Constraint Energy Minimizing Generalized Multiscale Discontinuous Galerkin Method

Siu Wun Cheung*   Eric T. Chung†   Wing Tat Leung‡

September 30, 2019

Abstract

Numerical simulation of flow problems and wave propagation in heterogeneous media has important applications in many engineering areas. However, numerical solutions on the fine grid are often prohibitively expensive, and multiscale model reduction techniques are introduced to efficiently solve for an accurate approximation on the coarse grid. In this paper, we propose an energy minimization based multiscale model reduction approach in the discontinuous Galerkin discretization setting. The main idea of the method is to extract the non-decaying component in the high conductivity regions by identifying dominant modes with small eigenvalues of local spectral problems, and define multiscale basis functions in coarse oversampled regions by constraint energy minimization problems. The multiscale basis functions are in general discontinuous on the coarse grid and coupled by interior penalty discontinuous Galerkin formulation. The minimal degree of freedom in representing high-contrast features is achieved through the design of local spectral problems, which provides the most compressed local multiscale space. We analyze the method for solving Darcy flow problem and show that the convergence is linear in coarse mesh size and independent of the contrast, provided that the oversampling size is appropriately chosen. Numerical results are presented to show the performance of the method for simulation on flow problem and wave propagation in high-contrast heterogeneous media.

1 Introduction

Many engineering applications require numerical simulation in heterogeneous media with multiple scales and high contrast. For example, Darcy flow equation in heterogeneous media is used to describe fluid flow in porous medium in reservoir simulation, and wave equation in heterogeneous media has been widely used for subsurface modeling. Numerical solutions on the fine grid are often prohibitively expensive in these complex multiscale problems.

To this end, extensive research effort had been devoted to developing efficient methods for solving multiscale problems at reduced expense, for example, numerical homogenization approaches [35, 39] and multiscale methods, including Multiscale Finite Element Methods (MsFEM) [27, 23, 4, 22, 7], Variational Multiscale Methods (VMS) [29, 30, 31, 3], Heterogeneous Multiscale Methods (HMM) [16, 1, 17] and and Generalized Multsicale Finite Element Methods (GMsFEM) [18, 15, 10, 8]. The common goal of these methods is to construct numerical solvers on the coarse grid, which is typically much coarser than the fine grid which captures all the

---

*Department of Mathematics, Texas A&M University, College Station, TX 77843, USA (tonycsw2905@math.tamu.edu)

†Department of Mathematics, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong SAR, China (tschung@math.cuhk.edu.hk)

‡Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, Texas, USA (wleungo@ices.utexas.edu)

heterogeneities in the medium properties. In numerical homogenization approaches, effective properties are computed and the global problem is formulated and solved on the coarse grid. However, these approaches are limited to the cases when the medium properties possess scale separation. On the other hand, multiscale methods construct of multiscale basis functions which are responsible for capturing the local oscillatory effects of the solution. Once the multiscale basis functions, coarse-scale equations are formulated. Moreover, fine-scale information can be recovered by the coarse-scale coefficients and mutliscale basis functions.

Many existing mutliscale methods, such as MsFEM, VMS and HMM, construct one basis function per local coarse region to handle the effects of local heterogeneities. However, for more complex multiscale problems, each local coarse region contains several high-conductivity regions and multiple multiscale basis functions are required to represent the local solution space. GMsFEM is developed to allow systematic enrichment of the coarse-scale space with fine-scale information and identify the underlying low-dimensional local structures for solution representation. The main idea of GMsFEM is to extract local dominant modes by carefully designed local spectral problems in coarse regions, and the convergence of the GMsFEM is related to eigenvalue decay of local spectral problems. For a more detailed discussion on GMsFEM, we refer the readers to [21, 18, 20, 15, 11, 8, 25, 6, 36, 38] and the references therein. Our method developed in this work is motivated by GMsFEM and achieves spectral convergence. Through the design of local spectral problems, our method results in the minimal degree of freedom in representing high-contrast features.

For typical mesh-based numerical discretization, such as the finite element method and finite difference method, the most important issue of the solution accuracy is mesh convergence. However, for multiscale problems, it is difficult to adjust coarse-grid size based on scales and contrast, and it is desirable to have convergence independent of these physical parameters. it becomes non-trivial to derive multiscale methods with convergence on coarse mesh size independent of scales and contrast. Recently, several multiscale methods with mesh convergence are developed. [34, 32, 33]. This idea can be combined with the use of local spectral problems for achieving both spectral convergence and mesh convergence [28, 13, 9, 5]. On the other hand, to overcome the difficulty of stability and conservation for convection-dominated problems and wave propagations in heterogeneous media, multiscale methods in the discontinuous Galerkin (DG) framework have been investigated [23, 2, 37, 19, 26, 24, 12, 14]. In these approaches, unlike conforming finite element formulations, multiscale basis functions are in general discontinuous on the coarse grid, and stabilization or penalty terms are added to ensure well-posedness of the global problem. Our goal in this work is to develop a robust multiscale method in interior penalty discontinuous Galerkin formulation, which exhibits both spectral convergence and mesh convergence.

