ADAPTIVITY AND ONLINE BASIS CONSTRUCTION FOR GENERALIZED MULTISCALE FINITE ELEMENT METHODS

A Dissertation

by

WING TAT LEUNG

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
Co-Chair of Committee,	Eric Chung
Committee Members,	Eduardo Gildin
	Peter Howard
	Raytcho Lazarov
Head of Department,	Emil Straube

August 2017

Major Subject: Mathematics

ABSTRACT

Many problems in application involve media with multiple scale, for example, in composite materials, porous media. These problems are usually computationally challenging since fine grid computation is extremely expensive. Therefore, one may need to develop a coarse grid model reduction for this type of problems. In this dissertation, we will consider a multiscale method called generalized multiscale finite element method (GMsFEM).

GMsFEM follows the framework of multiscale finite element method. Instead of using one basis function per coarse grid node, GMsFEM uses several basis functions for one coarse grid node. Since the media is highly heterogeneous and may involves high contrast, having more than one basis function per node is important to reduce the error significantly.

Due to the varying heterogeneity in the domain, we may require different numbers of basis functions in different regions. Then the question is how to determine the number of basis functions in each region. In this dissertation, we will discuss an adaptive enrichment algorithm for enriching basis functions for the regions with large error. We will consider two different types of basis function for enrichment. One is using the pre-computed offline basis functions. We call this method offline adaptive enrichment. The other method uses online constructed basis functions called online adaptive enrichment.

In applications, non-conforming basis functions can give us more flexibility on gridding. The discontinuous Galerkin method also makes the mass matrix block diagonal, which enhances the computation speed in solving time-dependent problem with an explicit scheme. In this dissertation, we will discuss offline and online adaptive methods for the generalized multiscale discontinuous Galerkin method (GMsDGM). We will also discuss using GMsDGM for simulating wave propagation in heterogeneous media.

ACKNOWLEDGMENTS

I would like to thank the Department of Mathematics, Texas A&M University. It will be difficult for me to purse my doctorate degree without the support of the department. The Department of Mathematics gives me many opportunities to attend different seminars, conferences and workshops. It helps me to explore future development of my research. It also allow me to make connection which potentially lead to interesting collaborative work.

Next, I would like to thank Dr. Yalchin Efendiev, my supervisor, for his guidance and advices in both academic and personal life. He gives me many ideas in my research and always be helpful when I am facing any problems. I am also thankful to Dr. Eric T. Chung. We have lots of discussions in my research topic. He always gives many useful suggestions.

Besides, I would like to thank Dr. Eduardo Gildin, Dr. Peter Howard and Dr. Raytcho Lazarov for serving as my committee and giving suggestions for my dissertation. I would also like to thank my colleagues, especially to Dr. Victor M. Calo, Dr. Richard L Gibson Jr and Dr. Maria Vasilyeva.

Last but not least, I want to show my gratitude to my parents and my brother for their support. Without their support, I will not be able to focus on my study.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

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

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

Funding Sources

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

NOMENCLATURE

MsFEM	Multiscale Finite Element Method
GMsFEM	Generalized Multiscale Finite Element Method
GMsDGM	Generalized Multiscale Discontinuous Galerkin Method
Ω	Computational domain
$\mathcal{T}^h$	A partition of Ω into fine finite element
$\mathcal{T}^H$	A partition of Ω into coarse finite element
h	The fine mesh size
H	The coarse mesh size
x_i	A coarse grid vertex
ω_i	a subdomain with a common vertex at x_i
ω_i^+	a oversampled subdomain for ω_i
K_i	a coarse grid element
K_i^+	a oversampled subdomain for K_i
u_h	The fine-grid solution
u_H	The coarse-grid solution
V_h	The fine-scale space
N_c	The number of coarse-grid nodes
N	The number of coarse-grid elements
χ_i	The partition of unity
$\psi_k^{(i)}$	The snapshot multiscale basis functions
$\phi_k^{(i)}$	The offline multiscale basis functions

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION

1.1 Motivation

In real life applications, many multiscale problems are involved in different areas (for example, in engineering, mathematics, physics, chemistry, computational biology and computer science problems). One can use an extremely fine mesh to solve these multiscale problems, however, it is computationally experience. Therefore, some types of model reduction technique is required for solving these problems efficiently. There are many different mutliscale model reduction methods. They are basically fall into two types of approach. One is called global model reduction [1, 2, 3]. The other is called local model reduction techniques [4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 3, 19, 20, 21, 22, 23].

Global model reduction techniques basically use global basis functions or global information to construct a reduced global model to approximate the original model. Since these methods using some global information in the offline stage, it may require a huge amount of computational power and also lack local adaptivity. In this dissertation, we will mostly focus on local model reduction techniques. In order to capture the global information of the model, we will present an online basis construction technique which uses the residual information obtained from solving the global equation. We remark that these online basis functions are constructed locally, therefore, it is time effective and compute these online basis function in parallel.

There are many different local multiscale model reduction techniques developed previously. They solve some underlying fine-scale problems on a coarse grid. These approaches can be classified into two categories. One is upscaling type approach. The other is multiscale method type approach. Upscaling approaches [4, 5] basically involve computation of some effective mediums in the coarse grid level which give an approximate model. Mul-

tiscale methods [7, 9, 16, 17, 18, 3, 19, 20, 21, 22] basically involve construction of some local basis functions in coarse grid level. These basis functions are constructed to capture the local multiscale information of the problem. We will briefly discuss some examples of upscaling and multiscale approach in Chapter 1 later.

There are many papers [24, 25, 26] discussed on using few local multiscale basis functions to approximate the solution. In this dissertation, we will mainly focus on a recent developed local model reduction method called Generalized Multiscale Finite Element method (GMsFEM) [27, 20, 22, 28, 29]. GMsFEM can be viewed as a generalization of the Multiscale Finite Element Method (MsFEM) ([30]). It gives a systematic way to enrich the basis functions in the coarse grid space. Enriching these basis functions are essential to reduce the error significantly. GMsFEM involves two stages of computation, which are offline stage and online stage. We will construct a small dimensional multiscale finite element space in the offline stage and use it to compute an approximate solution with a given source term and boundary condition in the online stage. We remark that one can use the multiscale finite element space for solving the problem with different source term and boundary condition without recomputing the basis functions. The most important part of the construction of offline spaces is the choice of local spectral problems and the choice of the snapshot space. We will discuss it in Chapter 2 in detail.

In practice, one may need to use different number of basis functions in different coarse elements (coarse neighborhoods) to obtain an accurate representation of the solution. We can use a small amount of basis functions in the regions with less heterogeneous property. In contrast, we will need to use more basis functions in the regions with more heterogeneities and high contrast. Therefore, we will need an adaptive enrichment algorithm to determine how many basis functions are required in each coarse region. In [31], the authors proposed a residual based local error estimator for adaptive enrichment of GMsFEM. In this dissertation, we will discuss on using residual based local error estimator to preform

basis enrichment for generalized multiscale discontinuous Galerkin method (GMsDGM). We will also use this residual based local error estimator to preform our online basis construction algorithm for GMsFEM and GMsDGM. There are many existing research papers [32, 33, 31] discussing on using local error estimator to preform mesh refinement or basis enrichment. Our analysis follows the same idea and do not consider the error coming from the fine grid discretization. We will discuss the adaptivity enrichment method for GMsDGM in detail in Chapter 3. We remark that there are many related activities in designing a-posteriori error estimates [34, 35, 36, 37, 38, 39] for global reduced models. The main difference is that our error estimators are based on special local eigenvalue problem and use the eigenstructure of the offline space. We also discuss an approach that adaptively selects multiscale basis functions from the offline space by selecting a basis with the most correlation to the local residual (cf. [40]). Adaptivity is important for local multiscale methods as it identifies regions with large errors. However, after adding some initial basis functions, one needs to take into account some global information as the distant effects can be important. In this dissertation, we discuss the development of online basis functions for both GMsFEM and GMsDGM that substantially accelerate the convergence of the method. The online basis functions are constructed based on a residual and motivated by the analysis. We will discuss the online basis construction in the adaptivity enrichment method for GMsFEM in Chapter 2 and GMsDGM in Chapter 4.

GMsDGM has many applications in different areas, for example, flow problem, wave simulation problem. In Chapter 5, we will discuss on using GMsDGM to simulate wave propagation in heterogeneous media. The discontinuous Galerkin framework give us a block diagonal mass matrix and more flexibility on gridding process which improve the computational performance and handle complex geometry of the computational domain.

1.2 Introduction of upscaling method and multiscale finite element method

In this section, we will briefly discuss about homogenization, numerical homogenization and multiscale finite element method for the flow problem with periodic media parameter, that is, solving u_ϵ satisfying

$$-\nabla \cdot (\kappa(\frac{x}{\epsilon})\nabla u_\epsilon) = f(x), \quad x \in \Omega$$

with periodic parameter $\kappa(\frac{x}{\epsilon})$ in the computational domain $\Omega = [0,1]^d \subset \mathbb{R}^d$, where ϵ is assumed to be extremely small.

1.2.1 Homogenization

In this subsection, we will introduce the homogenization technique for the flow problem. The basic idea of homogenization is to approximate the solution u_ϵ by a solution u_0 of the flow problem with a constant coefficient κ^*, which is,

$$-\nabla \cdot (\kappa^* \nabla u_0) = f(x), \quad x \in \Omega.$$

Next, the question is how to find out the constant coefficient κ^*. We assume u_ϵ can written as a asymptotic expansion which is

$$u_\epsilon = u_0(x,y) + \epsilon u_1(x,y) + \epsilon^2 u_2(x,y) + \cdots$$

where $y = \frac{x}{\epsilon}$ is fast variable and the functions $u_i(x,y)$ are periodic in y with period T. It is easy to see $\nabla = \nabla_x + \epsilon^{-1}\nabla_y$. By comparing the coefficient of ϵ in $O(\epsilon^{-2})$ order, we have,

$$-\nabla_y \cdot \Big(\kappa(y)\nabla_y u_0(x,y)\Big) = 0.$$

By solving the above equation with periodic boundary condition in y, we have u_0 is independent of y, that is, $u_0(x,y) = u_0(x)$. Next, by comparing the coefficient of ϵ in $O(\epsilon^{-1})$ order, we have

$$-\nabla_y \cdot \Big(\kappa(y)\nabla_y u_1(x,y)\Big) = \nabla_y \cdot \Big(\kappa(y)\nabla_x u_0(x)\Big) = \Big(\nabla_y \kappa(y)\Big) \cdot \Big(\nabla_x u_0(x)\Big) \tag{1.1}$$

We consider $\mathcal{N}_j = \mathcal{N}_j(y)$ to be the solution to the following "cell problem":

$$-\nabla_y \cdot \Big(\kappa(y)\nabla_y \mathcal{N}_j\Big) = \frac{\partial}{\partial y_j}\kappa(y),$$

with periodic boundary condition in y. Then, the solution of (1.1) can be written in form of

$$u_1(x,y) = \mathcal{N}_j \frac{\partial}{\partial x_j} u + \tilde{u}_1(x).$$

By comparing the coefficient of ϵ in $O(1)$ order, we have

$$-\nabla_x \cdot \Big(\kappa(y)(\nabla_x u_0(x) + \nabla_y u_1(x,y))\Big) - \nabla_y \cdot \Big(\kappa(y)(\nabla_x u_1(x,y) + \nabla_y u_2(x,y))\Big) = f(x) \tag{1.2}$$

Finally, integrating (1.2) over periodic cell Y, we have the homogenized equation is given as

$$-\nabla_x \cdot \Big(\kappa^* \nabla_x u_0(x)\Big) = f$$

where

$$\kappa^*_{ij} = \frac{1}{|Y|}\int_Y \kappa(\delta_{ij} + \frac{\partial}{\partial y_j}\mathcal{N}_i)dy.$$

We remark that even if the parameter coefficient $\kappa(y)$ is a scalar coefficient, the homogenized coefficient κ^* can be a symmetric tensor.

1.2.2 Numerical Homogenization

In this subsection, we will briefly introduce numerical homogenization technique. Similar to homogenization, we would like to compute the homogenized coefficient for the problem. However, it is not easily to determine the periodic cell Y in practice. Sometimes, the parameter coefficient is even not periodic. Therefore, we may need to fix the coarse regions of the "cell problem" numerically. In numerical homogenization approach, we will first partition the computational domain into a coarse grid (see 1.1). Using this

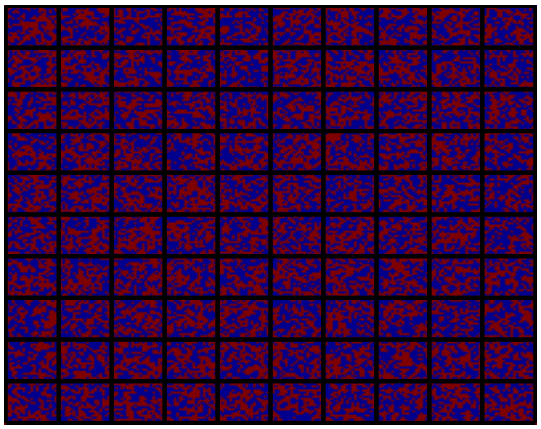

Figure 1.1: Illustration of the coarse grid for numerical homogenization.

coarse grid, we will compute the "cell problem" in each coarse element K where the "cell problem" is defined as: finding $\mathcal{N}_j^{nh}$ with

$$-\nabla(\kappa(\frac{x}{\epsilon})\nabla\mathcal{N}_j^{nh}) = 0 \;\; x \in K$$

subjecting to the boundary condition $\mathcal{N}_j^{nh} = x_j$ on ∂K.

Then the numerical homogenized coefficient $\kappa_{ij}^{*,\text{nh}}$ is defined by

$$\kappa_{ij}^{*,\text{nh}} = \frac{1}{|K|}\int_K \kappa \nabla \mathcal{N}_i^{nh}.$$

We remark that the numerical homogenized coefficient $\kappa_{ij}^{*,\text{nh}}$ is a symmetric tensor.

In fact, the "cell problem" in numerical homogenization is similar to the "cell problem" in homogenization case in the sense of $\mathcal{N}_j^{nh} - x_j$ satisfying

$$-\nabla_y \cdot \Big(\kappa(y)\nabla_y(\mathcal{N}_j^{nh} - x_j)\Big) = \frac{\partial}{\partial y_j}\kappa(y),$$

with zero Dirichlet boundary condition and

$$\kappa_{ij}^{*,\text{nh}} = \frac{1}{|K|}\int_K \kappa(\delta_{ij} + \frac{\partial}{\partial y_j}(\mathcal{N}_i^{nh} - x_i))dy.$$

That is, instead of solving the local problem with periodic boundary condition in periodic cell Y, we solve the local problem with boundary condition (can be Dirichlet, Mixed Dirichlet-Neumann or even periodic) in each coarse element K.

1.2.3 Multiscale finite element method (MsFEM)

In this subsection, we briefly introduce multiscale finite element method. In contrast to numerical homogenization, multiscale finite element method constructs the basis functions instead of the homogenized coefficient. Multiscale finite element method follows the framework of the standard finite element method, for example, it considers the variational

problem for flow problem: find $u \in H^1(\Omega)$ satistying

$$\int_\Omega \kappa \nabla u \cdot \nabla v = \int_\Omega f v \; \forall v \in H^1(\Omega).$$

Multicale finite element method will approximate the variational solution $u \in H^1(\Omega)$ by a solution u_{ms} satisfying the variational equation in a finite dimensional multsicale subspace $V_{ms} \subset H^1$, that is, finding $u_{ms} \in V_{ms}$ such that

$$\int_\Omega \kappa \nabla u_{ms} \cdot \nabla v = \int_\Omega f v \; \forall v \in V_{ms}. \tag{1.3}$$

The main idea of multiscale finite element method is constructing a multiscale basis function per coarse grid node which can capture the multiscale feature of the solution u.

We consider the computational domain Ω is partition into a coarse grid $\mathcal{T}^H$ which is similar to the coarse grid in numerical homogenization. For each coarse vertices x_i, the coarse neighborhood ω_i is defined by

$$\omega_i = \bigcup \{K_j \in \mathcal{T}^H; \quad x_i \in \overline{K}_j\}, \tag{1.4}$$

that is, ω_i is the union of the coarse elements $K \in \mathcal{T}^H$ containing the coarse grid node x_i. In Figure 1.2, we show an illustration of the coarse mesh and coarse neighborhood.

To construct the multiscale basis functions $\phi^{(i)}$, we will solve a local problem in each coarse neighborhood ω_i which is:

$$\begin{aligned} -\nabla \cdot \left(\kappa(x) \nabla \phi^{(i)}\right) &= 0 \; \forall K \in \mathcal{T}^H \\ \phi^{(i)}(x_j) &= \delta_{i,j} \;\; \text{for all coarse node } x_j, \end{aligned} \tag{1.5}$$

and $\phi^{(i)}$ is piecewise linear in each coarse edge, where $\delta_{i,j}$ is the Kronecker delta function.

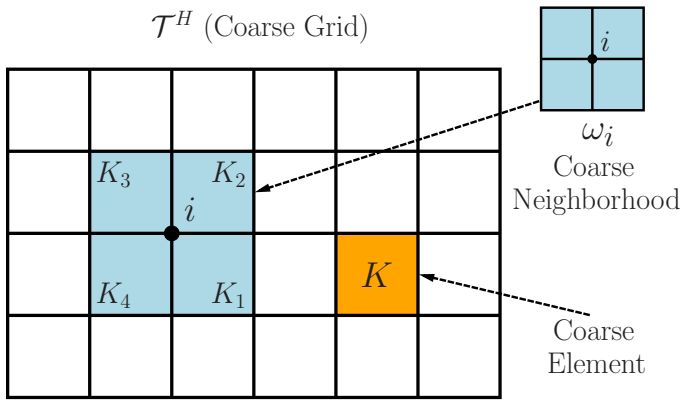

Figure 1.2: Illustration of a coarse neighborhood and a coarse element.

The multiscale finite element space is defined as the span of these local basis functions, namely,

$$V_{ms} = \text{span}\{\phi^{(i)} |\ \ 1 \leq i \leq N_c\}$$

where N_c is the number of coarse-grid nodes. After constructing the multiscale finite element space, one can use this space repeatedly for solving (1.3) with different source terms and boundary conditions.

The methods introduced in this section can basically handle the problems with scale separation quite well. However, in some applications, one may need to handle some more complicated cases, for example, high contrast media with multiple inclusions and channels. In these cases, we will require a more complex coarse grid upscaling model or enriching the multiscale finite element space to capture the complex multiscale feature of the solution. In the next chapter, we will discuss on the GMsFEM which is a generalization of the multiscale finite element method for handling the flow problem in high contrast medium.

2. RESIDUAL-BASED ONLINE GENERALIZED MULTISCALE FINITE ELEMENT METHODS

2.1 GMsFEM for high contrast flow

2.1.1 Overview

In this section, we will introduce the GMsFEM ([27, 20, 22, 28, 29]) for flow problem. Let Ω be the computational domain. We are considering the flow problem in highly heterogeneous media defined as

$$-\nabla \cdot \big(\kappa \nabla u\big) = f \quad \text{in} \quad \Omega, \tag{2.1}$$

$$u = 0 \quad \text{on} \quad \partial\Omega \tag{2.2}$$

where κ is the permeability coefficient in $L^\infty(\Omega)$, f is a source function in $L^2(\Omega)$. We assume the coefficient κ is highly heterogeneous, which can involve multiple non-separable scale and high contrast.

The GMsFEM starts with partitioning the computational domain Ω into a coarse mesh $\mathcal{T}^H$ with mesh size H. The coarse mesh is then refined into a fine mesh $\mathcal{T}^h$ with fine mesh size $h << H$. We assume this fine mesh is fine enough to restore the multiscale properties of the problem. To simplify the discussion, we will consider the coarse mesh and fine mesh are both uniform rectangular mesh. Let N_c be the number of nodal points of the coarse mesh. We consider the set $N_x = \{x_i \,|1 \leq i \leq N_c\}$ containing all of the coarse

Reprinted with permission from "Residual-driven online generalized multiscale finite element methods" by Eric T Chung, Yalchin Efendiev and Wing Tat Leung, 2015. Journal of Computational Physics, 302, 176-190, Copyright [2017] by Elsevier.

grid nodes. For each coarse grid node x_i, the coarse neighborhood ω_i is defined by

$$\omega_i = \bigcup \{K_j \in \mathcal{T}^H; \quad x_i \in \overline{K}_j\}, \tag{2.3}$$

that is, ω_i is the union of the coarse elements $K \in \mathcal{T}^H$ containing the coarse grid node x_i. In Figure 1.2, we show an illustration of the coarse mesh and coarse neighborhood.

Next, we denote the conforming piecewise bilinear fine-scale finite element space as V_h, that is $V_h = \{v \in H^1(\Omega) |\ v|_K \in Q^1(K),\ \forall K \in \mathcal{T}^H\}$. This finite element space is used to compute some local problems for constructing the multsicale finite element space and the fine-scale reference solution for comparison. The finite-scale reference solution u_h is defined by the following variational problem: find $u_h \in V_h$ such that

$$a(u_h, v) = (f, v), \quad \text{for all } v \in V_h, \tag{2.4}$$

where a is the bilinear form which is defined as $a(u, v) = \int_\Omega \kappa \nabla u \cdot \nabla v dx$. We define the energy norm, $\|v\|_a$, by $\|v\|_a^2 = a(v, v)$. Since the fine mesh is fine enough to resolve the multiscale properties of the problem, we notice that the reference solution is a good approximation of the exact solution. We would like to construct a multiscale finite element space V_{ms} which is a subspace of the the fine-scale finite element space V_h in the coarse mesh with multiscale basis function resolve the local multiscale behaviors of the exact solution.

Next, we will present the general framework of GMsFEM ([27, 20, 22, 28, 29]). It follows the standard continuous Galerkin framework. We seek for a multiscale numerical solution $u_{ms} \in V_{ms}$ such that

$$a(u_{ms}, v) = (f, v), \quad \text{for all } v \in V_{ms} \subset V_h. \tag{2.5}$$

The basis functions of the multiscale finite element space are constructed in coarse neighborhoods, and therefore, they are nodal based and supposed on coarse neighborhoods. That is, $V_{ms} = \text{span}\{\phi_j^{(i)} \in V_h | \text{ supp}\{\phi_j^{(i)}\} \subset \omega_i, \text{ for } 1 \leq i \leq N\}$. The most important part of GMsFEM is how to construct these local basis functions in order to obtain a good approximation of the solution. We will discuss it in detail in the latter subsection. By [27, 20, 22, 28, 29], we will use some basis functions which is called offline basis function to span our initial multiscale finite element space. These basis functions are called offline basis function since they are constructed in the offline stage. These basis functions are not always sufficient to give us a good approximation of the solution. Therefore, we may need to construct some local basis functions in online stage to give a rapid convergence to the reference solution.

2.1.2 Construction of offline basis functions for GMsFEM

In this section, we discuss how to construct the offline basis function in GMsFEM (see e.g. [27, 20, 22, 28, 29]). For each coarse neighbourhood ω, we consider the local fine-scale finite element space $V_h^{(i)}$ to be the restriction of the global finite-scale finite element space V_h, namely, $V_h^{(i)} = \{v| \in H^1(\omega)| \ v = w|_\omega \text{ for some } w \in V_h\}$. Next, we will introduce a subspace of the local fine-scale finite element space which is called the local snapshot space $V_{\text{snap}}^{(i)}$. The local snapshot space is considered to be a space containing all of the important components of the reference solution in the coarse neighbourhood ω. The snapshot space is usually too large that will not be used directly to be our multiscale finite element space. We will use a spectral problem to select some dominant modes in order to obtain a low dimensional subspace, and our multiscale finite element space is consider to be the sum of low dimensional subspace. Usually, there are two choices for the snapshot space. One is using the local fine-scale finite element space directly. Another choice is using the space containing all κ - harmonic extensions.

Next, we discuss the construction of harmonic extension snapshot space $V_{\text{snap}}^{(i)}$. Let $J_h(\omega)$ be the set of all finite grid nodal points on the coarse neighbourhood boundary $\partial\omega$. For each finite grid nodal point $x_j \in J_h(\omega)$, we construct a local snapshot basis function $\psi_{\text{snap},j}$ by solving

$$\begin{aligned} -\text{div}(\kappa(x)\nabla\psi_{\text{snap},j}) &= 0, \quad \text{in } \omega, \\ \psi_{\text{snap},j}(x_k) &= \delta_{k,j}^h, \quad \text{for all } x_k \in J_h(\omega). \end{aligned} \tag{2.6}$$

where δ_j^h is the discrete delta function defined by

$$\delta_{k,j}^h = \begin{cases} 1, & k = j \\ 0, & k \neq j \end{cases}.$$

The local snapshot space $V_{\text{snap}}^{(i)}$ is the space spanned by these snapshot basis functions, namely, $V_{\text{snap}}^{(i)} = \text{span}\{\psi_{\text{snap},j} |\ \forall x_j \in J_h(\omega)\}$. Therefore, the dimension of the local snapshot space is equal the number of finite grid nodes on on the coarse neighbourhood boundary $\partial\omega$. To construct this snapshot space, we need to solve a number of local problems. To reduce the computational cost of the construction, we can use randomized snapshots with oversampleing technique [41]. For each coarse neighbourhood ω, we can define an oversampled domain ω^+ by extending several fine grid layers around ω. We will construct K snapshot basis functions $\psi_{\text{snap},j}^+$ by solving the κ - harmonic extension with random boundary conditions, that is, solving

$$\begin{aligned} -\text{div}(\kappa(x)\nabla\psi_{\text{snap},j}^+) &= 0, \quad \text{in } \omega^+, \\ \psi_{\text{snap},j}^+ &= r_j, \quad \text{on } \partial\omega^+, . \end{aligned} \tag{2.7}$$

where r_j are independent identically distributed (i.i.d.) Gaussian random vectors on the

fine-grid nodes on $\partial\omega^+$. The local snapshot space $V_{\text{snap}}^{(i)}$ is the space spanned by these snapshot basis functions restricted in ω, namely, $V_{\text{snap}}^{(i)} = \text{span}\{\psi_{\text{snap},j}^+|_\omega| \text{ for } 1 \leq j \leq K\}$. K can be chosen as the number of offline basis function per coarse neighbourhood plus a small number p_{bf}, for example, $p_{bf} = 4$. Therefore, we only need to solve about $n + 4$ local problems if we need n offline basis functions.

Next, we will discuss about the construction of offline basis functions. We will use a spectral problem to perform a local model reduction. From the analysis in [16], we consider the spectral problem to be: find $(\phi, \lambda) \in V_{\text{snap}}^{(i)} \times \mathbb{R}$ such that

$$\int_\omega \kappa(x)\nabla\phi_j^{(i)} \cdot \nabla v\, dx = \lambda_j^{(i)} \int_\omega \widetilde{\kappa}(x)\psi_j^{(i)} v\, dx, \quad \forall v \in V_{\text{snap}}^{(i)} \tag{2.8}$$

where $\widetilde{\kappa}(x)$ is defined by

$$\widetilde{\kappa} = \kappa \sum_{i=1}^{N_c} H^2 |\nabla\chi_i|^2,$$

with χ_i being the nodal multiscale basis function for the coarse node x_i.

We assume the eigenvalues $\lambda_j^{(i)}$ from (2.8) are in ascending order. For each ω_i, the first l_i eigenfunctions corresponding to the small eigenvalues are chosen to construct the offline space. The offline space V_{off} is spanned by the basis $\chi_i\phi_j^{(i)}$, namely $V_{\text{off}} = \text{span}\{\chi_i\phi_j^{(i)}|1 \leq i \leq N,\ 1 \leq j \leq l_i\}$. After constructing the offline space, we can set the multiscale finite element space to be this offline space and compute the numerical solution by solving (2.5).

2.2 Offline Adaptive GMsFEM

In the previous section, we introduced a multiscale offline space to a coarse-scale approximation of the exact solution. However, using a pre-defined number of basis functions l_i in the coarse neighborhood ω_i may not be enough to restore all of the scale of the exact solution. Therefore, we may need to use an adaptive approach to select the number of basis functions used in each coarse neighborhood. In [31], there is an adaptive enrichment

algorithm for choosing l_i is introduced and analyzed. The general idea of this adaptive enrichment algorithm is using the local residual of the equation to indicate the regions which require more basis functions. We will first compute a numerical solution by using the offline multiscale space and use it to compute the local residual for each coarse neighborhood. Next, for each coarse neighborhood with large residual, we enrich more basis functions by using the next eigenfunctions in (2.8), namely adding $\chi_i\phi_{l_i+1}^{(i)}, \ldots, \chi_i\phi_{l_i+p}^{(i)}$ to finite element space. We repeat the basis enrichment process until the residual is smaller than a tolerance. We call this algorithm the offline adaptive GMsFEM since we only add the predefined offline basis functions to multiscale finite element space.

Next, we will discuss the offline adaptive GMsFEM in more details. We assume $u_{\text{ms}} \in V_{\text{off}}$ to be the solution obtained in (2.5). For a given coarse neighborhood, we can define the local residual functional, $R_i : V_h^{(i)} \to \mathbb{R}$, by

$$R_i(v) := \int_{\omega_i} fv - \int_{\omega_i} a\nabla u_{\text{ms}} \cdot \nabla v$$

The operator norm of the residual functional is defined by $\|R_i\| := \sup \dfrac{|R_i(v)|}{\|v\|_V}$. Follow the result in [31], we have a posteriori error estimate for GMsFEM which is

$$\|u_h - u_{\text{ms}}\|_V^2 \leq C_{\text{err}} \sum_{i=1}^{N_c} \|R_i\|^2 (\lambda_{l_i+1}^{(i)})^{-1}, \tag{2.9}$$

where C_{err} is a uniform constant, and $\lambda_{l_i+1}^{(i)}$ is the (l_i+1)-th eigenvalue for the spectral problem (2.8) which corresponds to the first eigenfunction that do not used in the offline basis construction.

Next, we will present the adaptive enrichment algorithm. The index $m \geq 1$is used to denote the enrichment level. To enrich the multiscale finite element space, we will use the

Dorlfer's bulk marking strategy [34] to select the number of enriching local regions. We first define a constant $0 < theta < 1$ independent of m. At each enrichment level m, we compute the numerical solution u_{ms}^m by computing the corresponding solution obtained in (2.5) with $V_{\text{ms}} = V_{\text{ms}}^m$ where V_{ms}^m is the corresponding multiscale finite element space in m-th level. Next, we compute the local error indicator η_i for each coarse neighborhood ω_i with

$$\eta_i^2 = \|R_i\|^2 (\lambda_{l_i^m+1}^{\omega_i})^{-1}, \qquad i = 1, 2, \cdots, N,$$

where l_i^m is the number of basis functions used in the coarse neighborhood ω_i and $R_i(v)$ is defined using u_{ms}^m, namely,

$$R_i(v) = \int_{\omega_i} fv - \int_{\omega_i} a\nabla u_{\text{ms}}^m \cdot \nabla v, \quad \forall v \in V_i.$$

We will use the error indicator to identify the regions where basis enrichment is required. We first rearrange the order of the coarse neighborhoods such that the error indicator η_i^2 are in decreasing order, that is, $\eta_1^2 \geq \eta_2^2 \geq \cdots \geq \eta_N^2$. Let k be the smallest integer such that

$$\theta \sum_{i=1}^{N_c} \eta_i^2 \leq \sum_{i=1}^{k} \eta_i^2. \tag{2.10}$$

We then add the next eigenfunction $\chi_i \phi_{l_i^m+1}^{(i)}$ to the multiscale finite element space for $1 \leq i \leq k$ where i is in the new arranged order. We call the enriched space V_{ms}^{m+1}, namely, $V_{\text{ms}}^{m+1} = V_{\text{ms}}^m + \text{span}\{\chi_i \phi_{l_i^m+1}^{(i)} |\ 1 \leq i \leq k\}$.

In [31], the above enrichment algorithm can give a fast convergent rate if the eigenvalues in (2.8) increase rapidly. Although rapid eigenvalue growth occur in many cases, this offline enrichment algorithm may be give a rapid convergence in some situations since the offline constructed basis functions may not contain the global information of the problem. Thus, we may need to construct some basis functions involving global information

to accelerate the convergence. In next section, we will introduce an adaptive algorithm by using online constructed basis functions where these basis functions can capture the global behaviors of the solution and thus give a rapid convergent rate.

2.3 Residual based online adaptive GMsFEM

In previous sections, we discussed about the offline adaptive enrichment method and mentioned that online constructed basis functions may be important in some cases. In this section, we will present the online basis function construction algorithm in more details. We call the adaptive algorithm using the online constructed basis functions online adaptive GMsFEM. Based on the analysis and numerical result, we show that using sufficient number of offline basis functions is necessary to give a rapid convergence for online adaptive GsFEM.

Basically, the online adaptive GMsFEM following a similar framework as the offline adaptive GMsFEM. Thus, we will use similar notations as in the previous section. In online adaptive GMsFEM, instead of, using the next eigenfunction constructed in 2.8, we will add a online constructed basis function $\phi_{\text{on}}^{(i)}$ to the multiscale finite element space V_{ms}. Therefore, the resulting multiscale finite element space contain both offline constructed and online constructed basis functions.

Next, we will discuss how to construct the online basis functions. In the m- th enrichment level, similar to the offline adaptive GMsFEM, we will obtain the numerical solution u_{ms}^m by solving (2.5) with $V_{\text{ms}} = V_{\text{ms}}^m$. The main point is how to find a suitable basis function ϕ_{on} with support in ω_i such that the next level multiscale finite element space $V_{\text{ms}}^{m+1} = V_{\text{ms}}^m + \text{span}\{\phi_{on}\}$ can give a good approximation of the reference solution.

Let u_{ms}^{m+1} be the numerical solution for the next enrichment level. By using Galerkin

orthogonal property of the numerical solution. We have

$$\|u_h - u_{\text{ms}}^{m+1}\|_a^2 = \inf_{v \in V_{\text{ms}}^{m+1}} \|u_h - v\|_a^2.$$

By taking $v = u_{\text{ms}}^m + \phi_{on}$, we have

$$\|u_h - u_{\text{ms}}^{m+1}\|_V^2 \leq \|u_h - u_{\text{ms}}^m - \alpha\phi_{on}\|_V^2 = \|u_h - u_{\text{ms}}^m\|_V^2 - 2a(u_h - u_{\text{ms}}^m, \phi_{on}) + a(\phi_{on}, \phi_{on}).$$

To reduce the error for next level enrichment, we will minimize the last two terms in the above inequality. That is, we will find $\phi_{on} \in V_h^{(i)}$ such that,

$$\phi_{on} = \text{argmin}_{\phi \in V_h^{(i)}} \{-2a(u_h - u_{\text{ms}}^m, \phi) + a(\phi, \phi)\} \tag{2.11}$$

Since $R_i(\phi) = \int_{\omega_i} f\phi - \int_{\omega_i} a\nabla u_{\text{ms}}^m \cdot \nabla\phi = a(u_h - u_{\text{ms}}^m, \phi)$, we have

$$\phi_{on} = \text{argmin}_{\phi \in V_h^{(i)}} \{-2R_i(\phi) + a(\phi, \phi)\}$$

and by simple variation argument, we have $\phi_{on} \in V_h^{(i)}$ satisfying

$$a(\phi_{on}, v) = R_i(v) \; \forall v \in V_h^{(i)}. \tag{2.12}$$

Moreover, we know the error of next level is reduced by $2R_i(\phi_{on}) - a(\phi_{on}, \phi_{on}) = a(\phi_{on}, \phi_{on}) = \|R_i\|^2$, namely,

$$\|u_h - u_{\text{ms}}^{m+1}\|^2 \leq \|u_h - u_{\text{ms}}^m\|^2 - \|R_i\|^2. \tag{2.13}$$

We consider the spaces V_{ms}^m contains the first n_i offline basis functions for the coarse

neighborhood ω_i for all level m. Therefore, by (2.9), we obtain

$$\|u - u_{\text{ms}}^m\|_V^2 \leq C_{\text{err}} \sum_{i=1}^{N_c} \|R_i\|^2 (\lambda_{n_i+1}^{\omega_i})^{-1}. \tag{2.14}$$

Combining (2.13) and (2.14), we obtain

$$\|u - u_{\text{ms}}^{m+1}\|_V^2 \leq \Big(1 - \frac{\lambda_{n_i+1}^{\omega_i}}{C_{\text{err}}} \frac{\|R_i\|^2 (\lambda_{n_i+1}^{\omega_i})^{-1}}{\sum_{j=1}^{N_c} \|R_j\|^2 (\lambda_{n_j+1}^{\omega_j})^{-1}}\Big) \|u - u_{\text{ms}}^m\|_V^2.$$

The above inequality show the error decay rate for enriching one basis function with support in ω_i per iteration. To increase the rate of convergence of the method, we can enrich multiple basis functions with non-overlapping support in each iteration. For example, we can choose a index set of some non-overlapping course neighborhoods $I \subset \{1, \cdots, N\}$, that is $\omega_i \cap \omega_j = \emptyset$ for all $i, j \in I$ with $i \neq j$. Next, we can construct a set of basis functions, $\{\phi_{on}^{(i)} \in V_h^{(i)} |\ \forall i \in I\}$, by solving (2.12) for each coarse neighborhood ω_i with $i \in I$. It is easy to check that, using the same argument as above, we can obtain the following inequality

$$\|u - u_{\text{ms}}^{m+1}\|_V^2 \leq \|u - u_{\text{ms}}^m\|_V^2 - \sum_{i \in I} \|R_i\|^2. \tag{2.15}$$

if the coarse neighborhoods $\omega_i,\ \forall i \in I$ are non-overlapping. Consequently, we have

$$\|u - u_{\text{ms}}^{m+1}\|_V^2 \leq \Big(1 - \frac{\Lambda_{\text{min}}^{(I)}}{C_{\text{err}}} \frac{\sum_{i \in I} \|R_i\|^2 (\lambda_{n_i+1}^{\omega_i})^{-1}}{\sum_{j=1}^{N_c} \|R_j\|^2 (\lambda_{n_j+1}^{\omega_j})^{-1}}\Big) \|u - u_{\text{ms}}^m\|_V^2 \tag{2.16}$$

where

$$\Lambda_{\text{min}}^{(I)} = \min_{i \in I} \lambda_{n_i+1}^{\omega_i}. \tag{2.17}$$

Form the above inequality, we can see that the convergence of the online enrichment

method is determined by $\Lambda_{\text{min}}^{(I)}$ which related to the number of basis functions of the offline space. To obtain a rapid convergence, we need to take enough offline basis function such that $\Lambda_{\text{min}}^{(I)}$ is large enough and

$$\frac{\Lambda_{\text{min}}^{(I)}}{C_{\text{err}}} \frac{\sum_{i\in I} r_i^2 (\lambda_{n_i+1}^{\omega_i})^{-1}}{\sum_{j=1}^{N_c} r_j^2 (\lambda_{n_j+1}^{\omega_j})^{-1}} \geq \theta_0$$

for some $0 < \theta_0 < 1$ which is independent of the contrast in $\kappa(x)$.

