GMSFEM FOR NONLINEAR PROBLEMS & SPACE-TIME GMSFEM

A Dissertation

by

SHUAI YE

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

DOCTOR OF PHILOSOPHY

Chair of Committee,	Yalchin Efendiev
Committee Members,	Raytcho Lazarov
	Jianxin Zhou
	Partha Mukherjee
Head of Department,	Emil Straube

December 2016

Major Subject: Mathematics

ABSTRACT

Many engineering and scientific applications deal with models that have multiple spatial scales, and these scales can be non-separable. Many of these processes can exhibit nonlinearities and have a tight coupling with the temporal scales. Because of the scale disparity, modeling these processes in the fine-scale approaches often involves the use of a large number of degrees of freedom, and thus can be prohibitively expensive. As such, some efficient model reduction methods are required to handle the multiscale problems.

In the dissertation, we are solving multiscale nonlinear problems and time-dependent problems. Existing methods to solve these problems include numerical homogenization, multiscale finite element methods, heterogeneous multiscale methods, and the Generalized Multiscale Finite Element Methods (GMsFEM). GMsFEM approaches propose a systematic enrichment, which calculates multiscale basis functions via local spectral decomposition in each coarse cell.

We first propose a multiscale model reduction framework within GMsFEM for nonlinear elliptic problems. We consider an exemplary problem, which consists of nonlinear p-Laplacian with heterogeneous coefficients. The main challenging feature of this problem is that local subgrid models are nonlinear involving the gradient of the solution. Our novel work includes re-casting the multiscale model reduction problem onto the boundaries of coarse cells, and introducing nonlinear eigenvalue problems in the snapshot space for these nonlinear "harmonic" functions. We also present convergence analysis and numerical results, which show that our approaches can recover the fine-scale solution with a few degrees of freedom. The proposed methods can, in general, be used for more general nonlinear problems, where one needs nonlinear local spectral decomposition.

We next consider solving problems with multiple scales in space and time. We de-

velop our approaches within the frameworks of GMsFEM and Generalized Multiscale Discontinuous Galerkin Methods (GMsDGM), separately, using space-time coarse cells. Previous research in developing multiscale spaces within GMsFEM or GMsDGM mainly considered spatial multiscale spaces and relevant ingredients only, which will usually lead to very high dimensional models. In the dissertation, we construct space-time offline spaces, where local spectral decomposition methods are designed based on our analysis. We also discuss adding the online stage, where we make use of the online residual information to construct online basis functions. Numerical results are presented to verify the theoretical findings and to show that using our proposed approaches, we can obtain an accurate solution with low dimensional coarse spaces.

ACKNOWLEDGEMENTS

Firstly, I would like to express my deepest gratitude to my advisor Prof. Yalchin Efendiev who has been the beacon of my research life. He introduced me to the world of scientific research and has been offering valuable and insightful guidance throughout my PhD life in Texas A&M University. Prof. Efendiev is a respectful mathematician with distinguished research achievements and the most hard-working one I have ever known. More importantly, he always understands his students and is able to offer advice when it's most needed. It has been my pleasure to work with him. Besides, Prof. Efendiev has been very generous in supporting me with Research Assistantship for several semesters, which allowed me to be more focused on my research work.

Secondly, I'm very grateful to Dr. Michael Presho, Prof. Eric Chung, Prof. Ke Shi and my co-worker Wing Tat Leung. All of them have been very patient in explaining research ideas to me and have given me many useful suggestions whenever I encounter difficulties. Their selfless help makes my PhD path much more smooth.

I also want to thank Prof. Raytcho Lazarov, Prof. Jianxin Zhou and Prof. Partha Mukherjee for serving on my dissertation committee. Moreover, I appreciate Department of Mathematics of Texas A&M for providing me an awesome environment to develop my skills and knowledge. Prof. Peter Howard and Mrs. Steward Monique along with all other staff in the department care a lot about the graduate students and their efforts to make the department a loving home for grads can never be overpraised. Special thanks also go to my friends and fellow students for sharing with me the ups and downs during my PhD life.

Last but not least, I want to thank my wonderful parents and brother for the invaluable love and support they have been giving me all these years. That always charges me up and pushes me to move forward.

NOMENCLATURE

$\mathbb{R}^d$	d-dimensional Euclidean space
Ω	Simply-connected bounded open subset of $\mathbb{R}^d$
$\mathcal{T}^h$	Partition of Ω into fine grid elements
h	Characteristic length of a fine grid element
$\mathcal{T}^H$	Partition of Ω into coarse grid elements
H	Characteristic length of a coarse grid element
x_i	Coarse grid vertex
ω_i	Coarse neighborhood with a common vertex at x_i
ω_i^+	Oversampled coarse neighborhood for ω_i
K	Coarse-grid block
K^+	Oversampled coarse-grid domain that contains K
u	Primal variable
u_h	Fine-grid solution
u_H	Coarse-grid solution
$u_H^{(n)}$	Coarse-grid space-time solution defined on (T_{n-1}, T_n)
V_h	Fine-scale space spanned by polynomials
χ	Partition of unity function

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

I. INTRODUCTION

I.1 Motivations

I.1.1 Multiscale and nonlinearity

Many processes in nature have multiscale nature and nonlinearities. The interaction between nonlinearities and multiple scales can be complex and non-separable. This occurs in many applications. To discuss some main concepts, we consider an example

$$\text{div}(a(x, u, \nabla u)) = f. \tag{I.1}$$

We assume $a(x, \cdot, \cdot)$ is highly heterogeneous with respect to x. In some nonlinear problems, the nonlinearities within coarse regions (a computational grid), that induce the change in the heterogeneities, can be parametrized with a low dimensional parameter, e.g., $a(x, u, \nabla u) = a_0(x, u)\nabla u$ (assuming smoothness and boundedness of a). Within each coarse region, one can approximate the solution u by a constant and thus, can handle these nonlinearities via a low dimensional parametrization (see, e.g., [34, 55] for homogenization and numerical homogenization discussions). If the nonlinearities and heterogeneities are separable in this case, i.e., $a(x, u, \nabla u) = a(x)b(u)\nabla u$, then, in fact, one can use a linear theory of multiscale methods (cf. [34, 55]). The situation is very different when $a(x, u, \nabla u) = a_1(x, \nabla u)\nabla u$. Because ∇u is highly heterogeneous, one can not use any low dimensional approximation and linear theories. This is true even for a separable case $a(x, u, \nabla u) = a_0(x)b(|\nabla u|)\nabla u$. These problems require nonlinear cell problems. In the dissertation, we focus on the case $a(x, u, \nabla u) = a(x, \nabla u)$.

Many previous research on multiscale methods have considered nonlinear problems. The approaches including homogenization [56, 46], numerical homogenization [4, 63, 24],

heterogeneous multiscale methods [25, 2, 40, 50], multiscale network approximations [9, 10, 8], multiscale finite element methods [24, 63, 5], variational multiscale methods [43, 42, 7, 44], polyharmonic homogenization [54, 11], generalized multiscale finite element methods [26, 28] have been developed and applied. These approaches approximate the solution of nonlinear PDEs on a coarse grid (see Figure II.1 for illustration of coarse and fine grids) by using subgrid models. Some common ingredients in these methods for linear problems are that local solutions are calculated and used to form equations on a coarse grid. GMsFEM approaches propose a systematic enrichment, which calculates multiscale basis functions via local spectral decomposition in each coarse cell. The extensions of these methods to nonlinear problems (as (I.1)) use nonlinear local problems.

The main idea of GMsFEM for linear problems is to form snapshot spaces and perform local spectral decomposition in the snapshot space. In this dissertation, we will follow the same general concept and introduce nonlinear eigenvalue problems. Previous approach [29] develops local nonlinear eigenvalue problems in each coarse cell. In Section III, we propose a systematic model reduction using nonlinear harmonic functions. The latter is important as it allows capturing the effects of separable scales. Without using nonlinear harmonic functions, one can not, in general, capture the effects of small separable scales. This is in contrast to linear problems, where one can construct one linear basis function per every coarse node that contains the effects of small scales. Using local solutions allows compressing the effects of small scales within a coarse block and we work with a system reduced to the boundaries of coarse cells. In this case, we can also guarantee that our approaches recover homogenization results when there is a scale separation (note that previous approaches [29] can not guarantee it). The proposed method is related to hybridization techniques [23, 32, 14, 33].

I.1.2 Multiscale in space and time

A wide variety of multiscale problems vary over multiple time and space scales. These time and space scales are often tightly coupled. For example, flow processes involving porous media can occur on multiple time scales over multiple spatial scales. Moreover, these scales can be non-separable. Reduced-order models for these problems require simultaneously treating spatial and temporal scales. Many previous approaches only handle spatial scales and spatial heterogeneities. These approaches have limitations when temporal heterogeneities arise. Some popular approaches for handling separable space and time scales are homogenization techniques [46, 55, 57, 36]. In these methods, one solves local problems These approaches work well in the scale separation cases, but do not provide accurate approximations when there is no scale separation.

Previous researchers developed a number of multiscale methods for solving space-time multiscale problems in the absence of scale separation. These approaches use Multiscale Finite Element Methods [41, 31, 47, 35], where one computes multiscale space-time basis functions, variational multiscale methods [45], and other approaches [60, 61, 52, 49] that are developed for stabilization. In [53], Owhadi and Zhang proposed a novel approach that uses global space-time information in computing multiscale basis functions. All these approaches use only a limited number of basis functions (one basis function) in each coarse block. We note that there has been a large body of works in space-time finite element methods.

In this dissertation, we use the GMsFEM framework and and GMsDGM framework, respectively, and develop a systematic approach for identifying multiscale basis functions. Previous approaches mainly considered multiscale spaces together with corresponding ingredients in space only when developing multiscale spaces within GMsFEM or GMsDGM. In our approaches, we develop snapshot spaces by solving local problems in local

space-time domains. We use the techniques of oversampling and randomization by solving local parabolic equations subject to random boundary and initial conditions, which can reduce the computational cost. We then solve local spectral problems to obtain offline multiscale basis functions in space-time domain. We discuss several choices for local spectral problems and present an analysis of the convergence rate of the proposed method.

The multiscale online basis functions are introduced in [15, 16]. The main idea of online basis functions is using the online residual information and add new multiscale basis functions to the existing offline space. One need to select a number of offline basis functions such that with only 1-2 online iterations, the error can be substantially reduced. In this dissertation, we will propose a possible online construction along with the space-time GMsFEM approach in Section IV.

I.2 Outline of the dissertation

In Section II, we present some preliminary background concepts that would be used in the march of this dissertation. We introduce the coarse-scale and fins-scale grids, which is the mesh we are using to solve for the discrete solutions. Numerical homogenization is then discussed to give some insights of the motivation of our proposed approaches. We also briefly explain reduced-order modeling via GMsFEM, which lay the framework for the methods developed in this dissertation.

Section III is focused on solving nonlinear PDEs. An exemplary problem, p-Laplacian with heterogeneous coefficients, is fully discussed. We propose a multiscale model reduction framework within GMsFEM, where we construct snapshot spaces and design nonlinear spectral problems in the spectral spaces to select dominated modes. One of the innovative parts is that we focus on re-casting the multiscale model reduction problem onto the boundaries of coarse cells via using the concept of p-harmonic extension. A detailed convergence analysis is presented and numerical results are shown to verify the theoretical

findings.

Parabolic equations with multiple scales in space and time are studied in Section IV. We develop a new approach within GMsFEM to solve this kind of problems by constructing space-time coarse cells. The local snapshot basis and local spectral basis are also defined in space-time coarse cells. We argue that this space-time GMsFEM approach can reduce the dimensions of the offline space substantially compared with the previous multiscale methods where only spatial multiscale basis functions are used. The convergence rate of the numerical solution is also analyzed. Besides, we consider adding a new enrichment procedure of computing online multiscale basis functions to the offline basis functions. The use of online basis functions gives a rapid convergence. Some numerical results are presented to exhibit the performance.

We continue exploring the space-time multiscale techniques by using discontinuous Galerkin methods in Section V. Compared with the conforming approaches in Section IV, we construct the local problems in each space-time coarse block instead of coarse neighborhood. The proposed method is within the framework of GMsDGM, and we construct local snapshot space and offline space in each space-time coarse block. A convergence analysis is presented and the numerical experiments verify the results in the analysis.

In Section VI, we conclude the dissertation by summarizing the contributions we have made as well as showing an outlook for future work.

II. PRELIMINARIES

In this section, we introduce some preliminary background concepts that would be helpful for the understanding of the approaches presented in this dissertation. We will explain the coarse-scale and fine-scale grids, numerical homogenization and reduced-order modeling via GMsFEM.

II.1 Coarse-scale and fine-scale grids

Multiscale approaches such as numerical homogenization [4, 63, 24], multiscale finite element methods [24, 63, 5], variational multiscale methods [43, 42, 7, 44] and generalized multiscale finite element methods [26, 28] require the construction of coarse-scale and fine-scale grids. For example, in GMsFEM, the local model reduction is performed in the level of coarse-scale grids. In this section, we will show in details how to define the two-scale level grids. We consider both spatial and time scales.

Let Ω be a bounded domain in $\mathbb{R}^2$ with a Lipschitz boundary $\partial\Omega$, and $[0, T]$ $(T > 0)$ be a time interval. We introduce $\mathcal{T}^h$ to be a partition of the domain Ω into fine finite elements where $h > 0$ is the fine mesh size. Then we form a coarse partition $\mathcal{T}^H$ of the domain Ω such that every element in $\mathcal{T}^H$ is a union of connected fine-mesh grid blocks, that is, $\forall K_j \in \mathcal{T}^H$, $K_j = \cup_{F\in I_j} F$ for some $I_j \subset \mathcal{T}^h$. We call $\mathcal{T}^H$ the coarse grid. $H > 0$ denote the coarse mesh size. In Figure II.1, we illustrate a multiscale discretization.

Let $\{y_i\}_{i=1}^{N_c}$ be the set of nodes in the coarse grid $\mathcal{T}^H$, where N_c is the number of coarse nodes. We denote the neighborhood of the node y_i by

$$\omega_i = \bigcup \{K_j \in \mathcal{T}^H : y_i \in \overline{K_j}\}.$$

Notice that ω_i is the union of all coarse elements $K_j \in \mathcal{T}^H$ sharing the coarse node y_i. See

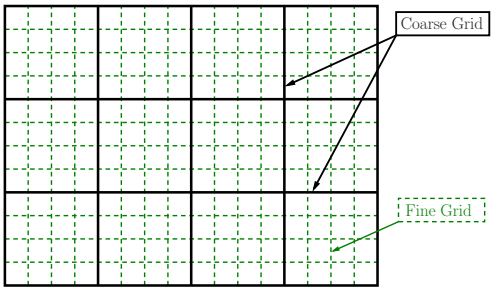

Figure II.1: Illustration of a multiscale discretization.

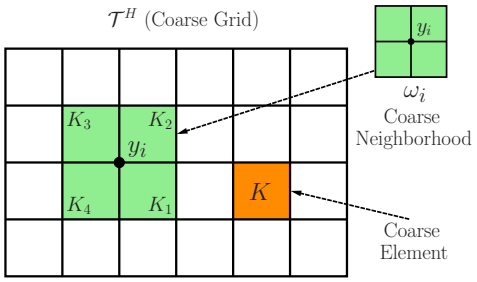

Figure II.2: Illustration of a coarse neighborhood and elements.

Figure II.2 for an illustration of a coarse neighborhood and elements. Inside each coarse neighborhood ω_i ($i = 1, \cdots, N_v$), we call the collection of the coarse edges with y_i being a common vertex the *cross* of y_i.

Next, let $\mathcal{T}^T = \{(T_{n-1}, T_n) | 1 \leq n \leq N\}$ be a coarse partition of the time interval $(0, T)$ where

$$0 = T_0 < T_1 < T_2 < \cdots < T_N = T$$

and we define a fine partition of $(0, T)$, $\mathcal{T}^t$ by refining the partition $\mathcal{T}^T$.

II.2 Numerical homogenization for nonlinear equations

When solving for numerical solutions of partial differential equations (PDEs) with rapidly oscillating coefficients (multiple scales), numerical homogenization methods are very useful techniques.

For example, we consider

$$-\text{div}(a(x, \nabla u)) = f \quad \text{in } \Omega,$$

with $u = 0$ on $\partial\Omega$. In numerical homogenization methods, one can use as a local problem in each coarse cell K,

$$-\text{div}(a(x, \nabla N_\xi)) = 0 \quad \text{in } K,$$

with the boundary conditions $N_\xi = \xi \cdot x$ on ∂K. The homogenized fluxes are computed by averaging the flux

$$a^*(\xi) = \langle a(x, \nabla \phi_\xi) \rangle = \frac{1}{|K|} \int_K a(x, \nabla N_\xi) dx.$$

The coarse-grid equation is given by

$$-\text{div}(a^*(x, \nabla u^*)) = f \quad \text{in } \Omega,$$

with $u^* = 0$ on $\partial\Omega$. These approaches follow homogenization theory ([39, 13, 3, 48], see also [2, 50] and, the references therein, for numerical homogenization), which is well developed.

II.3 Numerical homogenization for space-time problems

Numerical homogenization methods for space-time problems with separable spatial and temporal scales have been studied in the literature [46, 55, 57, 36]. In these methods, one solves local problems To give an example, we consider the following parabolic equation

$$\frac{\partial}{\partial t}u - \text{div}(\kappa(x, x/\epsilon^{\alpha}, t, t/\epsilon^{\beta})\nabla u) = f, \tag{II.1}$$

subject to smooth initial and boundary conditions. Here, ϵ is a small scale, and the spatial scale is ϵ^{α}, and the temporal scale is ϵ^{β}. One can show that (e.g., [46, 55]), the homogenized equation has the same form as (II.1), but with the smooth coefficients $\kappa^{*}(x, t)$. Due to the scale separation, the coefficients can be computed by using the solutions of local parabolic equations in the periodic cells. The local problems may or may not include time-dependent derivatives depending on the interplay between α and β since the cell problems are independent of ϵ. This homogenization procedure can be extended to numerical homogenization type methods [51, 1, 35, 37, 59, 62], where the local parabolic equations need to be solved in each coarse block and in each coarse time step. To compute the effective property, one averages the solutions of the local problems. These approaches work well in the scale separation cases, but do not provide accurate approximations when there is no scale separation. One of our goals in the dissertation is to design numerical methods for handling problems with non-separable space and time scales.

II.4 Reduced-order modeling via GMsFEM

In many multiscale finite element methods, they construct one basis function per coarse node. However, for many complex scale disparities, such as the case where several high-conductive regions are contained in the coarse region, one need multiple basis functions

per coarse node to capture the local information of the solution space. In this section, we will discuss how one can compute these multiscale basis functions and introduce Reduced-order modeling via GMsFEM.

To see the framework of GMsFEM clearly, we take the elliptic equation

$$L(u) = -\nabla \cdot (\kappa(x)\nabla u) = f \text{ in } \Omega \tag{II.2}$$

subject to Dirichlet boundary condition $u = g$ on $\partial\Omega$ as an example, where $\kappa(x)$ is a highly heterogeneous permeability field in Ω.

To propose the GMsFEM framework in the conforming setting, we first need to construct a set of partition of unity functions $\{\chi_i\}_{i=1}^{N_v}$. These functions are supported in coarse neighborhoods, and summed to one. Specifically, the support of χ_i is ω_i, and $\sum_{i=1}^{N_v} \chi_i = 1$. In addition, χ_i has value 1 at the vertex y_i. There are two commonly used sets of partition of unity functions, which are presented below.

- A bilinear partition of unity functions: χ_i is defined as bilinear functions χ_i^0 on ω_i, which equals 1 at node y_i and 0 on $\partial\omega_i$.

- A multiscale partition of unity functions: χ_i is defined by

$$L(u) = 0 \ \text{ in } K \in \omega_i, \quad \chi_i = \chi_i^0 \ \text{ on } \partial K, \ \text{ for all } K \in \omega_i.$$

In the following, we elaborate the steps of GMsFEM framework. And we note that the work in this dissertation basically follows this framework.

II.4.1 Step 1: Construct local snapshot space.

We consider one coarse neighborhood ω_i. The construction of the multiscale basis functions on ω_i starts with a snapshot space $V_{\text{snap}}^{\omega_i}$. There are two common choices of $V_{\text{snap}}^{\omega_i}$

in practice.

Using all possible fine-grid functions in ω_i is the first choice. This can provide accurate approximation for the solution space, while having large dimensions. The second choice consists of the use of harmonic extensions. That is, for each ω_i, we obtain snapshot basis functions $\psi_j^{\omega_i}$ by solving

$$\begin{aligned} L(\psi_j^{\omega_i}) &= 0 \quad \text{on } \omega_i \ (\omega_i^+), \\ \psi_j^{\omega_i}(x_k) &= \delta_{kj} \text{ for all fine grid node, } x_i, \text{ on } \partial\omega_i. \end{aligned}$$

Next, one can perform an appropriate spectral decomposition (cf. Proper Orthogonal Decomposition (POD)) to obtain some linear independent basis in ω_i. The local snapshot space $V_{\text{snap}}^{\omega_i}$ is defined by

$$V_{\text{snap}}^{\omega_i} = \text{span}\{\psi_j^{\omega_i} |\ 1 \leq j \leq p^{\omega_i}\}.$$

II.4.2 Step 2: Construct offline space.

The snapshot space $V_{\text{snap}}^{\omega_i}$ constructed in Step 1 contains all or most necessary components of the fine-scale solution restricted to ω_i. For the purpose of mode reduction, one propose a spectral problem in the local snapshot space and extract the dominant modes in the snapshot space. We use these dominant modes to obtain the offline basis functions and the offline space.

For each ω_i, one solves the following spectral problem: find $\{\phi_j^{\omega_i}\} \subset V_{\text{snap}}(\omega_i)$ such that

$$a_{\omega_i}(\phi_j^{\omega_i}, v) = \lambda_j^{\omega_i} s_{\omega_i}(\phi_j^{\omega_i}, v), \quad \text{for all } v \in V_{\text{snap}}^{\omega_i}.$$

For example, $a_{\omega_i}(u, v) := \int_{\omega_i} \kappa \nabla u \cdot \nabla v$, and $s_{\omega_i}(u, v) := \int_{\omega_i} \tilde{\kappa} uv$.

The multiscale offline space V_{off} is defined by

$$V_{\text{off}} = \text{span}\{\chi_i \phi_j^{\omega_i} \mid \forall \omega_i, 1 \leq j \leq p^{\omega_i}\}$$

where $\{\chi_i\}$ are partition of unity functions of Ω.

II.4.3 Step 4: Solve the global problem.

One can now solve the discrete weak form of (II.2) in the offline space for the multiscale solution. That is, find $u_H \in V_{\text{off}}$ such that

$$\int_\Omega \kappa \nabla u_H \cdot \nabla v = \int_\Omega f v, \text{ for all } v \in V_{\text{off}}.$$

III. GMSFEM FOR NONLINEAR PROBLEMS

In this section, we present a multiscale model reduction framework within Generalized Multiscale Finite Element Method (GMsFEM) for nonlinear elliptic problems [19]. We consider an exemplary problem, which consists of nonlinear p-Laplacian with heterogeneous coefficients. Our main objective is to develop snapshots and local spectral problems, which are the main ingredients of GMsFEM. In the process, we focus on re-casting the multiscale model reduction problem onto the boundaries of coarse cells. The sections are organized as follows. In Section III.1, we introduce the model problem and present a motivation. The description of our GMsFEM approach for nonlinear problem is presented in Section III.2. The convergence analysis of the method is given in Section III.3. We construct a numerical implementation in Section III.4, and present numerical results in Section III.5.

III.1 Model problem

Suppose Ω is a bounded open set in $\mathbb{R}^2$ with Lipschitz boundary $\partial\Omega$. We consider the following heterogeneous p-Laplacian equation

$$-\text{div}(a(x,\nabla u)) = f(x) \quad \text{in } \Omega, \qquad u = g \quad \text{on } \partial\Omega, \tag{III.1}$$

where $a(x,\nabla u) = \kappa(x)|\nabla u|^{p-2}\nabla u$, $p \geq 2$, $\kappa(x) \geq \kappa_0 > 0$ is a high-contrast coefficient (i.e., $\kappa_{\max}/\kappa_{\min}$ is large), $f \in W^{-1,q}(\Omega)$ $(1/p + 1/q = 1)$ is an external forcing term, and $g \in W^{1/q,p}(\Omega)$ is the Dirichlet boundary data.

The corresponding weak formulation is: ($\mathcal{P}$) Find $u \in W_g^{1,p}(\Omega) \equiv \{v \in W^{1,p}(\Omega) :$

$v = g$ on $\partial\Omega\}$ such that

$$\int_\Omega a(x, \nabla u) \cdot \nabla v = \int_\Omega fv, \qquad \forall v \in W_0^{1,p}(\Omega).$$

The well-posedness of $(\mathcal{P})$ is well established, and one can refer to, for example, Glowinski and Marrocco [38] or the account in Ciarlet [22]. Throughout the section, we define the energy norm of $u \in W^{1,p}(\Omega)$ as

$$\|u\|_{1,p(\Omega)} = \left(\int_\Omega \kappa(x) |\nabla u|^p dx \right)^{1/p}.$$

Next, we introduce the finite element approach for the problem. Let $\mathcal{T}^h$ be a fine triangulation. We denote by $V^h = V^h(\Omega)$ the usual finite element space containing continuous piecewise linear functions with respect to $\mathcal{T}^h$. We also define $V_0^h(\Omega)$ as the subset of $V^h(\Omega)$ containing functions that vanish on $\partial\Omega$. Similar notations, $V^h(K), V_0^h(K)$, are used for $K \subset \Omega$.

The discrete fine-scale problem is defined in the following: $(\mathcal{P}^h)$ Find $u^h \in V^h(\Omega)$ such that

$$\int_\Omega a(x, \nabla u^h) \cdot \nabla v = \int_\Omega fv, \qquad \forall v \in V_0^h(\Omega).$$

Additionally, we introduce the coarse-scale and fins-scale spatial grids and relevant concepts and notations such as coarse neighborhood in the domain Ω, see Section II.1.

III.1.1 Motivation

First, we introduce the concept of *p-harmonic extension*.

Definition III.1.1. *Let $u \in W^{1,p}(K)$ ($p \geq 2$) be a given function. Let $\tilde{u} \in W^{1,p}(K)$ be*

defined so that $\tilde{u} - u \in W_0^{1,p}(K)$*, and that* $\tilde{u}$ *satisfies:*

$$-div(a(x, \nabla \tilde{u})) = 0 \quad in\ K,$$

where $a(x, \nabla \tilde{u}) = \kappa(x)|\nabla \tilde{u}|^{p-2}\nabla \tilde{u}$*. Then* $\tilde{u}$ *is called the* p-harmonic extension *of* u*. We denote* $\tilde{u} := H_p(u)$*.*

Remark III.1.2. *The* p*-harmonic extension minimizes the energy norm, i.e.*

$$\int_K \kappa(x)|\nabla \tilde{u}|^p dx = \min_{v \in W_u^{1,p}(K)} \int_K \kappa(x)|\nabla v|^p dx,$$

where $W_u^{1,p}(K) = \{v \in W^{1,p}(K) \,|\, v = u \text{ on } \partial K\}$*.*

Remark III.1.3. *In this context, all* p*-harmonic extensions are accomplished coarse-element by coarse-element. Though, we might use the notation* H_p *directly on a larger domain such as a coarse neighborhoods* ω_i *or the whole domain* Ω*, it means that the* p*-harmonic extension is performed on each coarse element contained in* ω_i *or* Ω*.*

Our main idea is solving for the GMsFEM solution of Equation (III.1) on the *crosses* of the coarse mesh and then the solution in the whole domain can be approximated by p-harmonically extending the obtained cross values into the domain. This idea is motivated by the technique of numerical homogenization. Our goal is to show that our proposed GMsFEM approach recovers numerical homogenization.

