq \Omega^\epsilon$. Then, there exists a divergence-free projection $\mathbb{P}_{\mathcal{D}} : \mathbf{V}_0(\mathcal{D}) \to \mathbf{V}_0^{div}(\mathcal{D})$, where $\mathbf{V}_0^{div}(\mathcal{D}) := \{\mathbf{v} \in \mathbf{V}_0(\mathcal{D}) : b(\mathbf{v}, q) = 0$ for all $q \in L^2(\mathcal{D})\}$.*

Proof. Define a bilinear form on $\mathbf{V}_0(\mathcal{D})$ as follows: $(\mathbf{u}, \mathbf{v})_{as(\mathcal{D})} := a_{\mathcal{D}}(\mathbf{u}, \mathbf{v}) + s_{\mathcal{D}}(\mathbf{u}, \mathbf{v})$, where $a_{\mathcal{D}}(\cdot, \cdot)$ and $s_{\mathcal{D}}(\cdot, \cdot)$ are the restriction of $a(\cdot, \cdot)$ and $s(\cdot, \cdot)$ on the subregion $\mathcal{D}$. One can easily show that $(\cdot, \cdot)_{as}$ is an inner product defined on $\mathbf{V}_0(\mathcal{D})$.

Next, we show that $\mathbf{V}_0^{\text{div}}(\mathcal{D})$ is a closed subspace of $\mathbf{V}_0(\mathcal{D})$ with respect to the inner product $(\cdot, \cdot)_{as}$. Let $\{\mathbf{f}_n\}$ be a sequence in $\mathbf{V}_0^{\text{div}}(\mathcal{D})$ that converges to $\mathbf{f}$ in $\mathbf{V}_0(\mathcal{D})$. Since $\mathbf{f}_n \in \mathbf{V}_0^{\text{div}}(\mathcal{D})$, then we have $b(\mathbf{f}_n, g) = 0$ for all $g \in L^2(\mathcal{D})$. Then $\lim_{n\to\infty} b(\mathbf{f}_n, g) = b(\mathbf{f}, g) = 0$ for all $g \in L^2(\mathcal{D})$. It implies that $\mathbf{f} \in \mathbf{V}_0^{\text{div}}(\mathcal{D})$. Consequently, $\mathbf{V}_0^{\text{div}}(\mathcal{D})$ is a closed subspace of $\mathbf{V}_0(\mathcal{D})$. An application of Projection Theorem proves the desired result. □

Remark 3.3.6. *We denote $\|\cdot\|_{as(\mathcal{D})}$ the norm induced by the inner product $(\cdot, \cdot)_{as(\mathcal{D})}$. Then, we have $\|\mathbb{P}_{\mathcal{D}}(\mathbf{v})\|_{as(\mathcal{D})} \le \|\mathbf{v}\|_{as(\mathcal{D})}$ for any $v \in \mathbf{V}_0(\mathcal{D})$. We simply write $\|\cdot\|_{as}$ in short for $\|\cdot\|_{as(\mathcal{D})}$ when $\mathcal{D} = \Omega^\epsilon$. Moreover, the subscript $\mathcal{D}$ will be dropped from $\mathbb{P}_{\mathcal{D}}$ when there is no ambiguity.*

Lemma 3.3.7. *For any auxiliary function $v_{aux} \in V_{aux}$, there exists a function $z \in \mathbf{V}_0^{div}$ such that*

$$\pi(z) = v_{aux}, \quad \|z\|_a^2 \le D \|v_{aux}\|_s^2, \quad \text{and} \quad spt(z) \subseteq spt(v_{aux}).$$

Here, D is a generic constant depending only on the coarse mesh, the partition of unity, and the eigenvalues obtained in (3.4)*.*

Proof. Without loss of generality, we can assume that $v_{\text{aux}} \in V_{\text{aux}}^i$. Consider the following

variational problem: Find $z \in \mathbf{V}_0^{\mathrm{div}}(K_i)$ and $\mu \in V_{\mathrm{aux}}^i$ such that

$$\begin{aligned} a_i(z,v) + s_i(v,\mu) &= 0 \qquad && \text{for all } v \in \mathbf{V}_0^{\mathrm{div}}(K_i), \\ s_i(z,q) &= s_i(v_{\mathrm{aux}},q) && \text{for all } q \in V_{\mathrm{aux}}^i. \end{aligned} \tag{3.18}$$

Here, the bilinear forms $a_i(\cdot,\cdot)$ and $s_i(\cdot,\cdot)$ are defined in (3.5). We will show the well-posedness of the problem (3.18). It suffices to show that there is a function $z \in \mathbf{V}_0^{\mathrm{div}}(K_i)$ such that

$$s_i(z, v_{\mathrm{aux}}) \geq C_1 \left\| v_{\mathrm{aux}} \right\|_{s(K_i)}^2 \quad \text{and} \quad \left\| z \right\|_{a(K_i)}^2 \leq C_2 \left\| v_{\mathrm{aux}} \right\|_{s(K_i)}^2$$

for some generic constants C_1 and C_2. We denote $\mathcal{I}_{K_i} := \{j : x_j \text{ is a coarse vertex of } K_i\}$ and define $B := \prod_{j \in \mathcal{I}_{K_i}} \chi_j^{\mathrm{ms}}$. Taking $z = \mathbb{P}(Bv_{\mathrm{aux}})$, we have

$$s_i(z, v_{\mathrm{aux}}) = s_i(\mathbb{P}(Bv_{\mathrm{aux}}), v_{\mathrm{aux}}) = \int_{K_i} \tilde{\kappa} \mathbb{P}(Bv_{\mathrm{aux}}) v_{\mathrm{aux}} \, dx \geq C_\pi^{-1} \left\| v_{\mathrm{aux}} \right\|_{s(K_i)}^2 .$$

Here, the constant C_π is defined to be

$$C_\pi := \sup_{K \in \mathcal{T}^H,\ \mu \in V_{\mathrm{aux}}} \frac{\int_K \tilde{\kappa} \mu^2 \, dx}{\int_K \tilde{\kappa} \mathbb{P}(B\mu) \mu \, dx} > 0.$$

Note that $|B| \leq 1$ and $|\nabla B|^2 \leq C_{\mathcal{T}} \sum_{j \in \mathcal{I}_{K_i}} |\chi_j^{\mathrm{ms}}|^2$ with $C_{\mathcal{T}} := \max_{K \in \mathcal{T}^H} |\mathcal{I}_K|^2$. The following inequalities hold

$$\left\| z \right\|_{a(K_i)} \leq \left\| \mathbb{P}(Bv_{\mathrm{aux}}) \right\|_{as(K_i)} \leq \left\| Bv_{\mathrm{aux}} \right\|_{as(K_i)} \lesssim \left\| Bv_{\mathrm{aux}} \right\|_{a(K_i)} .$$

Then, we have

$$\left\| z \right\|_{a(K_i)}^2 \lesssim \left\| Bv_{\mathrm{aux}} \right\|_{a(K_i)}^2 \leq C_{\mathcal{T}} C_\pi (1 + \Gamma) \left\| v_{\mathrm{aux}} \right\|_{s(K_i)}^2 .$$

It shows the existence and uniqueness of the function z for a given auxiliary function $v_{\text{aux}} \in V^i_{\text{aux}}$. From the second equality in (3.18), we see that $\pi_i(z) = v_{\text{aux}}$. This completes the proof. □

The following lemma generalizes the result in [15, Lemma 5], where it shows that the global multiscale basis functions have a decay property. Here, we show that the multiscale basis functions for velocity has a similar decay property. We remark that the proof of exponential decay makes use of the techniques in the proof of [15, Lemma 5]. We emphasize that in our proof below, one has to construct test functions that are divergence-free based on the projection operator defined in Lemma 3.3.5, which differs from the proof for the elliptic case.

Lemma 3.3.8. *Let $\phi^i_j \in V_{aux}$ be a given auxiliary function. Suppose that $\psi^i_{j,ms}$ is a multiscale basis function obtained in* (3.7) *over the oversampling domain $K_{i,k}$ with $k \geq 2$ and ψ^i_j is the corresponding global basis function obtained in* (3.8)*. Then, the following estimate holds:*

$$\left\|\psi^i_j - \psi^i_{j,ms}\right\|^2_a + \left\|\pi(\psi^i_j - \psi^i_{j,ms})\right\|^2_s \leq E\left(\left\|\psi^i_j\right\|^2_a + \left\|\pi(\psi^i_j)\right\|^2_s\right),$$

where $E = 3\left(1+\Lambda^{-1}\right)\left(1+\left[6(1+\Lambda^{-1})\right]^{-1/2}\right)^{1-k}$ is a factor of exponential decay.

Proof. First, subtracting the first equation of (3.7) from that of (3.8), we obtain

$$a(\psi^i_j - \psi^i_{j,\text{ms}}, v) + s(\pi(\psi^i_j - \psi^i_{j,\text{ms}}), \pi(v)) + b(v, \xi^i_j - \xi^i_{j,\text{ms}}) = 0 \quad \text{for all } v \in \mathbf{V}_0(K_{i,k}).$$

Taking $v = w - \psi^i_{j,\text{ms}}$ with $w \in \mathbf{V}^{\text{div}}_0(K_{i,k})$, then we have

$$a(\psi^i_j - \psi^i_{j,\text{ms}}, \psi^i_{j,\text{ms}}) + s(\pi(\psi^i_j - \psi^i_{j,\text{ms}}), \pi(\psi^i_{j,\text{ms}})) = a(\psi^i_j - \psi^i_{j,\text{ms}}, w) + s(\pi(\psi^i_j - \psi^i_{j,\text{ms}}), \pi(w)). \tag{3.19}$$

Utilizing (3.19) and Cauchy-Schwarz inequality, one can show that

$$\left\|\psi_j^i - \psi_{j,\mathrm{ms}}^i\right\|_a^2 + \left\|\pi(\psi_j^i - \psi_{j,\mathrm{ms}}^i)\right\|_s^2 \leq \left\|\psi_j^i - w\right\|_a^2 + \left\|\pi(\psi_j^i - w)\right\|_s^2$$

for any $w \in \mathbf{V}_0^{\mathrm{div}}(K_{i,k})$. Let $w = \mathbb{P}(\chi_i^{k,k-1}\psi_j^i)$. Note that $\psi_j^i = \mathbb{P}(\psi_j^i)$. Then, we have

$$\begin{aligned}\left\|\psi_j^i - \psi_{j,\mathrm{ms}}^i\right\|_a^2 + \left\|\pi(\psi_j^i - \psi_{j,\mathrm{ms}}^i)\right\|_s^2 &\leq \left\|\psi_j^i - \mathbb{P}(\chi_i^{k,k-1}\psi_j^i)\right\|_a^2 + \left\|\pi(\psi_j^i - \mathbb{P}(\chi_i^{k,k-1}\psi_j^i))\right\|_s^2 \\ &\leq \left\|(1-\chi_i^{k,k-1})\psi_j^i\right\|_a^2 + \left\|(1-\chi_i^{k,k-1})\psi_j^i\right\|_s^2.\end{aligned} \tag{3.20}$$

Using (iv) and (v) of Lemma 3.3.2, we have

$$\left\|\psi_j^i - \psi_{j,\mathrm{ms}}^i\right\|_a^2 + \left\|\pi(\psi_j^i - \psi_{j,\mathrm{ms}}^i)\right\|_s^2 \leq 3(1+\Lambda^{-1})\left(\left\|\psi_j^i\right\|_{a(\Omega^\epsilon \setminus K_{i,k-1})}^2 + \left\|\pi(\psi_j^i)\right\|_{s(\Omega^\epsilon \setminus K_{i,k-1})}^2\right) \tag{3.21}$$

Next, we estimate the term $\left\|\psi_j^i\right\|_{a(\Omega^\epsilon \setminus K_{i,k-1})}^2 + \left\|\pi(\psi_j^i)\right\|_{s(\Omega^\epsilon \setminus K_{i,k-1})}^2$. We claim that it can be bounded by the term $F^2 := \left\|\psi_j^i\right\|_{a(K_{i,k-1} \setminus K_{i,k-2})}^2 + \left\|\pi(\psi_j^i)\right\|_{s(K_{i,k-1} \setminus K_{i,k-2})}^2$. This recursive property is crucial in our convergence estimate.

Note that $\mathrm{spt}(1-\chi_i^{k-1,k-2}) \subseteq \Omega^\epsilon \setminus K_{i,k-2}$ and $\mathrm{spt}(\phi_j^i) \subseteq K_i$. So $s(\phi_j^i, \pi\mathbb{P}((1-\chi_i^{k-1,k-2})\psi_j^i)) = 0$. Choosing test function $v = \mathbb{P}((1-\chi_i^{k-1,k-2})\psi_j^i)$ in the variational formulation (3.8), we have

$$a\left(\psi_j^i, \mathbb{P}((1-\chi_i^{k-1,k-2})\psi_j^i)\right) + s\left(\pi(\psi_j^i), \pi(\mathbb{P}((1-\chi_i^{k-1,k-2})\psi_j^i))\right) = 0. \tag{3.22}$$

Note that

$$\begin{aligned}a\left(\psi_j^i, \mathbb{P}((1-\chi_i^{k-1,k-2})\psi_j^i)\right) &= \int_{\Omega^\epsilon \setminus K_{i,k-2}} \nabla\psi_j^i : \nabla\left(\mathbb{P}((1-\chi_i^{k-1,k-2})\psi_j^i)\right)\, dx \\ &= \int_{\Omega^\epsilon \setminus K_{i,k-2}} |\nabla\psi_j^i|^2\, dx - \int_{\Omega^\epsilon \setminus K_{i,k-2}} \nabla\psi_j^i : \nabla\left(\mathbb{P}(\chi_i^{k-1,k-2}\psi_j^i)\right)\, dx.\end{aligned}$$

Consequently, we have

$$\begin{aligned}\left\|\psi_j^i\right\|^2_{a(\Omega^\epsilon\setminus K_{i,k-1})} &\le \int_{\Omega^\epsilon\setminus K_{i,k-2}} |\nabla\psi_j^i|^2\,dx \\ &= a\left(\psi_j^i, \mathbb{P}((1-\chi_i^{k-1,k-2})\psi_j^i)\right) + \int_{\Omega^\epsilon\setminus K_{i,k-2}} \nabla\psi_j^i : \nabla\left(\mathbb{P}(\chi_i^{k-1,k-2}\psi_j^i)\right)\,dx \qquad (3.23)\\ &\le a\left(\psi_j^i, \mathbb{P}((1-\chi_i^{k-1,k-2})\psi_j^i)\right) + \left\|\psi_j^i\right\|_{a(K_{i,k-1}\setminus K_{i,k-2})} \left\|\mathbb{P}(\chi_i^{k-1,k-2}\psi_j^i)\right\|_{as(K_{i,k-1}\setminus K_{i,k-2})}.\end{aligned}$$

Note that $\chi_i^{k-1,k-2} \equiv 0$ in $\Omega^\epsilon \setminus K_{i,k-1}$. Thus, we have

$$\begin{aligned}s\left(\pi(\psi_j^i), \pi(\mathbb{P}((1-\chi_i^{k-1,k-2})\psi_j^i))\right) &= \left\|\pi(\psi_j^i)\right\|^2_{s(\Omega^\epsilon\setminus K_{i,k-1})} \\ &\quad + \int_{K_{i,k-1}\setminus K_{i,k-2}} \tilde{\kappa}\pi(\psi_j^i)\pi\left(\mathbb{P}((1-\chi_i^{k-1,k-2})\psi_j^i)\right)\,dx\end{aligned}$$

and

$$\begin{aligned}&\left\|\pi(\psi_j^i)\right\|^2_{s(\Omega^\epsilon\setminus K_{i,k-1})} \\ &= s\left(\pi(\psi_j^i), \pi(\mathbb{P}((1-\chi_i^{k-1,k-2})\psi_j^i))\right) - \int_{K_{i,k-1}\setminus K_{i,k-2}} \tilde{\kappa}\pi(\psi_j^i)\pi\left(\mathbb{P}((1-\chi_i^{k-1,k-2})\psi_j^i)\right)\,dx \qquad (3.24)\\ &\le s\left(\pi(\psi_j^i), \pi(\mathbb{P}((1-\chi_i^{k-1,k-2})\psi_j^i))\right) + \left\|\pi(\psi_j^i)\right\|_{s(K_{i,k-1}\setminus K_{i,k-2})} \left\|\mathbb{P}((1-\chi_i^{k-1,k-2})\psi_j^i)\right\|_{as(K_{i,k-1}\setminus K_{i,k-2})}.\end{aligned}$$

Using (iv) and (v) of Lemma 3.3.2, one can show that

$$\begin{aligned}\left\|\mathbb{P}((1-\chi_i^{k-1,k-2})\psi_j^i)\right\|^2_{as(K_{i,k-1}\setminus K_{i,k-2})} &\le 3(1+\Lambda^{-1})F^2, \\ \left\|\mathbb{P}(\chi_i^{k-1,k-2}\psi_j^i)\right\|^2_{as(K_{i,k-1}\setminus K_{i,k-2})} &\le 3(1+\Lambda^{-1})F^2.\end{aligned} \qquad (3.25)$$

Combining (3.22) and the inequalities (3.23) – (3.25), we have

$$\begin{aligned}&\left\|\psi_j^i\right\|^2_{a(\Omega^\epsilon\setminus K_{i,k-1})} + \left\|\pi(\psi_j^i)\right\|^2_{s(\Omega^\epsilon\setminus K_{i,k-1})} \\ &\le \left(\left\|\psi_j^i\right\|^2_{a(K_{i,k-1}\setminus K_{i,k-2})} + \left\|\pi(\psi_j^i)\right\|^2_{s(K_{i,k-1}\setminus K_{i,k-2})}\right)^{1/2} \left[6(1+\Lambda^{-1})\right]^{1/2} F \qquad (3.26)\\ &= \left[6(1+\Lambda^{-1})\right]^{1/2} F^2.\end{aligned}$$

Notice that, using the inequality (3.26), we have

$$
\begin{aligned}
&\left\|\psi_j^i\right\|^2_{a(\Omega^\epsilon\setminus K_{i,k-2})} + \left\|\pi(\psi_j^i)\right\|^2_{s(\Omega^\epsilon\setminus K_{i,k-2})} \\
&= \left\|\psi_j^i\right\|^2_{a(\Omega^\epsilon\setminus K_{i,k-1})} + \left\|\pi(\psi_j^i)\right\|^2_{s(\Omega^\epsilon\setminus K_{i,k-1})} + \left\|\psi_j^i\right\|^2_{a(K_{i,k-1}\setminus K_{i,k-2})} + \left\|\pi(\psi_j^i)\right\|^2_{s(K_{i,k-1}\setminus K_{i,k-2})} \\
&\geq \left(1+\left[6(1+\Lambda^{-1})\right]^{-1/2}\right)\left(\left\|\psi_j^i\right\|^2_{a(\Omega^\epsilon\setminus K_{i,k-1})} + \left\|\pi(\psi_j^i)\right\|^2_{s(\Omega^\epsilon\setminus K_{i,k-1})}\right).
\end{aligned}
$$

Using the above inequality recursively, we obtain

$$
\left\|\psi_j^i\right\|^2_{a(\Omega^\epsilon\setminus K_{i,k-1})} + \left\|\pi(\psi_j^i)\right\|^2_{s(\Omega^\epsilon\setminus K_{i,k-1})} \leq \left(1+\left[6(1+\Lambda^{-1})\right]^{-1/2}\right)^{1-k}\left(\left\|\psi_j^i\right\|^2_a + \left\|\pi(\psi_j^i)\right\|^2_s\right).
$$

This completes the proof. □

The above lemma shows that the global multiscale basis is localizable. We need the following result to show the convergence estimate.

Lemma 3.3.9. *With the same notations in Lemma 3.3.8, we have*

$$
\left\|\sum_{i=1}^N(\psi_j^i-\psi_{j,ms}^i)\right\|_a^2 + \left\|\sum_{i=1}^N\pi(\psi_j^i-\psi_{j,ms}^i)\right\|_s^2 \lesssim (k+1)^d\sum_{i=1}^N\left[\left\|\psi_j^i-\psi_{j,ms}^i\right\|_a^2 + \left\|\pi(\psi_j^i-\psi_{j,ms}^i)\right\|_s^2\right].
$$

Proof. Denote $w := \sum_{i=1}^N(\psi_j^i-\psi_{j,\mathrm{ms}}^i)$.

Noice that, for any $i\in\{1,\cdots,N\}$, it holds that

$$
a(\psi_j^i-\psi_{j,\mathrm{ms}}^i,v) + s(\pi(\psi_j^i-\psi_{j,\mathrm{ms}}^i),\pi(v)) = 0, \qquad \text{for all } v\in\mathbf{V}_0^{\mathrm{div}}(K_{i,k}). \tag{3.27}
$$

Choosing $v = \mathbb{P}((1-\chi_i^{k+1,k})w)$ in (3.27), we have

$$
a\left(\psi_j^i-\psi_{j,\mathrm{ms}}^i,\mathbb{P}((1-\chi_i^{k+1,k})w)\right) + s\left(\pi(\psi_j^i-\psi_{j,\mathrm{ms}}^i),\pi(\mathbb{P}((1-\chi_i^{k+1,k})w))\right) = 0.
$$

Note that $\mathbb{P}(w) = w$. Hence, we have

$$\begin{aligned}\|w\|_a^2 + \|\pi w\|_s^2 &= \sum_{i=1}^N a(\psi_j^i - \psi_{j,\mathrm{ms}}^i, w) + s(\pi(\psi_j^i - \psi_{j,\mathrm{ms}}^i), \pi(w)) \\ &= \sum_{i=1}^N a\left(\psi_j^i - \psi_{j,\mathrm{ms}}^i, \mathbb{P}(\chi_i^{k+1,k} w)\right) + s\left(\pi(\psi_j^i - \psi_{j,\mathrm{ms}}^i), \pi(\mathbb{P}(\chi_i^{k+1,k} w))\right).\end{aligned}$$

For each $i \in \{1, \cdots, N\}$, using the properties of the cutoff function $\chi_i^{k+1,k}$ and (ii) of Lemma 3.3.2, we have the following estimates:

$$\begin{aligned}\left\|\chi_i^{k+1,k} w\right\|_a^2 &\lesssim \|w\|_{s(K_{i,k+1})}^2 + \|w\|_{a(K_{i,k+1})}^2 \leq (1 + \Lambda^{-1})\left(\|w\|_{a(K_{i,k+1})}^2 + \|\pi(w)\|_{s(K_{i,k+1})}^2\right), \\ \left\|\pi(\chi_i^{k+1,k} w)\right\|_s^2 &\leq \left\|\chi_i^{k+1,k} w\right\|_{s(K_{i,k+1})}^2 \leq \Lambda^{-1}\|w\|_{a(K_{i,k+1})}^2 + \|\pi(w)\|_{s(K_{i,k+1})}^2.\end{aligned} \tag{3.28}$$

Furthermore, an application of (3.28) we arrive at the following estimate:

$$\left\|\chi_i^{k+1,k} w\right\|_{as}^2 = \left\|\chi_i^{k+1,k} w\right\|_a^2 + \left\|\pi(\chi_i^{k+1,k} w)\right\|_s^2 \lesssim \|w\|_{a(K_{i,k+1})}^2 + \|\pi(w)\|_{s(K_{i,k+1})}^2 \tag{3.29}$$

Combining (3.28) and (3.29), we have

$$\begin{aligned}\|w\|_a^2 + \|\pi(w)\|_s^2 &\leq \sum_{i=1}^N \left\|\psi_j^i - \psi_{j,\mathrm{ms}}^i\right\|_a \cdot \left\|\chi_i^{k+1,k} w\right\|_{as} + \left\|\pi(\psi_j^i - \psi_{j,\mathrm{ms}}^i)\right\|_s \cdot \left\|\chi_i^{k+1,k} w\right\|_{as} \\ &\lesssim \sum_{i=1}^N \left(\left\|\psi_j^i - \psi_{j,\mathrm{ms}}^i\right\|_a^2 + \left\|\pi(\psi_j^i - \psi_{j,\mathrm{ms}}^i)\right\|_s^2\right)^{1/2} \cdot \left(\|w\|_{a(K_{i,k+1})}^2 + \|\pi(w)\|_{s(K_{i,k+1})}^2\right)^{1/2} \\ &\lesssim \left(\sum_{i=1}^N \left\|\psi_j^i - \psi_{j,\mathrm{ms}}^i\right\|_a^2 + \left\|\pi(\psi_j^i - \psi_{j,\mathrm{ms}}^i)\right\|_s^2\right)^{1/2} \left(\sum_{i=1}^N \|w\|_{a(K_{i,k+1})}^2 + \|\pi(w)\|_{s(K_{i,k+1})}^2\right)^{1/2} \\ &\lesssim (k+1)^{d/2} \left(\sum_{i=1}^N \left\|\psi_j^i - \psi_{j,\mathrm{ms}}^i\right\|_a^2 + \left\|\pi(\psi_j^i - \psi_{j,\mathrm{ms}}^i)\right\|_s^2\right)^{1/2} \left(\|w\|_a^2 + \|\pi(w)\|_s^2\right)^{1/2}.\end{aligned}$$

Therefore, we have

$$\|w\|_a^2 + \|\pi(w)\|_s^2 \lesssim (k+1)^d \sum_{i=1}^N \left[\left\|\psi_j^i - \psi_{j,\mathrm{ms}}^i\right\|_a^2 + \left\|\pi(\psi_j^i - \psi_{j,\mathrm{ms}}^i)\right\|_s^2\right].$$

This completes the proof. □

Finally, we state and prove the main result of this work. It reads as follows.