Hence, we obtain the following convergence for the online adaptive GMsFEM:

$$\|u - u_{\text{ms}}^{m+1}\|_V^2 \leq (1-\theta_0)\|u - u_{\text{ms}}^{m}\|_V^2.$$

We note that $\Lambda_{\text{min}}^{(I)}$ can be very small when there are channels in the domain. This is extensively discussed in [42]. For this reason, we introduce a definition.

Definition 2.3.1. *We say V_{off} satisfies Online Error Reduction Property (ONERP) if*

$$\frac{\Lambda_{\text{min}}^{(I)}}{C_{err}} \frac{\sum_{i\in I} r_i^2 (\lambda_{n_i+1}^{\omega_i})^{-1}}{\sum_{i=1}^{N_c} r_i^2 (\lambda_{n_i+1}^{\omega_i})^{-1}} \geq \theta_0,$$

for some $\theta_0 > \delta > 0$, where δ is independent of physical parameters such as contrast.

We remark that if V_{off} is ONERP, then the error will decrease independent of physical parameters such as the contrast and scales. We will show in our numerical results that if we do not choose V_{off} with ONERP, the online basis functions will not decrease the error. One of easiest way to determine V_{off} being ONERP is to guarantee that $\Lambda_{\text{min}}^{(I)}$ is sufficiently large. In general, one can use the sizes of $\Lambda_{\text{min}}^{(I)}$ and $\sum_{i\in I} r_i^2 (\lambda_{n_i+1}^{\omega_i})^{-1}$ to determine the switching between offline and online.

We remark that one can derive apriori error estimate for $\|u - u_{\text{ms}}^{m+1}\|_V^2$. To do so, we can use an error estimate for the GMsFEM that uses initial multiscale basis functions ([31,

29, 16]) and obtain estimate for $\|u-u_{\text{ms}}^0\|_V^2$, where u_{ms}^0 denotes the multiscale solution that uses the initial basis functions. This convergence rate depends on the use of oversampling and is proportional to $1/\Lambda_*$, where Λ_* is the largest eigenvalue that the corresponding eigenvector is not used in constructing the initial multiscale space.

2.4 Numerical result

In this section, we will present some numerical examples to demonstrate the performance of the online adaptive enrichment method. We first briefly introduce the algorithm for online adaptive GMsFEM. As we mentioned before, the online adaptive enrichment will start constructing a offline space with a fixed number of offline basis functions for each coarse neighborhood. The offline space is denoted as V_{off}. We consider the initial multiscale finite element space V_{ms}^1 be the offline space V_{off}. Next, the coarse neighborhoods are denoted by $\omega_{i,j}$, where $i = 1, 2, \cdots, N_x$ and $j = 1, 2, \cdots, N_y$ and N_x and N_y are the number of coarse nodes in the x and y directions respectively. We consider $I_{x,\text{odd}}$ and $I_{x,\text{even}}$ as the set of odd and even indices from $\{1, 2, \cdots, N_x\}$. Similarly, $I_{y,\text{odd}}$ and $I_{y,\text{even}}$ are the set of odd and even indices from $\{1, 2, \cdots, N_y\}$. In each iteration of our online adaptive GMsFEM, we will perform 4 sub-iterations which add online basis functions in the non-overlapping coarse neighborhoods $\omega_{i,j}$ with $(i,j) \in I_{x,\text{odd}} \times I_{y,\text{odd}}$, $(i,j) \in I_{x,\text{odd}} \times I_{y,\text{even}}$, $(i,j) \in I_{x,\text{even}} \times I_{y,\text{odd}}$ and $(i,j) \in I_{x,\text{even}} \times I_{y,\text{even}}$ respectively.

The computational domain D is considered as $[0,1]^2$ and the coarse mesh is an uniform square mesh with mesh size $H = 1/16$. The fine mesh is an uniform square mesh with mesh size $H = 1/256$. In Figure 2.1, we will show the permeability field κ and the source function f. In the following sections, we will use the following error quantities to measure the accuracy of our algorithm.

$$e_2 = \frac{\|u - u_{\text{ms}}\|_{L^2(D)}}{\|u\|_{L^2(D)}}, \quad e_a = \frac{\|u - u_{\text{ms}}\|_V}{\|u\|_V},$$

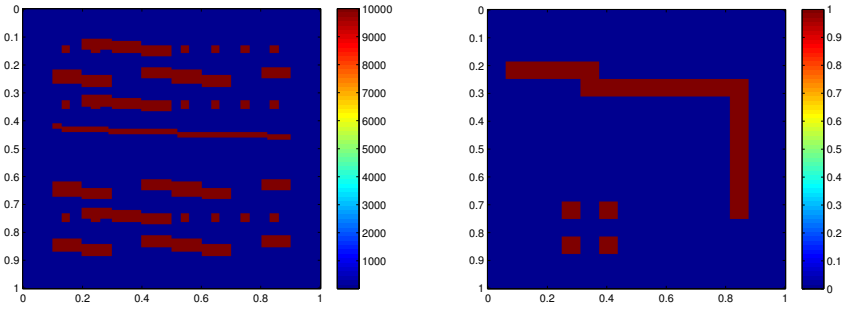

Figure 2.1: Left: Permeability field κ. Right: Source function f.

2.4.1 Comparison of using different number of initial basis

In this subsection, we will present the numerical result to compare the convergence rate of using different number of initial basis functions. In Table 2.1, we will show the result of using one basis function per coarse neighborhood for the offline space. In the table, we will show the number of basis functions used in each coarse neighborhood and the degrees of freedom (DOF), which is the total number of basis functions. In the table, we will consider two different contrast level. In the left table, we have the contrast $\frac{\max \kappa}{\min \kappa} = 10^4$. In the right table, we have a contrast $\frac{\max \kappa}{\min \kappa} = 10^6$ which is 100 times larger than the contrast in the left table. That is, we have a permeability field with permeability equal to 10^6 in the inclusion and channels in Figure 2.1. We know that the first few eigenvalue are dependent on the contrast([42]). For using 100 times contrast, we will have 100 times smaller eigenvalues for the first few eigenfunctions. By comparing the convergence history for using different level of contrast (in Table 2.1), we can see the convergence rate is dependent on the contrast if we do not have enough initial basis functions. Similarly, in Table 2.2, we can see the convergence rate is slower for the higher contrast case if we use two initial basis functions. However, in Table 2.3, we show that the convergence rate does

not dependent on the contrast since $\Lambda_{\min}$ is independent of the contrast. In Figure 2.2, we plot the relative errors in energy norm and the logarithm of relative errors in energy norm against total number of degrees of freedom. for various number of the initial basis functions for two different contrasts. We observed that the convergence rates are similar for using 3 and 4 initial basis functions. Therefore, we can conclude that the convergence rate is independent of contrast if we have enough initial basis functions. Moreover, we can see the error decay exponentially for using online adaptive method.

From Figure 2.2, we can observe the following facts. First, we observe if we choose the number of initial basis functions to be 2 (which satisfies ONERP), it will give the smallest error for a fixed coarse space dimension. Indeed, if we start with the smallest number of initial basis functions (that satisfy ONERP), then at every iteration of the online stage, the error will reduce more compared to the offline-stage basis addition (i.e., adding the basis functions computed in the offline stage). On the other hand, if we would like to reduce the online cost associated with computing basis functions, then it is more advantageous to choose more offline basis functions. Indeed, by choosing a larger number of initial offline basis functions will give us a better result, as we observe from Figure 2.2. For example, if we compare the errors for one online basis addition, we will find that the case with 4 initial offline basis functions gives smallest error. This is because the initial error for the four offline basis functions is smallest among those shown in Figure 2.2. In general, one needs to be careful as more offline basis functions will increase the dimension of the online system and, consequently, the cost of solving the coarse-grid system.

Next, we will present an example with a different medium parameter κ shown in Figure 2.3 to show the importance of ONERP. The source function f is taken as the constant 1. The domain D is divided into 8×8 coarse blocks consisting of uniform squares. Each coarse block is then divided into 32×32 fine blocks also consisting of uniform squares.

The convergence history for the use of one initial basis and the corresponding total

number of basis (DOF)	e_a	e_2
1(225)	60.71%	33.87%
2(450)	33.10%	13.38%
3(675)	14.38%	3.25%
4(900)	4.28%	1.02%
5(1125)	1.33%	0.24%
6(1350)	0.065%	0.0028%
7(1575)	0.00083%	2.96e-05%
8(1800)	1.59e-05%	4.87e-07%
9(2025)	2.35e-07%	2.10e-08%

number of basis (DOF)	e_a	e_2
1(225)	60.90%	34.15%
2(450)	35.90%	15.87%
3(675)	35.00%	15.29%
4(900)	25.77%	8.77%
5(1125)	14.17%	4.39%
6(1350)	7.79%	2.78%
7(1575)	6.83%	2.06%
8(1800)	4.15%	1.20%
9(2025)	2.60%	0.64%

Table 2.1: Convergence history for the permeability field in Figure 2.1 and for the case with one initial basis. Left: Lower contrast($1e4$). Right: Higher contrast($1e6$).

number of degrees of freedom (DOF) are shown in Table 2.4. In this case, $\Lambda_{\text{min}} = 0.0033$, which is considered to be very small, and we observe very slow convergence of the online adaptive procedure. In Table 2.5, we present the convergence history for the use of two to five initial basis, where we only show the results for the last 4 iterations. We see that the values of Λ_{min} increase as we increase the number of initial basis. We also observe that the convergence rate increase when we raise the number of initial basis from 2 to 4. For the use of 5 initial basis, we again see rapid convergence and a faster convergence compared when using 4 initial basis functions. In particular, we observe (based on 3 iterations following the initial one) that the error decays at 130-fold when 5 initial basis functions are selected, while the error decay is about 90-fold when 4 initial basis functions

num of basis(DOF)	e_a	e_2
2 (450)	26.60%	6.92%
3 (675)	1.46%	0.060%
4 (900)	0.017%	0.000079%
5 (1125)	0.000021%	1.06e-05%
6 (1350)	3.56e-06%	1.65e-07%

num of basis (DOF)	e_a	e_2
2 (450)	27.17%	7.53%
3 (675)	4.99%	0.79%
4 (900)	0.20%	0.0073%
5 (1125)	0.0017%	8.16e-05
6 (1350)	2.71e-05%	1.09e-06%

Table 2.2: Convergence history for the permeability field in Figure 2.1 and for the case with two initial basis. Left: Lower contrast($1e4$). Right: Higher contrast($1e6$).

are selected. A comparison of error decay for the use of 1 to 5 initial basis functions is shown in Figure 2.4. We have also tested harmonic basis functions and the results are similar, i.e., the convergence rate is very slow unless sufficient number of offline basis functions is selected.

In conclusion, we observe

- If V_{ms} does not satisfy ONERP, then the error decay is slower as the contrast becomes larger.
- If V_{ms} does not satisfy ONERP, in some cases, we have observed the error does not decrease as we add online basis functions (see Table 2.4, 2.5).
- If V_{ms} satisfies ONERP, then we observe a fast convergence, which is independent of contrast.

num of basis (DOF)	e_a	e_2
3 (675)	16.95%	2.53%
4 (900)	0.54%	0.023%
5 (1125)	0.011%	0.00040%
6 (1350)	9.07e-05%	3.79e-06%
7 (1575)	1.38e-06%	6.05e-08%
num of basis (DOF)	e_a	e_2
3 (675)	16.96%	2.54%
4 (900)	0.54%	0.023%
5 (1125)	0.011%	0.00041%
6 (1350)	9.07e-05%	3.79e-06%
7 (1575)	1.58e-06%	5.49e-07%

Table 2.3: Convergence history for the permeability field in Figure 2.1 and for the case with three initial basis. Left: Lower contrast($1e4$). Right: Higher contrast($1e6$).

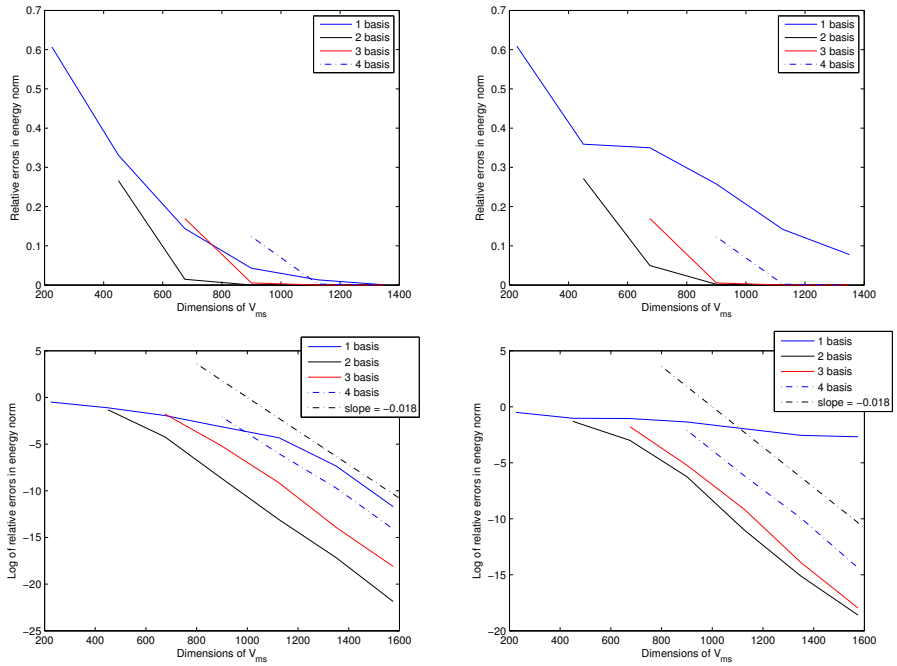

Figure 2.2: Comparison for the permeability field in Figure 2.1 with different choices of the number of initial basis. Left: The contrast is 10^4. Right: The contrast is 10^6.

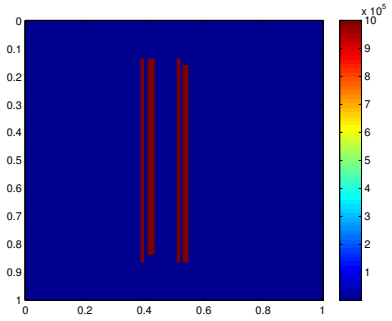

Figure 2.3: Permeability field κ for second case.

DOF	e_a	e_2
81	17.24%	4.35%
162	2.80%	1.02%
243	2.65%	0.88%
323	2.64%	0.87%
401	1.09%	0.081%
478	0.74%	0.094%
555	0.73%	0.090%
632	0.48%	0.039%
709	0.37%	0.026%

Table 2.4: Convergence history for the permeability field in Figure 2.3 and for the case with one initial basis. (The value of $\Lambda_{\text{min}} = 0.0033$).

DOF	e_a	e_2
162	13.29%	2.90%
405	1.60%	0.17%
486	0.33%	0.025%
567	0.30%	0.022%

DOF	e_a	e_2
405	7.24%	0.92%
486	0.0684%	0.0028%
567	0.00049%	1.51e-05%
635	3.80e-06%	2.49e-06%

Table 2.5: Convergence history for the permeability field in Figure 2.3. Left: Two initial basis ($\Lambda_{\text{min}} = 0.026$). Right: Five initial basis ($\Lambda_{\text{min}} = 319.32$).

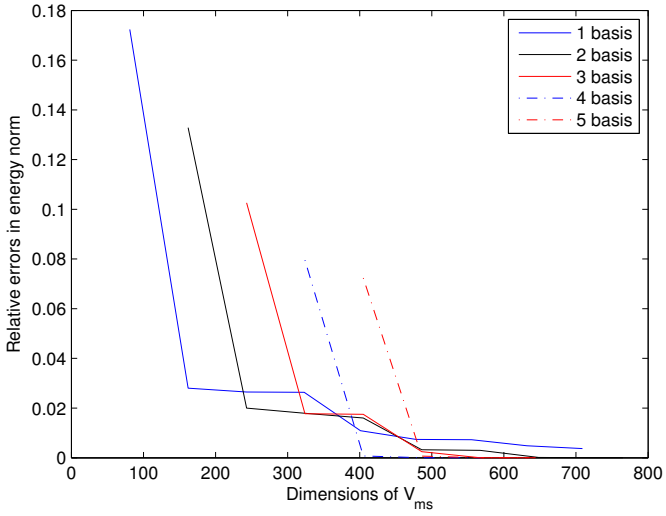

Figure 2.4: Comparison for the permeability field in Figure 2.3 with different numbers of initial basis functions.

2.4.2 Adaptive online enrichment

In this section, the online enrichment is performed only for regions with the residual that is larger than a certain threshold. In the first case, the online enrichment is performed for the coarse regions with a residual error bigger than a certain threshold which will be taken 10^{-3}, 10^{-4}, and 10^{-5}. In the second case, the online enrichment is performed for coarse regions that have cumulative residual that is θ fraction of the total residual. One of our objectives is to show that one can drive the error down to a number below a threshold, adaptively.

In our numerical results, we will consider three tolerances (tol) 10^{-3}, 10^{-4} and 10^{-5}. We will enrich coarse regions, if the H^{-1}-norm of the residual is bigger than the tolerance. In Table 2.6, we show the errors when using 1 initial basis function for tolerances 10^{-3}, 10^{-4} and 10^{-5}. We first observe a very slow reduction in errors similar to the results presented in the previous section. Another observation is that the energy error of the multiscale solution is in the same order of the tolerance, and the error cannot be further reduced if we perform more iterations. This allows us to compute a multiscale solution with a prescribed error level by choosing a suitable tolerance in the adaptive algorithm. In Table 2.7 and Table 2.8, we show the errors for the last three iterations when using 2 and 3 initial basis functions respectively for tolerances 10^{-3}, 10^{-4} and 10^{-5}. We observe that the convergences are much faster. In addition, the energy errors are again have the same magnitude as the tolerances. From these results, we obtain the following conclusions.

- Using smaller tolerances, we can reduce the final error below desired threshold errors.

- We have observed that the number of initial basis functions are important to achieve better results. For example, we observe a slow decay of the error when 1 initial basis function is selected. Moreover, if the contrast is higher, the decay becomes slower.

DOF	e_a	e_2
225	60.71%	33.87%
447	33.10%	13.39%
652	14.43%	3.28%
776	4.37%	1.06%
824	1.83%	0.37%
847	1.10%	0.25%
863	0.50%	0.029%

DOF	e_a	e_2
225	60.71%	33.87%
449	33.10%	13.38%
674	14.38%	3.25%
883	4.28%	1.02%
1031	1.33%	0.24%
1125	0.082%	0.0036%
1136	0.052%	0.0023%

DOF	e_a	e_2
225	60.701%	33.87%
450	33.10%	13.38%
675	14.38%	3.25%
899	4.28%	1.02%
1114	1.33%	0.24%
1275	0.065%	0.0028%
1338	0.0048%	0.00017%

Table 2.6: Convergence history with a fixed tolerance (tol) and one initial basis for the permeability field in Figure 2.1. Left: $tol = 10^{-3}$. Middle: $tol = 10^{-4}$. Right: $tol = 10^{-5}$.

In our next numerical example, the online enrichment is performed for coarse regions that have a cumulative residual that is θ fraction of the total residual. Assume that the local residuals are arranged so that

$$r_1 \geq r_2 \geq r_3 \geq \cdots .$$

Then, we only add the basis $\phi_1, \cdots, \phi_k$ for the coarse neighborhoods $\omega_1, \cdots, \omega_k$ such that k is the smallest integer with

$$\theta \sum_{i=1}^{N_c} r_i^2 \leq \sum_{i=1}^{k} r_i^2.$$

In Table 2.9, we present numerical results for the last 4 iterations when using 1, 2 and 3 initial basis functions with the tolerance 10^{-4} and $\theta = 0.7$. We observe that one can reduce the total number of basis functions compared to the previous case to achieve a similar error. Our conclusions regarding the importance of ONERP condition for V_{ms} is

DOF	e_a	e_2
450	26.60%	6.92%
649	1.49%	0.063%
666	0.53%	0.028%

DOF	e_a	e_2
450	26.60%	6.92%
674	1.46%	0.059%
802	0.048%	0.0022%

DOF	e_a	e_2
675	1.46%	0.060%
885	0.017%	0.00079%
925	0.0043%	0.00019%

Table 2.7: Convergence history with a fixed tolerance (tol) and two initial basis for the permeability field in Figure 2.1. Left: $tol = 10^{-3}$. Middle: $tol = 10^{-4}$. Right: $tol = 10^{-5}$.

DOF	e_a	e_2
675	16.96%	2.54%
863	0.63%	0.027%
867	0.44%	0.018%

DOF	e_a	e_2
675	16.96%	2.54%
898	0.054%	0.023%
993	0.046%	0.0015%

DOF	e_a	e_2
900	0.54%	0.023%
1087	0.011%	0.00043%
1106	0.0050%	0.00019%

Table 2.8: Convergence history with a fixed tolerance (tol) and three initial basis for the permeability field in Figure 2.1. Left: $tol = 10^{-3}$. Middle: $tol = 10^{-4}$. Right: $tol = 10^{-5}$.

the same as before.

DOG	e_a	e_2
620	1.10%	0.23%
709	0.49%	0.082%
787	0.050%	0.0024%
789	0.046%	0.0022%

DOF	e_a	e_2
450	26.60%	6.92%
576	1.94%	0.12%
690	0.20%	0.0099%
744	0.051%	0.0023%

DOF	e_a	e_2
675	16.96%	2.54%
827	1.02%	0.045%
957	0.091%	0.0033%
987	0.048%	0.0017%

Table 2.9: Convergence results using cumulative errors with $\theta = 0.7$, $tol = 10^{-4}$ and the permeability field in Figure 2.1. Left: One initial basis. Middle: Two initial basis. Right: Three initial basis.

3. AN ADAPTIVE GENERALIZED MULTISCALE DISCONTINUOUS GALERKIN METHOD (GMSDGM) FOR HIGH-CONTRAST FLOW PROBLEMS

3.1 GMsDGM for high-contrast flow problems

3.1.1 Overview

In this section, we will introduce the GMsDGM ([43]) for the flow problem. As in the chapter 2, we denote the computational domain by Ω. We will consider the flow problem (2.1) with high-contrast medium coefficient $\kappa(x)$ and Dirichlet boundary condition $u = g$ on boundary $\partial\Omega$. Instead of using continuous Galerkin approach, we would like to introduce an efficient coarse grid discontinuous Galerkin approach to capture the multiscale feature of the solution.

We denote the coarse grid by $\mathcal{T}^H$ as in the previous chapter. We then denote consider the fine grid $\mathcal{T}^h$ to be a refinement of the coarse grid. To simplify the discussion of the methodology, we will consider a conforming refinement only. We remark that the conforming refinement is not necessary in the discontinuous Galerkin case.

GMsDGM also consists of two main ingredients, which are the construction of local basis functions and the global coarse grid level coupling. For the local basis functions, a snapshot space $V_{\text{snap}}^{(i)}$ is first constructed for each coarse grid block $K_i \in \mathcal{T}^H$. The snapshot space contains a large library of basis functions, which can be used to obtain a fine scale approximate solution to (2.1). A spectral problem is then solved in the snapshot space $V_{\text{snap}}^{(i)}$ and eigenfunctions corresponding to dominant modes are used as the final basis functions. The resulting space is called the local offline space $V_{\text{off}}^{(i)}$ for the i-th coarse grid block K_i. The global offline space V_{off} is then defined as the linear span of all these $V_{\text{off}}^{(i)}$, for $i = 1, 2, \cdots, N$. This global offline space V_{off} will be used as the approximation space of

our discontinuous Galerkin method, which can be formulated as: find $u_H^{\text{DG}} \in V_{\text{off}}$ such that

$$a_{\text{DG}}(u_H^{\text{DG}}, v) = (f, v), \quad \forall v \in V_{\text{off}}, \tag{3.1}$$

where the bilinear form a_{DG} is defined as

$$a_{\text{DG}}(u, v) = a_H(u, v) - \sum_{E \in \mathcal{E}^H} \int_E \Big(\{\kappa \nabla u \cdot n_E\} [\![v]\!] + \{\kappa \nabla v \cdot n_E\} [\![u]\!] \Big) + \sum_{E \in \mathcal{E}^H} \frac{\gamma}{h} \int_E \overline{\kappa} [\![u]\!] [\![v]\!] \tag{3.2}$$

with

$$a_H(u, v) = \sum_{K \in \mathcal{T}_H} a_H^K(u, v), \quad a_H^K(u, v) = \int_K \kappa \nabla u \cdot \nabla v, \tag{3.3}$$

where $\gamma > 0$ is a penalty parameter, n_E is a fixed unit normal vector defined on the coarse edge $E \in \mathcal{E}^H$. Note that, in (3.2), the average and the jump operators are defined in the classical way. Specifically, consider an interior coarse edge $E \in \mathcal{E}^H$ and let K^+ and K^- be the two coarse grid blocks sharing the edge E. For a piecewise smooth function G, we define

$$\{G\} = \frac{1}{2}(G^+ + G^-), \qquad [\![G]\!] = G^+ - G^-, \qquad \text{on } E,$$

where $G^+ = G|_{K^+}$ and $G^- = G|_{K^-}$ and we assume that the normal vector n_E is pointing from K^+ to K^-. Moreover, on the edge E, we define $\overline{\kappa} = (\kappa_{K^+} + \kappa_{K^-})/2$ where $\kappa_{K^\pm}$ is the maximum value of κ over $K^\pm$. For a coarse edge E lying on the boundary ∂D, we define

$$\{G\} = [\![G]\!] = G, \quad \text{and} \quad \overline{\kappa} = \kappa_K \qquad \text{on } E,$$

where we always assume that n_E is pointing outside of D. We note that the DG coupling (3.1) is the classical interior penalty discontinuous Galerkin (IPDG) method [44] with our multiscale basis functions as the approximation space.

3.1.2 Construction of offline basis functions

In this subsection, we will give a detailed description of the method. We will first give the construction of the snapshot space, and then give the definitions of the local spectral problems for the construction of the offline space. Furthermore, a priori estimate of the method will be derived.

Let $K_i \in \mathcal{T}^H$ be a given coarse grid block. We will define two types of snapshot spaces. The first type of local snapshot space $V_{1,\text{snap}}^{(i)}$ for the coarse grid block K_i is defined as the linear span of all harmonic extensions. Specifically, given a function δ_k defined on ∂K_i, we find $\psi_k^{i,\text{snap}} \in V_h(K_i)$ by

$$\begin{aligned} \int_{K_i} \kappa \nabla \psi_k^{i,\text{snap}} \cdot \nabla v = 0, \quad \forall\ v \in V_h^0(K_i), \\ \psi_k^{i,\text{snap}} = \delta_k, \quad \text{on } \partial K_i, \end{aligned} \tag{3.4}$$

where $V_h(K_i)$ is the standard conforming piecewise polynomial finite element space with respect to the fine grid defined on K_i, $V_h^0(K_i)$ is the subspace of $V_h(K_i)$ containing functions vanishing on ∂K_i and δ_k is piecewise polynomial on ∂K_i with respect to the fine grid such that δ_k has the value one at the k-th fine grid degree of freedom and value zero at all the remaining fine grid degree of freedoms. The linear span of the above harmonic extensions is the local snapshot space $V_{1,\text{snap}}^{(i)}$, namely

$$V_{1,\text{snap}}^{(i)} = \text{span}\{\psi_k^{i,\text{snap}}, \quad k = 1, 2, \cdots, M^{i,\text{snap}}\},$$

where $M^{i,\text{snap}}$ is the number of basis functions in $V_{1,\text{snap}}^{(i)}$, which is also equal to the number of fine grid degree of freedoms on ∂K_i. The second type of local snapshot space $V_{2,\text{snap}}^{(i)}$ for the coarse grid block K_i is defined as $V_{2,\text{snap}}^{(i)} = V_h^0(K_i)$. It is easy to see that $V_h(K_i) = V_{1,\text{snap}}^{(i)} + V_{2,\text{snap}}^{(i)}$, namely the space $V_h(K_i)$ is decomposed as the sum of harmonic extensions

and functions vanishing on the boundary ∂K_i. Moreover, the global snapshot space $V_{1,\text{snap}}$ is defined so that any $v \in V_{1,\text{snap}}$ if $v|_{K_i} \in V_{1,\text{snap}}^{(i)}$. The space $V_{2,\text{snap}}$ is defined similarly.

We will perform dimension reductions on the above snapshot spaces by the use of some carefully selected spectral problems. Based on our analysis to be presented in this section, we define the spectral problem for $V_{1,\text{snap}}^{(i)}$ as finding eigenpairs $(\phi_k^{(i)}, \lambda_{1,k}^{(i)})$, $k = 1, 2, \cdots, M^{i,\text{snap}}$, such that

$$\int_{K_i} \kappa \nabla \phi_k^{(i)} \cdot \nabla v = \frac{\lambda_{1,k}^{(i)}}{H} \int_{\partial K_i} \widetilde{\kappa} \phi_k^{(i)} v, \qquad \forall v \in V_{1,\text{snap}}^{(i)}, \tag{3.5}$$

where $\widetilde{\kappa}$ is the maximum of $\overline{\kappa}$ over all coarse edges $E \in \partial K_i$. Moreover, we assume that

$$\lambda_{1,1}^{(i)} < \lambda_{1,2}^{(i)} < \cdots < \lambda_{1,M^{i,\text{snap}}}^{(i)}.$$

For the space $V_{2,\text{snap}}^{(i)}$, we define the spectral problem as finding eigenpairs $(\xi_k^{(i)}, \lambda_{2,k}^{(i)})$, $k = 1, 2, \cdots$, such that

$$\int_{K_i} \kappa \nabla \xi_k^{(i)} \cdot \nabla v = \frac{\lambda_{2,k}^{(i)}}{H^2} \int_{K_i} \kappa \xi_k^{(i)} v, \qquad \forall v \in V_{2,\text{snap}}^{(i)}, \tag{3.6}$$

where we also assume that

$$\lambda_{2,1}^{(i)} < \lambda_{2,2}^{(i)} < \cdots$$

In the spectral problems (3.5) and (3.6), we will take respectively the first $l_{1,i}$ and $l_{2,i}$ eigenfunctions to form the offline space for the coarse grid block K_i. The local offline spaces are then defined as

$$V_{1,\text{off}}^{(i)} = \text{span}\{\phi_l^{(i)}, \quad l = 1, 2, \cdots, l_{1,i}\},$$
$$V_{2,\text{off}}^{(i)} = \text{span}\{\xi_l^{(i)}, \quad l = 1, 2, \cdots, l_{2,i}\}.$$

We define $V_{\text{off}}^{(i)} = V_{1,\text{off}}^{(i)} + V_{2,\text{off}}^{(i)}$. The global offline space V_1^{off} is defined so that the restriction of any function $v \in V_1^{\text{off}}$ on the coarse grid block K_i belongs to $V_1^{i,\text{off}}$. The definition for V_2^{off} is defined similarly. In addition, we define $V_{\text{off}} = V_{1,\text{off}} + V_{2,\text{off}}$. This space is used as the approximation space in (3.1). We remark that we assume the eigenvalues in the spectral problems (3.5) and (3.6) are simple only to simplify the notations. If there is a non-simple eigenvalue, then there are multiple eigenfunctions correspond to this eigenvalue. In this case, we will include all these eigenfunctions in the space when this eigenvalue is chosen in our adaptive process.

Now we will analyze the method defined in (3.1). For any piecewise smooth function u, we define the DG-norm by

$$\|u\|_{\text{DG}}^2 = a_H(u,u) + \sum_{E\in\mathcal{E}_H} \frac{\gamma}{h}\int_E \overline{\kappa}[\![u]\!]^2\, ds.$$

Let K be a coarse grid block and let $n_{\partial K}$ be the unit outward normal vector on ∂K. We denote $V_h(\partial K)$ by the restriction of the conforming space $V_h(K)$ on ∂K. For any $u \in V_1^{\text{snap}}$, the normal flux $\kappa\nabla u \cdot n_{\partial K}$ is understood as an element in $V_h(\partial K)$ and is defined by

$$\int_{\partial K} (\kappa\nabla u \cdot n_{\partial K})\cdot v = \int_K \kappa\nabla u\cdot\nabla\widehat{v}, \quad v\in V^h(\partial K), \tag{3.7}$$

where $\widehat{v}\in V_h(K)$ is the harmonic extension of v in K. By the Cauchy-Schwarz inequality,

$$\int_{\partial K} (\kappa\nabla u \cdot n_{\partial K})\cdot v \le a_H^K(u,u)^{\frac{1}{2}}\, a_H^K(\widehat{v},\widehat{v})^{\frac{1}{2}}.$$

By an inverse inequality and the fact that $\widehat{v}$ is the harmonic extension of v

$$a_H^K(\widehat{v},\widehat{v}) \le \kappa_K C_{\text{inv}}^2 h^{-1}\int_{\partial K}|v|^2, \tag{3.8}$$

where we recall that κ_K is the maximum of κ over K and $C_{\text{inv}} > 0$ is the constant from inverse inequality. Thus,

$$\int_{\partial K} (\kappa \nabla u \cdot n_{\partial K}) \cdot v \leq \kappa_K^{\frac{1}{2}} C_{\text{inv}} h^{-\frac{1}{2}} \|v\|_{L^2(\partial K)} \, a_H^K(u,u)^{\frac{1}{2}}.$$

This shows that

$$\int_{\partial K} |\kappa \nabla u \cdot n_{\partial K}|^2 \leq \kappa_K C_{\text{inv}}^2 h^{-1} a_H^K(u,u). \tag{3.9}$$

We remark that the above steps show that the inequality (3.9) holds for any u in the snapshot space V_1^{snap}. On the other hand, for any function $u \in V_2^{\text{snap}}$, one can show that (3.9) holds, with possibly a different constant C_{inv}, by using arguments based on inverse inequality. See also Lemma 2 in [22]. Furthermore, the relation (3.7) will be used in the proof of Theorem 3.1.2.

Our first step in the development of an a priori estimate is to establish the continuity and the coercivity of the bilinear form (3.2) with respect to the DG-norm.

Lemma 3.1.1. *Assume that the penalty parameter γ is chosen so that $\gamma > C_{\text{inv}}^2$. The bilinear form a_{DG} defined in (3.2) is continuous and coercive, that is,*

$$a_{DG}(u,v) \leq a_1 \|u\|_{DG} \, \|v\|_{DG}, \tag{3.10}$$

$$a_{DG}(u,u) \geq a_0 \|u\|_{DG}^2, \tag{3.11}$$

for all u, v, where $a_0 = 1 - C_{\text{inv}} \gamma^{-\frac{1}{2}} > 0$ and $a_1 = 1 + C_{\text{inv}} \gamma^{-\frac{1}{2}}$.

Proof. By the definition of a_{DG}, we have

$$a_{\text{DG}}(u,v) = a_H(u,v) - \sum_{E \in \mathcal{E}^H} \int_E \Big(\{\kappa \nabla u \cdot n_E\} [\![v]\!] + \{\kappa \nabla v \cdot n_E\} [\![u]\!] \Big) + \sum_{E \in \mathcal{E}^H} \frac{\gamma}{h} \int_E \overline{\kappa} [\![u]\!] [\![v]\!].$$

Notice that

$$a_H(u,v) + \sum_{E\in\mathcal{E}^H} \frac{\gamma}{h}\int_E \overline{\kappa}[\![u]\!]\cdot[\![v]\!] \le \|u\|_{\text{DG}}\,\|v\|_{\text{DG}}.$$

For an interior coarse edge $E \in \mathcal{E}^H$, we let $K^+, K^- \in \mathcal{T}^H$ be the two coarse grid blocks having the edge E. By the Cauchy-Schwarz inequality, we have

$$\int_E \{\kappa\nabla u\cdot n_E\}\cdot[\![v]\!] \le \Big(h\int_E \{\kappa\nabla u\cdot n_E\}^2(\overline{\kappa})^{-1}\Big)^{\frac{1}{2}}\Big(\frac{1}{h}\int_E \overline{\kappa}[\![v]\!]^2\Big)^{\frac{1}{2}}. \tag{3.12}$$

Notice that

$$h\int_E \{\kappa\nabla u\cdot n_E\}^2(\overline{\kappa})^{-1} \le h\Big(\int_E (\kappa^+\nabla u^+\cdot n_E)^2(\kappa_{K^+})^{-1} + \int_E (\kappa^-\nabla u^-\cdot n_E)^2(\kappa_{K^-})^{-1}\Big)$$

where $u^\pm = u|_{K^\pm}$, $\kappa^\pm = \kappa|_{K^\pm}$. So, summing the above over all E and by (3.9), we have

$$h\sum_{E\in\mathcal{E}^H}\int_E \{\kappa\nabla u\cdot n_E\}^2(\overline{\kappa})^{-1} \le h\sum_{K\in\mathcal{T}_H}\int_{\partial K}(\kappa\nabla u\cdot n_{\partial K})^2(\kappa_K)^{-1} \le C_{\text{inv}}^2 a_H(u,u).$$

Thus we have

$$\sum_{E\in\mathcal{E}^H}\int_E \{\kappa\nabla u\cdot n_E\}[\![v]\!] \le C_{\text{inv}}a_H(u,u)^{\frac{1}{2}}\Big(\sum_{E\in\mathcal{E}^H}\frac{1}{h}\int_E \overline{\kappa}[\![v]\!]^2\,ds\Big)^{\frac{1}{2}}. \tag{3.13}$$

Similarly, we have

$$\sum_{E\in\mathcal{E}^H}\int_E \{\kappa\nabla v\cdot n_E\}[\![u]\!] \le C_{\text{inv}}a_H(v,v)^{\frac{1}{2}}\Big(\sum_{E\in\mathcal{E}^H}\frac{1}{h}\int_E \overline{\kappa}[\![u]\!]^2\,ds\Big)^{\frac{1}{2}}.$$

Summing the above two inequalities, we have

$$\sum_{E\in\mathcal{E}^H}\int_E \Big(\{\kappa\nabla u\cdot n_E\}[\![v]\!] + \{\kappa\nabla v\cdot n_E\}[\![u]\!]\Big) \le C_{\text{inv}}\gamma^{-\frac{1}{2}}\|u\|_{\text{DG}}\,\|v\|_{\text{DG}}. \tag{3.14}$$

This proves the continuity (3.10).