First, we describe a well known numerical homogenization technique. This method can be regarded as using a limited number of degrees of freedom per coarse element. Our objective is to show that the numerical homogenization is a finite element approximation on a coarse grid using p-harmonic extension with only one degree of freedom per edge. We consider

$$-\text{div}(a(x, \nabla u)) = f \quad \text{in } \Omega,$$

with $u = 0$ on $\partial\Omega$. We consider a coarse-grid block K and our goal for each coarse-grid block is to compute the effective property. This is done by solving local problem

$$-\text{div}(a(x, \nabla N_\xi)) = 0 \quad \text{in } K,$$

with boundary condition $N_\xi = \xi \cdot x$ on ∂K. According to the previous definition, we can write $N_\xi = H_p(\xi \cdot x)$. Then $a^*(\cdot)$ is defined as

$$a^*(\xi) = \frac{1}{|K|} \int_K a(y, \nabla N_\xi) dy.$$

The coarse-grid equation is given by

$$-\text{div}(a^*(x, \nabla u^*)) = f \quad \text{in } \Omega,$$

with $u^* = 0$ on $\partial\Omega$. Suppose $u^* = \sum c_k \phi_k$, where $\{\phi_k\}$ is a linear basis, then

$$F^{NH}(\vec{c}) = \int_\Omega a^*(x, \nabla \sum c_k \phi_k) \cdot \nabla \phi_j dx = \sum_{K \in \Omega} \int_K a^*(\sum c_k \nabla \phi_k) \cdot \nabla \phi_j dx,$$

At this step, we denote $\sum c_k \nabla \phi_k = \xi = constant$, then $N_\xi = H_p(\sum c_k \nabla \phi_k \cdot x)$ and

$$\begin{aligned}
F^{NH}(\vec{c}) &= \sum_{K \in \Omega} \int_K a^*(\xi) \cdot \nabla \phi_j dx \\
&= \sum_{K \in \Omega} \int_K \left(\frac{1}{|K|} \int_K a(x, \nabla N_\xi) dx \right) \cdot \nabla \phi_j dx \\
&= \int_\Omega \frac{1}{|K|} \left(\int_K a(x, \nabla H_p(\sum c_k \nabla \phi_k \cdot x)) dx \right) \cdot \nabla \phi_j dx \\
&= \int_\Omega \frac{1}{|K|} \left(\int_K a(x, \nabla H_p(\sum c_k \phi_k)) dx \right) \cdot \nabla \phi_j dx \\
&= \int_\Omega f \phi_j dx.
\end{aligned}$$

Compared with the numerical homogenization, GMsFEM seeks the approximation of the solution in the form $\sum_{i,k} \chi_i c_k^{\omega_i} \phi_k^{\omega_i}$, which solves

$$\begin{aligned} F(\vec{c}) &= \int_\Omega a(x, \nabla H_p(\sum_{i,k} \chi_i c_k^{\omega_i} \phi_k^{\omega_i}) \cdot \nabla \phi_j^{\omega_i} dx \\ &= \int_\Omega f \phi_j^{\omega_i} dx, \end{aligned}$$

where $\{\phi_k^{\omega_i}\}_{k=1}^{L_i}$ (L_i is the number of basis chosen in ω_i) are generalized multiscale basis constructed in each ω_i ($i = 1, \cdots, N_v$), $\{\chi_i\}_{i=1}^{N_v}$ is the set of partition of unity functions. Our main approach is to construct multiscale basis functions in a systematic way and provide a priori error. We see from the above discussion that GMsFEM can be thought as an extension of numerical homogenization, where we need to identify appropriate procedures for finding multiscale basis functions. In the following section, we will describe the details of constructing multiscale basis as well as partition of unity functions.

III.2 Generalized multiscale finite element basis

The goal of our proposed GMsFEM is to find a numerical approximation of the solution as well as employing the degree of freedoms only on the crosses in order to exhibit model reduction. Suppose the generalized multiscale finite element solution we are seeking for is $u_{ms} = H_p(\sum_i \sum_{k=1}^{L_i} \chi_i c_k^{\omega_i} \phi_k^{\omega_i})$, where $\{\phi_k^{\omega_i}\}_{k=1}^{L_i}$ are multiscale basis constructed in each coarse neighborhood ω_i, $\{\chi_i\}_{i=1}^{N_v}$ is the set of partition of unity functions, then the generalized multiscale finite element formulation for Equation (III.1) is the following: Find $\vec{c} = \{c_k^{\omega_i}\}_{i,k}$ such that

$$\int_\Omega a(x, \nabla H_p(\sum_i \sum_{k=1}^{L_i} \chi_i c_k^{\omega_i} \phi_k^{\omega_i}) \cdot \nabla \phi_j^{\omega_i} dx = \int_\Omega f \phi_j^{\omega_i} dx \qquad \text{for any } j, \tag{III.2}$$

where $a(x, \nabla u) = \kappa(x) |\nabla u|^{p-2} \nabla u$.

Remark III.2.1. *The GMsFEM numerical solution u_{ms} to (III.2) is uniquely defined. We suppose that $u_1 = H_p(\sum_i \sum_{k=1}^{L_i} \chi_i c_{k,1}^{\omega_i} \phi_k^{\omega_i})$ and $u_2 = H_p(\sum_i \sum_{k=1}^{L_i} \chi_i c_{k,2}^{\omega_i} \phi_k^{\omega_i})$ are two solutions to (III.2). Then, we have*

$$\int_\Omega (a(x, \nabla u_1) - a(x, \nabla u_2)) \cdot \nabla v dx = 0,$$

for any test function. Since u_1 and u_2 are harmonic in each coarse block, we have

$$\int_\Omega \kappa(x) |\nabla u_1 - \nabla u_2|^p dx \leq \int_\Omega (a(x, \nabla u_1) - a(x, \nabla u_2)) \cdot \nabla (u_1 - u_2) dx = 0,$$

which guarantees $u_1 = u_2$. So the solution to (III.2) is uniquely defined.

III.2.1 Snapshot space

We consider one coarse neighborhood ω_i. The construction of the multiscale basis functions on ω_i starts with a snapshot space $V_{\text{snap}}^{\omega_i}$. There are two common choices of $V_{\text{snap}}^{\omega_i}$ in practice.

Using all possible fine-grid functions in ω_i is the first choice. This can provide accurate approximation for the solution space, while having large dimensions. The second choice consists of the use of harmonic extensions. In particular, we denote by $M_h(\omega_i)$ the set of all nodes of the fine mesh $\mathcal{T}^h$ which lie on $\partial\omega_i$. For each fine-grid node $x_j \in M_h(\omega_i)$, we construct a discrete delta function $\delta_j^h(x)$ defined on $M_h(\omega_i)$ by

$$\delta_j^h(x_k) = \begin{cases} 1 & \text{for } k = j \\ 0 & \text{for } k \neq j \end{cases}, \qquad \forall x_k \in M_h(\omega_i).$$

Then the $j-$th snapshot basis function $\psi_j^{\omega_i}$ is defined as the solution of

$$\begin{aligned} -\text{div}(\kappa(x)\nabla\psi_j^{\omega_i}) &= 0 \quad \text{in } \omega_i, \\ \psi_j^{\omega_i} &= \delta_j^h \text{ on } \partial\omega_i. \end{aligned} \tag{III.3}$$

The dimension of $V_{\text{snap}}^{\omega_i}$ is equal to the size of $M_h(\omega_i)$. We remark that oversampling and randomization can be used to reduce the computational cost associated with the snapshot calculations. We refer to [12] for more details.

III.2.2 Offline space

The construction of generalized multiscale basis for solving p-Laplacian equation in the fashion of p-harmonic extension is based on the design of a proper nonlinear spectral problem which will be solved in the snapshot space. In each coarse neighborhood ω_i, we define the following nonlinear eigenvalue problem:

$$\begin{cases} \phi_1^{\omega_i} = c^{\omega_i}, & \lambda_1^{\omega_i} = 0, \\ \phi_k^{\omega_i} = \arg\min\limits_{v\in V_{\text{snap}}^{\omega_i}} \dfrac{G^{\omega_i}(v)}{G_\chi^{\omega_i}(v - P_{k-1}(v))}, & \lambda_k^{\omega_i} = \dfrac{G^{\omega_i}(\phi_k^{\omega_i})}{G_\chi^{\omega_i}(\phi_k^{\omega_i} - P_{k-1}(\phi_k^{\omega_i}))}, \text{ for } k \geq 2, \end{cases} \tag{III.4}$$

where $c^{\omega_i} \in V_{\text{snap}}^{\omega_i}$ is a constant function in ω_i, the functionals are given by

$$G^{\omega_i}(v) = \int_{\omega_i} \kappa(x)|\nabla H_p(v)|^p dx, G_\chi^{\omega_i}(v) = \int_{\omega_i} \kappa(x)|\nabla H_p(\chi_i v)|^p dx,$$

the projector $P_k(u) = \arg\min\limits_{v\in V_{k-1}^{\omega_i}} G^{\omega_i}(u-v)$, $V_{k-1}^{\omega_i} = \text{span}\{\phi_1^{\omega_i}, \cdots, \phi_{k-1}^{\omega_i}\}$. This nonlinear eigenvalue problem is a standard orthogonal subspace minimization method and is well-defined (see e.g., [64]).

The eigenfunctions $\{\phi_k^{\omega_i}\}_k$ in each coarse neighborhood ω_i will contribute as offline basis (or we call them generalized multiscale basis or eigenbasis) after being multiplied

by the associated partition of unity function χ_i. We choose the first L_i eigenfunctions on each ω_i and denote the offline space as

$$V^c = \text{span}\{\chi_i \phi_k^{\omega_i} : k = 1, \cdots, L_i;\ i = 1, \cdots, N_v\} \subseteq W^{1,p}(\Omega).$$

Recall that our solution assumes the form of $u_{ms} = H_p(\sum_i \sum_{k=1}^{L_i} c_k^{\omega_i} \chi_i \phi_k^{\omega_i})$, which means u_{ms} is obtained by p-harmonically extending $\sum_i \sum_{k=1}^{L_i} c_k^{\omega_i} \chi_i \phi_k^{\omega_i}$ in each coarse block K, thus only the values of $\sum_i \sum_{k=1}^{L_i} c_k^{\omega_i} \chi_i \phi_k^{\omega_i}$ on each coarse edge matter in this sense. If we consider one coarse neighborhood ω_i, for example the coarse neighborhood of an interior coarse vertex (see y_i and ω_i in Figure II.2), it is the restriction of $\sum_{k=1}^{L_i} c_k^{\omega_i} \chi_i \phi_k^{\omega_i}$ on the 12 coarse edges that will matter in the process of p-harmonic extension. Notice that the partition of unity function χ_i vanishes on and beyond the boundary of ω_i, thus merely the restriction of $\sum_{k=1}^{L_i} c_k^{\omega_i} \chi_i \phi_k^{\omega_i}$ on the *cross* (that is, the inside 4 coarse edges) makes an influence. Therefore, we can restrict $\chi_i \phi_k^{\omega_i}$ $(k = 1, \cdots, L_i)$ on the cross of ω_i and denote the restricted basis (which we call *cross basis* in this context) by $\hat{\phi}_k^{\omega_i}$. Then we can write $u_{ms} = H_p(\sum_i \sum_{k=1}^{L_i} c_k^{\omega_i} \hat{\phi}_k^{\omega_i})$. We denote

$$\widehat{V}^c = \text{span}\{\hat{\phi}_k^{\omega_i} : k = 1, \cdots, L_i;\ i = 1, \cdots, N_v\}.$$

In this way, we can focus on the degree of freedoms on the crosses and perform spectral decomposition on these crosses.

Remark III.2.2. *In the computation, we use a simpler eigenvalue problem*

$$\begin{cases} \phi_1^{\omega_i} = c^{\omega_i}, & \lambda_1^{\omega_i} = 0, \\ \phi_k^{\omega_i} = \textit{arg} \min\limits_{v \in X_k^{\omega_i}} \dfrac{G^{\omega_i}(v)}{G_\chi^{\omega_i}(v)}, & \lambda_k^{\omega_i} = \dfrac{G^{\omega_i}(\phi_k^{\omega_i})}{G_\chi^{\omega_i}(\phi_k^{\omega_i})}, \quad \textit{for} \quad k \geq 2, \end{cases} \tag{III.5}$$

where $X_k^{\omega_i}$ is a subspace of $V_{snap}^{\omega_i}$ and defined as $X_k^{\omega_i} = \left(span\{\phi_1^{\omega_i}, \cdots, \phi_{k-1}^{\omega_i}\}\right)^{\perp}$ where the orthogonality $\perp$ is defined with respect to the H^1 norm in $V_{snap}^{\omega_i}$.

III.3 Convergence analysis

To analyze the convergence of our proposed method, we first prove several lemmas. The first two lemmas are the direct applications of Lemmas 5.1 and 5.3 in Glowinski and Marrocco [38] and prove the monotonicity and continuity of p-Laplacian operator $a(x, \nabla u) = \kappa(x)|\nabla u|^{p-2}\nabla u$, respectively. In the following proof, we introduce the notation $F \preceq G$ to represent $F \leq \mathcal{C}G$ with a constant $\mathcal{C}$ independent of the mesh, contrast and the functions involved.

Lemma III.3.1. *$\forall u, v \in W^{1,p}(K)$, $p \geq 2$, the following inequality holds:*

$$\kappa(x)|\nabla u - \nabla v|^p \preceq (a(x, \nabla u) - a(x, \nabla v)) \cdot \nabla(u - v). \tag{III.6}$$

Proof. We use Lemma 5.1 in Glowinski and Marrocco [38], which proves the following inequality: $\forall p \geq 2$, $\exists \alpha > 0$ such that $\forall y, z \in \mathbb{R}^2$,

$$(|z|^{p-2}z - |y|^{p-2}y, z - y)_{\mathbb{R}^2} \geq \alpha|z - y|^p.$$

If we take $z = (\kappa(x))^{1/p}\nabla u, y = (\kappa(x))^{1/p}\nabla v$, then (III.6) is proved. □

Lemma III.3.2. *$\forall u, v \in W^{1,p}(K)$, $p \geq 2$, the following inequality holds:*

$$|a(x, \nabla u) - a(x, \nabla v))| \preceq M(x, u, v)|\nabla u - \nabla v|, \tag{III.7}$$

where $M(x, u, v) = \kappa(x)(|\nabla u| + |\nabla v|)^{p-2}$.

Proof. According to Lemma 5.3 in Glowinski and Marrocco [38], there holds the follow-

ing inequality: $\forall p \geq 2$, $\exists \beta > 0$ such that $\forall y, z \in \mathbb{R}^2$,

$$||z|^{p-2}z - |y|^{p-2}y| \leq \beta |z-y|(|z|+|y|)^{p-2}.$$

If we take $z = (\kappa(x))^{1/(p-1)}\nabla u, y = (\kappa(x))^{1/(p-1)}\nabla v$, then (III.7) is proved. □

Lemma III.3.3. *Suppose* $\tilde{u} = H_p(u_0), \tilde{v} = H_p(v_0), \tilde{w} = H_p(u_0 - v_0)$, *where* $u_0, v_0 \in W^{1,p}(K)$, $p \geq 2$. *Then we have*

$$\begin{aligned} &\int_K \kappa(x)|\nabla(\tilde{u}-\tilde{v})|^p dx \\ &\preceq \left(\int_K \kappa(x)|\nabla\tilde{w}|^p dx\right)^{\frac{q}{p}} \left(\int_K \kappa(x)|\nabla\tilde{u}|^p dx + \int_K \kappa(x)|\nabla\tilde{v}|^p dx\right)^{\frac{p-2}{p-1}}, \end{aligned} \tag{III.8}$$

where $1/p + 1/q = 1$.

Proof. Using Lemma III.3.1 and integrating by parts, we immediately obtain the following inequality:

$$\begin{aligned} \int_K \kappa(x)|\nabla(\tilde{u}-\tilde{v})|^p dx &\preceq \int_K (a(x,\nabla\tilde{u}) - a(x,\nabla\tilde{v}))\cdot\nabla(\tilde{u}-\tilde{v}) \\ &= \int_{\partial K} (a(x,\nabla\tilde{u}) - a(x,\nabla\tilde{v}))\cdot\vec{n}(\tilde{u}-\tilde{v})ds \\ &\quad - \int_K \nabla\cdot(a(x,\nabla\tilde{u}) - a(x,\nabla\tilde{v}))\,(\tilde{u}-\tilde{v})dx \\ &= \int_{\partial K} (a(x,\nabla\tilde{u}) - a(x,\nabla\tilde{v}))\cdot\vec{n}\,\tilde{w}ds \\ &\quad - \int_K \nabla\cdot(a(x,\nabla\tilde{u}) - a(x,\nabla\tilde{v}))\,\tilde{w}dx \\ &= \int_{\partial K} (a(x,\nabla\tilde{u}) - a(x,\nabla\tilde{v}))\cdot\nabla\tilde{w}dx \\ &\preceq \int_K M(x,\tilde{u},\tilde{v})|\nabla\tilde{u}-\nabla\tilde{v}||\nabla\tilde{w}|dx, \end{aligned} \tag{III.9}$$

where we used the continuity property of $a(x,\nabla u)$ proved in Lemma III.3.2 on the last

line. Applying Hölder's inequality to (III.9), we have

$$\int_K \kappa(x)|\nabla(\tilde{u}-\tilde{v})|^p dx \preceq \left(\int_K \kappa(x)|\nabla\tilde{u}-\nabla\tilde{v}|^p dx\right)^{\frac{1}{p}} \left(\int_K \kappa(x)|\nabla\tilde{w}|^p dx\right)^{\frac{1}{p}} \left(\int_K \left(\kappa(x)^{-\frac{2}{p}}\,|M(x,\tilde{u},\tilde{v})|\,\right)^{\frac{p}{p-2}} dx\right)^{\frac{p-2}{p}}.$$

Dividing both sides by $\left(\int_K \kappa(x)|\nabla\tilde{u}-\nabla\tilde{v}|^p dx\right)^{\frac{1}{p}}$ gives

$$\int_K \kappa(x)|\nabla(\tilde{u}-\tilde{v})|^p dx \preceq \left(\int_K \kappa|\nabla\tilde{w}|^p dx\right)^{\frac{q}{p}} \left(\int_K \left(\kappa^{-\frac{2}{p}}\,|M(x,\tilde{u},\tilde{v})|\,\right)^{\frac{p}{p-2}} dx\right)^{\frac{p-2}{p-1}}. \tag{III.10}$$

Recall from Lemma III.3.2 that $M(x,\tilde{u},\tilde{v}) = \kappa(x)(|\nabla\tilde{u}|+|\nabla\tilde{v}|)^{p-2}$, thus

$$\begin{aligned}\left(\int_K \left(\kappa(x)^{-\frac{2}{p}}\,|M(x,\tilde{u},\tilde{v})|\,\right)^{\frac{p}{p-2}} dx\right)^{\frac{p-2}{p-1}} &= \left(\int_K \kappa(x)(|\nabla\tilde{u}|+|\nabla\tilde{v}|)^p dx\right)^{\frac{p-2}{p-1}} \\ &\preceq \left(\int_K \kappa|\nabla\tilde{u}|^p dx + \int_K \kappa|\nabla\tilde{v}|^p dx\right)^{\frac{p-2}{p-1}}.\end{aligned} \tag{III.11}$$

We substitute (III.11) into (III.10), and we then see that (III.8) is proved.

□

Lemma III.3.4. *For any $u_0, v_0 \in W^{1,p}(K)$, $p \geq 2$, we have*

$$\int_K \kappa(x)|\nabla H_p(u_0+v_0)|^p dx \leq \int_K \kappa(x)|\nabla(H_p(u_0)+H_p(v_0))|^p dx. \tag{III.12}$$

Proof. Recall Remark III.1.2, we have

$$\int_K \kappa(x)|\nabla H_p(u)|^p dx = \min_{v\in W_u^{1,p}(K)} \int_K \kappa(x)|\nabla v|^p dx.$$

Therefore,

$$\begin{aligned}\int_K \kappa(x)|\nabla H_p(u_0+v_0)|^p dx &= \min_{w\in W^{1,p}_{u_0+v_0}(K)} \int_K \kappa(x)|\nabla w|^p dx \\ &= \min_{w_1+w_2\in W^{1,p}_{u_0+v_0}(K)} \int_K \kappa(x)|\nabla(w_1+w_2)|^p dx \\ &\leq \min_{w_1\in W^{1,p}_{u_0}(K),\, w_2\in W^{1,p}_{v_0}(K)} \int_K \kappa(x)|\nabla(w_1+w_2)|^p dx.\end{aligned}$$

Taking $w_1 = H_p(u_0), w_2 = H_p(v_0)$, we obtain

$$\int_K \kappa(x)|\nabla H_p(u_0+v_0)|^p dx \leq \int_K \kappa(x)|\nabla(H_p(u_0)+H_p(v_0))|^p dx.$$

□

We will next prove Lemma III.3.5 which approximates the error of GMsFEM solution by using functions from the offline space V^c.

Lemma III.3.5. *Suppose u is the exact solution of Equation (III.1), u_{ms} is the GMsFEM solution from Equation (III.2), then for any $p \geq 2$, we have*

$$\|u-u_{ms}\|_{1,p(\Omega)} \preceq \|u-H_p(v_H)\|^{\frac{q}{p}}_{1,p(\Omega)} \|u\|^{\frac{p-2}{p-1}}_{1,p(\Omega)} \quad \textit{for any } v_H \in V^c,$$

where $1/p+1/q=1$, $\|u\|_{1,p(\Omega)} = \left(\int_\Omega \kappa|\nabla u|^p dx\right)^{1/p}$ is the energy norm.

Proof. Using Lemma III.3.1, we immediately obtain that for any $v_H \in V^c$,

$$\begin{aligned}\int_\Omega \kappa|\nabla(u-u_{ms})|^p dx &\preceq \int_\Omega (a(x,\nabla u)-a(x,\nabla u_{ms}))\cdot\nabla(u-u_{ms}) \\ &= \int_\Omega (a(x,\nabla u)-a(x,\nabla u_{ms}))\cdot\nabla(u-H_p(v_H)) \\ &\preceq \int_\Omega M(x,u,u_{ms})|\nabla u-\nabla u_{ms}||\nabla u-\nabla H_p(v_H)|dx.\end{aligned}$$

Applying Hölder's inequality, we have

$$\int_\Omega \kappa|\nabla(u-u_{ms})|^p dx \preceq \left(\int_\Omega \kappa|\nabla(u-u_{ms})|^p dx\right)^{\frac{1}{p}} \left(\int_\Omega \kappa|\nabla(u-H_p(v_H))|^p dx\right)^{\frac{1}{p}} \left(\int_\Omega \left(\kappa^{-\frac{2}{p}}\,|M(x,u,u_{ms})|\right)^{\frac{p}{p-2}} dx\right)^{\frac{p-2}{p}}.$$

Dividing both sides by $\left(\int_\Omega \kappa(x)|\nabla(u-u_{ms})|^p dx\right)^{\frac{1}{p}}$ gives

$$\begin{aligned}\int_\Omega \kappa|\nabla(u-u_{ms})|^p dx &\preceq \left(\int_\Omega \kappa|\nabla(u-H_p(v_H))|^p dx\right)^{\frac{q}{p}} \left(\int_\Omega \left(\kappa^{-\frac{2}{p}}\,|M|\right)^{\frac{p}{p-2}} dx\right)^{\frac{p-2}{p-1}} \\ &= \left(\int_\Omega \kappa|\nabla(u-H_p(v_H))|^p dx\right)^{\frac{q}{p}} \left(\int_\Omega \kappa(|\nabla u|+|\nabla u_{ms}|)^p dx\right)^{\frac{p-2}{p-1}} \\ &\preceq \left(\int_\Omega \kappa|\nabla(u-H_p(v_H))|^p dx\right)^{\frac{q}{p}} \left(\int_\Omega \kappa(x)|\nabla u|^p dx\right)^{\frac{p-2}{p-1}}.\end{aligned}$$

It follows immediately

$$\|u-u_{ms}\|_{1,p(\Omega)} \preceq \|u-H_p(v_H)\|^{\frac{q}{p}}_{1,p(\Omega)} \|u\|^{\frac{p-2}{p-1}}_{1,p(\Omega)}.$$

□

Lemma III.3.6. *Suppose u is the exact solution of Equation (III.1), K is any coarse block of size H, $p \geq 2$, then we have*

$$\int_K \kappa(x)|\nabla(u-H_p(u))|^p dx \preceq H^q \int_K |f|^q dx, \tag{III.13}$$

where $1/p + 1/q = 1$.

Proof. It's clear that

$$-\text{div}(a(x,\nabla u) - a(x,\nabla H_p(u))) = f.$$

Thus,

$$\int_K (a(x,\nabla u) - a(x,\nabla H_p(u))) \cdot \nabla(u - H_p(u))dx = \int_K (u - H_p(u))f dx.$$

By Lemma III.3.1,

$$\begin{aligned}\int_K \kappa(x)|\nabla(u - H_p(u))|^p dx &\preceq \int_K |u - H_p(u)||f|dx \\ &\preceq \left(\int_K |u - H_p(u)|^p dx\right)^{\frac{1}{p}} \left(\int_K |f|^q dx\right)^{\frac{1}{q}}.\end{aligned}$$

Using Poincaré's inequality, we get

$$\begin{aligned}\int_K \kappa(x)|\nabla(u - H_p(u))|^p dx &\preceq H \left(\int_K |\nabla(u - H_p(u))|^p dx\right)^{\frac{1}{p}} \left(\int_K |f|^q dx\right)^{\frac{1}{q}} \\ &= \kappa_0^{-\frac{1}{p}} H \left(\int_K \kappa_0 |\nabla(u - H_p(u))|^p dx\right)^{\frac{1}{p}} \left(\int_K |f|^q dx\right)^{\frac{1}{q}} \\ &\preceq H \left(\int_K \kappa(x)|\nabla(u - H_p(u))|^p dx\right)^{\frac{1}{p}} \left(\int_K |f|^q dx\right)^{\frac{1}{q}}.\end{aligned}$$

Dividing both sides by $\left(\int_K \kappa(x)|\nabla(u - H_p(u))|^p dx\right)^{\frac{1}{p}}$, we get

$$\int_K \kappa(x)|\nabla(u - H_p(u))|^p dx \preceq H^q \int_K |f|^q dx.$$

□

Remark III.3.7. *This local error estimate proved in Lemma III.3.6 immediately deduces*

the global error estimate:

$$\|u - H_p(u)\|_{1,p(\Omega)}^p \preceq H^q \|f\|_{L^q(\Omega)}^q. \tag{III.14}$$

Now, we come to the main convergence theorem.

Theorem III.3.8. *Suppose u is the exact solution of Equation (III.1), u_{ms} is the GMsFEM solution from Equation (III.2), then for any $p \geq 2$, we have*

$$\|u - u_{ms}\|_{1,p(\Omega)} \preceq \|u\|_{1,p(\Omega)}^{\frac{p-2}{p-1}} \left\{ H^{\frac{1}{(p-1)^2}} \|f\|_{L^q(\Omega)}^{\frac{1}{(p-1)^2}} + \left(\frac{1}{\Lambda_*}\right)^{\frac{1}{p(p-1)^2}} \|u\|_{1,p(\Omega)}^{\frac{1}{p-1}} \right\}, \tag{III.15}$$

where $\Lambda_* = \min_{\omega_i} \lambda_{L_i+1}^{\omega_i}$, $\{\lambda_j^{\omega_i}\}$ are the eigenvalues defined by (III.4) in Section III.2.2, L_i is the number of eigenbasis chosen in each coarse neighborhood ω_i.