Theorem 3.3.10. *Let* $\mathbf{u}$ *be the solution of* (3.2) *and* $\mathbf{u}_{ms}$ *be the solution of* (3.10)*. Then, we have*

$$\left\|\mathbf{u}-\mathbf{u}_{ms}\right\|_a \lesssim \Lambda^{-1}\left\|\tilde{\kappa}^{-1/2}\mathbf{f}\right\| + \max\{\tilde{\kappa}\}(k+1)^{d/2}E^{1/2}(1+D)\left\|\mathbf{u}_{glo}\right\|_s,$$

where $\mathbf{u}_{glo}$ *is the solution of* (3.14)*. Moreover, if the oversampling parameter* k *is sufficiently large and* $\{\chi_i^{ms}\}_{i=1}^{N_c}$ *is a set of bilinear partition of unity, we have*

$$\left\|\mathbf{u}-\mathbf{u}_{ms}\right\|_a \lesssim H\Lambda^{-1}\left\|\mathbf{f}\right\|.$$

Proof. It follows from Galerkin orthogonality that $\left\|\mathbf{u}-\mathbf{u}_{\mathrm{ms}}\right\|_a \leq \left\|\mathbf{u}-\mathbf{v}\right\|_a$ for any $\mathbf{v}\in V_{\mathrm{ms}}$. We write $\mathbf{u}_{\mathrm{glo}} := \sum_{i=1}^{N}\sum_{j=1}^{\ell_i} c_{ij}\psi_j^i$ and define a function $\mathbf{v}$ such that $\mathbf{v} := \sum_{i=1}^{N}\sum_{j=1}^{\ell_i} c_{ij}\psi_{j,\mathrm{ms}}^i$. Then, we have

$$\left\|\mathbf{u}-\mathbf{u}_{\mathrm{ms}}\right\|_a \leq \left\|\mathbf{u}-\mathbf{v}\right\|_a \leq \left\|\mathbf{u}-\mathbf{u}_{\mathrm{glo}}\right\|_a + \left\|\mathbf{u}_{\mathrm{glo}}-\mathbf{v}\right\|_a.$$

The first term of the right-hand side can be estimated by the result of (3.17). It suffices to estimate the second term. By Lemmas 3.3.8 and 3.3.9, we have

$$\begin{aligned}
\left\|\mathbf{u}_{\mathrm{glo}}-\mathbf{v}\right\|_a^2 &= \left\|\sum_{i=1}^{N}\sum_{j=1}^{\ell_i} c_{ij}(\psi_j^i-\psi_{j,\mathrm{ms}}^i)\right\|_a^2 \\
&\leq C(k+1)^d\sum_{i=1}^{N}\left(\left\|\sum_{j=1}^{\ell_i} c_{ij}(\psi_j^i-\psi_{j,\mathrm{ms}}^i)\right\|_a^2 + \left\|\sum_{j=1}^{\ell_i} c_{ij}\pi(\psi_j^i-\psi_{j,\mathrm{ms}}^i)\right\|_s^2\right) \\
&\leq C(k+1)^d E\sum_{i=1}^{N}\sum_{j=1}^{\ell_i}(c_{ij})^2\left(\left\|\psi_j^i\right\|_a^2 + \left\|\pi\psi_j^i\right\|_s^2\right).
\end{aligned}$$

Choosing test function $v = \psi_j^i$ in (3.8), we obtain that $\left\|\psi_j^i\right\|_a^2 + \left\|\pi\psi_j^i\right\|_s^2 \leq \left\|\phi_j^i\right\|_s^2$. Therefore,

$$\left\|\mathbf{u}_{\text{glo}} - \mathbf{v}\right\|_a^2 \lesssim E(k+1)^d \sum_{i=1}^{N}\sum_{j=1}^{\ell_i}(c_{ij})^2 \left\|\phi_j^i\right\|_s^2 = E(k+1)^d \sum_{i=1}^{N}\sum_{j=1}^{\ell_i}(c_{ij})^2. \tag{3.30}$$

Next, we estimate the term $\sum_{i=1}^{N}\sum_{j=1}^{\ell_i}(c_{ij})^2$. Note that $\pi(\mathbf{u}_{\text{glo}}) = \sum_{i=1}^{N}\sum_{j=1}^{\ell_i} c_{ij}\pi(\psi_j^i)$. Using the variational formulation (3.8) with test function $v = \psi_j^i$, we obtain

$$b_{\ell k} := s\left(\pi(\mathbf{u}_{\text{glo}}), \phi_k^\ell\right) = \sum_{i=1}^{N}\sum_{j=1}^{\ell_i} c_{ij} s(\pi(\psi_j^i), \phi_k^\ell) = \sum_{i=1}^{N}\sum_{j=1}^{\ell_i} c_{ij}\left(\underbrace{a(\psi_j^i, \psi_k^\ell) + s(\pi(\psi_j^i), \pi(\psi_k^\ell))}_{=:a_{ij,\ell k}}\right).$$

If we denote $\mathbf{b} = (b_{\ell k}) \in \mathbb{R}^{\mathcal{N}}$ and $\mathbf{c} = (c_{ij}) \in \mathbb{R}^{\mathcal{N}}$ with $\mathcal{N} := \sum_{i=1}^{N} \ell_i$, then we have

$$\mathbf{b} = A\mathbf{c} \quad \text{and} \quad \|\mathbf{c}\|_2 \leq \left\|A^{-1}\right\|_2 \|\mathbf{b}\|_2,$$

where $A := (a_{ij,\ell k}) \in \mathbb{R}^{\mathcal{N}\times\mathcal{N}}$ and $\|\cdot\|_2$ denotes the standard Euclidean norm for vectors in $\mathbb{R}^{\mathcal{N}}$ and its induced matrix norm in $\mathbb{R}^{\mathcal{N}\times\mathcal{N}}$. By the definition of $\pi : \mathbf{V} \to V_{\text{aux}}$, we have

$$\pi(\mathbf{u}_{\text{glo}}) = \pi\left(\pi(\mathbf{u}_{\text{glo}})\right) = \sum_{i=1}^{N}\sum_{j=1}^{\ell_i} s(\pi(\mathbf{u}_{\text{glo}}), \phi_j^i)\phi_j^i = \sum_{i=1}^{N}\sum_{j=1}^{\ell_i} b_{ij}\phi_j^i.$$

Thus, we have $\|\mathbf{b}\|_2 = \|\pi(\mathbf{u}_{\text{glo}})\|_s$. We define $\phi := \sum_{i=1}^{N}\sum_{j=1}^{\ell_i} c_{ij}\phi_j^i$. Note that $\|\phi\|_s = \|\mathbf{c}\|_2$. Consequently, by Lemma 3.3.7, there exists a function $z \in \mathbf{V}_0^{\text{div}}$ such that $\pi(z) = \phi$ and $\|z\|_a^2 \leq D\|\phi\|_s^2$. Since the multiscale basis ψ_j^i satisfies (3.8) and $\mathbf{u}_{\text{glo}}$ is a linear combination of ψ_j^i's, we have

$$a(\mathbf{u}_{\text{glo}}, v) + s(\pi(\mathbf{u}_{\text{glo}}), \pi v) = s(\phi, \pi v) \quad \text{for all } v \in \mathbf{V}_0^{\text{div}}(\Omega^\epsilon). \tag{3.31}$$

Picking $v = z$ in (3.31), we arrive at

$$\begin{aligned}\|\phi\|_s^2 &= a(\mathbf{u}_{\text{glo}}, z) + s(\pi(\mathbf{u}_{\text{glo}}), \pi z) \leq \|\mathbf{u}_{\text{glo}}\|_a \cdot D^{1/2} \|\phi\|_s + \|\pi\mathbf{u}_{\text{glo}}\|_s \cdot \|\phi\|_s \\ &\leq (1+D)^{1/2} \|\phi\|_s \left(\|\mathbf{u}_{\text{glo}}\|_a^2 + \|\pi\mathbf{u}_{\text{glo}}\|_s^2\right)^{1/2}.\end{aligned}$$

Therefore, we have

$$\|\mathbf{c}\|_2^2 = \|\phi\|_s^2 \leq (1+D)\left(\|\mathbf{u}_{\text{glo}}\|_a^2 + \|\pi\mathbf{u}_{\text{glo}}\|_s^2\right) = (1+D)\mathbf{c}^T A\mathbf{c}.$$

From the above, we see that the largest eigenvalue of A^{-1} is bounded by $(1+D)$ and we have the following estimate

$$\|\mathbf{c}\|_2 \leq (1+D)\|\mathbf{b}\|_2 = (1+D)\|\mathbf{u}_{\text{glo}}\|_s.$$

As a result, we have

$$\|\mathbf{u}_{\text{glo}} - \mathbf{v}\|_a^2 \leq (k+1)^d E(1+D)^2 \|\mathbf{u}_{\text{glo}}\|_s^2.$$

It remains to estimate the term $\|\mathbf{u}_{\text{glo}}\|_s$. In particular, we have

$$\|\mathbf{u}_{\text{glo}}\|_s^2 \lesssim \max\{\tilde{\kappa}\}\|\mathbf{u}_{\text{glo}}\|_a^2 = \max\{\tilde{\kappa}\}\langle \mathbf{f}, \mathbf{u}_{\text{glo}}\rangle \leq \max\{\tilde{\kappa}\}\left\|\tilde{\kappa}^{-1/2}\mathbf{f}\right\| \|\mathbf{u}_{\text{glo}}\|_s.$$

Therefore, we have

$$\|\mathbf{u} - \mathbf{u}_{\text{ms}}\|_a \lesssim \Lambda^{-1}\left\|\tilde{\kappa}^{-1/2}\mathbf{f}\right\| + \max\{\tilde{\kappa}\}(k+1)^{d/2}E^{1/2}(1+D)\left\|\tilde{\kappa}^{-1/2}\mathbf{f}\right\|.$$

If we take $k = O(\log(H^{-1}))$ and assume that $\{\chi_i^{\text{ms}}\}_{i=1}^{N_c}$ is a set of bilinear partition of unity,

then we have

$$\|\mathbf{u} - \mathbf{u}_{\mathrm{ms}}\|_a \lesssim H\Lambda^{-1} \|\mathbf{f}\| .$$

This completes the proof. □

4. WAVELET-BASED EDGE MULTISCALE PARAREAL ALGORITHM FOR PARABOLIC EQUATIONS WITH HETEROGENEOUS COEFFICIENTS AND ROUGH INITIAL DATA

We consider in this chapter a new efficient multiscale parareal algorithm for parabolic problems with heterogeneous coefficients. We incorporate the parareal algorithm into Wavelet-based Edge Multiscale Finite Element Method (WEMsFEM) to numerically calculate the time evolution problems efficiently.

The main idea of WEMsFEM is to utilize wavelets as the basis functions over the coarse edges, and transform the approximation properties from the edges to each local region. Then the Partition of Unity Method (PUM) [59] is applied to derive the global convergence rate. The motivation for using wavelets as the ansatz space over the coarse edges origins from the low regularity of the solution to (4.1) in the spatial domain due to the existence of heterogeneity in the coefficient κ, which makes its approximation use the standard basis functions, e.g., the element-wise polynomials, infeasible or even prohibitive. Further, the multiresolution analysis enables the approximation of functions with low regularities using wavelets. We will apply this method in this chapter to handle the heterogeneity in the spatial domain.

The parareal algorithm facilitates speeding up the numerical solver to time dependent equations on the condition of sufficient processors [9], which is an iterative solver based on a cheap inaccurate sequential coarse-scale time solver and expensive accurate fine-scale time solvers that can be performed in parallel. If it converges sufficiently fast, then the parareal algorithm could result in less wall-clock time than sequentially computing.

Our proposed algorithm is called WEMP Algorithm, c.f. Algorithm 2. This algorithm is divided into two steps: a multiscale space $V_{\mathrm{ms},\ell}^{\mathrm{EW}}$ based on WEMsFEM with ℓ as the

wavelets level parameter is constructed in the first step, and then we apply the parareal algorithm by using $V_{\mathrm{ms},\ell}^{\mathrm{EW}}$ as the ansatz space in the second step to obtain the solution more efficiently.

The chapter is organized as follows. We summarize the basics on the fully discretization of Problem (4.1), the framework of WEMsFEMs in Section 4.1. Our main proposed algorithm is presented in Section 4.2. The convergence of WEMsFEM and WEMP algorithms are derived in Section 4.3. Extensive numerical tests are presented in Section 4.4.

4.1 Problem setting and the construction of multiscale space

In this section, we will mainly introduce the full discretization of problem (4.1), and its multiscale model reduction in the spatial domain D.

We first formulate the heterogeneous parabolic problems to present our new multiscale methods. Let $D \subset \mathbb{R}^d$ $(d = 1, 2, 3)$ be an open bounded Lipschitz domain. We seek a function $u(\cdot, t) \in V := H_0^1(D)$ such that

$$\begin{aligned} \frac{\partial u}{\partial t} - \nabla \cdot (\kappa \nabla u) &= f \qquad && \text{in } D \times (0, T] \\ u(\cdot, 0) &= u_0 \qquad && \text{in } D \\ u &= 0 \qquad && \text{on } \partial D \times [0, T], \end{aligned} \tag{4.1}$$

where the force term $f \in L^\infty([0,T]; \dot{H}^2(D))$ satisfying $\partial_t f \in L^1([0,T]; L^2(D))$, the initial data $u_0 \in L^2(D)$ and the permeability coefficient $\kappa \in C^\infty(D)$ with $\alpha \le \kappa(x) \le \beta$ almost everywhere for some lower bound $\alpha > 0$ and upper bound $\beta > \alpha$. Here, $\dot{H}^s(D) \subset L^2(D)$ is a Hilbert space to be defined in (4.2). We denote by $\Lambda := \frac{\beta}{\alpha}$ the ratio of these bounds, which reflects the contrast of the coefficient κ. To simplify the notation, let $I := [0, T]$. Note that the existence of multiple scales in the coefficient κ rends directly solving

Problem (4.1) challenging, since resolving the problem to the finest scale would incur huge computational cost.

4.1.1 Full discretization

We present in this subsection the discretization of problem (4.1). Firstly, we define the Hilbert space $\dot{H}^s(D)$, which is analogous to [60, Chapter 3].

Let $\{(\lambda_m, \phi_m)\}_{m=1}^{\infty}$ be the eigenpairs of the following eigenvalue problems with the eigenvalues arranged in a nondecreasing order,

$$\begin{aligned} \mathcal{L}\phi_m := -\nabla \cdot (\kappa \nabla \phi_m) &= \lambda_m \phi_m && \text{in } D \\ \phi_m &= 0 && \text{on } \partial D. \end{aligned}$$

Note that the eigenfunctions $\{\phi_m\}_{m=1}^{\infty}$ form an orthonormal basis in $L^2(D)$, and consequently, each $v \in L^2(D)$ admits the representation $v = \sum_{m=1}^{\infty} (v, \phi_m)_D \phi_m$ with $(\cdot,\cdot)_D$ being the inner product in $L^2(D)$. The Hilbert space $\dot{H}^s(D) \subset L^2(D)$ is defined by

$$\dot{H}^s(D) = \{v \in L^2(D) : \sum_{m=1}^{\infty} \lambda_m^s |(v, \phi_m)_D|^2 < \infty\}. \tag{4.2}$$

The associated norm in $\dot{H}^s(D)$ is $|v|_s = (\sum_{m=1}^{\infty} \lambda_m^s |(v, \phi_m)_D|^2)^{1/2}$.

Remark 4.1.1. *Since the initial data $u_0 \in \dot{H}^3(D) \cap H_1^0(D)$, we obtain*

$$\|\mathcal{L}u_0\|_{L^2(D)} = |u_0|_2. \tag{4.3}$$

Indeed, u_0 allows the expression

$$u_0 = \sum_{m=1}^{\infty} (u_0, \phi_m)_D \phi_m.$$

Taking $L^2(D)$-norm after operating $\mathcal{L}$ on both sides and utilize the definition (4.2)*, we obtain the desired assertion* (4.3)*.*

To discretize problem (4.1), we first introduce fine and coarse grids. Let $\mathcal{T}_H$ be a regular partition of the domain D into finite elements (triangles, quadrilaterals, tetrahedral, etc.) with a mesh size H. We refer to this partition as coarse grids, and its elements as the coarse elements. Then each coarse element is further partitioned into a union of connected fine grid blocks. The fine-grid partition is denoted by $\mathcal{T}_h$ with h being its mesh size. Let $\mathcal{F}_h$ (or $\mathcal{F}_H$) be the collection of all edges in $\mathcal{T}_h$ (or $\mathcal{T}_H$). Over the fine mesh $\mathcal{T}_h$, let V_h be the conforming piecewise linear finite element space:

$$V_h := \{v \in V : v|_E \in \mathcal{P}_1(E) \text{ for all } E \in \mathcal{T}_h\},$$

where $\mathcal{P}_1(E)$ denotes the space of linear polynomials on the fine element $E \in \mathcal{T}_h$.

The time interval $I := [0, T]$ is decomposed into a sequence of coarse subintervals $[T^n, T^{n+1}]$ for $n = 0, 1, \cdots, M_\Delta$ of size ΔT with $\Delta T := T/M_\Delta$ for some $M_\Delta \in \mathbb{N}_+$ and $T^0 := 0$. Each coarse time interval $[T^n, T^{n+1}]$ is further discretized with a fine time step δt. Let $t_n = n \times \delta t$ for $n = 0, 1, \cdots, M_\delta$ with $M_\delta := T \times \delta t^{-1}$. Note that $\Delta T \gg \delta t$. To simplify the notations, backward Euler method is utilized to discretize the time variable, and we use conforming Galerkin method for the discretization in the spatial variable throughout this chapter. Then the fine-scale solution $U_h^n \in V_h$ for $n = 1, 2, \cdots, M_\delta$ satisfies

$$\begin{cases} (\dfrac{U_h^n - U_h^{n-1}}{\delta t}, v_h)_D + a(U_h^n, v_h) = (f(\cdot, t_n), v_h)_D \quad \text{for all } v_h \in V_h, \\ U_h^0 = I_h u_0. \end{cases} \tag{4.4}$$

Here, the bilinear form $a(\cdot,\cdot)$ on $V \times V$ is defined by

$$a(v_1, v_2) \coloneqq \int_D \kappa \nabla v_1 \cdot \nabla v_2 \, \mathrm{d}x \text{ for all } v_1, v_2 \in V.$$

I_h is a proper projection from V to V_h. Furthermore, we define the energy norm $\|v\|_{H^1_\kappa(D)} \coloneqq \sqrt{a(v,v)}$ for all $v \in V$.

The fine-scale solution u_h^n will serve as a reference solution in Section 4.4. Note that due to the presence of multiple scales in the coefficient κ, the fine-scale mesh size h should be commensurate with the smallest scale and thus it can be very small in order to obtain an accurate solution. This necessarily involves huge computational complexity, and more efficient methods are in great demand.

4.1.2 Multiscale model reduction in the spatial domain D

We present in the section the multiscale model reduction to Problem (4.4) in the spatial domain D.

4.1.2.1 Multiscale solver in the spatial domain

The multiscale method we are investigating aim at solving Problem (4.1) on the coarse mesh $\mathcal{T}_H$, which, meanwhile, maintains a certain accuracy compared to the fine-scale solution U_h^n to Problem (4.4). To provide a brief overview, we first recap a few definitions.

The vertices of $\mathcal{T}_H$ are denoted by $\{O_i\}_{i=1}^N$, with N being the total number of coarse nodes. The coarse neighborhood associated with the node O_i is denoted by

$$\omega_i \coloneqq \bigcup \{K_j \in \mathcal{T}_H : \quad O_i \in \overline{K}_j\}. \tag{4.5}$$

We refer to Figure 4.1 for an illustration of neighborhoods and elements subordinated to the coarse discretization $\mathcal{T}_H$. Throughout, we use ω_i to denote a coarse neighborhood.

Furthermore, let $\mathcal{F}_h(\partial\omega_i)$ (or $\mathcal{F}_H(\partial\omega_i)$) be the restriction of $\mathcal{F}_h$ on $\partial\omega_i$ (or $\mathcal{F}_H$ on $\partial\omega_i$).

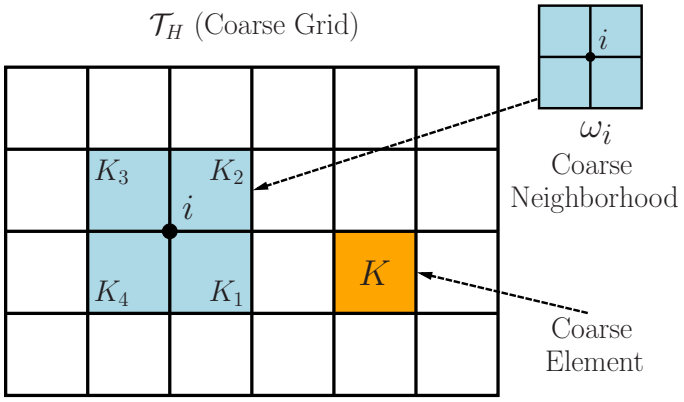

Figure 4.1: Illustration of a coarse neighborhood and coarse element.

Let V_{ms} be the multiscale finite element space to be defined in Section 4.1.2.2. The multiscale solution $U^n_{\text{ms}} \in V_{\text{ms}}$ for $n = 1, \cdots, M_\delta$ satisfies

$$\begin{cases} (\dfrac{U^n_{\text{ms}} - U^{n-1}_{\text{ms}}}{\delta t}, v_{\text{ms}})_D + a(U^n_{\text{ms}}, v_{\text{ms}}) = (f(\cdot, t_n), v_{\text{ms}})_D \quad \text{for all } v_{\text{ms}} \in V_{\text{ms}}, \\ U^0_{\text{ms}} = I_{\text{ms}} u_0, \end{cases} \tag{4.6}$$

where V_{ms} denotes the multiscale space spanned by these multiscale basis functions and I_{ms} is a $L^2(D)$-projection operator from V to V_{ms}.

Note that we need a very tiny fine-scale time step δt to guarantee a reasonable approximation property of u^n_{ms} to $u(\cdot, t_n)$ for $n = 1, \cdots, M_\delta$ due to, e.g., the singularity of the solution $u(\cdot, t)$ at $t = 0$ arising from the rough initial data u_0 or when the source term f fails to have certain regularity. Consequently, the computational complexity of the multiscale method (4.6) can be extremely expensive. For this reason, we present in Section 4.2 a multiscale algorithm incorporated with the parareal algorithm to reduce further this part

of computational cost.

We end this section with assumptions on the permeability field κ, which is required to obtain approximation properties of the multiscale finite element space V_{ms} in the energy norm, c.f., (4.14):

Assumption 4.1.1 (Structure of D and κ)**.** *Let D be a domain with a $C^{1,\alpha}$ $(0 < \alpha < 1)$ boundary ∂D, and $\{D_i\}_{i=1}^m \subset D$ be m pairwise disjoint strictly convex open subsets, each with a $C^{1,\alpha}$ boundary $\Gamma_i := \partial D_i$, and denote $D_0 = D\backslash\overline{\cup_{i=1}^m D_i}$. Let the permeability coefficient κ be piecewise regular function defined by*

$$\kappa = \begin{cases} \eta_i(x) & \text{in } D_i, \\ 1 & \text{in } D_0. \end{cases} \tag{4.7}$$

Here $\eta_i \in C^{\mu}(\bar{D}_i)$ with $\mu \in (0,1)$ for $i = 1, \cdots, m$. Denote $\eta_{min} := \min_i\{\min_{x\in D_i}\{\eta_i(x)\}\} \geq 1$ and $\eta_{max} := \max_i\{\|\eta_i\|_{C_0(D_i)}\}$.

4.1.2.2 Multiscale space construction

This subsection is concerned with the construction of the multiscale space by means of the Wavelet-based Edge Multiscale Finite Element Methods (WEMsFEM) [47, 30]. The algorithm is presented in Algorithm 1. Given the level parameter $\ell \in \mathbb{N}$, and the type of wavelets on each edge of the coarse neighborhood ω_i, one can obtain the local multiscale space $V_{i,\ell}$ on ω_i by solving $2^{\ell+2}$ local problems in Step 2 using the fine-scale mesh and its associated proper finite element space. Those local problems $\mathcal{L}_i$ are homogeneous elliptic operators coupled with wavelets Dirichlet data $V_{i,\ell}$. In Step 3, we can use these local multiscale space to build the global multiscale space $V_{\text{ms},\ell}^{\text{EW}}$ by multiplying the partition of unity functions χ_i. Finally, we can solve (4.6) by backward Euler conforming Galerkin scheme using this global multiscale space, coupled with $I_\ell u_0$ as the initial condition. Here, I_ℓ denotes the $L^2(D)$-projection from $L^2(D)$ to $V_{\text{ms},\ell}^{\text{EW}}$.

The weighted coefficient appears in Step 3 is

$$\widetilde{\kappa} = H^2\kappa \sum_{i=1}^{N} |\nabla \chi_i|^2. \tag{4.8}$$

Further, its inverse $\widetilde{\kappa}^{-1}$ is

$$\widetilde{\kappa}^{-1}(x) := \begin{cases} \widetilde{\kappa}^{-1}, & \text{when } \widetilde{\kappa}(x) \neq 0, \\ 1, & \text{otherwise }, \end{cases} \tag{4.9}$$

which will be utilized in the analysis. The partition of unity functions χ_i used in Step 4 are the standard multiscale basis functions defined coarse elementwise. On each coarse element $K \in \mathcal{T}_H$, it satisfies

$$\begin{aligned} -\nabla \cdot (\kappa(x)\nabla \chi_i) = 0 \quad &\text{in} \quad K, \\ \chi_i = g_i \quad &\text{on } \partial K, \end{aligned} \tag{4.10}$$

where g_i is affine over ∂K with $g_i(O_j) = \delta_{ij}$ for all $i, j = 1, \cdots, N$. Recall that $\{O_j\}_{j=1}^N$ are the set of coarse nodes on $\mathcal{T}_H$. By its definition, χ_i is locally supported,

$$\text{supp}(\chi_i) \subset \omega_i.$$

Algorithm 1 Wavelet-based Edge Multiscale Finte Element Method (WEMsFEM)

Input: The level parameter $\ell \in \mathbb{N}$; coarse neighborhood ω_i and its four coarse edges $\Gamma_{i,k}$ with $k = 1, 2, 3, 4$, i.e., $\cup_{k=1}^{4}\Gamma_{i,k} = \partial\omega_i$; the subspace $V_{\ell,k}^{i} \subset L^2(\Gamma_{i,k})$ up to level ℓ on each coarse edge $\Gamma_{i,k}$.

Output: Multiscale solution $u_{\mathrm{ms},\ell}^{\mathrm{EW}}$.

1: Denote $V_{i,\ell} := \oplus_{k=1}^{4} V_{\ell,k}^{i}$. Then the number of basis functions in $V_{i,\ell}$ is $4 \times 2^{\ell} = 2^{\ell+2}$. Denote these basis functions as v_k for $k = 1, \cdots, 2^{\ell+2}$.