For the coercivity (3.11), we have

$$a_{\text{DG}}(u,u) = \|u\|^2_{\text{DG}} - \sum_{E\in\mathcal{E}^H}\int_E \Big(\{\kappa\nabla u\cdot n_E\}\cdot[\![u]\!] + \{\kappa\nabla u\cdot n_E\}\cdot[\![u]\!]\Big).$$

By (3.14), we have

$$a_{\text{DG}}(u,u) \geq (1 - C_{\text{inv}}\gamma^{-\frac{1}{2}})\|u\|^2_{\text{DG}},$$

which gives the desired result.

□

In the following, we will prove an a priori estimate of the method (3.1). First, we let

$$V^h_{\text{DG}} = \{v \in L^2(D) \,:\, v|_K \in V_h(K)\}. \tag{3.15}$$

Let $u_h \in V_h^{\text{DG}}$ be the fine grid solution which satisfies

$$a_{\text{DG}}(u_h, v) = (f, v), \quad \forall v \in V_h^{\text{DG}}. \tag{3.16}$$

It is well-known that u_h converges to the exact solution u in the DG-norm as the fine mesh size $h \to 0$. Next, we define a projection $u_1 \in V_{1,\text{snap}}$ of u_h in the snapshot space by the following construction. For each coarse grid block K_i, the restriction of u_1 on K_i is defined as the harmonic extension of u_h, that is,

$$\begin{aligned} \int_{K_i} \kappa\nabla u_1\cdot\nabla v = 0, \qquad & \forall\, v \in V_h^0(K_i) \\ u_1 = u_h, \quad & \text{on } \partial K_i. \end{aligned} \tag{3.17}$$

The following theorem gives an a priori estimate for the GMsDGM (3.1).

Theorem 3.1.2. *Let $u_h \in V_h^{DG}$ be the fine grid solution defined in (3.16) and u_H be the GMsDGM solution defined in (3.1). Then we have*

$$\|u_h - u_H\|_{DG}^2 \leq C\Big(\sum_{i=1}^{N} \frac{H}{\widetilde{\kappa}\lambda_{1,l_{1,i}+1}^{(i)}}(1+\frac{\gamma H}{h\lambda_{1,l_{1,i}+1}^{(i)}})\int_{\partial K_i}(\kappa\nabla u_1\cdot n_{\partial K})^2 + \sum_{K\in\mathcal{T}_H}\frac{H^2}{\lambda_{2,l_{2,i}+1}^{(i)}}\|f\|_{L^2(K)}^2 + C_{\text{inv}}^2\sum_{E\in\mathcal{E}^H}\frac{1}{h}\int_E \overline{\kappa}[\![u_h]\!]^2\Big),$$

where u_1 is defined in (3.17).

Proof. First, we write $u_h = u_1 + u_2$ where $u_2 = u_h - u_1$. Notice that, on each coarse grid block K_i, the functions u_1 and u_2 can be represented by

$$u_1 = \sum_{l=1}^{M_i} c_l\phi_l^{(i)} \quad \text{and} \quad u_2 = \sum_{l\geq 1} d_l\xi_l^{(i)} \tag{3.18}$$

where $M_i = M^{i,\text{snap}}$ and we assume that the functions $\phi_l^{(i)}$ and $\xi_l^{(i)}$ are normalized so that

$$\int_{\partial K_i}\widetilde{\kappa}(\phi_l^{(i)})^2 = 1 \quad \text{and} \quad \int_{K_i}\kappa(\xi_l^{(i)})^2 = 1.$$

Notice that, the functions u_1 and u_2 belong to the snapshot spaces $V_{1,\text{snap}}$ and $V_{2,\text{snap}}$ respectively. We will need two functions $\widehat{u}_1$ and $\widehat{u}_2$, which belong to the offline spaces $V_{1,\text{off}}$ and $V_{2,\text{off}}$ respectively. These functions are defined by

$$\widehat{u}_1 = \sum_{l=1}^{l_{1.i}} c_l\phi_l^{(i)} \quad \text{and} \quad \widehat{u}_2 = \sum_{l=1}^{l_{2.i}} d_l\xi_l^{(i)} \qquad \text{on } K_i.$$

We remark that $\widehat{u}_1$ and $\widehat{u}_2$ are the truncation of u_1 and u_2 up to the eigenfunctions selected to form the offline space.

Next, we will find an estimate of $\|u_1 - \widehat{u}_1\|_{\text{DG}}$. Let $K_i \in \mathcal{T}^H$ be a given coarse grid

block. Recall that the spectral problem to form $V_{1,\text{off}}^{(i)}$ is

$$\int_{K_i} \kappa \nabla \phi_k^{(i)} \cdot \nabla v = \frac{\lambda_{1,k}^{(i)}}{H} \int_{\partial K_i} \widetilde{\kappa} \phi_k^{(i)} v, \qquad \forall v \in V_{1,\text{snap}}^{(i)}.$$

By the definition of the flux defined in (3.7), the above spectral problem can be represented as

$$\int_{\partial K_i} (\kappa \nabla \phi_k^{(i)} \cdot n_{\partial K_i}) v \, ds = \frac{\lambda_{1,k}^{(i)}}{H} \int_{\partial K_i} \widetilde{\kappa} \phi_k^{(i)} v.$$

By the definition of the DG-norm, the error $\|u_1 - \widehat{u}_1\|_{\text{DG}}$ can be estimated by

$$\|\widehat{u}_1 - u_1\|_{\text{DG}}^2 \leq \sum_{i=1}^{N} \Big(\int_{K_i} \kappa |\nabla(\widehat{u}_1 - u_1)|^2 + \frac{\gamma}{h} \int_{\partial K_i} \widetilde{\kappa} (\widehat{u}_1 - u_1)^2 \Big).$$

Note that, by (3.18), we have

$$\int_{K_i} \kappa |\nabla(\widehat{u}_1 - u_1)|^2 = \sum_{l=l_{1,i}+1}^{M_i} \frac{\lambda_{1,l}^{(i)}}{H} c_l^2 \leq \frac{H}{\lambda_{1,l_{1,i}+1}^{(i)}} \sum_{l=l_{1,i}+1}^{M_i} (\frac{\lambda_{1,l}^{(i)}}{H})^2 c_l^2$$

and

$$\frac{1}{h} \int_{\partial K_i} \widetilde{\kappa} (\widehat{u}_1 - u_1))^2 = \frac{1}{h} \sum_{l=l_{1,i}+1}^{M_i} c_l^2 \leq \frac{H^2}{h(\lambda_{1,l_{1,i}+1}^{(i)})^2} \sum_{l=l_{1,i}+1}^{M_i} (\frac{\lambda_{1,l}^{(i)}}{H})^2 c_l^2.$$

Furthermore,

$$\sum_{l=l_{1,i}+1}^{M_i} (\frac{\lambda_{1,l}^{(i)}}{H})^2 c_l^2 \leq \sum_{l=1}^{M_i} (\frac{\lambda_{1,l}^{(i)}}{H})^2 c_l^2 = (\widetilde{\kappa})^{-1} \int_{\partial K_i} (\kappa \nabla u_1 \cdot n_{\partial K_i})^2.$$

Consequently, we obtain the following bound

$$\|u_1 - \widehat{u}_1\|_{\text{DG}}^2 \leq \sum_{i=1}^{N} \frac{H}{\widetilde{\kappa} \lambda_{1,l_{1,i}+1}^{(i)}} (1 + \frac{\gamma H}{h \lambda_{1,l_{1,i}+1}^{(i)}}) \int_{\partial K_i} (\kappa \nabla u_1 \cdot n_{\partial K})^2.$$

Next, we will find an estimate of $\|u_2 - \widehat{u}_2\|_{\text{DG}}$. By definition of the bilinear form a_{DG},

$$a_{DG}(u_2, v) = -a_{DG}(u_1, v) + (f, v) = \sum_{E \in \mathcal{E}^H} \int_E \Big(\{\kappa \nabla v \cdot n_E\} [\![u_1]\!] \Big) + (f, v)$$

which holds for any $v \in V_{2,\text{snap}}$. In addition, by the fact that any function in $V_{2,\text{snap}}$ is zero on boundaries of coarse grid blocks, we have

$$\|u_2 - \widehat{u}_2\|^2_{\text{DG}} = a_{DG}(u_2 - \widehat{u}_2, u_2 - \widehat{u}_2) = a_{DG}(u_2, u_2 - \widehat{u}_2),$$

where it follows from the fact that the eigenfunctions of (3.6) are κ-orthogonal on every coarse grid block. Therefore we have

$$\|u_2 - \widehat{u}_2\|^2_{\text{DG}} = \sum_{E \in \mathcal{E}^H} \int_E \Big(\{\kappa \nabla (u_2 - \widehat{u}_2) \cdot n_E\} [\![u_1]\!] \Big) + (f, u_2 - \widehat{u}_2). \tag{3.19}$$

The second term on the right hand side of (3.19) can be estimated as

$$(f, u_2 - \widehat{u}_2) \leq \sum_{K \in \mathcal{T}^H} \|f\|_{L^2(K)} \|\kappa^{\frac{1}{2}} (u_2 - \widehat{u}_2)\|_{L^2(K)}.$$

By (3.6), for every $K_i \in \mathcal{T}^H$, we have

$$\int_{K_i} \kappa |(u_2 - \widehat{u}_2)|^2 = \sum_{l \geq l_{2.i}+1} d_l^2 \leq \Big(\frac{H^2}{\lambda^{(i)}_{2,l_{2,i}+1}} \Big) \sum_{l = l_{2.i}+1} \frac{\lambda^{(i)}_{2,l}}{H^2} d_l^2 = \frac{H^2}{\lambda^{(i)}_{2,l_{2,i}+1}} \int_{K_i} \kappa |\nabla (u_2 - \widehat{u}_2)|^2.$$

For the first term on the right hand side of (3.19), we use inequality (3.13) to conclude that

$$\sum_{E \in \mathcal{E}^H} \int_E \Big(\{\kappa \nabla (u_2 - \widehat{u}_2) \cdot n_E\} [\![u_1]\!] \Big) \leq C_{\text{inv}} \gamma^{-\frac{1}{2}} \|u_2 - \widehat{u}_2\|_{\text{DG}} \Big(\sum_{E \in \mathcal{E}^H} \frac{\gamma}{h} \int_E \overline{\kappa} [\![u_1]\!]^2 \Big)^{\frac{1}{2}}.$$

Consequently, from (3.19) and the fact that $[\![u_1]\!] = [\![u_h]\!]$ for all coarse edges, we obtain the

following bound

$$\|u_2 - \widehat{u}_2\|_{\text{DG}}^2 \leq C\Big(C_{\text{inv}}^2 \sum_{E\in\mathcal{E}^H} \frac{1}{h}\int_E \overline{\kappa}[\![u_h]\!]^2 + \sum_{K\in\mathcal{T}_H} \frac{H^2}{\lambda_{2,l_{2,i}+1}^{(i)}}\|f\|_{L^2(K)}^2\Big).$$

Finally, we will prove the required error bound. By coercivity,

$$\begin{aligned} a_0\|\widehat{u}_1 + \widehat{u}_2 - u_H\|_{\text{DG}}^2 &\leq a_{\text{DG}}(\widehat{u}_1 + \widehat{u}_2 - u_H, \widehat{u}_1 + \widehat{u}_2 - u_H) \\ &= a_{\text{DG}}(\widehat{u}_1 + \widehat{u}_2 - u_H, \widehat{u}_1 + \widehat{u}_2 - u_h) \\ &\quad + a_{\text{DG}}(\widehat{u}_1 + \widehat{u}_2 - u_H, u_h - u_H). \end{aligned}$$

Note that $a_{\text{DG}}(\widehat{u}_1 + \widehat{u}_2 - u_H, u_h - u_H) = 0$ since $\widehat{u}_1 + \widehat{u}_2 - u_H \in V_{\text{off}}$. Using the above results,

$$\begin{aligned} &\|\widehat{u}_1 + \widehat{u}_2 - u_H\|_{\text{DG}}^2 \\ &\leq C\Big(\sum_{i=1}^{N} \frac{H}{\widetilde{\kappa}\lambda_{1,l_{1,i}+1}^{(i)}}(1 + \frac{\gamma H}{h\lambda_{1,l_{1,i}+1}^{(i)}})\int_{\partial K_i}(\kappa\nabla u_1 \cdot n_{\partial K})^2 \\ &\quad + \sum_{K\in\mathcal{T}_H} \frac{H^2}{\lambda_{2,l_{2,i}+1}^{(i)}}\|f\|_{L^2(K)}^2 + C_{\text{inv}}^2 \sum_{E\in\mathcal{E}^H} \frac{1}{h}\int_E \overline{\kappa}[\![u_h]\!]^2\Big). \end{aligned}$$

The desired bound is then obtained by the triangle inequality

$$\|u_h - u_H\|_{\text{DG}} \leq \|u_h - \widehat{u}\|_{\text{DG}} + \|\widehat{u} - u_H\|_{\text{DG}},$$

where $\widehat{u} = \widehat{u}_1 + \widehat{u}_2$. This completes the proof.

□

We remark that, the term

$$\sum_{i=1}^{N} \frac{H}{\widetilde{\kappa}\lambda_{1,l_{1,i}+1}^{(i)}}(1+\frac{\gamma H}{h\lambda_{1,l_{1,i}+1}^{(i)}})\int_{\partial K_i}(\kappa\nabla u_1\cdot n_{\partial K})^2 \tag{3.20}$$

corresponds to the error for the space $V_{1,\text{off}}$ and the term

$$\sum_{K\in\mathcal{T}_H}\frac{H^2}{\lambda_{2,l_{2,i}+1}^{(i)}}\|f\|_{L^2(K)}^2$$

corresponds to the error for the space $V_{2,\text{off}}$. Moreover, the term

$$C_{\text{inv}}^2\sum_{E\in\mathcal{E}^H}\frac{1}{h}\int_E\overline{\kappa}[\![u_h]\!]^2$$

is the error in the fine grid solution u_h. This is the irreducible error, and an estimate of this can be derived following standard DG frameworks. In particular, the fine grid solution u_h satisfies (3.16), which is the standard interior penalty discontinuous Galerkin (IPDG) method with the use of the DG space defined in (3.15). One can find the error analysis of this method, for example, in [44]. More precisely, one can show that $\sum_{E\in\mathcal{E}^H}\frac{1}{h}\int_E\overline{\kappa}[\![u_h]\!]^2 \leq C_{\text{err}}h^{2p}$ where p is the degree of the piecewise polynomial in $V_h(K)$. We note that the constant C_{err} depends on the derivatives of the true solution and can be large. However, we assume that the fine mesh size h is small enough to resolve all scales of the true solution so that this error term is small.

h	Λ_K^{snap}
1/48	1.0021e+03
1/96	1.0193e+03
1/192	1.3094e+03

h	Λ_K^{snap}
1/48	7.7650e+05
1/96	1.6569e+06
1/192	3.3254e+06

Table 3.1: Left: oversampling basis, Right: no-oversampling basis

Remark 3.1.3. *It is important to note that one can also replace (3.8) by*

$$a_H^K(\widehat{v},\widehat{v}) \leq \Lambda_K^{snap}\widetilde{\kappa}\int_{\partial K}|v|^2, \tag{3.21}$$

where Λ_K^{snap} is the largest eigenvalue for the spectral problem (3.5). Therefore, (3.9) can be replaced by

$$\int_{\partial K}|\kappa\nabla u\cdot n_{\partial K}|^2 \leq \Lambda_K^{snap}\widetilde{\kappa}\, a_H^K(u,u). \tag{3.22}$$

By following the above steps, we see that one can choose γ in (3.20) so that

$$\gamma > C_\kappa h \max_{K\subset\mathcal{T}^H}\Lambda_K^{snap}$$

where the constant C_κ is defined as

$$C_\kappa = \max_{K\subset\mathcal{T}^H}\frac{\max_{E\subset\partial K}\overline{\kappa}}{\min_{E\subset\partial K}\overline{\kappa}}.$$

We remark that this constant C_κ is order one if we assume that every coarse element has a high contrast region.

One can take smaller values of γ if oversampling is used (oversampling method is discussed in Section 3.3). The main idea of the oversampling is to choose larger regions for computing snapshot vectors. For every coarse block K_i, we choose an enlarged region K_i^+, and find oversampling snapshot functions $\psi_k^{i,over}$ by solving (3.34). We have performed numerical experiments and computed Λ_K^{snap} with and without oversampling. Denote $\Lambda_{K^+}^{snap}$ to be the largest eigenvalue corresponding to the oversampled problem. In our numerical results (see Table 3.1), we have removed linearly dependent snapshot vectors with respect to the inner product corresponding to $\int_{\partial K}|v|^2$ before computing the largest eigenvalue. Our numerical results show that one can have about three orders of magnitude

smaller value for $\Lambda^{snap}_{K^+}$ compared to Λ^{snap}_{K}. Moreover, our numerical results show a weak h-dependence for $\Lambda^{snap}_{K^+}$ as we decrease h, while Λ^{snap}_{K} behaves as h^{-1} (when no-oversampling is used).

Our error analysis holds when oversampling snapshot space is used. The term in (3.20) will become

$$\sum_{i=1}^{N} \frac{H}{\widetilde{\kappa}\lambda^{(i)}_{1,l_{1,i}+1}}(1+\frac{\alpha C_\kappa \max_{K\subset\mathcal{T}^H}\Lambda^{snap}_{K^+}H}{\lambda^{(i)}_{1,l_{1,i}+1}})\int_{\partial K_i}(\kappa\nabla u_1\cdot n_{\partial K})^2 \tag{3.23}$$

when $\gamma = \alpha C_\kappa h \max_{K\subset\mathcal{T}^H}\Lambda^{snap}_{K^+}$. If $\Lambda^{snap}_{K^+}$ is a weak function of h, e.g., if it is bounded with respect to h, then the terms involving $\Lambda^{snap}_{K^+}$ doesnt influence the error and the error is dominated by the first term. We emphasize that our discussions in this Remark are based on our numerical studies and their analytical studies are difficult because it requires interior estimates for solutions. We plan to study them in future.

Remark 3.1.4. *From the remark above as well as (3.21) and (3.22), one sees that the penalty term in (3.2) can be replaced by*

$$\beta\Lambda^{snap}_{max}\int_E \overline{\kappa}[\![u]\!][\![v]\!]$$

where β is an $O(1)$ constant and $\Lambda^{snap}_{max} := \max_{K\subset\mathcal{T}^H}\Lambda^{snap}_{K}$. Moreover, it is easy to show that the result in Theorem 3.1.2 becomes

$$\begin{aligned}\|u_h-u_H\|^2_{DG} \le C\Big(&\sum_{i=1}^{N}\frac{1}{\widetilde{\kappa}\lambda^{(i)}_{1,l_{1,i}+1}}(1+\frac{\beta\Lambda^{snap}_{max}}{\lambda^{(i)}_{1,l_{1,i}+1}})\int_{\partial K_i}(\kappa\nabla u_1\cdot n_{\partial K})^2\\ &+\sum_{K\in\mathcal{T}_H}\frac{H^2}{\lambda^{(i)}_{2,l_{2,i}+1}}\|f\|^2_{L^2(K)}+\sum_{E\in\mathcal{E}^H}\Lambda^{snap}_{max}\int_E\overline{\kappa}[\![u_h]\!]^2\Big).\end{aligned}$$

Remark 3.1.5. *We note that while the solution can be well approximated by using our*

basis functions within coarse elements, the jumps of the solutions across coarse edges may not be well controlled (particularly, for a small number of multiscale basis functions). One reason for this is because we use a constant penalty parameter $\gamma > 0$ and a very large penalty factor (scaled as h^{-1}). In general, the solution has multiscale structure on coarse edges, and the use of multiscale penalty (instead of constant penalty) is desirable and is important in order to capture the jumps of the solution (e.g., similar to HDG methods). Despite these limitations, the IPDG formulation performs well in practical applications. The method is easy to implement in parallel using unstructured coarse meshes and avoids overlaps.

3.2 A-posteriori error estimate and adaptive enrichment

In this section, we will derive an a-posteriori error indicator for the error $u_h - u_H$ in energy norm. We will then use the error indicator to develop an adaptive enrichment algorithm. The a-posteriori error indicator gives an estimate of the local error on the coarse grid blocks K_i, and we can then add basis functions to improve the solution. Our indicator consists of two components, which correspond to the errors made in the spaces $V_{1,\text{snap}}$ and $V_{2,\text{snap}}$. By using the indicator, one can determine adaptively which space has to be enriched. This section is devoted to the description of the a-posteriori error indicator and the corresponding adaptive enrichment algorithm. The convergence analysis of the method will be given in the Section 3.4.

Recall that V_{DG}^h is the fine scale DG finite element space, and the fine scale solution u_h satisfies

$$a_{\text{DG}}(u_h, v) = (f, v) \quad \text{for all } \ v \in V_{\text{DG}}^h. \tag{3.24}$$

Moreover, the GMsDGM solution u_H satisfies

$$a_{\text{DG}}(u_H, v) = (f, v) \quad \text{for all } \ v \in V_{\text{off}}. \tag{3.25}$$

We remark that $V_{\text{off}} \subset V_h^{\text{DG}}$. Next we will give the definitions of the residuals.

Definitions of residuals:

Let K_i be a given coarse grid block. We will define two residuals corresponding to the two types of snapshot spaces. First, on the space $V_{1,\text{snap}}^{(i)}$, we define the following linear functional

$$R_{1,i}(v) = \int_{K_i} fv - a_{\text{DG}}(u_H, v), \quad v \in V_{1,\text{snap}}^{(i)}. \tag{3.26}$$

Similarly, on the space $V_{2,\text{snap}}^{(i)}$, we define the following linear functional

$$R_{2,i}(v) = \int_{K_i} fv - a_{\text{DG}}(u_H, v), \quad v \in V_{2,\text{snap}}^{(i)}. \tag{3.27}$$

These residuals measure how well the solution u_H satisfies the fine-scale equation (3.24). Furthermore, on the snapshot spaces $V_{1,\text{snap}}^{(i)}$ and $V_{2,\text{snap}}^{(i)}$, we define the following norms

$$\|v\|_{V_1(K_i)}^2 = H^{-1} \int_{\partial K_i} \widetilde{\kappa} v^2 \quad \text{and} \quad \|v\|_{V_2(K_i)}^2 = H^{-2} \int_{K_i} \kappa v^2 \tag{3.28}$$

respectively. The norms of the linear functionals $R_{1,i}$ and $R_{2,i}$ are defined in the standard way, namely

$$\|R_{1,i}\| = \sup_{v \in V_{1,\text{snap}}^{(i)}} \frac{|R_{1,i}(v)|}{\|v\|_{V_1(K_i)}} \quad \text{and} \quad \|R_{2,i}\| = \sup_{v \in V_{2,\text{snap}}^{(i)}} \frac{|R_{2,i}(v)|}{\|v\|_{V_2(K_i)}}. \tag{3.29}$$

The norms $\|R_{1,i}\|$ and $\|R_{2,i}\|$ give estimates on the sizes of fine-scale residual errors with respect to the spaces $V_{1,\text{snap}}^{(i)}$ and $V_{2,\text{snap}}^{(i)}$.

We recall that, for each coarse grid block K_i, the eigenfunctions of the spectral problem (3.5) corresponding to the eigenvalues $\lambda_{1,1}^{(i)}, \cdots, \lambda_{1,l_{1,i}}^{(i)}$ and the eigenfunctions of the spectral problem (3.6) corresponding to the eigenvalues $\lambda_{2,1}^{(i)}, \cdots, \lambda_{2,l_{2,i}}^{(i)}$ are used in the construction of V_{off}. In addition, the energy error in this section and Section 3.4 is mea-

sured by $\|u\|_a^2 = a_{\text{DG}}(u,u)$, which is equivalent to the DG norm.

In Section 3.4, we will prove the following theorem, and we see that the norms $\|R_{j,i}\|$ give indications on the size of the energy norm error $\|u_h - u_H\|_a$.

Theorem 3.2.1. *Let u_h and u_H be the solutions of* (3.24) *and* (3.25) *respectively. Then*

$$\|u_h - u_H\|_a^2 \;\leq\; C_{err} \sum_{i=1}^{N} \sum_{j=1}^{2} \|R_{j,i}\|^2 (\lambda^{(i)}_{j,l_{j,i}+1})^{-1}. \tag{3.30}$$

where C_{err} is a uniform constant.

We will now present the adaptive enrichment algorithm. We use $m \geq 1$ to represent the enrichment level, $V_{\text{off}}(m)$ to represent the solution space at level m and u_H^m to represent the GMsDGM solution at the enrichment level m. For each coarse grid block K_i, we use $l_{j,i}^m$ to represent the number of eigenfunctions in $V_{j,\text{off}}^{(i)}$ used at the enrichment level m for the coarse region K_i. Assume that the initial offline space $V_{\text{off}}(0)$ is given.

Adaptive enrichment algorithm: Choose $0 < \theta < 1$. For each $m = 0, 1, \cdots$,

1. Step 1: Find the solution in the current space. That is, find $u_H^m \in V_{\text{off}}(m)$ such that

$$a_{\text{DG}}(u_H^m, v) = (f, v) \quad \text{for all } \; v \in V_{\text{off}}(m). \tag{3.31}$$

2. Step 2: Compute the local residuals. For each coarse grid block K_i, we compute

$$\eta_{j,i}^2 = \|R_{j,i}\|^2 (\lambda^{(i)}_{j,l_{j,i}^m+1})^{-1}, \quad j = 1, 2.$$

where the residuals and their norms are defined in (3.26)-(3.29). In particular, the residuals are the L^2 projections with respect to the inner products (3.28) and their norms are L^2 norms of these projections. After we computed the residual norms, we re-enumerate the above $2N$ residuals in the decreasing order, that is, $\eta_1^2 \geq \eta_2^2 \geq$

$\cdots \geq \eta_{2N}^2$, where we adopted single index notations. That is, each single index k corresponds to a double index (j,i) for the i-th coarse block K_i and the space $V_{j,\text{off}}^{(i)}$.

3. Step 3: Find the coarse grid blocks and spaces where enrichment is needed. We choose the smallest integer k such that

$$\theta \sum_{J=1}^{2N} \eta_J^2 \leq \sum_{J=1}^{k} \eta_J^2. \tag{3.32}$$

   The above is the standard Dorfler marking strategy. It is also a criterion to select the coarse regions with large errors. The numbers of regions and spaces to enrich are determined by the parameter θ.

4. Step 4: Enrich the space. For each $J = 1, 2, \cdots, k$, chosen by the above criterion, we add basis functions in $V_{j,\text{off}}^{(i)}$ according to the following rule. Let s be the smallest positive integer such that $\lambda_{j,l_{j,i}^m+s+1}^{(i)}$ is large enough (see the proof of Theorem 3.2.2) compared with $\lambda_{j,l_{j,i}^m+1}^{(i)}$. Then we will add the eigenfunctions $\phi_k^{(i)}$ or $\xi_k^{(i)}$ from the spectral problem (3.5) or (3.6) for $k = l_{j,i}^m+1, \cdots, l_{j,i}^m+s$, depending on the value of j, in the construction of the basis functions. We note that $j = 1$ or 2 corresponds to the space defined by eigenfunctions the spectral problem (3.5) or (3.6). The resulting space is denoted as $V_{\text{off}}(m+1)$.

We remark that the choice of s above will ensure the convergence of the enrichment algorithm, and in practice, the value of s is easy to obtain. Moreover, contrary to classical adaptive refinement methods, the total number of basis functions that we can add is bounded by the dimension of the snapshot space. Thus, the condition (3.32) can be modified as follows. We choose the smallest integer k such that

$$\theta \sum_{J=1}^{2N} \eta_J^2 \leq \sum_{J \in I} \eta_J^2,$$

where the index set I is a subset of $\{1, 2, \cdots, k\}$.

Finally, we state the convergence theorem.

Theorem 3.2.2. *There are sequence $\{L_m\}_{m=1}^M$ and positive constants τ, ϵ_0, ρ independent of m such that the following contracting property holds*

$$\|u_h - u_H^{m+1}\|_a^2 + \frac{\tau}{1+\tau L_{m+1}} \sum_{J=1}^{2N} (S_J^{m+1})^2 \leq \varepsilon \Big(\|u_h - u_H^m\|_a^2 + \frac{\tau}{1+\tau L_m} \sum_{J=1}^{2N} (S_J^m)^2 \Big).$$

Note that $0 < \varepsilon < 1$ and

$$\varepsilon = 1 - \epsilon_0 \theta^2 \frac{\tau}{2C_{\text{err}}(1+\tau L_M)}$$

where we can take any value of θ with the following condition

$$0 < \theta < \min\Big\{1, \Big(\frac{C_{\text{err}}(1+\tau L_M)}{\tau}\Big)^{\frac{1}{2}}\Big\}. \tag{3.33}$$

We remark that the precise definitions of S_J^m as well as the constants τ, ϵ_0, ρ and the sequence L_m are given in Section 3.4.

3.3 Numerical Results

In this section, we will present some numerical examples to demonstrate the performance of the adaptive enrichment algorithm. The domain Ω is taken as the unit square $[0,1]^2$ and is divided into 16×16 coarse blocks consisting of uniform squares. Each coarse block is then divided into 32×32 fine blocks consisting of uniform squares. Consequently, the whole domain is partitioned by a 512×512 fine grid blocks. We will use the following error quantities to compare the accuracy of our algorithm

$$e_2 = \frac{\|u_H - u_h\|_{L^2(\Omega)}}{\|u_h\|_{L^2(\Omega)}}, \; e_a = \frac{\sqrt{a_{\text{DG}}(u_H - u_h, u_H - u_h)}}{\sqrt{a_{\text{DG}}(u_h, u_h)}}$$

$$e_2^{\text{snap}} = \frac{\|u_H - u_{\text{snap}}\|_{L^2(\Omega)}}{\|u_{\text{snap}}\|_{L^2(\Omega)}},\ e_a^{\text{snap}} = \frac{\sqrt{a_{\text{DG}}(u_H - u_{\text{snap}}, u_H - u_{\text{snap}})}}{\sqrt{a_{\text{DG}}(u_{\text{snap}}, u_{\text{snap}})}}$$

where u_H and u_h are the GMsDGM and the fine grid solutions respectively. Moreover, u_{snap} is the snapshot solution obtained by using all snapshot functions generated by an oversampling strategy, see below.

We consider the permeability field κ which is shown in Figure 3.1. The boundary condition is set to be bi-linear, $g = x_1 x_2$. We will consider two examples with two different source functions f. We will compare the result of V_1 enrichment, $V_1 - V_2$ enrichment, oversampling basis enrichment, uniform enrichment and the exact indicator enrichment. The following gives the details of these enrichments.

- V_1 **enrichment**: We use the error indicator, $\eta_{1,i}^2$ to perform the adaptive algorithm by enriching the basis functions in V_1 space only, that is, basis functions obtained by the first spectral problem (3.5). We use 4 basis functions from (3.5) and zero basis function from (3.6) in the initial step.

- $V_1 - V_2$ **enrichment**: We use both the error indicators, $\eta_{1,i}^2, \eta_{2,i}^2$ to perform the adaptive algorithm by enriching the basis functions in both V_1 and V_2 spaces, that is, basis functions from both spectral problems (3.5) and (3.6). We use 4 basis functions from (3.5) and zero basis function from (3.6) in the initial step.

- **Oversampling enrichment:** For every coarse block K_i, we choose an enlarged region K_i^+ (in the examples presented below, we enlarge the coarse block in each direction by a length H, that is K_i^+ is a 3×3 coarse blocks with K_i at the center).

Then we find oversampling snapshot functions $\psi_k^{i,\text{over}} \in V_h(K_i^+)$ by solving

$$\begin{aligned} \int_{K_i} \kappa \nabla \psi_k^{i,\text{over}} \cdot \nabla v = 0, \quad \forall\, v \in V_h^0(K_i^+), \\ \psi_k^{i,\text{over}} = \delta_k, \quad \text{on } \partial K_i^+. \end{aligned} \tag{3.34}$$

The linear span of these snapshot functions is called $V^{i,\text{over}}$. Then we choose 40 dominant oversampling basis functions by POD method. Specifically, we solve the following eigenvalue problem

$$\int_{K_i} \psi_k^{i,\text{over}} v = \lambda_k^i \int_{\partial K_i^+} \psi_k^{i,\text{over}} v, \quad \forall v \in V^{i,\text{over}}.$$

and choose the first 40 eigenfunctions with largest eigenvalues. Then we use these 40 functions as boundary conditions in (3.4) and repeat the remaining construction of the offline space.

- **Uniform enrichment**: We enrich the basis functions in V_1 space uniformly with 4 basis functions from the V_1 space in the initial step.

- **The exact indicator enrichment**: We use the exact error as the error indicator to perform the adaptive algorithm by enriching the basis functions in V_1 space only with 4 basis functions in the space V_1 in the initial step. Here, the exact error is defined as $\|u - u_H\|_a$.

3.3.1 Example 1

In our first example, we take the source function $f = 1$. The fine grid solution is shown in Figure 3.1. In Table 3.2 and Table 3.3, we present the convergence history of our algorithm for enriching in V_1 space only, enriching in both V_1 and V_2 spaces and enriching by the oversampling basis functions. We remark that, in the presentation of our results, DOF

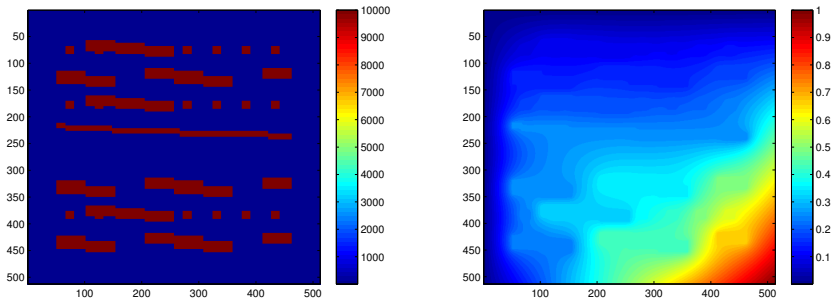

Figure 3.1: Left: Permeability field κ. Right: Fine grid solution with $f = 1$.

means the total number of basis functions used in the whole domain. We see from Table 3.2 that the behaviour of enriching in V_1 space only and enriching in both V_1 and V_2 spaces are similar. The is due to the fact that the source function f is a constant function, and the space V_2 will not help to improve the solution. This is in consistent with classical theory that basis functions obtained by harmonic extensions are good enough to approximate the solution. In Table 3.3, the convergence behaviour is shown for the oversampling case, and we see again that a clear convergence is obtained. For this case, we use 40 snapshot basis functions per coarse grid block giving a total DOF of 10240, and the corresponding snapshot errors (that is, the difference between the solution obtained by these 10240 basis functions and the solution u_h) of 4.5195×10^{-4} and 9.8935×10^{-4} in relative L^2 norm and relative a-norm respectively. In addition, we observe that the oversampling basis provides more efficient representation of the solution than the non-oversampling basis. To further demonstrate the efficiency of our algorithm, we compare our result with the uniform enrichment scheme. The convergence history for using uniform enrichment is shown in Table 3.3, and we see that our adaptive enrichment algorithm performs much better than uniform enrichment. Finally, a comparison among all the above cases and the enrichment by exact error is shown in Figure 3.2, in which the energy error is plotted against DOF. From the figure, we clearly see that our enrichment algorithm performs much better than uniform

enrichment. Moreover, our enrichment algorithm performs equally well compared with enrichment by the exact error. This shows that our indicator is both reliable and efficient.

DOF	e_2	e_a
1024	0.1082	0.0479
1769	0.0456	0.0178
2403	0.0156	0.0105
3135	0.0070	0.0067
5607	0.0016	0.0031

DOF	e_2	e_a
1024	0.1082	0.0479
1639	0.0802	0.0239
2584	0.0194	0.0114
3822	0.0061	0.0063
5660	0.0021	0.0037

Table 3.2: Convergence history with $\theta = 0.4$ for Example 1. Left: Enrich in V_1 space only. Right: Enrich in both V_1 and V_2 spaces.