Proof. We first define the interpolation of u onto the offline space V^c as

$$I_0 u = \arg\min_{v \in V^c} \{\|u - H_p(v)\|_{1,p(\Omega)}\}.$$

Since $I_0 u \in V^c$, we denote $I_0 u = \sum_i \chi_i u_0^{\omega_i}$, where $u_0^{\omega_i} = \sum_{k=1}^{L_i} c_k^{\omega_i} \phi_k^{\omega_i}$. By Lemma III.3.5 and (III.14), it follows that

$$\begin{aligned}
\|u - u_{ms}\|_{1,p(\Omega)} &\preceq \|u\|_{1,p(\Omega)}^{\frac{p-2}{p-1}} \|u - H_p(I_0 u)\|_{1,p(\Omega)}^{\frac{q}{p}} \\
&\preceq \|u\|_{1,p(\Omega)}^{\frac{p-2}{p-1}} \left(\|u - H_p(u)\|_{1,p(\Omega)}^{\frac{q}{p}} + \|H_p(u) - H_p(I_0 u)\|_{1,p(\Omega)}^{\frac{q}{p}} \right) \\
&\preceq \|u\|_{1,p(\Omega)}^{\frac{p-2}{p-1}} \left(H^{\frac{1}{(p-1)^2}} \|f\|_{L^q(\Omega)}^{\frac{1}{(p-1)^2}} + \|H_p(u) - H_p(I_0 u)\|_{1,p(\Omega)}^{\frac{1}{p-1}} \right).
\end{aligned} \tag{III.16}$$

In the following, we will estimate $\|H_p(u) - H_p(I_0 u)\|_{1,p(\Omega)}^{\frac{1}{p-1}}$. Using Lemma III.3.3, we have

$$
\begin{aligned}
&\|H_p(u) - H_p(I_0 u)\|^p_{1,p(\Omega)} \\
&\preceq \left(\int_\Omega \kappa(x)|\nabla H_p(u - I_0 u)|^p dx\right)^{\frac{q}{p}} \left(\int_\Omega \kappa(x)|\nabla H_p(u)|^p dx + \int_\Omega \kappa(x)|\nabla H_p(I_0 u)|^p dx\right)^{\frac{p-2}{p-1}} \\
&\preceq \left(\int_\Omega \kappa(x)|\nabla H_p(\sum_{\omega_i} \chi_i(u - u_0^{\omega_i}))|^p dx\right)^{\frac{q}{p}} \left(\int_\Omega \kappa(x)|\nabla u|^p dx\right)^{\frac{p-2}{p-1}}. \qquad \text{(III.17)}
\end{aligned}
$$

Applying the property of $H_p(\cdot)$ proved in Lemma III.3.4 to (III.17), we achieve

$$
\begin{aligned}
&\|H_p(u) - H_p(I_0 u)\|^p_{1,p(\Omega)} \\
&\preceq \left(\int_\Omega \sum_{\omega_i} \kappa(x)|\nabla H_p(\chi_i(u - u_0^{\omega_i}))|^p dx\right)^{\frac{q}{p}} \left(\int_\Omega \kappa(x)|\nabla u|^p dx\right)^{\frac{p-2}{p-1}} \\
&\preceq \left(\sum_{\omega_i} \int_{\omega_i} \kappa(x)|\nabla H_p(\chi_i(u - u_0^{\omega_i}))|^p dx\right)^{\frac{q}{p}} \left(\int_\Omega \kappa(x)|\nabla u|^p dx\right)^{\frac{p-2}{p-1}}. \qquad \text{(III.18)}
\end{aligned}
$$

Recall that $u_0^{\omega_i} = \sum_{k=1}^{L_i} c_k^{\omega_i} \phi_k^{\omega_i}$ with $\{\phi_k^{\omega_i}\}$ being eigenfunctions defined by (III.4) in Section III.2.2. We have the following inequality

$$
\int_{\omega_i} \kappa(x)|\nabla H_p(\chi_i(u - u_0^{\omega_i}))|^p dx \preceq \frac{1}{\lambda_{L_i+1}} \int_{\omega_i} \kappa(x)|\nabla H_p(u - u_0^{\omega_i})|^p dx. \qquad \text{(III.19)}
$$

Define $\Lambda_* = \min_{\omega_i} \lambda_{L_i+1}^{\omega_i}$, then through (III.18) and (III.19) we get

$$
\begin{aligned}
&\|H_p(u) - H_p(I_0 u)\|^p_{1,p(\Omega)} \\
&\preceq \left(\sum_{\omega_i} \frac{1}{\lambda_{L_i+1}} \int_{\omega_i} \kappa|\nabla H_p(u - u_0^{\omega_i})|^p dx\right)^{\frac{q}{p}} \left(\int_\Omega \kappa|\nabla u|^p dx\right)^{\frac{p-2}{p-1}} \\
&\preceq \left(\frac{1}{\Lambda_*}\right)^{\frac{q}{p}} \left(\int_\Omega \kappa|\nabla H_p(u)|^p dx\right)^{\frac{q}{p}} \left(\int_\Omega \kappa|\nabla u|^p dx\right)^{\frac{p-2}{p-1}}. \qquad \text{(III.20)}
\end{aligned}
$$

Using the energy minimization property of $H_p(\cdot)$ claimed in Remark III.1.2, we obtain

from (III.20) that

$$\begin{aligned}
&\|H_p(u) - H_p(I_0u)\|^p_{1,p(\Omega)} \\
&\preceq \left(\frac{1}{\Lambda_*}\right)^{\frac{q}{p}} \left(\int_\Omega \kappa(x)|\nabla u|^p dx\right)^{\frac{q}{p}} \left(\int_\Omega \kappa(x)|\nabla u|^p dx\right)^{\frac{p-2}{p-1}} \\
&= \left(\frac{1}{\Lambda_*}\right)^{\frac{q}{p}} \left(\int_\Omega \kappa(x)|\nabla u|^p dx\right)^{\frac{q}{p}+\frac{p-2}{p-1}} \\
&= \left(\frac{1}{\Lambda_*}\right)^{\frac{1}{p-1}} \int_\Omega \kappa(x)|\nabla u|^p dx.
\end{aligned}$$

This gives

$$\|H_p(u) - H_p(I_0u)\|^{\frac{1}{p-1}}_{1,p(\Omega)} \preceq \left(\frac{1}{\Lambda_*}\right)^{\frac{1}{p(p-1)^2}} \|u\|^{\frac{1}{p-1}}_{1,p(\Omega)}. \tag{III.21}$$

Substituting (III.21) into (III.16), we obtain

$$\|u - u_{ms}\|_{1,p(\Omega)} \preceq \|u\|^{\frac{p-2}{p-1}}_{1,p(\Omega)} \left\{ H^{\frac{1}{(p-1)^2}} \|f\|^{\frac{1}{(p-1)^2}}_{L^q(\Omega)} + \left(\frac{1}{\Lambda_*}\right)^{\frac{1}{p(p-1)^2}} \|u\|^{\frac{1}{p-1}}_{1,p(\Omega)} \right\}.$$

□

Remark III.3.9. *We notice that Λ^* will increase to infinity. Considering a function u with highly oscillating boundary conditions, the value of $\dfrac{G^{\omega_i}(u)}{G^{\omega_i}_\chi(u)}$ will be very large. More specifically, for u having highly oscillating boundary conditions, the value of $G^{\omega_i}(u)$ is large. But $\chi_i u$ will have less oscillation on the cross since u solves the harmonic problem. Therefore, $G^{\omega_i}_\chi(u)$ will be small and the ratio of $G^{\omega_i}(u)$ over $G^{\omega_i}_\chi(u)$ will be large.*

Besides, we can improve the offline convergence rate by using multiple oversampled spectral problems. To be simple, we start with two eigenvalue problems. We denote $I^{\omega_i}_0(u) = u^{\omega_i}_0$, $\widetilde{\chi_i} = \frac{\chi_i}{\chi_i^+}$, $\tilde{u} = \chi_i^+(u - I^{\omega_i}_0(u))$, where χ_i^+ is the partition of unity func-

tion on the oversampled domain ω_i^+. Following inequality (III.18),

$$\begin{aligned}
\sum_{\omega_i}\int_{\omega_i}\kappa(x)|\nabla H_p(\chi_i(u-I_0^{\omega_i}(u)))|^p dx &= \sum_{\omega_i}\int_{\omega_i}\kappa(x)|\nabla H_p(\widetilde{\chi}_i(\tilde{u}-I_0^{\omega_i}(\tilde{u})))|^p dx \\
&\preceq \sum_{\omega_i}\frac{1}{\lambda_{L_i+1}^{\omega_i}}\int_{\omega_i}\kappa(x)|\nabla H_p(\tilde{u}-I_0^{\omega_i}(\tilde{u}))|^p dx \\
&\preceq \frac{1}{\Lambda_*}\sum_{\omega_i^+}\frac{1}{\lambda_{L_i+1}^{\omega_i^+}}\int_{\omega_i^+}\kappa|\nabla H_p(u-I_0^{\omega_i}(u))|^p dx \\
&\preceq \frac{1}{\Lambda_*}\frac{1}{\Lambda_*^+}\sum_{\omega_i^+}\int_{\omega_i^+}\kappa(x)|\nabla H_p(u-I_0^{\omega_i}(u))|^p dx \\
&\preceq \frac{1}{\Lambda_*}\frac{1}{\Lambda_*^+}\int_{\Omega}\kappa(x)|\nabla u|^p dx,
\end{aligned}$$

where $\Lambda_* = \min_{\omega_i}\lambda_{L_i+1}^{\omega_i}$ and $\Lambda_*^+ = \min_{\omega_i^+}\lambda_{L_i+1}^{\omega_i^+}$. This result can be easily extended to multiple oversampled eigenvalue problems (instead of two eigenvalue problems), and the result would be

$$\sum_{\omega_i}\int_{\omega_i}\kappa(x)|\nabla H_p(\chi_i(u-I_0^{\omega_i}(u)))|^p dx \preceq \frac{1}{\Lambda_*}\frac{1}{\Lambda_*^+}\cdots\frac{1}{\Lambda_*^{+N}}\int_{\Omega}\kappa(x)|\nabla u|^p dx,$$

where $\Lambda_*^{+N} = \min_{\omega_i^+}\lambda_{L_i+1}^{\omega_i^{+N}}$, ω_i^{+N} is a N layers oversampled domain (ω_i^+ is a 1 layer oversampled domain). We note that if we choose all these $\frac{1}{\Lambda_*}$ and $\frac{1}{\Lambda_*^{+k}}$'s to be less than some δ ($0<\delta<1$), then $\frac{1}{\Lambda_*}\frac{1}{\Lambda_*^+}\cdots\frac{1}{\Lambda_*^{+N}} < \delta^{N+1}$ and the offline error would be exponential decay as the number of oversampled layers increases.

III.4 Numerical implementation

In this part, we exhibit the process of numerically implementing the proposed method for p-Laplacian equation. From Glowinski and Marrocco [38], or Ciarlet [22], ($\mathcal{P}$) is equivalent to the following minimization problem: ($\mathcal{Q}$) Find $u \in W_g^{1,p}(\Omega) \equiv \{v \in$

$W^{1,p}(\Omega) : v = g \text{ on } \partial\Omega\}$ such that

$$J_\Omega(u) = \inf_{v \in W_g^{1,p}(\Omega)} J_\Omega(v), \tag{III.22}$$

where $J_\Omega(u) = \frac{1}{p}\int_\Omega \kappa(x)|\nabla u|^p dx - \int_\Omega f u dx$.

It is easily established that $J_\Omega(u)$ is strictly convex and continuous on $W_g^{1,p}(\Omega)$. Besides, $J_\Omega(u)$ is Gateaux differentiable with

$$J_\Omega^{'}(u)(w) = \int_\Omega \kappa(x)|\nabla u|^{p-2}\nabla u \cdot \nabla w dx - \int_\Omega f w dx \quad \forall w \in W_0^{1,p}(\Omega).$$

Hence, there exists a unique solution to ($\mathcal{Q}$), and ($\mathcal{Q}$) is equivalent to its Euler equation ($\mathcal{P}$). The corresponding discrete problem of ($\mathcal{Q}$) is: ($\mathcal{Q}^h$) Find $u^h \in V^h(\Omega)$ such that

$$J_\Omega(u^h) = \min_{v^h \in V_0^h(\Omega)} J_\Omega(v^h). \tag{III.23}$$

The well-posedness of ($\mathcal{Q}^h$) = ($\mathcal{P}^h$) follows in an analogous way to that of ($\mathcal{Q}$) and ($\mathcal{P}$), see Glowinski and Marrocco [38] or Ciarlet [22].

Recall the discussion in Section III.2.2, we can represent the GMsFEM solution by $u^h = H_p(\sum_i \sum_{k=1}^{L_i} c_k^{\omega_i} \hat{\phi}_k^{\omega_i})$. For simplicity, we use a single-index notation and denote $u^h = H_p(\sum_{j=1}^{N} c_j \hat{\phi}_j)$. Then we apply *Broyden's method* (which is a Quasi-Newton's method) to solve the minimization problem ($\mathcal{Q}^h$), see Algorithm 1.

III.5 Numerical results

In this section, we present a number of representative numerical results to verify the proposed methods in the previous sections. In particular, we solve Equation (III.1) using the proposed method. We also solve (III.1) for the fine-grid solution on the unit square $\Omega = [0,1] \times [0,1]$ using a uniform fine grid of 100×100 square finite elements which

Algorithm 1 A Quasi-Newton algorithm

Initialization: An initial guess $\vec{c}^{(0)} = \left(c_j^{(0)}\right)_{j=1}^N$ and $B^{(0)} \in \mathbb{R}^{N\times N}$

(1) Compute the gradient vector $\vec{g}^{(0)}(\vec{c}^{(0)}) = \nabla J_\Omega(H_p(\sum_{j=1}^N c_j^{(0)}\hat{\phi}_j))$.

(2) Compute the stepsize $\tau^{(0)}$.

(3) Set: $\vec{c}^{(1)} = \vec{c}^{(0)} - \tau^{(0)} B^{(0)}\vec{g}^{(0)}$.

(4) **If** $\|\vec{c}^{(1)} - \vec{c}^{(0)}\| < \delta$, where $\|\cdot\|$ is a suitable norm, **return**.

for $k = 1$ to N: **do**

(1) Compute the gradient vector $\vec{g}^{(k)}(\vec{c}^{(k)}) = \nabla J_\Omega(H_p(\sum_{j=1}^N c_j^{(k)}\hat{\phi}_j))$.

(2) Compute the approximation of the inverse of Hessian matrix:

$$B^{(k)} = B^{(k-1)} + \frac{[(\vec{c}^{(k)} - \vec{c}^{(k-1)}) - B^{(k-1)}(\vec{g}^{(k)} - \vec{g}^{(k-1)})](\vec{c}^{(k)} - \vec{c}^{(k-1)})^T B^{(k-1)}}{(\vec{c}^{(k)} - \vec{c}^{(k-1)})^T B^{(k-1)}(\vec{g}^{(k)} - \vec{g}^{(k-1)})}.$$

(3) Compute the stepsize $\tau^{(k)}$.

(4) Set: $\vec{c}^{(k+1)} = \vec{c}^{(k)} - \tau^{(k)} B^{(k)}\vec{g}^{(k)}$.

(5) **If** $\|\vec{c}^{(k+1)} - \vec{c}^{(k)}\| < \delta$, **return**.

end for

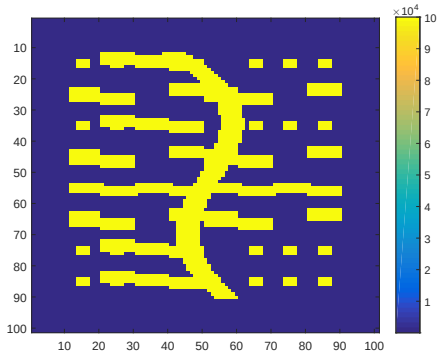

Figure III.1: Illustration of the high-contrast permeability field $\kappa_1(x)$.

is divided into 10×10 square coarse elements uniformly. The forcing term is chosen as $f = 1$ and we impose a linear Dirichlet boundary condition $u(x_1, x_2) = x_1 + x_2$. The high-contrast permeability field $\kappa_1(x)$ used in our experiments is shown in Figure III.1, with high-contrast ratio $\kappa_{\max}/\kappa_{\min}$ being 10^5.

III.5.1 Accuracy of GMsFEM using different numbers of basis functions

We use both fine-grid (FEM) and coarse-grid (GMsFEM) methods to solve the model problem (III.1). To compare the respective approaches, we introduce relative L_p errors and relative energy errors, which are defined as

$$\begin{aligned} L_p \text{ error} &= \frac{\|u - u_{ms}\|_{L^p(\Omega)}}{\|u\|_{L^p(\Omega)}} \times 100\%, \\ \text{Energy error} &= \frac{\|u - u_{ms}\|_{1,p(\Omega)}}{\|u\|_{1,p(\Omega)}} \times 100\%, \end{aligned} \tag{III.24}$$

where we recall that u denotes the FEM solution and u_{ms} denotes the GMsFEM solution.

For the first set of experiments, we take $p = 3, 4, 5, 6$ separately and use different numbers of cross basis (L_i for each ω_i) for each fixed value of p. Then we check the relative errors of the GMsFEM solutions. Numerical results are shown in Table III.1 and Figure III.2. Note that in the first column of each sub-table, we show the numbers of basis functions used in each coarse neighborhood ω_i, and the numbers in parentheses are the degrees of freedom (DOF) of offline space. To visually observe the accuracy of GMsFEM, we plot the solutions obtained by both FEM and GMsFEM in the case $p = 3$ using 4 cross basis functions in each coarse neighborhood, see Figure III.3.

By observing the columns in Table III.1 (or the curves in Figure III.2), we can clearly see that for each p, the relative error decays as we use more cross basis functions. We note that as L_i increases, the the value of $(L_i + 1)$'s eigenvalue increases, and the error bound $1/\Lambda_*$ will correspondingly decrease. Through a more careful examination, we notice that for each p, when 4 or more than 4 cross basis are chosen in each coarse neighborhood (i.e. $L_i \geq 4$ for each ω_i), the errors are much smaller. This might suggest that if we use 4 or more than 4 cross basis in each coarse neighborhood, we would get a better convergence. We will explore this in more details in the following subsection.

L_i (DOF)	$p = 3$	
	L_p error	Energy error
1(81)	9.52 %	41.03 %
2(162)	6.45 %	34.38 %
3(243)	5.76 %	27.76 %
4(324)	0.52 %	6.55 %
5(405)	0.45 %	5.15 %

L_i (DOF)	$p = 4$	
	L_p error	Energy error
1(81)	10.88 %	42.35 %
2(162)	6.47 %	32.93 %
3(243)	5.12 %	24.13 %
4(324)	0.92 %	8.57 %
5(405)	0.82 %	6.65 %

L_i (DOF)	$p = 5$	
	L_p error	Energy error
1(81)	10.12 %	40.46 %
2(162)	7.71 %	34.05 %
3(243)	5.17 %	27.88 %
4(324)	0.94 %	9.94 %
5(405)	0.81 %	7.92 %

L_i (DOF)	$p = 6$	
	L_p error	Energy error
1(81)	8.95 %	39.68 %
2(162)	6.94 %	30.92 %
3(243)	4.37 %	23.85 %
4(324)	1.07 %	8.70 %
5(405)	0.91 %	7.08 %

Table III.1: Relative errors for $p = 3, 4, 5, 6$ using different numbers of cross basis.

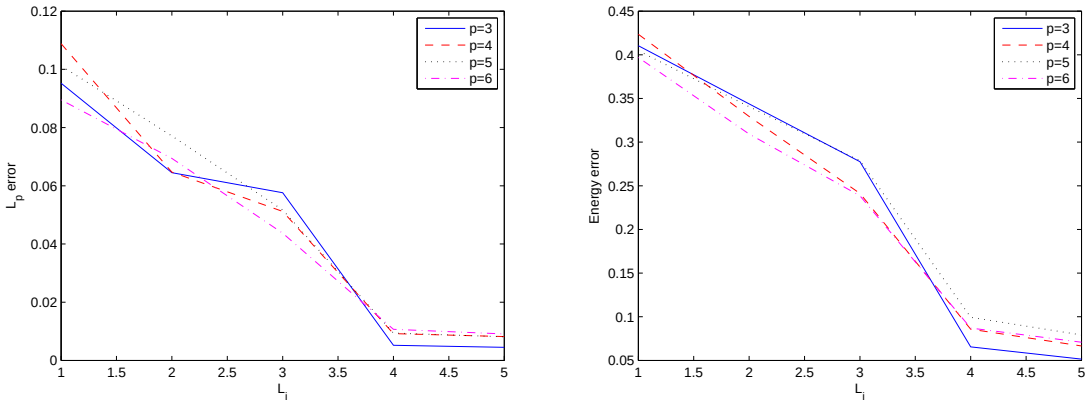

Figure III.2: Relative error vs L_i for $p = 3, 4, 5, 6$.

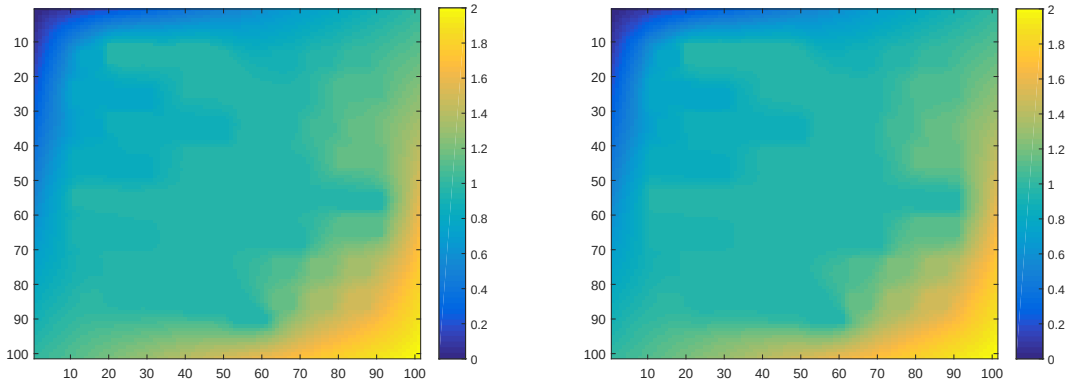

Figure III.3: FEM v.s. GMsFEM node-wise solutions, $p = 3$, DOF $= 324$.

III.5.2 Correlation between errors and eigenvalues

Aside from the accuracy of our proposed method, we are interested in determining how many cross basis (or DOFs) should be used. As we mentioned earlier, there is a "jump" in the relative energy errors when we take 4 cross basis in each coarse neighborhood (i.e. $L_i = 4$ for each ω_i, see Table III.1). Thus, $L_i = 4$ might be a good choice. According to our analysis in Section III.3, that is probably due to a sudden decrease in the quantity of $1/\Lambda_*$, where $\Lambda_* = \min_{\omega_i} \lambda^{\omega_i}_{L_i+1}$, $\{\lambda^{\omega_i}_j\}$ are the eigenvalues defined in (III.5) in Section III.2.2. To verify this theory, we calculate the corresponding $1/\Lambda_*$ for each L_i in the case of $p = 3$. The results are shown in Table III.2. In this table, we see the jump in Λ_* and $1/\Lambda_*$ at $L_i = 4$, which explains our earlier inference. Hence, we conclude that the proper number of cross basis is chosen at the spot where there is a sudden increase in the values of Λ_* (or a sudden decrease in the values of $1/\Lambda_*$). We would like to remark that an adaptive method can be employed to determine a best choice of L_i for each coarse neighborhood ω_i. Moreover, to see a more quantitative relationship between the relative errors and the values of Λ_* as well as being inspired by the result in Theorem III.3.8, we calculate the cross-correlation coefficient between the relative energy errors and the corresponding values of

$(\frac{1}{\Lambda_*})^{\frac{1}{p(p-1)^2}}$ for the case $p = 3$. We recall that the quantity $(\frac{1}{\Lambda_*})^{\frac{1}{p(p-1)^2}}$ comes from (III.15) in Theorem III.3.8. The evaluated cross-correlation coefficient is 0.99. This indicates a linear relationship between the relative energy error and the corresponding $(\frac{1}{\Lambda_*})^{\frac{1}{p(p-1)^2}}$, which verifies our result in (III.15).

L_i	Λ_*	$1/\Lambda_*$
1	8.86e-4	1.13e3
2	2.59e-3	3.86e2
3	4.46e-3	2.24e2
4	1.55e2	6.44e-3
5	4.01e2	2.50e-3

Table III.2: Values of Λ_* and $1/\Lambda_*$ when $p = 3$.

Remark III.5.1. *We note that the choice of L_i is not empirical. We refer to the discussion in Section 3 in [30], which states that the choice of the proper L_i is highly related to the number of "channels" and "inclusions" in each coarse neighborhood. In more details, if there are m inclusions and channels in a coarse neighborhood ω_i , then one can observe m small, asymptotically vanishing, eigenvalues. In the example presented in this section, we can see that there are at most 4 inclusions and channels in each coarse neighborhood. This suggests the choice of $L_i = 4$, and we verify this choice by observing the values of Λ_i^*.*

Since L_i is the number of eigen-pairs solved from the nonlinear eigenvalue problem, it's a finite number and can not grow to infinity. We can only guarantee that as L_i grows (not necessary to be a large number), the error will decay to a significant small level, which is verified by our numerical results and more importantly exhibits the effect of model reduction. This result is similar to solving problems in the linear setting.

L_i	Energy error	$1/\Lambda_*$
1	44.15 %	1.42e3
2	36.44 %	4.04e2
3	27.99 %	2.35e2
4	6.77 %	6.49e-3
5	5.30 %	2.50e-3

Table III.3: Relative energy errors and values of $1/\Lambda_*$ using $\kappa_2(x)$, $p = 3$.

III.5.3 Numerical tests with more permeability fields

To verify that our proposed method is applicable to more situations, we examine other choices of permeability field $\kappa(x)$. First, we would like to check that the GMsFEM solution errors do not depend on the high-contrast ratio $\kappa_{\max}/\kappa_{\min}$. To see this, we increase the high-contrast ratio of $\kappa_1(x)$, which is used in the previous subsections, from 10^5 to 10^7. We denote the new permeability field by $\kappa_2(x)$. Then we solve Equation (III.1) using both FEM and proposed GMsFEM, and calculate the relative errors and the error bound quantity $1/\Lambda_*$. Numerical results for $p = 3$ are shown in Table III.3. Comparing these results with the top left sub-table in Table III.1 and Table III.2, we can observe similar trend inside the columns as well as a slight increase in the values of both relative energy errors and $1/\Lambda_*$'s. The jump at $L_i = 4$ still occurs. The cross-correlation coefficient between the relative energy errors and $(\frac{1}{\Lambda_*})^{\frac{1}{p(p-1)^2}}$ is calculated to be 0.98.

We also consider a different high-contrast permeability field $\kappa_3(x)$, see Figure III.4. We solve Equation (III.1) for $p = 3$ and the results are presented in Table III.4. The cross-correlation coefficient between the relative energy errors and $(\frac{1}{\Lambda_*})^{\frac{1}{p(p-1)^2}}$ is calculated to be 0.94. Similar conclusions as made in Section III.5.1 and III.5.2 can be drawn for this new choice of permeability field. We can see that our proposed method works well for this permeability field.

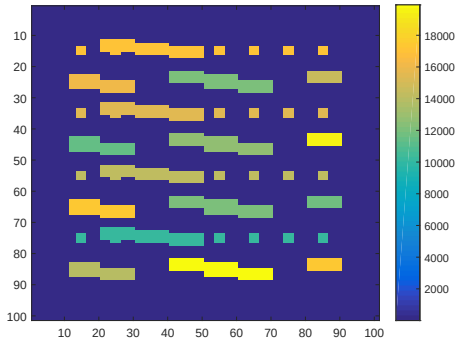

Figure III.4: Illustration of the high-contrast permeability field $\kappa_3(x)$.

L_i	Energy error	$1/\Lambda_*$
1	47.08 %	1.85e1
2	27.68 %	4.64e0
3	20.81 %	2.68e0
4	4.33 %	2.26e-3
5	2.69 %	1.01e-3

Table III.4: Relative energy errors and values of $1/\Lambda_*$ using $\kappa_3(x)$, $p = 3$.