2: Calculate local multiscale basis $\mathcal{L}_i^{-1}(v_k)$ for all $k = 1, \cdots, 2^{\ell+2}$. Here, $\mathcal{L}_i^{-1}(v_k) := v$ satisfies:
$$\begin{cases} \mathcal{L}_i v := -\nabla \cdot (\kappa \nabla v) = 0 & \text{in } \omega_i, \\ v = v_k & \text{on } \partial\omega_i. \end{cases}$$

3: Solve one local problem.
$$\begin{cases} -\nabla \cdot (\kappa \nabla v^i) = \dfrac{\widetilde{\kappa}}{\int_{\omega_i} \widetilde{\kappa}\, \mathrm{d}x} & \text{in } \omega_i, \\ -\kappa \dfrac{\partial v^i}{\partial n} = |\partial\omega_i|^{-1} & \text{on } \partial\omega_i. \end{cases}$$

4: Build global multiscale space. $V_{\mathrm{ms},\ell}^{\mathrm{EW}} := \mathrm{span}\{\chi_i \mathcal{L}_i^{-1}(v_k), \chi_i v^i : 1 \le i \le N,\ 1 \le k \le 2^{\ell+2}\}$.

5: Solve for (4.6) by backward Euler conforming Galerkin method in $V_{\mathrm{ms},\ell}^{\mathrm{EW}}$ with $U_{\mathrm{ms},\ell}^{\mathrm{EW},1} = I_\ell u_0$ to obtain $U_{\mathrm{ms},\ell}^{\mathrm{EW},n}$ for $n = 1, \cdots, M_\delta$.

4.2 Wavelet-based Edge Multiscale Parareal Algorithm

We construct in this section the Wavelet-based Edge Multiscale Parareal (WEMP) Algorithm, cf. Algorithm 2, which is divided into two main steps. In the first step, the multiscale space $V_{\mathrm{ms},\ell}^{\mathrm{EW}}$ for $\ell \in \mathbb{N}_+$ is built based on Section 4.1.2.2. This multiscale space serves as the trial space and test space for our conforming Galerkin method, cf. (4.6). Then the parareal algorithm is utilized in the second step to solve the problem.

We first recap a few terminologies commonly appeared in parareal algorithm.

The one step coarse solver on the time domain $(0, T)$ is

$$\begin{aligned} U^{n+1} &= E_\Delta(T^n, U^n), \quad U^0 = I_h u_0, \\ U^{n+1}_{\mathrm{ms},\ell} &= E^{\mathrm{ms},\ell}_\Delta(T^n, U^n_{\mathrm{ms},\ell}), \quad U^0_{\mathrm{ms},\ell} = I_{\mathrm{ms}} u_0, \end{aligned} \tag{4.11}$$

which yields U^{n+1} (or $U^{n+1}_{\mathrm{ms},\ell}$) as a coarse approximation to $u(\cdot, T^{n+1})$, provided with an approximation U^n (or $U^n_{\mathrm{ms},\ell}$) of $u(\cdot, T^n)$. In matrix form, it reads

$$\begin{aligned} U^{n+1} &= (M + \Delta T \times A)^{-1} M (U^n + \Delta T \times F^{n+1}), \\ U^{n+1}_{\mathrm{ms},\ell} &= \Phi_{\mathrm{ms},\ell} (\Phi^T_{\mathrm{ms},\ell} M \Phi_{\mathrm{ms},\ell} + \Delta T \times \Phi^T_{\mathrm{ms},\ell} A \Phi_{\mathrm{ms},\ell})^{-1} \Phi^T_{\mathrm{ms},\ell} M (U^n_{\mathrm{ms},\ell} + \Delta T \times F^{n+1}). \end{aligned}$$

Here, A and M are the mass matrices and stiffness matrices corresponding to the discretization of the elliptic operator $-\nabla\cdot(\kappa\nabla\cdot)$ in the finite element space $V_h := \mathrm{span}\{\phi_1, \cdots, \phi_{\mathrm{dof}_h}\}$. Here, dof_h denotes the dimension of V_h. $(F^{n+1})_i := \int_D f(\cdot, t_{n+1})\phi_i \,\mathrm{d}x$ for all $i = 1, \cdots, \mathrm{dof}_h$. $\Phi_{\mathrm{ms},\ell}$ denotes a matrix with columns composed of the coefficients of multiscale basis functions in $V^{\mathrm{EW}}_{\mathrm{ms},\ell}$ in the finite element space V_h.

The one step fine solver

$$\begin{aligned} \psi &= \mathcal{F}_\delta(s, \sigma, \phi), \\ \psi_{\mathrm{ms},\ell} &= \mathcal{F}^{\mathrm{ms},\ell}_\delta(s, \sigma, \phi), \end{aligned} \tag{4.12}$$

yields an approximation $\psi(\cdot, s + \sigma)$ (or $\psi_{\mathrm{ms},\ell}(\cdot, s + \sigma)$) to the solution $u(\cdot, s + \sigma)$ with the initial condition $\psi(\cdot, s) = \phi$ (or $\psi_{\mathrm{ms},\ell}(\cdot, s) = \mathcal{P}_\ell(\phi)$) and a uniform discrete time step δ for all $s \in (0, T)$ and $\sigma \in (0, T - s)$ in the infinite dimensional space V (or in the ansatz space $V^{\mathrm{EW}}_{\mathrm{ms},\ell}$) with $s/\delta t \in \mathbb{N}_+$.

We also define the semi-discretization in space solver

$$u_{\mathrm{ms},\ell}(\cdot, s+\sigma) = \mathcal{F}^{\mathrm{ms},\ell}(s,\sigma,\phi), \tag{4.13}$$

which yields an approximation $u_{\mathrm{ms},\ell}(\cdot, s+\sigma)$ to the solution $u(\cdot, s+\sigma)$ with initial condition $u_{\mathrm{ms},\ell}(\cdot, s) = \mathcal{P}_\ell(\phi)$ for all $s \in (0,T)$, $\sigma \in (0, T-s)$ in the ansatz space $V^{\mathrm{EW}}_{\mathrm{ms},\ell}$. We will denote $\bar{E}^{\mathrm{ms},\ell}_{\Delta}(T^n, U^n_{\mathrm{ms},\ell})$ as the one step coarse solver with $f = 0$. We will define $\bar{\mathcal{F}}^{\mathrm{ms},\ell}(s,\sigma,\phi)$ and $\bar{\mathcal{F}}^{\mathrm{ms},\ell}_{\delta}(s,\sigma,\phi)$ analogously.

Note that the cheap multiscale coarse solver $E^{\mathrm{ms},\ell}_{\Delta}$ is sequentially utilized over the global time interval I to provide a rough approximation to $u(\cdot, T^{n+1})$, while the expensive accurate multiscale fine solver $\mathcal{F}^{\mathrm{ms},\ell}_{\delta}$ is applied in each subinterval $[T^n, T^{n+1}]$ for $n = 0, 1, \cdots, M_\Delta - 1$ independently. This local fine solver will embed more detailed information to the approximation of $u(\cdot, T^{n+1})$, which usually differs from the one obtained from the global coarse solver. In the process of parareal algorithm, a correction operator is very important to improve the approximation to $u(\cdot, T^{n+1})$ based on the discrepancy between the coarse solver and fine solver, which is defined by

$$\mathcal{S}(T^n, U^n_{\mathrm{ms},\ell}) := \mathcal{F}^{\mathrm{ms},\ell}_{\delta}(T^n, \Delta T, U^n_{\mathrm{ms},\ell}) - E^{\mathrm{ms},\ell}_{\Delta}(T^n, U^n_{\mathrm{ms},\ell}) \quad \text{and} \quad U^0_{\mathrm{ms},\ell} = I_{\mathrm{ms}} u_0$$

for all $n = 0, 1, \cdots, M_\Delta - 1$.

Now we are ready to present our main algorithm, i.e., Algorithm 2. To obtain a good approximation to the solution of (4.1) at discrete time points $\{T^n\}$ for $n = 1, \cdots, M_\Delta$, we first construct a proper multiscale space $V^{\mathrm{EW}}_{\mathrm{ms},\ell}$ based on the WEMsFEM, i.e., Algorithm 1, which corresponds to Steps 1 to 3. This allows one to solve (4.6) using the constructed multiscale space $V^{\mathrm{EW}}_{\mathrm{ms},\ell}$ and obtain an intermediate solution $U^{\mathrm{EW},n}_{\mathrm{ms},\ell}$ with certain accuracy depending on the spatial coarse mesh size H and level parameter ℓ. This solution will only

be utilized in the convergence analysis.

In order to further reduce the computational cost, we apply the parareal algorithm in the following. Given the iteration parameter k, we apply the global coarse solver (4.11) in Step 6 to obtain U_k^{n+1}, which is an approximation to the intermediate solution $U_{\mathrm{ms},\ell}^{\mathrm{EW},n+1}$ from Algorithm 1. Using the coarse solution U_k^n as the initial condition, the fine solver (4.12) subsequently is used to calculate the fine solution U_k^{n+1} in parallel on each local time subinterval $[T^n, T^{n+1}]$. Then we calculate the discrepancy between the coarse solution and the fine solution in Step 8 on each discrete coarse time point T^n for $n = 1, 2, \cdots, M_\Delta$, and denote it as $\mathcal{S}(T^{n-1}, U_k^{n-1})$. Subsequently, this jump term is utilized in Step 9 to update the coarse solution via the global coarse solver (4.11). This process will be performed iteratively until certain tolerance on the jump terms is satisfied.

The multiscale solution in Algorithm 1 is calculated using a fine time step $\delta t \ll \Delta T$, which solves a linear system of size N_{ms} a number of $\frac{T}{\delta t}$ times, where N_{ms} denotes the dimension of the multiscale space $V_{\mathrm{ms},\ell}^{\mathrm{EW}}$. In each iteration, a linear system of size N_{ms} is solved a number of $\frac{T}{\Delta T}$ times sequentially and a number of $\frac{\Delta T}{\delta t}$ times simultaneously using $\frac{T}{\Delta T}$ processors. In total, Algorithm 2 involves solving a linear system of size N_{ms} a number of $k \times \frac{T}{\Delta T}$ times sequentially after k iteration. Algorithm 2 converges within a shorter wall-clock time compared with Algorithm 1 when $k \ll \frac{T}{\Delta T}$, which, indeed, can be supported by extensive numerical experiments presented in Section 4.4.

4.3 Convergence study

This section is concerned with the theoretical study of Algorithm 1 and Algorithm 2. The proof of the former follows from [60, Theorems 7.7 and 8.5], where the approximation properties of the multiscale space $V_{\mathrm{ms},\ell}^{\mathrm{EW}}$ and the convergence rate of the associated solver $\mathcal{L}_\ell^{-1}$ are needed. The error of the latter can be decomposed as the summation of the error from the multiscale space and the parareal error.

Algorithm 2 Wavelet-based Edge Multiscale Parareal (WEMP) Algorithm

Input: The initial data u_0, the source term f; tolerance ϵ; the level parameter $\ell \in \mathbb{N}$; coarse neighborhood ω_i and its four coarse edges $\Gamma_{i,j}$ with $j = 1, 2, 3, 4$, i.e., $\cup_{j=1}^{4} \Gamma_{i,j} = \partial\omega_i$; the subspace $V_{\ell,j}^{i} \subset L^2(\Gamma_{i,j})$ up to level ℓ on each coarse edge $\Gamma_{i,j}$.

Output: U.

1: Denote $V_{i,\ell} := \oplus_{k=1}^{4} V_{\ell,k}^{i}$. Then the number of basis functions in $V_{i,\ell}$ is $4 \times 2^{\ell} = 2^{\ell+2}$. Denote these basis functions as v_k for $k = 1, \cdots, 2^{\ell+2}$.

2: Calculate local multiscale basis $\mathcal{L}_i^{-1}(v_m)$ for all $m = 1, \cdots, 2^{\ell+2}$. Here, $\mathcal{L}_i^{-1}(v_m) := v$ satisfies: $\begin{cases} \mathcal{L}_i v := -\nabla \cdot (\kappa \nabla v) = 0 & \text{in } \omega_i, \\ v = v_m & \text{on } \partial\omega_i. \end{cases}$

3: Build global multiscale space. $V_{\mathrm{ms},\ell}^{\mathrm{EW}} := \mathrm{span}\{\chi_i \mathcal{L}_i^{-1}(v_k), \chi_i v^i : \ 1 \le i \le N, \ 1 \le k \le 2^{\ell+2}\}$.

4: $k = 0$, err= 1.

5: **while** err> ϵ **do**

6: Compute U_k^{n+1} for $n = 0, \cdots, M_\Delta - 1$:

$$U_k^{n+1} = E_\Delta^{\mathrm{ms},\ell}(T^n, U_k^n),$$
$$U_k^0 = \mathcal{P}_\ell u_0.$$

7: Compute u_k^{n+1} for $n = 0, \cdots, M_\Delta - 1$ on each local time subinterval [T^n,T^{n+1}]:

$$u_k^{n+1} = \mathcal{F}_\delta^{\mathrm{ms},\ell}(T^n, \Delta T, U_k^n).$$

8: Compute the jumps for $n = 1, \cdots, M_\Delta$:

$$\mathcal{S}(T^{n-1}, U_k^{n-1}) := u_k^n - U_k^n.$$

9: Compute the corrected coarse solutions U_{k+1}^{n+1} for $n = 0, \cdots, M_\Delta - 1$:

$$U_{k+1}^{n+1} = \mathcal{S}(T^n, U_k^n) + E_\Delta^{\mathrm{ms},\ell}(T^n, U_{k+1}^n),$$
$$U_{k+1}^0 = \mathcal{P}_\ell u_0.$$

10: Calculate:

$$\mathrm{err} := 1/M_\Delta \sum_{n=1}^{M_\Delta} \|U_{k+1}^n - U_k^n\|_{\ell_2}.$$

$k \leftarrow k + 1$

11: **end while**

12: $U_n := U_k^n$ and $U := [U_0, \cdots, U_{M_\Delta}]$.

4.3.1 Convergence for Algorithm 1

We first derive in the section the properties of the numerical operator that approximate the differential operator in Algorithm 1, then present the approximation properties of this numerical operator in the multiscale space $V_{\mathrm{ms},\ell}^{\mathrm{EW}}$.

Let $\mathcal{L} := -\nabla \cdot (\kappa \nabla \cdot)$ be the elliptic operator defined on V, and let its discrete operator $\mathcal{L}_\ell : V_{\mathrm{ms},\ell}^{\mathrm{EW}} \to L^2(D)$ be

$$(\mathcal{L}_\ell w_\ell, v_\ell) := (\mathcal{L} w_\ell, v_\ell) = a(w_\ell, v_\ell) \text{ for all } v_\ell \text{ and } w_\ell \in V_{\mathrm{ms},\ell}^{\mathrm{EW}}.$$

Then the inverse operator $\mathcal{L}_\ell^{-1}$ exists, which is self-adjoint, positive semi-definite on $L^2(D)$, and positive definite on $V_{\mathrm{ms},\ell}^{\mathrm{EW}}$. Further, let $\mathcal{R}_\ell$ be the Riesz operator associated to $\mathcal{L}$ in the multiscale space $V_{\mathrm{ms},\ell}^{\mathrm{EW}}$, i.e.,

$$\forall v \in V \text{ and } w_{\mathrm{ms}} \in V_{\mathrm{ms},\ell}^{\mathrm{EW}} : a(v - \mathcal{R}_\ell v, w_{\mathrm{ms}}) = 0.$$

Then it holds

$$\mathcal{L}_\ell^{-1} = \mathcal{R}_\ell \mathcal{L}^{-1}.$$

The approximation property of $\mathcal{L}_\ell^{-1}$ in energy norm is derived in [29, Proposition 5.2]. For any $v \in L^2(D)$, it holds

$$\left\| \mathcal{L}^{-1} v - \mathcal{L}_\ell^{-1} v \right\|_{H^1_\kappa(D)} \lesssim \left(H \left\| \widetilde{\kappa} \right\|_{L^\infty(D)}^{1/2} + 2^{-\ell/2} \left\| \kappa \right\|_{L^\infty(\mathcal{F}_H)} \right) \left\| v \right\|_{L^2(D)}. \tag{4.14}$$

Together with the duality argument, we can derive the approximation property of $\mathcal{L}_\ell^{-1}$ in $L^2(D)$-norm.

Lemma 4.3.1 (Approximation property of $\mathcal{L}_\ell^{-1}$ in $L^2(D)$-norm)**.** *For all $v \in L^2(D)$, there holds*

$$\left\|\mathcal{L}^{-1}v - \mathcal{L}_\ell^{-1}v\right\|_{L^2(D)} \lesssim \left(H\,\|\widetilde{\kappa}\|_{L^\infty(D)}^{1/2} + 2^{-\ell/2}\|\kappa\|_{L^\infty(\mathcal{F}_H)}\right)\left\|\mathcal{L}^{-1}v\right\|_{H^1_\kappa(D)}. \tag{4.15}$$

Proof. This assertion can be derived from the duality argument together with (4.14). Indeed, let $w \in V$ and $w_{\mathrm{ms}} \in V_{\mathrm{ms},\ell}^{\mathrm{EW}}$ satisfy

$$\mathcal{L}w = \mathcal{L}^{-1}v - \mathcal{L}_\ell^{-1}v$$
$$\mathcal{L}_\ell w_{\mathrm{ms}} = \mathcal{L}^{-1}v - \mathcal{L}_\ell^{-1}v.$$

Then it holds

$$\begin{aligned}\left\|\mathcal{L}^{-1}v - \mathcal{L}_\ell^{-1}v\right\|_{L^2(D)}^2 &= a(w - w_{\mathrm{ms}}, \mathcal{L}^{-1}v - \mathcal{L}_\ell^{-1}v)\\ &\le \|w - w_{\mathrm{ms}}\|_{H^1_\kappa(D)}\left\|\mathcal{L}^{-1}v - \mathcal{L}_\ell^{-1}v\right\|_{H^1_\kappa(D)}\end{aligned}$$

Together with the estimate (4.14), we derive

$$\begin{aligned}&\left\|\mathcal{L}^{-1}v - \mathcal{L}_\ell^{-1}v\right\|_{L^2(D)}^2\\ &\le \left(H\,\|\widetilde{\kappa}\|_{L^\infty(D)}^{1/2} + 2^{-\ell/2}\|\kappa\|_{L^\infty(\mathcal{F}_H)}\right)\left\|\mathcal{L}^{-1}v - \mathcal{L}_\ell^{-1}v\right\|_{L^2(D)}\left\|\mathcal{L}^{-1}v - \mathcal{L}_\ell^{-1}v\right\|_{H^1_\kappa(D)}.\end{aligned} \tag{4.16}$$

The stability of Riesz projection $\mathcal{R}_\ell$ implies

$$\left\|\mathcal{L}^{-1}v - \mathcal{L}_\ell^{-1}v\right\|_{H^1_\kappa(D)} \le \left\|\mathcal{L}^{-1}v\right\|_{H^1_\kappa(D)},$$

then together with (4.16), this shows the desired assertion. □

Using the properties of the discrete operator $\mathcal{L}_\ell$, we can obtain the error estimate of

Algorithm 1.

Proposition 4.3.2 (Pointwise-in-time error estimate in $L^2(D)$-norm for Algorithm 1)**.** *For all $m = 1, 2, \cdots, M_\delta$, there holds*

$$\begin{aligned}\left\|u(\cdot,t_m) - U_{ms,\ell}^{EW,m}\right\|_{L^2(D)} \lesssim & \Big(H\left\|\widetilde{\kappa}\right\|_{L^\infty(D)}^{1/2} + 2^{-\ell/2}\|\kappa\|_{L^\infty(\mathcal{F}_H)} + \delta t\Big) t_m^{-1}\left\|u_0\right\|_{L^2(D)} \\ & + \Big(H\left\|\widetilde{\kappa}\right\|_{L^\infty(D)}^{1/2} + 2^{-\ell/2}\|\kappa\|_{L^\infty(\mathcal{F}_H)}\Big) t_m \sup_{s\le t_m}|f(\cdot,s)|_2 \\ & + \delta t\Big(t_m \sup_{s\le t_m}|f(\cdot,s)|_2 + \int_0^{t_m}\|\partial_s f(\cdot,s)\|_{L^2(D)}\mathrm{d}s\Big). \end{aligned} \tag{4.17}$$

Proof. This result can be obtained from [60, Theorems 7.7 and 8.5]. □

4.3.2 Convergence for Algorithm 2

We present in this section the convergence analysis for Algorithm 2. To this end, we first prove the boundedness and Lipschitz continuity properties of the coarse solver $E_\Delta^{\mathrm{ms},\ell}$ and the jump operator $\mathcal{S}$ in the multiscale space $V_{\mathrm{ms},\ell}^{\mathrm{EW}}$:

Lemma 4.3.3. *For all $n \in \{1, \cdots, M_\Delta - 1\}$, the following properties hold.*

1. *The one step coarse solver $E_\Delta^{ms,\ell}$ is Lipschitz in $V_{ms,\ell}^{\mathrm{EW}}$. For all $v_1, v_2 \in V_{ms,\ell}^{\mathrm{EW}}$, there holds*

$$\left\|E_\Delta^{ms,\ell}(T^n, v_1) - E_\Delta^{ms,\ell}(T^n, v_2)\right\|_{L^2(D)} \le \|v_1 - v_2\|_{L^2(D)}.$$

2. *The jump operator $\mathcal{S}$ is an approximation of order 1 with Lipschitz regularity. For all $v_1, v_2 \in V_{ms,\ell}^{\mathrm{EW}} \cap \dot{H}^2(D)$ and any $\epsilon > 0$, there holds*

$$\|\mathcal{S}(T^n, v_1) - \mathcal{S}(T^n, v_2)\|_{L^2(D)} \lesssim \Delta T |v_1 - v_2|_{2+\epsilon}. \tag{4.18}$$

Proof. 1. Let $e_{\mathrm{ms}}^{n+1} := E_{\Delta}^{\mathrm{ms},\ell}(T^n, v_1) - E_{\Delta}^{\mathrm{ms},\ell}(T^n, v_2)$, then it holds

$$\forall w_{\mathrm{ms}} \in V_{\mathrm{ms},\ell}^{\mathrm{EW}} : \int_D e_{\mathrm{ms}}^{n+1} w_{\mathrm{ms}} \,\mathrm{d}x + \Delta T \int_D \kappa \nabla e_{\mathrm{ms}}^{n+1} \cdot \nabla w_{\mathrm{ms}} \,\mathrm{d}x = \int_D (v_1 - v_2) w_{\mathrm{ms}} \,\mathrm{d}x.$$

Choosing $w_{\mathrm{ms}} := e_{\mathrm{ms}}^{n+1}$ leads to

$$\left\| e_{\mathrm{ms}}^{n+1} \right\|_{L^2(D)}^2 + \Delta T \left\| e_{\mathrm{ms}}^{n+1} \right\|_{H_\kappa^1(D)}^2 = \int_D e_{\mathrm{ms}}^{n+1} (v_1 - v_2) \,\mathrm{d}x.$$

Finally an application of the Young's inequality proves the first assertion.

2. To prove the second assertion, let

$$\begin{aligned}
e_{\mathrm{ms}}^{n+1} &:= \mathcal{S}(T^n, v_1) - \mathcal{S}(T^n, v_2) \\
&= \left(\mathcal{F}_{\delta}^{\mathrm{ms},\ell}(T^n, \Delta T, v_1) - \mathcal{F}_{\delta}^{\mathrm{ms},\ell}(T^n, \Delta T, v_2) \right) - \left(E_{\Delta}^{\mathrm{ms},\ell}(T^n, v_1) - E_{\Delta}^{\mathrm{ms},\ell}(T^n, v_2) \right) \\
&= \bar{\mathcal{F}}_{\delta}^{\mathrm{ms},\ell}(T^n, \Delta T, v_1 - v_2) - \bar{E}_{\Delta}^{\mathrm{ms},\ell}(T^n, v_1 - v_2) \\
&= \left(\bar{\mathcal{F}}_{\delta}^{\mathrm{ms},\ell}(T^n, \Delta T, v_1 - v_2) - \bar{\mathcal{F}}^{\mathrm{ms},\ell}(T^n, \Delta T, v_1 - v_2) \right) - \left(\bar{E}_{\Delta}^{\mathrm{ms},\ell}(T^n, v_1 - v_2) \right. \\
&\quad \left. - \bar{\mathcal{F}}^{\mathrm{ms},\ell}(T^n, \Delta T, v_1 - v_2) \right) \\
&=: e_{\mathrm{ms},\delta}^{n+1} - e_{\mathrm{ms},\Delta}^{n+1}.
\end{aligned}$$

To estimate e_{ms}^{n+1}, we only need to derive the estimate for $e_{\mathrm{ms},\delta}^{n+1}$ and $e_{\mathrm{ms},\Delta}^{n+1}$, separately.

To this end, let $v_{\mathrm{ms},i}^{n+1} := v_{\mathrm{ms},i}(\cdot, T^{n+1}) := \mathcal{F}^{\mathrm{ms},\ell}(T^n, \Delta T, v_i)$ for $i = 1, 2$, we first construct the equation for $e_{\mathrm{ms},\Delta}^{n+1}$ by the definitions of the coarse solver (4.11) and fine solver (4.12). There holds

$$\forall w_{\mathrm{ms}} \in V_{\mathrm{ms},\ell}^{\mathrm{EW}} : \int_D e_{\mathrm{ms},\Delta}^{n+1} w_{\mathrm{ms}} \,\mathrm{d}x + \Delta T \int_D \kappa \nabla e_{\mathrm{ms},\Delta}^{n+1} \cdot \nabla w_{\mathrm{ms}} \,\mathrm{d}x = \int_D w_0 \cdot w_{\mathrm{ms}} \,\mathrm{d}x.$$

Here,

$$\begin{aligned} w_0 &:= \Delta T\Big(-\partial_t v_{\mathrm{ms},1}|_{t=T^{n+1}} + \frac{v_{\mathrm{ms},1}^{n+1} - v_1}{\Delta T} + \partial_t v_{\mathrm{ms},2}|_{t=T^{n+1}} - \frac{v_{\mathrm{ms},2}^{n+1} - v_1}{\Delta T}\Big) \\ &= -\int_{T^n}^{T^{n+1}} (s - T^n)\partial_{ss}(v_{\mathrm{ms},1} - v_{\mathrm{ms},2})(\cdot, s)\,\mathrm{d}s. \end{aligned}$$

Note that

$$\|w_0\|_{L^2(D)} \le \Delta T \int_{T^n}^{T^{n+1}} \|\partial_{ss}(v_{\mathrm{ms},1} - v_{\mathrm{ms},2})(\cdot, s)\|_{L^2(D)}\,\mathrm{d}s.$$

An adaptation of the proof to [60, Lemma 3.2] shows

$$\|\partial_{tt}(v_{\mathrm{ms},1} - v_{\mathrm{ms},2})(\cdot, t)\|_{L^2(D)} \lesssim (t - T^n)^{-1+\epsilon/2}|v_1 - v_2|_{2+\epsilon} \text{ for all } t > 0.$$

Consequently, we derive

$$\|w_0\|_{L^2(D)} \lesssim \Delta T|v_1 - v_2|_{2+\epsilon}.$$

Choosing $w_{\mathrm{ms}} := e_{\mathrm{ms},\Delta}^{n+1}$ leads to

$$\left\|e_{\mathrm{ms},\Delta}^{n+1}\right\|_{L^2(D)}^2 + \Delta T \left\|e_{\mathrm{ms},\Delta}^{n+1}\right\|_{H_\kappa^1(D)}^2 = \int_D e_{\mathrm{ms},\Delta}^{n+1} w_0\,\mathrm{d}x.$$

Consequently, an application of the Young's inequality implies

$$\left\|e_{\mathrm{ms},\Delta}^{n+1}\right\|_{L^2(D)} \lesssim \Delta T|v_1 - v_2|_{2+\epsilon}.$$

Analogously, we can obtain the estimate for $e_{\mathrm{ms},\delta}^{n+1}$, which reads

$$\left\|e_{\mathrm{ms},\delta}^{n+1}\right\|_{L^2(D)} \le \delta t|v_1 - v_2|_{2+\epsilon}.$$

Note that $\delta t \ll \Delta T$, then a combination of the two estimates above with the triangle

inequality, shows the second assertion. □

Remark 4.3.4. *Lemma 4.3.3 indicates that the approximation property of the jump operator $\mathcal{S}(T^n,\cdot)$ deteriorates when T^n is small.*

We present in the next theorem the convergence rate of Algorithm 2 to Problems (4.1) in pointwise-in-time in $L^2(D)$-norm. To derive it, we first decompose the error from Algorithm 2 as a summation of the error from WEMsFEM and the error from parareal algorithm. Then we estimate the former by Proposition 4.3.2, and the latter can be estimated by mathematical induction. This result relies on the followiing assumption.