In this paper, we present Constraint Energy Minimizing Generalized Multiscale Discontinuous Galerkin Method (CEM-GMsDGM). There are two key ingredients of the presented approach. The first main ingredient is the local spectral problems in each coarse block for identification of auxiliary basis functions. The low-energy dominant modes, which are eigenvectors corresponding to small eigenvalues of local spectral problems, are used as auxiliary basis functions for further construction. The auxiliary basis functions possess the information related to high conductivity channels and it suffices to use the same number of auxiliary basis functions as the number of channels in a coarse block. The second ingredient is the constraint energy minimization problems for definition of multiscale basis functions. Each of the auxiliary basis functions sets up an independent constraint and uniquely defines a corresponding multiscale basis function. The multiscale basis functions will then be used to span the multiscale space and used to solve the coarse-scale global problem in IPDG formulation. We remark that the local spectral problems and the constraint energy minimization problems are carefully designed and supported by our analysis. Thanks to the design of local spectral problems, the auxiliary space is of minimal dimension for representing high-contrast features and obtaining a contrast-independent convergence. Due to the fact that the dimensions of the auxiliary space and the multiscale space are identical, the multiscale space is of minimum dimension as well. In the construction of multiscale basis

functions, the constraints are responsible for handling non-decaying components represented by the auxiliary basis functions in the high conductivity regions and achieving linear convergence in coarse mesh size. On the other hand, the multiscale basis functions are supported in oversampled coarse regions and allowed to have discontinuity on the coarse grid. Therefore, the IPDG bilinear form is also used to define the energy term in the constraint energy minimization problems. The advantages of the method is verified both theoretically and numerically. We analyze the method for solving Darcy flow problem and establish a criterion for the oversampling size which is sufficient for linear coarse-mesh convergence independent of the contrast. Numerical results are presented to show the performance of the method for simulation on flow problem and wave propagation in high-contrast heterogeneous media.

The paper is organized as follows. In Section 2, we will introduce the notions of grids, and essential discretization details such as DG finite element spaces and IPDG formulation on the coarse grid. The details of the proposed method will be presented in Section 3. The method will be analyzed in Section 4. Numerical results will be provided in Section 5. Finally, a conclusion will be given in Section 6.

2   Preliminaries

We consider the following high-contrast flow problem

$$- \operatorname{div}(\kappa \nabla u) = f \text{ in } \Omega, \tag{1}$$

subject to the homogeneous Dirichlet boundary condition $u = 0$ on $\partial\Omega$, where $\Omega \subset \mathbb{R}^2$ is the computational domain and f is a given source term. We assume that the permeability field κ is highly heterogeneous with very high contrast $\kappa_0 \leq \kappa \leq \kappa_1$.

Next, we introduce the notions of coarse and fine meshes. We start with a usual partition $\mathcal{T}^H$ of Ω into finite elements, which does not necessarily resolve any multiscale features. The partition $\mathcal{T}^H$ is called a coarse grid and a generic element K in the partition $\mathcal{T}^H$ is called a coarse element. Moreover, $H > 0$ is called the coarse mesh size. We let N_c be the number of coarse grid nodes and N be the number of coarse elements. We also denote the collection of all coarse grid edges by $\mathcal{E}^H$. We perform a refinement of $\mathcal{T}^H$ to obtain a fine grid $\mathcal{T}^h$, where $h > 0$ is called the fine mesh size. It is assumed that the fine grid is sufficiently fine to resolve the solution. An illustration of the fine grid and the coarse grid and a coarse element are shown in Figure 1.
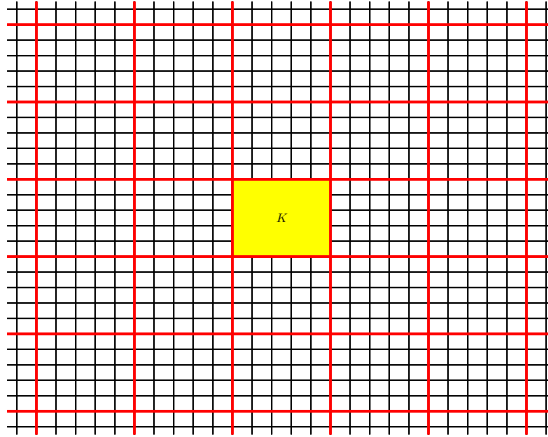

Figure 1: An illustration of the fine grid and the coarse grid and a coarse element.