III.5.4 Comments on the computational cost

The online cost is independent of fine mesh parameters, while it will grow as the spectral basis parameters increase. We note that the online cost is proportional to that of solving homogeneous p-Laplacian equation with polynomial basis. In practise, we usually only use a few spectral basis, so the online cost is close to that of solving homogeneous p-Laplacian equation with low order polynomial basis. We note that solving the nonlinear eigenvalue problem in each coarse neighborhood is one source of the computational cost. However, this is an offline step, which means when dealing with different forcing terms and boundary conditions we only need to solve this nonlinear eigenvalue problem for a single time. Thus, the computation of this eigenvalue problem will not affect the online

cost of our method. We also note that in Algorithm 1, we need to iteratively update the values of the nonlinear function $N(\nabla u) = \kappa(x)|\nabla u|^{p-2}$ at the fine-grid update level. This can be time-consuming due to the large size of the fine grid. To decrease this computational cost, one can use the discrete empirical interpolation method (DEIM). Within DEIM, the nonlinear function on the fine grid can be approximated by only evaluating at a few carefully selected points. We refer to [6] [58] for more detailed discussions on the use of DEIM. Moreover, by comparing the degrees of freedoms listed in Table III.1 with the size of the fine-scale finite element system $N_f = 10201$, we see that we obtain a reduced sized system by applying GMsFEM, which will reduce the computational cost.

Remark III.5.2. *Compared with the online cost, the offline cost depends on the fine mesh parameter, considering that each local snapshot problem is solved on the local coarse neighborhood consisting of fine grids. We note that the cost of numerical homogenization is high because the local problem* $-div(a(x, \nabla N_{\xi_i})) = 0 \quad$ *in* K *with boundary condition* $N_{\xi_i} = \xi_i \cdot x$ *on* ∂K *(*K *is a coarse block) is solved for all* $\xi_1, \xi_2, \cdots, \xi_N$*. Similarly, the local problem in offline stage is solved for all possible boundary conditions, which is consistent to numerical homogenization. So the offline cost is high. However, as mentioned in Section 3.2.1, the techniques of oversampling and randomization can be adopted to help reduce the offline computational cost associated with the calculation of snapshots. Also, one can select and compute eigenbasis adaptively in each coarse neighborhood which can eliminate the use of non-dominated modes and reduce offline cost. Besides, as mentioned above, DEIM is introduced to reduce offline computational cost when it comes to the evaluation of nonlinear functions.*

Remark III.5.3. *To illustrate how the error of DEIM affects the global error estimate of Theorem 4.8, we adopt the notation* $a_{DEIM}(x, \nabla u) = N_1(\nabla u)\nabla u$*, where* $N_1(\nabla u)(\approx \kappa(x)|\nabla u|^{p-2})$ *is evaluated using DEIM. We denote* $\tilde{u}_{ms}$ *the DEIM solution which is ob-*

tained by solving

$$\int_\Omega a_{DEIM}(x, \nabla \tilde{u}_{ms}) \cdot \nabla v = \int_\Omega f v, \qquad \forall v \in V_0^h(\Omega).$$

Then we have

$$\begin{aligned}
\|u_{ms} - \tilde{u}_{ms}\|_{1,p(\Omega)}^p &= \int_\Omega a(x, \nabla(u_{ms} - \tilde{u}_{ms})) \cdot \nabla(u_{ms} - \tilde{u}_{ms}) \\
&\leq \int_\Omega a(x, \nabla u_{ms}) \cdot \nabla(u_{ms} - \tilde{u}_{ms}) - \int_\Omega a(x, \nabla \tilde{u}_{ms}) \cdot \nabla(u_{ms} - \tilde{u}_{ms}) \\
&= \int_\Omega (u_{ms} - \tilde{u}_{ms}) f - \int_\Omega a(x, \nabla \tilde{u}_{ms}) \cdot \nabla(u_{ms} - \tilde{u}_{ms}) \\
&= \int_\Omega (a_{DEIM}(x, \nabla \tilde{u}_{ms}) - a(x, \nabla \tilde{u}_{ms})) \cdot \nabla(u_{ms} - \tilde{u}_{ms}).
\end{aligned}$$

We note that $\|a_{DEIM}(x, \nabla \tilde{u}_{ms}) - a(x, \nabla \tilde{u}_{ms})\|$ is the DEIM error, which can be assumed to be a small quantity (ref. [5]). Then from the above inequality, it follows

$$\|u_{ms} - \tilde{u}_{ms}\|_{1,p(\Omega)}^p \leq \delta^{p-1} \|u_{ms} - \tilde{u}_{ms}\|_{1,p(\Omega)}, \tag{III.25}$$

for some small $\delta (0 < \delta < 1)$.

Combining (III.25) with Theorem 4.8, we obtain

$$\begin{aligned}
\|u - \tilde{u}_{ms}\|_{1,p(\Omega)} &\leq \|u - u_{ms}\|_{1,p(\Omega)} + \|u_{ms} - \tilde{u}_{ms}\|_{1,p(\Omega)} \\
&\preceq \delta + \|u\|_{1,p(\Omega)}^{\frac{p-2}{p-1}} \left\{ H^{\frac{1}{(p-1)^2}} \|f\|_{L^q(\Omega)}^{\frac{1}{(p-1)^2}} + \left(\frac{1}{\Lambda_*}\right)^{\frac{1}{p(p-1)^2}} \|u\|_{1,p(\Omega)}^{\frac{1}{p-1}} \right\}.
\end{aligned}$$

IV. SPACE-TIME GMSFEM

In this section, we consider a local multiscale model reduction for problems with multiple scales in space and time [18]. Our main objective is to develop a multiscale model reduction framework within GMsFEM that uses space-time coarse cells. We organize the sections as following. In Section IV.1, we present the underlying problem, the motivation of space-time approach, and the space-time GMsFEM formulations. In Section IV.2, we construct the multiscale basis by introducing the snapshot and spectral problems. In Section IV.3, we present the convergence analysis for our proposed method. In Section IV.5, we present the new enrichment procedure of computing online multiscale basis functions. We present numerical results for offline GMsFEM and online GMsFEM in Section IV.4 and Section IV.6, separately.

IV.1 Model problem

Suppose Ω is a bounded domain in $\mathbb{R}^2$ with a Lipschitz boundary $\partial\Omega$, and $[0,T]$ $(T>0)$ be a time interval. In this section, we consider the following parabolic differential equation

$$\begin{aligned} \frac{\partial}{\partial t}u - \text{div}(\kappa(x,t)\nabla u) &= f \qquad && \text{in } \Omega\times(0,T), \\ u &= 0 && \text{on } \partial\Omega\times(0,T), \\ u(x,0) &= \beta(x) && \text{in } \Omega, \end{aligned} \tag{IV.1}$$

where $\kappa(x,t)$ is a time dependent heterogeneous media (for example, a time dependent high-contrast permeability field), f is a source term, $\beta(x)$ is the initial condition. Our main objective is to develop space-time multiscale model reduction within GMsFEM and we use the time-dependent parabolic equation as an example. The proposed methods can

be used for other models that require space-time multiscale model reduction.

We will introduce the space-time GMsFEM in this section, where time dependent multiscale basis functions are constructed on the coarse space-time cell. We recall that the mesh used in this section is as introduced in Section II.1.

We first compute the fine-scale solution by using the standard conforming piecewise linear finite element method. One can use discontinuous Galerkin coupling also [27, 20, 21]. Specifically, we define the finite element space V_h with respect to $\mathcal{T}^h \times (0,T)$ as

$$\begin{aligned} V_h &= \{v \in L^2((0,T); C^0(\Omega)) \mid v = \phi(x)\psi(t), \text{ where} \\ &\quad \phi|_K \in Q_1(K)\ \forall K \in \mathcal{T}^h,\ \psi|_\tau \in C^0(\tau)\ \forall \tau \in \mathcal{T}^T \\ &\quad \text{and } \psi|_\tau \in P_1(\tau)\ \forall \tau \in \mathcal{T}^t\}, \end{aligned}$$

then the fine-scale solution $u_h \in V_h$ is obtained by solving the following variational problem

$$\begin{aligned} &\int_0^T \int_\Omega \frac{\partial u_h}{\partial t} v + \int_0^T \int_\Omega \kappa \nabla u_h \cdot \nabla v + \sum_{n=0}^{N-1} \int_\Omega [u_h(x,T_n)] v(x,T_n^+) \\ &= \int_0^T \int_\Omega f v + \int_\Omega \beta(x) v(x,T_0^+), \end{aligned} \tag{IV.2}$$

$\forall v \in V_h$, where $[\cdot]$ is the jump operator such that

$$\begin{cases} [u_h(x,T_n)] = u_h(x,T_n^+) - u_h(x,T_n^-) & \text{for } n \geq 1, \\ [u_h(x,T_n)] = u_h(x,T_0^+) & \text{for } n = 0. \end{cases}$$

Now we use the space-time finite element method to solve problem (IV.1) on the coarse

grid. That is, we find $u_H \in V_H$ such that

$$\begin{aligned} &\int_0^T \int_\Omega \frac{\partial u_H}{\partial t} v + \int_0^T \int_\Omega \kappa \nabla u_H \cdot \nabla v + \sum_{n=0}^{N-1} \int_\Omega [u_H(x, T_n)] v(x, T_n^+) \\ =& \int_0^T \int_\Omega f v + \int_\Omega \beta(x) v(x, T_0^+), \end{aligned} \tag{IV.3}$$

$\forall v \in V_H$, where V_H is the multiscale finite element space which will be introduced in the following subsections.

The computational cost for solving the equation (IV.3) is huge since we need to compute the solution u_H in the whole time interval $(0, T)$ at one time. To fix this issue, we assume the solution space V_H can be decomposed into a direct sum of subspaces and each subspace only contains functions defined on one single coarse time interval (T_{n-1}, T_n), that is,

$$V_H = \oplus_{n=1}^N V_H^{(n)},$$

where $V_H^{(n)}$ is defined as

$$V_H^{(n)} = \{v(x,t) | v(\cdot, t) = 0 \text{ for } t \in (0,T) \backslash (T_{n-1}, T_n)\}.$$

Then the coarse problem (IV.3) can be decomposed into a sequence of sub-problems: find $u_H^{(n)} \in V_H^{(n)}$ such that

$$\begin{aligned} &\int_{T_{n-1}}^{T_n} \int_\Omega \frac{\partial u_H^{(n)}}{\partial t} v + \int_{T_{n-1}}^{T_n} \int_\Omega \kappa \nabla u_H^{(n)} \cdot \nabla v + \int_\Omega u_H^{(n)}(x, T_{n-1}^+) v(x, T_{n-1}^+) \\ =& \int_{T_{n-1}}^{T_n} \int_\Omega f v + \int_\Omega g_H^{(n)}(x) v(x, T_{n-1}^+), \ \forall v \in V_H^{(n)}, \end{aligned} \tag{IV.4}$$

where

$$g_H^{(n)}(\cdot) = \begin{cases} u_H^{(n-1)}(\cdot, T_{n-1}^-) & \text{for } n \geq 1, \\ \beta(\cdot) & \text{for } n = 0. \end{cases}$$

Then, the solution u_H of the problem (IV.3) is the direct sum of all these $u_H^{(n)}$'s, that is $u_H = \oplus_{n=1}^N u_H^{(n)}$.

In the next, we will discuss space-time multiscale basis functions. First, we will construct multiscale basis functions in the offline mode without using the residual. Next, in Section IV.5, we will discuss online space-time multiscale basis construction.

IV.2 Construction of offline basis functions

IV.2.1 Snapshot space

Let ω be a given coarse neighborhood in space. To simplify the notations, we omit the coarse node index for now. Within GMSFEM framework, we need first to construct a snapshot space V_{snap}^{ω} (or $V_{\text{snap}}^{\omega(n)}$) on coarse time interval (T_{n-1}, T_n) in order to get offline basis functions. There are two common choices of V_{snap}^{ω} in practice.

Using all possible fine-grid functions in $\omega \times (T_{n-1}, T_n)$ is the first choice. This can provide accurate approximation for the solution space, while having very large dimensions. The second choice involves solving local problems using all possible boundary conditions. In particular, we define ψ_j as the solution of the following local problem

$$\begin{aligned} &\frac{\partial}{\partial t}\psi_j - \text{div}(\kappa(x,t)\nabla\psi_j) = 0 \ \text{ in } \omega \times (T_{n-1}, T_n), \\ &\psi_j(x,t) = \delta_j(x,t) \ \text{ on } \ \partial\left(\omega \times (T_{n-1}, T_n)\right). \end{aligned} \tag{IV.5}$$

Here $\delta_j(x,t)$ is a fine-grid delta function and $\partial\left(\omega \times (T_{n-1}, T_n)\right)$ denotes the boundaries $t = T_{n-1}$ and on $\partial\omega \times (T_{n-1}, T_n)$. In general, the computations of these snapshots are expensive because in each ω, one need to solve the number of $O(M_n^{\partial\omega})$ local problems,

where $M_n^{\partial\omega}$ is the number of fine grids on the boundaries $t = T_{n-1}$ and on $\partial\omega \times (T_{n-1}, T_n)$. Thus, in order to build a more efficient multiscale method, one needs a smaller yet accurate snapshot space. We can take an advantage of randomized oversampling concepts [12] and compute only a few snapshot vectors, which will reduce the computational cost remarkably while keeping required accuracy. Next, we introduce randomized snapshots.

Firstly, we describe the notations for oversampled regions. We denote by ω^+ the oversampled space region of $\omega \subset \omega^+$, defined by adding several fine- or coarse-grid layers around ω. Time interval (T^*_{n-1}, T_n) is defined as the left-side oversampled time region for (T_{n-1}, T_n). We will construct inexpensive snapshots by imposing random boundary conditions on the local region $\omega^+ \times (T^*_{n-1}, T_n)$. In particular, one solves the following local problems:

$$\begin{aligned}
&\frac{\partial}{\partial t}\psi_j^+ - \text{div}(\kappa(x,t)\nabla\psi_j^+) = 0 \ \text{ in } \omega^+ \times (T^*_{n-1}, T_n),\\
&\psi_j^+(x,t) = r_l \ \text{ on } \ \partial\left(\omega^+ \times (T^*_{n-1}, T_n)\right),
\end{aligned}$$

where r_l is a random vector. Then the local snapshot space on $\omega^+ \times (T^*_{n-1}, T_n)$ is

$$V^{\omega^+}_{\text{snap}} = \text{span}\{\psi_j^+(x,t) | j = 1, \cdots, L^\omega + p^\omega_{\text{bf}}\},$$

where L^ω and p^ω_{bf} are the number of local offline basis in ω and the buffer number, respectively.

IV.2.2 Offline space

Now, we are going to construct the offline basis functions. This is done by performing a space reduction in each local space-time snapshot space through some well-defined spectral problems. In particular, we propose the following eigenvalue problem on $\omega^+ \times (T_{n-1}, T_n)$:

Find $(\phi, \lambda) \in V_{\text{snap}}^{\omega^+} \times \mathbb{R}$ such that

$$A_n(\phi, v) = \lambda S_n(\phi, v), \quad \forall v \in V_{\text{snap}}^{\omega^+}, \tag{IV.6}$$

where the bilinear operators $A_n(\phi, v)$ and $S_n(\phi, v)$ are defined by

$$\begin{aligned} A_n(\phi, v) =& \frac{1}{2}\left(\int_{\omega^+} \phi(x, T_n)v(x, T_n) + \int_{\omega^+} \phi(x, T_{n-1})v(x, T_{n-1})\right) \\ &+ \int_{T_{n-1}}^{T_n}\int_{\omega^+} \kappa(x, t)\nabla\phi \cdot \nabla v, \\ S_n(\phi, v) =& \int_{\omega_+} \phi(x, T_{n-1})v(x, T_{n-1}) + \int_{T_{n-1}}^{T_n}\int_{\omega^+} \widetilde{\kappa}^+(x, t)\phi v, \end{aligned} \tag{IV.7}$$

where the weighted function $\widetilde{\kappa}^+(x, t)$ is defined by

$$\widetilde{\kappa}^+(x, t) = \kappa(x, t)\sum_{i=1}^{N_c} |\nabla\chi_i^+|^2.$$

We note that $\{\chi_i^+\}_{i=1}^{N_c}$ are partition of unity functions associated with $\{\omega_i^+\}_{i=1}^{N_c}$ and satisfies $|\nabla\chi_i^+| \geq |\nabla\chi_i|$ on ω_i, while $\{\chi_i\}_{i=1}^{N_c}$ are partition of unity functions on $\{\omega_i\}_{i=1}^{N_c}$ and defined by the following local problems:

$$\begin{aligned} -\text{div}(\kappa(x, T_{n-1})\nabla\chi_i) = 0, \quad &\text{in } K \in \omega_i, \\ \chi_i = g_i, \quad &\text{on } \partial K, \end{aligned} \tag{IV.8}$$

for all $K \in \omega_i$, where g_i is a continuous function on ∂K and is linear on each edge of ∂K.

The eigenvalues $\{\lambda_j^{\omega^+} | j = 1, 2, \cdots L^\omega + p_{\text{bf}}^\omega\}$ from (IV.6) are then arranged in the ascending order, and we select the first L^ω eigenfunctions, which are associated with the first L^ω ordered eigenvalues, and denote them by $\{\Psi_1^{\omega^+,\text{off}}, \cdots, \Psi_{L^\omega}^{\omega^+,\text{off}}\}$. Using these eigen-

functions, we can define

$$\psi_j^{\omega^+}(x,t) = \sum_{k=1}^{L^\omega + p_{\text{bf}}^\omega} (\Psi_j^{\omega^+,\text{off}})_k \psi_k^+(x,t), \qquad j = 1, 2, \cdots, L^\omega,$$

where $(\Psi_j^{\omega^+,\text{off}})_k$ denotes the k-th component of $\Psi_j^{\omega^+,\text{off}}$, and $\psi_k^+(x,t)$ is the snapshot basis function computed on $\omega^+ \times (T_{n-1}^*, T_n)$ as in the previous subsection. Then we can obtain the snapshots $\psi_j^\omega(x,t)$ on the target region $\omega \times (T_{n-1}, T_n)$ by restricting $\psi_j^{\omega^+}(x,t)$ onto $\omega \times (T_{n-1}, T_n)$. Finally, the offline basis functions on $\omega \times (T_{n-1}, T_n)$ are defined by $\phi_j^\omega(x,t) = \chi \psi_j^\omega(x,t)$, where χ is the standard multiscale basis function from (IV.8) for a generic coarse neighborhood ω.

The local offline space on $\omega \times (T_{n-1}, T_n)$ can be defined as

$$V_{\text{off}}^\omega = \text{span}\{\phi_j^\omega(x,t) | j = 1, \cdots, L^\omega\}.$$

Note that one can take $V_H^{(n)}$ in (IV.4) as $V_H^{(n)} = V_{\text{off}}^{(n)} = \text{span}\{\phi_j^{\omega_i}(x,t) | 1 \le i \le N_c, 1 \le j \le L_i\}$. As a result, $V_H = V_{\text{off}} = \oplus_{n=1}^N V_H^{(n)}$.

Remark IV.2.1. *For the convenience of convergence analysis in Section IV.3, we also denote by* $\{\Psi_1^{\omega^+,off}, \cdots, \Psi_{L^\omega + p_{bf}^\omega}^{\omega^+,off}\}$ *all the eigenfunctions from (IV.6) corresponding to the ordered eigenvalues, and define*

$$\psi_j^{\omega^+}(x,t) = \sum_{k=1}^{L^\omega + p_{bf}^\omega} (\Psi_j^{\omega^+,off})_k \psi_k^+(x,t),$$

for $j = 1, 2, \cdots, L^\omega + p_{bf}^\omega$.

We note that the snapshot space on $\omega^+ \times (T_{n-1}^*, T_n)$ *can be rewritten as*

$$V_{snap}^{\omega^+} = span\{\psi_j^{\omega^+}(x,t) | j = 1, \cdots, L^\omega + p_{bf}^\omega\},$$

and the snapshot space on $\omega \times (T_{n-1}, T_n)$ can be written as

$$V^{\omega}_{snap} = span\{\psi^{\omega}_j(x,t) | j = 1, \cdots, L^{\omega} + p^{\omega}_{bf}\},$$

where each $\psi^{\omega}_j(x,t)$ is the restriction of $\psi^{\omega^+}_j(x,t)$ onto $\omega \times (T_{n-1}, T_n)$. By collecting all local snapshot spaces on each $\omega \times (T_{n-1}, T_n)$, we can obtain the snapshot space $V^{(n)}_{snap}$ on $\Omega \times (T_{n-1}, T_n)$.

The offline space can be rewritten as

$$V^{\omega}_{off} = span\{\chi\psi^{\omega}_j(x,t) | j \leq L^{\omega}\}.$$

Remark IV.2.2. *One can use a more general spectral problem in (IV.6) with*

$$\begin{aligned} A(\phi, v) =& \frac{1}{2}\left(\int_{\omega} \phi(x, T_n)v(x, T_n) + \int_{\omega} \phi(x, T_{n-1})v(x, T_{n-1})\right) \\ &+ \int_{T_{n-1}}^{T_n}\int_{\omega} \kappa(x,t)\nabla\phi \cdot \nabla v + \int_{T_{n-1}}^{T_n}\int_{\omega} \kappa(x,t)(z_\phi z_v + \nabla z_\phi \cdot \nabla z_v), \\ S(\phi, v) =& \int_{\omega} \phi(x, T_{n-1})v(x, T_{n-1}) + \int_{T_{n-1}}^{T_n}\int_{\omega} \widetilde{\kappa}(x,t)\phi v + \int_{T_{n-1}}^{T_n}\int_{\omega} \kappa|\nabla\chi|^2 z_\phi z_v, \end{aligned} \tag{IV.9}$$

where for any $w \in V^{\omega}_{snap}$, z_w satisfies

$$-z_w(x,t) + \nabla \cdot (\kappa(x,t)\nabla z_w(x,t)) = \chi\frac{\partial w}{\partial t}, \quad \forall t \in (T_{n-1}, T_n).$$

With this spectral problem, one can simplify the proof presented in Section IV.3. However, the numerical implementation of this local spectral problem is more complicated.

IV.3 Convergence analysis

In this section, we will analyze the convergence of our proposed method. To start, we firstly define two norms that are used in the analysis. We define $\|\cdot\|^2_{V^{(n)}}$ and $\|\cdot\|^2_{W^{(n)}}$ by

$$\|u\|^2_{V^{(n)}} = \int_{T_{n-1}}^{T_n}\int_\Omega \kappa|\nabla u|^2 + \frac{1}{2}\int_\Omega u^2(x,T_n^-) + \frac{1}{2}\int_\Omega u^2(x,T_{n-1}^+),$$
$$\|u\|^2_{W^{(n)}} = \|u\|^2_{V^{(n)}} + \int_{T_{n-1}}^{T_n}\|u_t(\cdot,t)\|^2_{H^{-1}(\kappa,\Omega)},$$

where

$$\|u\|_{H^{-1}_{(\kappa,\Omega)}} = \sup_{v\in H^1_0(\Omega)} \frac{\int_\Omega uv}{(\int_\Omega \kappa|\nabla v|^2)^{\frac{1}{2}}}.$$

In the following, we will show the $V^{(n)}$-norm of the error $u_h - u_H$ can be bounded by the $W^{(n)}$-norm of the difference $u_h - w$ for any $w \in V_H^{(n)}$, where u_h is the fine scale solution from Eqn.(IV.2), u_H is the multiscale solution from Eqn.(IV.4), and $V_H^{(n)}$ is the multiscale space defined in the previous section.

Lemma IV.3.1. *Both the fine scale problem (IV.2) and the coarse scale problem (IV.4) have a unique solution for each H, Δt, and κ. Moreover, if u_h is the fine scale solution from Equation (IV.2) and u_H is the multiscale solution from Equation (IV.4), then we have the following estimate*

$$\|u_h - u_H\|^2_{V^{(n)}} \le \begin{cases} C\|u_h - w\|^2_{W^{(n)}} & \text{for } n = 1, \\ C\Big(\|u_h - w\|^2_{W^{(n)}} + \|u_h - u_H\|^2_{V^{(n-1)}}\Big) & \text{for } n > 1, \end{cases}$$

for any $w \in V_H^{(n)}$. If we define the $V^{(0)}$-norm to be 0, then we can write the above as

$$\|u_h - u_H\|^2_{V^{(n)}} \le C\Big(\|u_h - w\|^2_{W^{(n)}} + \|u_h - u_H\|^2_{V^{(n-1)}}\Big) \quad \text{for } n \ge 1,$$

for any $w \in V_H^{(n)}$.

Proof. We will prove the existence of a unique solution for the coarse scale system (IV.4). The proof for the fine scale system is similar. To simplify notations, we define

$$a(u_H, v) = \int_{T_{n-1}}^{T_n} \int_\Omega \frac{\partial u_H}{\partial t} v + \int_{T_{n-1}}^{T_n} \int_\Omega \kappa \nabla u_H \cdot \nabla v + \int_\Omega u_H(x, T_{n-1}^+) v(x, T_{n-1}^+)$$

and

$$F_H(v) = \int_{T_{n-1}}^{T_n} \int_\Omega f v + \int_\Omega g_H^{(n)}(x) v(x, T_{n-1}^+).$$

Using integration by parts for the time variable, it is easy to see that $a(u_H, u_H) = \|u_H\|^2_{V^{(n)}}$. Moreover, we have

$$a(u_H, v) \leq \|u_H\|_{W^{(n)}} \|v\|_{V^{(n)}}. \tag{IV.10}$$

For $u_H \in V_H^{(n)}$, we denote $u_H = \sum_{\omega_i} \chi_i u_i$, then

$$\int_\Omega \frac{\partial u_H}{\partial t} v = \sum_{\omega_i} \int_{\omega_i} \frac{\partial u_i}{\partial t} \chi_i v.$$

Notice that

$$\int_{\omega_i} \frac{\partial u_i}{\partial t} \phi + \int_{\omega_i} \kappa \nabla u_i \cdot \nabla \phi = 0, \forall \phi \in H_0^1(\omega_i),$$

we have

$$\begin{aligned}
&\int_{\omega_i} \frac{\partial u_i}{\partial t} \chi_i v = -\int_{\omega_i} \kappa (\nabla u_i \cdot \nabla \chi_i) v - \int_{\omega_i} \kappa \chi_i (\nabla u_i \cdot \nabla v) \\
&\leq \left(\int_{\omega_i} \kappa (\nabla u_i \cdot \nabla \chi_i)^2 \right)^{\frac{1}{2}} \left(\int_{\omega_i} \kappa v^2 \right)^{\frac{1}{2}} + \left(\int_{\omega_i} \kappa \chi_i^2 |\nabla u_i|^2 \right)^{\frac{1}{2}} \left(\int_{\omega_i} \kappa |\nabla v|^2 \right)^{\frac{1}{2}} \\
&\leq \left(\int_{\omega_i} \kappa (\nabla u_i \cdot \nabla \chi_i)^2 + \int_{\omega_i} \kappa \chi_i^2 |\nabla u_i|^2 \right)^{\frac{1}{2}} \left(\int_{\omega_i} \kappa v^2 + \int_{\omega_i} \kappa |\nabla v|^2 \right)^{\frac{1}{2}}.
\end{aligned}$$

Therefore,

$$\int_\Omega \frac{\partial u_H}{\partial t} v \leq C \left(\sum_{\omega_i} \int_{\omega_i} \kappa (\nabla u_i \cdot \nabla \chi_i)^2 + \sum_{\omega_i} \int_{\omega_i} \kappa \chi_i^2 |\nabla u_i|^2 \right)^{\frac{1}{2}} \left(\int_\Omega \kappa v^2 + \int_\Omega \kappa |\nabla v|^2 \right)^{\frac{1}{2}} \tag{IV.11}$$

for some constant C.