Assumption 4.3.1. *Let m be a positive integer such that $T^n = t_m$ for some integer n. For $\epsilon > 0$ be sufficiently small, we assume the following inequality holds*

$$|U_{ms,\ell}^{EW,m} - U_k^n|_{2+\epsilon} \lesssim (T^n)^{-1} \left\| U_{ms,\ell}^{EW,m} - U_k^n \right\|_{L^2(D)}.$$

We remark here this assumption is provable for the continuous problem [60, Lemma 3.2].

Theorem 4.3.5. *[Pointwise-in-time error estimate in $L^2(D)$-norm for Algorithm 2] Let Assumptions 4.1.1 and 4.3.1 hold. Assume that the source term $f \in L^\infty([0,T];\dot{H}^2(D))$ satisfying $\partial_t f \in L^1([0,T];L^2(D))$ and initial data $u_0 \in L^2(D)$. Let $\ell \in \mathbb{N}_+$ be the level parameter. The coarse time step size and fine time step size are ΔT and δt. Let $u(\cdot,t) \in V$ be the solution to Problem* (4.1) *and let U_k^n be the solution from Algorithm 2 with iteration*

$k \in \mathbb{N}$. *There holds*

$$\begin{aligned}\|u(\cdot,T^n)-U_k^n\|_{L^2(D)} \lesssim & \Big(H\,\|\widetilde{\kappa}\|_{L^\infty(D)}^{1/2}+2^{-\ell/2}\|\kappa\|_{L^\infty(\mathcal{F}_H)}+\delta t\Big)\frac{1}{T^n}\,\|u_0\|_{L^2(D)}\\ &+\Big(H\,\|\widetilde{\kappa}\|_{L^\infty(D)}^{1/2}+2^{-\ell/2}\|\kappa\|_{L^\infty(\mathcal{F}_H)}\Big)T^n\sup_{s\le T^n}|f(\cdot,s)|_2\\ &+\delta t\Big(T^n\sup_{s\le T^n}|f(\cdot,s)|_2+\int_0^{T^n}\|\partial_s f(\cdot,s)\|_{L^2(D)}\mathrm{d}s\Big)\\ &+\Big(\Pi_{j=0}^k\frac{1}{T^{n-j}}\Big)\Delta T^{k+1}\,\|u_0\|_{L^2(D)}\,.\end{aligned}$$

Proof. We first define the multiscale solution to Problem (4.1) using Algorithm 1. Find $U_{\mathrm{ms},\ell}^{\mathrm{EW},m}\in V_{\mathrm{ms},\ell}^{\mathrm{EW}}$ for $m=1,\cdots,M_\delta$, satisfying

$$\begin{aligned}\forall w_{\mathrm{ms}}\in V_{\mathrm{ms},\ell}^{\mathrm{EW}}:\big(\frac{U_{\mathrm{ms},\ell}^{\mathrm{EW},m}-U_{\mathrm{ms},\ell}^{\mathrm{EW},m-1}}{\delta t},w_{\mathrm{ms}}\big)_D+a(U_{\mathrm{ms},\ell}^{\mathrm{EW},m},w_{\mathrm{ms}})&=(f(\cdot,t_m),w_{\mathrm{ms}})_D\\ U_{\mathrm{ms},\ell}^{\mathrm{EW},0}&=I_\ell(u_0).\end{aligned}\tag{4.19}$$

Then we only need to estimate $\big\|u(\cdot,T^n)-U_{\mathrm{ms},\ell}^{\mathrm{EW},m}\big\|_{L^2(D)}$ and $\big\|U_{\mathrm{ms},\ell}^{\mathrm{EW},m}-U_k^n\big\|_{L^2(D)}$ for $m:=\Delta T/\delta t\times n$. Note that $T^n=t_m$. Therefore, we can replace T^n with t_m. Similarly, let $m':=\Delta T/\delta t\times(n-1)$, then it holds $t_{m'}=T^{n-1}$. The first term $\big\|u(\cdot,t_m)-U_{\mathrm{ms},\ell}^{\mathrm{EW},m}\big\|_{L^2(D)}$ can be estimated by Proposition 4.3.2. The second term $\big\|U_{\mathrm{ms},\ell}^{\mathrm{EW},m}-U_k^n\big\|_{L^2(D)}$ corresponds to the error induced by parareal algorithm in the multiscale method, and we will prove by mathematical induction:

$$e_k^n:=\big\|U_{\mathrm{ms},\ell}^{\mathrm{EW},m}-U_k^n\big\|_{L^2(D)}\lesssim\Big(\Pi_{j=0}^k\frac{1}{T^{n-j}}\Big)\Delta T^{k+1}\,\|u_0\|_{L^2(D)}\,.\tag{4.20}$$

Obviously, the inequality (4.20) holds when $k=0$. Assume that it holds for iteration k for some $k\in\mathbb{N}_+$. We will show that it holds for the next iteration $k+1$.

We can obtain from Algorithm 2:

$$e_{k+1}^{n} = \mathcal{S}(T^{n-1}, U_{\mathrm{ms},\ell}^{\mathrm{EW},m'}) - \mathcal{S}(T^{n-1}, U_{k}^{n-1}) + E_{\Delta}^{\mathrm{ms},\ell}(T^{n-1}, U_{\mathrm{ms},\ell}^{\mathrm{EW},m'}) - E_{\Delta}^{\mathrm{ms},\ell}(T^{n-1}, U_{k+1}^{n-1}).$$

Consequently, an application of Lemma 4.3.3 and Assumption 4.3.1 lead to

$$\begin{aligned}
e_{k+1}^{n} &\leq \left\| \mathcal{S}(T^{n-1}, U_{\mathrm{ms},\ell}^{\mathrm{EW},m'}) - \mathcal{S}(T^{n-1}, U_{k}^{n-1}) \right\|_{L^2(D)} + \left\| E_{\Delta}^{\mathrm{ms},\ell}(T^{n-1}, U_{\mathrm{ms},\ell}^{\mathrm{EW},m'}) - E_{\Delta}^{\mathrm{ms},\ell}(T^{n-1}, U_{k+1}^{n-1}) \right\|_{L^2(D)} \\
&\lesssim \Delta T |U_{\mathrm{ms},\ell}^{\mathrm{EW},m'} - U_{k}^{n-1}|_{2+\epsilon} + \left\| U_{\mathrm{ms},\ell}^{\mathrm{EW},m'} - U_{k+1}^{n-1} \right\|_{L^2(D)} \\
&\lesssim \frac{\Delta T}{T^n} |U_{\mathrm{ms},\ell}^{\mathrm{EW},m'} - U_{k}^{n-1}|_{2} + \left\| U_{\mathrm{ms},\ell}^{\mathrm{EW},m'} - U_{k+1}^{n-1} \right\|_{L^2(D)} \\
&= \frac{\Delta T}{T^n} e_{k}^{n-1} + e_{k+1}^{n-1}.
\end{aligned}$$

Note that $e_{n}^{k+1} = 0$ for all $n \leq k+1$. We can obtain

$$\begin{aligned}
e_{k+1}^{k+2} &\lesssim \frac{\Delta T}{T^{k+2}} \Big(\Pi_{j=0}^{k} \frac{1}{T^{k+1-j}} \Big) \Delta T^{k+1} \left\| u_0 \right\|_{L^2(D)} \\
&= \Big(\Pi_{j=0}^{k+1} \frac{1}{T^{k+2-j}} \Big) \Delta T^{k+2} \left\| u_0 \right\|_{L^2(D)}.
\end{aligned}$$

In a similar manner, we can utilize mathematical induction to prove

$$e_{k+1}^{k+m} \lesssim \Big(\Pi_{j=0}^{k+1} \frac{1}{T^{k+m-j}} \Big) \Delta T^{k+2} \left\| u_0 \right\|_{L^2(D)},$$

and this proves estimate (4.20) corresponding to the case with iteration $k+1$. Hence, the estimate (4.20) is proved. Finally, a combination with (4.17) results in the desired estimate. □

Theorem 4.3.5 indicates that the pointwise-in-time error estimate of Algorithm 2 to Problems (4.1) in $L^2(D)$-norm will deteriorate when the time step approaches the original $t = 0$. This blow-up of error is produced by the parareal algorithm (Step 2 in the proof

to Theorem 4.3.5), which essentially arises from the approximation property of the jump operator (4.18). This estimate can be improved to

$$\|\mathcal{S}(T^n, v_1) - \mathcal{S}(T^n, v_2)\|_{L^2(D)} \leq (\Delta T)^2 |v_1 - v_2|_4, \tag{4.21}$$

when $v_1, v_2 \in \dot{H}_4(D)$.

However, the estimate above has different norms on both sides of the inequality. This makes the argument in Step 2, proof to Theorem 4.3.5 invalid.

Remark 4.3.6. *Algorithm 2 outweighs Algorithm 1 only when the former achieves similar accuracy to the latter within a very few iteration $k \ll M_\Delta$. Therefore, we are not interested in the case when $k \geq M_\Delta$ or the error at time level T^n with $k \geq n$.*

4.4 Numerical results

In this section, we perform a series of numerical experiments to demonstrate the performance of the proposed WEMP Algorithm. In particular, we compare the performance of Algorithms 1 and 2 for each experiment. Furthermore, we investigate whether replacing backward Euler scheme by Crank-Nicolson scheme would reduce the iteration number. Motivated by the critical condition proposed in [63], we choose different values of $\frac{\Delta T}{\delta t}$ to test how they will influence the iteration number. It can be seen from Equation (4.19) that WEMP Algorithm would generate a solution of better accuracy when the source term being 0. In the last subsection, we conduct experiments to verify this.

We consider the parabolic equation (4.1) in the space domain $D := [0, 1]^2$ and the time domain $[0, T] = [0, 1]$. The permeability coefficient κ we choose has two distinct value: 1 and 1000. It is high-contrast and heterogenous. We refer to Figure 4.2 (left figure) for an illustration. The initial data tested in our numerical experiments is chosen to be a smooth function $u_0 := x(1 - x)y(1 - y)$. We refer to Figure 4.2 (right figure) for an illustration.

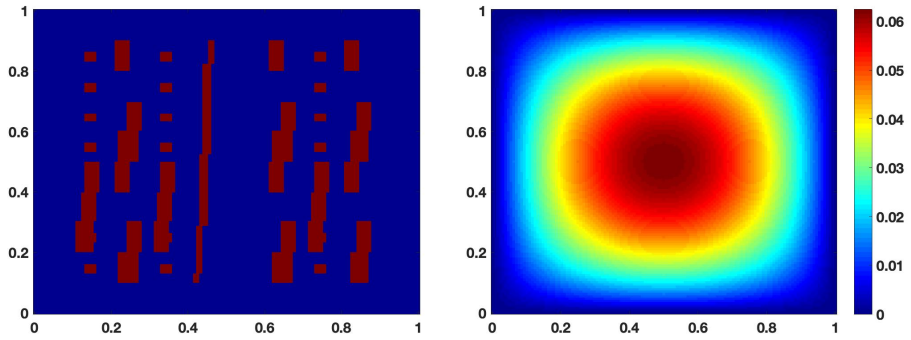

Figure 4.2: The heterogeneous permeability field κ and the initial data $u_0 = x(1-x)y(1-y)$.

Let $\mathcal{T}_H$ be a decomposition of the domain D into non-overlapping shape-regular rectangular elements with maximal mesh size $H := 2^{-4}$. These coarse rectangular elements are further partitioned into a collection of connected fine rectangular elements $\mathcal{T}_h$ using fine mesh size $h := 2^{-7}$. Similarly, we define V_h to be a conforming piecewise affine finite element associated with $\mathcal{T}_h$. In our numerical experiments, space meshes $\mathcal{T}_H$ and $\mathcal{T}_h$ are fixed. To keep our presentation concise, we will only present the numerical results with a fixed level parameter $\ell := 2$. The temporal discretization is presented in Section 4.1 with $T := 1$. The coarse time step size and fine time step size are ΔT and δt. Note that $\delta t \ll \Delta T$.

We introduce the following notations to calculate the errors. The relative errors for the multiscale solution in $L^2(D)$-norm and $H^1_\kappa(D)$-norm are

$$\text{Rel}^{\text{EW}}_{L^2}(t_m) := \frac{\left\|U^m_h - U^{\text{EW},m}_{\text{ms},\ell}\right\|_{L^2(D)}}{\left\|U^m_h\right\|_{L^2(D)}} \times 100 \quad \text{and} \quad \text{Rel}^{\text{EW}}_{H^1_\kappa}(t_m) := \frac{\left\|U^m_h - U^{\text{EW},m}_{\text{ms},\ell}\right\|_{H^1_\kappa(D)}}{\left\|U^m_h\right\|_{H^1_\kappa(D)}} \times 100.$$

Analogously, the relative errors for our proposed algorithm with iteration $k \in \mathbb{N}$ in $L^2(D)$-

norm and $H^1_\kappa(D)$-norm are

$$\mathrm{Rel}^k_{L^2}(T^n) := \frac{\left\|U^m_h - U^n_k\right\|_{L^2(D)}}{\left\|U^m_h\right\|_{L^2(D)}} \times 100 \quad \text{and} \quad \mathrm{Rel}^k_{H^1_\kappa}(T^n) := \frac{\left\|U^m_h - U^n_k\right\|_{H^1_\kappa(D)}}{\left\|U^m_h\right\|_{H^1_\kappa(D)}} \times 100$$

with $m := \Delta T/\delta t \times n$.

Our numerical experiments include testing nonzero source term in section 4.4.1 and zero source term in section 4.4.2. We investigate the influence of different tempo discretization schemes, e.g., backward Euler scheme and Crank-Nicolson Galerkin scheme, on the performance of our algorithm.

4.4.1 Numerical tests with nonzero source term

To define nonzero source term, we take time-dependent smooth function

$$f(x, y, t) := 200\pi^2 \sin(\pi x) \sin(\pi y) \sin(10\pi t x).$$

Since there is no analytic solution to system (4.1), we need to find an approximation of the exact solutions. To this end, we take time step size $\delta t = 10^{-4}$ and use backward Euler Galerkin Method in (4.4) to obtain the reference solutions U^n_h. Note that we use a much finer time step size to simulate the reference solution. We plot the reference solutions U^n_h for $n = 10^3, 3 \times 10^3, 5 \times 10^3$ and 10^4 in Figure 4.3.

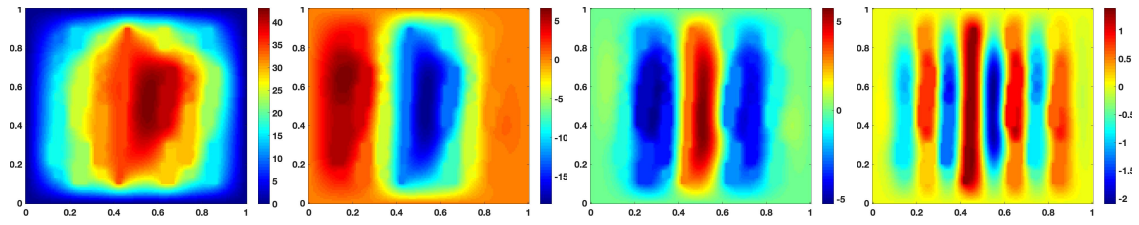

Figure 4.3: Numerical solution U^n_h to (4.4) for $n = 10^3, 3 \times 10^3, 5 \times 10^3$ and 10^4 with $\delta t = 10^{-4}$.

In the rest of this subsection, we will present numerical tests using backward Euler scheme with $\frac{\Delta T}{\delta t} = 100$ in Experiment 1, Crank-Nicolson scheme with $\frac{\Delta T}{\delta t} = 100$ in Experiment 2 and backward Euler scheme with $\frac{\Delta T}{\delta t} = 10$ in Experiment 3. For all the three experiments, our proposed algorithm, i.e. Algorithm 2, can generate numerical solutions by a few iterations at least of the same accuracy as the multiscale solutions from Algorithm 1. For the brevity of the chapter, we only present numerical solutions U_k^n from Algorithm 2 with iteration number $k = 0, 1$ and 2 and multiscale solutions $U_{\text{ms},\ell}^{\text{EW},m}$ from Algorithm 1 in Experiment 1.

Experiment 1: Backward Euler with $\frac{\Delta T}{\delta t} = 100$

We test in this experiment the performance of Algorithm 2 with a fine time step size $\delta t = 10^{-3}$ and a coarse time step size $\Delta T = 0.1$. The backward Euler scheme is utilized for the time discretization.

We present the numerical solutions U_k^n for $n = 1, 3, 5, 10$ from Algorithm 2 with iteration number $k = 0, 1$ and 2 in Figure 4.4. One can observe that U_k^n converges to the multiscale solution $U_{\text{ms},\ell}^{\text{EW},n}$ as the iteration k increases.

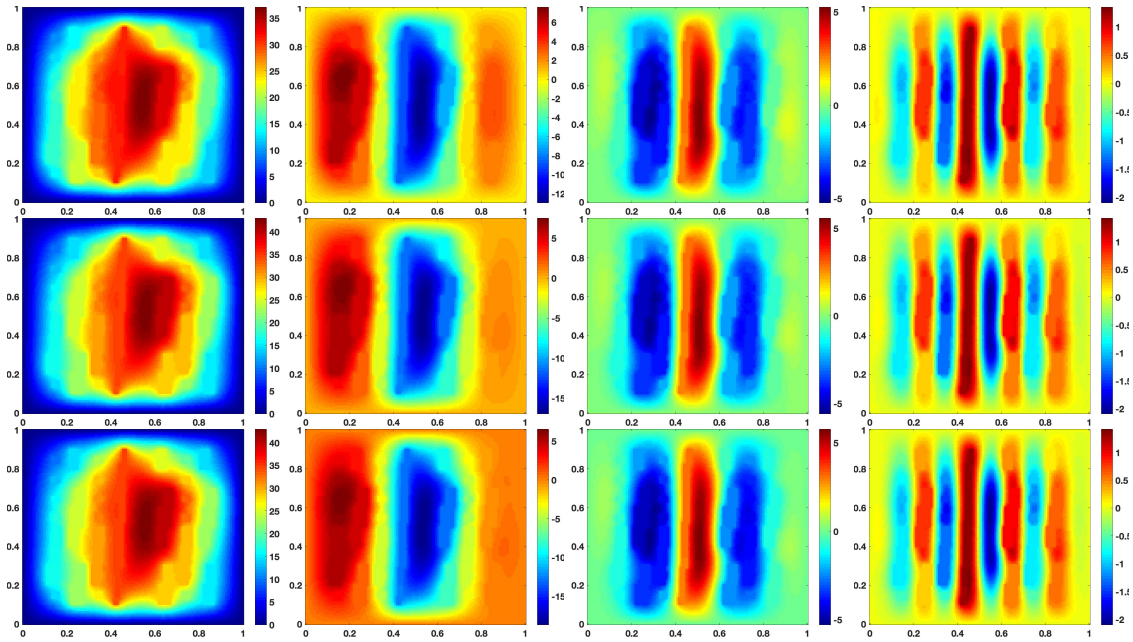

Figure 4.4: Numerical solutions U_k^n for $n = 1, 3, 5, 10$ from Algorithm 2 with $\Delta T = 0.1$ and $\delta t = 10^{-3}$, backward Euler scheme: iteration number $k = 0$ (top), $k = 1$ (middle) and $k = 2$ (bottom).

The convergence history of Algorithm 2 in relative $L^2(D)$-norm and relative $H^1_\kappa(D)$-norm are presented in Figure 4.5.

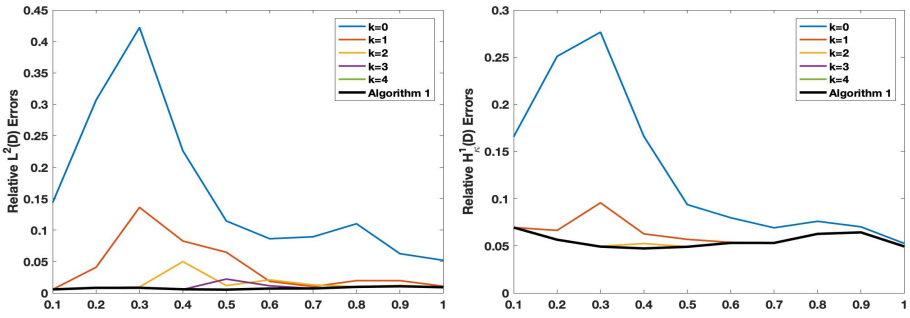

Figure 4.5: Convergence history of Algorithm 2 in relative $L^2(D)$ error and relative $H^1_\kappa(D)$ error for Experiment 1: backward Euler scheme with $\Delta T = 0.1$ and $\delta t = 10^{-3}$.

One observes from Figure 4.5 that 4 iterations are sufficient for Algorithm 2 to attain the same accuracy as Algorithm 1 for all discrete time steps under the $L^2(D)$-norm, while 2 iterations under the $H^1_\kappa(D)$-norm. Each iteration involves a number of $1/\Delta T = 10$ sequential solver and $\Delta T/\delta t = 100$ parallel solver. In comparison, Algorithm 1 involves a number of $1/\delta t = 1000$ sequential solver. Consequently, Algorithm 2 involves much less wall clock time with the aid of a sufficient number of processors.

Experiment 2: Crank-Nicolson with $\frac{\Delta T}{\delta t} = 100$

Since the backward Euler scheme is only first order accurate, a higher order accurate scheme can improve the performance of Algorithm 1 and Algorithm 2. This can be seen from the proof to Theorem 4.3.5. In this section, we will present the numerical tests with Crank-Nicolson scheme for both algorithms.

A direct application of Crank-Nicolson scheme as a time discretization fails to maintain second order accuracy due to the possible blow up of the eigenvalues of the elliptic operator $-\nabla \cdot (\kappa \nabla \cdot)$ when $\eta_{\max} \to \infty$. To improve its performance and maintain second order convergence rate, we use 3 steps of backward Euler scheme before Crank-Nicolson scheme kicks in [50, 60].

The convergence history of Algorithm 2 in $L^2(D)$-norm and $H^1_\kappa(D)$-norm is presented in Figure 4.6. Similar to Experiment 1, we observe that 4 iterations are sufficient for Algorithm 2 to reach the same accuracy as Algorithm 1 at all discrete time levels under the $L^2(D)$-norm, while 2 iterations under the $H^1_\kappa(D)$-norm. Comparing Figure 4.5 with Figure 4.6, one observes that Algorithm 2 with Crank-Nicolson scheme outperforms that with backward Euler scheme under $L^2(D)$-norm.

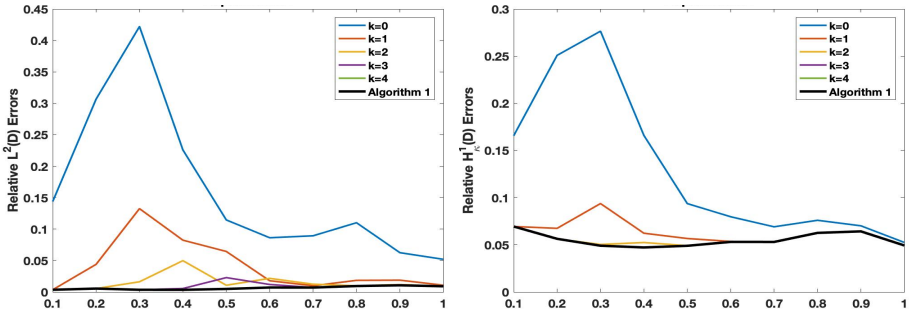

Figure 4.6: Convergence history of Algorithm 2 in relative $L^2(D)$-norm and relative $H^1_\kappa(D)$-norm for Experiment 2: Crank-Nicolson scheme with $\Delta T = 0.1$ and $\delta t = 10^{-3}$.

Experiment 3: backward Euler with $\frac{\Delta T}{\delta t} = 10$

We are also interested in studying how the coarse solver and fine solver affect the performance of our proposed WEMP algorithm. To this end, we choose $\Delta T = 10^{-2}$, $\delta t = 10^{-3}$ and utilize backward Euler scheme in time discretization. Note that the ratio between the coarse time step and fine time step is smaller than that in Experiment 1.

The convergence history of Algorithm 2 in $L^2(D)$-norm and $H^1_\kappa(D)$-norm is presented in Figure 4.7. Comparing Figure 4.5 with Figure 4.7, one can see 1 iteration is sufficient for the numerical solutions from Algorithm 2 to reach the same accuracy as multiscale solutions from Algorithm 1 under $L^2(D)$-norm and $H^1_\kappa(D)$-norm when the coarse time step $\Delta T = 10^{-2}$ becomes smaller. However, this involves more coarse solvers for each iteration. Furthermore, a decreased coarse time step is only practical when sufficient processors are available.