We are now going to discuss the discontinuous Galerkin (DG) discretization and the interior penalty discontinuous Galerkin (IPDG) global formulation. For the i-th coarse block K_i, we denote the restriction of the Sobolev space $H_0^1(\Omega)$ on K_i by $V(K_i)$. We let $V_h(K_i)$ be the conforming bilinear elements defined on the fine grid $\mathcal{T}^h$ in K_i, i.e.

$$V_h(K_i) = \left\{ v \in V(K_i) : v|_\tau \in \mathbb{Q}^1(\tau) \text{ for all } \tau \in \mathcal{T}^h \text{ and } \tau \subset K_i \right\}, \tag{2}$$

where $\mathbb{Q}^1(\tau)$ stands for the bilinear element on the fine grid block τ. The DG approximation space is then given by the space of coarse-scale locally conforming piecewise bilinear fine-grid basis functions, namely

$$V_h = \oplus_{i=1}^N V_h(K_i). \tag{3}$$

We remark that functions in V_h are continuous within coarse blocks, but discontinuous across the coarse grid edges in general. The global formulation of IPDG method then reads: find $u_h \in V_h$ such that

$$a_{DG}(u_h, w) = \int_\Omega f w \, dx \text{ for all } w \in V_h, \tag{4}$$

where the bilinear form a_{DG} is defined by:

$$a_{DG}(v, w) = \sum_{K \in \mathcal{T}^H} \int_K \kappa \nabla v \cdot \nabla w \, dx - \sum_{E \in \mathcal{E}^H} \int_E \{\kappa \nabla v \cdot n_E\} [\![w]\!] \, d\sigma$$
$$- \sum_{E \in \mathcal{E}^H} \int_E \{\kappa \nabla w \cdot n_E\} [\![v]\!] \, d\sigma + \frac{\gamma}{h} \sum_{E \in \mathcal{E}^H} \int_E \overline{\kappa} [\![v]\!] [\![w]\!] \, d\sigma, \tag{5}$$

where $\gamma > 0$ is a penalty parameter and n_E is a fixed unit normal vector defined on the coarse edge $E \in \mathcal{E}^H$. Note that, in (5), 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, where the unit normal vector n_E is pointing from K^+ to K^-. For a piecewise smooth function G with respect to the coarse grid $\mathcal{T}^H$, we define

$$\{G\} = \frac{1}{2} \left(G^+ + G^- \right),$$
$$[\![G]\!] = G^+ - G^-, \tag{6}$$

where $G^+ = G|_{K^+}$ and $G^- = G|_{K^-}$. 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 $\partial\Omega$, we define $\{G\} = [\![G]\!] = G$, and $\kappa = \kappa_K$ on E, where we always assume that n_E is pointing outside of Ω.

First, we define the energy norm on the space V of coarse-grid piecewise smooth functions by

$$\|w\|_a^2 = a_{DG}(w, w) \text{ for all } w \in V. \tag{7}$$

We also define the DG-norm on V by

$$\|w\|_{DG}^2 = \sum_{K \in \mathcal{T}^H} \int_K \kappa |\nabla w|^2 \, dx + \frac{\gamma}{h} \sum_{E \in \mathcal{E}^H} \int_E \overline{\kappa} [\![w]\!]^2 \, d\sigma \text{ for all } w \in V. \tag{8}$$

The two norms are equivalent on the subspace of piecewise bi-cubic polynomials in V: there exists $C_0 \geq 1$ such that

$$C_0^{-1} \|w\|_a \leq \|w\|_{DG} \leq C_0 \|w\|_a. \tag{9}$$

The continuity and coercivity results of the bilinear form a_{DG} with respect to the DG-norm is ensured by a sufficiently large penalty parameter γ. While the method works well for general highly heterogeneous field κ, we assume κ is piecewise constant on the fine grid $\mathcal{T}^h$ for the sake of simplicity in our analysis presented in Section 4.

3 Method description

In this section, we will present the construction of the multiscale basis functions. First, we will use the concept of GMsFEM to construct our auxiliary multiscale basis functions on a generic coarse block K in the coarse grid. We consider $V_h(K_i)$ as the snapshot space in K_i and perform a dimension reduction through a spectral problem, which is to find a real number $\lambda_j^{(i)}$ and a function $\phi_j^{(i)} \in V_h(K_i)$ such that

$$a_i\left(\phi_j^{(i)}, w\right) = \lambda_j^{(i)} s_i\left(\phi_j^{(i)}, w\right) \text{ for all } w \in V_h(K_i), \tag{10}$$

where a_i is a symmetric non-negative definite bilinear operator and s_i is a symmetric positive definite bilinear operators defined on $V_h(K_i) \times V_h(K_i)$. We remark that the above problem is solved on the fine mesh in the actual computations. Based on our analysis, we can choose

$$
\begin{aligned}
a_i\left(v, w\right) &= \int_{K_i} \kappa \nabla v \cdot \nabla w \, dx, \\
s_i\left(v, w\right) &= \int_{K_i} \widetilde{\kappa} v w \, dx,
\end{aligned}
\tag{11}
$$

where $\tilde{\kappa} = \sum_{j=1}^{N_c} \kappa |\nabla \chi_j^{ms}|^2$ and $\{\chi_j^{ms}\}_{j=1}^{N_c}$ are the standard multiscale finite element (MsFEM) basis functions. We let $\lambda_j^{(i)}$ be the eigenvalues of (10) arranged in ascending order in j, and use the first L_i eigenfunctions to construct our local auxiliary multiscale space