DOF	e_2	e_a	e_2^{snap}	e_a^{snap}
1024	0.0940	0.0469	0.0939	0.0469
1975	0.0204	0.0121	0.0202	0.0121
2648	0.0087	0.0077	0.0084	0.0077
3422	0.0046	0.0056	0.0043	0.0056
6748	0.0009	0.0022	0.0006	0.0020

DOF	e_2	e_a
1024	0.1082	0.0479
2048	0.0671	0.0199
3328	0.0423	0.0150
5888	0.0161	0.0059
8448	0.0128	0.0044

Table 3.3: Left: Convergence history for oversampling basis with $\theta = 0.4$ for Example 1 and enrichment in V_1 space only. Right: Convergence history for uniform enrichment in V_1 space only.

Next, we will numerically demonstrate in Figure 3.3 the result of Theorem 3.1.2. In particular, we will show that the error decay is inversely proportional to the smallest eigenvalue corresponding to the un-used eigenfunctions. To better show this relation, we will consider basis function enrichment only in the V_1 space and we fix the number of basis in the V_2 space. Thus, we consider only the relation between error decay and the reciprocal of eigenvalues from the V_1 space. In the left graph of Figure 3.3, the decay in relative error

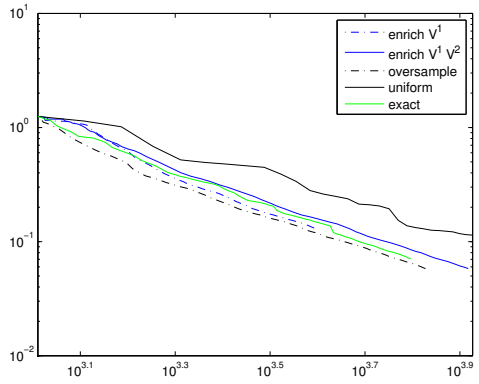

Figure 3.2: A comparison of different ways of enrichment for Example 1.

with respect to the number of basis functions per coarse element is shown, and in the right graph of Figure 3.3, the decay in reciprocal of eigenvalue with respect to the number of basis functions per coarse element is shown. We can see that the decay behaviours are very similar. Indeed, the correlation of these two curves is 0.9863, which confirms our result in Theorem 3.1.2. We remark that we conducted a similar test for the space V_2 and obtained the same conclusion.

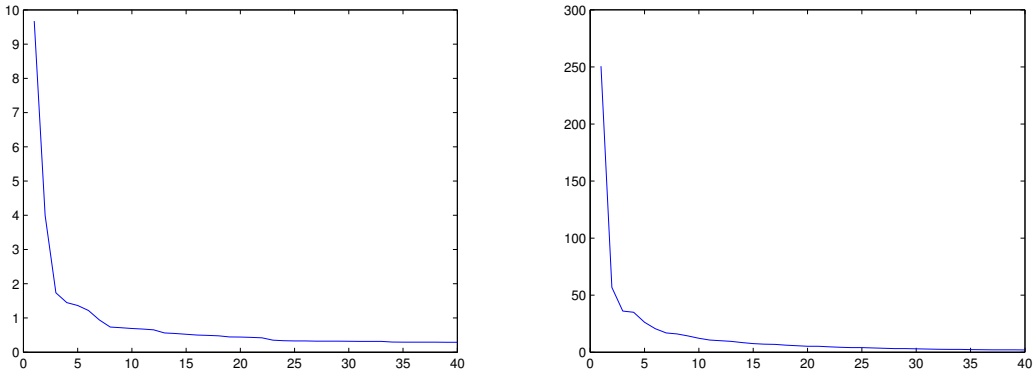

Figure 3.3: Relation of error and eigenvalue decays for Example 1. The correlation coefficient of these two curves is 0.9863. Left: Relative error decay. Right: Eigenvalue decay.

Finally, we will show numerically that the error indicator (3.26)-(3.27) gives a good approximation of the error. In Figure 3.4, we plot the ratio of error to residual with respect to the dimension of the approximation space, which contains both basis functions from V_1 and V_2. The error and the residual are the quantities appeared in (3.30). We see that the ratio remains bounded and $O(1)$, and is independent of the dimension of the approximation space. This shows that the constant in (3.30) is $O(1)$. This also shows that the residual gives a good indication of the level of error.

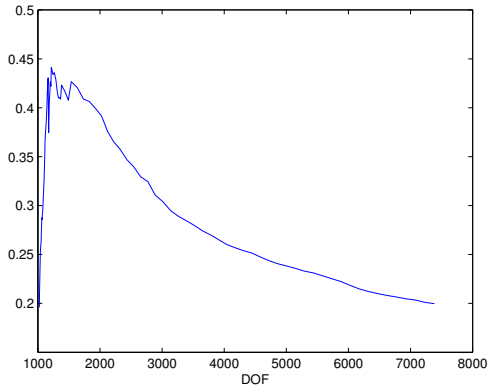

Figure 3.4: The ratio of error to residual with respect to the dimension of the approximation space for Example 1.

3.3.2 Example 2

In our second example, we will take the source f to be the function shown in the left plot of Figure 3.5 and the corresponding fine grid solution shown in the right plot of Figure 3.5. In Table 3.4 and Table 3.5, we present the convergence history of our algorithm for enriching in V_1 space only, enriching in both V_1 and V_2 spaces and enriching by oversampling basis. We see from Table 3.4 that enrichment in both V_1 and V_2 spaces provides much more efficient methods than enrichment in V_1 space only. In particular,

for an error level of approximately 1%, we see that enrichment in both V_1 and V_2 spaces requires 3144 DOF while enrichment in V_1 space only requires 3483 DOF. In Table 3.5, the convergence behaviour is shown for the oversampling case, and we see again that a clear convergence is obtained. For this case, we use 40 snapshot basis functions per coarse grid block giving a total DOF of 10240, and the corresponding snapshot errors (that is, the difference between the solution obtained by these 10240 basis functions and the solution u_h) of 0.0078 and 0.0093 in relative L^2 norm and relative a-norm respectively. In addition, we observe again that the oversampling basis provides more efficient representation of the solution than the non-oversampling basis. To further demonstrate the efficiency of our algorithm, we compare our results with the uniform enrichment scheme. The result for using uniform enrichment is shown in Table 3.5 and we clearly observe that our adaptive method is more efficient. Moreover, a comparison of the performance of various strategies is shown in Figure 3.6, where the errors against DOF are plotted. From the figure, we see that our method is much better than uniform enrichment. Furthermore, enrichment in both V_1 and V_2 spaces has the best performance, which suggests that both V_1 and V_2 spaces are important for more complicated source functions.

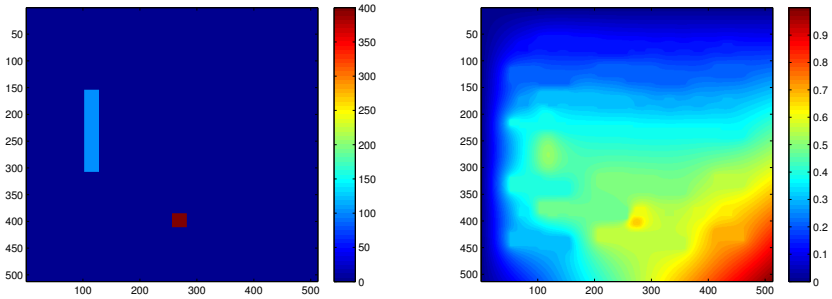

Figure 3.5: Left: The source function f for the second example. Right: The fine grid solution.

DOF	e_2	e_a
1024	0.2052	0.0554
2028	0.0362	0.0191
2717	0.0152	0.0140
3483	0.0111	0.0118
5116	0.0084	0.0102

DOF	e_2	e_a
1024	0.2052	0.0554
2023	0.0486	0.0206
3144	0.0113	0.0105
4456	0.0050	0.0066
7407	0.0013	0.0034

Table 3.4: Convergence history with $\theta = 0.4$ for Example 2. Left: Enrich in V_1 space only. Right: Enrich in both V_1 and V_2 spaces.

DOF	e_2	e_a	e_2^{snap}	e_a^{snap}
1024	0.1882	0.0540	0.1865	0.0532
1926	0.0296	0.0182	0.0269	0.0156
2626	0.0137	0.0135	0.0098	0.0098
3368	0.0105	0.0116	0.0057	0.0070
6677	0.0080	0.0097	0.0007	0.0025

DOF	e_2	e_a
1024	0.2052	0.0554
2048	0.0923	0.0282
3328	0.0659	0.0215
5888	0.0278	0.0135
8448	0.0226	0.0121

Table 3.5: Left: Convergence history for oversampling basis with $\theta = 0.4$ for Example 2 and enrichment in V_1 space only. Right: Convergence history for uniform enrichment in V_1 space only.

3.3.3 Adaptive enrichment algorithm

3.3.3.1 Adaptive enrichment algorithm with basis removal

In our adaptive enrichment algorithm, we can add basis functions to the offline space by using the error indicators. However, the addition of the basis functions must follow the ordering of the eigenfunctions. There may be cases that some of the intermediate eigenfunctions are not required in the representation of the solution. Therefore, we propose a numerical strategy to remove basis functions that do not contribute or contribute less to the representation of the solution. In the following, we will present this numerical strategy. Basically, the first 4 steps are the same as before. We introduce Step 5 below to remove unnecessary basis functions.

Adaptive enrichment algorithm with basis removal: Choose $0 < \theta < 1$. For each

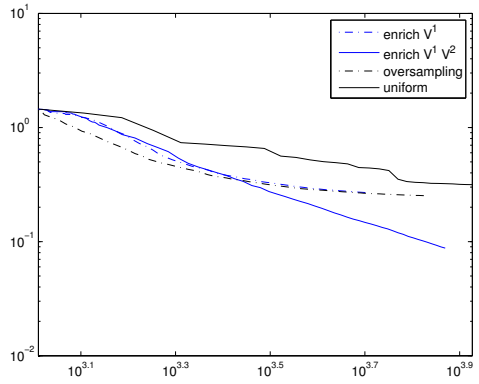

Figure 3.6: A comparison of different ways of enrichment.

$m = 0, 1, \cdots,$

1. Step 1: Find the solution in the current space. That is, find $u_H^m \in V_{\text{off}}(m)$ such that

$$a_{\text{DG}}(u_H^m, v) = (f, v) \quad \text{for all} \ \ v \in V_{\text{off}}(m). \tag{3.35}$$

2. Step 2: Compute the local residuals. For each coarse grid block K_i, we compute

$$\eta_{j,i}^2 = \|R_{j,i}\|^2 (\lambda_{j,l_{j,i}^m+1}^{(i)})^{-1}, \quad j = 1, 2.$$

   Then we re-enumerate the $2N$ residuals in the decreasing order, that is, $\eta_1^2 \geq \eta_2^2 \geq \cdots \geq \eta_{2N}^2$, where we adopted single index notations.

3. Step 3: Find the coarse grid blocks where enrichment is needed. We choose the smallest integer k such that

$$\theta \sum_{J=1}^{2N} \eta_J^2 \leq \sum_{J=1}^{k} \eta_J^2. \tag{3.36}$$

4. Step 4: Enrich the space. For each $J = 1, 2, \cdots, k$, we add basis function in $V_{j,\text{off}}^{(i)}$

according to the following rule. Let s be the smallest positive integer such that $\lambda^{(i)}_{j,l^m_{j,i}+s+1}$ is large enough (see the proof of Theorem 3.2.2) compared with $\lambda^{(i)}_{j,l^m_{j,i}+1}$. Then we include the eigenfunctions in the construction of the basis functions. The resulting space is denoted as $\widehat{V}^{\text{off}}(m+1)$. Note that this is the offline space without basis removal.

5. Step 5: Remove basis. In this step, we will remove basis functions that have little contribution to the solution. For each coarse grid block K_i, we can write the restriction of the current solution u^m_H on K_i as

$$\sum_{l=1}^{l^m_{1,i}} \alpha_{1,l}\phi^{(i)}_l + \sum_{l=1}^{l^m_{2,i}} \alpha_{2,l}\xi^{(i)}_l.$$

Fixed a tolerance $\varepsilon > 0$, which specify the importance of basis functions. Then the basis function $\phi^{(i)}_l$ or $\xi^{(i)}_l$ is removed from $\widehat{V}^{\text{off}}(m+1)$ if

$$\alpha^2_{1,l} < \varepsilon\Big(\sum_{l=1}^{l^m_{1,i}} \alpha^2_{1,l} + \sum_{l=1}^{l^m_{2,i}} \alpha^2_{2,l}\Big) \quad \text{or} \quad \alpha^2_{2,l} < \varepsilon\Big(\sum_{l=1}^{l^m_{1,i}} \alpha^2_{1,l} + \sum_{l=1}^{l^m_{2,i}} \alpha^2_{2,l}\Big)$$

is satisfied. The resulting space is called $V_{\text{off}}(m+1)$.

To test this strategy, we consider our second example with the source function f defined in Figure 3.5. We will consider three choices of ε, with values 10^{-12}, 10^{-13} and 10^{-14}. The convergence history of these cases are shown in Table 3.6. We can see that our basis removal strategy gives more efficient representation of the solution. For example, comparing the errors with DOF of around 2000 with basis removal (Table 3.6) and without basis removal (Table 3.4), we see that the method with basis removal gives a solution with smaller errors in both L^2 norm and a-norm. On the other hand, we see that the choice of $\varepsilon = 10^{-14}$ performs better than $\varepsilon = 10^{-12}$. In particular, for DOF of around 2200, the

error with $\varepsilon = 10^{-14}$ is around 2% while the error with $\varepsilon = 10^{-12}$ is around 4%. However, one expects that smaller choices of ε are not as economical as larger choices of ε.

DOF	L_2-error	a-error
1024	0.2052	0.0554
951	0.1824	0.0502
1074	0.1158	0.0415
1742	0.0461	0.0174
2218	0.0404	0.0153

DOF	L_2-error	a-error
1024	0.2052	0.0554
996	0.1767	0.0501
1107	0.1236	0.0431
2006	0.0266	0.0154
2824	0.0192	0.0123

DOF	L_2-error	a-error
1024	0.2052	0.0554
1048	0.1774	0.0500
1185	0.1280	0.0434
2235	0.0223	0.0150
3275	0.0147	0.0117

Table 3.6: Enrichment with $\theta = 0.4$ and basis removal as well as enrichment in V_1 space only. Left: $\varepsilon = 10^{-12}$. Middle: $\varepsilon = 10^{-13}$. Right: $\varepsilon = 10^{-14}$

3.3.3.2 Adaptive enrichment using local basis pursuit

In this section, we discuss an algorithm that follows basis pursuit ideas [40] and identify the basis functions which need to be added based on the residual. The main idea is to find multiscale basis functions that correlate to the residual the most and add those basis functions. More precisely, we identify basis functions that has the largest correlation coefficient with respect to the residual and add those basis functions. In the following, we will present the details of the numerical algorithm. The key components of the algorithm are Step 2 and Step 3, which determine the basis functions that correlate to the residual the most.

Adaptive enrichment algorithm using local basis pursuit: Choose $0 < \theta < 1$. For each $m = 0, 1, \cdots,$

1. Step 1: Find the solution in the current space. That is, find $u_H^m \in V_{\text{off}}(m)$ such that

$$a_{\text{DG}}(u_H^m, v) = (f, v) \quad \text{for all } \ v \in V_{\text{off}}(m). \tag{3.37}$$

2. Step 2: Compute the local residuals. For each coarse grid block K_i, we compute

$$\zeta_{j,i,l}^2 = \frac{|R_{j,i}(v_l)|^2}{\|v_l\|_{V_j(K_i)}^2}, \quad j = 1, 2, \quad \forall v_l \in V_j(K_i).$$

   Note that $|R_{j,i}(v_l)|$ is the inner-product that identifies the basis functions that have the largest correlation to the residual. More precisely,

$$|R_{j,i}(v_l)| = |\int_{K_i} f v_l - a_{DG}(u_H^m, v_l)|$$

   which is just the local inner-product of the residual vector and basis function v_l. We remark that v_l denotes generically the l-th basis functions in the space $V_j(K_i)$. The above quantity $\zeta_{j,i,l}^2$ has triple indices. But we re-enumerate these residuals using single indices in the decreasing order, that is, $\zeta_1^2 \geq \zeta_2^2 \geq \cdots$. More precisely, each single-index residual η_k^2 corresponds to a triple-index residual $\zeta_{j,i,l}^2$ for some (j, i, l).

3. Step 3: Find the coarse grid blocks where enrichment is needed. We choose the smallest integer k such that

$$\zeta_k \geq \theta \zeta_1. \tag{3.38}$$

   More precisely, if ζ_k^2 is selected by (3.38), then by Step 2, η_k^2 corresponds to a triple-index residual $\zeta_{j,i,l}^2$ for some (j, i, l). We will then add the basis function v_l in the space $V_j(K_i)$ to the new approximation space. (See Step 4).

4. Step 4: Enrich the space. For each $J = 1, 2, \cdots, k$, we add the basis function

$v_l \in V_j(K_i)$ corresponding to ζ_J. The resulting space is denoted as $\widehat{V}^{\text{off}}(m+1)$. Note that this is the offline space without basis removal.

5. Step 5: Remove basis. For each coarse grid block K_i, we can write the restriction of the current solution u_H^m on K_i as

$$\sum_{l=1}^{l_{1,i}^m} \alpha_{1,l}\phi_l^{(i)} + \sum_{l=1}^{l_{2,i}^m} \alpha_{2,l}\xi_l^{(i)}.$$

Fixed a tolerance $\varepsilon > 0$. Then the basis function $\phi_l^{(i)}$ or $\xi_l^{(i)}$ is removed if

$$\alpha_{1,l}^2 < \varepsilon\Big(\sum_{l=1}^{l_{1,i}^m} \alpha_{1,l}^2 + \sum_{l=1}^{l_{2,i}^m} \alpha_{2,l}^2\Big) \quad \text{or} \quad \alpha_{2,l}^2 < \varepsilon\Big(\sum_{l=1}^{l_{1,i}^m} \alpha_{1,l}^2 + \sum_{l=1}^{l_{2,i}^m} \alpha_{2,l}^2\Big)$$

is satisfied. The resulting space is called $V_{\text{off}}(m+1)$.

To demonstrate the performance of this strategy, we will consider two examples. In the first example, the source function f is defined as in Figure 3.5 and the rest of the parameters as in the Example 2. In the second example, we will take the solution (see Figure 3.7) which only contain the component of the 1st, 17th and 30th eigen-basis. The boundary conditions are as in Example 2 and the source term is calculated based on this sparse solution. The convergence history for the first example is shown in Table 3.7. Comparing these results to Table 3.6, we can see that the adaptive enrichment provides a better convergence. The convergence history is substantially improved if we consider the sparse solution as in our second example. The numerical results are shown in Table 3.8.

3.4 Convergence analysis

In this section, we will provide the proofs for the a-posteriori error estimates (Theorem 3.2.1) and the convergence of the adaptive enrichment algorithm (Theorem 3.2.2).

For each coarse grid block K_i, $i = 1, 2, \cdots, N$, we define two projection operators

DOF	e_2	e_a
1024	0.2052	0.0554
1036	0.1474	0.0405
1259	0.0585	0.0230
2096	0.0129	0.0125
2643	0.0099	0.0111

Table 3.7: Enrichment using basis pursuit with $\theta = 0.8$ and basis removal as well as enrichment in V_1 space only.

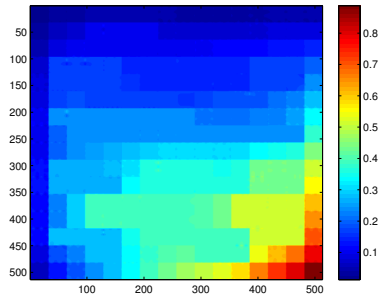

Figure 3.7: Solution with sparse coefficient

$P_{j,i} : V^{(i)}_{j,\text{snap}} \to V^{(i)}_{j,\text{off}}$, $j = 1, 2$, from the local snapshot spaces to the corresponding local offline spaces by

$$\begin{aligned} \int_{\partial K_i} \widetilde{\kappa} P_{1,i}(v) w &= \int_{\partial K_i} \widetilde{\kappa} v w \qquad \forall w \in V^{(i)}_{1,\text{off}}, \\ \int_{K_i} \kappa P_{2,i}(v) w &= \int_{K_i} \kappa v w \qquad \forall w \in V^{(i)}_{2,\text{off}}. \end{aligned}$$

For any $v \in V^{(i)}_{1,\text{snap}}$, we can write $v = \sum_{l=1}^{M_i} c_{1,l} \phi^{(i)}_l$. By orthogonality of eigenfunctions, we have $P_{1,i}(v) = \sum_{l=1}^{l_{1,i}} c_{1,l} \phi^{(i)}_l$. Therefore, by the equivalence of $\|\cdot\|_a$ and $\|\cdot\|_{\text{DG}}$, we have

$$\|P_{1,i}(v)\|_a^2 \leq a_1 \left(\int_{K_i} \kappa |\nabla P_{1,i}(v)|^2 + \frac{\gamma}{h} \int_{\partial K_i} \widetilde{\kappa} P_{1,i}(v)^2 \right).$$

DOF	e_2	e_a	DOF	e_2	e_a
1024	0.0150	0.0424	1024	0.0150	0.0424
941	0.0069	0.0286	997	0.0150	0.0424
934	0.0032	0.0135	1327	0.0108	0.0412
688	0.0001	0.0010	2447	0.0023	0.0154
744	1.20e-06	2.13e-05	889	0.0003	0.0022

Table 3.8: Enrichment with $\theta = 0.8$. Left: basis pursuit. Right: standard enrichment

By the spectral problem (3.5) and the fact that the eigenvalues are ordered increasingly, we have

$$\begin{aligned}\|P_{1,i}(v)\|_a^2 &\le a_1 \left(\sum_{l=1}^{l_{1,i}} \frac{\lambda_{1,l}^{(i)}}{H} c_{1,l}^2 + \frac{\gamma}{h} \int_{\partial K_i} \widetilde{\kappa} P_{1,i}(v)^2 \right) \\ &\le a_1 \left(\frac{\lambda_{1,l_{1,i}}^{(i)}}{H} + \frac{\gamma}{h} \right) \sum_{l=1}^{l_{1,i}} c_{1,l}^2 = a_1 \left(\lambda_{1,l_{1,i}}^{(i)} + \frac{\gamma H}{h} \right) \|v\|_{V_1(K_i)}^2.\end{aligned}$$

Similarly, for $v = \sum_{l\ge 1} c_{2,l}\xi_l^{(i)}$, we have $P_{2,i}(v) = \sum_{l=1}^{l_{2,i}} c_{2,l}\xi_l^{(i)}$. Therefore, by the equivalence of $\|\cdot\|_a$ and $\|\cdot\|_{\text{DG}}$, we have

$$\|P_{2,i}(v)\|_a^2 \le a_1 \left(\int_{K_i} \kappa |\nabla P_{2,i}(v)|^2 \right).$$

By the spectral problem (3.6) and the fact that the eigenvalues are ordered increasingly, we have

$$\|P_{2,i}(v)\|_a^2 \le a_1 \left(\sum_{l=1}^{l_{2,i}} \frac{\lambda_{2,l}^{(i)}}{H^2} c_{2,l}^2 \right) \le a_1 \left(\frac{\lambda_{2,l_{2,i}}^{(i)}}{H^2} \right) \sum_{l=1}^{l_{2,i}} c_{2,l}^2 = a_1 \lambda_{2,l_{2,i}}^{(i)} \|v\|_{V_2(K_i)}^2.$$

Thus the projections $P_{j,i}$ satisfy the following stability bound

$$||P_{j,i}(v)||_a \leq a_1^{\frac{1}{2}} \left(\lambda_{j,l_{j,i}}^{(i)} + \frac{\gamma H}{h} \right)^{\frac{1}{2}} \|v\|_{V_j(K_i)}, \quad j = 1, 2, \quad i = 1, 2, \cdots, N. \tag{3.39}$$

Next, we will establish some approximation properties for the projection operators $P_{j,i}$. Indeed, by the definitions of the operators $P_{j,i}$, for any $v \in V_{j,\text{snap}}^{(i)}$,

$$\begin{aligned} \|v - P_{j,i}(v)\|_{V_j(K_i)}^2 &= H^{-j} \sum_{l \geq l_{j,i}+1} c_{j,l}^2 \leq (\lambda_{l_{j,i}+1}^{(i)})^{-1} \sum_{l \geq l_{j,i}+1} \frac{\lambda_{j,l}^{(i)}}{H^j} c_{j,l}^2 \\ &= (\lambda_{j,l_{j,i}+1}^{(i)})^{-1} \int_{K_i} \kappa |\nabla v|^2, \end{aligned}$$

and therefore the following convergence result holds

$$\|v - P_{j,i}(v)\|_{V_j(K_i)} \leq \left(\lambda_{j,l_{j,i}+1}^{(i)} \right)^{-\frac{1}{2}} \Big(\int_{K_i} \kappa |\nabla v|^2 \Big)^{\frac{1}{2}}. \tag{3.40}$$

For the analysis presented below, we define the projection $\Pi : V_{\text{snap}} \to V_{\text{off}}$ by $\Pi v = \sum_{i=1}^{N} \sum_{j=1}^{2} P_{j,i}(v)$.

3.4.1 Proof of Theorem 3.2.1

Let $v \in V_h^{\text{DG}}$ be an arbitrary function in the space V_h^{DG}. Using (3.24), we have

$$a_{\text{DG}}(u_h - u_H, v) = a_{\text{DG}}(u_h, v) - a_{\text{DG}}(u_H, v) = (f, v) - a_{\text{DG}}(u_H, v).$$

Therefore,

$$\begin{aligned} a_{\text{DG}}(u_h - u_H, v) &= (f, v) - a_{\text{DG}}(u_H, v) \\ &= (f, v - \Pi v) + (f, \Pi v) - a_{\text{DG}}(u_H, \Pi v) - a_{\text{DG}}(u_H, v - \Pi v). \end{aligned}$$

Thus, using (3.25), we have

$$a_{\text{DG}}(u_h - u_H, v) = (f, v - \Pi v) - a_{\text{DG}}(u_H, v - \Pi v). \tag{3.41}$$

Since the space V_h^{DG} is the same as V_{snap}, we can write $v = \sum_{i=1}^{N} \sum_{j=1}^{2} v_j^{(i)}$ with $v_j^{(i)} \in V_{j,\text{snap}}^{(i)}$. Hence, (3.41) becomes

$$a_{\text{DG}}(u_h - u_H, v) = \sum_{i=1}^{N} \sum_{j=1}^{2} \Big(\int_{K_i} f(v_j^{(i)} - P_{j,i} v_j^{(i)}) - a_{\text{DG}}(u_H, v_j^{(i)} - P_{j,i} v_j^{(i)}) \Big). \tag{3.42}$$

We remark that, in the computation of the term $a_{\text{DG}}(u_H, v_j^{(i)} - P_{j,i} v_j^{(i)})$ in (3.42), we assume that the second argument is zero outside the coarse grid block K_i.

Using the definition of $R_{j,i}$, we see that (3.42) can be written as

$$a_{\text{DG}}(u_h - u_H, v) = \sum_{i=1}^{N} \sum_{j=1}^{2} R_{j,i}(v_j^{(i)} - P_{j,i} v_j^{(i)}).$$

Thus, we have

$$a_{\text{DG}}(u_h - u_H, v) \leq \sum_{i=1}^{N} \sum_{j=1}^{2} \|R_{j,i}\| \|(v_j^{(i)} - P_{j,i} v_j^{(i)})\|_{V_j(K_i)}.$$

Using (3.40),

$$a_{\mathrm{DG}}(u_h - u_H, v) \leq \sum_{i=1}^{N}\sum_{j=1}^{2} \|R_{j,i}\| \Big(\lambda^{(i)}_{j,l_{j,i}+1}\Big)^{-\frac{1}{2}} \Big(\int_{K_i} \kappa |\nabla v_j^{(i)}|^2 \Big)^{\frac{1}{2}}.$$

The inequality (3.30) is then followed by taking $v = u_h - u_H$ and

$$\sum_{i=1}^{N}\sum_{j=1}^{2} \int_{K_i} \kappa |\nabla v_j^{(i)}|^2 \leq \|v\|^2_{\mathrm{DG}} \leq a_0 \|v\|^2_a.$$

3.4.2 An auxiliary lemma

In this section, we will derive an auxiliary lemma which will be used for the proof of the convergence of the adaptive enrichment algorithm stated in Theorem 3.2.2. We use the notation $P^m_{j,i}$ to denote the projection operator $P_{j,i}$ at the enrichment level m.

In Theorem 3.2.1, we see that $\|R_{j,i}\|$ gives an upper bound of the energy error $\|u_h - u_H\|_a$. We will first show that, $\|R_{j,i}\|$ is also a lower bound up to a correction term (see Lemma 3.4.1). To state this precisely, we define

$$S^m_{j,i} = (\lambda^{(i)}_{j,l^m_{j,i}+1})^{-\frac{1}{2}} \sup_{v \in V^{(i)}_{j,\mathrm{snap}}} \frac{|R_{j,i}(v - P^m_{j,i}(v))|}{\|v\|_{V_j(K_i)}}. \tag{3.43}$$

Notice that the residual $R_{j,i}$ is computed using the solution u^m_H obtained at enrichment level m. We omit the index m in $R_{j,i}$ to simplify notations. Next, we will obtain

$$(S^m_{j,i})^2 = \|R_{j,i}\|^2 (\lambda^{(i)}_{j,l^m_{j,i}+1})^{-1}. \tag{3.44}$$

Indeed, by the fact that $P^m_{j,i}(v) \in V^{(i)}_{j,\mathrm{off}}$,

$$R_{j,i}(P^m_{j,i}(v)) = \int_{K_i} f P^m_{j,i}(v) - a_{\mathrm{DG}}(u^m_H, P^m_{j,i}(v)) = 0.$$

Thus,

$$S_{j,i}^m = (\lambda_{j,l_{j,i}^m+1}^{(i)})^{-\frac{1}{2}} \sup_{v \in V_{j,\text{snap}}^{(i)}} \frac{|R_{j,i}(v - P_{j,i}^m(v))|}{\|v\|_{V_j(K_i)}}$$
$$= (\lambda_{j,l_{j,i}^m+1}^{(i)})^{-\frac{1}{2}} \sup_{v \in V_{j,\text{snap}}^{(i)}} \frac{|R_{j,i}(v)|}{\|v\|_{V_j(K_i)}} = (\lambda_{j,l_{j,i}^m+1}^{(i)})^{-\frac{1}{2}} \|R_{j,i}\|.$$

This implies (3.44).

To prove Theorem 3.2.2, we will need the following recursive properties for $S_{j,i}^m$ (see Lemma 3.4.1). Notice that, the notation $\|u\|_{a,K_i}$ is defined as

$$\|u\|_{a,K_i}^2 = a_{\text{DG}}(u,u) = \int_{K_i} \kappa |\nabla u|^2 - \sum_{E \in \partial K_i} 2 \int_E \{\kappa \nabla u \cdot n_E\} [\![u]\!] + \sum_{E \in \partial K_i} \frac{\gamma}{h} \int_E \overline{\kappa} [\![u]\!]^2.$$

Lemma 3.4.1. *For any $\alpha > 0$, we have*

$$(S_{j,i}^{m+1})^2 \leq (1+\alpha) \frac{\lambda_{j,l_{j,i}^m+1}^{(i)}}{\lambda_{j,l_{j,i}^{m+1}+1}^{(i)}} (S_{j,i}^m)^2 + (1+\alpha^{-1}) a_1 D \|u_H^{m+1} - u_H^m\|_{a,K_i}^2 \tag{3.45}$$

where the enrichment level dependent constant D is defined by by

$$D = \left(\frac{\Lambda_{j,i}}{\lambda_{j,l_{j,i}^{m+1}+1}^{(i)}} + \frac{\gamma H}{h \lambda_{j,l_{j,i}^{m+1}+1}^{(i)}} \right)$$

with $\Lambda_{j,i} = \max_l \lambda_{j,l}^{(i)}$.

Proof. By direct calculations, we have

$$\begin{aligned}
&\int_{K_i} f(v - P_{j,i}^{m+1}(v)) - a_{\mathrm{DG}}(u_H^{m+1}, v - P_{j,i}^{m+1}(v)) \\
&= \int_{K_i} fv - a_{\mathrm{DG}}(u_H^{m+1}, v) \\
&= \int_{K_i} fv - a_{\mathrm{DG}}(u_H^{m}, v) + a_{\mathrm{DG}}(u_H^{m} - u_H^{m+1}, v) \\
&= \int_{K_i} f(v - P_{j,i}^{m}(v)) - a_{\mathrm{DG}}(u_H^{m}, v - P_{j,i}^{m}(v)) + a_{\mathrm{DG}}(u_H^{m} - u_H^{m+1}, v).
\end{aligned} \tag{3.46}$$

By definition of $S_{j,i}^m$, we have

$$S_{j,i}^m = (\lambda_{j,l_{j,i}^m+1}^{(i)})^{-\frac{1}{2}} \sup_{v \in V_{j,\mathrm{snap}}^{(i)}} \frac{|\int_{K_i} f(v - P_{j,i}^m(v)) - a_{\mathrm{DG}}(u_H^m, v - P_{j,i}^m(v))|}{\|v\|_{V_j(K_i)}}. \tag{3.47}$$

Multiplying (3.46) by $(\lambda_{j,l_{j,i}^m+1}^{(i)})^{-\frac{1}{2}} \|v\|_{V_j(K_i)}^{-1}$ and taking supremum with respect to v, we have

$$S_{j,i}^{m+1} \leq \Big(\frac{\lambda_{j,l_{j,i}^m+1}^{(i)}}{\lambda_{j,l_{j,i}^{m+1}+1}^{(i)}} \Big)^{\frac{1}{2}} S_{j,i}^m + I, \tag{3.48}$$

where

$$I = (\lambda_{j,l_{j,i}^m+1}^{(i)})^{-\frac{1}{2}} \sup_{v \in V_{j,\mathrm{snap}}^{(i)}} \frac{|a_{\mathrm{DG}}(u_H^m - u_H^{m+1}, v)|}{\|v\|_{V_j(K_i)}}.$$

To estimate I, we note that

$$a_{\mathrm{DG}}(u_H^m, P_{j,i}^m(v)) = \int_{K_i} f P_{j,i}^m(v) = a_{\mathrm{DG}}(u_H^{m+1}, P_{j,i}^m(v)).$$

Therefore, we have

$$a_{\mathrm{DG}}(u_H^m - u_H^{m+1}, v) = a_{\mathrm{DG}}(u_H^m - u_H^{m+1}, v - P_{j,i}^m(v)) \leq \|u_H^m - u_H^{m+1}\|_{a,K_i} \|v - P_{j,i}^m(v)\|_{a,K_i},$$

where we remark that v has value zero outside K_i. By the stability bound (3.39),

$$\|v - P_{j,i}^m(v)\|_a \leq a_1^{\frac{1}{2}} \left(\Lambda_{j,i} + \frac{\gamma H}{h}\right)^{\frac{1}{2}} \|v - P_{j,i}^m(v)\|_{V_j(K_i)} \leq a_1^{\frac{1}{2}} \left(\Lambda_{j,i} + \frac{\gamma H}{h}\right)^{\frac{1}{2}} \|v\|_{V_j(K_i)}.$$

Thus we have

$$I \leq a_1^{\frac{1}{2}} (\lambda_{j,l_{j,i}^m+1}^{(i)})^{-\frac{1}{2}} \left(\Lambda_{j,i} + \frac{\gamma H}{h}\right)^{\frac{1}{2}} \|u_H^{m+1} - u_H^m\|_{a,K_i}.$$

Using (3.48), we get

$$S_{j,i}^{m+1} \leq \Big(\frac{\lambda_{j,l_{j,i}^m+1}^{(i)}}{\lambda_{j,l_{j,i}^{m+1}+1}^{(i)}}\Big)^{\frac{1}{2}} S_{j,i}^m + a_1^{\frac{1}{2}} \left(\frac{\Lambda_{j,i}}{\lambda_{j,l_{j,i}^{m+1}+1}^{(i)}} + \frac{\gamma H}{h\lambda_{j,l_{j,i}^{m+1}+1}^{(i)}}\right)^{\frac{1}{2}} \|u_{\text{ms}}^{m+1} - u_{\text{ms}}^m\|_{a,K_i}.$$

Hence, (3.45) is proved. □

3.4.3 Proof of Theorem 3.2.2

In this section, we prove the convergence of the adaptive enrichment algorithm. First of all, we recall that

$$\eta_{j,i}^2 = \|R_{j,i}\|^2 (\lambda_{j,l_{j,i}^m+1}^{(i)})^{-1} = (S_{j,i}^m)^2. \tag{3.49}$$

We let the single index $J = (j,i)$, and we will use the single index notation η_J and S_J^m for $\eta_{j,i}$ and $S_{j,i}^m$ respectively. Note that there is an one-to-one mapping of the set of single indices J and the set of integers $\{1, 2 \cdots, 2N\}$.