Recall the definition of $\| \cdot \|_{H^{-1}(\kappa,\Omega)}$, using (IV.11) we obtain

$$\| \frac{\partial u_H}{\partial t} \|^2_{H^{-1}(\kappa,\Omega)} \leq C C_1 C_2 \int_\Omega \kappa |\nabla u_H|^2, \tag{IV.12}$$

where

$$\begin{aligned} C_1(t) &= \max_{\sum_i \chi_i u_i \in V_H^{(n)}} \frac{\sum_i \int_{\omega_i} \kappa (\nabla u_i \cdot \nabla \chi_i)^2 + \sum_i \int_{\omega_i} \kappa \chi_i^2 |\nabla u_i|^2}{\int_\Omega \kappa |\nabla \sum_i \chi_i u_i|^2}, \\ C_2(t) &= \max_{v \in V_H^{(n)}} \frac{\int_\Omega \kappa |\nabla v|^2 + \int_\Omega \kappa v^2}{\int_\Omega \kappa |\nabla v|^2} \end{aligned}$$

are two continuous functions of t. Note that it is difficult to get an upper bound for $C_1(t)$ that is independent of discretization and physical parameters. Thus, our results are applicable for a fixed discretization and physical parameters. We plan to investigate a different approach that can provide results are robust w.r.t. discretization and physical parameters. Then we get

$$\begin{aligned} \int_{T_{n-1}}^{T_n} \| \frac{\partial u_H}{\partial t} \|^2_{H^{-1}(\kappa,\Omega)} &\leq C \tilde{C}_1 \tilde{C}_2 \int_{T_{n-1}}^{T_n} \int_\Omega \kappa |\nabla u_H|^2 \\ &\leq C \tilde{C}_1 \tilde{C}_2 \| u_H \|^2_{V^{(n)}}, \end{aligned}$$

where $\tilde{C}_1$ and $\tilde{C}_2$ are two constants. This inequality immediately implies that

$$\| u_H \|^2_{W^{(n)}} \leq C_0 \| u_H \|^2_{V^{(n)}}$$

for a constant C_0.

Recall (IV.10), we get the continuity of the bilinear form $a(\cdot,\cdot)$, which is

$$a(u_H, v) \leq C_0 \|u_H\|_{V^{(n)}} \|v\|_{V^{(n)}}.$$

In addition,

$$F_H(v) \leq C\Big(\|f\|_{L^2((T_{n-1},T_n);\Omega)} + \|g_H^{(n)}\|_{V^{(n)}}\Big)\|v\|_{V^{(n)}}.$$

Hence, by the standard Lax-Milgram theory, the problem (IV.4) has a unique solution.

Next, we prove the error bound. Notice that the coarse scale system (IV.4) is written as

$$a(u_H^{(n)}, v) = F_H(v).$$

On the other hand, the fine scale system (IV.2) is written as

$$a(u_h^{(n)}, v) = F_h(v),$$

where

$$F_h(v) = \int_{T_{n-1}}^{T_n} \int_\Omega fv + \int_\Omega g_h^{(n)}(x) v(x, T_{n-1}^+)$$

and

$$g_h^{(n)}(\cdot) = \begin{cases} u_h^{(n-1)}(\cdot, T_{n-1}^-) & \text{for } n \geq 1, \\ \beta(\cdot) & \text{for } n = 0. \end{cases}$$

Since the multiscale space is conforming, we obtain

$$a(u_h^{(n)} - u_H^{(n)}, v) = F_h(v) - F_H(v).$$

For any $w \in V_H^{(n)}$, we take $v = w - u_H^{(n)}$ and the above implies

$$\begin{aligned}\|w - u_H^{(n)}\|_{V^{(n)}}^2 =& a(w - u_H^{(n)}, w - u_H^{(n)}) \\ =& a(w - u_h^{(n)}, w - u_H^{(n)}) + F_h(w - u_H^{(n)}) - F_H(w - u_H^{(n)}).\end{aligned}$$

Since

$$\begin{aligned}F_h(w - u_H^{(n)}) - F_H(w - u_H^{(n)}) &= \int_\Omega \big(g_h^{(n)}(x) - g_H^{(n)}(x)\big)\big(w - u_H^{(n)}\big) \\ &= \int_\Omega \big(u_h^{(n-1)}(x, T_{n-1}^-) - u_H^{(n-1)}(x, T_{n-1}^-)\big)\big(w - u_H^{(n)}\big) \\ &\le C\|u_h - u_H\|_{V^{(n-1)}} \|w - u_H^{(n)}\|_{V^{(n)}}.\end{aligned}$$

Thus, we obtain

$$\|w - u_H^{(n)}\|_{V^{(n)}} \le C\Big(\|w - u_h^{(n)}\|_{W^{(n)}} + \|u_h - u_H\|_{V^{(n-1)}}\Big).$$

The proof is complete by noting that $\|u_h^{(n)} - u_H^{(n)}\|_{V^{(n)}} \le \|u_h^{(n)} - w\|_{V^{(n)}} + \|w - u_H^{(n)}\|_{V^{(n)}}$.

□

By the above lemma, we can estimate the error of our multiscale solution. In particular, it suffices to find a function w in $V_H^{(n)}$ such that $\|u_h - w\|_{W^{(n)}}$is small. Besides Lemma IV.3.1, we need the following lemma, which can be proved easily by multiplying the equation by χ_i and using integration by parts.

Lemma IV.3.2. *For any v satisfying*

$$\frac{\partial}{\partial t} v - \mathit{div}(\kappa(x,t)\nabla v) = 0 \ \ \mathit{in} \ \ \omega_i \times (T_{n-1}, T_n),$$

we have

$$\int_{\omega_i} \chi_i^2 v^2(x,T_n) + \int_{T_{n-1}}^{T_n} \int_{\omega_i} \kappa |\chi_i^2||\nabla v|^2 \preceq \int_{\omega_i} \chi_i^2 v^2(x,T_{n-1}) + \int_{T_{n-1}}^{T_n} \int_{\omega_i} \kappa |\nabla \chi_i|^2 v^2, \tag{IV.13}$$

where the notation $F \preceq G$ means $F \leq \mathcal{C}G$ with a constant $\mathcal{C}$ independent of the mesh, contrast and the functions involved.

Next, we will state and prove our main result in this section. We will define a few constants that will appear in our error estimate. Given a coarse block K, we define M_K as the number of coarse neighborhoods having non-empty intersection with K and define $M = \max_{K \in \mathcal{T}^H} M_K$. For each coarse neighborhood ω_i, we define

$$D_i = \sup_{v \in H_0^1(\Omega)} \frac{\int_{\omega_i} \kappa |\nabla \chi_i|^2 v^2 + \int_{\omega_i} \kappa \chi_i^2 |\nabla v|^2}{\int_{\omega_i} \kappa |\nabla v|^2 + \int_{\omega_i} \kappa v^2}, \quad F_i = \frac{1}{\min_{x \in \omega_i} \{|\chi_i^+(x)|^2\}}$$

and $D = \max_{1 \leq i \leq N_c} D_i$ and $F = \max_{1 \leq i \leq N_c} F_i$. Note that the constants D and F are independent of the multiscale space and the basis functions. In addition, we define the constant E by

$$E = \sup_{w \in H_0^1(\Omega)} \frac{\int_\Omega \kappa |\nabla w|^2 + \int_\Omega \kappa w^2}{\int_\Omega \kappa |\nabla w|^2}.$$

The constant E is related to the best constant C_Ω in a weighted Poincare inequality by $E = 1 + C_\Omega$ where $\int_\Omega \kappa w^2 \leq C_\Omega \int_\Omega \kappa |\nabla w|^2$ for all $w \in H_0^1(\Omega)$. Recall that u_h is the fine scale solution. We define $\tilde{u}_h$ as the best approximation of u_h within the snapshot space, namely,

$$\tilde{u}_h = \text{argmin}_{v \in V_{\text{snap}}^{(n)}} \|u_h - v\|_{W^{(n)}}.$$

We remark that the norm $\|u_h - \tilde{u}_h\|_{W^{(n)}}$ represents an irreducible error, and it is small by the construction of the snapshot space. Notice that we can write $\tilde{u}_h = \sum_i \chi_i \tilde{u}_{h,i}$ with

$\tilde{u}_{h,i} = \sum_j c_{i,j}\psi_j^{\omega_i}$. Finally, we define $\tilde{u}_{h,i}^+ = \sum_j c_{i,j}\psi_j^{\omega_i^+}$ and the $\|\cdot\|_{V^{(n)}(\omega_i^+)}$ norm in the oversampled region ω_i^+ by

$$\|v\|^2_{V^{(n)}(\omega_i^+)} = \int_{T_{n-1}}^{T_n}\int_{\omega_i^+} \kappa|\nabla v|^2 + \frac{1}{2}\int_{\omega_i^+} v^2(x, T_n^-) + \frac{1}{2}\int_{\omega_i^+} v^2(x, T_{n-1}^+).$$

The main convergence result is presented in the following theorem.

Theorem IV.3.3. *Let u_h be the fine scale solution from Equation (IV.2), u_H be the multi-scale solution from Equation (IV.4). Then we have the following error bound*

$$\begin{aligned}\|u_h - u_H\|^2_{V^{(n)}} \preceq & M(DEF+1)\sum_{i=1}^{N_c}\left(\frac{1}{\lambda^{\omega_i^+}_{L_i+1}}\|\tilde{u}^+_{h,i}\|^2_{V^{(n)}(\omega_i^+)}\right) + \|u_h - \tilde{u}_h\|^2_{W^{(n)}} \\ & + \|u_h - u_H\|^2_{V^{(n-1)}}.\end{aligned} \tag{IV.14}$$

Proof. By Lemma IV.3.1,

$$\|u_h - u_H\|^2_{V^{(n)}} \preceq \inf_{w\in V_H^{(n)}} \|u_h - w\|^2_{W^{(n)}} + \|u_h - u_H\|^2_{V^{(n-1)}}. \tag{IV.15}$$

Therefore, we need to estimate $\inf_{w\in V_H^{(n)}} \|u_h - w\|^2_{W^{(n)}}$.

Note that $\tilde{u}_h = \sum_i \chi_i\tilde{u}_{h,i} = \sum_i\sum_j c_{i,j}\chi_i\psi_j^{\omega_i}$. Using this expression, we can define a projection of $\tilde{u}_h$ into $V_H^{(n)}$ by $P(\tilde{u}_h) = \sum_i\sum_{j\le L_i} c_{i,j}\chi_i\psi_j^{\omega_i}$. Then

$$\begin{aligned}\inf_{w\in V_H^{(n)}} \|u_h - w\|^2_{W^{(n)}} &\le \|u_h - P(\tilde{u}_h)\|^2_{W^{(n)}} \\ &\le \|u_h - \tilde{u}_h\|^2_{W^{(n)}} + \|\tilde{u}_h - P(\tilde{u}_h)\|^2_{W^{(n)}}.\end{aligned} \tag{IV.16}$$

We will estimate $\|\tilde{u}_h - P(\tilde{u}_h)\|^2_{W^{(n)}}$.

By the definition of $\|\cdot\|_{W^{(n)}}$, we have

$$\|\tilde{u}_h - P(\tilde{u}_h)\|^2_{W^{(n)}} = \|\sum_i \chi_i(\tilde{u}_{h,i} - P(\tilde{u}_{h,i}))\|^2_{V^{(n)}} + \int_{T_{n-1}}^{T_n} \|\frac{\partial(\sum_i \chi_i(\tilde{u}_{h,i} - P(\tilde{u}_{h,i})))}{\partial t}\|^2_{H^{-1}(\kappa,\Omega)} ,$$

where $\tilde{u}_{h,i} = \sum_j c_{i,j}\psi_j^{\omega_i}$ and $P(\tilde{u}_{h,i}) = \sum_{j\leq L_i} c_{i,j}\psi_j^{\omega_i}$. Let $e_i = \tilde{u}_{h,i} - P(\tilde{u}_{h,i})$, then $\tilde{u}_h - P(\tilde{u}_h) = \sum_i \chi_i e_i$. Therefore,

$$\|\tilde{u}_h - P(\tilde{u}_h)\|^2_{W^{(n)}} = \|\sum_i \chi_i e_i\|^2_{V^{(n)}} + \int_{T_{n-1}}^{T_n} \|\frac{\partial(\sum_i \chi_i e_i)}{\partial t}\|^2_{H^{-1}(\kappa,\Omega)}. \tag{IV.17}$$

In the following, we will estimate the two terms on the right hand side of (IV.17), separately. Then the proof is done.

First, we estimate the term $\|\sum_i \chi_i e_i\|^2_{V^{(n)}}$. We define the local norm $\|\cdot\|_{V^{(n)}(K)}$ by

$$\|v\|^2_{V^{(n)}(K)} = \int_{T_{n-1}}^{T_n}\int_K \kappa|\nabla v|^2 + \frac{1}{2}\int_K v^2(x,T_n^-) + \frac{1}{2}\int_K v^2(x,T_{n-1}^+).$$

Then we have

$$\|\sum_i \chi_i e_i\|^2_{V^{(n)}} \leq \sum_K \|\sum_i \chi_i e_i\|^2_{V^{(n)}(K)}.$$

Moreover,

$$\|\sum_i \chi_i e_i\|^2_{V^{(n)}(K)} \leq M_K \sum_i \|\chi_i e_i\|^2_{V^{(n)}(K)},$$

where M_K is the number of coarse neighborhoods ω_i's which have nonempty intersection with K.

Therefore,

$$\|\sum_i \chi_i e_i\|^2_{V^{(n)}} \leq \sum_K M_K \sum_i \|\chi_i e_i\|^2_{V^{(n)}(K)} \leq M \sum_i \|\chi_i e_i\|^2_{V^{(n)}(\omega_i)}, \tag{IV.18}$$

where we recall that $M = \max_K\{M_K\}$. Now, we need to estimate the term $\|\chi_i e_i\|^2_{V^{(n)}(\omega_i)}$. Since $\nabla(\chi_i e_i) = e_i\nabla\chi_i + \chi_i\nabla e_i$, we obtain

$$\begin{aligned}\|\chi_i e_i\|^2_{V^{(n)}(\omega_i)} \;\leq\; & 2\int_{T_{n-1}}^{T_n}\int_{\omega_i} \kappa|\nabla\chi_i|^2 e_i^2 + 2\int_{T_{n-1}}^{T_n}\int_{\omega_i}\kappa\chi_i^2|\nabla e_i|^2 \\ & + \frac{1}{2}\int_{\omega_i}\chi_i^2 e_i^2(x,T_n^-) + \frac{1}{2}\int_{\omega_i}\chi_i^2 e_i^2(x,T_{n-1}^+).\end{aligned}$$

Using Lemma IV.3.2, we have

$$\begin{aligned}\|\chi_i e_i\|^2_{V^{(n)}(\omega_i)} &\preceq \int_{T_{n-1}}^{T_n}\int_{\omega_i}\kappa|\nabla\chi_i|^2 e_i^2 + \int_{\omega_i}\chi_i^2 e_i^2(x,T_{n-1}^+) \\ &\preceq \int_{T_{n-1}}^{T_n}\int_{\omega_i}\kappa|\nabla\chi_i|^2 e_i^2 + \int_{\omega_i} e_i^2(x,T_{n-1}^+).\end{aligned}$$

Now we introduce notations in ω_i^+ and denote $e_i^+ = \tilde{u}_{h,i}^+ - P(\tilde{u}_{h,i}^+)$, where $\tilde{u}_{h,i}^+ = \sum_j c_{i,j}\psi_j^{\omega_i^+}$ and $P(\tilde{u}_{h,i}^+) = \sum_{j\leq L_i} c_{i,j}\psi_j^{\omega_i^+}$. It is obvious that $\tilde{u}_{h,i}^+|_{\omega_i} = \tilde{u}_{h,i}$, $P(\tilde{u}_{h,i}^+)|_{\omega_i} = P(\tilde{u}_{h,i})$ and $e_i^+|_{\omega_i} = e_i$. And there holds the following two inequalities,

$$\int_{T_{n-1}}^{T_n}\int_{\omega_i}\kappa|\nabla\chi_i|^2 e_i^2 \leq \int_{T_{n-1}}^{T_n}\int_{\omega_i^+}\kappa|\nabla\chi_i^+|^2|e_i^+|^2, \tag{IV.19}$$

and

$$\int_{\omega_i} e_i^2(x,T_{n-1}^+) \leq \int_{\omega_i^+}|e_i^+(x,T_{n-1}^+)|^2. \tag{IV.20}$$

Thus,

$$\|\chi_i e_i\|^2_{V^{(n)}(\omega_i)} \preceq \int_{T_{n-1}}^{T_n} \int_{\omega_i^+} \kappa |\nabla \chi_i^+|^2 |e_i^+|^2 + \int_{\omega_i^+} |e_i^+(x, T_{n-1}^+)|^2. \tag{IV.21}$$

Substituting (IV.21) into (IV.18), we immediately obtain

$$\| \sum_i \chi_i e_i \|^2_{V^{(n)}} \preceq M \sum_i \left(\int_{T_{n-1}}^{T_n} \int_{\omega_i^+} \kappa |\nabla \chi_i^+|^2 |e_i^+|^2 + \int_{\omega_i^+} |e_i^+(x, T_{n-1}^+)|^2 \right). \tag{IV.22}$$

Next, we will estimate the term

$$\int_{T_{n-1}}^{T_n} \| \frac{\partial (\sum_i \chi_i e_i)}{\partial t} \|^2_{H^{-1}(\kappa,\Omega)}.$$

By definition, we have

$$\begin{aligned} \int_{T_{n-1}}^{T_n} \| \frac{\partial (\sum_i \chi_i e_i)}{\partial t} \|^2_{H^{-1}(\kappa,\Omega)} &= \int_{T_{n-1}}^{T_n} \sup_{w \in H_0^1(\Omega)} \frac{\left(\int_\Omega \sum_i \chi_i \frac{\partial e_i}{\partial t} w \right)^2}{\int_\Omega \kappa |\nabla w|^2} \\ &\leq \int_{T_{n-1}}^{T_n} \sup_{w \in H_0^1(\Omega)} \frac{\left(\sum_i | \int_{\omega_i} \chi_i \frac{\partial e_i}{\partial t} w | \right)^2}{\int_\Omega \kappa |\nabla w|^2}. \end{aligned} \tag{IV.23}$$

Since e_i satisfies the equation

$$\frac{\partial}{\partial t} e_i - \text{div}(\kappa(x,t) \nabla e_i) = 0 \text{ in } \omega_i \times (T_{n-1}, T_n),$$

we have

$$\begin{aligned} \int_{\omega_i} \chi_i \frac{\partial e_i}{\partial t} w &= - \int_{\omega_i} \kappa(x,t) \nabla e_i \cdot \nabla (\chi_i w) \\ &= - \int_{\omega_i} \kappa(x,t) w \nabla e_i \cdot \nabla \chi_i - \int_{\omega_i} \kappa(x,t) \chi_i \nabla e_i \cdot \nabla w. \end{aligned}$$

Moreover,

$$\begin{aligned}\left|\int_{\omega_i}\chi_i\frac{\partial e_i}{\partial t}w\right| =& \left|-\int_{\omega_i}\kappa w\nabla e_i\cdot\nabla\chi_i-\int_{\omega_i}\kappa\chi_i\nabla e_i\cdot\nabla w\right| \\ \leq& \left(\int_{\omega_i}\kappa w^2|\nabla\chi_i|^2\right)^{\frac{1}{2}}\left(\int_{\omega_i}\kappa|\nabla e_i|^2\right)^{\frac{1}{2}}+\left(\int_{\omega_i}\kappa\chi_i^2|\nabla w|^2\right)^{\frac{1}{2}}\left(\int_{\omega_i}\kappa|\nabla e_i|^2\right)^{\frac{1}{2}} \\ \leq& 2\left(\int_{\omega_i}\kappa w^2|\nabla\chi_i|^2+\int_{\omega_i}\kappa\chi_i^2|\nabla w|^2\right)^{\frac{1}{2}}\left(\int_{\omega_i}\kappa|\nabla e_i|^2\right)^{\frac{1}{2}}. \end{aligned} \tag{IV.24}$$

Using the definition

$$D_i=\sup_{v\in H_0^1(\Omega)}\frac{\int_{\omega_i}\kappa|\nabla\chi_i|^2v^2+\int_{\omega_i}\kappa\chi_i^2|\nabla v|^2}{\int_{\omega_i}\kappa|\nabla v|^2+\int_{\omega_i}\kappa v^2},$$

from (IV.24) we obtain

$$\left|\int_{\omega_i}\chi_i\frac{\partial e_i}{\partial t}w\right|\leq 2D_i^{\frac{1}{2}}\left(\int_{\omega_i}\kappa|\nabla w|^2+\int_{\omega_i}\kappa w^2\right)^{\frac{1}{2}}\left(\int_{\omega_i}\kappa|\nabla e_i|^2\right)^{\frac{1}{2}}.$$

Therefore,

$$\begin{aligned}\sum_i\left|\int_{\omega_i}\chi_i\frac{\partial e_i}{\partial t}w\right| \leq& 2\sum_i D_i^{\frac{1}{2}}\left(\int_{\omega_i}\kappa|\nabla w|^2+\int_{\omega_i}\kappa w^2\right)^{\frac{1}{2}}\left(\int_{\omega_i}\kappa|\nabla e_i|^2\right)^{\frac{1}{2}} \\ \leq& 2\left(\sum_i D_i(\int_{\omega_i}\kappa|\nabla w|^2+\int_{\omega_i}\kappa w^2)\right)^{\frac{1}{2}}\left(\sum_i\int_{\omega_i}\kappa|\nabla e_i|^2\right)^{\frac{1}{2}} \\ \leq& 2D^{\frac{1}{2}}M^{\frac{1}{2}}\left(\int_{\Omega}\kappa|\nabla w|^2+\int_{\Omega}\kappa w^2\right)^{\frac{1}{2}}\left(\sum_i\int_{\omega_i}\kappa|\nabla e_i|^2\right)^{\frac{1}{2}}, \end{aligned} \tag{IV.25}$$

where $D=\max\{D_i\}$. Combining (IV.23) with (IV.25), we have

$$\int_{T_{n-1}}^{T_n}\|\frac{\partial(\sum_i\chi_ie_i)}{\partial t}\|^2_{H^{-1}(\kappa,\Omega)}\leq 4DME\left(\sum_i\int_{T_{n-1}}^{T_n}\int_{\omega_i}\kappa|\nabla e_i|^2\right). \tag{IV.26}$$

where we use the definition

$$E = \sup_{w \in H_0^1(\Omega)} \frac{\int_\Omega \kappa |\nabla w|^2 + \int_\Omega \kappa w^2}{\int_\Omega \kappa |\nabla w|^2}.$$

Now, we substitute (IV.26) and (IV.22) into (IV.17), then we have

$$\begin{aligned} &\|\tilde{u}_h - P(\tilde{u}_h)\|^2_{W^{(n)}} \\ \preceq & M \sum_i \left(DE \int_{T_{n-1}}^{T_n} \int_{\omega_i} \kappa |\nabla e_i|^2 + \int_{T_{n-1}}^{T_n} \int_{\omega_i^+} \kappa |\nabla \chi_i^+|^2 |e_i^+|^2 + \int_{\omega_i^+} |e_i^+(x, T_{n-1}^+)|^2 \right). \end{aligned} \tag{IV.27}$$

Note that

$$\begin{aligned} \int_{T_{n-1}}^{T_n} \int_{\omega_i} \kappa |\nabla e_i|^2 &\leq \int_{T_{n-1}}^{T_n} \frac{1}{\min_{x \in \omega_i} \{|\chi_i^+(x)|^2\}} \int_{\omega_i} \kappa |\chi_i^+|^2 |\nabla e_i|^2 \\ &\leq \frac{1}{\min_{x \in \omega_i} \{|\chi_i^+(x)|^2\}} \int_{T_{n-1}}^{T_n} \int_{\omega_i^+} \kappa |\chi_i^+|^2 |\nabla e_i^+|^2. \end{aligned}$$

Applying Lemma IV.3.2 for ω_i^+ then implies

$$\begin{aligned} &\int_{T_{n-1}}^{T_n} \int_{\omega_i} \kappa |\nabla e_i|^2 \\ \leq & \frac{1}{\min_{x \in \omega_i} \{|\chi_i^+(x)|^2\}} \left(\int_{T_{n-1}}^{T_n} \int_{\omega_i^+} \kappa |\nabla \chi_i^+|^2 |e_i^+|^2 + \int_{\omega_i^+} |\chi_i^+|^2 |e_i^+(x, T_{n-1}^+)|^2 \right) \\ \leq & F_i \left(\int_{T_{n-1}}^{T_n} \int_{\omega_i^+} \kappa |\nabla \chi_i^+|^2 |e_i^+|^2 + \int_{\omega_i^+} |e_i^+(x, T_{n-1}^+)|^2 \right), \end{aligned} \tag{IV.28}$$

where

$$F_i = \frac{1}{\min_{x \in \omega_i} \{|\chi_i^+(x)|^2\}}.$$

Substituting (IV.28) into (IV.27) gives

$$\|\tilde{u}_h - P(\tilde{u}_h)\|^2_{W^{(n)}} \tag{IV.29}$$

$$\preceq M \sum_i (DEF_i + 1) \left(\int_{T_{n-1}}^{T_n} \int_{\omega_i^+} \kappa |\nabla \chi_i^+|^2 |e_i^+|^2 + \int_{\omega_i^+} |e_i^+(x, T_{n-1}^+)|^2 \right)$$

$$\preceq M(DEF + 1) \sum_i \left(\int_{T_{n-1}}^{T_n} \int_{\omega_i^+} \widetilde{\kappa}^+(x,t) |e_i^+|^2 + \int_{\omega_i^+} |e_i^+(x, T_{n-1}^+)|^2 \right), \tag{IV.30}$$

where $F = \max\{F_i\}$. Using the spectral problem, we have

$$\|\tilde{u}_h - P(\tilde{u}_h)\|^2_{W^{(n)}} \preceq M(DEF + 1) \sum_i \left(\frac{1}{\lambda^{\omega_i^+}_{L_i+1}} \|\tilde{u}^+_{h,i}\|^2_{V^{(n)}(\omega_i^+)} \right). \tag{IV.31}$$

Combine (IV.15), (IV.16) and (IV.31), and we finally obtain

$$\|u_h - u_H\|^2_{V^{(n)}} \preceq M(DEF + 1) \sum_i \left(\frac{1}{\lambda^{\omega_i^+}_{L_i+1}} \|\tilde{u}^+_{h,i}\|^2_{V^{(n)}(\omega_i^+)} \right) + \|u_h - \tilde{u}_h\|^2_{W^{(n)}}$$
$$+ \|u_h - u_H\|^2_{V^{(n-1)}}.$$

□

IV.4 Numerical results. Offline step.

We present representative numerical results in this section to show the performance of the proposed method. In particular, we solve Equation (IV.1) using the space-time GMsFEM. We take the space domain Ω to be the unit square $[0, 1] \times [0, 1]$ and divide it into 10×10 uniform coarse square blocks. Each coarse block is then divided into 10×10 uniform fine square blocks. That is, Ω is partitioned by 100×100 square fine-grid blocks. The whole time interval is $[0, 1.6]$ (i.e., $T = 1.6$) and is divided into two uniform coarse time intervals and each coarse time interval is then divided into 8 fine time intervals. The source

term is chosen as $f = 1$ and a continuous initial condition $\beta(x_1, x_2) = \sin(\pi x_1)\sin(\pi x_2)$ is imposed. We employ three different high-contrast permeability fields $\kappa(x,t)$'s to examine our method, which will be shown in the following three cases separately. In each case, we first solve for u_h from Equation (IV.2) to obtain the fine-grid solution. Then we solve for the multiscale solution u_H using the space-time GMsFEM. To compare the accuracy, we will use the following error quantities:

$$e_1 = \left(\frac{\int_0^T \|u_H(t) - u_h(t)\|^2_{L^2(\Omega)}}{\int_0^T \|u_h(t)\|^2_{L^2(\Omega)}} \right)^{1/2}, \quad e_2 = \left(\frac{\int_0^T \int_\Omega \kappa |\nabla(u_H(t) - u_h(t))|^2}{\int_0^T \int_\Omega \kappa |\nabla u_h(t)|^2} \right)^{1/2}. \tag{IV.32}$$

Since we are using the technique of randomized oversampling in the computation of the snapshot space, we would like to introduce the concept of *snapshot ratio*, which is defined as the ratio of the number of chosen randomized snapshots over the number of the full snapshots on one coarse neighborhood ω_i. In the following experiment with 100×100 fine-grid mesh, this number of the full snapshots on each coarse neighborhood is calculated by $n^{\text{snap}}_{\text{total}} = 21 \times 21 + 40 \times 8 = 761$.