$$V_{aux}^{(i)} = \text{span}\{\phi_j^{(i)} : 1 \le j \le L_i\}. \tag{12}$$

The global auxiliary multiscale space V_{aux}^h is then defined as the sum of these local auxiliary multiscale spaces

$$V_{aux} = \oplus_{i=1}^{N} V_{aux}^{(i)}. \tag{13}$$

For the local auxiliary multiscale space $V_{aux}^{(i)}$, the bilinear form s_i in (11) defines an inner product with norm $\|v\|_{s(K_i)} = s(v, v)^{\frac{1}{2}}$. These local inner products and norms provide a natural definitions of inner product and norm for the global auxiliary multiscale space V_{aux}, which are defined by

$$
\begin{aligned}
s\left(v, w\right) &= \sum_{i=1}^{N} s_i\left(v, w\right) \text{ for all } v, w \in V_{aux}, \\
\|v\|_s &= s\left(v, v\right)^{\frac{1}{2}} \text{ for all } v \in V_{aux}.
\end{aligned}
\tag{14}
$$

We note that $s(v, w)$ and $\|v\|_s$ are also an inner product and norm for the space V_h. Before we move on to discuss the construction of multiscale basis functions, we introduce some tools which will be used to describe our method and analyze the convergence. We first introduce the concept of ϕ-orthogonality. For $1 \le i \le N$ and $1 \le j \le L_i$, in coarse block K_i, given auxiliary basis function $\phi_j^{(i)} \in V_{aux}$, we say that $\psi \in V_h$ is $\phi_j^{(i)}$-orthogonal if

$$s\left(\psi, \phi_{j'}^{(i')}\right) = \delta_{i,i'}\delta_{j,j'} \text{ for all } 1 \le j' \le L_{i'} \text{ and } 1 \le i' \le N. \tag{15}$$

We also introduce a projection operator $\pi : V_h \to V_{aux}$ by $\pi = \sum_{i=1}^{N} \pi_i$, where

$$\pi_i(v) = \sum_{j=1}^{L_i} \frac{s_i\left(v, \phi_j^{(i)}\right)}{s_i\left(\phi_j^{(i)}, \phi_j^{(i)}\right)} \phi_j^{(i)} \text{ for all } v \in V_h, \text{ for all } i = 1, 2, \dots, N. \tag{16}$$

Next, we construct our global multiscale basis functions in V_h. The global multiscale basis function $\psi_j^{(i)} \in V_h$ is defined as the solution of the following constrained energy minimization problem

$$\psi_j^{(i)} = \operatorname{argmin} \left\{ a_{DG} \left(\psi, \psi \right) : \psi \in V_h \text{ is } \phi_j^{(i)}\text{-orthogonal} \right\}. \tag{17}$$

By introducing a Lagrange multiplier, the minimization problem (17) is equivalent to the following variational problem: find $\psi_j^{(i)} \in V_h$ and $\mu_j^{(i)} \in V_{aux}^{(i)}$ such that

$$a_{DG} \left(\psi_j^{(i)}, \psi \right) + s_i \left(\psi, \mu_j^{(i)} \right) = 0 \text{ for all } \psi \in V_h,$$
$$s_i \left(\psi_j^{(i)} - \phi_j^{(i)}, \mu \right) = 0 \text{ for all } \mu \in V_{aux}^{(i)}. \tag{18}$$

Now we discuss the construction our localized multiscale basis functions. We first denote by $K_{i,m}$ an oversampled domain formed by enlarging the coarse grid block K_i by m coarse grid layers. An illustration of an oversampled domain is shown in Figure 2. We introduce the subspace $V_h \left(K_{i,m} \right)$, which contains restriction of fine-scale basis functions in V_h on the oversampled domain $K_{i,m}$. We also define $V_{h,0} \left(K_{i,m} \right) = V_h(K_{i,m}) \cap H_0^1(K_{i,m})$ by the subspace of functions in $V_h \left(K_{i,m} \right)$ vanishing on the boundary of the oversampled domain $K_{i,m}$. Motivated by the construction of
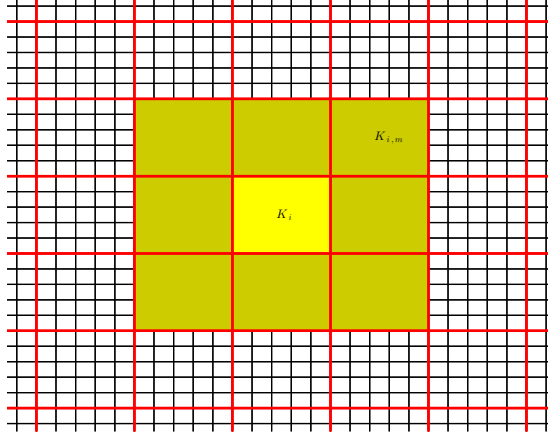

Figure 2: An illustration of an oversampled domain formed by enlarging K_i with 1 coarse grid layer.