Let $0 < \theta < 1$. We choose an index set I so that

$$\theta^2 \sum_{J=1}^{2N} \eta_J^2 \leq \sum_{J \in I} \eta_J^2. \tag{3.50}$$

We will then add basis function from $V^{(i)}_{j,\text{snap}}$ with $J = (j,i) \in I$. Then, using Theorem 3.2.1 and (3.50), we have

$$\theta^2 \|u_h - u_H^m\|_a^2 \leq \theta^2 C_{\text{err}} \sum_{J=1}^{2N} \eta_J^2 \leq C_{\text{err}} \sum_{J \in I} \eta_J^2.$$

By (3.49), we also have

$$\|u_h - u_H^m\|_a^2 \leq \frac{C_{\text{err}}}{\theta^2} \sum_{J \in I} (S_J^m)^2. \tag{3.51}$$

On the other hand,

$$\sum_{J=1}^{N} (S_J^{m+1})^2 = \sum_{J \in I} (S_J^{m+1})^2 + \sum_{J \notin I} (S_J^{m+1})^2.$$

By Lemma 3.4.1, for any $J \in I$, we have

$$(S_J^{m+1})^2 \leq (1+\alpha) \frac{\lambda^{(i)}_{j,l^m_{j,i}+1}}{\lambda^{(i)}_{j,l^{m+1}_{j,i}+1}} (S_J^m)^2 + (1+\alpha^{-1}) a_1 D \|u_H^{m+1} - u_H^m\|^2_{a,K_i}.$$

If $J \notin I$, then there is no enrichment involved. So $\lambda^{(i)}_{j,l^m_{j,i}+1} = \lambda^{(i)}_{j,l^{m+1}_{j,i}+1}$, which implies

$$(S_J^{m+1})^2 \leq (1+\alpha)(S_J^m)^2 + (1+\alpha^{-1}) a_1 D \|u_H^{m+1} - u_H^m\|^2_{a,K_i}.$$

Combining the above two cases, we have

$$\begin{aligned} \sum_{J=1}^{2N} (S_J^{m+1})^2 \;\leq\; & \sum_{J \in I} \left((1+\alpha) \frac{\lambda^{(i)}_{j,l^m_{j,i}+1}}{\lambda^{(i)}_{j,l^{m+1}_{j,i}+1}} (S_J^m)^2 + (1+\alpha^{-1}) a_1 D \|u_H^{m+1} - u_H^m\|^2_{a,K_i} \right) \\ & + \sum_{J \notin I} \left((1+\alpha)(S_J^m)^2 + (1+\alpha^{-1}) a_1 D \|u_H^{m+1} - u_H^m\|^2_{a,K_i} \right). \end{aligned}$$

We assume the enrichment is obtained so that

$$\delta := \max_{J\in I} \frac{\lambda^{(i)}_{j,l^m_{j,i}+1}}{\lambda^{(i)}_{j,l^{m+1}_{j,i}+1}} \leq \delta_0 < 1,$$

where δ_0 is independent of m. We then have

$$\begin{aligned}\sum_{J=1}^{2N}(S_J^{m+1})^2 \;\leq\; & \sum_{J\in I}\Big((1+\alpha)\delta_0(S_J^m)^2\Big) + \sum_{J\notin I}\Big((1+\alpha)(S_J^m)^2\Big) \\ & + \sum_{J=1}^{2N}(1+\alpha^{-1})a_1 D\|u_H^{m+1}-u_H^m\|^2_{a,K_i}.\end{aligned}$$

Since $\delta_0 = 1-(1-\delta_0)$, the above can be written as

$$\sum_{J=1}^{2N}(S_J^{m+1})^2 \leq (1+\alpha)\sum_{J=1}^{2N}(S_J^m)^2 - (1+\alpha)(1-\delta_0)\sum_{J\in I}(S_J^m)^2 + L_{m+1}\|u_H^{m+1}-u_H^m\|^2_a,$$

where

$$L_{m+1} = N_E(1+\alpha^{-1})a_1\Big(\max_{1\leq i\leq N}\max_{1\leq j\leq 2} D\Big), \tag{3.52}$$

where N_E is the maximum number of edges of coarse grid blocks, and we also emphasise that L_{m+1} depends on m. By (3.50),

$$\sum_{J=1}^{2N}(S_J^{m+1})^2 \leq (1+\alpha)\sum_{J=1}^{2N}(S_J^m)^2 - (1+\alpha)(1-\delta_0)\theta^2\sum_{J=1}^{2N}(S_J^m)^2 + L_{m+1}\|u_H^{m+1}-u_H^m\|^2_a.$$

Let $\rho = (1+\alpha)(1-(1-\delta_0)\theta^2)$. We choose $\alpha > 0$ small enough so that $0<\rho<1$. We remark that α is fixed if we fix the values of δ_0 and θ, which are user inputs. The above inequality can then be written as

$$\sum_{J=1}^{2N}(S_J^{m+1})^2 \leq \rho\sum_{J=1}^{2N}(S_J^m)^2 + L_{m+1}\|u_H^{m+1}-u_H^m\|^2_a. \tag{3.53}$$

Note that, by Galerkin orthogonality, we have

$$\|u_H^{m+1} - u_H^m\|_a^2 = \|u_h - u_H^m\|_a^2 - \|u_h - u_H^{m+1}\|_a^2. \tag{3.54}$$

We let $\tau > 0$ which will be specified in the following. Then, by using (3.53) and (3.54), we have

$$\|u_h - u_H^{m+1}\|_a^2 + \tau \sum_{J=1}^{2N} (S_J^{m+1})^2 \leq \|u_h - u_H^m\|_a^2 + \tau\rho \sum_{J=1}^{2N} (S_J^m)^2 + \tau L_{m+1} \|u_H^{m+1} - u_H^m\|_a^2.$$

Using (3.54), the above can be written as

$$\begin{aligned} \|u_h - u_H^{m+1}\|_a^2 + \tau \sum_{J=1}^{2N} (S_J^{m+1})^2 \quad &\leq \quad \|u_h - u_H^m\|_a^2 + \tau\rho \sum_{J=1}^{2N} (S_J^m)^2 \\ &\quad + \tau L_{m+1} \Big(\|u_h - u_H^m\|_a^2 - \|u_h - u_H^{m+1}\|_a^2 \Big) \end{aligned}$$

which implies

$$\|u_h - u_H^{m+1}\|_a^2 + \frac{\tau}{1 + \tau L_{m+1}} \sum_{J=1}^{2N} (S_J^{m+1})^2 \leq \|u_h - u_H^m\|_a^2 + \frac{\tau\rho}{1 + \tau L_{m+1}} \sum_{J=1}^{2N} (S_J^m)^2.$$

Next, we let $0 < \beta < 1$ which will be specified in the following. Using (3.51), we have

$$\|u_h - u_H^{m+1}\|_a^2 + \frac{\tau}{1 + \tau L_{m+1}} \sum_{J=1}^{2N} (S_J^{m+1})^2 \tag{3.55}$$

$$\leq \quad \Big(1 - \beta\Big) \|u_h - u_H^m\|_a^2 + \Big(\frac{\beta C_{\text{err}}}{\theta^2} + \frac{\tau\rho}{1 + \tau L_{m+1}} \Big) \sum_{J=1}^{2N} (S_J^m)^2. \tag{3.56}$$

We choose β by the following

$$\beta = \theta^2\Big(\theta^2 + C_{\text{err}}\tau^{-1}(1+\tau L_m)\Big)^{-1}\Big(1-\rho\frac{1+\tau L_m}{1+\tau L_{m+1}}\Big).$$

Then (3.55) becomes

$$\|u_h - u_H^{m+1}\|_a^2 + \frac{\tau}{1+\tau L_{m+1}}\sum_{J=1}^{2N}(S_J^{m+1})^2 \leq \Big(1-\beta\Big)\|u_h - u_H^m\|_a^2 + \frac{\tau(1-\beta)}{1+\tau L_m}\sum_{J=1}^{2N}(S_J^m)^2. \tag{3.57}$$

Finally, we will see that one can choose τ such that $\beta > 0$. Since $\{L_m\}$ is a decreasing sequence, we have

$$\frac{1+\tau L_m}{1+\tau L_{m+1}} \leq 1+\tau L_1.$$

Then we have

$$1-\rho\frac{1+\tau L_m}{1+\tau L_{m+1}} \geq \epsilon_0 > 0 \qquad \Longleftrightarrow \qquad \tau < \frac{1-\rho-\epsilon_0}{L_1}$$

where we take $\epsilon_0 < 1-\rho$. Using the above choice of τ, we see that $\beta > 0$. On the other hand, the condition $\beta < 1$ is obvious. To complete the proof, we let M be the maximum number of adaptive iterations, so that we have

$$L_M = a_1 N_E(1+\alpha^{-1}) \max_{1\leq i\leq N}\max_{1\leq j\leq 2}(1+\gamma H h^{-1}\Lambda_{j,i}^{-1}).$$

Using the condition (3.33), we obtain

$$\beta \geq \epsilon_0\theta^2\Big(C_{\text{err}}\tau^{-1}(1+\tau L_M) + C_{\text{err}}\tau^{-1}(1+\tau L_m)\Big)^{-1} \geq \epsilon_0\theta^2\Big(2C_{\text{err}}\tau^{-1}(1+\tau L_M)\Big)^{-1}$$

which gives the required convergence rate ϵ. This completes the proof of Theorem 3.2.2.

4. AN ONLINE GENERALIZED MULTISCALE DISCONTINUOUS GALERKIN METHOD (GMSDGM) FOR FLOWS IN HETEROGENEOUS MEDIA

4.1 Overview

In this chapter, we will discuss about the online basis construction for GMsDGM for the high-contrast flow problem (2.1). In order to obtain a good convergence result for the online adaptive enrichment method, we need to use different multiscale offline space as in the Chapter 3. Instead of using a different offline space, the offline stage of the method is basically the same as in Chapter 3. Therefore, we will use the same set of notations defined before.

Let we recall the framework of our GMsDGM. The methodology consists of two main ingredients, namely, the construction of local basis functions and the global coarse grid level coupling. For the coarse grid level coupling, we will apply the interior penalty discontinuous Galerkin (IPDG) method. That is, we find $u_H \in V_H$ such that

$$a_{\text{DG}}(u_H, v) = (f, v), \quad \forall v \in V_H, \tag{4.1}$$

where the bilinear form a_{DG} is defined as

$$a_{\text{DG}}(u, v) = a_H(u, v) - \sum_{E \in \mathcal{E}^H} \int_E \Big(\{\kappa \nabla u \cdot n_E\} [\![v]\!] + \{\kappa \nabla v \cdot n_E\} [\![u]\!] \Big) + \sum_{E \in \mathcal{E}^H} \frac{\gamma}{h} \int_E \overline{\kappa} [\![u]\!] [\![v]\!] \tag{4.2}$$

with

$$a_H(u, v) = \sum_{K \in \mathcal{T}_H} a_H^K(u, v), \quad a_H^K(u, v) = \int_K \kappa \nabla u \cdot \nabla v, \tag{4.3}$$

In the previous chapter, we already proved the continuity and coercivity of the bilinear form a_{DG} in DG-norm $\| \cdot \|_{DG}$. Therefore, this bilinear form is well pose for any offline

subspace $V_{\text{off}} \subset V_h^{DG}$.

One main result of the chapter is a convergence estimate of an adaptive procedure for the problem (4.1). For this purpose, we will compare the multiscale solution u_H to a fine-scale solution u_h defined in the following way. We first let

$$V_h^{\text{DG}} = \{v \in L^2(D) \,:\, v|_K \in V^h(K)\},$$

where $V^h(K)$ is the space of continuous piecewise bilinear functions defined on K with respect to the fine grid. The fine-scale solution $u_h \in V_{\text{DG}}^h$ is defined as the solution of the following

$$a_{\text{DG}}(u_h, v) = (f, v), \quad \forall v \in V_{\text{DG}}^h. \tag{4.4}$$

It is well-known that u_h gives a good approximation to the exact solution u up to a coarse grid discretization error.

The second main component of our method is the construction of local basis functions, which contains two stages, namely the offline stage and the online stage. In the offline stage, a snapshot space $V_{\text{snap}}^{(i)}$ is first constructed for each coarse grid block $K_i \in \mathcal{T}^H$. The snapshot space contains a rich space of basis functions, which can be used to approximate the fine-scale solution defined (4.4) with a good accuracy. A spectral problem is then solved in the snapshot space $V_{\text{snap}}^{(i)}$ and eigenfunctions corresponding to dominant modes are used as the basis functions. The resulting space is called the local offline space $V_{\text{off}}^{(i)}$ for the i-th coarse grid block K_i. The global offline space V_{off} is then defined as the linear span of all these $V_{\text{off}}^{(i)}$, for $i = 1, 2, \cdots, N$. This global offline space V_{off} will be used as the initial space of our method. We denote this initial space as $V_H^{(0)}$. Using the initial space, an initial solution $u_H^{(0)}$ can be computed by solving (4.1). Local residuals in coarse grid blocks can then be computed based on the initial solution $u_H^{(0)}$. In coarse grid blocks with

large residuals, new basis functions are computed and added to the approximation space. This procedure is continued until certain tolerance is reached. Next, we present a general outline of the method.

Assume that the initial space $V_H^{(0)}$ is given and the initial solution $u_H^{(0)}$ is computed. For any $m \geq 0$, we repeat the following until the solution $u_H^{(m)}$ satisfies certain tolerance requirement.

Step 1: Solve (4.1) using the space $V_H^{(m)}$ to obtain the solution $u_H^{(m)} \in V_H^{(m)}$.

Step 2: Compute local residuals based on the solution $u_H^{(m)}$.

Step 3: Construct new basis functions in regions, where the residuals are large.

Step 4: Add these basis functions to $V_H^{(m)}$ to form a new space $V^{(m+1)}$.

In the following, we will give the details of Step 2 and Step 3. We will also explain how one chooses the initial space $V_H^{(0)}$.

4.2 Locally online adaptivity

In this section, we will give details of our locally online adaptivity for the problem (4.1). As presented in the general outline of the method from the previous section, our adaptivity idea contains the choice of initial space as well as construction of new local multiscale basis functions. In the following, we will give the construction of these in detail.

4.2.1 Initial space

We present the definition of the initial space $V_H^{(0)}$. Let x_i be a node in the coarse grid $\mathcal{T}^H$, referred to as the i-th coarse node, for $i = 1, 2, \cdots, N_c$, where N_c is the number of nodes in the coarse grid $\mathcal{T}^H$. We will then define the i-th coarse neighbourhood ω_i as the union of all coarse grid blocks having the node x_i, see Figure 1.2. Moreover, for each

coarse grid block $K \in \mathcal{T}^H$, we let $\chi^K_{(j)}$, $j = 1, 2, 3, 4$, be the partition of unity functions, having value 1 at one vertex y_j and value 0 at the remaining three vertices, where y_j, $j = 1, 2, 3, 4$, are the four vertices of K. Note that there is exactly one value of j such that the vertex y_j is the same as the vertex x_i. In the case, we write $\chi^K_{(j)} = \chi^K_i$. One can use the standard multiscale basis functions or bilinear functions as the partition of unity functions. Note that we do not require any continuity of these partition of unity functions across coarse grid edges. The partition of unity functions are all supported on coarse grid blocks. Furthermore, we define the space $V^h(\omega_i)$ by

$$V^h(\omega_i) = \{v \in L^2(\omega_i) \, : \, v|_K \in V^h(K), \; K \in \mathcal{T}^H, \; K \subset \omega_i\}.$$

where $V^h(K)$ is the space of continuous piecewise bilinear functions defined on K. We remark that functions in $V^h(\omega_i)$ are supported in ω_i and belong to the space $V^h(K)$ for each coarse grid block $K \subset \omega_i$. Note that there is no continuity condition across boundaries of coarse grid blocks. We consider $V^h(\omega_i)$ as the snapshot space in ω_i, that is $V^{(i)}_{\text{snap}} = V^h(\omega_i)$, and perform a dimension reduction through a spectral problem. For this purpose, we define $\mathcal{E}^H_i$ be the set of coarse grid edges lying in the interior of ω_i, and the following bilinear form

$$a_{\omega_i}(u, v) = \sum_{K \in \mathcal{T}^H, K \subset \omega_i} a^K_H(u, v) + \sum_{E \in \mathcal{E}^H_i} \frac{\gamma}{h} \int_E \overline{\kappa} [\![u]\!][\![v]\!], \quad \forall u, v \in V^h(\omega_i). \tag{4.5}$$

Based on our analysis to be presented next, we solve the following spectral problem

$$a_{\omega_i}(u, v) = \lambda s_{\omega_i}(u, v), \quad \forall v \in V^h(\omega_i), \tag{4.6}$$

where

$$s_{\omega_i}(u,v) = \sum_{K\in\mathcal{T}^H, K\subset\omega_i} \int_K \kappa |\nabla \chi_i^K|^2 u\, v + \sum_{E\in\mathcal{E}_i^H} \frac{\gamma}{h}\int_E \overline{\kappa}[\![\chi_i^K]\!]^2\{u\}\,\{v\}, \quad \forall u, v \in V^h(\omega_i). \tag{4.7}$$

We use the notations $\lambda_k^{\omega_i}$ and $\Phi_k^{\omega_i}$ to denote the k-th eigenvalue and the k-th eigenvector of the above spectral problem (4.6). Each eigenfunction $\Phi_k^{\omega_i}$ corresponds to a function in $\phi_k^{\omega_i} \in V^h(\omega_i)$ defined by

$$\phi_k^{\omega_i} = \sum_{j=1}^{n_i} (\Phi_k^{\omega_i})_j w_j^{\omega_i}, \tag{4.8}$$

where n_i is the dimension of $V^h(\omega_i)$ and $\{w_j^{\omega_i}\}_{j=1}^{n_i}$ is a basis for $V^h(\omega_i)$. In the above definition, $(\Phi_k^{(i)})_j$ is the j-th component of the eigenvector $\Phi_k^{\omega_i}$.

For each ω_i, we solve the spectral problem (4.6) and the first L_i eigenfunctions are used to form the initial space. Each eigenfunction $\phi_k^{\omega_i}$ will be first multiplied by the partition of unity function χ_i^K, for each $K \subset \omega_i$, and is then decoupled across coarse grid edges to form 4 basis functions. In particular, the 4 new basis functions have support in one of the coarse grid block forming ω_i and are zero in the other three coarse grid blocks forming ω_i. For example, if $K \subset \omega_i$, the basis function is $\chi_i^K \phi_k^{\omega_i}$. We write $V_{\text{off}}^{(i)}$ as the space spanned by $I_K(\chi_i^K \phi_k^{\omega_i})$, for all $K \subset \omega_i$, where I_K is the standard bilinear interpolation operator. Therefore, $V_{\text{off}}^{(i)}$ is a subspace of V^h. See Figure 4.1 for an illustration. The initial space $V_H^{(0)}$ is obtained by the linear span of all functions constructed in the above procedure. Notice that $V_H^{(0)} \subset V_h^{\text{DG}}$.

4.2.2 Construction of online basis functions

In this section, we will discuss the construction of our local online basis functions. The purpose is to add basis functions locally in some coarse neighborhoods to obtain rapidly decaying errors. Assume that the space $V_H^{(m)}$ at the m-th iteration and the corresponding

solution $u_H^{(m)}$ are given. For each coarse neighborhood ω_i, we define the local residual by

$$R_i^{(m)}(v) = (f, v) - a_{\text{DG}}(u_H^{(m)}, v), \quad v \in V_0^h(\omega_i) \tag{4.9}$$

where $V_0^h(\omega_i) \subset V^h(\omega_i)$ contains functions that are zero on $\partial\omega_i$ and

$$V^h(\omega_i) = \{v \in L^2(\omega_i) \,:\, v|_K \in V^h(K),\ K \in \mathcal{T}^H,\ K \subset \omega_i\},$$

The residual $R_i^{(m)}$ can be seen as a linear functional defined on $V_0^h(\omega_i)$ with norm $\|R_i^{(m)}\|$ defined by

$$\|R_i^{(m)}\| = \sup_{v \in V_0^h(\omega_i)} \frac{|R_i^{(m)}(v)|}{\|v\|_{\omega_i}}, \tag{4.10}$$

and $\|v\|_{\omega_i}^2 = a_{\omega_i}(v, v)$. We will then find the new online basis function $\phi \in V_0^h(\omega_i)$ by solving

$$a_{\omega_i}(\phi, v) = R_i^{(m)}(v), \quad \forall v \in V_0^h(\omega_i). \tag{4.11}$$

The new online basis function ϕ is added to $V_H^{(m)}$ to form $V_H^{(m+1)}$.

The motivation of finding a new basis function ϕ by solving (4.11) can be explained as follows. We define the A-norm by

$$\|u\|_A^2 = a_{\text{DG}}(u, u), \quad \forall u \in V_H.$$

We notice that the A-norm is equivalent to the DG-norm $\|u\|_{\text{DG}}$ by Lemma 3.1.1. From (4.1) and (4.4), we have the following Galerkin orthogonality condition

$$a_{\text{DG}}(u_h - u_H^{(m)}, v) = 0, \quad \forall v \in V_H^{(m)}. \tag{4.12}$$

Thus, we see that the following optimal error bound holds

$$\|u_h - u_H^{(m)}\|_A^2 \leq \|u_h - \widetilde{u}\|_A^2, \quad \forall \widetilde{u} \in V_H^{(m)}. \tag{4.13}$$

Notice that (4.12) and (4.13) hold for any $m \geq 0$.

We will enrich the space $V_H^{(m)}$ by adding a basis function ϕ in the space $V_0^h(\omega_i)$ to form $V_H^{(m+1)}$. First, (4.13) implies

$$\|u_h - u_H^{(m+1)}\|_A^2 \leq \|u_h - \widetilde{u}\|_A^2, \quad \forall \widetilde{u} \in V_H^{(m+1)}. \tag{4.14}$$

Taking $\widetilde{u} = u_H^{(m)} + \alpha\phi$, for some scalar α, in (4.14), we have

$$\|u_h - u_H^{(m+1)}\|_A^2 \leq \|u_h - u_H^{(m)} - \alpha\phi\|_A^2$$

which implies

$$\|u_h - u_H^{(m+1)}\|_A^2 \leq \|u_h - u_H^{(m)}\|_A^2 - 2\alpha\, a_{\text{DG}}(u_h - u_H^{(m)}, \phi) + \alpha^2 \|\phi\|_A^2.$$

Taking $\alpha = a_{\text{DG}}(u_h - u_H^{(m)}, \phi)/\|\phi\|_A^2$, we obtain

$$\|u_h - u_H^{(m+1)}\|_A^2 \leq \|u_h - u_H^{(m)}\|_A^2 - \frac{a_{\text{DG}}(u_h - u_H^{(m)}, \phi)^2}{\|\phi\|_A^2}. \tag{4.15}$$

By the definition of the residual $R_i^{(m)}$ in (4.9), we see that (4.15) becomes

$$\|u_h - u_H^{(m+1)}\|_A^2 \leq \|u_h - u_H^{(m)}\|_A^2 - \frac{(R_i^{(m)}(\phi))^2}{\|\phi\|_A^2}. \tag{4.16}$$

From (4.16), we see that the quantity $(R_i^{(m)}(\phi))^2/\|\phi\|_A^2$ measures the amount of reduction in error when the basis function ϕ is added in $V_H^{(m)}$ to form $V_H^{(m+1)}$. We will construct the

function $\phi \in V_0^h(\omega_i)$ to obtain the most reduction in error. Thus, we find $\phi \in V_0^h(\omega_i)$ that maximizes $R_i^{(m)}(\phi)/\|\phi\|_A$. Equivalently, we find $\phi \in V_0^h(\omega_i)$ by solving

$$a_{\omega_i}(\phi, v) = R_i^{(m)}(v), \quad \forall v \in V_0^h(\omega_i).$$

Notice that, we have used the fact that $\|\phi\|_A = \|\phi\|_{\omega_i}$ when $\phi \in V_0^h(\omega_i)$.

Next, we summarize our online GMsDGM:

- *Initialization:* We will construct the initial space $V_H^{(0)}$. For each coarse neighborhood ω_i, we solve the following spectral problem in the space $V^h(\omega_i)$:

  $$a_{\omega_i}(u, v) = \lambda\, s_{\omega_i}(u, v).$$

  Then we obtain the eigenfunctions $\psi_k^{\omega_i}$ by (4.8), and we will consider the restriction of $\psi_k^{\omega_i}$ in each coarse element in ω_i as an individual function. Using the first L_i of these eigenfunctions, we obtain the basis functions $I_K(\chi_i^L \psi_k^{\omega_i})$.

- *Iteration:* For each $m \geq 0$, we perform the following iterations.

  1. Find $u_H^{(m)} \in V_H^{(m)}$ by solving

     $$a_{\text{DG}}(u_H^{(m)}, v) = (f, v), \quad \forall v \in V_H^{(m)}.$$

  2. Choose non-overlapping coarse neighborhoods ω_i and find new basis functions $\phi \in V_0^h(\omega_i)$ by

     $$a_{\omega_i}(\phi, v) = R_i^{(m)}(v), \quad \forall v \in V_0^h(\omega_i).$$

  3. Add those basis functions to $V_H^{(m)}$ to form $V_H^{(m+1)}$.

4.2.3 Convergence of the adaptive procedure

In this section, we analyze the convergence of the above online enrichment procedure. Before proving the convergence, we need the following lemma which will can be proved by standard techniques.

Lemma 4.2.1. *There exists a positive constant C_I such that*

$$\|I(v)\|_{\text{DG}} \leq C_I \|v\|_{\text{DG}}, \quad \forall v \in V_h^{DG} \tag{4.17}$$

where $I(v)|_K = I_K(v)$ for all $K \in \mathcal{T}^H$.

We begin the convergence analysis at the inequality (4.16). Notice that, this inequality can be written as

$$\|u_h - u_H^{(m+1)}\|_A^2 \leq \|u_h - u_H^{(m)}\|_A^2 - \|R_i^{(m)}\|^2, \tag{4.18}$$

when the basis function ϕ is obtained as in (4.11).

On the other hand, we will show that the error $\|u_h - u_H^{(m)}\|_A$ can be controlled by the residual norm $\|R_i^{(m)}\|$. To do so, we consider an arbitrary function $v \in V_h^{\text{DG}}$. Let $v_i \in V^h(\omega_i)$ be the restriction of v in ω_i, and let $v_i^{(0)} \in V_{\text{off}}^{(i)}$ be the component of v_i in the offline space $V_{\text{off}}^{(i)}$. By the GMsDGM (4.1), the fine-grid problem (4.4) and the Galerkin orthogonality (4.12), we have

$$a_{\text{DG}}(u_h - u_H^{(m)}, v) = a_{\text{DG}}(u_h - u_H^{(m)}, v - v^{(0)}), \quad \forall v^{(0)} \in V_H^{(0)},$$

where we define $v^{(0)} = \sum_{i=1}^{N_c} v_i^{(0)} \in V_H^{(0)}$ and use the fact that $V_H^{(0)} \subset V_H^{(m)}$ for all $m \geq 0$. By (4.4), we have

$$a_{\text{DG}}(u_h - u_H^{(m)}, v) = (f, v - v^{(0)}) - a_{\text{DG}}(u_H^{(m)}, v - v^{(0)}).$$

Using the property $\sum_{j=1}^{4} \chi_{(j)}^{K} = 1$ for all $K \in \mathcal{T}^H$ and the linearity of interpolation operator I,

$$a_{\text{DG}}(u_h - u_H^{(m)}, v) = \sum_{K \in \mathcal{T}^H} \sum_{j=1}^{4} \Big((f, I_K(\chi_{(j)}^K (v - v_i^{(0)}))) - a_{\text{DG}}(u_H^{(m)}, I_K(\chi_{(j)}^K (v - v_i^{(0)}))) \Big).$$

Writing the above sum over coarse neighborhoods ω_i, we have

$$a_{\text{DG}}(u_h - u_H^{(m)}, v) = \sum_{i=1}^{N_c} \sum_{K \subset \omega_i} \Big((f, I_K(\chi_i^K (v - v_i^{(0)}))) - a_{\text{DG}}(u_H^{(m)}, I_K(\chi_i^K (v - v_i^{(0)}))) \Big).$$

For each coarse neighborhood ω_i, we define the following modified local residual by

$$\widetilde{R}_i^{(m)}(v) = \sum_{K \subset \omega_i} \Big((f, I_K(\chi_i^K v)) - a_{\text{DG}}(u_H^{(m)}, I_K(\chi_i^K v)) \Big), \quad v \in V^h(\omega_i). \tag{4.19}$$

The modified residual $\widetilde{R}_i^{(m)}$ can be seen as a linear functional defined on $V^h(\omega_i)$ with norm $\|\widetilde{R}_i^{(m)}\|$ defined in the following way

$$\|\widetilde{R}_i^{(m)}\| = \sup_{v \in V^h(\omega_i)} \frac{|\widetilde{R}_i^{(m)}(v)|}{\| \sum_{K \subset \omega_i} I_K(\chi_i^K v)\|_{\omega_i}}$$

In the above definitions, χ_i^K is considered to be defined only on K, and has zero value outside K.

Using the definition of the modified residual $\widetilde{R}_i^{(m)}$, we have

$$a_{\text{DG}}(u_h - u_H^{(m)}, v) \leq \sum_{i=1}^{N_c} \|\widetilde{R}_i^{(m)}\| \, \| \sum_{K \subset \omega_i} I_K(\chi_i^K (v - v_i^{(0)}))\|_A \tag{4.20}$$

where we used the fact that $\sum_{K\subset\omega_i} I_K(\chi_i^K(v-v_i^{(0)}))$ is zero on $\partial\omega_i$. Using Lemma 3.1.1,

$$\|\sum_{K\subset\omega_i} I_K(\chi_i^K(v-v_i^{(0)}))\|_A \leq a_1^{\frac{1}{2}}\|\sum_{K\subset\omega_i} I_K(\chi_i^K(v-v_i^{(0)}))\|_{\mathrm{DG}}. \tag{4.21}$$

By the definition of the DG-norm and Lemma 4.2.1,

$$\begin{aligned}
\|I_K(\chi_i^K(v-v_i^{(0)}))\|_{\mathrm{DG}}^2 &\leq C_I^2\|\chi_i^K(v-v_i^{(0)})\|_{\mathrm{DG}}^2 && (4.22)\\
&= C_I^2\Big(\sum_{K\subset\omega_i}\int_K \kappa|\nabla(\chi_i^K(v-v_i^{(0)}))|^2 && (4.23)\\
&\quad +\frac{\gamma}{h}\sum_e\int_e \overline{\kappa}[\![\chi_i^K(v-v_i^{(0)})]\!]^2\Big). && (4.24)
\end{aligned}$$

For each $K\subset\omega_i$, we have

$$\int_K \kappa|\nabla(\chi_i^K(v-v_i^{(0)}))|^2 \leq 2\int_K \kappa\chi_i^2|\nabla(v-v_i^{(0)})|^2 + 2\int_K \kappa|\nabla\chi_i^K|^2(v-v_i^{(0)})^2. \tag{4.25}$$

For each $e\in\mathcal{E}_i^H$, we have

$$\int_e \overline{\kappa}[\![\chi_i^K(v-v_i^{(0)})]\!]^2 \leq 2\int_e \overline{\kappa}\{\chi_i^K\}^2[\![v-v_i^{(0)}]\!]^2 + 2\int_e \overline{\kappa}[\![\chi_i^K]\!]^2\{v-v_i^{(0)}\}^2. \tag{4.26}$$

Combining inequalities (4.25) and (4.26) in (4.24), we have

$$\begin{aligned}
&\|I_K(\chi_i^K(v-v_i^{(0)}))\|_{\mathrm{DG}}^2 && (4.27)\\
\leq\ & 2C_I^2\|v-v_i^{(0)}\|_{A_i}^2 && (4.28)\\
&+2C_I^2\Big(\int_{\omega_i}\overline{\kappa}|\nabla\chi_i^K|^2(v-v_i^{(0)})^2 + \frac{\gamma}{h}\sum_{e\in\mathcal{E}_i^H}\int_e \overline{\kappa}[\![\chi_i^K]\!]^2\{v-v_i^{(0)}\}^2\Big) && (4.29)
\end{aligned}$$

where $\|v\|^2_{A_i} = a_{\omega_i}(v,v)$. Using the spectral problem (4.6), we have

$$\|v - v_i^{(0)}\|^2_{A_i} = a_{\omega_i}(v_i - v_i^{(0)}, v_i - v_i^{(0)}) \leq a_{\omega_i}(v_i, v_i) = \|v\|^2_{A_i}$$

and

$$\begin{aligned} & \int_{\omega_i} \overline{\kappa} |\nabla \chi_i^K|^2 (v - v_i^{(0)})^2 + \frac{\gamma}{h} \sum_{e \in \mathcal{E}_i^H} \int_e \overline{\kappa} [\![\nabla \chi_i^K]\!]^2 \{v - v_i^{(0)}\}^2 \\ = \;& s_{\omega_i}(v_i - v_i^{(0)}, v_i - v_i^{(0)}) \leq \frac{1}{\lambda^{\omega_i}_{L_i+1}} \|v\|^2_{A_i}. \end{aligned}$$

Thus, (4.29) and (4.21) impleis

$$\|I_K(\chi_i^K(v - v_i^{(0)}))\|^2_A \leq 2a_1 C_I^2 \Big(1 + \frac{1}{\lambda^{\omega_i}_{L_i+1}}\Big) \|v\|^2_{A_i}.$$

Hence, (4.20) becomes

$$a_{\mathrm{DG}}(u_h - u_H^{(m)}, v) \leq \Big(\sum_{i=1}^{N_c} 2a_1 C_I^2 (1 + \frac{1}{\lambda^{\omega_i}_{L_i+1}}) \|\widetilde{R}_i^{(m)}\|^2 \Big)^{\frac{1}{2}} \Big(\sum_{i=1}^{N_c} \|v\|^2_{A_i} \Big)^{\frac{1}{2}}.$$

We remark that the above inequality holds for any $v \in V_h^{\mathrm{DG}}$. Taking $v = v_h - v_H^{(m)}$ and using Lemma 3.1.1, we finally obtain

$$\|u_h - u_H^{(m)}\|^2_A \leq 2a_0^{-1} a_1 C_0 C_I^2 \sum_{i=1}^{N_c} \Big(1 + \frac{1}{\lambda^{\omega_i}_{L_i+1}}\Big) \|\widetilde{R}_i^{(m)}\|^2, \tag{4.30}$$

where $C_0 = \max_{K \in \mathcal{T}^H} n_K$ and n_K is the number of vertices of the coarse grid block K.

We define

$$\theta = \|R_i^{(m)}\|^2 / \eta^2, \quad \text{and} \quad \eta^2 = 2a_0^{-1} a_1 C_0 C_I^2 \sum_{i=1}^{N_c} \Big(1 + \frac{1}{\lambda^{\omega_i}_{L_i+1}}\Big) \|\widetilde{R}_i^{(m)}\|^2. \tag{4.31}$$

From (4.18) and (4.30), we see that the following convergence holds

$$\|u_h - u_H^{(m+1)}\|_A^2 \leq (1-\theta)\|u_h - u_H^{(m)}\|_A^2.$$

We summarize the above results in the following theorem.

Theorem 4.2.2. *Let u_h be the solution of (4.4) and $u_H^{(m)}$, $m \geq 0$, be the solution of (4.1) in the m-th iteration. Then the following residual bound holds*

$$\|u_h - u_H^{(m)}\|_A^2 \leq 2a_0^{-1}a_1 C_0 C_I^2 \sum_{i=1}^{N_c} \Big(1 + \frac{1}{\lambda_{L_i+1}^{\omega_i}}\Big)\|\widetilde{R}_i^{(m)}\|^2. \tag{4.32}$$

Moreover, the following convergence holds

$$\|u_h - u_H^{(m+1)}\|_A^2 \leq (1-\theta)\|u_h - u_H^{(m)}\|_A^2 \tag{4.33}$$

where θ is defined in (4.31).

We remark that one can derive a priori error estimate for the error $\|u_h - u_H^{(m)}\|_{\text{DG}}$, for every $m \geq 0$. Since the purpose of this chapter is an a posteriori error estimate (4.32) and the convergence of an adaptive enrichment algorithm (4.33), we will not derive a priori error estimate.

Finally, we remark that by using more basis functions in the initial space $V_H^{(0)}$, the values of the eigenvalues $\lambda_{L_i+1}^{\omega_i}$ are larger. Thus, the value of θ is further away from zero, and this fact enhances the convergence rate. In particular, the convergence rate is affected by the quantity $\Lambda_{\text{min}} = \min_{1\leq i\leq N_c} \lambda_{L_i+1}^{\omega_i}$. The convergence is slow when Λ_{min} is small (cf. [45, 46]). We also remark that one can add online basis functions in multiple coarse neighborhoods to speed up the convergence. Let S be the index set for which online basis

functions are added in ω_i for $i \in S$. By using similar arguments as above, we obtain

$$\|u_h - u_H^{(m+1)}\|_A^2 \leq (1 - \widetilde{\theta})\|u_h - u_H^{(m)}\|_A^2$$

where

$$\widetilde{\theta} = \sum_{i \in S} \|R_i^{(m)}\|^2 / \eta^2.$$

4.3 Numerical Results

In this section, we will present some numerical examples to show the performance of the proposed method. The implementation procedure of online adaptive GMsDGM is described below. First, we choose a fixed number of functions for every coarse neighborhood by solving the local spectral problem. This fixed number for every coarse neighborhood is called the number of initial basis. After that, we split these functions into the basis functions of the offline space such that each basis function is supported in one coarse grid block. We denote this offline space as V_{off} and set $V_H^{(0)} = V_{\text{off}}$.

The coarse neighborhoods are denoted by $\omega_{i,j}$, where $i = 1, 2, \cdots, N_x$ and $j = 1, 2, \cdots, N_y$ and N_x and N_y are the number of coarse nodes in the x and y directions respectively. We consider $I_{x,\text{odd}}$ and $I_{x,\text{even}}$ as the set of odd and even indices from $\{1, 2, \cdots, N_x\}$. Similarly, $I_{y,\text{odd}}$ and $I_{y,\text{even}}$ are the set of odd and even indices from $\{1, 2, \cdots, N_y\}$. In each iteration of our online adaptive GMsDGM, we will perform 4 sub-iterations which add online basis functions in the non-overlapping coarse neighborhoods $\omega_{i,j}$ with $(i,j) \in I_{x,\text{odd}} \times I_{y,\text{odd}}$, $(i,j) \in I_{x,\text{odd}} \times I_{y,\text{even}}$, $(i,j) \in I_{x,\text{even}} \times I_{y,\text{odd}}$ and $(i,j) \in I_{x,\text{even}} \times I_{y,\text{even}}$ respectively.