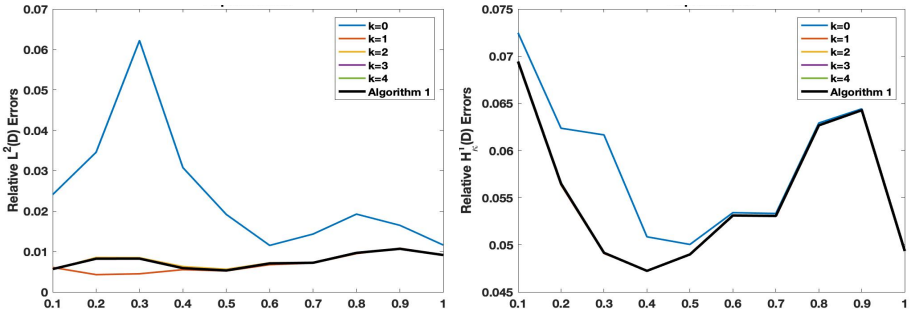

Figure 4.7: Convergence history of Algorithm 2 in relative $L^2(D)$ error and relative $H^1_\kappa(D)$ error for Experiment 3: backward Euler scheme with $\Delta T = 10^{-2}$ and $\delta t = 10^{-3}$.

4.4.2 Numerical tests with a vanishing source term

To avoid the complicated requirement on the source term in Theorem 4.3.5, we test in this section the performance of Algorithm 2 for Problem (4.1) with a vanishing source term $f := 0$ using a backward Euler scheme and Crank-Nicolson scheme. Consequently, the solution decays rapidly to 0. To generate solutions with reasonable size, we set the final time $T = 0.1$, the coarse time step $\Delta T := 10^{-2}$ and the fine time step $\delta t = 10^{-3}$. The initial data and permeability are the same as in the previous section. We use backward Euler scheme with time step 10^{-3} to obtain the reference solutions U_h^n. The reference solutions U_h^n for $n = 10, 30, 50, 100$ are plotted in Figure 4.8.

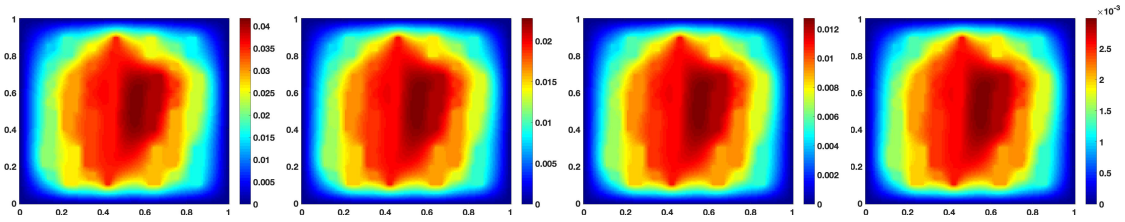

Figure 4.8: Numerical solutions U_h^n to (4.4) with $f = 0$ for $n = 10, 30, 50$ and 100 with $\delta t = 10^{-3}$.

We present the numerical solutions U_k^n for $n = 1, 3, 5, 10$ from Algorithm 2 with iteration number $k = 0, 1, 2$ in Figure 4.9. One can observe U_k^n converges to $U_{\text{ms},\ell}^{\text{EW},n}$ as the iteration number k increases.

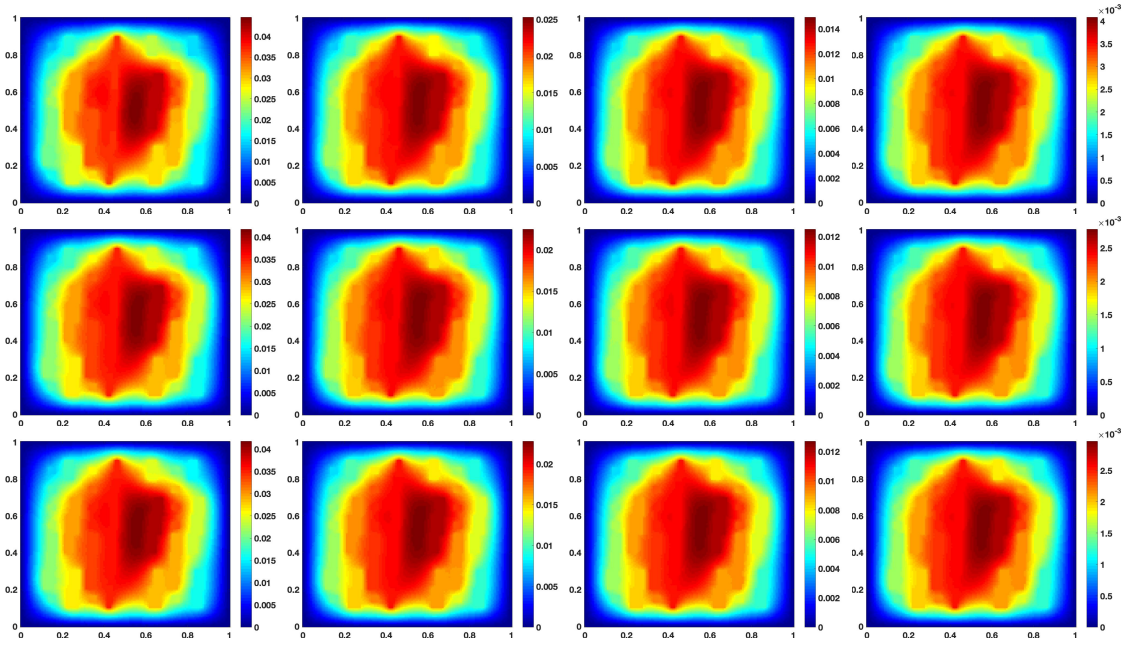

Figure 4.9: Numerical solutions U_k^n for $n = 1, 3, 5, 10$ from Algorithm 2 with $\Delta T = 10^{-2}$ and $\delta t = 10^{-3}$, backward Euler scheme: iteration number $k = 0$ (top), $k = 1$ (middle) and $k = 2$ (bottom).

The convergence history of Algorithm 2 in $L^2(D)$-norm and $H^1_\kappa(D)$-norm is presented in Figure 4.10 From the figure, one can see that 1 iteration is sufficent for the numerical solutions from Algorithm 2 with backward Euler to converge under $L^2(D)$-norm and $H^1_\kappa(D)$-norm. We can conclude that our proposed algorithm with backward Euler scheme is effective in solving Problem (4.1) with zero source term.

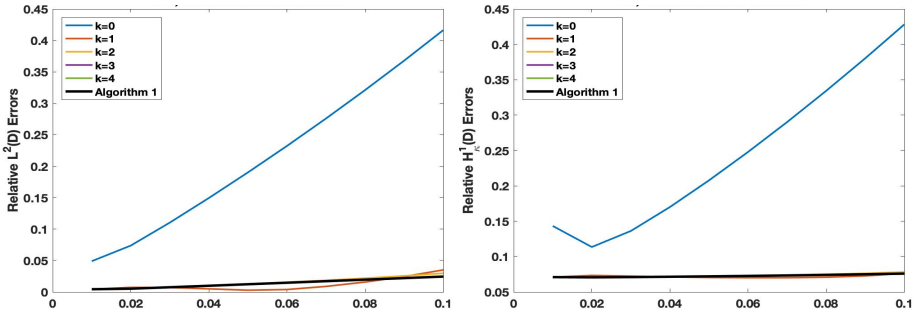

Figure 4.10: Convergence history of Algorithm 2 in relative $L^2(D)$-norm and relative $H^1_\kappa(D)$-norm for $f = 0$: backward Euler scheme with $\Delta T = 10^{-2}$ and $\delta t = 10^{-3}$.

Our last experiment is replacing backward Euler scheme by Crank-Nicolson scheme for the above problem. We observe the same convergence behavior as in the previous experiment that the numerical solutions from Algorithm 2 converges to the multiscale solutions from Algorithm 1. For brevity of the chapter, we do not present these figures.

The convergence history of Algorithm 2 in $L^2(D)$-norm and $H^1_\kappa(D)$-norm is presented in Figure 4.11. One observes that it takes 4 iterations to converge under $L^2(D)$-norm and 3 iterations to converge under $H^1_\kappa(D)$-norm when using Algorithm 2 with Crank-Nicolson scheme. Comparing Figure 4.10 with Figure 4.11, we can see that Algorithm 2 with the Crank-Nicolson scheme yields a better accuracy than that with the backward Euler scheme. We conclude that Algorithm 2 with backward Euler scheme converges faster than that with Crank-Nicolson scheme, while Algorithm 2 with Crank-Nicolson scheme generate solutions with a higher accuracy for Problem (4.1) with zero source term.

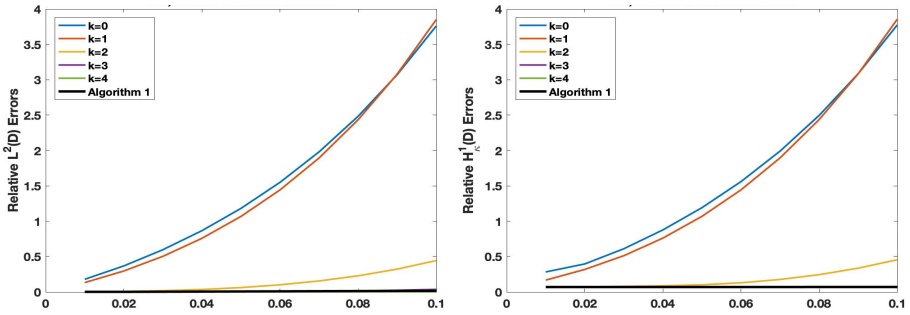

Figure 4.11: Convergence history of Algorithm 2 in relative $L^2(D)$-norm and relative $H^1_\kappa(D)$-norm for $f = 0$: Crank-Nicolson scheme with $\Delta T = 10^{-2}$ and $\delta t = 10^{-3}$.

5. SPACE-TIME NON-LOCAL MULTI-CONTINUA UPSCALING FOR PARABOLIC EQUATIONS WITH MOVING CHANNELIZED MEDIA

In the chapter, we will develop and analyze a novel multiscale method for parabolic problems with time-dependent multiscale coefficients. Our approach is based on Non-local multi-continua (NLMC) upscaling method and space-time finite element method. We assume that one knows each separate channel within each space-time coarse block and follow a general procedure in [19] to construct a multiscale basis functions. NLMC identifies space-time multi-continua parameters and defines a piece-wise constant functions as local auxiliary functions. Next, multiscale basis functions are sought in the oversampled region subject to a constraint that the minimizer is orthogonal to the auxiliary space. These multiscale functions are shown to decay exponentially outside the corresponding local oversampling regions. This exponential decay property plays a vital role in the convergence of the proposed method and justifies the use of the local multiscale basis functions. In this paper, we construct local space-time ansatz spaces to approximate the global space-time ansatz spaces.

The remainder of this chapter is organized as follows. In Section 5.1, we introduce the parabolic model problem, standard space-time weak formulation and functional spaces that will be used in this work. We develop local and global NLMC upscaling method in Section 5.2. Convergence analysis of our proposed method is studied in Section 5.3. We present numerical experiments in Section 5.4 to demonstrate the approximation of our local and global multiscale spaces.

5.1 Problem setting

In this section, we present some preliminaries of the model problems and introduce the necessary notations. Our aim is to develop an efficient numerical upscaling method

for parabolic problems with time-dependent heterogeneous coefficients. Let $\Omega \subset \mathbb{R}^d$ ($d \in \{2, 3\}$) be a bounded domain with a sufficiently smooth boundary $\partial\Omega$. Let $T > 0$ be a given positive terminal time. We seek a function $u = u(t, x)$ such that it solves the following initial boundary value problem:

$$\begin{cases} \partial_t u(t,x) - \nabla \cdot (\kappa(t,x)\nabla u(t,x)) = f(t,x), & (t,x) \in (0,T) \times \Omega, \\ u(0,x) = 0, & x \in \Omega, \\ u(t,x) = 0, & (t,x) \in [0,T] \times \partial\Omega, \end{cases} \tag{5.1}$$

where $\kappa = \kappa(t, x)$ is a high-contrast time-dependent permeability field and $f \in L^2(0, T; L^2(\Omega))$ is a source function. We assume that there exist two positive constants κ_0 and κ_1 such that $0 < \kappa_0 \leq \kappa(t, x) \leq \kappa_1$ for any $(t, x) \in \Omega_T := [0, T] \times \Omega$.

In this work, we will mainly focus on the case when κ is a so-called channelized-moving medium. In particular, we assume that κ is a piecewise constant function such that

$$\kappa(x,t) = \begin{cases} \kappa_m & \text{if } (t,x) \in D_m, \\ \kappa_i & \text{if } (t,x) \in D_{c,i}, \end{cases}$$

where κ_m and κ_i are two positive constants between κ_0 and κ_1 such that the ratio κ_i/κ_m is very large. Here, the space-time domain Ω_T is divided into two non-overlapping sets of regions in $\mathbb{R}^{d+1}$ with

$$\Omega_T = D_m \bigcup_{i=1}^{\mathcal{I}_c} D_{c,i}.$$

The set D_m is called the matrix region of the coefficient κ; $D_{c,i}$ is called the i-th channel of the coefficient κ and $\mathcal{I}_c$ is the total number of channels in the coefficient κ. In practice, the space-time *volume* of the matrix D_m is much larger than that of the channelized region $D_c := \bigcup_{i=1}^{\mathcal{I}_c} D_{c,i}$.

5.1.1 Space-time variational formulation and space-time discretization

Let $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_d)$ be a multi-index with non-negative integers α_i for $i = 1, 2, \cdots, d$. We use $|\alpha|$ to denote the sum of its elements, that is, $|\alpha| = \sum_{i=1}^{d} \alpha_i$. For non-negative integers l and k, we define a Sobolev space on the space-time domain Ω_T as $H^{l,k}(\Omega_T) := \{u \in L^2(\Omega_T) : \partial_x^\alpha u \in L^2(\Omega_T) \text{ for all } \alpha \text{ with } 0 \leq |\alpha| \leq l, \text{ and } \partial_t^i u \in L^2(\Omega_T) \text{ for } i = 0, 1, \cdots, k\}$. Moreover, we define $H_0^{1,0}(\Omega_T) = \{u \in H^{1,0}(\Omega_T) : u(t,x) = 0 \text{ for } x \in \partial\Omega\}$ and $H_{0,0}^{1,1}(\Omega_T) := \{u \in H^{1,1}(\Omega_T) : u(t,x) = 0 \text{ for } x \in \partial\Omega, \text{ and } u(0,x) = 0 \text{ for } x \in \Omega\}$. The weak space-time variational formulation of (5.1) reads as follows: find $u \in H_{0,0}^{1,1}(\Omega_T)$ such that

$$b(u,v) + a(u,v) = (f,v) \ \forall v \in H_0^{1,0}(\Omega_T), \tag{5.2}$$

where $b(u,v) = \int_{\Omega_T} \partial_t u v$, $a(u,v) = \int_{\Omega_T} \kappa \nabla_x u \nabla_x v$ and $(f,v) = \int_{\Omega_T} f v$.

To discretize the variational problem (5.2), let $\mathcal{T}_H$ be a partition of space domain Ω into non-overlapping shape-regular rectangular elements with maximal mesh size H. The time domain $(0,T]$ is partitioned into $\mathcal{T}_{\Delta t} = \{(t_i, t_{i+1}]\}_{i=0}^{N_T-1}$ with the maximal temporal mesh size $\Delta t := \max_{0 \leq i \leq N_T-1}\{t_{i+1} - t_i\}$. A space-time coarse element $K^{(n,i)}$ is then defined by $(t_n, t_{n+1}] \times K^i$ for $K^i \in \mathcal{T}_H$ and $(t_n, t_{n+1}] \in \mathcal{T}_{\Delta T}$. Furthermore, let $\mathcal{T}_h$ be a refinement of $\mathcal{T}_H$ and $\mathcal{T}_{\delta t}$ a refinement of $\mathcal{T}_{\Delta T}$.

For each coarse space element K^i, we define the oversampled region $K_{k_i}^i \subseteq \Omega$ by enlarging K^i by $k_i \in \mathbb{N}$ layer(s), i.e.,

$$K_0^i := K_i, \quad K_{k_i}^i := \bigcup \{K \in \mathcal{T}_H : K \cap K_{k_i-1}^i \neq \varnothing\} \quad \text{for } k_i = 1, 2, \cdots.$$

For simplicity, we denote K_+^i a generic oversampling region related to the coarse element K^i with a specific oversampling parameter k_i. See Figure 5.1 for an illustration of K_1^i. For each space-time coarse element $K^{(n,i)}$, its oversampling region is defined as the re-

gion enlarging K^i by some coarse spatial layers and some temporal layers. For example, letting $t_n^- = t_{\max\{n-M,0\}}$, the oversampling region of $K^{(n,i)}$ with N_s spatial and M temporal oversampling layers is defined as $(t_n^-, t_{n+1}] \times K^i_{N_s}$. Similarly, we denote $K_+^{(n,i)}$ a generic oversampling region related to the coarse space-time element $K^{(n,i)}$.

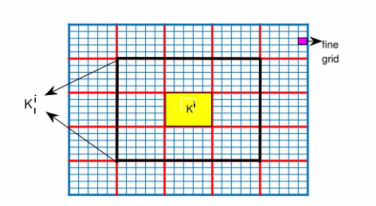

Figure 5.1: Illustration of oversampling space domain K_1^i.

5.1.2 Functional spaces and bilinear forms

In this subsection, we introduce some functional spaces and bilinear forms used throughout the chapter. For each $\omega \subset \Omega$ and $0 \le t_n < t_m \le T$, we define the following functional spaces.

$$
\begin{aligned}
&V(t_n, t_m; \omega) := L^2(t_n, t_m; H^1(\omega)) \cap H^1(t_n, t_m; L^2(\Omega)),\\
&W(t_n, t_m; \omega) := L^2(t_n, t_m; H^1(\omega)),\\
&V_0(t_n, t_m; \omega) := \{v \in V(t_n, t_m; \omega) |\ v(t_n, \cdot) = 0,\ v(t,x) = 0\ \forall (t,x) \in (t_n, t_m) \times \partial\omega\},\\
&V_d(t_n, t_m; \omega) := \{v \in L^2(t_n, t_m; H^1(\omega)) |\ v|_{(t_k, t_{k+1}) \times \omega} \in H^1(t_k, t_{k+1}; L^2(\Omega)),\ \forall k : n \le k \le m-1,\\
&\qquad\qquad \text{and } v(t,x) = 0\ \forall (t,x) \in (t_n, t_m) \times \partial\omega\}.
\end{aligned}
$$

To shorten notations, we use V, V_0, V_1 and W to denote $V(0,T;\Omega)$, $V_0(0,T;\Omega)$, $V_1(0,T;\Omega)$, respectively.

Next, we will introduce some auxiliary functions $\psi_j^{(n,i)}$'s corresponding to different

continua of the problem. Consider an oversampling region $K_+^{(n,i)}$ of the coarse space-time block $K^{(n,i)}$. For any coarse space-time block $K^{(n',i')} \subset K_+^{(n,i)}$, we denote $F^{(n',i')} = \{f_k^{(n',i')} | f_k^{(n',i')} = D_m \cap K^{(n',i')} \neq \varnothing\}$ as a set containing discrete channels inside $K^{(n',i')}$. Set $L^{(n',i')} = |F^{(n',i')}|$. The functions $\psi_j^{(n,i)}$ for $j = 0, 1, \cdots, L^{(n,i)}$ are defined as follows:

$$\int_{K^{(n',i')}} \tilde{\kappa}\psi_0^{(n,i)} = \delta_{n,n'}\delta_{i,i'} \text{ and } \int_{f_k^{(n',i')}} \tilde{\kappa}\psi_\ell^{(n,i)} = \delta_{n,n'}\delta_{i,i'}\delta_{\ell,k}.$$

We notice that $\psi_j^{(n,i)}$ is supported in $K^{(n,i)}$. Let $V_{aux}^{(n,i)} = \text{span}_j\{\psi_j^{(n,i)}\}$ be the local auxiliary multiscale space corresponding to the coarse space-time block $K^{(n,i)}$. For any $\omega \in \Omega$ and $0 \leq t_n < t_m \leq T$, we denote $V_{aux}(t_n, t_m; \omega) = \bigcup\{V_{aux}^{(k,i)} : K^i \subset \omega \text{ and } (t_k, t_{k+1}) \subset (t_n, t_m)\}$. For simplicity, we shall use V_{aux} to denote $V_{aux}(0, T; \Omega)$. We denote $N := \dim(V_{aux})$.

We now define $s(\cdot,\cdot)$ as a weighted L^2 inner production with weighting function $\tilde{\kappa} := \sum_j \kappa|\nabla\chi_j|^2$, that is

$$s(u, v) = \int_0^T \int_\Omega \tilde{\kappa}uv.$$

Here, χ_j's are the standard multiscale basis functions defined coarse elementwise.

Next, for each $\omega \subset \Omega$ and $0 \leq t_n \leq t_m \leq T$, we will define several bilinear operators $a(t_n, t_m; \omega; \cdot, \cdot) : V(t_n, t_m; \omega) \times W(t_n, t_m; \omega) \to \mathbb{R}$, $b(t_n, t_m; \omega; \cdot, \cdot) : V(t_n, t_m; \omega) \times W(t_n, t_m; \omega) \to \mathbb{R}$ and $e(t_n, t_m; \omega; \cdot, \cdot) : V(t_n, t_m; \omega) \times V(t_n, t_m; \omega) \to \mathbb{R}$ such that

$$a(t_n, t_m; \omega; v, w) = \int_{t_n}^{t_m} \int_\omega \kappa\nabla v \cdot \nabla w,$$

$$e(t_n, t_m; \omega; v, w) = \sum_{i=n}^{m-1} \int_{t_i}^{t_{i+1}} \int_\omega \tilde{\kappa}^{-1}\partial_t v \partial_t w,$$

$$b(t_n, t_m; \omega; v, w) = \int_{t_n}^{t_m} \int_\omega \partial_t v w,$$

and

$$s(t_n, t_m; \omega; u, v) = \int_{t_n}^{t_m} \int_{\omega} \tilde{\kappa} u v.$$

Then we will define $c(t_n, t_m; \omega; \cdot, \cdot)$ and $d(t_n, t_m; \omega; \cdot, \cdot)$ as

$$c(t_n, t_m; \omega; v, w) = b(t_n, t_m; \omega; v, w) + a(t_n, t_m; \omega; v, w),$$

and

$$d(t_n, t_m; \omega; v, w) = c(t_n, t_m; \omega; v, w) + e(t_n, t_m; \omega; v, w).$$

Furthermore, we can define several norms related to the above bilinear operators. For any $\omega \subset \Omega$ and $0 \leq t_n < t_m \leq T$, we define:

$$\|v\|^2_{L^2(\kappa,\omega)} = \int_{\omega} \kappa v^2,$$

$$\|v\|^2_{s(t_n,t_m;\omega)} = \int_{t_n}^{t_m} \int_{\omega} \tilde{\kappa} v^2,$$

$$\|v\|^2_{V(t_n,t_m;\omega)} = \int_{t_n}^{t_m} \|\nabla v\|^2_{L^2(\kappa,\omega)} + \int_{t_n}^{t_m} \|\tilde{\kappa}^{-\frac{1}{2}} \partial_t v\|^2_{L^2(\omega)}$$

and

$$\|v\|^2_{W(t_n,t_m;\omega)} = \int_{t_n}^{t_m} \int_{\omega} \kappa |\nabla v|^2.$$

We note that $\|v\|^2_{V(t_n,t_m;\omega)} = \|v\|^2_{W(t_n,t_m;\omega)} + \int_{t_n}^{t_m} \|\tilde{\kappa}^{-\frac{1}{2}} \partial_t v\|^2_{L^2(\omega)}$. To simplify the notations, we denote $\|v\|_{L^2(\omega)} = \|v\|_{L^2(0,T;\omega)}$, $\|v\|_{s(\omega)} = \|v\|_{s(0,T;\omega)}$, $\|v\|_{W(\omega)} = \|v\|_{W(0,T;\omega)}$, $\|v\|_{V(\omega)} = \|v\|_{V(0,T;\omega)}$, $\|v\|_{L^2(t_n,t_m)} = \|v\|_{L^2(t_n,t_m;\Omega)}$, $\|v\|_{s(t_n,t_m)} = \|v\|_{s(t_n,t_m;\Omega)}$, $\|v\|_{W(t_n,t_m)} = \|v\|_{W(t_n,t_m;\Omega)}$, $\|v\|_{V(t_n,t_m)} = \|v\|_{V(t_n,t_m;\Omega)}$, $\|v\|_{L^2} = \|v\|_{L^2(0,T;\Omega)}$, $\|v\|_{s} = \|v\|_{s(0,T;\Omega)}$, $\|v\|_{W} = \|v\|_{W(0,T;\Omega)}$ and $\|v\|_{V} = \|v\|_{V(0,T;\Omega)}$.

5.2 Space-time NLMC

In this section, we present the space-time NLMC upscaling method. First, we construct global and the localized space-time downscale operators, which can be used to define global space-time multiscale basis functions and local space-time multiscale basis functions. Then, we present the formulation of the coarse-grid solution.

5.2.1 Global multiscale space

We present the construction of the global downscale operator and the corresponding global numerical solution. We define the global downscale operator $F : \mathbb{R}^N \to V_0 \times V_{aux}$ by $U \mapsto (F_1(U), F_2(U))$ and

$$\begin{aligned} d(0,T;\Omega;F_1(U),w) - s(F_2(U),w) &= 0, && \forall w \in V_d, \\ s(F_1(U),\psi_j^{(n,i)}) &= U_j^{(n,i)}, && \forall \psi_j^{(n,i)} \in V_{aux}. \end{aligned}$$

We remark here that the global downscale operator also defines the global basis functions. Next, we can define the global coarse grid problem as: finding $U \in \mathbb{R}^N$ such that

$$s(F_2(U),\psi_j^{(n,i)}) = \int_0^T \int_\Omega f\psi_j^{(n,i)}, \ \forall \psi_j^{(n,i)} \in V_{aux} \tag{5.3}$$

and the global numerical solution u_{glo} is defined by

$$u_{glo} := F_1(U). \tag{5.4}$$

5.2.2 Localization of global multiscale basis functions

In this subsection, we will introduce the localized downscale operator $F_{ms} = (F_{ms,1}, F_{ms,2})$ and the localized coarse grid problem. For each space-time coarse block $K^{(n,i)}$ and

its oversampled region $K_+^{(n,i)} = [t_n^-, t_{n+1}] \times K_+^i$, we define a local downscale operator $F_{loc}^{(n,i)} : \mathbb{R}^N \to V_0(t_n^-, t_{n+1}; K_+^i) \times V_{aux}(t_n^-, t_{n+1}; K_+^i)$ by $U \mapsto (F_{loc,1}(U), F_{loc,2}(U))$ and

$$\begin{aligned} d(t_n^-, t_{n+1}; K_+^i; F_{loc(U),1}^{(n,i)}, w) - s(F_{loc,2}^{(n,i)}(U), w) = 0, \qquad & \forall w \in V_d(t_n^-, t_{n+1}; K_+^i), \\ s(F_{loc,1}^{(n,i)}(U), \psi_j^{(m,l)}) = U_j^{(m,l)}, \quad & \forall \psi_j^{(m,l)} \in V_{aux}(t_n^-, t_{n+1}, K_+^i). \end{aligned}$$

Then the localized downscale operator is defined by $F_{ms}(U) = \sum_{n,i} \chi^{(n,i)} F_{loc}^{(n,i)}(U)$ where $\chi^{(n,i)}$ is a partition of unity such that $\sum_{n,i} \chi^{(n,i)} \equiv 1$.