our global multiscale basis functions, the method for construction of the localized multiscale basis functions are as follows: The localized multiscale basis function $\psi_{j,ms}^{(i)} \in V_h \left(K_{i,m} \right)$ is defined as the solution of the following constrained energy minimization problem

$$\psi_{j,ms}^{(i)} = \operatorname{argmin} \left\{ a_{DG} \left(\psi, \psi \right) : \psi \in V_h \left(K_{i,m} \right) \text{ is } \phi_j^{(i)}\text{-orthogonal} \right\}. \tag{19}$$

Using the method of Lagrange multiplier, the minimization problem (19) is equivalent to the following variational problem: find $\psi_{j,ms}^{(i)} \in V_h \left(K_{i,m} \right)$ and $\mu_j^{(i)} \in V_{aux}^{(i)}$ such that

$$a_{DG} \left(\psi_{j,ms}^{(i)}, \psi \right) + s_i \left(\psi, \mu_j^{(i)} \right) = 0 \text{ for all } \psi \in V_h \left(K_{i,m} \right),$$
$$s_i \left(\psi_{j,ms}^{(i)} - \phi_j^{(i)}, \mu \right) = 0 \text{ for all } \mu \in V_{aux}^{(i)}. \tag{20}$$

We use the localized multiscale basis functions to construct the multiscale DG finite element space, which is defined as

$$V_{ms} = \operatorname{span}\{\psi_{j,ms}^{(i)} : 1 \le j \le L_i, 1 \le i \le N\}. \tag{21}$$

6

We remark that the multiscale finite element space V_{ms} is a subspace of V_h. After the multiscale DG finite element space is constructed, the multiscale solution u_{ms} is given by: find $u_{ms} \in V_{ms}$ such that

$$a_{DG}(u_{ms}, w) = \int_\Omega fw\, dx \text{ for all } w \in V_{ms}. \tag{22}$$

4 Analysis

In this section, we will analyze the proposed method. Besides the energy norm and the DG norm, we also define the s-norm on V by

$$\|w\|_s^2 = \sum_{K \in \mathcal{T}^H} \int_K \widetilde{\kappa}|w|^2\, dx. \tag{23}$$

Given a subdomain $\Omega' \subseteq \Omega$ formed by a union of coarse blocks $K \in \mathcal{T}^H$, we also define the local s-norm by

$$\|w\|_{s(\Omega')}^2 = \sum_{K \subseteq \Omega'} \int_K \widetilde{\kappa}|w|^2\, dx. \tag{24}$$

The flow of our analysis goes as follows. First, we prove the convergence using the global multiscale basis functions. With the global multiscale basis functions constructed, the global multiscale finite element space is defined by

$$V_{glo} = \text{span}\{\psi_j^{(i)} : 1 \le j \le L_i, 1 \le i \le N\}, \tag{25}$$

and an approximated solution $u_{glo} \in V_{glo}$ is given by

$$a_{DG}(u_{glo}, w) = \int_\Omega fw\, dx \text{ for all } w \in V_{glo}. \tag{26}$$

We remark that the construction of global multiscale basis functions motivates the construction of localized multiscale basis functions. The approximated solution u_{glo} will also be used in our convergence analysis. Next, we give an estimate of the difference between the global multiscale functions $\psi_j^{(i)}$ and the localized multiscale basis functions $\psi_{j,ms}^{(i)}$, in order to show that using the multiscale solution u_{ms} provide similar convergence results as the global solution u_{glo}. For this purpose, we denote the kernel of the projection operator π by $\widetilde{V}_h$. Then, for any $\psi_j^{(i)} \in V_{glo}$, we have

$$a_{DG}\left(\psi_j^{(i)}, w\right) = 0 \text{ for all } w \in \widetilde{V}_h, \tag{27}$$

which implies $\widetilde{V}_h \subseteq V_{glo}^\perp$, where $V_{glo}^\perp$ is the orthogonal complement of V_{glo} with respect to the inner product $a_{DG}(\cdot, \cdot)$. Moreover, since $\dim(V_{glo}) = \dim(V_{aux})$, we have $\widetilde{V}_h = V_{glo}^\perp$ and $V_h = V_{glo} \oplus \widetilde{V}_h$.

4.1 Convergence result

The convergence analysis will start with the following lemma, which concerns about the convergence of the approximated solution by the global multiscale basis functions.

Lemma 1. *Let $u_h \in V_h$ be the solution of (4) and $u_{glo} \in V_{glo}$ be the solution of (26) with the global multiscale basis functions defined by the constrained energy minimization problem (17). Then we have $u_h - u_{glo} \in \widetilde{V}_h$ and*

$$\|u_h - u_{glo}\|_a \le \Lambda^{-\frac{1}{2}}\|\widetilde{\kappa}^{-\frac{1}{2}}f\|_{L^2(\Omega)}, \tag{28}$$

7

where

$$\Lambda = \min_{1 \le i \le N} \lambda_{L_i+1}^{(i)}. \tag{29}$$

Moreover, if we replace the multiscale partition of unity $\{\chi_j^{ms}\}$ by the bilinear partition of unity, we have

$$\|u_h - u_{glo}\|_a \le CH\Lambda^{-\frac{1}{2}} \|\kappa^{-\frac{1}{2}} f\|_{L^2(\Omega)}. \tag{30}$$