IV.4.1 High-contrast medium translated in time

We construct a time dependent high-contrast permeability field $\kappa(x,t)$ by uniformly translating the permeability field after every other fine time step. High-contrast permeability fields at the initial and final time steps are shown in Figure IV.1. Next, we consider applying the space-time GMsFEM to Equation (IV.1) and solve for the multiscale solution u_H. Recall the procedures that are described in the Section IV.2, where we need to construct the snapshot spaces in the first place.

First, we fix $L_i = 11$ for all ω_i's and examine the influences of various buffer numbers on the solution errors e_1 and e_2. The results are displayed in the left table of Table IV.1. It is observed that when increasing the buffer numbers, one can get more accurate solutions,

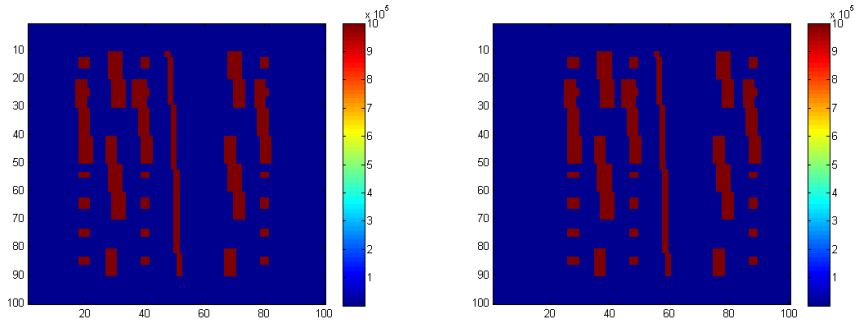

Figure IV.1: High-contrast permeability field 1. Left: initial. Right: final.

which is as expected. But the error decays very slowly, which indicates that using different buffer numbers doesn't affect the convergence rate too much. Based on this observation, it is not necessary to choose a large buffer number in order to improve convergence rate. Then we consider the choice of L_i, the number of eigenbasis in a neighborhood. With the fixed buffer number $p_{\text{bf}} = 8$, we examine the convergence behaviors of using different L_i's. Relative errors of multiscale solutions are shown in the right table of Table IV.1. We observe that as adding more offline basis, the relative errors decay. We note that the error decay is becoming slow after using more than 40 offline basis per space-time course neighborhood. This is because space-time problems have very large degrees of freedom, while we are using a small part of them at this stage. We will further explore a fast decay technique by adding online basis functions in Section IV.5.

To see a more quantitative relationship between the relative offline errors and the values of L_i as well as being inspired by the result in Theorem IV.3.3, we inspect the values of $1/\Lambda_*$ and the corresponding squared errors (see Table IV.2 and Figure IV.2), where $\Lambda_* = \min_{\omega_i} \lambda^{\omega_i}_{L_i+1}$ and $\{\lambda^{\omega_i}_j\}$ are the eigenvalues associated with the eigenbasis computed by spectral problem (IV.6) in each ω_i. We note that when plotting Figure IV.2, we don't use the values of case $L_i = 2$, because in this case as in the case with one basis function per node, the method does not converge as we do not have sufficient number of basis func-

p_{bf}	Snapshot ratio	e_1	e_2
1	0.0158	6.18%	53.90%
4	0.0197	5.66%	48.04%
8	0.0250	5.17%	45.86%
12	0.0302	5.16%	43.83%
20	0.0407	4.71%	41.14%
30	0.0539	4.35%	38.68%
40	0.0670	4.23%	37.60%

L_i	Snapshot ratio	e_1	e_2
2	0.0131	17.03%	129.14%
6	0.0184	8.11%	62.59%
10	0.0237	6.97%	54.85%
20	0.0368	4.81%	41.18%
30	0.0499	3.29%	31.64%
40	0.0631	2.28%	24.43%
50	0.0762	1.54%	18.45%

Table IV.1: First permeability field. Left: errors with the fixed number of offline basis $L_i = 11$. Right: errors with the fixed buffer number $p_{\text{bf}} = 8$.

tions. We note that the two curves in Figure IV.2 track each other somewhat closely. This indicates that $1/\Lambda_*$'s and e_2^2's are correlated and we calculate for the correlation coefficient to be $corrcoef(1/\Lambda_*, e_2^2) = 0.9778$.

Observing the dimensions of the offline spaces V_{off}, one can see that compared with the traditional fine-scale finite element method, the proposed space-time GMsFEM uses much fewer degrees of freedom while achieving an accurate solution. Also, by inspecting the snapshot ratios, one can see that the use of randomization can reduce the dimension of snapshot spaces substantially. We would like to comment that oversampling technique is necessary for the randomization. For example, in the case $L_i = 6$ and $p_{\text{bf}} = 8$, if without oversampling the errors e_1 and e_2 are 11.19% and 88.42%, respectively, which are worse than the errors obtained with oversampling.

IV.4.2 Four channels translated in time

In this subsection, we consider a more structured high-contrast permeability field $\kappa(x, t)$, which has four channels inside and these four channels are translated uniformly in time. High-contrast permeability fields at the initial and final time steps are shown in Figure IV.3. We repeat our steps from the previous example by fixing L_i and p_{bf}, separately. The results are shown in Table IV.3. One can still observe that increasing the buffer numbers

L_i	$1/\Lambda_*$	e_1^2	e_2^2
2	0.2734	2.90%	166.78%
6	0.0120	0.66%	39.17%
10	0.0085	0.49%	30.08%
20	0.0061	0.23%	16.96%
30	0.0053	0.11%	10.01%
40	0.0048	0.05%	5.97%
50	0.0042	0.02%	3.40%

Table IV.2: $1/\Lambda_*$ values and errors.

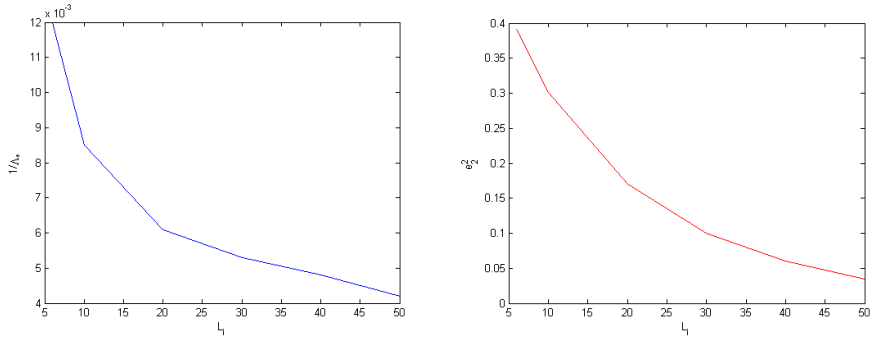

Figure IV.2: Left: $1/\Lambda_*$ vs L_i; Right: e_2^2 vs L_i.

will slowly reduce the relative errors and when fixing the buffer number, the errors decay as more offline basis are being used. Using a similar approach, we can also get the cross-correlation coefficient between e_2^2 and $1/\Lambda_*$, which is 0.9863. This suggests a linear relationship between e_2^2 and $1/\Lambda_*$ and verifies Theorem IV.3.3.

IV.4.3 Four channels rotated in time

In the third example, we consider another structured high-contrast permeability field $\kappa(x,t)$ which has four channels inside and these four channels are rotated anticlockwise around the center by 11.25 degrees after each fine time step. High contrast permeability fields at the initial time step is shown in Figure IV.4. We repeat the same procedures as in

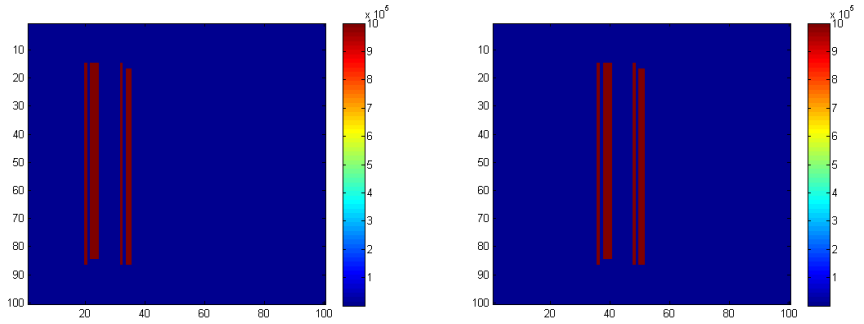

Figure IV.3: High-contrast permeability field 2. Left: initial. Right: final.

p_{bf}	Snapshot ratio	e_1	e_2
1	0.0158	7.42%	61.87%
4	0.0197	7.30%	58.95%
8	0.0250	7.14%	57.30%
12	0.0302	7.00%	54.01%
20	0.0407	6.81%	50.85%
30	0.0539	6.61%	49.30%
40	0.0670	6.43%	48.26%

L_i	Snapshot ratio	e_1	e_2
2	0.0131	11.91%	104.95%
6	0.0184	8.33%	70.82%
10	0.0237	7.25%	58.25%
20	0.0368	5.67%	43.10%
30	0.0499	3.90%	32.75%
40	0.0631	2.73%	27.08%
50	0.0762	1.86%	20.70%

Table IV.3: Second permeability field. Left: errors with the fixed number of offline basis $L_i = 11$. Right: errors with the fixed buffer number $p_{\text{bf}} = 8$.

the previous two examples. The results are shown in Table IV.4 and one can draw similar conclusions as before. The cross-correlation coefficient between e_2^2 and $1/\Lambda_*$ is calculated as 0.9959. This shows a linear relationship between e_2^2 and $1/\Lambda_*$ (see Theorem IV.3.3).

IV.5 Residual based online adaptive procedure

As we observe in the previous examples, the offline errors do not decrease rapidly after several multiscale functions are selected. In these cases, online basis functions can help to reduce the error and get an accurate approximation of the fine-scale solution [17]. Next, we will derive a framework for the online multiscale method.

We use the index $m \geq 1$ to represent the online enrichment level. At m's enrichment

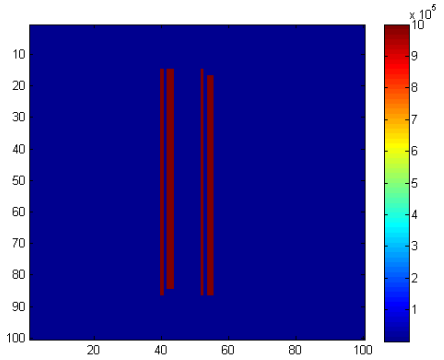

Figure IV.4: High-contrast permeability field 3 at the initial time.

p_{bf}	Snapshot ratio	e_1	e_2
1	0.0158	8.68%	72.86%
4	0.0197	8.67%	71.67%
8	0.0250	8.56%	71.42%
12	0.0302	8.44%	68.87%
20	0.0407	8.18%	65.88%
30	0.0539	7.96%	61.56%
40	0.0670	7.58%	57.58%

L_i	Snapshot ratio	e_1	e_2
2	0.0131	10.41%	109.40%
6	0.0184	9.40%	83.60%
10	0.0237	8.63%	70.84%
20	0.0368	7.42%	57.66%
30	0.0499	6.14%	47.78%
40	0.0631	4.75%	39.89%
50	0.0762	3.29%	30.11%

Table IV.4: Third permeability field. Left: errors with the fixed number of offline basis $L_i = 11$. Right: errors with the fixed buffer number $p_{\text{bf}} = 8$.

level, we define V_{ms}^m as the corresponding space-time GMsFEM space and u_{ms}^m the corresponding solution to (IV.4) in space V_{ms}^m. We remark that both offline and online basis functions can be contained in V_{ms}^m, and define $V_{ms}^0 = V_{\text{off}}$. We will show how to update V_{ms}^{m+1} from V_{ms}^m iteratively.

Now we consider constructing online basis functions. We first note that the computation of online basis occurs in the iterative process and involves the use of the residual. This is the different from offline basis functions, of which the computation happens before the iterative process. In particular, we use local residuals for the multiscale solution u_{ms}^m to construct the online basis functions for enrichment level $m + 1$. For brevity, we denote

the left hand side of (IV.4) by $a(u_{ms}^{(n)}, v)$ and the right hand side $G(v)$. That is, the solution $u_{ms} = \oplus_{n=1}^{N} u_{ms}^{(n)}$ where $u_{ms}^{(n)}$ satisfies

$$a(u_{ms}^{(n)}, v) = G(v), \quad \forall v \in V_H^{(n)}.$$

We consider a given coarse neighborhood ω_i. Suppose that at m's enrichment level, we need to add an online basis function $\phi \in V_h$ in ω_i. Then the required $\phi = \oplus_{n=1}^{N} \phi^{(n)}$ satisfies that $\phi^{(n)}$ solves the equation

$$a(\phi^{(n)}, v) = R^{(n)}(v), \quad \forall v \in V_h,$$

where $R^{(n)}(v) = G(v) - a(u_{ms}^{m(n)}, v)$ is the online residual at the coarse time interval $[T_{n-1}, T_n]$.

In the following, we would like to form a residual based online algorithm in each coarse time interval $[T_{n-1}, T_n]$, see Algorithm 2. We will omit the time index (n) on the spaces and solutions in this description. We note that online enrichment are performed on non-overlapping coarse neighborhoods. Thus, we divide the $\{\omega_i\}_{i=1}^{N_c}$ into P non-overlapping groups and denote each group by $\{\omega_i\}_{i \in I_p}$, $p = 1, ..., P$. We denote by M the number of online iterations.

To further improve the efficiency of the online method, we can adopt an online adaptive procedure. In this adaptive approach, we only perform the online enrichment in coarse neighborhoods that have a cumulative residual that is θ fraction of the total residual. More precisely, assume that the $V^{(n)}$ norm of local residuals on $\{\omega_i | i \in I_p\}$, denoted by $\{r_i | i \in I_p\}$, are arranged so that

$$r_{p_1} \geq r_{p_2} \geq r_{p_3} \geq \cdots \geq r_{p_J},$$

where we suppose $I_p = \{p_1, p_2, p_3, \cdots, p_J\}$. Instead of adding $\{\phi_i | i \in I_p\}$ into V_{ms}^m at

Algorithm 2 Residual based online algorithm

1: **Initialization:** Offline space $V_{ms}^0 = V_{\text{off}}$, offline solution $u_{ms}^0 = u_{\text{H}}$.
2: **for** $m = 0$ to M: **do**
3: **for** $p = 1$ to P **do**
4: (1) On each $\omega_i (i \in I_p)$, compute residual $R^m(v) = a(u_{ms}^m, v) - G(v), \quad v \in V_h$.
5: (2) For each i, solve $a(\phi_i, v) = R^m(v), \quad \forall v \in V_h$.
6: (3) Set $V_{ms}^m = V_{ms}^m \cup \{\phi_i | i \in I_p\}$.
7: (4) Solve for a new $u_{ms}^m \in V_{ms}^m$ satisfying $a(u_{ms}^m, v) = G(v), \quad \forall v \in V_{ms}^m$.
8: **end for**
9: Set $V_{ms}^{m+1} = V_{ms}^m$, and $u_{ms}^{m+1} = u_{ms}^m$.
10: **end for**

step 6 in Algorithm 2, we only add the basis $\{\phi_1, \cdots, \phi_k\}$ for the corresponding coarse neighborhoods such that k is the smallest integer satisfying

$$\Sigma_{i=1}^k r_{p_i}^2 \geq \theta \Sigma_{i=1}^J r_{p_i}^2.$$

In the examples below, we will see that the proposed adaptive procedure gives a better convergence and is more efficient.

IV.6 Numerical results. Online step.

We conduct numerical experiments to demonstrate the proposed online method in solving Equation (IV.1). To implement the space-time online GMsFEM, a fixed number of offline basis functions need to be constructed in each coarse neighborhood. Then we conduct the online process by following Algorithm 2. In this experiment, we use the same space-time domain and mesh (coarse and fine), the same source term f and initial condition $\beta(x_1, x_2)$, the same definitions of relative errors e_1 and e_2, as in Section IV.4. The permeability field $\kappa(x, t)$ is chosen as the high-contrast permeability field 1 from Section IV.4.1. The buffer number in the computation of snapshot space is chosen to be 8.

First, we implement the space-time online GMsFEM by choosing different numbers of offline basis functions ($L_i = 1, 2, 3, 4, 5$) on every coarse neighborhood. The relative errors of online solutions are presented in Table IV.5 and Table IV.6. Note that in the first column, we show the number of basis functions used for each coarse neighborhood ω_i along with the degrees of freedom (DOF) of multiscale space on each coarse time interval which are the numbers in parentheses, after online enrichment. For example, $2(162)$ in the first column means that after online enrichment, 2 multiscale basis are used on each ω_i and the DOF of multiscale space on each coarse time interval is 162. And if we initially choose $L_i = 1$, then it means 1 online iteration is performed, which add 1 online basis to each ω_i. If $L_i = 2$ initially, then it means we do not perform any online iteration and 2 multiscale basis are offline basis functions. By observing each column, one can see that the errors decay fast with more online iterations being performed. This is observed for both e_1 and e_2 when $L_i \geq 4$. This suggests that in this specific setting, we can get a fast online convergence with 4 offline basis chosen on each ω_i. After a small number of online iterations, the relative errors decrease to a significantly small level. We consider reducing the high contrast of the permeability field $\kappa(x, t)$ from 10^6 to 100. Then we look at the relative errors of online multiscale solutions (see Table IV.7 and Table IV.8). The same phenomena can be observed except that the fast online convergence rate can be achieved for any choice of L_i. This implies that the number of offline basis functions used to guarantee a fast online convergence rate is related to the high contrast of the permeability field.

Next, we perform online adaptive basis construction procedure with $\theta = 0.7$. The numerical results for using 3, 4, and 5 offline basis per coarse neighborhood are shown in Table IV.9. Notice that "$M1 + M2$" in the DOF columns means $M1$ degrees of freedom are used on the first coarse time interval and $M2$ degrees of freedom on the second coarse time interval. To compare the behaviors of online processes with and without adaptivity,

DOF	$e_1(L_i = 1)$	$e_1(L_i = 2)$	$e_1(L_i = 3)$	$e_1(L_i = 4)$	$e_1(L_i = 5)$
1(81)	97.57%	-	-	-	-
2(162)	93.20%	96.71%	-	-	-
3(243)	44.24%	23.22%	21.27%	-	-
4(324)	15.37%	6.53%	7.17e-1%	10.20%	-
5(405)	8.65%	3.69%	2.06e-1%	2.58e-1%	5.20%
6(486)	5.15%	1.71%	5.41e-2%	1.75e-2%	1.06e-1%
7(567)	2.58%	3.11e-1%	5.54e-3%	6.12e-4%	2.99e-3%

Table IV.5: Relative online errors e_1, with the different numbers of offline basis functions. High contrast = 10^6.

DOF	$e_2(L_i = 1)$	$e_2(L_i = 2)$	$e_2(L_i = 3)$	$e_2(L_i = 4)$	$e_2(L_i = 5)$
1(81)	138%	-	-	-	-
2(162)	113%	114%	-	-	-
3(243)	84.93%	139%	104%	-	-
4(324)	82.48%	82.08%	11.43%	73.50%	-
5(405)	69.15%	51.13%	3.29%	4.78%	48.26%
6(486)	51.17%	34.00%	1.01%	3.53e-1%	1.86%
7(567)	37.93%	7.81%	1.05e-1%	9.89e-3%	4.75e-2%

Table IV.6: Relative online errors e_2, with the different numbers of offline basis functions. High contrast = 10^6.

we plot out the log values of e_2 against DOFs. See Figure IV.5. We observe that to achieve a certain error, fewer online basis functions are needed with adaptivity. This indicates that the proposed adaptive procedure gives us better convergence and is more efficient.

DOF	e_1(1 basis)	e_1(2 basis)	e_1(3 basis)	e_1(4 basis)	e_1(5 basis)
1(81)	19.28%	-	-	-	-
2(162)	1.97%	13.03%	-	-	-
3(243)	2.81e-1%	9.81e-1%	9.27%	-	-
4(324)	3.48e-2%	1.24e-1%	2.23e-1%	8.34%	-
5(405)	1.89e-3%	1.11e-2%	9.70e-2%	2.09e-1%	7.38%
6(486)	2.67e-5%	1.33e-4%	2.07e-4%	8.71e-3%	1.56e-1%
7(567)	2.51e-7%	9.32e-7%	1.45e-6%	1.16e-4%	8.62e-3%

Table IV.7: Relative online errors e_1, with the different numbers of offline basis functions. High contrast = 100.

DOF	e_2(1 basis)	e_2(2 basis)	e_2(3 basis)	e_2(4 basis)	e_2(5 basis)
1(81)	219%	-	-	-	-
2(162)	14.75%	123%	-	-	-
3(243)	3.35%	8.37%	81.80%	-	-
4(324)	4.03e-1%	1.11%	2.63%	67.86%	-
5(405)	2.11e-2%	1.01e-1%	1.68e-1%	2.29%	59.93%
6(486)	5.61e-4%	1.64e-3%	3.71e-3%	1.35e-1%	1.77%
7(567)	4.57e-6%	1.72e-5%	2.29e-5%	2.08e-3%	1.41e-1%

Table IV.8: Relative online errors e_2 with the different numbers of offline basis functions. High contrast = 100.

3 offline basis		4 offline basis		5 offline basis	
DOF	e_2	DOF	e_2	DOF	e_2
243+243	104%	324+324	73.50%	405+405	48.26%
323+322	10.57%	399+401	3.56%	471+473	1.95%
403+392	1.49%	468+466	2.13e-1%	533+536	1.21e-1%
480+465	9.81e-2%	541+529	1.03e-2%	599+603	6.81e-3%
552+533	4.24e-3%	611+601	5.00e-4%	670+669	3.41e-4%

Table IV.9: Relative online adaptive errors e_2 with different numbers of offline basis functions.

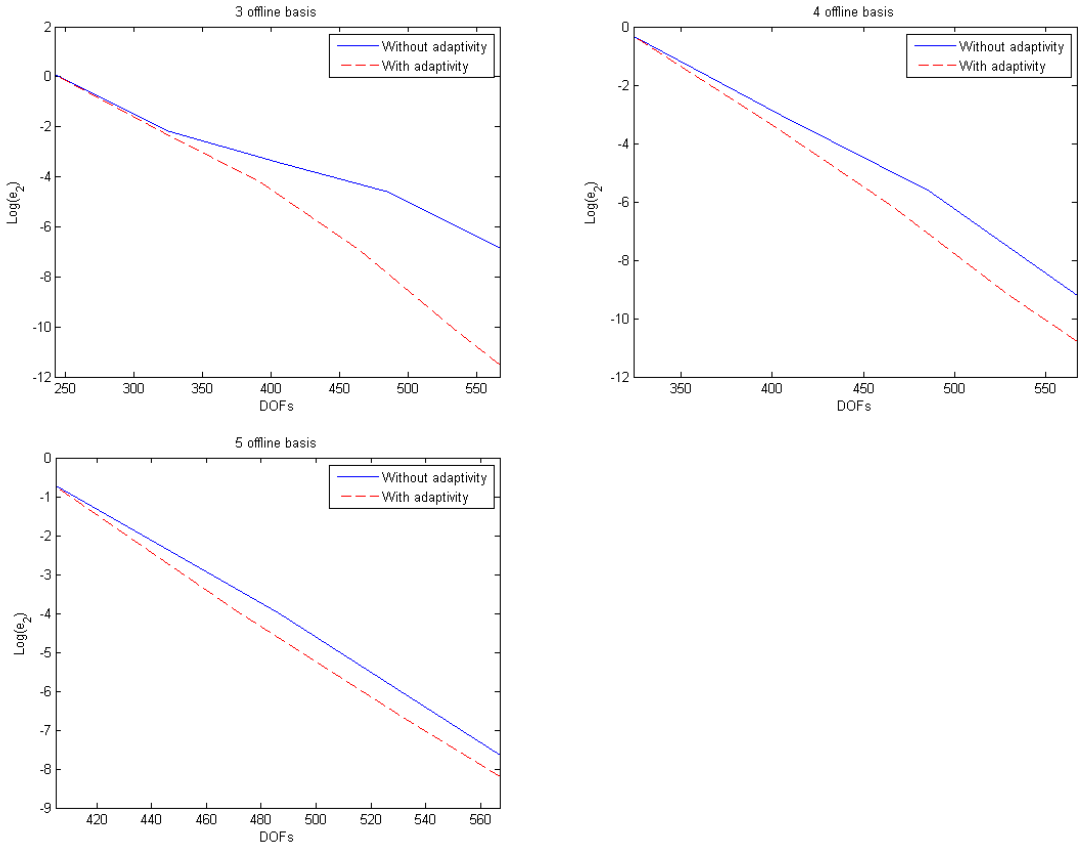

Figure IV.5: Adaptivity v.s. no adaptivity.

V. SPACE-TIME GMSDGM

This section is a continuing exploration in space-time multiscale method for solving parabolic equations. The main objective is to develop a multiscale model reduction framework within Generalized Multiscale Discontinuous Galerkin Method (GMsDGM) that uses space-time coarse cells. There are several benefits in exploring the space-time multiscale methods in the fashion of discontinuous Galerkin, such as small local problems, and feasible for parallel computing. We organize the sections as following. In Section V.1, we present the underlying model problem and the space-time GMsDGM formulations. In Section V.2, we construct the multiscale basis by introducing the snapshot and spectral problems. In Section V.3, we present representative numerical examples. In Section V.4, we demonstrate the convergence analysis for our proposed method.

V.1 Model problem

Suppose Ω is a bounded domain in $\mathbb{R}^2$ with a Lipschitz boundary $\partial\Omega$, and $[0,T]$ $(T > 0)$ ia a time interval. We consider the same model problem as presented in Section IV.1, which is the following parabolic equation

$$\begin{aligned} \frac{\partial}{\partial t}u - \text{div}(\kappa(x,t)\nabla u) &= f \qquad && \text{in } \Omega \times (0,T), \\ u &= 0 \qquad && \text{on } \partial\Omega \times (0,T), \\ u(x,0) &= \beta(x) \qquad && \text{in } \Omega, \end{aligned} \tag{V.1}$$

where $\kappa(x,t)$ is a time dependent heterogeneous media, $\beta(x)$ is the initial condition.

Now we present the general idea of space-time GMsDGM. We will use the space-time finite element method to solve problem (V.1) on the coarse grid. That is, we find $u_H \in V_H$

such that

$$\int_0^T \sum_{K\in\Omega}\int_K \frac{\partial u_H}{\partial t}v + \int_0^T \sum_{K\in\Omega}\int_K \kappa\nabla u_H\cdot\nabla v$$
$$-\int_0^T \sum_{E\in\varepsilon^H}\int_E (\{\kappa\nabla u_H\cdot n_E\}[\![v]\!] + \{\kappa\nabla v\cdot n_E\}[\![u_H]\!]) + \int_0^T \sum_{E\in\varepsilon^H}\frac{\gamma}{h}\int_E \bar{\kappa}[\![u_H]\!][\![v]\!]$$
$$+\sum_{n=0}^{N-1}\sum_{K\in\Omega}\int_K [u_H(x,T_n)]v(x,T_n^+) = \int_0^T \sum_{K\in\Omega}\int_K fv + \sum_{K\in\Omega}\int_K \beta(x)v(x,T_0^+), \tag{V.2}$$

$\forall v \in V_H$ where V_H is the multiscale finite element space which will be introduced in the following subsections, n_E is a fixed unit normal vector defined on the coarse edge $E \in \varepsilon^H$, $\gamma > 0$ is a penalty parameter. Note that, in (V.2), the spatial average operator $\{\cdot\}$ and spatial jump operator $[\![\cdot]\!]$ are defined in the classical way. The time jump operator $[\cdot]$ is defined such that

$$\begin{cases} [u_H(x,T_n)] = u_H(x,T_n^+) - u_H(x,T_n^-) & \text{for } n \geq 1, \\ [u_H(x,T_n)] = u_H(x,T_0^+) & \text{for } n = 0. \end{cases}$$

The computational cost for solving the equation (V.2) is huge since we need to compute the solution u_H in the whole time interval $(0,T)$ at one time. To fix this issue, we assume the solution space V_H can be decomposed into a direct sum of subspaces and each subspace only contains functions defined on one single coarse time interval (T_{n-1},T_n), that is,

$$V_H = \oplus_{n=1}^N V_H^{(n)},$$

where $V_H^{(n)}$ is defined as

$$V_H^{(n)} = \{v(x,t) | v(\cdot,t) = 0 \text{ for } t \in (0,T)\backslash(T_{n-1},T_n)\}.$$

Then the coarse problem (V.2) can be decomposed into a sequence of sub-problems: find $u_H^{(n)} \in V_H^{(n)}$ such that

$$\begin{aligned}
&\int_{T_{n-1}}^{T_n} \sum_{K\in\Omega} \int_K \frac{\partial u_H^{(n)}}{\partial t} v + \int_{T_{n-1}}^{T_n} \sum_{K\in\Omega} \int_K \kappa \nabla u_H^{(n)} \cdot \nabla v \\
&- \int_{T_{n-1}}^{T_n} \sum_{E\in\varepsilon^H} \int_E \Big(\{\kappa \nabla u_H^{(n)} \cdot n_E\} [\![v]\!] + \{\kappa \nabla v \cdot n_E\} [\![u_H^{(n)}]\!] \Big) + \int_{T_{n-1}}^{T_n} \sum_{E\in\varepsilon^H} \frac{\gamma}{h} \int_E \bar{\kappa} [\![u_H^{(n)}]\!] [\![v]\!] \\
&+ \sum_{K\in\Omega} \int_K u_H^{(n)}(x, T_{n-1}^+) v(x, T_{n-1}^+) = \int_{T_{n-1}}^{T_n} \sum_{K\in\Omega} \int_K f v + \sum_{K\in\Omega} \int_K g_H^{(n)}(x) v(x, T_{n-1}^+),
\end{aligned} \tag{V.3}$$

$\forall v \in V_H^{(n)}$, where

$$g_H^{(n)}(\cdot) = \begin{cases} u_H^{(n-1)}(\cdot, T_{n-1}^-) & \text{for } n \geq 1, \\ \beta(\cdot) & \text{for } n = 0. \end{cases}$$

Then, the solution u_H of the problem (V.2) is the direct sum of all these $u_H^{(n)}$'s, that is $u_H = \oplus_{n=1}^N u_H^{(n)}$.