The downscale operator also defines multiscale basis functions with support being $K_+^{(n,i)}$. The coarse grid problem is then defined as: finding $U \in \mathbb{R}^N$ such that

$$s(F_{ms,2}(U), \psi_j^{(n,i)}) = \int_0^T \int_\Omega f \psi_j^{(n,i)}, \quad \forall \psi_j^{(n,i)} \in V_{aux}$$

and the localized numerical solution u_{ms} is defined by

$$u_{ms} := F_{ms,1}(U). \tag{5.5}$$

5.3 Convergence analysis

In this section, we will present a convergence analysis of the proposed method. We first prove in Theorem 5.3.3 that the global numerical solution is a good approximation of the solution. Then we prove that the global downscale operators have a decay property with respect to the temporal oversampling layers and the local downscale operators have a decay property with respect to the spatial oversampling layers in Lemma 5.3.7 and Lemma 5.3.8, respectively. In this chapter, we write $a \lesssim b$ if there exists a generic constant $C > 0$ such that $a \leq Cb$.

We first define a projection operator $\pi : L^2(0,T;L^2(\Omega)) \to V_{aux}$ such that

$$s(\pi(v),w) = s(v,w)\ \forall w \in V_{aux}.$$

Remark: It is easy to prove that there exists a constant C_0 such that for all $w \in V$

$$\frac{\|w-\pi(w)\|_s}{\|w\|_V} \leq C_0\Big(1+(\frac{\Delta t}{H^2})^{\frac{1}{2}}\Big). \tag{5.6}$$

We present the following result of the projection operator π.

Lemma 5.3.1. *Let $K \in \mathcal{T}_H$ be any coarse spatial element and $[t_{n-1},t_n] \subset [0,T]$. Then there exists a constant $C_s > 0$ such that for all $v_{aux} \in V_{aux}(t_{n-1},t_n;K)$, there exist $w \in V_0(t_{n-1},t_n;K)$ satisfying*

$$\pi(w) = v_{aux},\ \|w\|_V \leq C_s\|v_{aux}\|_s.$$

Lemma 5.3.1 can be proved using a similar technique in Lemma 3.2 [64]. For brevity of this article, we omit the proof.

Next, we establish the following estimates for later use in the analysis.

Lemma 5.3.2. *For any $\omega \subset \Omega$ and $0 \leq t_n < t_m \leq T$, the following inequalities hold for any $u \in V_0(t_n,t_m;\omega)$,*

$$\|u\|^2_{W(t_n,t_m;\omega)} \leq c(t_n,t_m;\omega;u,u), \tag{5.7}$$

$$\|u\|^2_{V(t_n,t_m;\omega)} \leq d(t_n,t_m;\omega;u,u). \tag{5.8}$$

Proof. It follows from the definitions of bilinear operators and $u(t_n) = 0$ that we have

$$\begin{aligned} c(t_n, t_m; \omega; u, u) &= \int_{t_n}^{t_m} \int_{\omega} \kappa \nabla u \cdot \nabla u + \int_{t_n}^{t_m} \int_{\omega} (\partial_t u) u \\ &= \int_{t_n}^{t_m} \int_{\omega} \kappa \nabla u \cdot \nabla u + \frac{1}{2} \|u(t_m)\|^2_{L^2(\omega)} \\ &\geq \|u\|^2_{W(t_n, t_m; \omega)} \end{aligned}$$

and

$$\begin{aligned} e(t_n, t_m; \omega; u, u) &= \sum_{i=n}^{m-1} \int_{t_i}^{t_{i+1}} \int_{\omega} \tilde{\kappa}^{-1} \partial_t u \partial_t u \\ &= \|\tilde{\kappa}^{-\frac{1}{2}} \partial_t u\|^2_{L^2(t_n, t_m; \omega)}. \end{aligned}$$

Therefore, we have

$$\|u\|^2_{V(t_n, t_m; \omega)} \leq c(t_n, t_m; \omega; u, u) + e(t_n, t_m; \omega; u, u) = d(t_n, t_m; \omega; u, u),$$

which proves the second inequality. This completes the proof. □

To prove the convergence result of the proposed method, we first show the convergence result of using the global multiscale basis functions.

Theorem 5.3.3. *Let u be the exact solution of* (5.1) *and u_{glo} be the solution of* (5.4)*. We have*

$$\|u - u_{glo}\|_V \leq C_0 (1 + (\frac{\Delta t}{H^2})^{\frac{1}{2}}) \|\tilde{\kappa}^{-\frac{1}{2}} f\|_{L^2} + \|\tilde{\kappa}^{-\frac{1}{2}} \partial_t u\|_{L^2}.$$

Moreover, if the multiscale partition of unity χ_i is replaced by the bilinear partition of

unity in the definition of $\tilde{\kappa}$, we have

$$\|u - u_{glo}\|_V \lesssim C_0(H + (\Delta t)^{\frac{1}{2}})\|\kappa^{-\frac{1}{2}} f\|_{L^2} + H\|\kappa^{-\frac{1}{2}}\partial_t u\|_{L^2}.$$

Proof. Set $\tilde{f} := \frac{f}{\tilde{\kappa}}$. It follows directly from (5.3) that we have

$$s(F_2(U), v) = s(\tilde{f}, v), \ \forall v \in V_{aux}.$$

Therefore, we have

$$F_2(U) = \pi(\tilde{f}).$$

We have

$$d(0, T; \Omega; u_{glo}, v) = s(F_2(U), v) \ \forall v \in W$$

and

$$c(0, T; \Omega; u, v) = s(\tilde{f}, v) \ \forall v \in W.$$

Then for any $v \in W$, the following equalities hold:

$$\begin{aligned} d(0, T; \Omega; u - u_{glo}, v) &= s(\tilde{f} - \pi(\tilde{f}), v) + e(0, T; \Omega; u, v) \\ &= s(\tilde{f} - \pi(\tilde{f}), v - \pi(v)) + e(0, T; \Omega; u, v) \\ &= (f, v - \pi(v)) + e(0, T; \Omega; u, v). \end{aligned}$$

Choosing $v = u - u_{glo}$ and utilizing Lemma 5.3.2 and (5.6), we have

$$\begin{aligned}\|u - u_{glo}\|_V &\leq \|\tilde{\kappa}^{-\frac{1}{2}} f\|_{L^2} \frac{\|(u - u_{glo}) - \pi(u - u_{glo})\|_s}{\|u - u_{glo}\|_V} + \|\tilde{\kappa}^{-\frac{1}{2}} \partial_t u\|_{L^2} \\ &\leq C_0(1 + (\frac{\Delta t}{H^2})^{\frac{1}{2}})\|\tilde{\kappa}^{-\frac{1}{2}} f\|_{L^2} + \|\tilde{\kappa}^{-\frac{1}{2}} \partial_t u\|_{L^2}.\end{aligned}$$

The second part of the theorem follows from the definition of $\tilde{\kappa}$ and $|\nabla \chi_i| = O(H^{-1})$. □

Theorem 5.3.3 justifies the use of global downscale operators. Moreover, it also implies that the coarse time step size should be at most $O(H^2)$ to ensure a good accuracy. To prove our main theorem, we need two important lemmas. We first show in Lemma 5.3.7 that the global downscale operators have a decay property with respect to the temporal oversampling layers. Then we prove in Lemma 5.3.8 that the local downscale operators have a decay property with respect to the spatial oversampling layers. Our main theorem shall be presented in Theorem 5.3.9. We first prove the following lemma, which will be frequently used in proofs.

Lemma 5.3.4. *For any $K \in \mathcal{T}_H$ and $\Delta t \leq t_n \leq T$, if $v_1 \in V(t_{n-1}, t_n; K)$ and $v_2 \in V_{aux}(t_{n-1}, t_n; K)$ satisfy*

$$d(t_{n-1}, t_n; K; v_1, w) = s(v_2, w) \ \forall w \in V_d(t_{n-1}, t_n; K),$$

then we have

$$\|v_2\|_{s(t_{n-1},t_n;K)} \leq \sigma \|v_1\|_{V(t_{n-1},t_n;K)},$$

where $\sigma = \Big(C_s\Big(C_0\Big(1 + (\frac{\Delta t}{H^2})^{\frac{1}{2}}\Big) + 1\Big) + 1\Big)$.

Proof. It follows from Lemma 5.3.1 that there exists $w \in V_0(t_{n-1}, t_n; K)$ such that

$$\pi(w) = v_2, \ \|w\|_V \leq C_s \|v_2\|_s.$$

Therefore, we have

$$\begin{aligned}\|v_2\|_s^2 &= s(v_2, w)\\ &= d(t_{n-1}, t_n; K; v_1, w)\\ &\leq \int_{t_{n-1}}^{t_n}\int_K \partial_t v_1 w + \int_{t_{n-1}}^{t_n}\int_K \kappa\nabla v_1\cdot\nabla w + \int_{t_{n-1}}^{t_n}\int_K \tilde{\kappa}^{-1}\partial_t v_1\partial_t w\\ &\leq \|v_1\|_{V(t_{n-1},t_n;K)}\Big(\|w\|_{s(t_{n-1},t_n;K)} + \|w\|_{V(t_{n-1},t_n;K)}\Big).\end{aligned} \tag{5.9}$$

Notice that

$$\begin{aligned}\|w\|_{s(t_{n-1},t_n;K)} &= \|w-\pi(w)\|_{s(t_{n-1},t_n;K)} + \|v_2\|_{s(t_{n-1},t_n;K)}\\ &\leq C_0\Big(1+(\frac{\Delta t}{H^2})^{\frac{1}{2}}\Big)\|w\|_{V(t_{n-1},t_n;K)} + \|v_2\|_{s(t_{n-1},t_n;K)}.\end{aligned}$$

Then the following inequalities hold true:

$$\begin{aligned}\Big(\|w\|_{s(t_{n-1},t_n;K)} + \|w\|_{V(t_{n-1},t_n;K)}\Big) &\leq \Big(C_0\Big(1+(\frac{\Delta t}{H^2})^{\frac{1}{2}}\Big)+1\Big)\|w\|_{V(t_{n-1},t_n;K)} + \|v_2\|_{s(t_{n-1},t_n;K)}\Big)\\ &\leq \Big(C_s\Big(C_0\Big(1+(\frac{\Delta t}{H^2})^{\frac{1}{2}}\Big)+1\Big)+1\Big)\|v_2\|_{s(t_{n-1},t_n;K)}\end{aligned} \tag{5.10}$$

A combination of (5.9) and (5.10) completes the proof. □

Before deriving the error between the global and localized downscale operators, we introduce some notions to be used in the analysis. We first define two cut-off functions: cut-off function in temporal variable $\chi_{k,m}(t)$ and cut-off function in spatial variable $\chi^s_{k,m}(x)$.

Definition 5.3.5. *For two non-negative integers k, m with $0 \leq k < m$,*

- *the cut-off function in time* $\chi_{k,m}(t)$ *is defined as*

$$\chi_{k,m}(t) := \begin{cases} 1, & if\ t > t_m, \\ \dfrac{t - t_k}{t_m - t_k}, & if\ t_k \le t \le t_m, \\ 0, & if\ t \le t_k; \end{cases}$$

- *the cut-off function in space* $\chi^s_{k,m}(x)$ *is defined as a smooth function such that*

  (a) $\chi^s_{k,m}(x) \in [0,1]$,

  (b)

$$\chi^s_{k,m}(x) = \begin{cases} 1, & on\ K_k, \\ 0, & on\ K_m, \end{cases}$$

  (c) $|\nabla \chi^s_{k,m}|^2 \le C_\chi \sum_i |\nabla \chi_i|^2$ *for some constant* C_χ.

Note that $\chi_{k,m}(t) \in [0,1]$. To simply the notations, for $0 \le k \le n$, we denote $\chi_k(t) = \chi_{n-k,n-k+1}(t)$.

Next, we shall define a temporal localized downscale operator $\tilde{F}_{loc} = (\tilde{F}_{loc,1}, \tilde{F}_{loc,2})$.

Definition 5.3.6. *The temporal localized downscale operator* $\tilde{F}^{(n)}_{loc} : \mathbb{R}^N \to V_0(t_n^-, t_{n+1}; \Omega) \times V_{aux}(t_n^-, t_{n+1}; \Omega)$ *are defined by* $U \mapsto (\tilde{F}^{(n)}_{loc,1}, \tilde{F}^{(n)}_{loc,2})$ *and*

$$d(t_n^-, t_{n+1}; \Omega; \tilde{F}^{(n)}_{loc,1}(U), w) - s(\tilde{F}^{(n)}_{loc,2}(U), w) = 0, \qquad \forall w \in V_d(t_n^-, t_{n+1}; \Omega),$$
$$s(\tilde{F}^{(n)}_{loc,1}(U), \psi^{(m,l)}_j) = U^{(m,l)}_j, \quad \forall \psi^{(m,l)}_j \in V_{aux}(t_n^-, t_{n+1}; \Omega).$$

We prove in the following lemma that the global downscale operator has a decay property with respect to the temporal oversampling layers. This also implies that the global

multiscale basis functions has a decay property with respect to the temporal oversampling layers.

Lemma 5.3.7. *Let M be the number of temporal oversampling layers. For any space-time element $K = K^{(n,i)}$, $t_n^- = t_{n-M}$ and $U \in \mathcal{R}^N$, we have*

$$\|F_1(U) - \tilde{F}^{(n)}_{loc,1}(U)\|_{V(t_n,t_{n+1})} \lesssim (1+\tilde{E}^{-1})^{1-M}\Big(\|F_1(U)\|^2_{V(t_n^-,t_{n-M+1})} + \|F_1(U)\|^2_{s(t_n^-,t_{n-M+1})}\Big),$$

where $\tilde{E} = \dfrac{C_0}{2}\Big(\Big(1+\frac{1}{\Delta t \min\{\bar{\kappa}\}}\Big)+\sigma^2\Big)\Big(1+(\frac{\Delta t}{H^2})^{\frac{1}{2}}\Big)$.

Proof. First, since

$$\begin{aligned} d(0,T;\Omega;F_1(U),w) - s(F_2(U),w) &= 0, && \forall w \in V_d,\\ s(F_1(U),\psi_j^{(n,i)}) &= U_j^{(n,i)}, && \forall \psi_j^{(n,i)} \in V_{aux}, \end{aligned}$$

and $V_d(t_{n-M},t_{n+1}) \subset V_d$, we have

$$\begin{aligned} d(t_{n-M},t_{n+1};\Omega;F_1(U),w) - s(F_2(U),w) &= 0, && \forall w \in V_d(t_{n-M},t_{n+1}),\\ s(F_1(U),\psi_j^{(m,l)}) &= U_j^{(m,l)}, && \forall \psi_j^{(m,l)} \in V_{aux}(t_{n-M},t_{n+1}). \end{aligned}$$

We define $\tilde{\eta} := F(U) - \tilde{F}^{(n)}_{loc}(U)$, $\tilde{\eta_1} := F_1(U) - \tilde{F}^{(n)}_{loc,1}(U)$ and $\tilde{\eta_2} := F_2(U) - \tilde{F}^{(n)}_{loc,2}(U)$. Then the following equalities hold:

$$\begin{aligned} d(t_{n-M},t_{n+1};\Omega;\tilde{\eta}_1,w) - s(\tilde{\eta}_2,w) &= 0, && \forall w \in V_d(t_{n-M},t_{n+1}),\\ s(\tilde{\eta}_1,\psi_j^{(m,l)}) &= 0, && \forall \psi_j^{(m,l)} \in V_{aux}(t_{n-M},t_{n+1}). \end{aligned}$$

We will estimate $\|\tilde{\eta}\|^2_{V(t_n,t_{n+1};\Omega)}$ in three steps.

Step 1: We will prove

$$\|\tilde{\eta}_1\|^2_{V(t_{n-k+1},t_{n+1})} \le \tilde{E}\|\tilde{\eta}_1\|^2_{V(t_{n-k},t_{n-k+1})} \text{ for } 1 \le k \le M-1. \tag{5.11}$$

Let $w = \chi_k\tilde{\eta}_1$, for $k \le M-1$. Since $\tilde{\eta}_1 \in V_0$ and $1-\chi_k(t) = 0$ if $t \ge t_{n-k+1}$, then $w \in V_d(t_{n-M}, t_{n+1})$. Then we have

$$\begin{aligned} d(t_{n-M}, t_{n+1}; \Omega; \tilde{\eta}_1, \chi_k\tilde{\eta}_1) - s(\tilde{\eta}_2, \chi_k\tilde{\eta}_1) = 0, \\ s(\tilde{\eta}_1, \psi_j^{(m,l)}) = 0 \ \forall \psi_j^{(m,l)} \in V_{aux}(t_{n-k}, t_{n+1}). \end{aligned} \tag{5.12}$$

Notice that

$$\begin{aligned} \int_{t_{n-k}}^{t_{n+1}} \int_\Omega (\tilde{\eta}_1)_t \chi_k \tilde{\eta}_1 &= -\int_{t_{n-k}}^{t_{n+1}} \int_\Omega (\chi_k\tilde{\eta}_1)_t \tilde{\eta}_1 + \int_\Omega \tilde{\eta}_1^2(t_{n+1}, \cdot) \\ &= -\int_{t_{n-k}}^{t_{n+1}} \int_\Omega (\tilde{\eta}_1)_t \chi_k \tilde{\eta}_1 - \frac{1}{\Delta t}\int_{t_{n-k}}^{t_{n-k+1}} \int_\Omega \tilde{\eta}_1^2 + \int_\Omega \tilde{\eta}_1^2(t_{n+1}, \cdot). \end{aligned}$$

This gives

$$\int_{t_{n-k}}^{t_{n+1}} \int_\Omega (\tilde{\eta}_1)_t \chi_k \tilde{\eta}_1 = -\frac{1}{2\Delta t}\int_{t_{n-k}}^{t_{n-k+1}} \int_\Omega \tilde{\eta}_1^2 + \frac{1}{2}\int_\Omega \tilde{\eta}_1^2(t_{n+1}, \cdot). \tag{5.13}$$

Combining (5.12) and (5.13), we arrive at the following estimate:

$$\begin{aligned} \|\eta_1\|^2_{V(t_{n-k+1},t_{n+1};\Omega)} + \frac{1}{2}\int_\Omega \tilde{\eta}_1^2(t_{n+1}, \cdot) &\le \frac{1}{2\Delta t}\int_{t_{n-k}}^{t_{n-k+1}} \int_\Omega \tilde{\eta}_1^2 + s(\tilde{\eta}_2, \chi_k\tilde{\eta}_1) \\ &= \frac{1}{2\Delta t}\int_{t_{n-k}}^{t_{n-k+1}} \int_\Omega \tilde{\eta}_1^2 + \int_{t_{n-k}}^{t_{n-k+1}} \int_\Omega \tilde{\kappa}\chi_k\tilde{\eta}_2\tilde{\eta}_1. \end{aligned} \tag{5.14}$$

Utilizing (5.14) and Cauchy-Schwarz inequality, one can show that

$$\|\tilde{\eta}_1\|^2_{V((t_{n-k+1},t_{n+1})} \leq \frac{1}{2}\Big((1+\frac{1}{\Delta t \min\{\tilde{\kappa}\}})\|\tilde{\eta}_1\|^2_{s((t_{n-k},t_{n-k+1})} + \|\tilde{\eta}_2^2\|_{s((t_{n-k},t_{n-k+1})}\Big).$$

Since $d(t_{n-k}, t_{n-k+1}; \Omega; \tilde{\eta}_1, w) = s(\tilde{\eta}_2, w)$ for any $w \in V_d(t_{n-k}, t_{n-k+1})$, it follows from Lemma 5.3.4 that

$$\|\tilde{\eta}_2^2\|_{s(t_{n-k},t_{n-k+1})} \leq \sigma^2 \|\tilde{\eta}_1\|^2_{s(t_{n-k},t_{n-k+1})}.$$

Therefore, we have

$$\|\tilde{\eta}_1\|^2_{V(t_{n-k+1},t_{n+1})} \leq \tilde{E}\|\tilde{\eta}_1\|^2_{V(t_{n-k},t_{n-k+1})},$$

where $\tilde{E} = \frac{C_0}{2}\Big(\Big(1+\frac{1}{\Delta t \min\{\tilde{\kappa}\}}\Big)+\sigma^2\Big)\Big(1+(\frac{\Delta t}{H^2})^{\frac{1}{2}}\Big)$.

Step 2: We will prove

$$\|\tilde{\eta}_1\|^2_{V(t_n,t_{n+1})} \leq (1+\tilde{E}^{-1})^{1-M}\|\tilde{\eta}_1\|^2_{(t_{n-M+1},t_{n+1})}. \tag{5.15}$$

Using Inequality (5.11), we have the following estimate: for $1 \leq k \leq M-1$,

$$\begin{aligned}\|\tilde{\eta}_1\|^2_{V(t_{n-k},t_{n+1})} &= \|\tilde{\eta}_1\|^2_{V(t_{n-k+1},t_{n+1})} + \|\tilde{\eta}_1\|^2_{V(t_{n-k},t_{n-k+1})} \\ &\geq (1+\tilde{E}^{-1})\|\tilde{\eta}_1\|^2_{V(t_{n-k+1},t_{n+1})},\end{aligned}$$

Using the above inequality recursively, we obtain (5.15).

Step 3: We will prove

$$\|\tilde{\eta}_1\|^2_{V(t_{n-M+1},t_{n+1})} \leq \Big(1+\frac{1}{\Delta t \min\{\tilde{\kappa}\}}\Big)\|F_1(U)\|^2_{s(t_{n-M},t_{n-M+1})} + \|F_1(U)\|^2_{V(t_{n-M},t_{n-M+1})}.$$

Since $\tilde{\eta}_1 \in V_d(t_{n-M}, t_{n+1})$, the following equalities hold true:

$$\begin{aligned} d(t_{n-M}, t_{n+1}; \Omega; \tilde{\eta}_1, \tilde{\eta}_1) &= d(t_{n-M}, t_{n+1}; \Omega; \tilde{\eta}_1, \tilde{\eta}_1) - s(\tilde{\eta}_2, \tilde{\eta}_1) \\ &= 0. \end{aligned}$$

Using a similar derivation to obtain (5.13), one can also show that

$$\begin{aligned} \frac{1}{2}\Big(\int_\Omega \tilde{\eta}_1^2(t_{n+1}, \cdot) - \int_\Omega \tilde{\eta}_1^2(t_{n-M}, \cdot)\Big) + \|\tilde{\eta}_1\|^2_{V(t_{n-M}, t_{n+1})} &= d(t_{n-M}, t_{n+1}; \Omega; \tilde{\eta}_1, \tilde{\eta}_1) \\ &= 0. \end{aligned}$$

Notice that

$$\int_\Omega \tilde{\eta}_1^2(t_{n-M}, \cdot) = \int_\Omega F_1^2(t_{n-M}, \cdot) = -2 \int_{t_{n-M}}^{t_{n-M+1}} \int_\Omega \partial_t((1-\chi_M)F_1(U))(1-\chi_M)F_1(U). \tag{5.16}$$

Utilizing (5.16), we have

$$\begin{aligned} \|\tilde{\eta}_1\|^2_{V(t_{n-M}, t_{n+1})} &\le \frac{1}{2} \int_\Omega \tilde{\eta}_1^2(t_{n-M}, \cdot) \\ &= - \int_{t_{n-M}}^{t_{n-M+1}} \int_\Omega \partial_t((1-\chi_M)F_1(U))(1-\chi_M)F_1(U) \\ &\le \frac{1}{\Delta t \min\{\tilde{\kappa}\}} \|F_1(U)\|^2_{s(t_{n-M}, t_{n-M+1})} + \|F_1(U)\|_{V(t_{n-M}, t_{n-M+1})} \|F_1(U)\|_{s(t_{n-M}, t_{n-M+1})} \\ &\le \Big(1 + \frac{1}{\Delta t \min\{\tilde{\kappa}\}}\Big) \|F_1(U)\|^2_{s(t_{n-M}, t_{n-M+1})} + \|F_1(U)\|^2_{V(t_{n-M}, t_{n-M+1})}. \end{aligned}$$

The proof is completed using $\|\tilde{\eta}_1\|^2_{V(t_{n-M}, t_{n+1})} \ge \|\tilde{\eta}_1\|^2_{V(t_{n-M+1}, t_{n+1})}$ together with Step 2. □

Define the constant

$$C_\kappa = \sup_{v\in V} \frac{\|v\|_s}{\|v\|_V}.$$

We now prove in the following lemma that local downscale operators have a decay property with respect to the spatial oversampling layers.

Lemma 5.3.8. *Let N_s be the number of the oversampling layers in space. For any coarse space element $K = K^i \in \mathcal{T}_H$ and time element $[t_n, t_{n+1}] \in \mathcal{T}_{\Delta t}$, we have*

$$\|\tilde{F}^{(n)}_{loc,1}(U) - F^{(n,i)}_{loc,1}(U)\|^2_{V(t_n^-,t_{n+1};K)} \le C_\chi E^{1-N_s}\Big(\|\tilde{F}^{(n)}_{loc,1}(U)\|^2_{V(t_n^-,t_{n+1};K_N\setminus K_{N-1})} + \|\tilde{F}^{(n)}_{loc,1}(U)\|^2_{s(t_n^-,t_{n+1};K_N\setminus K_{N-1})}\Big),$$

where $E := 1 + \dfrac{1}{2C_\chi^2 + \sigma C_\kappa}$.