Proof. By the definitions of u_h in (4) and u_{glo} in (26), we have

$$a_{DG}(u_h, w) = \int_\Omega fw \, dx \text{ for all } w \in V_h,$$
$$a_{DG}(u_{glo}, w) = \int_\Omega fw \, dx \text{ for all } w \in V_{glo}. \tag{31}$$

Since $V_{glo} \subseteq V_h$, this yields the Galerkin orthogonality property.

$$a_{DG}(u_h - u_{glo}, w) = 0 \text{ for all } w \in V_{glo}, \tag{32}$$

which implies $u_h - u_{glo} \in V_{glo}^\perp = \widetilde{V}_h$. In particular, if we take $w = u_{glo}$ in (32), together with (4), we have

$$\begin{aligned}\|u_h - u_{glo}\|_a^2 &= a_{DG}(u_h, u_h - u_{glo}) \\ &= (f, u_h - u_{glo})_{0,\Omega} \\ &\le \|\widetilde{\kappa}^{-\frac{1}{2}} f\|_{L^2(\Omega)} \|u_h - u_{glo}\|_s.\end{aligned} \tag{33}$$

Since $u_h - u_{glo} \in \widetilde{V}_h$, we have $\pi(u_h - u_{glo}) = 0$. Furthermore, since K_i are disjoint, we have $\pi_i(u_h - u_{glo}) = 0$ for all $i = 1, 2, \ldots, N$. This implies

$$\begin{aligned}\|u_h - u_{glo}\|_s^2 &= \sum_{i=1}^N \|u_h - u_{glo}\|_{s(K_i)}^2 \\ &= \sum_{i=1}^N \|(I - \pi_i)(u_h - u_{glo})\|_{s(K_i)}^2\end{aligned} \tag{34}$$

By the s_i-orthogonality of the eigenfunctions $\phi_j^{(i)}$, we have

$$\begin{aligned}\|(I - \pi_i)(u_h - u_{glo})\|_{s(K_i)}^2 &\le \left(\lambda_{L_i+1}^{(i)}\right)^{-1} a_i(u_h - u_{glo}, u_h - u_{glo}) \\ &\le \Lambda^{-1} a_i(u_h - u_{glo}, u_h - u_{glo}).\end{aligned} \tag{35}$$

Therefore, we have

$$\begin{aligned}\|u_h - u_{glo}\|_s^2 &\le \Lambda^{-1} \sum_{i=1}^N a_i(u_h - u_{glo}, u_h - u_{glo}) \\ &\le \Lambda^{-1} \|u_h - u_{glo}\|_a^2.\end{aligned} \tag{36}$$

Using (33) and (36), we obtain our desired result. The second part of the result follows from the property $|\nabla\chi_j| = O(H^{-1})$ of the bilinear partition of unity. $\square$

The next step is to prove the global basis functions are indeed localizable. This makes use of the following lemma, which states some approximation properties of the projection operator π. In the analysis, we will make use of the Lagrange interpolation operator and a bubble function in

8

the coarse grid. We define the Lagrange interpolation operator $I_h : C^0(\Omega) \cap H_0^1(\Omega) \to C^0(\Omega) \cap V_h$ by: for all $u \in C^0(\Omega) \cap H_0^1(\Omega)$, the interpolant $I_h u \in C^0(\Omega) \cap V_h$ is defined piecewise on each fine block $\tau \in \mathcal{T}^h$ by

$$(I_h u)(x) = u(x) \text{ for all vectices } x \text{ of } \tau, \tag{37}$$

which satisfies the standard approximation properties: there exists $C_I \geq 1$ such that for every $u \in C^0(\Omega) \cap H_0^1(\Omega)$,

$$\left\| \widetilde{\kappa}^{\frac{1}{2}} (u - I_h u) \right\|_{L^2(\tau)} + h \left\| \kappa^{\frac{1}{2}} \nabla (u - I_h u) \right\|_{L^2(\tau)} \leq C_I h \left\| \kappa^{\frac{1}{2}} \nabla u \right\|_{L^2(\tau)}, \tag{38}$$

on each fine block $\tau \in \mathcal{T}^h$. For any coarse grid block K, we define a bubble function B on K, i.e. $B(x) = 0$ for all $x \in \partial K$ and $B(x) > 0$ for all $x \in \text{int}(K)$. More precisely, we take $B = \prod_j \chi_j^{ms}$, where the product is taken over all the coarse grid nodes lying on the boundary ∂K. We can then define the constant

$$C_\pi = \sup_{K \in \mathcal{T}^H, \mu \in V_{aux}} \frac{\int_K \widetilde{\kappa} \mu^2}{\int_K \widetilde{\kappa} B \mu^2}. \tag{39}$$