V.2 Construction of multiscale basis functions

V.2.1 Snapshot space

We will use the techniques of oversampling and randomization in the construction of snapshot spaces. Firstly, we describe the notations for oversampled regions. We denote by K^+ the oversampled space region of $K \subset K^+$, defined by adding several fine-grid layers around K. Time interval (T_{n-1}^*, T_n) is defined as the left-side oversampled time region for (T_{n-1}, T_n). We will construct inexpensive snapshots by imposing random boundary conditions on the local region $K^+ \times (T_{n-1}^*, T_n)$. In particular, one solves the following

local problems:

$$\frac{\partial}{\partial t}\psi_j^+ - \text{div}(\kappa(x,t)\nabla\psi_j^+) = 0 \text{ in } K^+ \times (T_{n-1}^*, T_n),$$
$$\psi_j^+(x,t) = r_l \text{ on } \partial\left(K^+ \times (T_{n-1}^*, T_n)\right),$$

where r_l is a random vector.

Then the local snapshot space on $K^+ \times (T_{n-1}^*, T_n)$ is

$$V_{\text{snap}}^{K^+} = \text{span}\{\psi_j^+(x,t) | j = 1, \cdots, L^K + p_{\text{bf}}^K\},$$

where L^K and p_{bf}^K are the number of local offline basis in K and the buffer number, respectively.

V.2.2 Offline space

Now, we are going to construct the offline basis functions. This is done by performing a space reduction in each local space-time snapshot space through some well-defined spectral problems. In particular, we propose the following eigenvalue problem on $K^+ \times (T_{n-1}, T_n)$:

Find $(\phi, \lambda) \in V_{\text{snap}}^{K^+} \times \mathbb{R}$ such that

$$A_n(\phi, v) = \lambda S_n(\phi, v), \quad \forall v \in V_{\text{snap}}^{K^+}, \tag{V.4}$$

where the bilinear operators $A_n(\phi, v)$ and $S_n(\phi, v)$ are defined by

$$\begin{aligned} A_n(\phi, v) &= \int_{T_{n-1}}^{T_n}\int_{K^+} \kappa(x,t)\nabla\phi\cdot\nabla v + H\int_{T_{n-1}}^{T_n}\int_{\partial K^+} \kappa(x,t)(\nabla\phi\cdot\vec{n})(\nabla v\cdot\vec{n}), \\ S_n(\phi, v) &= \int_{T_{n-1}}^{T_n}\int_{\partial K^+} \kappa(x,t)\phi v + \int_{K^+}\phi(x, T_{n-1})v(x, T_{n-1}), \end{aligned} \tag{V.5}$$

where H is the size of coarse elements.

The eigenvalues $\{\lambda_j^{K^+} | j = 1, 2, \cdots L^K + p_{\text{bf}}^K\}$ from (V.4) are arranged in the ascending order, and we select the first L^K eigenfunctions, which are associated with the first L^K ordered eigenvalues, and denote them by $\{\Psi_1^{K^+,\text{off}}, \cdots, \Psi_{L^K}^{K^+,\text{off}}\}$. Using these eigenfunctions, we can define

$$\phi_j^{K^+}(x,t) = \sum_{k=1}^{L^K + p_{\text{bf}}^K} (\Psi_j^{K^+,\text{off}})_k \psi_k^+(x,t), \qquad j = 1, 2, \cdots, L^K,$$

where $(\Psi_j^{K^+,\text{off}})_k$ denotes the k-th component of $\Psi_j^{K^+,\text{off}}$, and $\psi_k^+(x,t)$ is the snapshot basis function computed on $K^+ \times (T_{n-1}^*, T_n)$ as in the previous section. Then we can obtain the offline basis functions $\phi_j^K(x,t)$ on the target region $K \times (T_{n-1}, T_n)$ by restricting $\phi_j^{K^+}(x,t)$ onto $K \times (T_{n-1}, T_n)$.

The local offline space on $K \times (T_{n-1}, T_n)$ can be defined as

$$V_{\text{off}}^K = \text{span}\{\phi_j^K(x,t) | j = 1, \cdots, L^K\}.$$

Note that one can take $V_H^{(n)}$ in (V.3) as $V_H^{(n)} = V_{\text{off}}^{(n)} = \text{span}\{\phi_j^{K_i}(x,t) | 1 \le i \le N_c, 1 \le j \le L_i\}$. As a result, $V_H = V_{\text{off}} = \oplus_{n=1}^N V_H^{(n)}$.

V.3 Numerical result

We take the space domain Ω to be the unit square $[0,1] \times [0,1]$ and divide it into 8×8 uniform coarse square blocks. Each coarse block is then divided into 10×10 uniform fine square blocks. That is, Ω is partitioned by 80×80 square fine-grid blocks. The whole time interval is $[0,2]$ (i.e., $T = 2$) and is divided into two uniform coarse time intervals and each coarse time interval is then divided into 10 fine time intervals. The source term is chosen as $f = 1$ and a continuous initial condition $\beta(x_1, x_2) = \sin(\pi x_1)\sin(\pi x_2)$ is imposed.

We first solve for the FEM solution u_h, then solve for the multiscale solution u_H using the space-time GMsDGM. To compare the accuracy, we will use the following error quantities:

$$e_1 = \left(\frac{\int_0^T \|u_H(t) - u_h(t)\|^2_{L^2(\Omega)}}{\int_0^T \|u_h(t)\|^2_{L^2(\Omega)}} \right)^{1/2}, \quad e_2 = \left(\frac{\int_0^T \int_\Omega \kappa |\nabla(u_H(t) - u_h(t))|^2}{\int_0^T \int_\Omega \kappa |\nabla u_h(t)|^2} \right)^{1/2}. \tag{V.6}$$

Since we are using the technique of randomized oversampling in the computation of the snapshot space, we would like to introduce the concept of *snapshot ratio*, which is defined as the ratio of the number of chosen randomized snapshots over the number of the full snapshots on one coarse block K. In the following experiment with 80×80 fine-grid mesh, this number of the full snapshots on each coarse neighborhood is calculated by $n^{\text{snap}}_{\text{total}} = 11 \times 11 + 40 \times 10 = 521$.

V.3.1 High-contrast medium translated in time

We construct a time dependent high-contrast permeability field $\kappa(x, t)$ by uniformly translating the permeability field after every other fine time step. High-contrast permeability fields at the initial is shown in Figure V.1. The number of local offline basis that will be used in each K_i, denoted by L_i, and the buffer number p_{bf} is fixed to be 30. Relative errors of multiscale solutions are shown in Table V.1. We observe that as adding more offline basis, the relative errors decay. However, we note that the error decay is becoming slow eventually, which is quite similar to what we have observed in Section IV.4. We emphasize that we are dealing with space-time problems, which have large degrees of freedom. Taking this experiment as an example, at each fixed time, there are $81 \times 81 = 6561$ fine-scale degrees of freedom, which results in 65610 degrees of freedom in each coarse time interval. But due to our space-time GMsDGM approach, we only use 3200 degrees of freedom in each coarse time interval to obtain a relative energy error of 19.44%, which is a great

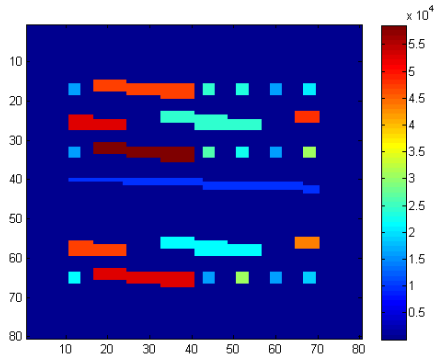

Figure V.1: Initial high-contrast permeability field 1.

L_i	$dim(V_{\text{off}})$	Snapshot ratio	e_1	e_2
2	128	0.0997	47.84%	97.42%
6	384	0.1121	18.85%	84.20%
10	640	0.1246	9.57%	63.55%
20	1280	0.1558	4.56%	40.30%
30	1920	0.1869	3.02%	28.98%
40	2560	0.2181	2.03%	22.89%
50	3200	0.2492	1.64%	19.44%

Table V.1: Permeability field 1: relative errors using different L_i's.

saving and exhibit the effect of model reduction.

To see a more quantitative relationship between the relative errors and the values of L_i as well as being inspired by the result in Theorem V.4.2, we inspect the values of $1/\Lambda_*$ and the corresponding squared errors (see Table V.2, where $\Lambda_* = \min_{K_i} \lambda_{L_i+1}^{K_i}$ and $\{\lambda_j^{K_i}\}$ are the eigenvalues associated with the eigenbasis computed by spectral problem (V.4) in each K_i. We calculate for the correlation coefficient to be

$$corrcoef(1/\Lambda_*, e_2^2) = 0.9617.$$

This indicates that $1/\Lambda_*$'s are correlated to e_2^2's.

L_i	$1/\Lambda_*$	e_2^2
2	0.2653e-4	0.9491
6	0.1483e-4	0.7089
10	0.1172e-4	0.4038
20	0.0881e-4	0.1624
30	0.0674e-4	0.0840
40	0.0527e-4	0.0524
50	0.0456e-4	0.0378

Table V.2: Permeability field 1: $1/\Lambda_*$ values and errors.

V.3.2 Four channels translated in time

In this subsection, we consider a more structured high-contrast permeability field $\kappa(x,t)$, which has four channels inside and these four channels are translated uniformly in time. High-contrast permeability fields at the initial time steps is shown in Figure V.2. We repeat our steps from the previous example by fixing L_i and p_{bf}, separately. The results are shown in Table V.3. One can still observe that when fixing the buffer number, the errors decay as more offline basis are being used. Using a similar approach, we can also get the cross-correlation coefficient between e_2^2 and $1/\Lambda_*$, which is

$$corrcoef(1/\Lambda_*, e_2^2) = 0.9672.$$

This suggests a linear relationship between e_2^2 and $1/\Lambda_*$.

V.3.3 Four channels rotated in time

In the third example, we consider another structured high-contrast permeability field $\kappa(x,t)$ which has four channels inside and these four channels are rotated anticlockwise around the center by 18 degrees after each fine time step. High contrast permeability fields at the initial time step is shown in Figure V.3. We repeat the same procedures as in the

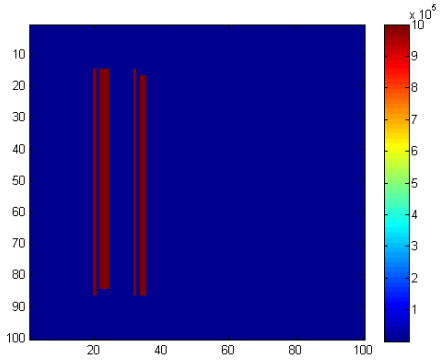

Figure V.2: Initial high-contrast permeability field 2.

L_i	$dim(V_{\text{off}})$	Snapshot ratio	e_1	e_2
2	128	0.0997	23.52%	96.83%
6	384	0.1121	16.89%	80.17%
10	640	0.1246	11.25%	71.53%
20	1280	0.1558	5.57%	46.96%
30	1920	0.1869	2.43%	31.62%
40	2560	0.2181	1.70%	22.03%
50	3200	0.2492	1.31%	18.07%

Table V.3: Permeability field 2: relative errors using different L_i's.

previous two examples. The results are shown in Table V.4 and one can draw similar conclusions as before. The cross-correlation coefficient between e_2^2 and $1/\Lambda_*$ is calculated as 0.9937.

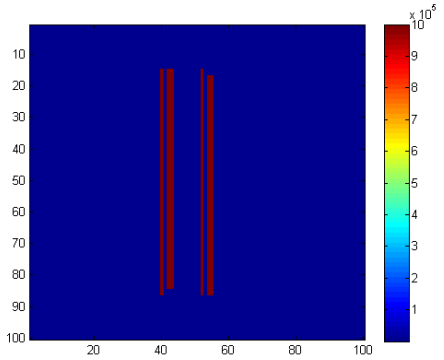

Figure V.3: Initial high-contrast permeability field 3.

L_i	$dim(V_{\text{off}})$	Snapshot ratio	e_1	e_2
2	128	0.0997	26.86%	99.32%
6	384	0.1121	17.64%	79.92%
10	640	0.1246	12.71%	68.59%
20	1280	0.1558	6.53%	48.32%
30	1920	0.1869	4.63%	36.64%
40	2560	0.2181	3.74%	30.77%
50	3200	0.2492	2.74%	24.87%

Table V.4: Permeability field 3: relative errors using different L_i's.

V.4 Convergence analysis

In this section, we will analyze the convergence of our proposed method. To start, we firstly define two norms $\|\cdot\|^2_{V^{(n)}}$ and $\|\cdot\|^2_{W^{(n)}}$ on each $V_H^{(n)}$ by

$$\|u\|^2_{V^{(n)}} = \int_{T_{n-1}}^{T_n} a_{dg}(u,u) + \frac{1}{2}\sum_{K\in\Omega}\int_K u^2(x,T_n^-) + \frac{1}{2}\sum_{K\in\Omega}\int_K u^2(x,T_{n-1}^+),$$

$$\|u\|^2_{W^{(n)}} = \|u\|^2_{V^{(n)}} + \int_{T_{n-1}}^{T_n} \|u_t(x,t)\|^2_{H^{-1}_{dg}(\kappa,\Omega)},$$

where the bilinear form $a_{dg}(\cdot,\cdot)$ is defined as

$$a_{dg}(u,v) = \sum_{K\in\Omega}\int_K \kappa\nabla u\cdot\nabla v - \sum_{E\in\varepsilon^H}\int_E (\{\kappa\nabla u\cdot n_E\}[\![v]\!] + \{\kappa\nabla v\cdot n_E\}[\![u]\!]) + \sum_{E\in\varepsilon^H}\frac{\gamma}{h}\int_E \bar{\kappa}[\![u]\!][\![v]\!],$$

$\|u_t(\cdot,t)\|_{H_{dg}^{-1}(\kappa,\Omega)}$ is defined as

$$\|u\|_{H_{dg}^{-1}(\kappa,\Omega)} = \sup_{v|_K\in H_0^1(K)\forall K\in\Omega}\frac{\sum_{K\in\Omega}\int_K uv}{(a_{dg}(v,v))^{\frac{1}{2}}}.$$

In the following, we will show the $V^{(n)}$-norm of the error $u_h - u_H$ can be bounded by the $W^{(n)}$-norm of the difference $u_h - w$ for any $w \in V_H^{(n)}$, where u_h is the solution by solving Eqn.(V.3) using all fine-scale basis (u_h is a proxy of the exact solution), u_H is the solution by solving Eqn.(V.3) using space-time multiscale basis constructed in Section V.2, and $V_H^{(n)}$ is the multiscale space defined in in Section V.2.

Lemma V.4.1. *Both the coarse scale problem (V.3) and the fine scale problem associated with (V.3) have a unique solution. Moreover, if u_h is the fine-scale solution by solving Eqn.(V.3) using all fine-scale basis (a proxy of the exact solution), u_H be the coarse scale solution by solving Eqn.(V.3) using space-time multiscale basis constructed in Section V.2. We have the following estimate*

$$\|u_h - u_H\|^2_{V^{(n)}} \leq \begin{cases} C\|u_h - w\|^2_{W^{(n)}} & \textit{for } n = 1, \\ C(\|u_h - w\|^2_{W^{(n)}} + \|u_h - u_H\|^2_{V^{(n-1)}}) & \textit{for } n > 1, \end{cases}$$

for any $w \in V_H^{(n)}$. *If we define the* $V^{(0)}$*-norm to be* 0*, then we can write*

$$\|u_h - u_H\|^2_{V^{(n)}} \leq C(\|u_h - w\|^2_{W^{(n)}} + \|u_h - u_H\|^2_{V^{(n-1)}}) \quad \textit{for } n \geq 1,$$

for any $w \in V_H^{(n)}$.

Proof. We will prove the existence of a unique solution for the coarse scale system (V.3). The proof for the fine scale system is similar. To simplify notations, we define

$$a(u_H, v) = \int_{T_{n-1}}^{T_n} \sum_{K\in\Omega} \int_K \frac{\partial u_H}{\partial t} v + \int_{T_{n-1}}^{T_n} a_{dg}(u_H, v) + \sum_{K\in\Omega} \int_K u_H(x, T_{n-1}^+) v(x, T_{n-1}^+)$$

and

$$F_H(v) = \int_{T_{n-1}}^{T_n} \sum_{K\in\Omega} \int_K f v + \sum_{K\in\Omega} \int_K g_H^{(n)}(x) v(x, T_{n-1}^+).$$

Using integration by parts for the time variable, it is easy to see that $a(u_H, u_H) = \|u_H\|^2_{V^{(n)}}$.

Moreover, we have

$$a(u_H, v) \leq \|u_H\|_{W^{(n)}} \|v\|_{V^{(n)}}. \tag{V.7}$$

For $u_H \in V_H^{(n)}$, we denote $u_H = \sum_{K_i} u_i$, then u_i solves the following local problem

$$\int_{K_i} \frac{\partial u_i}{\partial t} v - \int_{K_i} \text{div}(\kappa \nabla u_i) v = 0, \ \forall v \in H_0^1(K_i).$$

Using integration by parts, we have

$$\begin{aligned}
\int_{K_i} \frac{\partial u_i}{\partial t} v &= \int_{\partial K_i} \kappa(\nabla u_i \cdot \vec{n}) v - \int_{K_i} \kappa \nabla u_i \cdot \nabla v \\
&\leq \left(H \int_{\partial K_i} \kappa |\nabla u_i \cdot \vec{n}|^2 \right)^{\frac{1}{2}} \left(\frac{1}{H} \int_{\partial K_i} \kappa v^2 \right)^{\frac{1}{2}} + \left(\int_{K_i} \kappa |\nabla u_i|^2 \right)^{\frac{1}{2}} \left(\int_{K_i} \kappa |\nabla v|^2 \right)^{\frac{1}{2}} \\
&\leq \left(H \int_{\partial K_i} \kappa |\nabla u_i \cdot \vec{n}|^2 \right)^{\frac{1}{2}} \left(\int_{K_i} \kappa |\nabla v|^2 \right)^{\frac{1}{2}} + \left(\int_{K_i} \kappa |\nabla u_i|^2 \right)^{\frac{1}{2}} \left(\int_{K_i} \kappa |\nabla v|^2 \right)^{\frac{1}{2}},
\end{aligned} \tag{V.8}$$

where on the last step we use trace theorem.

Combining the temporal integration and using integration by parts for the temporal variable, we get

$$
\begin{aligned}
&\int_{T_{n-1}}^{T_n}\int_{K_i}\frac{\partial u_i}{\partial t}v\\
\leq&\int_{T_{n-1}}^{T_n}\left(H\int_{\partial K_i}\kappa|\nabla u_i\cdot\vec{n}|^2\right)^{\frac{1}{2}}\left(\int_{K_i}\kappa|\nabla v|^2\right)^{\frac{1}{2}}+\int_{T_{n-1}}^{T_n}\left(\int_{K_i}\kappa|\nabla u_i|^2\right)^{\frac{1}{2}}\left(\int_{K_i}\kappa|\nabla v|^2\right)^{\frac{1}{2}}\\
\leq&\left[\left(H\int_{T_{n-1}}^{T_n}\int_{\partial K_i}\kappa|\nabla u_i\cdot\vec{n}|^2\right)^{\frac{1}{2}}+\left(\int_{T_{n-1}}^{T_n}\int_{K_i}\kappa|\nabla u_i|^2\right)^{\frac{1}{2}}\right]\left(\int_{T_{n-1}}^{T_n}\int_{K_i}\kappa|\nabla v|^2\right)^{\frac{1}{2}}\\
\leq&\sqrt{2}\left(H\int_{T_{n-1}}^{T_n}\int_{\partial K_i}\kappa|\nabla u_i\cdot\vec{n}|^2+\int_{T_{n-1}}^{T_n}\int_{K_i}\kappa|\nabla u_i|^2\right)^{\frac{1}{2}}\left(\int_{T_{n-1}}^{T_n}\int_{K_i}\kappa|\nabla v|^2\right)^{\frac{1}{2}}.
\end{aligned}
$$

Using the eigenvalue problem (V.4), we have

$$
\begin{aligned}
&\int_{T_{n-1}}^{T_n}\int_{K_i}\frac{\partial u_i}{\partial t}v\leq\sqrt{2\lambda_i}\left(\int_{T_{n-1}}^{T_n}\int_{\partial K_i}\kappa u_i^2+\int_{K_i}u_i(T_{n-1}^+)^2\right)^{\frac{1}{2}}\left(\int_{T_{n-1}}^{T_n}\int_{K_i}\kappa|\nabla v|^2\right)^{\frac{1}{2}}\\
\leq&\sqrt{2\lambda_i}\left(H\int_{T_{n-1}}^{T_n}\int_{K_i}\kappa|\nabla u_i|^2+\int_{K_i}u_i(T_{n-1}^+)^2\right)^{\frac{1}{2}}\left(\int_{T_{n-1}}^{T_n}\int_{K_i}\kappa|\nabla v|^2\right)^{\frac{1}{2}}\\
\leq&\sqrt{2\lambda_i}\left[\sqrt{H}\left(\int_{T_{n-1}}^{T_n}\int_{K_i}\kappa|\nabla u_i|^2\right)^{\frac{1}{2}}+\sqrt{2}\left(\frac{1}{2}\int_{K_i}u_i(T_{n-1}^+)^2\right)^{\frac{1}{2}}\right]\left(\int_{T_{n-1}}^{T_n}\int_{K_i}\kappa|\nabla v|^2\right)^{\frac{1}{2}}.
\end{aligned}
\tag{V.9}
$$

This immediately leads to

$$
\begin{aligned}
&\int_{T_{n-1}}^{T_n}\sum_i\int_{K_i}\frac{\partial u_i}{\partial t}v\\
\leq&C_0\sqrt{\Lambda}\left[\sqrt{H}\left(\int_{T_{n-1}}^{T_n}\sum_i\int_{K_i}\kappa|\nabla u_i|^2\right)^{\frac{1}{2}}+\sqrt{2}\left(\frac{1}{2}\sum_i\int_{K_i}u_i(T_{n-1}^+)^2\right)^{\frac{1}{2}}\right]\\
&\left(\int_{T_{n-1}}^{T_n}\sum_i\int_{K_i}\kappa|\nabla v|^2\right)^{\frac{1}{2}}\\
\leq&C_0\sqrt{\Lambda}(\sqrt{H}+\sqrt{2})\|u_H\|_{V^{(n)}}\|v\|_{V^{(n)}}
\end{aligned}
\tag{V.10}
$$

for some constant C_0, where $\Lambda = \max_i \lambda_i$.

Recall the definition of $\|u_t(\cdot,t)\|_{H_{dg}^{-1}(\kappa,\Omega)}$, by using (V.10) we obtain

$$\begin{aligned}\int_{T_{n-1}}^{T_n} \|\frac{\partial u_H}{\partial t}(\cdot,t)\|^2_{H_{dg}^{-1}(\kappa,\Omega)} =& \sup_{v|_K\in H_0^1(K)\forall K\in\Omega} \int_{T_{n-1}}^{T_n} \frac{(\sum_i \int_{K_i} \frac{\partial u_i}{\partial t} v)^2}{a_{dg}(v,v)} \\ \leq& C_1\Lambda(H+2)\|u_H\|^2_{V^{(n)}},\end{aligned} \tag{V.11}$$

for some constant C_1. Thus we have $\|u_H\|_{W^{(n)}} \leq \sqrt{1+C_1\Lambda(H+2)}\|u_H\|_{V^{(n)}}$.

Recall (V.7), we get

$$a(u_H,v) \leq \sqrt{1+C_1\Lambda(H+2)}\|u_H\|_{V^{(n)}}\|v\|_{V^{(n)}}.$$

In addition,

$$F_H(v) \leq C\Big(\|f\|_{L^2((T_{n-1},T_n);\Omega)} + \|g_H^{(n)}\|_{V^{(n)}}\Big)\|v\|_{V^{(n)}}.$$

Hence, by Lax-Milgram theory, the problem (V.3) has a unique solution.