Proof. Notice that $V_d(t_n^-, t_{n+1}; K^i_+) \subset V_d(t_n^-, t_{n+1}; \Omega)$. It follows from the definitions of $\tilde{F}_{loc}$ and $F^{(n,i)}_{loc}$ that the following equalities hold true for any $v \in V_d(t_n^-, t_{n+1}; K^i_+)$ and $\psi^{(n,i)}_j \in V_{aux}(t_n^-, t_{n+1}; K^i_+)$:

$$\begin{aligned} d(t_n^-, t_{n+1}; K^i_+; F^{(n,i)}_{loc,1}(U) - \tilde{F}^{(n)}_{loc,1}(U), v) + s(F^{(n,i)}_{loc,2}(U) - \tilde{F}^{(n)}_{loc,2}(U), v) = 0,\\ s(F^{(n,i)}_{loc,1}(U) - \tilde{F}^{(n)}_{loc,1}(U), \psi^{(n,i)}_j) = 0. \end{aligned} \tag{5.17}$$

In this proof, we denote $\|\cdot\|_{V(\omega)}$, $\|\cdot\|_{W(\omega)}$ and $\|\cdot\|_{s(\omega)}$ by

$$\|v\|_{V(\omega)} := \|v\|_{V(t_n^-,t_{n+1};\omega)},\ \|v\|_{W(\omega)} := \|v\|_{W(t_n^-,t_{n+1};\omega)},\ \|v\|_{s(\omega)} := \|v\|_{s(t_n^-,t_{n+1};\omega)}.$$

We then define $\eta = F^{(n,i)}_{loc}(U) - \tilde{F}^{(n)}_{loc}(U)$ and $\eta_j = F^{(n,i)}_{loc,j}(U) - \tilde{F}^{(n)}_{loc,j}(U)$ for $j = 1, 2$. For

$k = 1, 2, \cdots, N_s$, we denote $\chi_{k,k-1} := 1 - \chi_{k-1,k}$. Then we have

$$\begin{aligned} &\|\eta_1(t_{n+1})\|^2_{L^2(K_{k-1})} + \|\eta_1\|^2_{V(K_{k-1})} \\ \leq & d(t_n^-, t_{n+1}; K_k; \eta_1, \chi_{k,k-1}\eta_1) - a(t_n^-, t_{n+1}; K_k \backslash K_{k-1}; \eta_1, \chi_{k,k-1}\eta_1). \end{aligned} \tag{5.18}$$

Notice that $\pi(\chi_{k,k-1}\eta_1)|_{K_{k-1}} = \pi(\eta_1)|_{K_{k-1}} = 0$. Choosing $v = \chi_{k,k-1}\eta_1$ in (5.17) and utilizing Cauchy-Schwartz Inequality, we have

$$\begin{aligned} d(t_n^-, t_{n+1}; K_k; \eta_1, \chi_{k,k-1}\eta_1) &= -s(\eta_2, \chi_{k,k-1}\eta_1) \\ &\leq \|\eta_2\|_{s(K_k \backslash K_{k-1})} \|\chi_{k,k-1}\eta_1\|_{s(K_k \backslash K_{k-1})} \\ &\leq \|\eta_2\|_{s(K_k \backslash K_{k-1})} \|\eta_1\|_{s(K_k \backslash K_{k-1})}. \end{aligned}$$

By Lemma 5.3.4, we have

$$\|\eta_2\|_{s(K_k \backslash K_{k-1})} \leq \sigma \|\eta_1\|_{V(K_k \backslash K_{k-1})}.$$

Moreover, we have $\|\eta_1\|_{s(K_k \backslash K_{k-1})} \leq C_\kappa \|\eta_1\|_{V(K_k \backslash K_{k-1})}$. Therefore

$$d(t_n^-, t_{n+1}; K_k; \eta_1, \chi_{k,k-1}\eta_1) \leq \sigma C_\kappa \|\eta_1\|^2_{V(K_k \backslash K_{k-1})}. \tag{5.19}$$

Since $\nabla(\chi_{k,k-1}\eta_1) = \eta_1 \nabla(\chi_{k,k-1}) + \chi_{k,k-1}\nabla(\eta_1)$ and $|\nabla \chi_{k,k-1}|^2 \leq C_\chi \sum_i |\nabla \chi_i|^2$, we have

$$\begin{aligned} \int_{t_n^-}^{t_{n+1}} \int_{K_k \backslash K_{k-1}} \kappa \nabla \eta_1 \cdot \nabla\Big(\chi_{k,k-1}\eta_1\Big) &\leq \|\eta_1\|_{W(K_k \backslash K_{k-1})} \|\chi_{k,k-1}\eta_1\|_{W(K_k \backslash K_{k-1})} \\ &\leq C_\chi \|\eta_1\|_{W(K_k \backslash K_{k-1})} \Big(\|\eta_1\|_{W(K_k \backslash K_{k-1})} + \|\eta_1\|_{s(K_k \backslash K_{k-1})} \Big) \\ &\leq C_\chi \|\eta_1\|_{W(K_k \backslash K_{k-1})} \Big(\|\eta_1\|_{W(K_k \backslash K_{k-1})} + \|\eta_1\|_{V(K_k \backslash K_{k-1})} \Big). \end{aligned}$$

Using $\|\eta_1\|_{W(K_k\setminus K_{k-1})} \le \|\eta_1\|_{V(K_k\setminus K_{k-1})}$, we obtain

$$\int_{t_n^-}^{t_{n+1}} \int_{K_k\setminus K_{k-1}} \kappa\nabla\eta_1\cdot\nabla\Big(\chi_{k,k-1}\eta_1\Big) \le 2C_\chi\|\eta_1\|^2_{V(K_k\setminus K_{k-1})}. \tag{5.20}$$

By a combination of (5.18), (5.19) and (5.20), we arrive at

$$\begin{aligned}\|\eta_1(t_{n+1})\|^2_{L^2(K_{k-1})} + \|\eta_1\|^2_{V(K_{k-1})} &\le (2C_\chi^2+\sigma C_\kappa)\|\eta_1\|^2_{V(K_k\setminus K_{k-1})}\\ &= (2C_\chi^2+\sigma C_\kappa)\Big(\|\eta_1\|^2_{V(K_k)} - \|\eta_1\|^2_{V(K_{k-1})}\Big),\end{aligned}$$

which gives

$$\|\eta_1\|^2_{V(K_{k-1})} \le \Big(1+\frac{1}{2C_\chi^2+\sigma C_\kappa}\Big)^{-1}\|\eta_1\|^2_{V(K_k)},\ \text{for } 1\le k\le N_s-1.$$

Denote $E := 1+\dfrac{1}{2C_\chi^2+\sigma C_\kappa}$. Using above inequality recursively, we obtain

$$\|\eta_1\|^2_{V(K)} \le E^{1-N_s}\|\eta_1\|^2_{V(K_{N_s-1})}.$$

It remains to estimate $\|\eta_1\|^2_{V(K_{N_s-1})}$. We shall prove:

$$\|\eta_1\|^2_{V(K_{N_s-1})} \le C_\chi\Big(\|\tilde{F}^{(n)}_{loc,1}(U)\|^2_{V(t_n^-,t_{n+1};K_{N_s}\setminus K_{N_s-1})} + \|\tilde{F}^{(n)}_{loc,1}(U)\|^2_{s(t_n^-,t_{n+1};K_{N_s}\setminus K_{N_s-1})}\Big).$$

Notice that

$$\begin{aligned}
&\frac{1}{2}\|\eta_1(t_{n+1})\|^2 + \|\eta_1\|^2_{V(K_{N_s})}\\
=&d(t_n^-, t_{n+1}; K_{N_s}; \eta_1, \eta_1)\\
=&d(t_n^-, t_{n+1}; K_{N_s}; \eta_1, F^{(n,i)}_{loc,1}(U) - \chi_{N_s,N_s-1}\tilde{F}^{(n)}_{loc,1}(U))\\
&+ d(t_n^-, t_{n+1}; K_{N_s}; \eta_1, (\chi_{N_s,N_s-1} - 1)\tilde{F}^{(n)}_{loc,1}(U)).
\end{aligned} \tag{5.21}$$

We next estimate each of the above two terms. Choosing $v = F^{(n,i)}_{loc,1}(U) - \chi_{N_s,N_s-1}\tilde{F}^{(n)}_{loc,1}(U)$ in (5.17) and using Cauchy-Schwartz Inequality, we have the following estimate:

$$\begin{aligned}
&d(t_n^-, t_{n+1}; K_{N_s}; \eta_1, F^{(n,i)}_{loc,1}(U) - \chi_{N_s,N_s-1}\tilde{F}^{(n)}_{loc,1}(U))\\
=&-s(\eta_2, F^{(n,i)}_{loc,1}(U) - \chi_{N_s,N_s-1}\tilde{F}^{(n)}_{loc,1}(U))\\
=&-s(\eta_2, (1 - \chi_{N_s,N_s-1})\tilde{F}^{(n)}_{loc,1}(U))\\
\leq&\|\eta_2\|_{s(K_{N_s}\setminus K_{N_s-1})}\|\tilde{F}^{(n)}_{loc,1}(U)\|_{s(K_{N_s}\setminus K_{N_s-1})}.
\end{aligned}$$

Furthermore, since $\|\eta_2\|_{s(K_{N_s}\setminus K_{N_s-1})} \leq \sigma\|\eta_1\|_{V(K_{N_s}\setminus K_{N_s-1})}$, we have the following estimate:

$$\begin{aligned}
&d(t_n^-, t_{n+1}; K_{N_s}; \eta_1, F^{(n,i)}_{loc,1}(U) - \chi_{N_s,N_s-1}\tilde{F}^{(n)}_{loc,1}(U))\\
\leq&\sigma\|\eta_1\|_{V(K_{N_s}\setminus K_{N_s-1})}\|\tilde{F}^{(n)}_{loc,1}(U)\|_{s(K_{N_s}\setminus K_{N_s-1})}.
\end{aligned} \tag{5.22}$$

We also have

$$\begin{aligned}
&d(t_n^-, t_{n+1}; K_{N_s}; \eta_1, (\chi_{N_s,N_s-1} - 1)\tilde{F}^{(n)}_{loc,1}(U))\\
\leq&\|\tilde{\kappa}^{-\frac{1}{2}}\partial_t\eta_1\|_{L^2(K_{N_s}\setminus K_{N_s-1})}(\|\tilde{F}^{(n)}_{loc,1}(U)\|_{s(K_{N_s}\setminus K_{N_s-1})} + \|\tilde{\kappa}^{-\frac{1}{2}}\partial_t\tilde{F}^{(n)}_{loc,1}(U)\|_{L^2(K_{N_s}\setminus K_{N_s-1})})\\
&+ \|\eta_1\|_{W(K_{N_s}\setminus K_{N_s-1})}\|(\chi_{N_s,N_s-1} - 1)\tilde{F}^{(n)}_{loc,1}(U)\|_{W(K_{N_s}\setminus K_{N_s-1})}.
\end{aligned}$$

Notice that

$$\|(\chi_{N_s,N_s-1}-1)\tilde{F}^{(n)}_{loc,1}(U)\|_{W(K_{N_s}\setminus K_{N_{s-1}})} \le C_\chi\Big(\|\tilde{F}^{(n)}_{loc,1}(U)\|_{W(K_{N_s}\setminus K_{N_{s-1}})} + \|\tilde{F}^{(n)}_{loc,1}(U)\|_{s(K_{N_s}\setminus K_{N_{s-1}})}\Big).$$

We obtain the following estimate:

$$\begin{aligned} &d(t_n^-, t_{n+1}; K_{N_s}; \eta_1, (\chi_{N_s,N_s-1}-1)\tilde{F}^{(n)}_{loc,1}(U)) \\ \le& C_\chi\|\eta_1\|_{V(K_{N_s}\setminus K_{N_{s-1}})}\Big(\|\tilde{F}^{(n)}_{loc,1}(U)\|_{V(K_{N_s}\setminus K_{N_{s-1}})} + \|\tilde{F}^{(n)}_{loc,1}(U)\|_{s(K_{N_s}\setminus K_{N_{s-1}})}\Big). \end{aligned} \tag{5.23}$$

Combing (5.22) and (5.23), we arrive at

$$\frac{1}{2}\|\eta_1(t_{n+1})\|^2 + \|\eta_1\|^2_{V(K_{N_s})} \le C_\chi\|\eta_1\|_{V(K_{N_s}\setminus K_{N_{s-1}})}\Big(\|\tilde{F}^{(n)}_{loc,1}(U)\|_{V(K_{N_s}\setminus K_{N_{s-1}})} + \|\tilde{F}^{(n)}_{loc,1}(U)\|_{s(K_{N_s}\setminus K_{N_{s-1}})}\Big).$$

Therefore,

$$\|\eta_1\|^2_{V(K_{N_s})} \le C_\chi\Big(\|\tilde{F}^{(n)}_{loc,1}(U)\|^2_{V(K_{N_s}\setminus K_{N_{s-1}})} + \|\tilde{F}^{(n)}_{loc,1}(U)\|^2_{s(K_{N_s}\setminus K_{N_{s-1}})}\Big).$$

This completes the proof. □

Finally, we state and prove the main result of this work. It reads as follows.

Theorem 5.3.9. *Let u be the solution of* (5.1) *and $F_{ms,1}(U_{ms})$ be the solution of* (5.5) *with the numbers of spatial and temporal oversampling layers being N_s and M, respectively.*

We have

$$\|u - F_{ms,1}(U_{ms})\|_V \leq C_0(1 + (\frac{\Delta t}{H^2})^{\frac{1}{2}})\|\tilde{\kappa}^{-\frac{1}{2}} f\|_{L^2} + \|\tilde{\kappa}^{-\frac{1}{2}}\partial_t u\|_{L^2}$$
$$+C\Big((1+\tilde{E}^{-1})^{1-M} + (1+E^{-1})^{1-N}\Big)^{\frac{1}{2}} H^{-\frac{d}{2}}\Big(\|F_1(U_{ms})\|_V + \|F_1(U_{ms})\|_s\Big).$$

Moreover, if $C_\kappa(1+C_\kappa^2)^{\frac{1}{2}} C^{\frac{1}{2}}\Big((1+\tilde{E}^{-1})^{1-M} + (1+E^{-1})^{1-N}\Big)^{\frac{1}{2}} H^{-\frac{d}{2}} \leq \frac{1}{2}$*, we have*

$$\begin{aligned}\|u - F_{ms,1}(U_{ms})\|_V \quad &\leq \quad C_0(1 + (\frac{\Delta t}{H^2})^{\frac{1}{2}})\|\tilde{\kappa}^{-\frac{1}{2}} f\|_{L^2} + \|\tilde{\kappa}^{-\frac{1}{2}}\partial_t u\|_{L^2} \\ &\quad +C\Big((1+\tilde{E}^{-1})^{1-M} + (1+E^{-1})^{1-N}\Big)^{\frac{1}{2}} H^{-\frac{d}{2}}\Big(\|F_1(U_{glo})\|_V\Big).\end{aligned}$$

Proof. Notice that $u - F_{ms,1}(U_{ms}) = u - F_1(U_{glo}) + F_1(U_{glo}) - F_1(U_{ms}) + F_1(U_{ms}) - F_{ms,1}(U_{ms})$, where $F_1(U_{glo})$ is the solution to (5.4). Using triangle inequality, we obtain

$$\|u - F_{ms,1}(U_{ms})\|_V \leq \|u - F_1(U_{glo})\|_V + \|F_1(U_{glo}) - F_1(U_{ms})\|_V + \|F_1(U_{ms}) - F_{ms,1}(U_{ms})\|_V.$$

We will estimate the above three terms separately. By Theorem 5.3.3, we obtain the estimate for the first term:

$$\|u - F_1(U_{glo})\|_V \leq C_0(1 + (\frac{\Delta t}{H^2})^{\frac{1}{2}})\|\tilde{\kappa}^{-\frac{1}{2}} f\|_{L^2} + \|\tilde{\kappa}^{-\frac{1}{2}}\partial_t u\|_{L^2}.$$

To estimate $\|F_1(U_{ms}) - F_{ms,1}(U_{ms})\|_V$, we utilize Lemma 5.3.7 and Lemma 5.3.8 to obtain

the following estimate.

$$
\begin{aligned}
&\|F_1(U_{ms}) - F_{ms,1}(U_{ms})\|_V^2 \\
=&\sum_{n,K^i} \|F_1(U_{ms}) - F_{ms,1}(U_{ms})\|^2_{V(t_n,t_{n+1};K^i)} \\
\leq& C \sum_{n,K^i} \Big((1+\tilde{E}^{-1})^{1-M} + (1+E^{-1})^{1-N_s}\Big) \|F_1(U_{ms})\|^2_{\tilde{V}(t_{n-M},t_{n+1};K^i_{N_s}\setminus K^i_{N_s-1})} \\
\leq& C\Big((1+\tilde{E}^{-1})^{1-M} + (1+E^{-1})^{1-N_s}\Big) H^{-d} \|F_1(U_{ms})\|^2_{\tilde{V}},
\end{aligned}
$$

where $\|.\|_{\tilde{V}}$ is defined as $\|v\|^2_{\tilde{V}} := \|v\|^2_V + \|v\|^2_s$. Finally, we only need to estimate $\|F_1(U_{glo}) - F_1(U_{ms})\|_V$. Using Cauchy-Schwartz Inequality and the definition of C_κ, we have

$$
\begin{aligned}
\|F_1(U_{glo}) - F_1(U_{ms})\|_V^2 &\leq s(F_2(U_{glo}) - F_2(U_{ms}), F_1(U_{glo}) - F_1(U_{ms})) \\
&\leq \|F_2(U_{glo}) - F_2(U_{ms})\|_s \|F_1(U_{glo}) - F_1(U_{ms})\|_s \\
&\leq C_\kappa \|F_2(U_{glo}) - F_2(U_{ms})\|_s \|F_1(U_{glo}) - F_1(U_{ms})\|_V.
\end{aligned}
$$

For any $K^i \in \mathcal{T}_H$ and $0 \leq t_n < t_{n+1} \leq T$, by Lemma 5.3.4 we have

$$
\begin{aligned}
&\|F_2(U_{glo}) - F_2(U_{ms})\|^2_{s(t_n,t_{n+1};K^i)} = \|F^{(n,i)}_{loc,2}(U_{ms}) - F_2(U_{ms})\|^2_{s(t_n,t_{n+1};K^i)} \\
\leq& C\Big((1+\tilde{E}^{-1})^{1-M} + (1+E^{-1})^{1-N_s}\Big)\Big(\|F_1(U_{ms})\|^2_{\tilde{V}(t_{n-M},t_{n+1};K^i_{N_s})} \\
&- \|F_1(U_{ms})\|^2_{\tilde{V}(t_{n-M+1},t_{n+1};K^i_{N_s-1})}\Big).
\end{aligned}
$$

Finally, we obtain

$$
\|F_1(U_{glo}) - F_1(U_{ms})\|_V^2 \leq C_\kappa^2 C\Big((1+\tilde{E}^{-1})^{1-M} + (1+E^{-1})^{1-N_s}\Big) H^{-d} \|F_1(U_{ms})\|^2_{\tilde{V}}.
$$

Since $\|v\|_s \leq C_\kappa \|v\|_V$ for any $v \in V$, we have

$$\|v\|_{\tilde{V}} \leq (1 + C_\kappa^2)^{\frac{1}{2}} \|v\|_V.$$

Therefore, we have

$$\begin{aligned}\|F_1(U_{ms}) - F_1(U_{glo})\|_V &\leq C_\kappa (1 + C_\kappa^2)^{\frac{1}{2}} C^{\frac{1}{2}} \Big((1 + \tilde{E}^{-1})^{1-M} + (1 + E^{-1})^{1-N_s} \Big)^{\frac{1}{2}} H^{-\frac{d}{2}} \|F_1(U_{ms})\|_V \\ &\leq C_\kappa (1 + C_\kappa^2)^{\frac{1}{2}} C^{\frac{1}{2}} \Big((1 + \tilde{E}^{-1})^{1-M} + (1 + E^{-1})^{1-N_s} \Big)^{\frac{1}{2}} H^{-\frac{d}{2}} \\ &\quad \times \Big(\|F_1(U_{ms}) - F_1(U_{glo})\|_V + \|F_1(U_{glo})\|_V \Big).\end{aligned}$$

If $C_\kappa (1 + C_\kappa^2)^{\frac{1}{2}} C^{\frac{1}{2}} \Big((1 + \tilde{E}^{-1})^{1-M} + (1 + E^{-1})^{1-N} \Big)^{\frac{1}{2}} H^{-\frac{d}{2}} \leq \dfrac{1}{2}$, we have

$$\|F_1(U_{ms}) - F_1(U_{glo})\|_V \leq 2 C_\kappa (1 + C_\kappa^2)^{\frac{1}{2}} C^{\frac{1}{2}} \Big((1 + \tilde{E}^{-1})^{1-M} + (1 + E^{-1})^{1-N} \Big)^{\frac{1}{2}} H^{-\frac{d}{2}} \|F_1(U_{glo})\|_V.$$

□

Remark: If the multiscale partition of unity χ_i is replaced by the bilinear partition of unity in the definition of $\tilde{\kappa}$, one can easily prove that with an appropriate choice of the spatial and temporal oversampling layers, we have

$$\|u - F_{ms,1}(U_{ms})\|_V \lesssim CH \|\kappa f\|_{L^2} + CH \|\kappa \partial_t u\|_{L^2} + CH \|F_1(U_{glo})\|_V.$$

5.4 Numerical results

In this section, we present numerical results for the proposed numerical method. We shall solve the system (5.1) in the unit square $\Omega = [0, 1]^2$ with total time $T = 1.0$. The source term $f(t, x)$ is chosen to be a smooth function $f(t, x) := x_1 x_2 t$. The permeability

filed $\kappa(t,x)$ is time-dependent. We will test our numerical methods with two kinds of permeability field: slow moving permeability in Experiment 1 and faster moving permeability in Experiment 2.

Let $\mathcal{T}_H \times \mathcal{T}_{\Delta t}$ be a decomposition of the space-time domain $\Omega \times [0,T]$ into non-overlapping shape-regular cubic elements with maximal spatial mesh size H and temporal mesh size Δt. These coarse cubic elements are further partitioned into a collection of connected fine cubic elements $\mathcal{T}_h \times \mathcal{T}_{\delta t}$ using fine spatial mesh size h and temporal mesh size δt. Similarly, we define $V_{h,\delta}$ to be a conforming piecewise affine finite element associated with $\mathcal{T}_h \times \mathcal{T}_{\delta t}$. Since there is no analytic solution to system (5.1) we need to find an approximation of the exact solutions. To this end, we use the constructed fine mesh and conforming space-time finite element method to obtain the reference solutions $U_{h,\delta}$. The multiscale solutions $U_{\text{ms}}^{\ell_x,\ell_t}$ are obtained using our proposed space-time NLMC method with spatial oversampling layers number being ℓ_x and temporal oversampling layers number being ℓ_t. We use $\tilde{U}_{h,\delta,t}$ to denote the snapshot of the reference solutions at time t and $\tilde{U}_{\text{ms},t}^{\ell_x,\ell_t}$ to denote the snapshot of multiscale solutions using spatial oversampling layer ℓ_x and temporal oversampling layer ℓ_t at time t. To simply notations, we use $\tilde{U}_{\text{ms},t}^{\ell}$ to denote $\tilde{U}_{\text{ms},t}^{\ell_x,\ell_t}$ when the number of spatial oversampling layers equals that of temporal oversampling layers, that is, $\ell := \ell_x = \ell_t$.

We introduce the following notations to calculate the errors. The relative errors for the multiscale solution in L^2-norm and H^1_κ-norm are

$$\text{Rel}_{L^2}^{\ell} := \frac{\left\|U_{h,\delta} - U_{\text{ms}}^{\ell_x,\ell_t}\right\|_{L^2}}{\left\|U_{h,\delta}\right\|_{L^2}} \times 100 \quad \text{and} \quad \text{Rel}_{H^1_\kappa}^{\ell} := \frac{\left\|U_{h,\delta} - U_{\text{ms}}^{\ell_x,\ell_t}\right\|_{H^1_\kappa}}{\left\|U_{h,\delta}\right\|_{H^1_\kappa}} \times 100.$$

5.4.1 Experiment 1: Slow moving permeability

In this experiment, we choose the permeability with 1 channel moving slowly in horizontal direction. Let

$$S_1 := \{(x_1, x_2, t) : 0.375 < x_1 < 0.6094, 0.50 < x_2 < 0.5156, 0 \leq t < 0.5\}$$

and

$$S_2 := \{(x_1, x_2, t) : 0.3906 < x_1 < 0.6250, 0.50 < x_2 < 0.5156, 0.5 \leq t \leq 1.0\}.$$

The permeability $\kappa(x_1, x_2, t)$ is defined as

$$\kappa(x_1, x_2, t) := \begin{cases} 1000, \text{ if } (x_1, x_2, t) \in S_1 \cup S_2, \\ 1, \text{ otherwise.} \end{cases}$$

We present the permeability field at time $t = 0$ and $t = 0.5$ in Figure 5.2 for an illustration.

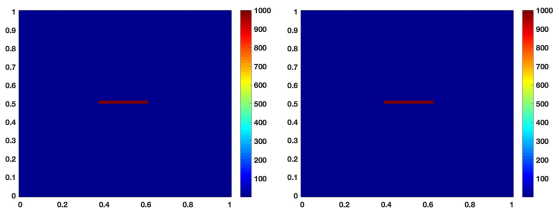

Figure 5.2: Permeability field $\kappa(x_1, x_2, 0)$ and $\kappa(x_1, x_2, 0.5)$ for Experiment 1.

We choose fine spatial mesh size $h = 2^{-6}$, fine temporal mesh size $\delta t = 0.01$, coarse spatial mesh size $H = 2^{-3}$ and coarse temporal mesh size $\Delta t = 0.1$. The number of spatial and

temporal oversampling layers ℓ_x and ℓ_t are chosen to be $\ell = \ell_x = \ell_t \in \{1, 2, \cdots, 5\}$. The snapshot of reference solutions $\tilde{U}_{h,\delta,t}$ and multiscale solutions $\tilde{U}^{\ell}_{\mathrm{ms},t}$ for $t = 0.25, 0.5, 0.75, 1.0$ and $\ell = 1, 2, 3$ are plotted in Figure 5.3 and Figure 5.4, respectively.

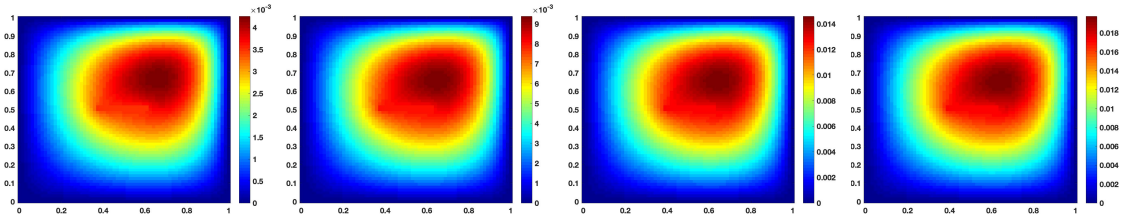

Figure 5.3: Snapshot of the reference solutions $\tilde{U}_{h,\delta,t}$ for $t = 0.25, 0.5, 0.75, 1.0$.