In our following analysis, we will assume the following smallness criterion on the fine mesh size h:

$$C_\pi C_I (C_{\mathcal{T}}^2 + \lambda_{max}) \|\Theta\|_{L^\infty(\Omega)}^{\frac{1}{2}} h < 1, \tag{40}$$

where $C_{\mathcal{T}}$ is the maximum number of vertices over all coarse elements $K \in \mathcal{T}^H$ and

$$\lambda_{max} = \max_{1 \leq i \leq N} \lambda_{L_i}^{(i)},$$
$$\Theta = \sum_j |\nabla \chi_j^{ms}|^2. \tag{41}$$

Lemma 2. *Assume the smallness criterion* (40) *on the fine mesh size* h. *For any* $v_{aux} \in V_{aux}$, *there exists a function* $v \in C^0(\Omega) \cap V_h$ *such that*

$$\pi(v) = v_{aux}, \quad \|v\|_a^2 \leq D \|v_{aux}\|_s^2, \quad supp(v) \subseteq supp(v_{aux}), \tag{42}$$

where the constant D *is defined by*

$$D = \left(\frac{2 C_\pi (1 + C_I^2) (C_{\mathcal{T}}^2 + \lambda_{max})}{1 - C_\pi C_I (C_{\mathcal{T}}^2 + \lambda_{max}) \|\Theta\|_{L^\infty(K_i)}^{\frac{1}{2}} h} \right)^2. \tag{43}$$

Proof. Let $v_{aux} \in V_{aux}^{(i)}$. We consider the following constraint minimization problem on the block K_i:

$$v = \text{argmin} \left\{ a_{DG}(v, v) : v \in V_{h,0}(K_i), \quad s_i(v, \nu) = s_i(v_{aux}, \nu) \text{ for all } \nu \in V_{aux}^{(i)} \right\}. \tag{44}$$

The minimization problem (44) is equivalent to the following variational problem: find $(v, \mu) \in V_{h,0}(K_i) \times V_{aux}^{(i)}$ such that

$$a_i(v, w) + s_i(w, \mu) = 0 \text{ for all } w \in V_{h,0}(K_i),$$
$$s_i(v - v_{aux}, \nu) = 0 \text{ for all } \nu \in V_{aux}^{(i)}. \tag{45}$$

The existence of solution of (45) is based on an inf-sup condition:

$$\inf_{\nu \in V_{aux}^{(i)}} \sup_{w \in V_{h,0}(K_i) \backslash \{0\}} \frac{s_i(w, \nu)}{a_i(w, w)} \geq \beta, \tag{46}$$

9

where $\beta > 0$ is a constant independent to be determined. Pick any $\nu \in V_{aux}^{(i)}$. We take $w = I_h(B\nu) \in C^0(\Omega) \cap V_h$. Since $\nu \in V_h(K_i)$ and $B(x) = 0$ for all vertices of K_i, we have $w \in V_{h,0}(K_i)$. First, we see that

$$
\begin{aligned}
s_i(w, \nu) &= \int_{K_i} \widetilde{\kappa} B\nu^2 + \int_{K_i} \widetilde{\kappa}(I_h(B\nu) - B\nu)\nu \\
&\geq C_\pi^{-1} \|\nu\|_{s(K_i)}^2 - \|I_h(B\nu) - B\nu\|_s \|\nu\|_{s(K_i)} \\
&\geq C_\pi^{-1} \|\nu\|_{s(K_i)}^2 - C_I h \left\|\widetilde{\kappa}^{\frac{1}{2}} \nabla(B\nu)\right\|_{L^2(K_i)} \|\nu\|_{s(K_i)} \\
&\geq C_\pi^{-1} \|\nu\|_{s(K_i)}^2 - C_I \|\Theta\|_{L^\infty(K_i)}^{\frac{1}{2}} h \left\|\kappa^{\frac{1}{2}} \nabla(B\nu)\right\|_{L^2(K_i)} \|\nu\|_{s(K_i)}.
\end{aligned}
\tag{47}
$$

On the other hand, we observe that

$$
\begin{aligned}
a_i(w, w) &\leq 2 \left(\left\|\kappa^{\frac{1}{2}} \nabla(B\nu)\right\|_{L^2(K_i)}^2 + \left\|\kappa^{\frac{1}{2}} \nabla(B\nu - I_h(B\nu))\right\|_{L^2(K_i)}^2 \right) \\
&\leq 2(1 + C_I^2) \left\|\kappa^{\frac{1}{2}} \nabla(B\nu)\right\|_{L^2(K_i)}.
\end{aligned}
\tag{48}
$$

It remains to estimate the term $\left\|\kappa^{\frac{1}{2}} \nabla(B\nu)\right\|_{L^2(K_i)}$. Since $0 \leq \chi_j^{ms} \leq 1$, we have $0 \leq B \leq 1$ and $|\nabla B|^2 \leq C_{\mathcal{T}}^2 \Theta$. Using these facts together with $\nabla(B\nu) = (\nabla B)\nu + B(\nabla \nu)$, we imply

$$
\begin{aligned}
\left\|\kappa^{\frac{1}{2}} \nabla(B\nu)\right\|_{L^2(K_i)}^2 &\leq C_{\mathcal{T}}^2 \|\nu\|_{s(K_i)}^2 + a_i(\nu, \nu) \\
&\leq \left(C_{\mathcal{T}}^2 + \lambda_{max}\right) \|\nu\|_{s(K_i)}^2
\end{aligned}
\tag{49}
$$