Next, we prove the error bound. By the definition of $\|\cdot\|_{V^{(n)}}$,

$$\begin{aligned}&\|u_h-u_H\|^2_{V^{(n)}} \\ =&\frac{1}{2}\int_\Omega (u_h-u_H)^2|_{t=T_n^-} + \frac{1}{2}\int_\Omega (u_h-u_H)^2|_{t=T_{n-1}^+} + \int_{T_{n-1}}^{T_n} a_{dg}(u_h-u_H,u_h-u_H) \\ =&\int_{T_{n-1}}^{T_n}\int_\Omega \frac{\partial(u_h-u_H)}{\partial t}(u_h-u_H) + \int_\Omega (u_h-u_H)^2|_{t=T_{n-1}^+} \\ &+\int_{T_{n-1}}^{T_n} a_{dg}(u_h-u_H,u_h-u_H) \\ =&\int_{T_{n-1}}^{T_n}\int_\Omega \frac{\partial(u_h-u_H)}{\partial t}(u_h-w) + \int_\Omega (u_h-u_H)(u_h-w)|_{t=T_{n-1}^+} \\ &+\int_{T_{n-1}}^{T_n} a_{dg}(u_h-u_H,u_h-w) + \int_{T_{n-1}}^{T_n}\int_\Omega \frac{\partial(u_h-u_H)}{\partial t}(w-u_H) \\ &+\int_\Omega (u_h-u_H)(w-u_H)|_{t=T_{n-1}^+} + \int_{T_{n-1}}^{T_n} a_{dg}(u_h-u_H,w-u_H).\end{aligned} \tag{V.12}$$

From (V.3) and the similar formulation for fine scale solution u_h, we have

$$\begin{aligned}&\int_{T_{n-1}}^{T_n}\int_\Omega \frac{\partial(u_h-u_H)}{\partial t}v + \int_{T_{n-1}}^{T_n} a_{dg}(u_h-u_H,v) + \int_\Omega (u_h-u_H)v|_{t=T_{n-1}^+}\\ =&\int_\Omega \left(g_h^{(n)}-g_H^{(n)}\right)v(x,T_{n-1}^+),\ \forall v\in V_H^{(n)}.\end{aligned} \tag{V.13}$$

Therefore, taking $v=w-u_H$ and combining the equation (V.12) and (V.13), we obtain

$$\begin{aligned}\|u_h-u_H\|^2_{V^{(n)}} \quad=\quad & \int_{T_{n-1}}^{T_n}\int_\Omega \frac{\partial(u_h-u_H)}{\partial t}(u_h-w) + \int_\Omega (u_h-u_H)(u_h-w)|_{t=T_{n-1}^+}\\ &+\int_{T_{n-1}}^{T_n} a_{dg}(u_h-u_H,u_h-w) + \int_\Omega \left(g_h^{(n)}-g_H^{(n)}\right)(w-u_H)|_{t=T_{n-1}^+}.\end{aligned}$$

Using integration by parts, we have

$$\begin{aligned}&\int_{T_{n-1}}^{T_n}\int_\Omega \frac{\partial(u_h-u_H)}{\partial t}(u_h-w) + \int_\Omega (u_h-u_H)(u_h-w)|_{t=T_{n-1}^+}\\ =&-\int_{T_{n-1}}^{T_n}\int_\Omega \frac{\partial(u_h-w)}{\partial t}(u_h-u_H) + \int_\Omega (u_h-u_H)(u_h-w)|_{t=T_n^-}.\end{aligned}$$

Thus,

$$\begin{aligned}\|u_h-u_H\|^2_{V^{(n)}} =&-\int_{T_{n-1}}^{T_n}\int_\Omega \frac{\partial(u_h-w)}{\partial t}(u_h-u_H) + \int_\Omega (u_h-u_H)(u_h-w)|_{t=T_n^-}\\ &+\int_{T_{n-1}}^{T_n} a_{dg}(u_h-u_H,u_h-w) + \int_\Omega \left(g_h^{(n)}-g_H^{(n)}\right)(u_h-u_H)|_{t=T_{n-1}^+}\\ &+\int_\Omega \left(g_h^{(n)}-g_H^{(n)}\right)(w-u_h)|_{t=T_{n-1}^+}\\ \le&\left(\int_{T_{n-1}}^{T_n}\frac{|\int_\Omega \frac{\partial(u_h-w)}{\partial t}(u_h-u_H)|^2}{a_{dg}(u_h-u_H,u_h-u_H)}\right)^{\frac{1}{2}}\left(\int_{T_{n-1}}^{T_n} a_{dg}(u_h-u_H,u_h-u_H)\right)^{\frac{1}{2}}\\ &+\|(u_h-u_H)(\cdot,T_n^-)\|_{L^2(\Omega)}\|(u_h-w)(\cdot,T_n^-)\|_{L^2(\Omega)}\\ &+\|u_h-w\|_{L^2((T_{n-1},T_n);\kappa)}\|u_h-u_H\|_{L^2((T_{n-1},T_n);\kappa)}\\ &+\|g_h^{(n)}-g_H^{(n)}\|_{L^2(\Omega)}(\|(u_h-u_H)(\cdot,T_{n-1}^+)\|_{L^2(\Omega)}\\ &+\|(u_h-w)(\cdot,T_{n-1}^+)\|_{L^2(\Omega)}).\end{aligned}$$

Using Young's inequality, we have

$$\|u_h - u_H\|^2_{V^{(n)}} \leq \frac{1}{2}\|u_h - u_H\|^2_{V^{(n)}} + 2\left(C\|u_h - w\|^2_{W^{(n)}} + \|g_h^{(n)} - g_H^{(n)}\|^2_{L^2(\Omega)}\right).$$

and

$$\|g_h^{(n)} - g_H^{(n)}\|^2_{L^2(\Omega)} = \begin{cases} 0 & \text{for } n = 1 \\ \|u_h^{(n-1)}(\cdot, T_{n-1}^-) - u_H^{(n-1)}(\cdot, T_{n-1}^-)\|^2_{L^2(\Omega)} & \text{for } n > 1 \end{cases}$$
$$\leq \begin{cases} 0 & \text{for } n = 1 \\ \|u_h - u_H\|^2_{V^{(n-1)}} & \text{for } n > 1 \end{cases}.$$

□

The main convergence result is presented in the following theorem.

Theorem V.4.2. *Let u_h be the solution by solving Eqn.(V.3) using all fine-scale basis (a proxy of the exact solution), u_H be the solution by solving Eqn.(V.3) using space-time multiscale basis constructed in Section V.2. Let $\tilde{u}_h = \text{argmin}_{v \in V_{snap}^{(n)}}\{\|u_h - v\|_{W^{(n)}}\}$ and we denote $\tilde{u}_h = \sum_i \tilde{u}_{h,i}$ with $\tilde{u}_{h,i} = \sum_j c_{i,j}\psi_j^{K_i}$. There holds*

$$\|u_h - u_H\|^2_{V^{(n)}} \preceq C \sum_i \left(\left(\frac{1}{h\lambda_{L_i+1}^{K_i}} + 1\right)\frac{4}{3\lambda_{L_i+1}^{K_i} - 1} + H\right)\|\kappa \nabla \tilde{u}_h \cdot n\|^2_{L^2(\partial K_i)} \Bigg)$$
$$+ \|u_h - \tilde{u}_h\|^2_{W^{(n)}} + \|u_h - u_H\|^2_{V^{(n-1)}}.$$

Proof. By Lemma V.4.1,

$$\|u_h - u_H\|^2_{V^{(n)}} \preceq \inf_{w \in V_H^{(n)}} \|u_h - w\|^2_{W^{(n)}} + \|u_h - u_H\|^2_{V^{(n-1)}}. \tag{V.14}$$

Therefore, we need to estimate $\inf_{w\in V_H^{(n)}} \|u_h - w\|^2_{W^{(n)}}$. Note $\tilde{u}_h = \sum_i \tilde{u}_{h,i} = \sum_i \sum_j c_{i,j}\psi_j^{K_i}$. Using this expression, we can define a projection of $\tilde{u}_h$ into $V_H^{(n)}$ by

$$P(\tilde{u}_h) = \sum_i \sum_{j\le L_i} c_{i,j}\psi_j^{K_i}.$$

Then

$$\begin{aligned}\inf_{w\in V_H^{(n)}} \|u_h - w\|^2_{W^{(n)}} &\le \|u_h - P(\tilde{u}_h)\|^2_{W^{(n)}} \\ &\le \|u_h - \tilde{u}_h\|^2_{W^{(n)}} + \|\tilde{u}_h - P(\tilde{u}_h)\|^2_{W^{(n)}}. \end{aligned} \tag{V.15}$$

We will estimate $\|\tilde{u}_h - P(\tilde{u}_h)\|^2_{W^{(n)}}$.

By the definition of $\|\cdot\|_{W^{(n)}}$, we can obtain

$$\begin{aligned}&\|\tilde{u}_h - P(\tilde{u}_h)\|^2_{W^{(n)}} \\ \le&\|\sum_i (\tilde{u}_{h,i} - P(\tilde{u}_{h,i}))\|^2_{V^{(n)}} + \left(\int_{T_{n-1}}^{T_n} \frac{\sum_i |\int_{K_i} \frac{\partial(\tilde{u}_h - P(\tilde{u}_h))}{\partial t}(u_h - u_H)|^2}{a_{dg}(u_h - u_H, u_h - u_H)}\right),\end{aligned}$$

where $\tilde{u}_{h,i} = \sum_j c_{i,j}\psi_j^{K_i}$ and $P(\tilde{u}_{h,i}) = \sum_{j\le L_i} c_{i,j}\psi_j^{K_i}$. Let $e_i = \tilde{u}_{h,i} - P(\tilde{u}_{h,i})$, then $\tilde{u}_h - P(\tilde{u}_h) = \sum_i e_i$. Thus,

$$\|\tilde{u}_h - P(\tilde{u}_h)\|^2_{W^{(n)}} \le \|\sum_i e_i\|^2_{V^{(n)}} + \left(\int_{T_{n-1}}^{T_n} \frac{\sum_i |\int_{K_i} \frac{\partial e_i}{\partial t}(u_h - u_H)|^2}{a_{dg}(u_h - u_H, u_h - u_H)}\right). \tag{V.16}$$

In the following, we will estimate the two terms on the right hand side of (V.16), separately. Then the proof is done.

Since $\frac{\partial}{\partial t}e_i - \nabla\cdot\kappa\nabla e_i = 0$ in $K_i \times (T_{n-1}, T_n)$, it deduces

$$\int_{T_{n-1}}^{T_n}\int_{K_i} (\frac{\partial}{\partial t}e_i)e_i + \int_{T_{n-1}}^{T_n}\int_{K_i} \kappa|\nabla e_i|^2 = \int_{T_{n-1}}^{T_n}\int_{\partial K_i} \kappa\nabla e_i \cdot n e_i,$$

which implies

$$\int_{T_{n-1}}^{T_n}\int_{K_i}\kappa|\nabla e_i|^2=\int_{T_{n-1}}^{T_n}\int_{\partial K_i}\kappa\nabla e_i\cdot ne_i+\frac{1}{2}\int_{K_i}e_i^2(x,T_{n-1}^+)-\frac{1}{2}\int_{K_i}e_i^2(x,T_n^-).$$

Therefore,

$$\begin{aligned}
&\int_{T_{n-1}}^{T_n}\int_{K_i}\kappa|\nabla e_i|^2+\frac{1}{2}\int_{K_i}e_i^2(x,T_{n-1}^+)+\frac{1}{2}\int_{K_i}e_i^2(x,T_n^-)\\
=&\int_{T_{n-1}}^{T_n}\int_{\partial K_i}\kappa\nabla\tilde{u}_h\cdot ne_i+\int_{K_i}e_i^2(x,T_{n-1}^+)\\
\leq&\frac{1}{\lambda_{L_i+1}^{K_i}}\|\kappa^{\frac{1}{2}}\nabla\tilde{u}_h\cdot n\|_{L^2(\partial K_i)}^2+\frac{\lambda_{L_i+1}^{K_i}}{4}\|\kappa^{\frac{1}{2}}e_i\|_{L^2(\partial K_i)}^2+\int_{K_i}e_i^2(x,T_{n-1}^+)\\
\leq&(\frac{3}{4}-\frac{1}{\lambda_{L_i+1}^{K_i}})^{-1}\frac{1}{\lambda_{L_i+1}^{K_i}}\|\kappa^{\frac{1}{2}}\nabla\tilde{u}_h\cdot n\|_{L^2(\partial K)}^2\\
=&\frac{4}{3\lambda_{L_i+1}^{K_i}-1}\|\kappa\nabla\tilde{u}_h\cdot n\|_{L^2(\partial K)}^2.
\end{aligned}$$

Now, we estimate the term

$$\frac{1}{h}\sum_E\int_{T_{n-1}}^{T_n}\int_E\kappa[\![e_i]\!]^2,$$

which is the following,

$$\begin{aligned}
\frac{1}{h}\sum_E\int_{T_{n-1}}^{T_n}\int_E\kappa[\![e_i]\!]^2\leq\frac{C}{h}\sum_i\int_{T_{n-1}}^{T_n}\int_{\partial K_i}\kappa e_i^2&\leq\sum_i\frac{C}{h\lambda_{L_i+1}^{K_i}}\int_{T_{n-1}}^{T_n}\int_{K_i}\kappa|\nabla e_i|^2\\
&\leq\sum_i\frac{C}{h\lambda_{L_i+1}^{K_i}}\frac{4}{3\lambda_{L_i+1}^{K_i}-1}\|\kappa\nabla\tilde{u}_h\cdot n\|_{L^2(\partial K)}^2.
\end{aligned}$$

We will also need to estimate the term

$$\int_{T_{n-1}}^{T_n}\frac{\sum_i|\int_{K_i}\frac{\partial e_i}{\partial t}(\widehat{u_h-u_H})|^2}{a_{dg}(u_h-u_H,u_h-u_H)},$$

where

$$\widehat{(u_h - u_H)}|_{K_i} = (u_h - u_H)|_{K_i} - \frac{1}{|K_i|}\int_{K_i}(u_h - u_H).$$

Since e_i satisfies the equation

$$\frac{\partial}{\partial t}e_i - \nabla\cdot(\kappa\nabla e_i) = 0 \text{ in } K_i \times (T_{n-1}, T_n),$$

we have

$$\begin{aligned}\int_{K_i}\frac{\partial e_i}{\partial t}w &= -\int_{K_i}\kappa\nabla e_i\nabla w + \int_{\partial K_i}(\kappa\nabla\tilde{u}_h\cdot n)w\\ \leq &(\int_{K_i}\kappa|\nabla e_i|^2)^{\frac{1}{2}}(\int_{K_i}\kappa|\nabla w|^2)^{\frac{1}{2}} + (H\int_{\partial K_i}\kappa(\nabla\tilde{u}_h\cdot n)^2)^{\frac{1}{2}}(\frac{1}{H}\int_{\partial K_i}\kappa w^2)^{\frac{1}{2}}.\end{aligned}$$

Moreover, we have

$$\frac{1}{H}\int_{\partial K_i}w^2 \leq C\int_{K_i}\kappa|\nabla w|^2$$

for all $w \in H^1(K_i)$ with $\int_{K_i} w = 0$.

Therefore,

$$\int_{K_i}\frac{\partial e_i}{\partial t}\widehat{(u_h - u_H)} \leq C\big(\int_{K_i}\kappa|\nabla e_i|^2 + H\int_{\partial K_i}(\kappa\nabla\tilde{u}_h\cdot n)^2\big)^{\frac{1}{2}}\big(\int_{K_i}\kappa|\nabla\widehat{(u_h - u_H)}|^2\big)^{\frac{1}{2}},$$

and

$$\begin{aligned}\int_{T_{n-1}}^{T_n}\frac{\sum_i|\int_{K_i}\frac{\partial e_i}{\partial t}\widehat{(u_h - u_H)}|^2}{a_{dg}(u_h - u_H, u_h - u_H)} &\leq C\big(\int_{T_{n-1}}^{T_n}\int_\Omega\kappa|\nabla e_i|^2 + \sum_e H\int_{T_{n-1}}^{T_n}\int_E(\kappa\nabla\tilde{u}_h\cdot n)^2\big)\\ &\leq C\sum_i(\frac{4}{3\lambda_{L_i+1}^{K_i} - 1} + H)\|\kappa\nabla\tilde{u}_h\cdot n\|^2_{L^2(\partial K_i)}.\end{aligned}$$

□

VI. CONCLUSION

In this dissertation, the model reduction methods for two types of partial differential equations, nonlinear equations and parabolic equations, in high-contrast heterogeneous media are investigated. The work is partially motivated by numerical homogenization, and proceeds within the framework of GMsFEM which is introduced in Section II.

In Section III, we develop a multiscale model reduction method for nonlinear problems within GMsFEM. Making use of the nonlinear harmonic extension (p-harmonic extension, for example), we recast the problem on the boundaries of coarse elements and only use the degrees of freedom defined on the boundaries. This will reduce the dimensions of the model compared to other approaches that use basis in the whole domain. In our approach, we construct the local snapshot space and propose a local nonlinear spectral decomposition, which select dominant modes in these nonlinear snapshot spaces. Via these local solutions, we can capture the effects of small separable scales. The convergence of the multiscale solution is studied and we are able to relate the convergence rate to the coarse cell size and the eigenvalues. We also design a numerical implementation of our proposed method by applying Broyden's method to solve the derived optimization problem.

In Section IV and Section V, we study space-time heterogeneous parabolic equations. We focus on the construction of the space-time methods to solve these equations. We propose two different approaches within GMsFEM framework and GMsDGM framework, respectively. The main ingredients of our approaches are (1) the construction of space-time snapshot vectors, (2) the local spectral decomposition in the snapshot space. To perform local spectral decomposition, we discuss a couple of choices for local eigenvalue problems motivated by the analysis. We present convergence analysis of the proposed methods. In particular, we consider examples where the space-time permeability fields have high

contrast and these high-conductivity regions move in the space. If only spatial multiscale basis functions are used, it will require a large dimensional space. Thanks to the space-time multiscale space, we can approximate the problem with a fewer degrees of freedom. Our numerical results show that one can obtain accurate solutions. The proposed concepts can be used for other applications, where one needs space-time multiscale basis functions.

We also discuss online procedures in space-time GMsFEM approach in Section IV, where new multiscale basis functions are constructed using the residual. These basis functions are computed in each local space-time domain. Using online basis functions adaptively, one can reduce the error substantially at a cost of online computations. We will investigate the convergence rate of the online procedures in our future work.

REFERENCES

[1] A. Abdulle and M. Huber. Finite element heterogeneous multiscale method for nonlinear monotone parabolic homogenization problems. Technical report, 2014.

[2] A. Abdulle and G. Vilmart. Analysis of the finite element heterogeneous multiscale method for quasilinear elliptic homogenization problems. *Mathematics of Computation*, 83(286):513–536, 2014.

[3] G. Allaire. Homogenization and two-scale convergence. *SIAM Journal on Mathematical Analysis*, 23(6):1482–1518, 1992.

[4] G. Allaire and R. Brizzi. A multiscale finite element method for numerical homogenization. *SIAM J. Multiscale Modeling and Simulation*, 4(3):790–812, 2005.

[5] T. Arbogast. Analysis of a two-scale, locally conservative subgrid upscaling for elliptic problems. *SIAM J. Numer. Anal.*, 42(2):576–598 (electronic), 2004.

[6] M. Barrault, Y. Maday, N. Nguyen, and A. Patera. An empirical interpolationmethod: application to efficient reduced-basis discretization of partial differential equations. *Comptes Rendus Mathematique*, 339(9):667–672, 2004.

[7] Y. Bazilevs, V. Calo, J. Cottrell, T. Hughes, A. Reali, and G. Scovazzi. Variational multiscale residual-based turbulence modeling for large eddy simulation of incompressible flows. *Computer Methods in Applied Mechanics and Engineering*, 197(1):173–201, 2007.

[8] L. Berlyand, Y. Gorb, and A. Novikov. Discrete network approximation for highly-packed composites with irregular geometry in three dimensions. In *Multiscale methods in science and engineering*, pages 21–57. Springer, 2005.

[9] L. Berlyand, A. Kolpakov, and A. Novikov. *Introduction to the network approximation method for materials modeling*. Number 148. Cambridge University Press, 2013.

[10] L. Berlyand and A. Novikov. Error of the network approximation for densely packed composites with irregular geometry. *SIAM journal on mathematical analysis*, 34(2):385–408, 2002.

[11] L. Berlyand and H. Owhadi. Flux norm approach to finite dimensional homogenization approximations with non-separated scales and high contrast. *Archive for rational mechanics and analysis*, 198(2):677–721, 2010.

[12] V. Calo, Y. Efendiev, J. Galvis, and G. Li. Randomized oversampling for generalized multiscale finite element methods. *arXiv preprint arXiv:1409.7114*, 2014.

[13] V. ChiadòPiat and A. Defranceschi. Homogenization of monotone operators. *Nonlinear Analysis: Theory, Methods & Applications*, 14(9):717–732, 1990.

[14] E. Chung, B. Cockburn, and G. Fu. The staggered dg method is the limit of a hybridizable dg method. *SIAM Journal on Numerical Analysis*, 52(2):915–932, 2014.

[15] E. Chung, Y. Efendiev, and W. Leung. Residual-driven online generalized multiscale finite element methods. *To appear in J. Comput. Phys.*

[16] E. Chung, Y. Efendiev, and W. Leung. An online generalized multiscale discontinuous Galerkin method (GMsDGM) for flows in heterogeneous media. *arXiv preprint arXiv:1504.04417*, 2015.

[17] E. Chung, Y. Efendiev, and W. Leung. Residual-driven online generalized multiscale finite element methods. *J. Comput. Phys.*, 302(C):176–190, December 2015.

[18] E. Chung, Y. Efendiev, W. Leung, and S. Ye. Generalized multiscale finite element methods for space-time heterogeneous parabolic equations. *Submitted*, 2016.

[19] E. Chung, Y. Efendiev, K. Shi, and S. Ye. A multiscale model reduction method for nonlinear monotone elliptic equations in heterogeneous media. *Submitted*, 2016.

[20] E.T. Chung, Y. Efendiev, and W.T. Leung. An adaptive generalized multiscale discontinuous galerkin method (gmsdgm) for high-contrast flow problems. *Submitted to SIAM MMS*.

[21] E.T. Chung, Y. Efendiev, and W.T. Leung. Generalized multiscale finite element method for wave propagation in heterogeneous media. *to appear in SIAM MMS*.

[22] P. Ciarlet. *The finite element method for elliptic problems*, volume 40. Siam, 2002.

[23] B. Cockburn, J. Gopalakrishnan, and R. Lazarov. Unified hybridization of discontinuous galerkin, mixed, and continuous galerkin methods for second order elliptic problems. *SIAM Journal on Numerical Analysis*, 47(2):1319–1365, 2009.

[24] L.J. Durlofsky. Numerical calculation of equivalent grid block permeability tensors for heterogeneous porous media. *Water Resour. Res.*, 27:699–708, 1991.

[25] W. E and B. Engquist. Heterogeneous multiscale methods. *Comm. Math. Sci.*, 1(1):87–132, 2003.

[26] Y. Efendiev, J. Galvis, and T. Hou. Generalized multiscale finite element methods. *Journal of Computational Physics*, 251:116–135, 2013.

[27] Y. Efendiev, J. Galvis, R. Lazarov, M. Moon, and M. Sarkis. Generalized multiscale finite element method. Symmetric interior penalty coupling. *Journal of Computational Physics*, 255:1–15, 2013.

[28] Y. Efendiev, J. Galvis, G. Li, and M. Presho. Generalized multiscale finite element methods. Oversampling strategies. *International Journal for Multiscale Computational Engineering, accepted*, 2013.

[29] Y. Efendiev, J. Galvis, M. Presho, and J. Zhou. A multiscale enrichment procedure for nonlinear monotone operators. *ESAIM: Mathematical Modelling and Numerical Analysis*, 48(02):475–491, 2014.

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

[31] Y. Efendiev and T. Hou. *Multiscale Finite Element Methods: Theory and Applications*. Springer, 2009.

[32] Y. Efendiev, R. Lazarov, M. Moon, and K. Shi. A spectral multiscale hybridizable discontinuous Galerkin method for second order elliptic problems. *Computer Methods in Applied Mechanics and Engineering*, 292:243–256, 2015.

[33] Y. Efendiev, R. Lazarov, and K. Shi. A multiscale hdg method for second order elliptic equations. part i. polynomial and homogenization-based multiscale spaces. *SIAM Journal on Numerical Analysis*, 53(1):342–369, 2015.

[34] Y. Efendiev and A. Pankov. Numerical homogenization and correctors for nonlinear elliptic equations. *SIAM J. Appl. Math.*, 65(1):43–68, 2004.

[35] Y. Efendiev and A. Pankov. Numerical homogenization of nonlinear random parabolic operators. *Multiscale Modeling & Simulation*, 2(2):237–268, 2004.

[36] Y. Efendiev and A. Pankov. Homogenization of nonlinear random parabolic operators. *Advances in Differential Equations*, 10(11):1235–1260, 2005.

[37] J. Fish and W. Chen. Space–time multiscale model for wave propagation in heterogeneous media. *Computer Methods in applied mechanics and engineering*, 193(45):4837–4856, 2004.

[38] R. Glowinski and A. Marroco. Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de dirichlet non linéaires. *Revue française d'automatique, informatique, recherche opérationnelle. Analyse numérique*, 9(2):41–76, 1975.

[39] P. Henning. Heterogeneous multiscale finite element methods for advection-diffusion and nonlinear elliptic multiscale problems. *Münster: Univ. Münster, Mathematisch-Naturwissenschaftliche Fakultät, Fachbereich Mathematik und Informatik (Diss.). ii*, page 63, 2011.

[40] P. Henning and M. Ohlberger. Error control and adaptivity for heterogeneous multiscale approximations of nonlinear monotone problems. *Discrete and Continuous Dynamical Systems-Serie S. Special Issue on Numerical Methods based on Homogenization and Two-Scale Convergence*, 2015.

[41] T. Hou and X.H. Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *J. Comput. Phys.*, 134:169–189, 1997.

[42] T. Hughes. Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods. *Computer methods in applied mechanics and engineering*, 127(1):387–401, 1995.

[43] T. Hughes, G. Feijoo, L. Mazzei, and J. Quincy. The variational multiscale method—a paradigm for computational mechanics. *Comput. Methods Appl. Mech. Engrg.*, 166:3–24, 1998.

[44] T. Hughes and G. Sangalli. Variational multiscale analysis: the fine-scale Green's function, projection, optimization, localization, and stabilized methods. *SIAM Journal on Numerical Analysis*, 45(2):539–557, 2007.

[45] T. Hughes and J. Stewart. A space-time formulation for multiscale phenomena. *Journal of Computational and Applied Mathematics*, 74(1):217–229, 1996.

[46] V. Jikov, S. Kozlov, and O. Oleinik. *Homogenization of differential operators and integral functionals*. Springer Science & Business Media, 2012.

[47] R. Künze and I. Lunati. An adaptive multiscale method for density-driven instabilities. *Journal of Computational Physics*, 231(17):5557–5570, 2012.

[48] J. Lions, D. Lukkassen, L. Persson, and P. Wall. Reiterated homogenization of nonlinear monotone operators. *Chinese Annals of Mathematics*, 22(01):1–12, 2001.

[49] A. Masud and T. Hughes. A space-time galerkin/least-squares finite element formulation of the navier-stokes equations for moving domain problems. *Computer Methods in Applied Mechanics and Engineering*, 146(1):91–126, 1997.

[50] P. Ming and P. Zhang. Analysis of the heterogeneous multiscale method for elliptic homogenization problems. *Journal of the American Mathematical Society*, 18(1):121–156, 2005.

[51] P. Ming and P. Zhang. Analysis of the heterogeneous multiscale method for parabolic homogenization problems. *Mathematics of Computation*, 76(257):153–177, 2007.

[52] H. Nguyen and J. Reynen. A space-time least-square finite element scheme for advection-diffusion equations. *Computer Methods in Applied Mechanics and Engineering*, 42(3):331–342, 1984.

[53] H. Owhadi and L. Zhang. Homogenization of parabolic equations with a continuum of space and time scales. *SIAM Journal on Numerical Analysis*, 46(1):1–36, 2007.

[54] H. Owhadi, L. Zhang, and L. Berlyand. Polyharmonic homogenization, rough polyharmonic splines and sparse super-localization. *ESAIM: Mathematical Modelling and Numerical Analysis*, 48(02):517–552, 2014.

[55] A. Pankov. *G-convergence and homogenization of nonlinear partial differential operators*, volume 422. Springer Science & Business Media, 2013.

[56] G. Papanicolau, A. Bensoussan, and J. Lions. *Asymptotic analysis for periodic structures*. Elsevier, 1978.

[57] G. Pavliotis and A. Stuart. *Multiscale methods: averaging and homogenization*. Springer Science & Business Media, 2008.

[58] M. Presho and S. Ye. Reduced-order multiscale modeling of nonlinear p-Laplacian flows in high-contrast media. *Computational Geosciences*, 19(4):921–932, 2015.

[59] S. Srinivasan, R. Lazarov, and P. Minev. Multiscale direction-splitting algorithms for parabolic equations with highly heterogeneous coefficients. *Computers & Mathematics with Applications*, 72(6):1641–1654, 2016.

[60] K. Takizawa and T. Tezduyar. Multiscale space–time fluid–structure interaction techniques. *Computational Mechanics*, 48(3):247–267, 2011.

[61] T. Tezduyar, M. Behr, S. Mittal, and A. Johnson. Computation of unsteady incompressible flows with the stabilized finite element methods: space-time formulations, iterative strategies and massively parallel implementations. *Asme Pressure Vessels Piping Div Publ PVP*, 246:7–24, 1992.

[62] C. Timofte. Multiscale analysis of diffusion processes in composite media. *Computers & Mathematics with Applications*, 66(9):1573–1580, 2013.

[63] X.H. Wu, Y. Efendiev, and T.Y. Hou. Analysis of upscaling absolute permeability. *Discrete and Continuous Dynamical Systems, Series B.*, 2:158–204, 2002.

[64] E. Zeidler. *Nonlinear Functional Analysis and Its Applications: III: Variational Methods and Optimization*. Springer Science & Business Media, 2013.