We will take $\gamma = 2$ and $D = [0,1]^2$. The domain is divided into 10×10 uniform square coarse blocks. Each coarse block is then divided into 10×10 fine blocks consisting of uniform squares. Namely, the whole domain is partitioned by 100×100 fine grid blocks.

The medium parameter κ is shown in Figure 4.2. The source function f is taken as the constant 1. To compare the accuracy, we will use the following error quantities

$$e_2 = \frac{\|u_h - u_H\|_{L^2(D)}}{\|u_h\|_{L^2(D)}}, \quad \text{and} \quad e_a = \frac{\|u_h - u_H\|_{\text{DG}}}{\|u_h\|_{\text{DG}}}.$$

4.3.1 Comparison of using different number of initial basis

In Table 4.1, we present the convergence history of our algorithm for using one, two, three, four initial basis per coarse neighborhood. Notice that, in the presentation of our results, DOF means the total number of basis functions used in the whole domain and m denotes the number of iterations. We use the continuous multiscale basis functions as the initial partition of unity. In the tables, we obtain a fast error decay which give us a numerical solution with error smaller than 0.1% in two or three iterations. We can see the error decay of using one initial basis is slower than the error decay of using two or more initial basis since Λ_{min} for using one initial basis is too small.

m	DOF	e_a	e_2
0	324	44.50%	24.88%
1	648	9.92%	2.18%
2	972	0.78%	7.54e-2%
3	1296	3.24e-2%	2.13e-3%
4	1620	2.42e-4%	1.10e-5%

m	DOF	e_a	e_2
0	648	17.73%	3.58%
1	972	0.31%	1.80e-2%
2	1296	3.52e-3%	1.62e-4%
3	1620	1.81e-5%	8.58e-7%
4	1948	1.04e-7%	4.68e-9%

m	DOF	e_a	e_2
0	972	11.30%	1.72%
1	1296	0.45%	2.44e-2%
2	1620	3.05e-3%	1.37e-4%
3	1944	1.06e-5%	4.08e-7%
4	2240	4.59e-8%	2.14e-9%

m	DOF	e_a	e_2
0	1296	8.38%	1.00%
1	1620	7.98e-2%	3.13e-3%
2	1944	9.93e-4%	3.57e-5%
3	2268	1.39e-5%	5.15e-7%
4	2540	4.23e-8%	1.55e-9%

Table 4.1: Top-left: One initial basis ($\Lambda_{\text{min}} = 4.89e-4$). Top-right: Two initial basis ($\Lambda_{\text{min}} = 0.9504$). Bottom-left: Three initial basis ($\Lambda_{\text{min}} = 1.4226$). Bottom-right: Four initial basis ($\Lambda_{\text{min}} = 2.2045$).

To further study the importance of the initial basis, we will present another example with a different medium parameter κ shown in Figure 4.3. The domain D is divided into 5×5 coarse blocks consisting of uniform squares. Each coarse block is then divided into 40×40 fine blocks also consisting of uniform squares. The convergence history for the use of one, two, three, four initial basis and the corresponding total number of degrees of freedom (DOF) are shown in Table 4.2, Table 4.3, Table 4.4, Table 4.5 respectively. We consider two different contrasts. On the right table, we increase the contrast by 100 times. More precisely, the conductivity of inclusions and channels in Figure 2 (left figure) is multiplied by 100. In this case, the first 4 eigenvalue that are in the regions with channels become 100 times smaller. The decrease in the eigenvalues will slow down the error decay. In Table 4.2, we can observe that the error decay for the lower contrast case is much faster than the higher contrast case. In the higher contrast case, the error stop decreasing in some iterations. Similar observations are obtained when we use 2 or 3 initial basis. For using four initial basis, we observe a rapid convergence for both higher and lower contrast case.

m	DOF	e_a	e_2
0	64	25.44%	6.67%
1	128	1.20%	0.23%
2	192	0.47%	0.10%
3	256	0.26%	5.79e-2%
4	320	0.10%	2.30e-2%
5	384	6.22e-2%	1.02e-2%
6	448	3.70e-4%	1.57e-5%

m	DOF	e_a	e_2
0	64	25.45%	6.67%
1	128	1.45%	0.27%
2	192	1.39%	0.27%
3	256	0.84%	0.15%
4	320	0.34%	7.98e-2%
5	384	0.34%	7.91e-2%
6	448	0.15%	3.71e-2%

Table 4.2: One initial basis. Left: Lower contrast(1e4)($\Lambda_{\text{min}} = 0.0062$). Right: Higher contrast(1e6)($\Lambda_{\text{min}} = 6.22e-5$).

m	DOF	e_a	e_2
0	128	18.22%	4.42%
1	192	1.14%	0.12%
2	256	0.50%	4.95e-2%
3	320	4.17e-2%	2.06e-3%
4	384	5.73e-3%	5.38e-4%
5	448	7.12e-4%	2.89e-5%

m	DOF	e_a	e_2
0	128	18.56%	4.62%
1	192	1.37%	0.16%
2	256	1.25%	0.14%
3	320	1.23%	0.13%
4	384	0.41%	3.22e-2%
5	448	3.63e-2%	3.56e-3%

Table 4.3: Two initial basis. Left: Lower contrast(1e4)($\Lambda_{\text{min}} = 0.027$). Right: Higher contrast(1e6)($\Lambda_{\text{min}} = 2.72e-4$).

m	DOF	e_a	e_2
0	192	10.69%	1.86%
1	256	0.80%	6.66e-2%
2	320	0.34%	2.24e-2%
3	384	1.51e-2%	6.24e-4%
4	448	2.25e-4%	1.61e-5%
5	508	1.72e-6%	6.70e-8%

m	DOF	e_a	e_2
0	192	11.55%	2.14%
1	256	1.13%	0.10%
2	320	0.98%	8.85e-2%
3	384	0.96%	8.95e-2%
4	448	0.30%	1.39e-2%
5	508	2.00e-3%	8.39e-5%

Table 4.4: Three initial basis. Left: Lower contrast(1e4)($\Lambda_{\text{min}} = 0.0371$). Right: Higher contrast(1e6)($\Lambda_{\text{min}} = 3.75e-4$).

4.3.2 Setting tolerance for the residual

In this section, we will show the performance for the online enrichment implementing it only for regions with a residual error bigger than a certain threshold. We consider the medium parameter shown in Figure 4.2. We show the results for using three different tolerances (tol) 10^{-3}, 10^{-4} and 10^{-5}. We will enrich for the coarse regions with residual larger than the tolerance. In Table 4.6, we show the errors when using 1 initial basis function for tolerances 10^{-3}, 10^{-4} and 10^{-5}. We can see that the convergence history in the first few iteration is similar to the result shown in previous section. Moreover, the energy error of the multiscale solution is in the same order of the tolerance and the error will stop decreasing even if we perform more iterations. Therefore, we can compute a

m	DOF	e_a	e_2
0	248	7.92%	1.14%
1	312	0.25%	2.42e-2%
2	376	5.09e-3%	2.72e-4%
3	440	5.18e-5%	2.62e-6%
4	484	1.39e-6%	6.40e-8%

m	DOF	e_a	e_2
0	242	9.63%	1.59%
1	306	0.51%	5.40e-2%
2	370	1.38e-2%	9.46e-4%
3	434	2.10e-4%	1.59e-5%
4	494	1.74e-6%	1.27e-7%

Table 4.5: Four initial basis. Left: Lower contrast(1e4)($\Lambda_{\text{min}} = 0.4472$). Right: Higher contrast(1e6)($\Lambda_{\text{min}} = 0.3844$).

multiscale solution with a prescribed error level by choosing a suitable tolerance in the adaptive algorithm. In Table 4.7 and Table 4.8, we show the errors for the last three iterations when using 2 and 3 initial basis functions respectively for tolerances 10^{-3}, 10^{-4} and 10^{-5}. We have the same observation that the energy errors have the same magnitude as the tolerances.

m	DOF	e_a	e_2
0	324	44.50%	24.88%
1	648	9.92%	2.18%
2	924	0.81%	7.72e-2%
3	976	0.29%	2.49e-2%

m	DOF	e_a	e_2
0	324	44.50%	24.88%
1	648	9.92%	2.18%
2	972	0.78%	7.54e-2%
3	1176	4.12e-2%	2.88e-3%
4	1184	2.65e-2%	1.57e-3%

m	DOF	e_a	e_2
0	324	44.50%	24.88%
1	648	9.92%	2.18%
2	972	0.78%	7.54e-2%
3	1284	3.24e-2%	2.13e-3%
4	1364	2.56e-3%	1.55e-4%

Table 4.6: One initial basis. Left: $tol = 10^{-3}$. Middle: $tol = 10^{-4}$. Right: $tol = 10^{-5}$.

m	DOF	e_a	e_2
0	648	17.73%	3.58%
1	964	0.33%	1.85e-2%
2	972	0.30%	1.63e-2%

m	DOF	e_a	e_2
0	648	17.73%	3.58%
1	972	0.31%	1.80e-2%
2	1136	2.53e-2%	1.24e-3%

m	DOF	e_a	e_2
0	972	0.31%	1.80e-2%
1	1248	3.99e-3%	1.85e-4%
2	1276	2.49e-3%	1.19e-4%

Table 4.7: Two initial basis. Left: $tol = 10^{-3}$. Middle: $tol = 10^{-4}$. Right: $tol = 10^{-5}$.

m	DOF	e_a	e_2
0	972	11.30%	1.72%
1	1248	0.50%	2.57e-2%
2	1276	0.24%	9.98e-3%

m	DOF	e_a	e_2
0	972	11.30%	1.72%
1	1296	0.45%	2.44e-2%
2	1436	2.60e-2%	9.70e-4%

m	DOF	e_a	e_2
0	1296	0.45%	2.44e-2%
1	1564	3.52e-3%	1.56e-4%
2	1576	2.49e-3%	1.04e-4%

Table 4.8: Three initial basis. Left: $tol = 10^{-3}$. Middle: $tol = 10^{-4}$. Right: $tol = 10^{-5}$.

4.3.3 Adaptive online enrichment

In this section, we will show the performance for the online enrichment implementing it only for regions that have a cumulative residual that is θ fraction of the total residual. We consider the medium parameter shown in Figure 4.3 (4 channels medium).

Assume that the local residuals are arranged such that

$$r_1 \geq r_2 \geq r_3 \geq \cdots$$

where r_i is the norm of the local residual $\|R_i\|$. We only add the basis $\phi_1, \cdots, \phi_k$ for the

coarse neighborhoods $\omega_1, \cdots, \omega_k$ such that k is the smallest integer with

$$\theta \sum_{i=1}^{N_c} r_i^2 \leq \sum_{i=1}^{k} r_i^2.$$

We will also consider two ways to compute the local residual r_i, one is based on L_2 norm and the other is based on an energy norm. For the L_2 norm based approach, we will compute $r_i = \|R_i\|$ where the norm is the L_2 norm. This is an inexpensive way to compute r_i. The energy norm based residual is to compute $r_i = \|R_i\|$ using an energy norm, which is defined in (4.10). In particular, we will solve a local problem to compute this norm. This requires a larger computational cost, but it provides a better solution.

In Table 4.3.3, we present the error for the last 5 iterations when using 1 initial basis functions with the tolerance 10^{-5} and $\theta = 0.5$. We also present a comparison of using the L_2 based residual and the energy norm based residual. Comparing the result to the previous case, we can observe that this can use less number of basis functions to achieve a similar error. We also see that the energy norm based residual produces better solutions. In Figure 4.4, we present the distribution of number of basis functions in coarse blocks, and see that the number of basis functions is larger near the channels (c.f. Figure 4.3). Thus, online basis functions can be adaptively added in some regions using an error indicator.

m	DOF	e_a	e_2
0	336	0.35%	3.80e-2%
1	352	0.27%	1.74e-2%
2	368	7.71e-3%	5.88e-4%
3	384	2.47e-3%	2.15e-4%
4	396	8.77e-4%	6.94e-5%

m	DOF	e_a	e_2
0	348	0.16%	5.19e-2%
1	368	0.10%	4.03e-2%
2	392	7.13e-2%	9.34e-3%
3	412	6.04e-3%	6.60e-4%
4	424	1.51e-3%	1.25e-4%

Table 4.9: The results using cumulative errors with $\theta = 0.5$, $tol = 10^{-5}$ and 1 initial basis. Left: using energy norm based residual. Right: using L_2 norm based residual.

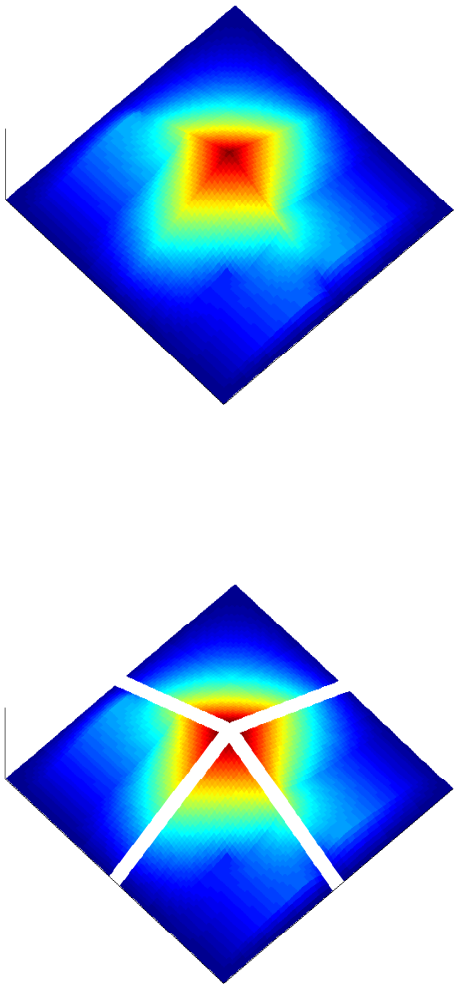

Figure 4.1: Illustration of the initial basis construction. Left: An eigenfunction $\chi_i^K \psi_k^{\omega_i}$ is defined in ω_i. Right: 4 basis functions are obtained by splitting $\chi_i^K \psi_k^{\omega_i}$ into 4 pieces, and each has support in $K \subset \omega_i$.

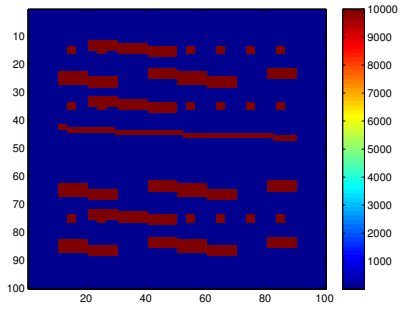

Figure 4.2: Permeability field κ for first case.

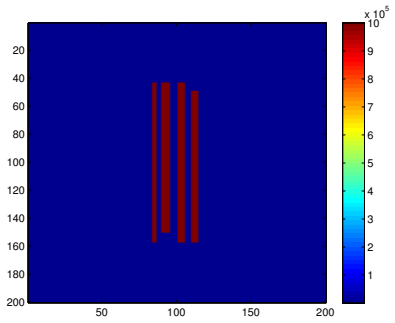

Figure 4.3: Permeability field κ for second case.

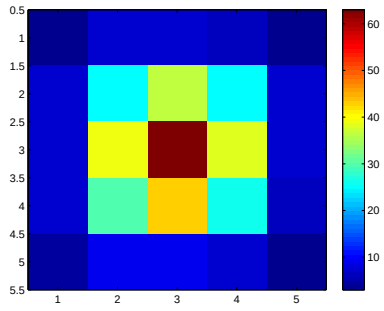

Figure 4.4: Distribution of number of basis functions in coarse blocks.

5. GENERALIZED MULTISCALE FINITE ELEMENT METHODS FOR WAVE PROPAGATION IN HETEROGENEOUS MEDIA

5.1 Overview

Numerical modeling of wave propagation is important in many applications that include geophysics, material science, and so on. For example, in geophysics applications, wave propagation simulations play an important role in determining subsurface properties [47, 48, 49, 50, 51, 52]. These approaches include finite difference methods, finite element methods, and spectral methods that use polynomials basis [53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64]. While these methods have different strengths and weaknesses, all of them tend to have limitations associated with discretization, especially in 3-D applications as frequency content of the simulated wavefield increases. Though the solutions to the wave equation have been shown to be accurate when the grid is fine enough [65], the practical limitations in discretization caused by limitations in computational power restrict this accuracy. An example of an application where this may be important is in the modeling of fractured media, where establishing reliable and accurate relationships between the properties of reflected seismic wavefields and variations in the density, orientation and compliance of fractures may help provide important constraints for hydrocarbon production. While more general finite element and spectral element methods may be able to address some problems by adapting grids to conform to heterogeneous structures, there are basic limitations associated with representing fine-scale features, and there is therefore a need to find approaches that reliably and accurately incorporate fine-scale features in a coarsely gridded model.

In this chapter, we present GMsDGM for wave propagation simulations on a coarse grid. We will focus our discussions on two-dimensional problems. The methodology can be applied to three-dimensional problems without significant conceptual difference. Let $\Omega \subset \mathbb{R}^2$ be a bounded domain of two dimensions. The aim is to develop a multiscale method for the following wave equation

$$\frac{\partial^2 u}{\partial t^2} = \nabla \cdot (a \nabla u) + f \qquad \text{in} \quad [0,T] \times \Omega \tag{5.1}$$

with the homogeneous Dirichlet boundary condition $u = 0$ on $[0,T] \times \partial\Omega$. The function $f(x,t)$ is a given source. The problem (5.1) is supplemented with the following initial conditions

$$u(x,0) = g_0(x), \qquad u_t(x,0) = g_1(x).$$

We assume that the coefficient $a(x)$ is highly oscillatory, representing the complicated model in which the waves are simulated. It is well-known that solving (5.1) by standard methods requires a very fine mesh, which is computationally prohibited. Thus a coarse grid solution strategy is needed. Next we present our fine scale solver. The fine scale solution is considered as the exact solution when we discuss the convergence of our multiscale method in the following sections. We assume that the domain Ω is partitioned by a set of rectangles, called fine mesh, with maximum side length $h > 0$. We denote the resulting mesh by $\mathcal{T}^h$ and the set of all edges and vertices by $\mathcal{E}^h$ and $\mathcal{N}^h$ respectively. We assume that the fine-mesh discretization of the wave equation provides an accurate approximation of the solution. The fine scale solver is the standard conforming bilinear finite element method. Let V_h be the standard conforming piecewise bilinear finite element space. We find $u_h \in V_h$ such that

$$(\frac{\partial^2 u_h}{\partial t^2}, v) + a(u_h, v) = (f, v), \quad \forall v \in V_h, \tag{5.2}$$

where the bilinear form a is defined by

$$a(u,v) = \int_{\Omega} a\nabla u \cdot \nabla v, \quad \forall u, v \in V_h \tag{5.3}$$

and $(\cdot,\cdot)$ represents the standard L^2 inner product defined on Ω.

The numerical results are presented for several representative examples. We investigate the GMsDGM's accuracy and, in particular, how choosing modes from different snapshot spaces can affect the accuracy. Our numerical results show that choosing the basis functions from interior modes can improve the accuracy of GMsDGM substantially for wave equations. These results differ from those we observe for flow equations [27].

The chapter is organized as follows. In Section 5.2, we will present the multiscale method. Numerical results are shown in Section 5.3. Stability and spectral convergence of the semi-discrete scheme are proved in Section 5.4. In Section 5.5, the convergence of the fully-discrete scheme is also proved. Finally, conclusions are presented.

5.2 The generalized multiscale discontinuous Galerkin method

In this section, we will give a detailed description of our generalized multiscale discontinuous Galerkin method for wave equation. The method gives a numerical solver on a coarse grid, providing an efficient way to simulate waves in complicated media.

5.2.1 Global IPDG solver

We will apply the standard symmetric IPDG approach as in Chapter 3 to solve (5.1) on the coarse grid $\mathcal{T}^H$. The method follows the standard framework as discussed in [67, 44], but the finite element space will be replaced by the space spanned by our multiscale basis functions. We emphasize that the use of the IPDG approach is an example of the global coupling of our local multiscale basis functions, and other choices of coarse grid methods are equally good. The key to our proposed method's success of is the construction of our

local multiscale basis functions.

Let V_H be a finite dimensional function space which consists of functions that are smooth on each coarse grid blocks but are in general discontinuous across coarse grid edges. We can then state the IPDG method as: find $u_H(t,\cdot) \in V_H$ such that

$$(\frac{\partial^2 u_H}{\partial t^2}, v) + a_{DG}(u_H, v) = l(v), \qquad \forall\, v \in V_H, \tag{5.4}$$

where the bilinear form $a_{DG}(u,v)$ and the linear functional $l(v)$ are defined by

$$\begin{aligned} a_{DG}(u,v) &= \sum_{K\in\mathcal{T}^H}\int_K a\nabla u\cdot\nabla v + \sum_{e\in\mathcal{E}^H}\Big(-\int_e \{a\nabla u\cdot n\}_e\,[v]_e \\ &\quad -\int_e \{a\nabla v\cdot n\}_e\,[u]_e + \frac{\gamma}{h}\int_e a[u]_e\,[v]_e\Big) \\ l(v) &= (f,v) \end{aligned}$$

where $\gamma > 0$ is a penalty parameter and n denotes the unit normal vector on e. The jump and average operator are defined as Chapter 3. The initial conditions for the problem (5.4) are defined by $u_H(0) = P_H(g_0)$ and $(u_H)_t(0) = P_H(g_1)$, where P_H is the L^2-projection operator into V_H.

Let $T > 0$ be a fixed time and $\Delta t = T/N$ be the time step size. The time discretization is done in the standard way, we find $u_H^{n+1} \in V_H$ such that

$$(u_H^{n+1}, v) = 2(u_H^n, v) - (u_H^{n-1}, v) - \Delta t^2\Big(a_{DG}(u_H^n, v) - l(v)\Big), \qquad \forall\, v \in V_H \tag{5.5}$$

in each time step. Throughout the chapter, the notation u^n represents the value of the

function u at time t_n. The initial conditions are obtained as follows

$$u_H^0 = P_H(g_0),$$

$$u_H^1 = u_H^0 + \Delta t\, P_H(g_1) + \frac{\Delta t^2}{2}\tilde{v},$$

where $\tilde{v} \in V_H$ is defined by

$$(\tilde{v}, v) = (f(0), v) - a_{DG}(g_0, v), \qquad \forall\, v \in V_H.$$

We will use the offline space defined in Chapter 3 to be our multiscale finite element space V_H. We recall that the basis construction process follows the offline construction in Chapter 3 in Section 3.1.2. That is, we will construct V_H as a sum of two spaces $V1, \text{off}$ and $V_{2,\text{off}}$, namely,

$$V_H = V_{1,\text{off}} + V_{2,\text{off}}.$$

where $V_{1,\text{off}}$ and $V_{2,\text{off}}$ are the sum of the local offline space $V_{1,\text{off}}^{(i)}$ and $V_{2,\text{off}}^{(i)}$, where

$$V_{1,\text{off}}^{(i)} = \text{span}\{\widetilde{w}_{l,K_i}, \quad l = 1, 2, \cdots, l_{1,i}\},$$
$$V_{2,\text{off}}^{(i)} = \text{span}\{z_{l,K_i}, \quad l = 1, 2, \cdots, l_{2,i}\}.$$

with $\phi_k^{(i)}$ and $\xi_k^{(i)}$ satisfying the spectral problem in (3.5) and (3.6). That is,

$$\int_{K_i} \kappa \nabla \widetilde{w}_{l,K_i} \cdot \nabla v = \frac{\lambda_{k,K}}{H} \int_{\partial K_i} \widetilde{\kappa} \widetilde{w}_{l,K_i} v, \qquad \forall v \in V_{1,\text{snap}}^{(i)}, \tag{5.6}$$

and

$$\int_{K_i} \kappa \nabla z_{l,K_i} \cdot \nabla v = \frac{\mu_{k,K}}{H^2} \int_{K_i} \kappa z_{l,K_i} v, \qquad \forall v \in V_{2,\text{snap}}^{(i)}, . \tag{5.7}$$

Given a function $w \in V_{1,\text{off}}^{(i)}$, we can define a discrete normal flux $a\frac{\partial w}{\partial n} \in \partial\mathcal{T}^h(K_i)$ on ∂K_i by

$$\int_{\partial K_i} a\frac{\partial w}{\partial n}\, v = \int_{K_i} a\nabla w \cdot \nabla \hat{v}, \quad v \in \partial\mathcal{T}^h(K_i) \tag{5.8}$$

where $\hat{v}$ is any a-harmonic extension of v in K_i. This is well-defined since w is also obtained by an a-harmonic extension.

Quasi-orthogonality of $V_{1,\text{off}}^{(i)}$ and $V_{2,\text{off}}^{(i)}$. Finally, we point out the following quasi-orthogonality condition which will be used in our analysis. For any $v \in V_{1,\text{off}}^{(i)}$ and $u \in V_{2,\text{off}}^{(i)}$, we conclude by (3.4) that

$$\int_{K_i} a\nabla v \cdot \nabla u = 0. \tag{5.9}$$

5.3 Numerical Results

In this section, we will present some numerical examples to show the performance of our multiscale method. The media that we will consider is a heterogeneous field which is a modified Marmousi model (see the left plot of Figure 5.1). We have also considered more regular periodic highly heterogeneous fields and observed similar results. We will compare both the accuracy and efficiency of our method with the direct fine scale simulation defined in (5.2). To compare the accuracy, we will use the following error quantities

$$e_2 = \frac{\|u_H - u_h\|_{L^2(\Omega)}}{\|u_h\|_{L^2(\Omega)}}, \quad \overline{e_2} = \frac{\sqrt{\sum_K |\int_K u_H - \int_K u_h|^2}}{\sqrt{\sum_K |\int_K u_h|^2}}, \quad e_{H^1} = \frac{\|\nabla(u_H - u_h)\|_{L^2(\Omega)}}{\|\nabla u_h\|_{L^2(\Omega)}}$$

which are the relative L^2 norm error, the relative L^2 norm error for coarse grid averages and the relative L^2 norm error of the gradient. We will also consider the jump error on

coarse grid edges defined by

$$e_{Jump} = \sum_{e \in \mathcal{E}^H} \int_e [u_H]_e^2.$$

Moreover, we let t_{off} be the time needed for offline computations and t_{on} be the online computational time. These quantities are used to compare the efficiency of our method with direct fine scale simulation. To perform a fair comparison, we will use the same time step size for both of our GMsFEM and the fine scale method, since we only consider spatial upscaling in this chapter. However, we note that multiscale basis functions can be used for different source terms and boundary conditions which will provide a substantial computational saving. Furthermore, we will take $\gamma = 2$ and $\Omega = [0, 1]^2$ for all of our examples. The initial conditions g_0 and g_1 are zero. Throughout the chapter, all computational times are measured in seconds.

The Ricker wavelet with frequency $f_0 = 20$

$$f(x, y, t) = (10)^2 e^{-10^2((x-0.5)^2+(y-0.5)^2)}(1 - 2\pi^2 f_0^2(t - 2/f_0)^2)e^{-\pi^2 f_0^2(t-2/f_o)^2}$$

is used as the source term. We will compute the solution at time $T = 0.4$. The coarse mesh size is taken as $H = 1/16$. Each coarse grid block is divided into a 32×32 grid, that is, $n = 32$. Thus, the fine mesh size $h = 1/512$ and there are totally 128 and 961 local basis functions in the space $V_{1,\text{off}}^{(i)}$ and $V_{2,\text{off}}^{(i)}$ respectively on each coarse grid block. The time step size for both GMsFEM and the fine grid solver is taken as $\Delta t = h/80$ in order to meet the stability requirement and the computation time for fine grid solution is 55.06. We will compare the accuracy and efficiency of our method using the solution computed at the time $T = 0.2$, which is shown in the right figure of Figure 5.1.

In Table 5.1 and Table 5.2, we present the errors and computational times for the case with $m = 1$, that is, we only use the first eigenfunction in the space $V_{2,\text{off}}$. We see that if we

use 80% of the total energy, the number of basis functions is between 33 and 40 on each coarse grid while the computational time for the offline procedure is 1019.06 and the time for online computations is 32.43. Note that the online computational time is about 59% of that of the online computational time of the direct fine grid simulation. The relative L^2 error and the relative error for cell averages are only 3.92% and 2.74% respectively. In addition, the relative error for the gradient is 14.86% and the jump error is 0.003. When 75% of the total energy is used, the number of basis functions is reduced to a number between 24 and 29 while the computational time for the offline procedures is 326.83 and the time for online computations is 18.21. The time for the online computation is 33% of the time required for direct fine grid simulation. The relative L^2 error and the relative error for cell averages are increased slightly to 4.23% and 3.12%, respectively. In Table 5.1, we also present the values of μ_{min} for the space $V_{1,\text{off}}$. Moreover, the eigenvalues are shown in Figure 5.3. The numerical solutions for these cases are shown in Figure 5.2. We note that the error decay is not fast mostly due to the error contribution because of the modes corresponding to the interior. Even though the error between the GMsFEM solution and the solution computed using the entire snapshot space $V_{1,\text{off}}$ is very small, the overall error between the GMsFEM solution and the fine-scale solution may not be small because we have only used one basis function in $V_{2,\text{off}}$. Next, we will add more basis functions from $V_{2,\text{off}}$ and compare the errors.

Next, we will investigate the use of more eigenfunctions in the space $V_{2,\text{off}}$ that will allow reducing the overall error. To do so, we consider the first case where 75% energy in the space $V_{1,\text{off}}$ is used and we consider using various number of eigenfunctions in $V_{2,\text{off}}$. The errors and computational times are shown in Table 5.3 and Table 5.4. In general, we obtain better numerical approximations as more eigenfunctions are used. When two eigenfunctions are used (this corresponds to using less than 3% of the total local degrees of freedom in constructing all GMsFEM basis functions), the relative error is 3.52% and

the online computational time is 18.64. When five eigenfunctions are used, the relative error is 1.93% and the online computational time is 18.21. Thus, we see that adding a few eigenfunctions in the space $V_{2,\text{off}}$ will improve the multiscale solution. This indicates that for the multiscale wave simulations, the modes that represent the interior nodes can improve the accuracy of the method and play an important role in obtaining an accurate solution. We also report the largest eigenvalue used in Table 5.3. Furthermore, for the use of the central frequency $f_0 = 20$, 75% energy in the space V_H^1 and $m = 5$ in the space $V_{2,\text{off}}$, there are about 35 basis functions per coarse-grid block and the number of points per wavelength is approximately 30, where the number of points per wavelength is computed as $2\pi\sqrt{N_{\text{total}}}/f_0$ and N_{total} is the total number of basis functions. On the other hand, since the fine grid solution is computed on a 512×512 fine grid, the number of points per wavelength is about 160. Therefore, the fine grid solution has sufficient resolution.

We would like to remark that the computational gain will be higher when implicit methods are used or we employ finer grids to resolve the problem. In the latter case, the CPU time for coarse-grid simulations will not change.

5.3.1 The use of oversampling

In this section, we present the performance of the method when the basis functions in the space $V_{1,\text{off}}$ are obtained by oversampling. We consider the previous example. The oversampling technique is used and the harmonic extension problems are solved on enlarged coarse grids, which are obtained by extending the original coarse grids by $H/16$ on each side. The results for using one basis functions in $V_{2,\text{off}}$ and various number of basis functions in $V_{1,\text{off}}$ are shown in Table 5.5. Moreover, we compute the errors using 73% energy for $V_{1,\text{off}}$ and various number of basis functions in $V_{2,\text{off}}$. The results are presented in Table 5.6. We observe that there is no improvement in this case. This is due to the error from the modes representing internal nodes.

5.4 Stability and convergence

In this section, we will prove the stability and convergence of the generalized multi-scale finite element method constructed in Section 5.2. We will first state and prove some preliminary results, and then prove the main convergence theorem for the semi-discrete scheme (5.4).

5.4.1 Preliminaries

Before we analyze the convergence of our GMsFEM, we first prove some basic results. To do so, we introduce some notations and state the assumptions required in our analysis. For functions $u, v \in H^1(\mathcal{T}^H)$, we define the bilinear form $a(\cdot,\cdot)$ by

$$a(u,v) = \sum_{K\in\mathcal{T}^H} \int_K a\nabla u \cdot \nabla v.$$

Moreover, for any function $u \in H^1(\mathcal{T}^H)$, we define the a-norm by

$$\|u\|_a = \left(a(u,u) + \frac{\gamma}{h} \sum_{e\in\mathcal{E}^H} \|a^{\frac{1}{2}}[u]_e\|^2_{L^2(e)} \right)^{\frac{1}{2}}$$

and the a-semi-norm by

$$|u|_a = a(u,u)^{\frac{1}{2}}.$$

Furthermore, the broken H^1-norm for $u \in H^1(\mathcal{T}^H)$ is defined as

$$\|u\|_{H^1(\mathcal{T}^H)} = \left(\sum_{K\in T^H} |u|^2_{H^1(K)} + \frac{\gamma}{h} \sum_{e\in\mathcal{E}^H} \|[u]_e\|^2_{L^2(e)} \right)^{\frac{1}{2}}.$$

Assumption 1. *The function $a(x)$ is bounded, that is, there exist positive numbers a_0 and*

a_1 *such that*

$$a_0 \leq a(x) \leq a_1, \qquad \forall\, x \in \Omega.$$

This assumption implies that the norms $\|\cdot\|_a$ *and* $\|\cdot\|_{H^1(\mathcal{T}^H)}$ *are equivalent.*

In the following, we will describe the consistency of the method (5.4). We define the consistency error by

$$R_{u_h}(v) = l(v) - (\frac{\partial^2 u_h}{\partial t^2}, v) - a_{DG}(u_h, v), \qquad \forall\, v \in V_H, \tag{5.10}$$

where u_h is the fine grid finite element solution defined in (5.2). Clearly, we have

$$R_{u_h}(v) = 0, \qquad \forall\, v \in V_{2,\text{off}} \tag{5.11}$$

since $V_{2,\text{off}} \subset V_h$. Thus, we only need to estimate $R_{u_h}(v)$ for $v \in V_{1,\text{off}}$. The following lemma states that the method (5.4) is consistent with the fine grid solution defined by (5.2). The proof will be presented in the Appendix.

Lemma 5.4.1. *Let* u_h *and* u *be the finite element solution defined in (5.2) and the exact solution of the wave propagation problem (5.1) respectively. Assume that* $u, u_t \in L^\infty(0,T;H^2(\Omega))$ *and that* $u \in W^{3,1}(0,T;H^1(\Omega))$*. We have*

$$|R_{u_h}(v)| \leq C(u)h\|v\|_a, \qquad v \in V_{1,\mathit{off}} \tag{5.12}$$

where $C(u)$ *is a constant which depends on the solution* u *but independent of the fine mesh size* h*. This inequality gives the consistency of our method.*

The coercivity and continuity conditions of the bilinear form a_{DG} are followed from Lemma 3.10 and 3.11 in Chapter 3.

Next, we will prove the convergence of the semi-discrete scheme (5.4). First, we define the following error quantities. Let

$$\eta = u_h - w_H, \quad \xi = u_H - w_H, \quad \text{and} \quad \varepsilon = u_h - u_H, \tag{5.13}$$

where $w_H \in V_H$ is defined by solving the following elliptic projection problem

$$a_{DG}(w_H, v) = a_{DG}(u_h, v) + R_{u_h}(v), \qquad \forall\, v \in V_H. \tag{5.14}$$

Notice that ε is the difference between the multiscale solution u_H and the fine grid finite element solution u_h. Moreover, η measures the difference between the fine grid solution u_h as its projection w_H. In the following, we will prove estimates for ε. First, we let

$$\|\varepsilon\|_{L^\infty([0,T];L^2(\Omega))} = \max_{0\le t\le T} \|\varepsilon\|_{L^2(\Omega)} \quad \text{and} \quad \|\varepsilon\|_{L^\infty([0,T];a)} = \max_{0\le t\le T} \|\varepsilon\|_a.$$

Then we will prove the following two inequalities, which estimate the error for the solution ε by the error for the projection η and the initial errors $\mathcal{I}_1$ and $\mathcal{I}_2$, which are defined in the statements of the theorems.

Theorem 5.4.2. *Let ε, η and ξ be the error quantities defined in (5.13). Then we have the following error bound*

$$\begin{aligned} &\|\varepsilon_t\|_{L^\infty([0,T];L^2(\Omega))} + \|\varepsilon\|_{L^\infty([0,T];a)} \\ &\le C\Big(\|\eta_t\|_{L^\infty([0,T];L^2(\Omega))} + \|\eta\|_{L^\infty([0,T];H^1(\mathcal{T}^H))} + \|\eta_{tt}\|_{L^1([0,T];L^2(\Omega))} + \mathcal{I}_1 \Big), \end{aligned} \tag{5.15}$$

where $\mathcal{I}_1 = \|\xi_t(0)\|_{L^2(\Omega)} + \|\xi(0)\|_{H^1(\mathcal{T}^H)}$.