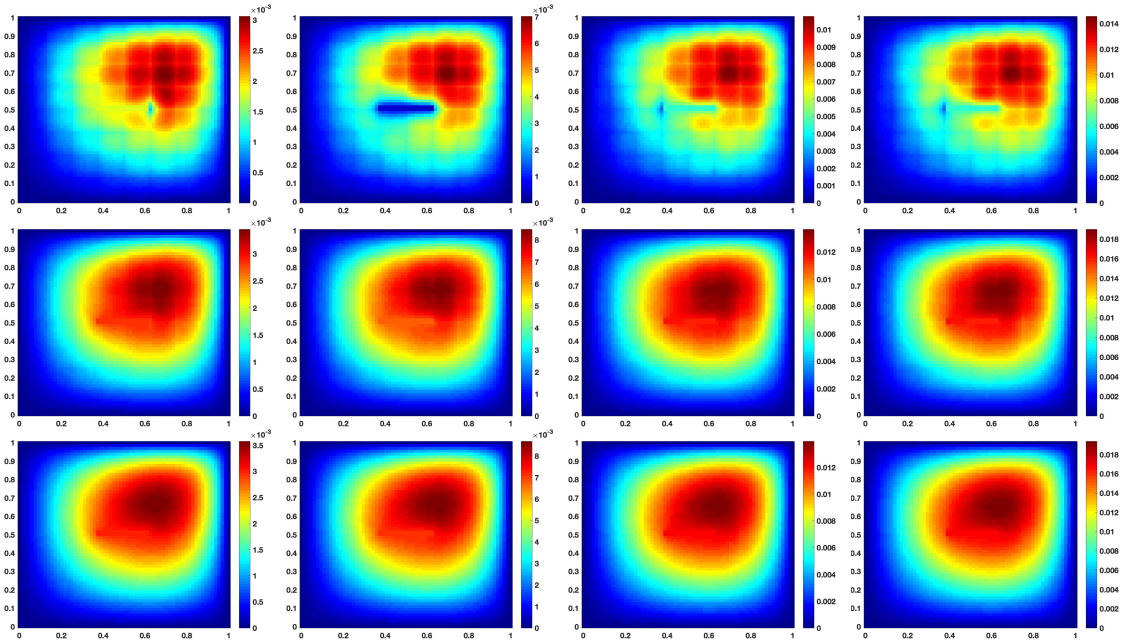

Figure 5.4: Snapshot of the multiscale solutions $\tilde{U}^{\ell}_{\mathrm{ms},t}$ at $t = 0.25, 0.5, 0.75, 1.0$ with oversampling layer $\ell = 1$ (top), $\ell = 2$ (middle), $\ell = 3$ (bottom).

The convergence history in relative L^2-norm and relative H^1_κ-norm with oversampling layers number $\ell = 1, 2, \cdots, 5$ are presented in Table 5.1.

ℓ	$\text{Rel}^\ell_{H^1_\kappa}$	$\text{Rel}^\ell_{L^2}$
1	53.6304	35.6654
2	15.2632	5.0203
3	7.2096	3.3863
4	4.3655	2.7838
5	3.4061	2.5349

Table 5.1: Convergence history of Experiment 1.

5.4.2 Experiment 2: Faster moving permeability

In this experiment, we choose a permeability with faster moving channels. To define the permeability for this experiment, we introduce 4 sets S_1, S_2, S_3 and S_4 as follows.

$$
\begin{aligned}
S_1 := \cup_{k=1}^{25}\{(x_1, x_2, t) : 0.09 + 0.01k \le x_1 \le 0.11 + 0.01k, 0.30 \le x_2 \le 0.70,\\
0.04(k-1) \le t \le 0.04k\},\\
S_2 := \cup_{k=1}^{20}\{(x_1, x_2, t) : 0.39 + 0.01k \le x_1 \le 0.79 + 0.01k, 0.15 \le x_2 \le 0.17,\\
0.05(k-1) \le t \le 0.05k\},\\
S_3 := \cup_{k=1}^{25}\{(x_1, x_2, t) : 0.29 + 0.01k \le x_1 \le 0.44 + 0.01k, 0.19 + 0.01k \le x_2 \le 0.21 + 0.01k,\\
0.04(k-1) \le t \le 0.04k\},\\
S_4 := \cup_{k=1}^{10}\{(x_1, x_2, t) : 0.59 + 0.01k \le x_1 \le 0.94 + 0.01k, 0.63 + 0.01k \le x_2 \le 0.65 + 0.01k,\\
0.1(k-1) \le t \le 0.1k\}.
\end{aligned}
$$

The permeability $\kappa(x_1, x_2, t)$ is defined as below:

$$\kappa(x_1, x_2, t) := \begin{cases} 1000, \text{ if } (x_1, x_2, t) \in S_1 \cup S_2 \cup S_3 \cup S_4, \\ \\ 1, \text{ otherwise.} \end{cases}$$

We present the permeability field at time $t = 0, 0.5, 0.8$ and $t = 1.0$ in Figure 5.5 for an illustration.

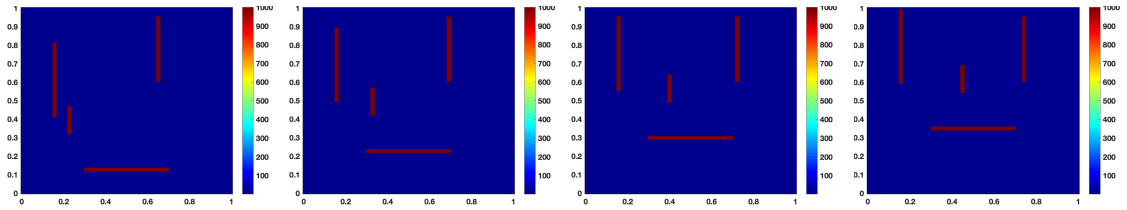

Figure 5.5: Permeability field $\kappa(x_1, x_2, 0)$, $\kappa(x_1, x_2, 0.5)$, $\kappa(x_1, x_2, 0.8)$ and $\kappa(x_1, x_2, 1.0)$.

The spatial and temporal fine mesh size we use to approximate the exact solution is $h = 0.01$ and $\delta t = 0.01$. The snapshot of reference solutions $\tilde{U}_h(t)$ to approximate exact solution $u(x, t)$ at time $t = 0.2, 0.5, 0.8, 1.0$ are plotted as below.

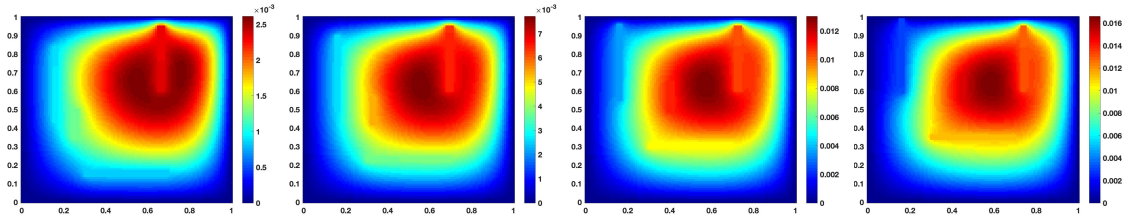

Figure 5.6: Snapshot of the reference solutions $\tilde{U}_{h,\delta,t}$ for $t = 0.2, 0.5, 0.8, 1.0$.

The coarse spatial and temporal mesh size we use is $H = 0.1$ and $\Delta t = 0.1$. The

number of spatial and temporal oversampling layers ℓ_x and ℓ_t are chosen to be $\ell = \ell_x = \ell_t \in \{1, 2, \cdots, 5\}$. We present the snapshot of numerical solutions $\tilde{U}^{\ell}_{\text{ms},t}$ for $t = 0.2, 0.5, 0.8, 1.0$ with the oversampling layer $\ell = 1, 2, 3$ in Figure 5.7.

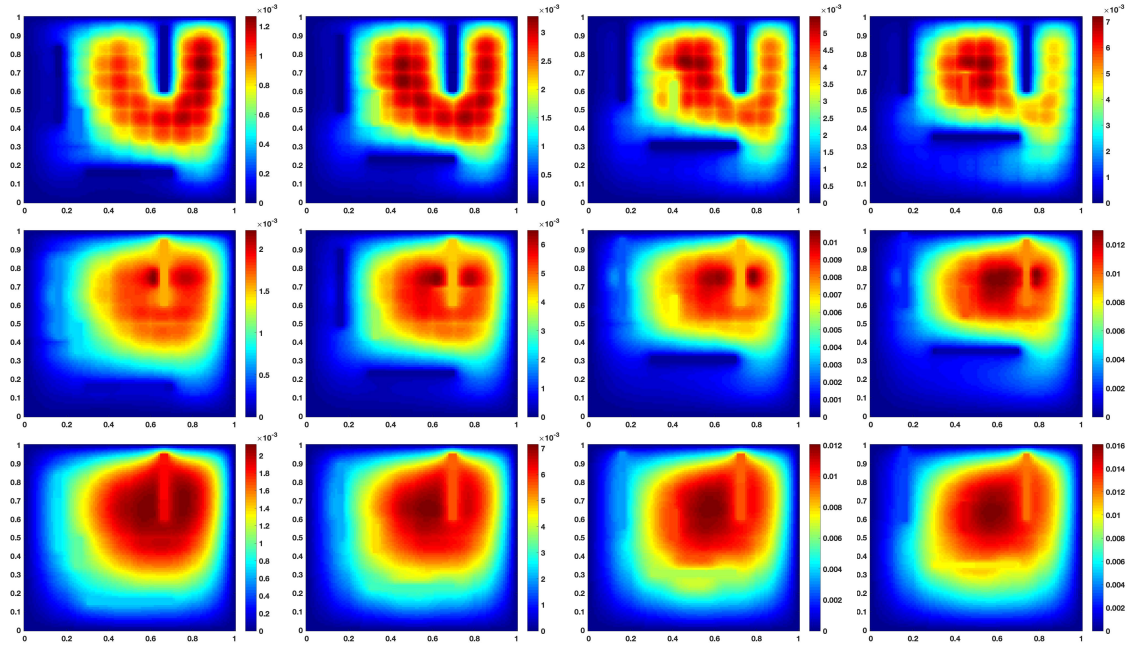

Figure 5.7: Snapshot of the multiscale solutions $\tilde{U}^{\ell}_{\text{ms},t}$ at $t = 0.2, 0.5, 0.8, 1.0$ with oversampling layer $\ell = 1$ (top), $\ell = 2$ (middle), $\ell = 3$ (bottom).

The convergence history in relative L^2-norm and relative H^1_κ-norm with oversampling layers number $\ell = 1, 2, \cdots, 5$ are presented in Table 5.2.

ℓ	$\mathrm{Rel}^{\ell}_{H^1_\kappa}$	$\mathrm{Rel}^{\ell}_{L^2}$
1	80.2825	68.3637
2	51.5355	22.0861
3	17.1313	5.1881
4	0.5724	0.0658
5	0.1876	0.04265

Table 5.2: Convergence history of Experiment 2.

6. SUMMARY AND CONCLUSIONS

We conclude this dissertation with a brief summary. Direct treatment of multiscale problems is challenging, since resolving the problems to the finest scale would incur huge computational cost. To reduce the computational expense, we make use of model reduction techniques and the increasing computational capacity of the current computers.

We study in Chapter 2 homogenization approach for the initial boundary value problem for the time-fractional diffusion equation with a homogeneous Dirichlet boundary condition and an inhomogeneous initial data in a bounded convex polyhedral domain. We analyze the homogenized solution under the assumption that the diffusion coefficient is smooth and periodic with period being a small parameter.

We next consider in Chapter 3 the incompressible Stokes flow problem in a perforated domain and employ the constraint energy minimizing generalized multiscale finite element method to solve this problem. The proposed method provides a flexible and systematical approach to construct crucial divergence-free multiscale basis functions for approximating the velocity field. These basis functions are constructed by solving a class of local energy minimization problems over the eigenspaces that contain local information on the heterogeneities. These multiscale basis functions are shown to have the property of exponential decay outside the corresponding local oversampling regions. By adopting the technique of oversampling, the spectral convergence of the method with error bounds related to the coarse mesh size is proved.

Parabolic equations are investigated in Chapter 4 and Chapter 5. In Chapter 4, we propose the Wavelet-based Edge Multiscale Parareal Algorithm to solve parabolic equations with heterogeneous time-independent coefficients. This algorithm combines the advantages of multiscale methods that can deal with heterogeneity in the spatial domain effec-

tively, and the strength of parareal algorithms for speeding up time evolution problems when sufficient processors are available. We derive the convergence rate of this algorithm in terms of the mesh size in the spatial domain, the level parameter used in the multiscale method, the coarse-scale time step and the fine-scale time step.

Finally, we consider the parabolic equations with time-dependent heterogeneous coefficients in Chapter 5. Many applied problems have coupled space and time heterogeneities. Their homogenization or upscaling requires cell problems that are formulated in space-time representative volumes for problems with scale separation. In problems without scale separation, local problems include multiple macroscopic variables and oversampled local problems, where these macroscopic parameters are computed. These approaches, called Non-local multi-continua, are proposed for problems with complex spatial heterogeneities in a number of previous papers. We extend this approach for space-time heterogeneities, by identifying macroscopic parameters in space-time regions. Our proposed method space-time NLMC is an efficient numerical solver to deal with time-dependent heterogeneous coefficients. It provides a flexible and systematic way to construct multiscale basis functions to approximate the solution. These multiscale basis functions are constructed by solving a local energy minimization problems in the oversampled space-time regions such that these multiscale basis functions decay exponentially outside the oversampled domain. Unlike the classical time-stepping methods combined with full-discretization technique, our space-time NLMC efficiently constructs the multiscale basis functions in a space-time domain and can provide a computational savings compared to space-only approaches.

REFERENCES

[1] Geoscale - direct reservoir simulation on geocellular models. https://www.sintef.no/projectweb/geoscale/results/msmfem/spe10/.

[2] Assyr Abdulle and Patrick Henning. Localized orthogonal decomposition method for the wave equation with a continuum of scales. *Mathematics of Computation*, 86(304):549–587, 2017.

[3] Assyr Abdulle and Gilles Vilmart. Analysis of the finite element heterogeneous multiscale method for quasilinear elliptic homogenization problems. *Mathematics of Computation*, 83(286):513–536, 2014.

[4] Grégoire Allaire. Homogenization of the stokes flow in a connected porous medium. *Asymptotic Analysis*, 2(3):203–222, 1989.

[5] Grégoire Allaire and Robert Brizzi. A multiscale finite element method for numerical homogenization. *Multiscale Modeling & Simulation*, 4(3):790–812, 2005.

[6] Todd Arbogast and Heather L Lehr. Homogenization of a darcy–stokes system modeling vuggy porous media. *Computational Geosciences*, 10(3):291–302, 2006.

[7] I. Babuška and J. M. Melenk. The partition of unity method. *Int. J. Numer. Meth. Engrg.*, 40:727–758, 1997.

[8] Ivo Babuška. Homogenization and its application. mathematical and computational problems. In *Numerical solution of partial differential equations–III*, pages 89–116. Elsevier, 1976.

[9] Guillaume Bal. On the convergence and the stability of the parareal algorithm to solve partial differential equations. In *Domain decomposition methods in science and engineering*, pages 425–432. Springer, 2005.

[10] Leonid Berlyand and Houman Owhadi. Flux norm approach to finite dimensional homogenization approximations with non-separated scales and high contrast. *Arch. Ration. Mech. Anal.*, 198(2):677–721, 2010.

[11] Charles-Edouard Bréhier. Analysis of an hmm time-discretization scheme for a system of stochastic pdes. *SIAM Journal on Numerical Analysis*, 51(2):1185–1210, 2013.

[12] Zhiming Chen and Thomas Hou. A mixed multiscale finite element method for elliptic problems with oscillating coefficients. *Mathematics of Computation*, 72(242):541–576, 2003.

[13] C-C Chu, Ivan Graham, and T-Y Hou. A new multiscale finite element method for high-contrast elliptic interface problems. *Mathematics of Computation*, 79(272):1915–1955, 2010.

[14] Eric Chung, Yalchin Efendiev, and Thomas Y. Hou. Adaptive multiscale model reduction with generalized multiscale finite element methods. *J. Comput. Phys.*, 320:69–95, 2016.

[15] Eric Chung, Yalchin Efendiev, and Wing Tat Leung. Constraint energy minimizing generalized multiscale finite element method. *Comput. Methods Appl. Mech. Engrg.*, 339:298–319, 2018.

[16] Eric T Chung, Yalchin Efendiev, and Chak Shing Lee. Mixed generalized multiscale finite element methods and applications. *Multiscale Modeling & Simulation*, 13(1):338–366, 2015.

[17] Eric T Chung, Yalchin Efendiev, and Wing Tat Leung. Generalized multiscale finite element methods for wave propagation in heterogeneous media. *Multiscale Modeling & Simulation*, 12(4):1691–1721, 2014.

[18] Eric T Chung, Yalchin Efendiev, and Wing Tat Leung. Residual-driven online generalized multiscale finite element methods. *Journal of Computational Physics*, 302:176–190, 2015.

[19] Eric T Chung, Yalchin Efendiev, Wing Tat Leung, Maria Vasilyeva, and Yating Wang. Non-local multi-continua upscaling for flows in heterogeneous fractured media. *Journal of Computational Physics*, 372:22–34, 2018.

[20] Philippe G Ciarlet. *Linear and Nonlinear Functional Analysis with Applications*. Society for Industrial and Applied Mathematics, 2013.

[21] Weinan E and Bjorn Engquist. The heterogeneous multiscale methods. *Commun. Math. Sci.*, 1(1):87–132, 2003.

[22] Y. Efendiev, J. Galvis, and T. Hou. Generalized multiscale finite element methods. *J. Comput. Phys.*, 251:116–135, 2013.

[23] Yalchin Efendiev, Juan Galvis, Guanglian Li, and Michael Presho. Generalized multiscale finite element methods: Oversampling strategies. *International Journal for Multiscale Computational Engineering*, 12(6), 2014.

[24] Yalchin Efendiev, Juan Galvis, and Xiao-Hui Wu. Multiscale finite element methods for high-contrast problems using local spectral basis functions. *Journal of Computational Physics*, 230(4):937–955, 2011.

[25] Yalchin Efendiev and Thomas Y Hou. *Multiscale finite element methods: theory and applications*, volume 4. Springer Science & Business Media, 2009.

[26] Yalchin R Efendiev, Thomas Y Hou, and Xiao-Hui Wu. Convergence of a nonconforming multiscale finite element method. *SIAM Journal on Numerical Analysis*, 37(3):888–910, 2000.

[27] Christian Engwer, Patrick Henning, Axel Målqvist, and Daniel Peterseim. Efficient implementation of the localized orthogonal decomposition method. *Computer Methods in Applied Mechanics and Engineering*, 350:123–153, 2019.

[28] Alexandre Ern and Jean-Luc Guermond. *Theory and practice of finite elements*, volume 159. Springer Science & Business Media, 2013.

[29] Shubin Fu, Eric Chung, and Guanglian Li. Edge multiscale methods for elliptic problems with heterogeneous coefficients. *J. Comput. Phys*, 369(1):228–242, 2019.

[30] Shubin Fu, Guanglian Li, Richard Craster, and Sebastien Guenneau. Wavelet-based edge multiscale finite element method for helmholtz problems in perforated domains. *Submitted*, 2019.

[31] Massimiliano Giona, Stefano Cerbelli, and H Eduardo Roman. Fractional diffusion equation and relaxation in complex viscoelastic materials. *Physica A: Statistical Mechanics and its Applications*, 191(1-4):449–453, 1992.

[32] Pierre Grisvard. *Elliptic problems in nonsmooth domains*, volume 69. SIAM, 2011.

[33] Shu Gu. Convergence rates in homogenization of stokes systems. *Journal of Differential Equations*, 260(7):5796–5815, 2016.

[34] Max D Gunzburger and Janet S Peterson. On conforming finite element methods for the inhomogeneous stationary Navier-Stokes equations. *Numerische Mathematik*, 42(2):173–194, 1983.

[35] Patrick Henning and Axel Målqvist. Localized orthogonal decomposition techniques for boundary value problems. *SIAM Journal on Scientific Computing*, 36(4):A1609–A1634, 2014.

[36] Thomas Hou and Xiao-Hui Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *J. Comput. Phys.*, 134(1):169–189, 1997.

[37] Thomas Hughes, Gonzalo Feijóo, Luca Mazzei, and Jean-Baptiste Quincy. The variational multiscale method—a paradigm for computational mechanics. *Comput. Methods Appl. Mech. Engrg.*, 166(1-2):3–24, 1998.

[38] V. V. Jikov, S. M. Kozlov, and O. A. Oleĭnik. *Homogenization of Differential Operators and Integral Functionals*. Springer-Verlag, Berlin, 1994. Translated from the Russian by G. A. Yosifian.

[39] Vasili Vasilievitch Jikov, Sergei M Kozlov, and Olga Arsenievna Oleinik. *Homogenization of differential operators and integral functionals*. Springer Science & Business Media, 2012.

[40] Bangti Jin, Raytcho Lazarov, Joseph Pasciak, and Zhi Zhou. Error analysis of semidiscrete finite element methods for inhomogeneous time-fractional diffusion. *IMA J. Numer. Anal.*, 35(2):561–582, 2015.

[41] Bangti Jin, Raytcho Lazarov, and Zhi Zhou. Numerical methods for time-fractional evolution equations with nonsmooth data: a concise overview. *Computer Methods in Applied Mechanics and Engineering*, 346:332–358, 2019.

[42] Volker John and Songul Kaya. A finite element variational multiscale method for the navier–stokes equations. *SIAM Journal on Scientific Computing*, 26(5):1485–1503, 2005.

[43] Volker John, Songul Kaya, and William Layton. A two-level variational multiscale method for convection-dominated convection–diffusion equations. *Computer Methods in Applied Mechanics and Engineering*, 195(33-36):4594–4603, 2006.

[44] Anatoly A. Kilbas, Hari M. Srivastava, and Juan J. Trujillo. *Theory and Applications of Fractional Differential Equations*, volume 204 of *North-Holland Mathematics Studies*. Elsevier Science B.V., Amsterdam, 2006.

[45] Samuel C Kou. Stochastic modeling in nanoscale biophysics: subdiffusion within proteins. *The Annals of Applied Statistics*, 2(2):501–535, 2008.

[46] Olga Aleksandrovna Ladyzhenskaia, Vsevolod Alekseevich Solonnikov, and Nina N Ural'tseva. *Linear and quasi-linear equations of parabolic type*, volume 23. American Mathematical Soc., 1988.

[47] Guanglian Li. On the convergence rates of gmsfems for heterogeneous elliptic problems without oversampling techniques. *Multiscale Modeling & Simulation*, 17(2):593–619, 2019.

[48] Guanglian Li, Daniel Peterseim, and Mira Schedensack. Error analysis of a variational multiscale stabilization for convection-dominated diffusion equations in two dimensions. *IMA J. Numer. Anal.*, 38(3):1229–1253, 2018.

[49] Yumin Lin and Chuanju Xu. Finite difference/spectral approximations for the time-fractional diffusion equation. *J. Comput. Phys.*, 225(2):1533–1552, 2007.

[50] Mitchell Luskin and Rolf Rannacher. On the smoothing property of the Crank-Nicolson scheme. *Applicable Anal.*, 14(2):117–135, 1982/83.

[51] Axel Målqvist and Daniel Peterseim. Localization of elliptic multiscale problems. *Mathematics of Computation*, 83(290):2583–2603, 2014.

[52] Eduard Marušić-Paloka and Andro Mikelić. An error estimate for correctors in the homogenization of the Stokes and the Navier-Stokes equations in a porous medium. *Boll. Un. Mat. Ital. A (7)*, 10(3):661–671, 1996.

[53] Bagus Putra Muljadi, Jacek Narski, Alexei Lozinski, and Pierre Degond. Nonconforming multiscale finite element method for stokes flows in heterogeneous media. Part I: methodologies and numerical experiments. *Multiscale Modeling & Simulation*, 13(4):1146–1172, 2015.

[54] RR Nigmatullin. The realization of the generalized transfer equation in a medium with fractal geometry. *Physica Status Solidi (b)*, 133(1):425–430, 1986.

[55] Svetlana Pastukhova. Estimates in homogenization of parabolic equations with locally periodic coefficients. *Asymptotic Analysis*, 66(3-4):207–228, 2010.

[56] Igor Podlubny. *Fractional Differential Equations*, volume 198 of *Mathematics in Science and Engineering*. Academic Press, Inc., San Diego, CA, 1999. An introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications.

[57] Kenichi Sakamoto and Masahiro Yamamoto. Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems. *J. Math. Anal. Appl.*, 382(1):426–447, 2011.

[58] Harvey Scher and Elliott W Montroll. Anomalous transit-time dispersion in amorphous solids. *Physical Review B*, 12(6):2455, 1975.

[59] K. Copps T. Strouboulis, I. Babuška. The design and analysis of the generalized finite element method. *Comput. Methods Appl. Mech. Engrg.*, 181:43–69, 2000.

[60] Vidar Thomée. *Galerkin finite element methods for parabolic problems*, volume 1054. Springer, 2006.

[61] AL Vanel, O Schnitzer, and RV Craster. Asymptotic network models of subwavelength metamaterials formed by closely packed photonic and phononic crystals. *EPL (Europhysics Letters)*, 119, 2017.

[62] M Vassaux, RA Richardson, and PV Coveney. The heterogeneous multiscale method applied to inelastic polymer mechanics. *Philosophical Transactions of the Royal Society A*, 377(2142):20180150, 2019.

[63] Shu-Lin Wu and Tao Zhou. Convergence analysis for three parareal solvers. *SIAM J. Sci. Comput.*, 37(2):A970–A992, 2015.

[64] Lina Zhao and Eric T Chung. An analysis of the nlmc upscaling method for high contrast problems. *Journal of Computational and Applied Mathematics*, 367:112480, 2020.

[65] V. V. Zhikov and S. E. Pastukhova. Estimates of homogenization for a parabolic equation with periodic coefficients. *Russian Journal of Mathematical Physics*, 13(2):224–237, Apr 2006.