Proof: First, using (5.4) and the definition of ξ, we have

$$(\xi_{tt}, v) + a_{DG}(\xi, v) = (f, v) - ((w_H)_{tt}, v) - a_{DG}(w_H, v).$$

Then by (5.12), we have

$$(\xi_{tt}, v) + a_{DG}(\xi, v) = (\eta_{tt}, v). \tag{5.16}$$

Taking $v = \xi_t$ in (5.16), we have

$$(\xi_{tt}, \xi_t) + a_{DG}(\xi, \xi_t) = (\eta_{tt}, \xi_t),$$

which implies

$$\frac{1}{2}\frac{d}{dt}\left(\|\xi_t\|^2_{L^2(\Omega)} + a_{DG}(\xi, \xi)\right) \leq \|\eta_{tt}\|_{L^2(\Omega)}\|\xi_t\|_{L^2(\Omega)}.$$

Integrating from $t = 0$ to $t = \tau$, we have

$$\begin{aligned}\|\xi_t(\tau)\|^2_{L^2(\Omega)} + \frac{1}{2}\|\xi(\tau)\|^2_a &\leq \|\xi_t(0)\|^2_{L^2(\Omega)} + 2\|\xi(0)\|^2_a + 2\int_0^\tau \|\eta_{tt}\|_{L^2(\Omega)}\|\xi_t\|_{L^2(\Omega)} \\ &\leq \|\xi_t(0)\|^2_{L^2(\Omega)} + 2\|\xi(0)\|^2_a + 2\max_{0\leq t\leq T}\|\xi_t\|_{L^2(\Omega)}\int_0^T \|\eta_{tt}\|_{L^2(\Omega)}.\end{aligned}$$

Therefore, we obtain

$$\|\xi_t\|^2_{L^\infty([0,T];L^2(\Omega))} + \|\xi\|^2_{L^\infty([0,T];a)} \leq C\Big(\|\xi_t(0)\|^2_{L^2(\Omega)} + \|\xi(0)\|^2_a + (\int_0^T \|\eta_{tt}\|_{L^2(\Omega)}\, dt)^2\Big).$$

Finally, (5.15) is proved by noting that $\varepsilon = \eta - \xi$.

□

Theorem 5.4.3. *Let ε, η and ξ be the error quantities defined in (5.13). Then we have the*

following error bound

$$\|\varepsilon\|_{L^\infty([0,T];L^2(\Omega))} \leq C\Big(\|\eta_t\|_{L^1([0,T],L^2(\Omega))} + \|\eta\|_{L^\infty([0,T];L^2(\Omega))} + \mathcal{I}_2\Big), \tag{5.17}$$

where $\mathcal{I}_2 = \|\xi(0)\|_{L^2(\Omega)}$.

Proof: Integrating by parts with respect to time in (5.16), we have

$$-(\xi_t, v_t) + \partial_t(\xi_t, v) + a_{DG}(\xi, v) = \partial_t(\eta_t, v) - (\eta_t, v_t).$$

Taking $v(x,t) = \int_t^\gamma \xi(x,\tau)d\tau$, we have $v_t = -\xi$ and $v(\gamma) = 0$. So,

$$(\xi_t, \xi) - \partial_t(\xi_t, v) - a_{DG}(v_t, v) = \partial_t(\eta_t, v) + (\eta_t, \xi),$$

which implies that

$$\frac{1}{2}\frac{d}{dt}\|\xi\|^2_{L^2(\Omega)} - \partial_t(\xi_t, v) - \frac{1}{2}\frac{d}{dt}a_{DG}(v, v) = \partial_t(\eta_t, v) + (\eta_t, \xi).$$

Integrating from $t = 0$ to $t = \gamma$, we have

$$\begin{aligned}&\frac{1}{2}\|\xi(\gamma)\|^2_{L^2(\Omega)} - \frac{1}{2}\|\xi(0)\|^2_{L^2(\Omega)} + (\xi_t(0), v(0)) + \frac{1}{2}a_{DG}(v(0), v(0))\\ =&(\eta_t(0), v(0)) + \int_0^\gamma (\eta_t, \xi).\end{aligned}$$

Since $\xi_t - \eta_t = (u_H - u_h)_t$, we obtain

$$(\xi_t(0) - \eta_t(0), v(0)) = ((u_H - u_h)_t(0), v(0)) = 0.$$

Using the coercivity of a_{DG}, we have

$$\begin{aligned}\|\xi(\gamma)\|^2_{L^2(\Omega)} &\leq \|\xi(0)\|^2_{L^2(\Omega)} + 2\int_0^\gamma \|\eta_t\|_{L^2(\Omega)} \|\xi\|_{L^2(\Omega)} \\ &\leq \|\xi(0)\|^2_{L^2(\Omega)} + 2 \max_{0\leq t\leq T} \|\xi\|_{L^2(\Omega)} \int_0^T \|\eta_t\|_{L^2(\Omega)}.\end{aligned}$$

Hence (5.17) is proved by noting that $\varepsilon = \eta - \xi$.

□

From Theorem 5.4.2 and Theorem 5.4.3, we see that, in order to estimate the error $\varepsilon = u_h - u_H$, we will need to find a bound for η given that the initial values $\xi_t(0)$ and $\xi(0)$ are sufficiently accurate.

5.4.2 Convergence analysis

In this section, we will derive an error bound for $\eta = u_h - w_H$. Notice that, on each coarse grid block K, we can express u_h as

$$u_h = \sum_{i=1}^{n} c_{i,K}\widetilde{w}_{i,K} + \sum_{i=1}^{n_0} d_{i,K} z_{i,K} = u_{1,K} + u_{2,K}$$

for some suitable coefficients $c_{i,K}$ and $d_{i,K}$ determined by a L^2-type projection, where n_0 is the dimension of $V_h^0(K)$. We write $u_h = u_1 + u_2$ with $u_i|_K = u_{i,K}$ for $i = 1, 2$. Moreover, we recall that $C(u, f)$, defined in (5.38), is the constant appearing in the consistency error estimate in Lemma 5.4.1. In the following theorem, we will give an estimate for the difference between the fine grid solution u_h and the projection of u_h into the coarse space V_H defined in (5.14). The theorem says that such difference is bounded by a best approximation error $\|u_h - v\|_a$ and a consistency error $hC(u, f)$. We emphasize that, even though the coarse mesh size H is fixed, but the fine mesh size h can be arbitrary small, and hence the consistency error is small compared with the best approximation error $\|u_h - v\|_a$.

Theorem 5.4.4. *Let $w_H \in V_H$ be the solution of (5.14) and u_h be the solution of (5.2). Then, for a fixed time $t \in [0, T]$, we have*

$$\|u_h - w_H\|_a \leq C(\|u_h - v\|_a + hC(u, f)), \qquad \forall\, v \in V_H. \tag{5.18}$$

Proof: By the definition of w_H, we have

$$a_{DG}(w_H, v) = a_{DG}(u_h, v) + R_{u_h}(v), \quad \forall\, v \in V_H.$$

So, we have

$$a_{DG}(w_H - v, w_H - v) = a_{DG}(u_h - v, w_H - v) + R_{u_h}(w_H - v).$$

By (3.11), (3.10) and (5.12), we get

$$\begin{aligned} \|w_H - v\|_a^2 &\leq 2a_{DG}(w_H - v, w_H - v) \\ &= 2a_{DG}(u_h - v, w_H - v) + 2R_{u_h}(w_H - v) \\ &\leq C(\|u_h - v\|_a + hC(u, f))\|w_H - v\|_a. \end{aligned}$$

Finally, we obtain

$$\begin{aligned} \|u_h - w_H\|_a &\leq \|u_h - v\|_a + \|w_H - v\|_a \\ &\leq C(\|u_h - v\|_a + hC(u, f)). \end{aligned}$$

□

From the above theorem, we see that the error $\|u_h - w_H\|_a$ is controlled by the quantity

$\|u_h - v\|_a$ for an arbitrary choice of the function $v \in V_H$. Thus, to obtain our final error bound, we only need to find a suitable function $v \in V_H$ to approximate the finite element solution u_h. In the following theorem, we will choose a specific v in Theorem 5.4.4 and prove the corresponding error estimate.

Theorem 5.4.5. *Let $u_h \in V_h$ be the finite element solution. Then, for a fixed time $t \in [0, T]$, we have*

$$\|u_h - \phi\|_a^2 \leq \sum_{K \in \mathcal{T}^H} \Big(\frac{H}{\mu_{p+1,K}} (1 + \frac{2a_1 \gamma H}{h \mu_{p+1,K}}) \int_{\partial K} (a \frac{\partial u_1}{\partial n})^2 + \frac{H^2}{\lambda_{m+1,K}} \|f - (u_h)_{tt}\|^2_{L^2(K)} \Big), \tag{5.19}$$

where the function $\phi \in V_H$ is defined as

$$\phi|_K = \sum_{i=1}^{p} c_{i,K} \widetilde{w}_{i,K} + \sum_{i=1}^{m} d_{i,K} z_{i,K} = \phi_{1,K} + \phi_{2,K}.$$

Proof: For a given coarse grid block K, using the orthogonality condition (5.9), we have

$$\int_K a |\nabla (u_h - \phi)|^2 = \int_K a |\nabla (u_1 - \phi_1)|^2 + \int_K a |\nabla (u_2 - \phi_2)|^2$$

which implies

$$\|u_h - \phi\|_a^2 = \|u_1 - \phi_1\|_a^2 + |u_2 - \phi_2|_a^2,$$

where we write $\phi = \phi_1 + \phi_2$ and $\phi_i|_K = \phi_{i,K}$, for $i = 1, 2$. We will first estimate

$\|u_1 - \phi_1\|_a^2$. By the definition of a-norm, we have

$$\begin{aligned}
\|u_1 - \phi_1\|_a^2 &= \sum_{K\in\mathcal{T}^H} \Big(\int_K a|\nabla(u_1-\phi_1)|^2 + \sum_{e\in\mathcal{E}^H} \frac{\gamma}{h}\int_e a|[(u_1-\phi_1)]_e|^2 \Big) \\
&\leq \sum_{K\in\mathcal{T}^H} \left(\int_K a|\nabla(u_1-\phi_1)|^2 + \frac{2\gamma}{h}\int_{\partial K} a|(u_1-\phi_1)|^2 \right) \\
&\leq \sum_{K\in\mathcal{T}^H} \left(\int_K a|\nabla(u_1-\phi_1)|^2 + \frac{2a_1\gamma}{h}\int_{\partial K} |(u_1-\phi_1)|^2 \right).
\end{aligned} \tag{5.20}$$

Next, we will estimate the right hand side of (5.20) for each K.

We note that the eigenvalue problem (3.5) is motivated by the right hand side of (5.20). In particular, based on the right hand side of (5.20), we consider

$$\int_K a\nabla w_\mu \cdot \nabla v + \frac{1}{H}\int_{\partial K} w_\mu v = \widehat{\mu}\int_K R(w_\mu)\cdot R(v), \qquad \forall\, v \in V_{1,\text{off}}^{(i)}, \tag{5.21}$$

where the choice of R, e.g., $R = \sqrt{a}\nabla w_\mu$, depends on how we would like to bound the error. Indeed, choosing the eigenvectors that correspond to the largest L_K eigenvalues, one can guarantee that the best L_K dimensional space *in the space of snapshots* is given by the first L_K dominant eigenvectors. The choice of $R(\cdot)$ is important and can influence the eigenvalue behavior. For example, the use of oversampling domains both for the snapshot space and the eigenvalue can provide a faster convergence. In this chapter, we take

$$R = \sqrt{a}\nabla w_\mu,$$

which allows estimating the right hand side of (5.20) by the energy norm. Note that, in (3.5), we use the smallest eigenvalues to determine the basis functions which is the same

as choosing the largest eigenvectors that correspond to the largest eigenvalues of (5.21) because $\widehat{\mu} = 1 + \frac{1}{\mu}$.

Note that the eigenvalue problem (3.5) is equivalent to

$$a\frac{\partial w_\mu}{\partial n} = \frac{\mu}{H} w_\mu \quad \text{on} \quad \partial K$$

where the normal flux $a\frac{\partial w_\mu}{\partial n}$ is defined discretely in (5.8). So, for each K,

$$\int_{\partial K} (a\frac{\partial u_{1,K}}{\partial n})^2 = \int_{\partial K} (a\frac{\partial}{\partial n}(\sum_{i=1}^{n} c_{i,K} w_{i,K}))^2 \tag{5.22}$$

$$= \int_{\partial K} (\sum_{i=1}^{n} \frac{\mu_{i,K}}{H} c_{i,K} w_{i,K})^2 = \sum_{i=1}^{n} (\frac{\mu_{i,K} c_{i,K}}{H})^2, \tag{5.23}$$

where we have used the fact that $\int_{\partial K} w_{i,K} w_{j,K} = \delta_{ij}$. Then, by using the eigenvalue problem defined in (3.5), we have

$$\frac{1}{h}\int_{\partial K} |(u_{1,K} - \phi_{1,K})|^2 = \frac{1}{h}\sum_{i=p+1}^{n} c_{i,K}^2 \leq \frac{H^2}{h\mu_{p+1,K}^2} \sum_{i=p+1}^{n} \left(\frac{\mu_{i,K}}{H}\right)^2 c_{i,K}^2$$

and

$$\int_K a|\nabla(u_{1,K} - \phi_{1,K})|^2 = \sum_{i=p+1}^{n} \frac{\mu_{i,K}}{H} c_{i,K}^2 \leq \frac{H}{\mu_{p+1,K}} \sum_{i=p+1}^{n} (\frac{\mu_{i,K}}{H})^2 c_{i,K}^2.$$

Note that, by using (5.23), we have,

$$\sum_{i=p+1}^{n} (\frac{\mu_{i,K}}{H})^2 c_{i,K}^2 \leq \sum_{i=1}^{n} (\frac{\mu_{i,K}}{H})^2 c_{i,K}^2 = \int_{\partial K} (a\frac{\partial u_{1,K}}{\partial n})^2.$$

Therefore

$$\begin{aligned}\|u_1-\phi_1\|_a^2 &\le \sum_{K\in\mathcal{T}^H}\Big(\frac{H}{\mu_{p+1,K}}(1+\frac{2a_1\gamma H}{h\mu_{p+1,K}})\sum_{i=p+1}^{n}(\frac{\mu_{i,K}}{H})^2c_{i,K}^2\Big)\\ &\le \sum_{K\in\mathcal{T}^H}\Big(\frac{H}{\mu_{p+1,K}}(1+\frac{2a_1\gamma H}{h\mu_{p+1,K}})\int_{\partial K}(a\frac{\partial u_1}{\partial n})^2\Big).\end{aligned} \tag{5.24}$$

Next, we will estimate $|u_2-\phi_2|_a^2$. Since u_h satisfies

$$\int_K a\nabla u_h\cdot\nabla v=\int_K\left(f-(u_h)_{tt}\right)v,\qquad \forall\, v\in V_h^0(K).$$

Putting $v=z_{i,K}$, we obtain

$$\frac{\lambda_{i,K}}{H^2}d_{i,K}=\int_K a\nabla u_h\cdot\nabla z_{i,K}=\int_K\left(f-(u_h)_{tt}\right)z_{i,K}.$$

We define $f_{i,K}=\int_K\left(f-(u_h)_{tt}\right)z_{i,K}$. Then we have $f_{i,K}=\dfrac{\lambda_{i,K}}{H^2}d_{i,K}$ and

$$\sum_{i=1}^{n_0}f_{i,K}^2\le\|f-(u_h)_{tt}\|_{L^2(K)}^2.$$

Hence,

$$
\begin{aligned}
|u_2-\phi_2|_a^2 &= \sum_{K\in\mathcal{T}^H}\int_K a|\nabla(u_2-\phi_2)|^2 \\
&= \sum_{K\in\mathcal{T}^H}\sum_{i\geq m+1}\frac{\lambda_{i,K}}{H^2}d_{i,K}^2 \\
&\leq \sum_{K\in\mathcal{T}^H}\frac{H^2}{\lambda_{m+1,K}}\sum_{i\geq m+1}\frac{\lambda_{i,K}^2}{H^4}d_{i,K}^2 \\
&= \sum_{K\in\mathcal{T}^H}\frac{H^2}{\lambda_{m+1,K}}\sum_{i\geq m+1}f_{i,K}^2 \\
&\leq \sum_{K\in\mathcal{T}^H}\frac{H^2}{\lambda_{m+1,K}}\|f-(u_h)_{tt}\|_{L^2(K)}^2.
\end{aligned}
$$

□

We note that, by the technique in [66], we can also derive a bound for $\|u_1-\phi_1\|_a$ as follows

$$\|u_1-\phi_1\|_a^2 \leq \sum_{K\in\mathcal{T}^H}\sum_{i\geq p+1}c_{i,K}^2.$$

This bound shows the decay of the error when more basis functions are used.

The bound in (5.19) gives the spectral convergence of our GMsFEM. Notice that, the term

$$H\sum_{K\in\mathcal{T}^H}\sum_{\partial K}(a\frac{\partial u_1}{\partial n})^2$$

is uniformly bounded and can be considered as a norm for u_1. Thus, (5.19) states that the error behaves as $O(\mu_{p+1,K}^{-1}+\lambda_{m+1,K}^{-1})$. We note that the eigenvalues increase (and go to the infinity as the fine mesh size decreases) and thus the error decreases as we increase the coarse space dimension.

Combining the results in Theorem 5.4.4 and Theorem 5.4.5, we obtain

$$\|\eta\|_a^2 \leq C \sum_{K\in\mathcal{T}^H} \Big(\frac{H^2}{\lambda_{m+1,K}} \|f - (u_h)_{tt}\|^2_{L^2(K)} \tag{5.25}$$

$$+ \frac{H}{\mu_{p+1,K}} (1 + \frac{2a_1\gamma H}{h\mu_{p+1,K}}) \int_{\partial K} (a\frac{\partial u_1}{\partial n})^2 \Big) + h^2 C(u)^2. \tag{5.26}$$

Similarly, we obtain

$$\begin{aligned}\|\eta_t\|_a^2 \leq\ & C \sum_{K\in\mathcal{T}^H} \Big(\frac{H^2}{\lambda_{m+1,K}} \|f_t - (u_h)_{ttt}\|^2_{L^2(K)} \\ & + \frac{H}{\mu_{p+1,K}} (1 + \frac{2a_1\gamma H}{h\mu_{p+1,K}}) \int_{\partial K} (a\frac{\partial (u_1)_t}{\partial n})^2 \Big) + h^2 C(u_t)^2.\end{aligned}$$

Finally, using these bounds for η, as well as the estimates proved in Theorem 5.4.2 and Theorem 5.4.3, we obtain estimates for the error ε.

5.5 Convergence of the fully discrete scheme

In this section, we will prove the convergence of the fully discrete scheme (5.5). To simplify the notations, we define the second order central difference operator δ^2 by

$$\delta^2(u^n) = \frac{u^{n+1} - 2u^n + u^{n-1}}{\Delta t^2}.$$

By the definition of the consistency error (5.10), at the time t_n, we have

$$((u_h)^n_{tt}, v) + a_{DG}(u^n_h, v) = (f^n, v) - R_{u^n_h}(v). \tag{5.27}$$

The fully discrete scheme (5.5) can be written using the operator δ^2 as

$$(\delta^2(u_H^n), v) + a_{DG}(u_H^n, v) = (f^n, v), \qquad \text{for } n \geq 1. \tag{5.28}$$

Moreover, we define

$$r^n = \begin{cases} u_{tt}^n - \delta^2(w_H^n), & \text{for } n \geq 1, \\ \Delta t^{-2}(\xi^1 - \xi^0), & \text{for } n = 0, \end{cases} \tag{5.29}$$

and

$$R^n = \Delta t \sum_{i=0}^{n} r^i.$$

In order to prove the convergence for the fully discrete scheme, we first prove the following lemma. The result will be needed in the derivation of an upper bound for the time step size Δt.

Lemma 5.5.1. *There exists a positive constant $\beta(h)$ such that*

$$a_{DG}(v, v) \leq \beta(h)^{-1} \|v\|_{L^2(\Omega)}^2, \quad \forall\, v \in V_H.$$

Moreover, the constant $\beta(h)$ can be taken as $h^2 a_1^{-1}(24 + 32\sqrt{3\Lambda} + 16\Lambda a_1^2 a_0^{-2})^{-1}$.

Proof: We first note that, if p is a linear function defined on the interval $I = [x_1 - h/2, x_1 + h/2]$, then we have

$$\|p\|_{L^\infty(I)}^2 \leq \frac{4}{h} \|p\|_{L^2(I)}^2 \tag{5.30}$$

$$|p|_{H^1(I)}^2 \leq \frac{12}{h^2} \|p\|_{L^2(I)}^2. \tag{5.31}$$

Then by the definition of a_{DG} and the Cauchy-Schwarz inequality, we have

$$
\begin{aligned}
& a_{DG}(v,v) \\
\leq & \sum_{K\in\mathcal{T}^H}\int_K a|\nabla v|^2 - 2\sum_{e\in\mathcal{E}^H}\Big(\int_e \{a\nabla v\cdot n\}_e\cdot[v]_e + \frac{\gamma}{h}\int_e a[v]_e^2\Big) \\
\leq & \sum_{K\in\mathcal{T}^H}\int_K a|\nabla v|^2 + 2(\sum_{K\in\mathcal{T}^H} h\|a\frac{\partial v}{\partial n_{\partial K}}\|^2_{L^2(\partial K)})^{\frac{1}{2}}(\sum_{e\in\mathcal{E}^H} h^{-1}\|[v]\|^2_{L^2(e)})^{\frac{1}{2}} \\
& + \frac{\gamma}{h}\sum_{e\in\mathcal{E}^H}\int a[v]_e^2.
\end{aligned}
$$

Then by using the discrete trace inequality , $a \leq a_1$ and estimating the jump terms by $L^2(\partial K)$ norms, we have

$$
\begin{aligned}
& a_{DG}(v,v) \\
\leq & a_1\Big(\sum_{K\in\mathcal{T}^H}\int_K |\nabla v|^2 + 2(\sum_{K\in\mathcal{T}^H}\Lambda\int_K|\nabla v|^2)^{\frac{1}{2}}(\sum_{e\in\mathcal{E}^H} h^{-1}\|[v]\|^2_{L^2(e)})^{\frac{1}{2}} + \frac{\gamma}{h}\sum_{e\in\mathcal{E}^H}\int[v]_e^2\Big) \\
\leq & a_1\Big(\sum_{K\in\mathcal{T}^H}\int_K |\nabla v|^2 + 4(\sum_{K\in\mathcal{T}^H}\Lambda\int_K|\nabla v|^2)^{\frac{1}{2}}(\sum_{K\in\mathcal{T}^H} h^{-1}\|v\|^2_{L^2(\partial K)})^{\frac{1}{2}} \\
& + \frac{2\gamma}{h}\sum_{K\in\mathcal{T}^H}\|v\|^2_{L^2(\partial K)}\Big).
\end{aligned}
\tag{5.32}
$$

Thus, it remains to estimate $\|\nabla v\|_{L^2(K)}$ and $\|v\|_{L^2(\partial K)}$ by the norm $\|v\|_{L^2(K)}$. The techniques used are the same as those used for standard finite element inverse inequalities. We include the derivation below for the explicit expression of $\beta(h)$.

We will estimate the term $\|\nabla v\|_{L^2(K)}$ first. For a given coarse grid block K, we can write it as the union of fine grid blocks $K = \cup_{F\subset K} F$, where we use F to represent a generic fine grid block. Since the fine grid blocks are rectangles, we can write F as a

tensor product of two intervals, namely, $F = I_x^F \times I_y^F$. For any $v \in V_H$ we can also write the restriction of v on F as $v(x, y) = v_{F,1}(x)v_{F,2}(y)$.

$$\begin{aligned}\int_K |\nabla v|^2 &= \sum_{F\subset K} \int_F |\nabla v|^2 \\ &= \sum_{F\subset K} \left(h(v'_{F,2})^2 \int_{I_x^F} (v_{F,1}(x))^2 + h(v'_{F,1})^2 \int_{I_y^F} (v_{F,2}(y))^2 \right) \\ &= \sum_{F\subset K} \left(\int_{I_y^F} (v'_{F,2}(y))^2 \int_{I_x^F} (v_{F,1}(x))^2 + \int_{I_x^F} (v'_{F,1}(x))^2 \int_{I_y^F} (v_{F,2}(y))^2 \right).\end{aligned}$$

Then, using (5.31), we have

$$\begin{aligned}\int_K |\nabla v|^2 &\leq 12h^{-2} \sum_{F\subset K} \left(\int_{I_y^F} (v_{F,2}(y))^2 \int_{I_x^F} (v_{F,1}(x))^2 + \int_{I_x^F} (v_{F,1}(x))^2 \int_{I_y^F} (v_{F,2}(y))^2 \right) \\ &= 24h^{-2} \sum_{F\subset K} \int_F |v|^2.\end{aligned}$$

Next, we estimate the term $\|v\|_{L^2(\partial K)}$. For a generic fine grid cell F, we write $I_x^F = [x_1, x_2]$ and $I_y^F = [y_1, y_2]$. Then, we define $B_{y_1} = \partial F \cap (I_x \times \{y_1\})$, $B_{y_2} = \partial F \cap (I_x \times \{y_2\})$, $B_{x_1} = \partial F \cap (\{x_1\} \times I_y)$ and $B_{x_2} = \partial F \cap (\{x_2\} \times I_y)$. By using (5.30),

$$
\begin{aligned}
&\|v\|^2_{L^2(\partial K)} = \sum_{F\subset K}\int_{\partial F\cap\partial K}(v_F)^2 \\
= \;& \sum_{F\subset K}\left(\int_{B_{y_1}}(v_{F,2}(y_1)v_{F,1}(x))^2 + \int_{B_{y_2}}(v_{F,2}(y_2)v_{F,1}(x))^2\right) \\
&+\sum_{F\subset K}\left(\int_{B_{x_1}}(v_{F,1}(x_1)v_{F,2}(y))^2 + \int_{B_{x_2}}(v_{F,1}(x_2)v_{F,2}(y))^2\right) \\
\leq \;& \frac{4}{h}\sum_{F\subset K}\Big(\int_{B_{y_1}}\int_{[y_1,y_1+h]} v_{F,2}(y)^2v_{F,1}(x)^2 + \int_{B_{y_2}}\int_{[y_2-h,y_2]} v_{F,2}(y)^2v_{F,1}(x)^2\Big) \\
&+\frac{4}{h}\sum_{F\subset K}\Big(\int_{B_{x_1}}\int_{[x_1,x_1+h]} v_{F,1}(x)^2v_{F,2}(y)^2 + \int_{B_{x_2}}\int_{[x_2-h,x_2]} v_{F,1}(x)v_{F,2}(y)^2\Big) \\
\leq \;& \frac{4}{h}\sum_{F\subset K}\left(\int_F(v_{F,2}(y)v_{F,1}(x))^2 + \int_F(v_{F,1}(x)v_{F,2}(y))^2\right) \\
= \;& \frac{8}{h}\|v\|^2_{L^2(K)}.
\end{aligned}
$$

Consequently, combining the above results and using (5.32),

$$
a_{DG}(v,v) \leq \frac{a_1}{h^2}\Big(24+32\sqrt{3\Lambda}+16\gamma\Big)\|v\|^2_{L^2(\Omega)}.
$$

Furthermore, using the lower bound of γ, we see that we can take $\beta(h)$ as

$$
\beta(h) = h^2a_1^{-1}(24+32\sqrt{3\Lambda}+16\Lambda a_1^2a_0^{-2})^{-1}.
$$

□

Finally, we will state and prove the convergence of the fully discrete scheme (5.5).

Theorem 5.5.2. *Assume that the time step size Δt satisfies the stability condition $\Delta t^2 < 4\beta(h)$. We have*

$$\max_{0\le n\le N} \|\varepsilon^n\|_{L^2(\Omega)} \le C \left(\|\varepsilon^0\|_{L^2(\Omega)} + \max_{0\le n\le N} \|\eta^n\|_{L^2(\Omega)} + \Delta t \sum_{n=0}^{N} \|R^n\|_{L^2(\Omega)} \right). \quad (5.33)$$

Proof: Subtracting (5.28) by (5.27), we have

$$(\delta^2(u_H^n) - (u_h)_{tt}^n, v) + a_{DG}(u_H^n - u_h^n, v) = R_{u_h^n}(v), \qquad \text{for } n \ge 1,$$

which implies

$$(\delta^2(u_H^n - w_H^n + w_H^n), v) + a_{DG}(u_H^n - u_h^n, v) = ((u_h)_{tt}^n, v) + R_{u_h^n}(v).$$

Using the fact that $\xi = u_H - w_H$ and the definition of the elliptic projection w_H given in (5.14), we have

$$(\delta^2(\xi^n), v) + a_{DG}(\xi^n, v) = (r^n, v), \qquad \text{for } n \ge 1,$$

where r^n is defined in (5.29). Using the definition of the operator δ^2, we have

$$(\frac{\xi^{n+1} - \xi^n}{\Delta t}, v) - (\frac{\xi^n - \xi^{n-1}}{\Delta t}, v) + \Delta t\, a_{DG}(\xi^n, v) = \Delta t (r^n, v).$$

Summing the above equations, we get for $n \ge 1$,

$$(\frac{\xi^{n+1} - \xi^n}{\Delta t}, v) - (\frac{\xi^1 - \xi^0}{\Delta t}, v) + \Delta t \sum_{i=1}^{n} a_{DG}(\xi^i, v) = \Delta t \sum_{i=1}^{n} (r^i, v). \quad (5.34)$$

To simplify the notations, we define

$$\Xi^n = \Delta t \sum_{i=1}^{n} \xi^i, \quad \text{for } n \geq 1; \quad \text{and} \quad \Xi^0 = 0.$$

Using the above definition and the definition of R^n, we can write (5.34) as

$$\left(\frac{\xi^{n+1} - \xi^n}{\Delta t}, v \right) + a_{DG}(\Xi^n, v) = (R^n, v), \quad n \geq 1.$$

Substituting $v = \xi^{n+1} + \xi^n$, we have

$$\|\xi^{n+1}\|^2_{L^2(\Omega)} - \|\xi^n\|^2_{L^2(\Omega)} + \Delta t\, a_{DG}(\Xi^n, \xi^{n+1} + \xi^n) = \Delta t (R^n, \xi^{n+1} + \xi^n),$$

and summing for all $n \geq 1$, we have

$$\|\xi^{n+1}\|^2_{L^2(\Omega)} - \|\xi^1\|^2_{L^2(\Omega)} + \Delta t \sum_{i=1}^{n} a_{DG}(\Xi^i, \xi^{i+1} + \xi^i) = \Delta t \sum_{i=1}^{n} (R^i, \xi^{i+1} + \xi^i). \quad (5.35)$$

By the definition of Ξ^n, we have $\Xi^{n+1} - \Xi^{n-1} = \Delta t(\xi^{n+1} + \xi^n)$ for $n \geq 1$. So

$$\begin{aligned}
\Delta t \sum_{i=1}^{n} a_{DG}(\Xi^i, \xi^{i+1} + \xi^i) &= \sum_{i=1}^{n} a_{DG}(\Xi^i, \Xi^{i+1} - \Xi^{i-1}) \\
&= \sum_{i=1}^{n} a_{DG}(\Xi^i, \Xi^{i+1}) - \sum_{i=0}^{n-1} a_{DG}(\Xi^i, \Xi^{i+1}) \\
&= a_{DG}(\Xi^n, \Xi^{n+1}).
\end{aligned}$$

Moreover,

$$a_{DG}(\Xi^n, \Xi^{n+1}) = a_{DG}(\frac{\Xi^n + \Xi^{n+1}}{2}, \frac{\Xi^n + \Xi^{n+1}}{2}) - a_{DG}(\frac{\Xi^n - \Xi^{n+1}}{2}, \frac{\Xi^n - \Xi^{n+1}}{2})$$
$$\geq -\frac{\Delta t^2}{4} a_{DG}(\xi^{n+1}, \xi^{n+1}).$$

So, (5.35) becomes

$$\|\xi^{n+1}\|^2_{L^2(\Omega)} - \frac{\Delta t^2}{4} a_{DG}(\xi^{n+1}, \xi^{n+1}) \leq \|\xi^1\|^2_{L^2(\Omega)} + \Delta t \sum_{i=1}^{n} (R^i, \xi^{i+1} + \xi^i), \quad n \geq 1.$$

Using the assumption $\Delta t^2 < 4\beta(h)$, we define $C_s = 1 - \dfrac{\Delta t^2}{4\beta(h)} > 0$. By Lemma 5.5.1,

$$\begin{aligned} C_s \|\xi^{n+1}\|^2_{L^2(\Omega)} &\leq \|\xi^1\|^2_{L^2(\Omega)} + \Delta t \sum_{i=1}^{n} (R^i, \xi^{i+1} + \xi^i) \\ &\leq \|\xi^1\|^2_{L^2(\Omega)} + 2\Delta t \max_{1 \leq i \leq n+1} \{\|\xi^i\|_{L^2(\Omega)}\} \sum_{i=1}^{n} \|R^i\|_{L^2(\Omega)} \\ &\leq \|\xi^1\|^2_{L^2(\Omega)} + \frac{C_s}{2} \max_{1 \leq i \leq n+1} \{\|\xi^i\|_{L^2(\Omega)}\}^2 + \frac{2}{C_s} \left(\Delta t \sum_{i=1}^{n} \|R^i\|_{L^2(\Omega)} \right)^2. \end{aligned}$$

Therefore, we have

$$\max_{1 \leq i \leq n+1} \{\|\xi^i\|_{L^2(\Omega)}\} \leq C \Big(\|\xi^1\|_{L^2(\Omega)} + \Delta t \sum_{i=1}^{n} \|R^i\|_{L^2(\Omega)} \Big).$$

Since $\xi^1 = \xi^0 + \Delta t^2 r^0$, we have

$$\max_{1 \leq i \leq n+1} \{\|\xi^i\|_{L^2(\Omega)}\} \leq C \left(\|\xi^0\|_{L^2(\Omega)} + \Delta t^2 \|r^0\|_{L^2(\Omega)} + \Delta t \sum_{i=1}^{n} \|R^i\|_{L^2(\Omega)} \right)$$

and using the definition of R^0,

$$\max_{1\leq i\leq n+1}\{\|\xi^i\|_{L^2(\Omega)}\} \leq C\left(\|\xi^0\|_{L^2(\Omega)} + \Delta t\sum_{i=0}^{n}\|R^i\|_{L^2(\Omega)}\right).$$

Thus,

$$\max_{0\leq i\leq n+1}\{\|\xi^i\|_{L^2(\Omega)}\} \leq C\left(\|\xi^0\|_{L^2(\Omega)} + \Delta t\sum_{i=0}^{n}\|R^i\|_{L^2(\Omega)}\right).$$

Finally, by using the relation $\varepsilon = \eta - \xi$, we obtain (5.33).

□

From Theorem 5.5.2, we see that the error ε^n of the fully discrete scheme mainly depends on two quantities, which are η^n and R^n. Recall that, η^n can be estimated as in (5.26). Therefore, it remains to get a bound for R^n. To do so, we will prove the following two lemmas for an upper bound of r^n.

Lemma 5.5.3. *We have*

$$\|r^0\|_{L^2(\Omega)} \leq C(\Delta t^{-1}\|\eta_t\|_{L^\infty([0,T];L^2(\Omega))} + \Delta t\|(u_h)_{ttt}\|_{C([0,T];L^2(\Omega))}).$$

Proof: By (5.29), we have $r^0 = \Delta t^{-2}(\xi^1 - \xi^0)$ and by the definition of u_H^0, we have

$$(u_H^0 - u^0, v) = 0, \quad \forall\, v \in V_H.$$

Then using the definitions of ξ^1 and ξ^0, we have

$$\begin{aligned}(\xi^1 - \xi^0, v) &= (u_H^1 - w_H^1, v) - (u_H^0 - w_H^0, v)\\ &= (u_h^1 - w_H^1, v) + (u_H^1 - u_h^1, v) - (u_h^0 - w_H^0, v)\\ &= ((u_h^1 - u_h^0) - (w_H^1 - w_H^0), v) + (u_H^1 - u_h^1, v).\end{aligned}$$

The first term can be estimated in the following way

$$|((u_h^1 - u_h^0) - (w_H^1 - w_H^0), v)| \leq \left|(\int_0^{t_1} \partial_t(u_h - w_H), v)\right| \\ \leq \Delta t \, \|\eta_t\|_{L^\infty([0,T];L^2(\Omega))} \|v\|_{L^2(\Omega)}.$$

To estimate the second term, by the Taylor's expansion, we get

$$u_h^1 = u_h^0 + \Delta t \, (u_h)_t^0 + \frac{\Delta t^2}{2}(u_h)_{tt}^0 + \frac{\Delta t^3}{6}(u_h)_{ttt}(\cdot, s), \qquad \text{where } 0 < s < t^1.$$

By the definition of u_H^1,

$$(u_H^1, v) = (u_h^1, v) = (u_h^0 + \Delta t (u_h)_t^0 + \frac{\Delta t^2}{2}(\tilde{v}, v)$$

Thus,

$$\begin{aligned}(u_H^1 - u_h^1, v) &= \frac{\Delta t^2}{2}(\tilde{v} - (u_h)_{tt}^0, v) - \frac{\Delta t^3}{6}((u_h)_{ttt}(\cdot, s), v) \\ &= \frac{\Delta t^2}{2}[(f^0, v) - a(u_h^0, v) + ((u_h)_{tt}^0, v)] - \frac{\Delta t^3}{6}((u_h)_{ttt}(\cdot, s), v) \\ &= -\frac{\Delta t^3}{6}((u_h)_{ttt}(\cdot, s), v)\end{aligned}$$

which proves the Lemma.

□

Lemma 5.5.4. *For $n \geq 1$, we have*

$$\|r^n\|_{L^2(\Omega)} \leq C(\Delta t^{-1} \int_{t_{n-1}}^{t_{n+1}} \|\eta_{tt}(\cdot, \tau)\|_{L^2(\Omega)} + \Delta t \int_{t_{n-1}}^{t_{n+1}} \|(u_h)_{tttt}(\cdot, \tau)\|_{L^2(\Omega)}).$$

Proof: By the definition of r^n,

$$\begin{aligned}\|r^n\|_{L^2(\Omega)} &= \|(u_h)_{tt}^n - \delta^2 w_H^n\|_{L^2(\Omega)} \\ &\leq \|\delta^2(w_H^n - u_h^n)\|_{L^2(\Omega)} + \|(u_h)_{tt}^n - \delta^2 u_h^n\|_{L^2(\Omega)}.\end{aligned}$$

Using the identity

$$v^{n+1} - 2v^n + v^{n-1} = \Delta t \int_{t_{n-1}}^{t_{n+1}} \left(1 - \frac{|\tau - t_n|}{\Delta t}\right) v_{tt}(\tau) d\tau,$$

the first term can be estimated as follows

$$\begin{aligned}(\delta^2(w_H^n - u_h^n), v) &= \frac{1}{\Delta t}\int_{t_{n-1}}^{t_{n+1}} (1 - \frac{|\tau - t_n|}{\Delta t}) \left(((w_H)_{tt} - (u_h)_{tt}), v\right)(\tau) d\tau \\ &\leq \frac{1}{\Delta t}\int_{t_{n-1}}^{t_{n+1}} \|\eta_{tt}(\cdot, \tau)\|_{\tilde{L}^2(\Omega)} \|v\|_{\tilde{L}^2(\Omega)} d\tau.\end{aligned}$$

To estimate the term $\|(u_h)_{tt}^n - \delta^2 u_h^n\|_{L^2(\Omega)}$, we use

$$\delta^2 u_h^n = (u_h)_{tt}^n + \frac{1}{6\Delta t^2}\int_{t_{n-1}}^{t_{n+1}} (\Delta t - |\tau - t_n|)^3 (u_h)_{tttt}(\cdot, \tau) d\tau.$$

This implies

$$\|(u_h)_{tt}^n - \delta^2 u_h^n\|_{L^2(\Omega)} \leq \frac{\Delta t}{6}\int_{t_{n-1}}^{t_{n+1}} \|(u_h)_{tttt}(\cdot, \tau)\|_{L^2(\Omega)} d\tau.$$

□

Using the definition of R^n and the above two lemma, we get

$$\begin{aligned}
&\|R^n\|_{L^2(\Omega)} \\
\leq\, & C\Big(\int_0^{t_n} \|\eta_{tt}(\cdot,\tau)\|_{L^2(\Omega)} + \|\eta_t\|_{L^\infty([0,T];L^2(\Omega))} \\
& + \Delta t^2 \int_0^{t_n} \|(u_h)_{tttt}(\cdot,\tau)\|_{L^2(\Omega)} + \Delta t^2 \|(u_h)_{ttt}\|_{C([0,T];L^2(\Omega))} \Big).
\end{aligned}$$

Hence we obtain

$$\begin{aligned}
&\Delta t \sum_{n=0}^{N} \|R^n\|_{L^2(\Omega)} \\
\leq\, & 2T \max_{0\leq n\leq N} \|R^n\|_{L^2(\Omega)} \\
\leq\, & C\Big(\int_0^{T} \|\eta_{tt}(\cdot,\tau)\|_{L^2(\Omega)} + \|\eta_t\|_{L^\infty([0,T];L^2(\Omega))} + \\
& \Delta t^2 \int_0^{T} \|(u_h)_{tttt}(\cdot,\tau)\|_{L^2(\Omega)} + \Delta t^2 \|(u_h)_{ttt}\|_{C([0,T];L^2(\Omega))} \Big).
\end{aligned}$$

Combining the estimates of η proved in Section 3 and Theorem 5.5.2, we obtain the error estimate for the fully discrete scheme (5.5).

5.6 Proof of the Lemma 5.4.1

In this section, we will prove Lemma 5.4.1. Let $v \in V_{1,\text{off}}$. By assumption, $u \in H^2(\Omega)$ for all time $t \in [0,T]$. Thus $a_{DG}(u,v)$ is well-defined, the normal flux is continuous across coarse edges and we have

$$\sum_{K\in\mathcal{T}^H} (\frac{\partial^2 u}{\partial t^2}, v)_{L^2(K)} + a_{DG}(u,v) = \sum_{K\in\mathcal{T}^H} (f,v)_{L^2(K)}, \quad \forall\, v \in H^1(\mathcal{T}^H). \tag{5.36}$$

Moreover, the following standard finite element error estimate holds

$$\begin{aligned} & \|\frac{\partial}{\partial t}(u-u_h)\|_{L^2(\Omega)} + |u-u_h|_{H^1(\Omega)} \\ \leq \;& Ch\Big(|u|_{L^\infty(0,T;H^2(\Omega))} + |u|_{W^{2,1}(0,T;H^1(\Omega))}\Big), \quad \forall\, t \in [0,T] \end{aligned}$$

and

$$\begin{aligned} & \|\frac{\partial^2}{\partial t^2}(u-u_h)\|_{L^2(\Omega)} + |\frac{\partial}{\partial t}(u-u_h)|_{H^1(\Omega)} \\ \leq \;& Ch\Big(|u_t|_{L^\infty(0,T;H^2(\Omega))} + |u|_{W^{3,1}(0,T;H^1(\Omega))}\Big), \quad \forall\, t \in [0,T]. \end{aligned}$$

We remark that, for simplicity, we assume the initial conditions belong to the fine space V_h. In the following derivations, we will fix the time variable t. By the definition of the consistency error and equation (5.36), we have

$$R_{u_h}(v) = (\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u_h}{\partial t^2}, v) + a_{DG}(u,v) - a_{DG}(u_h,v), \qquad \forall\, v \in H^1(\mathcal{T}^H). \tag{5.37}$$

Next, we define $v_c \in V_h$ in the following way. For each vertex in the triangulation, the value of v_c is defined as the average value of v at this vertex. Then by direct calculations, we have

$$\sum_{K\in\mathcal{T}^H} |v-v_c|^2_{H^1(K)} \leq C\frac{1}{h}\sum_{e\in\mathcal{E}^H} \|[v]\|^2_{L^2(e)}$$

and

$$\sum_{K\in\mathcal{T}^H} \|v-v_c\|^2_{L^2(K)} \leq Ch\sum_{e\in\mathcal{E}^H} \|[v]\|^2_{L^2(e)}.$$

Clearly, we have $[v - v_c]_e = [v]_e$ for all $e \in \mathcal{E}^H$ since $v_c \in C^0(\Omega)$. Therefore we get

$$\|v - v_c\|^2_{H^1(\mathcal{T}^H)} \leq C\frac{1}{h}\sum_{e\in\mathcal{E}^H} \|[v]\|^2_{L^2(e)}.$$

By (5.37) and (5.2) as well as the fact that $a_{DG}(u_h, v_c) = a(u_h, v_c)$, we have

$$R_{u_h}(v) = \sum_{K\in\mathcal{T}^H} (\frac{\partial^2(u - u_h)}{\partial t^2}, v - v_c)_{L^2(K)} + a_{DG}(u - u_h, v - v_c).$$

Next, we will estimate the two terms on the right hand side. For the first term, we have

$$\begin{aligned}
&\sum_{K\in\mathcal{T}^H} (\frac{\partial(u - u_h)}{\partial t^2}, v - v_c)_{L^2(K)} \\
\leq &\|\frac{\partial^2(u - u_h)}{\partial t^2}\|_{L^2(\Omega)} \left(\sum_{K\in\mathcal{T}^H} \|v - v_c\|^2_{L^2(K)}\right)^{\frac{1}{2}} \\
\leq &Ch\,\|\frac{\partial^2(u - u_h)}{\partial t^2}\|_{L^2(\Omega)} \left(\frac{1}{h}\sum_{e\in\mathcal{E}^H} \|[v]\|^2_{L^2(e)}\right)^{\frac{1}{2}} \\
\leq &Ch^2\Big(|u_t|_{L^\infty(0,T;H^2(\Omega))} + |u|_{W^{3,1}(0,T;H^1(\Omega))}\Big) \left(\frac{1}{h}\sum_{e\in\mathcal{E}^H} \|[v]\|^2_{L^2(e)}\right)^{\frac{1}{2}}.
\end{aligned}$$

For the second term, by the definition of a_{DG} and the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
& a_{DG}(u-u_h, v-v_c) \\
&= \sum_{K\in\mathcal{T}^H}\int_K a\nabla(u-u_h)\cdot\nabla(v-v_c) - \sum_{e\in\mathcal{E}^H}\int_e \{a\nabla(u-u_h)\cdot n\}[v] \\
&\leq \sum_{K\in\mathcal{T}^H} a_1 |u-u_h|_{H^1(K)}\ |v-v_c|_{H^1(K)} \\
&\quad + \left(\sum_{K\in\mathcal{T}^H}\int_{\partial K}(a\nabla(u-u_h)\cdot n)^2\right)^{\frac{1}{2}}\left(\sum_{e\in\mathcal{E}^H}\int_e [v]^2\right)^{\frac{1}{2}}.
\end{aligned}$$

To estimate the flux term above, we let I_K be the standard finite element interpolant. Then we have

$$\begin{aligned}
& \sum_{K\in\mathcal{T}^H}\int_{\partial K}(a\nabla(u-u_h)\cdot n)^2 \\
&\leq 2\left(\sum_{K\in\mathcal{T}^H}\int_{\partial K}(a\nabla(u-I_K(u))\cdot n)^2\right) + 2\left(\sum_{K\in\mathcal{T}^H}\int_{\partial K}(a\nabla(I_K(u)-u_h)\cdot n)^2\right) \\
&\leq Ca_1\left(h|u|^2_{H^2(K)} + \frac{1}{h}|I_K(u)-u_h|^2_{H^1(K)}\right) \\
&\leq Ca_1\left(h|u|^2_{H^2(K)} + \frac{1}{h}|I_K(u)-u|^2_{H^1(K)} + \frac{1}{h}|u-u_h|^2_{H^1(K)}\right) \\
&\leq Ca_1 h|u|^2_{H^2(K)}.
\end{aligned}$$

Next we will estimate the term $\sum_{K\in\mathcal{T}^H} a_1|u-u_h|_{H^1(K)}\ |v-v_c|_{H^1(K)}$. We have

$$\begin{aligned}
&\sum_{K\in\mathcal{T}^H} a_1|u-u_h|_{H^1(K)}\cdot|v-v_c|_{H^1(K)}\\
\leq& Ca_1(\sum_{K\in\mathcal{T}^H}|u-u_h|^2_{H^1(K)})^{\frac{1}{2}}(\sum_{K\in\mathcal{T}^H}|v-v_c|^2_{H^1(K)})^{\frac{1}{2}}\\
\leq& Ca_1(\frac{1}{h}\sum_{K\in\mathcal{T}^H}|u-u_h|^2_{H^1(K)})^{\frac{1}{2}}(\sum_{e\in\mathcal{E}^H}\|[v]\|^2_{L^2(e)})^{\frac{1}{2}}\\
\leq& Ca_1h|u|_{H^2(\Omega)}(\frac{1}{h}\sum_{e\in\mathcal{E}^H}\|[v]\|^2_{L^2(e)})^{\frac{1}{2}}.
\end{aligned}$$

Combining the above estimates, we get

$$\begin{aligned}
&|R_{u_h}(v)|\\
\leq&\frac{C}{a_0}h\Big(h\Big(|u_t|_{L^\infty(0,T;H^2(\Omega))}+|u|_{W^{3,1}(0,T;H^1(\Omega))}\Big)\\
&+a_1|u|_{L^\infty(0,T;H^2(\Omega))}\Big)\left(\frac{1}{h}\sum_{e\in\mathcal{E}^H}\|a[v]\|^2_{L^2(e)}\right)^{\frac{1}{2}},
\end{aligned}$$

for all $t\in[0,T]$. Finally, the constant $C(u)$ in the lemma can be chosen as

$$C(u)\approx|u_t|_{L^\infty(0,T;H^2(\Omega))}+|u|_{W^{3,1}(0,T;H^1(\Omega))}+|u|_{L^\infty(0,T;H^2(\Omega))}. \tag{5.38}$$

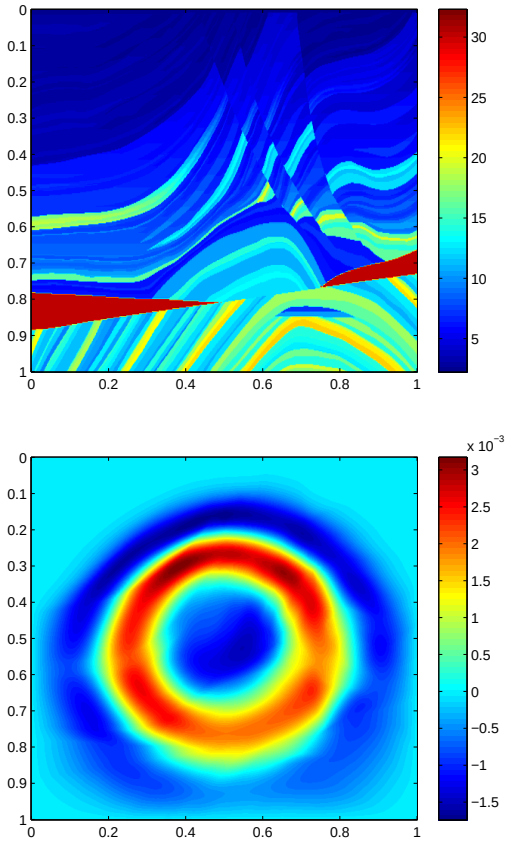

Figure 5.1: Left: a subset of the Marmousi model. Right: fine grid solution.

Energy	Number of basis	e_2	$\overline{e_2}$	e_{H^1}	e_{Jump}	μ_{min}
75%	24-29	0.0423	0.0312	0.1542	4.7304e-04	1.9414
80%	33-40	0.0392	0.0274	0.1486	3.0671e-04	2.9992

Table 5.1: Errors for various choices of energy for the space $V_{1,\text{off}}$.

Energy	t_{off}	t_{on}
75%	326.83	18.21
80%	1019.06	32.43

Table 5.2: Offline and online computational times.

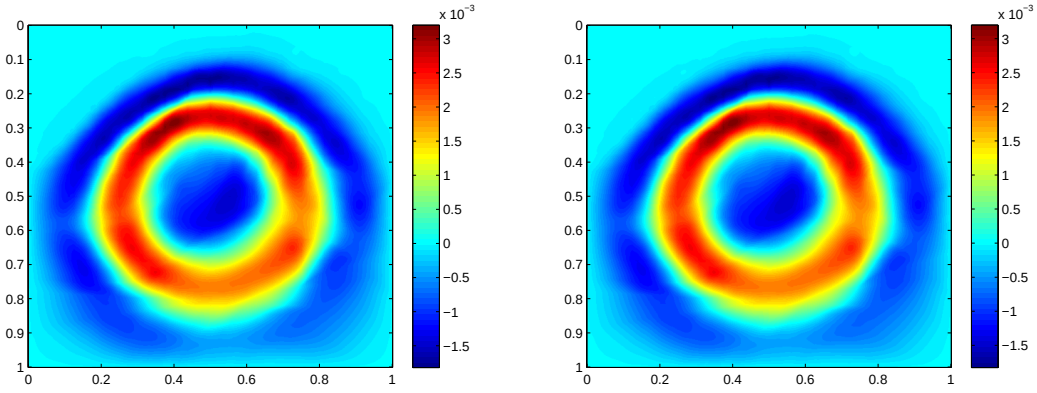

Figure 5.2: Left: 75% energy. Right: 80% energy.

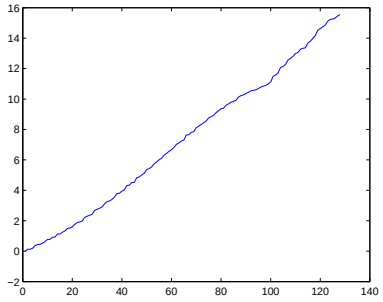

Figure 5.3: Eigenvalues for the space $V_{1,\text{off}}$.

m	e_2	$\overline{e_2}$	e_{H^1}	e_{Jump}	λ_{min}
1	0.0423	0.0312	0.1542	4.7304e-04	3.4805e+04
2	0.0352	0.0259	0.1346	4.7030e-04	3.4873e+04
3	0.0227	0.0187	0.0945	4.5931e-04	5.5906e+04
5	0.0193	0.0163	0.0833	4.5910e-04	6.9650e+04

Table 5.3: Errors for various number of eigenfunctions in $V_{2,\text{off}}$ for using 75% energy in $V_{1,\text{off}}$.

m	t_{off}	t_{on}
1	326.83	18.21
2	368.89	18.64
3	405.73	19.88
5	528.47	25.96

Table 5.4: Offline and online computational times for various number of eigenfunctions in $V_{2,\text{off}}$ for using 75% energy in $V_{1,\text{off}}$.

Energy	Number of basis	e_2	$\overline{e_2}$	e_{H^1}	e_{Jump}	μ_{min}
73%	24-30	0.0673	0.0583	0.1866	5.2038e-04	1.6755
79%	33-40	0.0640	0.0548	0.1827	3.4797e-04	2.5681
84%	45-55	0.0626	0.0534	0.1809	2.6388e-04	3.7918

Table 5.5: Simulation results with one basis function in $V_{2,\text{off}}$.

m	e_2	$\overline{e_2}$	e_{H^1}	e_{Jump}	λ_{min}
1	0.0673	0.0583	0.1866	5.2038e-04	3.4805e+04
2	0.0596	0.0524	0.1666	5.1865e-04	3.4873e+04
3	0.0488	0.0449	0.1332	5.0929e-04	5.5906e+04
5	0.0449	0.0419	0.1220	5.0793e-04	6.9650e+04

Table 5.6: Errors and computational times for various number of eigenfunctions in $V_{2,\text{off}}$ for using $E = 73\%$.

6. SUMMARY AND CONCLUSIONS

In this section, we will have a summary of this dissertation. In Chapter 2, we discuss about a residual-based online basis construction for GMsFEM. The main idea of the proposed method is to construct the online basis functions by solving local problems based on a computed residual. In particular, in each coarse region, an online multiscale basis function is constructed by solving local problems with a right hand side that is a residual computed at the current solution iterate. We show that the offline space needs to satisfy ONERP condition in order to guarantee that adding online basis function will decrease the error independent of the contrast and small scales. The online basis functions account for global effects that are missing in local GMsFEM basis functions. However, the first several GMsFEM basis functions are needed in order to guarantee that the online basis functions will decrease the error independent of the contrast. This method is applied in conjunction with an adaptivity where online basis functions are added in selective regions. The overall procedure results to a local multiscale approach where one can adaptively select regions and compute multiscale basis functions without resorting to global solves. We test our approaches on several examples and present some representative numerical results. Our numerical results show that with the offline spaces that satisfy ONERP, one can achieve a rapid decay of the error. We propose some strategies to reduce the computational cost associated with calculating the online basis functions.

In Chapter 3, we an offline adaptive Generalized Multiscale Discontinuous Galerkin Method (GMsDGM) for a class of high-contrast flow problems, and derive a-priori and a-posteriori error estimates for the method. Based on the a-posteriori error estimator, we de-

velop an adaptive enrichment algorithm for our GMsDGM and prove its convergence. The adaptive enrichment algorithm gives an automatic way to enrich the approximation space in regions where the solution requires more basis functions, which are shown to perform well compared with a uniform enrichment. We also discuss an approach that adaptively selects multiscale basis functions by correlating the residual to multiscale basis functions (cf. [40]). The proposed error indicators are L_2-based and can be inexpensively computed which makes our approach efficient. Numerical results are presented that demonstrate the robustness of the proposed error indicators.

In Chapter 4, we discuss about the online basis construction for GMsDGM. Though the use of offline basis functions is important for multiscale finite element methods, adding online basis functions in some regions can improve the convergence dramatically. The construction of online basis functions for various applications and discretizations require a careful analysis. In particular, as we have shown earlier [45] for GMsFEM within continuous Galerkin framework that one needs a certain number of offline basis functions in order to guarantee that the online basis functions can result to a convergence independent of physical parameters. In this chapter, we develop an online basis procedure for GMsDGM that can provide a convergence independent of the contrast and small scales. Because multiscale basis functions are discontinuous across coarse-grid boundaries, we construct a special offline space as well as online basis functions. We show that our construction will guarantee a convergence independent of the contrast and small scales if we select a certain number of offline basis functions based on a local spectral problem. Furthermore, we apply an adaptive procedure to add online basis functions in only some selected regions. Numerical results are presented to back up our theoretical findings.

In Chapter 5, we present a multiscale wave simulation method based on GMsDGM for highly heterogeneous media. This wave simulation method is based on GMsDGM which preform a local model reduction by solving a spectral problem in each local snap-

shot space. The local spectral problems is used to select the important modes in that local neighborhood. The local spectral problems are designed to achieve a high accuracy and motivated by the global coupling formulation. The use of multiple snapshot spaces and multiple spectral problems is one of the novelties of this work. Using the dominant modes from local spectral problems, multiscale basis functions are constructed to represent the solution space locally within each coarse block. These multiscale basis functions are coupled via the symmetric interior penalty discontinuous Galerkin method which provides a block diagonal mass matrix, and, consequently, results in fast computations in an explicit time discretization. Numerical examples are presented. In particular, we discuss how the modes from our snapshot spaces can affect the accuracy of the method. Our numerical results show that one can obtain an accurate approximation of the solution with GMsFEM using less than 3% of the total local degrees of freedom.

REFERENCES

[1] M. Hinze and S. Volkwein, "Proper orthogonal decomposition surrogate models for nonlinear dynamical systems: error estimates and suboptimal control," in *Dimension Reduction of Large-Scale Systems* (P. Benner, V. Mehrmann, and D. Sorensen, eds.), vol. 45 of *Lecture Notes in Computational Science and Engineering*, pp. 261–306, Springer Berlin Heidelberg, 2005.

[2] Y. Maday, "Reduced-basis method for the rapid and reliable solution of partial differential equations." Downloadable at http://hal.archives-ouvertes.fr/hal-00112152/, 7 2006.

[3] M. Ghommem, M. Presho, V. M. Calo, and Y. Efendiev, "Mode decomposition methods for flows in high-contrast porous media. globalï£¡-local approach," *Journal of Computational Physics, Vol. 253.*, pp. 226ï£¡–238.

[4] L. Durlofsky, "Numerical calculation of equivalent grid block permeability tensors for heterogeneous porous media," *Water Resour. Res.*, vol. 27, pp. 699–708, 1991.

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

[6] T. Arbogast, "Analysis of a two-scale, locally conservative subgrid upscaling for elliptic problems," *SIAM J. Numer. Anal.*, vol. 42, no. 2, pp. 576–598 (electronic), 2004.

[7] C.-C. Chu, I. G. Graham, and T.-Y. Hou, "A new multiscale finite element method for high-contrast elliptic interface problems," *Math. Comp.*, vol. 79, no. 272, pp. 1915–1955, 2010.

[8] A. Abdulle, "The finite element heterogeneous multiscale method: a computational strategy for multiscale pdes," *GAKUTO International Series Mathematical Sciences and Applications*, vol. 31, no. EPFL-ARTICLE-182121, pp. 135–184, 2009.

[9] W. E and B. Engquist, "Heterogeneous multiscale methods," *Comm. Math. Sci.*, vol. 1, no. 1, pp. 87–132, 2003.

[10] A. Anantharaman, R. Costaouec, C. Le Bris, F. Legoll, F. Thomines, *et al.*, "Introduction to numerical stochastic homogenization and the related computational challenges: some recent developments," *Multiscale modeling and analysis for materials simulation*, 2011.

[11] P. Henning and M. Ohlberger, "The heterogeneous multiscale finite element method for elliptic homogenization problems in perforated domains," *Numerische Mathematik*, vol. 113, no. 4, pp. 601–629, 2009.

[12] G. Allaire and R. Brizzi, "A multiscale finite element method for numerical homogenization," *Multiscale Modeling & Simulation*, vol. 4, no. 3, pp. 790–812, 2005.

[13] P. Jenny, S. H. Lee, and H. A. Tchelepi, "Adaptive fully implicit multi-scale finite-volume method for multi-phase flow and transport in heterogeneous porous media," *Journal of Computational Physics*, vol. 217, no. 2, pp. 627–641, 2006.

[14] P. Jenny, S. Lee, and H. Tchelepi, "Multi-scale finite-volume method for elliptic problems in subsurface flow simulation," *Journal of Computational Physics*, vol. 187, no. 1, pp. 47–67, 2003.

[15] H. Hajibeygi and P. Jenny, "Adaptive iterative multiscale finite volume method," *Journal of Computational Physics*, vol. 230, no. 3, pp. 628–643, 2011.

[16] Y. Efendiev, J. Galvis, and X. Wu, "Multiscale finite element methods for high-contrast problems using local spectral basis functions," *Journal of Computational*

Physics, vol. 230, pp. 937–955, 2011.

[17] Y. Efendiev and T. Hou, *Multiscale Finite Element Methods: Theory and Applications*, vol. 4 of *Surveys and Tutorials in the Applied Mathematical Sciences*. New York: Springer, 2009.

[18] Y. Efendiev, T. Hou, and V. Ginting, "Multiscale finite element methods for nonlinear problems and their applications," *Comm. Math. Sci.*, vol. 2, pp. 553–589, 2004.

[19] E. Chung and Y. Efendiev, "Reduced-contrast approximations for high-contrast multiscale flow problems," *Multiscale Model. Simul.*, vol. 8, pp. 1128–1153, 2010.

[20] E. Chung and W. T. Leung, "A sub-grid structure enhanced discontinuous galerkin method for multiscale diffusion and convection-diffusion problems," *Commun. Comput. Phys.*, vol. 14, pp. 370–392, 2013.

[21] E. Chung, Y. Efendiev, and R. Gibson, "An energy-conserving discontinuous multiscale finite element method for the wave equation in heterogeneous media," *Advances in Adaptive Data Analysis*, vol. 3, pp. 251–268, 2011.

[22] E. Chung, Y. Efendiev, and W. T. Leung, "Generalized multiscale finite element method for wave propagation in heterogeneous media," *Multiscale Model. Simul*, vol. 12, pp. 1691–1721, 2014.

[23] R. Gibson, K. Gao, E. Chung, and Y. Efendiev, "Multiscale modeling of acoustic wave propagation in two-dimensional media," *Geophysics*, vol. 79, pp. T61–T75, 2014.

[24] Y. Efendiev, T. Hou, and X. Wu, "Convergence of a nonconforming multiscale finite element method," *SIAM J. Numer. Anal.*, vol. 37, pp. 888–910, 2000.

[25] H. Owhadi and L. Zhang, "Metric-based upscaling," *Comm. Pure. Appl. Math.*, vol. 60, pp. 675–723, 2007.

[26] Y. Efendiev, V. Ginting, T. Hou, and R. Ewing, "Accurate multiscale finite element methods for two-phase flow simulations," *Journal of Computational Physics*, vol. 220, pp. 155–174, 2006.

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

[28] Y. Efendiev, J. Galvis, R. Lazarov, M. Moon, and M. Sarkis, "Generalized multiscale finite element method. Symmetric interior penalty coupling," *J. Comput. Phys.*, vol. 255, pp. 1–15, 2013.

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

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

[31] E. Chung, Y. Efendiev, and G. Li, "An adaptive GMsFEM for high contrast flow problems," *J. Comput. Phys.*, vol. 273, pp. 54–76, 2014.

[32] S. Brenner and L. Scott, *The Mathematical Theory of Finite Element Methods*. New York: Springer-Verlag, 2007.

[33] K. Mekchay and R. H. Nochetto, "Convergence of adaptive finite element method for general second order elliptic PDEs," *SIAM J. Numer. Anal.*, vol. 43, pp. 1803–1827, 2005.

[34] W. Dorfler, "A convergent adaptive algorithm for poisson's equation," *SIAM J.Numer. Anal.*, vol. 33, pp. 1106 – 1124, 1996.

[35] M. Drohmann, B. Haasdonk, and M. Ohlberger, "Reduced basis approximation for nonlinear parametrized evolution equations based on empirical operator interpolation," *SIAM J. Sci. Comput.*, vol. 34, no. 2, pp. A937–A969, 2012.

[36] A. Abdulle and Y. Bai, "Adaptive reduced basis finite element heterogeneous multiscale method," *Comput. Methods Appl. Mech. Engrg.*, vol. 257, pp. 203–220, 2013.

[37] D. B. P. Huynh, D. J. Knezevic, and A. T. Patera, "A static condensation reduced basis element method: approximation and *a posteriori* error estimation," *ESAIM Math. Model. Numer. Anal.*, vol. 47, no. 1, pp. 213–251, 2013.

[38] N. C. Nguyen, G. Rozza, D. B. P. Huynh, and A. T. Patera, "Reduced basis approximation and a posteriori error estimation for parametrized parabolic PDEs: application to real-time Bayesian parameter estimation," in *Large-scale inverse problems and quantification of uncertainty*, Wiley Ser. Comput. Stat., pp. 151–177, Chichester: Wiley, 2011.

[39] T. Tonn, K. Urban, and S. Volkwein, "Comparison of the reduced-basis and POD *a posteriori* error estimators for an elliptic linear-quadratic optimal control problem," *Math. Comput. Model. Dyn. Syst.*, vol. 17, no. 4, pp. 355–369, 2011.

[40] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, 2001. Reprinted from SIAM J. Sci. Comput. **2**0 (1998), no. 1, 33–61 (electronic) [MR1639094 (99h:94013)].

[41] V. Calo, Y. Efendiev, J. Galvis, and G. Li, "Randomized oversampling for generalized multiscale finite element methods," *http://arxiv.org/pdf/1409.7114.pdf*. to appear in SIAM MMS.

[42] Y. Efendiev and J. Galvis, "A domain decomposition preconditioner for multiscale high-contrast problems," in *Domain Decomposition Methods in Science and Engi-*

neering XIX (Y. Huang, R. Kornhuber, O. Widlund, and J. Xu, eds.), vol. 78 of *Lect. Notes in Comput. Science and Eng.*, pp. 189–196, Springer-Verlag, 2011.

[43] E. Chung, Y. Efendiev, and W. T. Leung, "An adaptive generalized multiscale discontinuous galerkin method (GMsDGM) for high-contrast flow problems," *arXiv preprint arXiv:1409.3474*, 2014.

[44] B. Riviere, *Discontinuous Galerkin Methods For Solving Elliptic And parabolic Equations: Theory and Implementation*. SIAM, 2008.

[45] E. Chung, Y. Efendiev, and T. Leung, "Residual-driven online generalized multiscale finite element methods," submitted, arXiv:1501.04565.

[46] J. Galvis and Y. Efendiev, "Domain decomposition preconditioners for multiscale flows in high contrast media. reduced dimension coarse spaces," *SIAM J. Multiscale Modeling and Simulation*, vol. 8, pp. 1621–1644, 2010.

[47] J. Virieux, "P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method," *Geophysics*, vol. 51, no. 4, pp. 889–901, 1986.

[48] J. Virieux, "SH-wave propagation in heterogeneous media: Velocity-stress finite-difference method," *Geophysics*, vol. 49, no. 11, pp. 1933–1942, 1984.

[49] E. H. Saenger, R. Ciz, O. S. Krüger, S. M. Schmalholz, B. Gurevich, and S. A. Shapiro, "Finite-difference modeling of wave propagation on microscale: A snapshot of the work in progress," *Geophysics*, vol. 72, no. 5, pp. SM293–SM300, 2007.

[50] B. Lombard, J. Piraux, C. Gélis, and J. Virieux, "Free and smooth boundaries in 2-D finite-difference schemes for transient elastic waves," *Geophysical Journal International*, vol. 172, no. 1, pp. 252–261, 2008.

[51] W. Symes, I. S. Terentyev, and T. Vdovina, "Getting it right without knowing the answer: Quality control in a large seismic modeling project," *SEG Technical Program Expanded Abstracts*, vol. 28, no. 1, pp. 2602–2606, 2009.

[52] Y. J. Masson and S. R. Pride, "Finite-difference modeling of Biot's poroelastic equations across all frequencies," *Geophysics*, vol. 75, no. 2, pp. N33–N41, 2010.

[53] J. D. De Basabe and M. K. Sen, "New developments in the finite-element method for seismic modeling," *The Leading Edge*, vol. 28, no. 5, pp. 562–567, 2009.

[54] M. Käser, C. Pelties, C. E. Castro, H. Djikpesse, and M. Prange, "Wavefield modeling in exploration seismology using the discontinuous Galerkin finite-element method on HPC infrastructure," *The Leading Edge*, vol. 29, no. 1, pp. 76–85, 2010.

[55] C. Pelties, M. Käser, V. Hermann, and C. E. Castro, "Regular versus irregular meshing for complicated models and their effect on synthetic seismograms," *Geophysical Journal International*, pp. no–no, 2010.

[56] V. Hermann, M. Käser, and C. E. Castro, "Non-conforming hybrid meshes for efficient 2-D wave propagation using the discontinuous Galerkin method," *Geophysical Journal International*, vol. 184, no. 2, pp. 746–758, 2011.

[57] P. Moczo, J. Kristek, M. Galis, E. Chaljub, and V. Etienne, "3-D finite-difference, finite-element, discontinuous-Galerkin and spectral-element schemes analysed for their accuracy with respect to p-wave to s-wave speed ratio," *Geophysical Journal International*, vol. 187, no. 3, pp. 1645–1667, 2011.

[58] D. Komatitsch, D. Göddeke, G. Erlebacher, and D. Michéa, "Modeling the propagation of elastic waves using spectral elements on a cluster of 192 GPUs," *Computer Science - Research and Development*, vol. 25, pp. 75–82, 2010. 10.1007/s00450-010-0109-1.

[59] D. Komatitsch and J. Tromp, "Spectral-element simulations of global seismic wave propagation—I. Validation," *Geophysical Journal International*, vol. 149, no. 2, pp. 390–412, 2002.

[60] D. Komatitsch and J. Tromp, "Introduction to the spectral element method for three-dimensional seismic wave propagation," *Geophysical Journal International*, vol. 139, no. 3, pp. 806–822, 1999.

[61] C. Morency, Y. Luo, and J. Tromp, "Acoustic, elastic and poroelastic simulations of CO_2 sequestration crosswell monitoring based on spectral-element and adjoint methods," *Geophysical Journal International*, pp. no–no, 2011.

[62] W. Zhang, L. Tong, and E. Chung, "Exact nonreflecting boundary conditions for three dimensional poroelastic wave equations.," *Comm. Math. Sci.*, vol. To appear.

[63] W. Zhang, L. Tong, and E. Chung, "A new high accuracy locally one-dimensional scheme for the wave equation.," *J. Comput. Appl. Math.*, vol. 236, pp. 1343–1353, 2011.

[64] W. Zhang, L. Tong, and E. Chung, "Efficient simulation of wave propagation with implicit finite difference schemes.," *Numer. Math. Theor. Meth. Appl.*, vol. 5, pp. 205–228, 2012.

[65] F. Delprat-Jannaud and P. Lailly, "Wave propagation in heterogeneous media: Effects of fine-scale heterogeneity," *Geophysics*, vol. 73, no. 3, pp. T37–T49, 2008.

[66] Y. Efendiev, J. Galvis, R. Lazarov, M. Moon, and M. Sarkis, "Generalized multiscale finite element method. IPDG." Submitted, 2013.

[67] M. Grote and D. Schotzau, "Optimal error estimates for the fully discrete interior penalty DG method for the wave equation.," *J. Sci. Comput.*, vol. 40, pp. 257–272, 2009.

[68] E. Chung and W. Leung, "A sub-grid structure enhanced discontinuous Galerkin method for multiscale diffusion and convection-diffusion problems.," *Commum. Comput. Phys.*, vol. 14, pp. 370–392, 2013.

[69] E. Chung and B. Engquist, "Optimal discontinuous Galerkin methods for wave propagation.," *SIAM J. Numer. Anal.*, vol. 44, pp. 2131–2158, 2006.

[70] E. Chung and B. Engquist, "Optimal discontinuous Galerkin methods for the acoustic wave equation in higher dimensions.," *SIAM J. Numer. Anal.*, vol. 47, pp. 3820–3848, 2009.

[71] H. Chan, E. Chung, and G. Cohen, "Stability and dispersion analysis of staggered discontinuous Galerkin method for wave propagation.," *Int. J. Numer. Anal. Model.*, vol. 10, pp. 233–256, 2013.

[72] E. Chung, P. C. Jr., and T. Yu, "Convergence and superconvergence of staggered discontinuous Galerkin methods for the three-dimensional Maxwell's equations on Cartesian grids.," *J. Comput. Phys.*, vol. 235, pp. 14–31, 2013.

[73] E. Chung, H. Kim, and O. Widlund, "Two-level overlapping Schwarz algorithms for a staggered discontinuous Galerkin method.," *SIAM J. Numer. Anal.*, vol. 51, pp. 47–67, 2013.

[74] E. Chung and C. Lee, "A staggered discontinuous Galerkin method for the convection diffusion equation.," *J. Numer. Math.*, vol. 20, pp. 1–31, 2012.

[75] E. Chung and C. Lee, "A staggered discontinuous Galerkin method for the curl-curl operator.," *IMA J. Numer. Anal.*, vol. 32, pp. 1241–1265, 2012.

[76] E. Chung and P. C. Jr., "A staggered discontinuous Galerkin method for wave propagation in media with dielectrics and meta-materials.," *J. Comput. Appl. Math.*, vol. 239, pp. 189–207, 2013.

[77] E. Chung, Y. Efendiev, and R. Gibson, "An energy-conserving discontinuous multiscale finite element method for the wave equation in heterogeneous media.," *Advances in Adaptive Data Analysis*, vol. 3, pp. 251–268, 2011.

[78] R. Gibson, K. Gao, E. Chung, and Y. Efendiev, "Multiscale modeling of acoustic wave propagation in two-dimensional media.," *To appear in Geophysics*.

[79] F. Nataf, H. Xiang, V. Dolean, and N. Spillane, "A coarse space construction based on local Dirichlet-to-Neumann maps.," *SIAM J. Sci. Comput.*, vol. 33, pp. 1623–1642, 2011.

[80] A. Taflove and S. C. Hagness, *Computational Electromagnetics: The finite-difference time-domain method*. Artech House, 2005.

[81] Y. Efendiev, J. Galvis, G. Li, and M. Presho, "Generalized multiscale finite element methods. oversampling strategies," to appear in International Journal for Multiscale Computational Engineering.