By taking the inf-sup constant

$$
\beta = \frac{1 - C_\pi C_I \left(C_{\mathcal{T}}^2 + \lambda_{max}\right) \|\Theta\|_{L^\infty(K_i)}^{\frac{1}{2}} h}{2C_\pi(1 + C_I^2)\left(C_{\mathcal{T}}^2 + \lambda_{max}\right)},
\tag{50}
$$

we prove the inf-sup condition (46) and therefore the existence of $(v, \mu) \in V_{h,0}(K_i) \times V_{aux}^{(i)}$ in (45). It is then direct to check that the solution $v \in V_{h,0}(K_i)$ satisfies the desired properties. $\qquad \square$

We remark that, without loss of generality, we can assume $D \geq C_0^2(1 + C_I^2)$. We are now going to establish an estimate of the difference between the global multiscale basis functions and localized multiscale basis functions. We will see that the global multiscale basis functions have a decay property, and their values are small outside a suitably large oversampled domain. We will make use of a cutoff function in our proof. For each coarse block K_i and $M > m$, the oversampling regions $K_{i,M}$ and $K_{i,m}$ define an outer neighborhood and an inner neighborhood respectively. We define $\chi_i^{M,m} \in \text{span}\{\chi_j^{ms}\}$ such that $0 \leq \chi_i^{M,m} \leq 1$ and

$$
\chi_i^{M,m} = 1 \text{ in } K_{i,m} \text{ and } \chi_i^{M,m} = 0 \text{ in } \Omega \setminus K_{i,M}.
\tag{51}
$$

Moreover, we define the following DG norm for $w \in V$ on $K_{i,M} \setminus K_{i,m}$.

$$
\|w\|_{DG(K_{i,M} \setminus K_{i,m})}^2 = \sum_{K_k \subset K_{i,M} \setminus K_{i,m}} a_k(w, w) + \frac{\gamma}{h} \sum_{E \in \mathcal{E}^H(K_{i,M} \setminus K_{i,m})} \int_E \overline{\kappa}[\![w]\!]^2 d\sigma,
\tag{52}
$$

where $\mathcal{E}^H(K_{i,M} \setminus K_{i,m})$ denotes the collection of all coarse grid edges in $\mathcal{E}^H$ which lie within in the interior of $K_{i,M} \setminus K_{i,m}$ and the boundary of $K_{i,M}$. We remark that the definition also applies to a region $\Omega \setminus K_{i,m}$ in the case when M is sufficiently large.

10

Lemma 3. *Assume the smallness criterion* (40) *on the fine mesh size* h. *Suppose* $m > 2$ *is the number of coarse grid layers in the oversampled domain* $K_{i,m}$ *extended from the coarse grid block* K_i. *Let* $\phi_j^{(i)} \in V_{aux}$ *be a given auxiliary multiscale basis function. Let* $\psi_j^{(i)} \in V_{glo}$ *be the global multiscale basis function obtained from* (17), *and* $\psi_{j,ms}^{(i)} \in V_h(K_{i,m})$ *be the localized multiscale basis function obtained from* (19). *Then we have*

$$\|\psi_j^{(i)} - \psi_{j,ms}^{(i)}\|_a^2 \leq E\|\phi_j^{(i)}\|_{s(K_i)}^2, \tag{53}$$

where $E = 40D^3(1 + \Lambda^{-1})\left(1 + 6D^{-2}\left(1 + \Lambda^{-\frac{1}{2}}\right)^{-1}\right)^{1-m}$.

Proof. By the variational formulations (18) and (20), we have

$$a_{DG}\left(\psi_j^{(i)} - \psi_{j,ms}^{(i)}, \psi\right) + s_i\left(\psi, \mu_j^{(i)} - \mu_{j,ms}^{(i)}\right) = 0 \text{ for all } \psi \in V_h(K_{i,m}). \tag{54}$$

By Lemma 2, there exists $\widetilde{\phi}_j^{(i)} \in V_h$ such that

$$\pi(\widetilde{\phi}_j^{(i)}) = \phi_j^{(i)}, \quad \|\widetilde{\phi}_j^{(i)}\|_a^2 \leq D\|\phi_j^{(i)}\|_{s(K_i)}^2, \quad \text{supp}\left(\widetilde{\phi}_j^{(i)}\right) \subseteq K_i. \tag{55}$$

We take $\eta = \psi_j^{(i)} - \widetilde{\phi}_j^{(i)} \in V_h$ and $\zeta = \widetilde{\phi}_j^{(i)} - \psi_{j,ms}^{(i)} \in V_h(K_{i,m})$. By definition, we have $\pi(\eta) = \pi(\zeta) = 0$ and therefore $\eta, \zeta \in \widetilde{V}_h$. Again, by Lemma 2, there exists $\rho \in V_h$ such that

$$\pi(\rho) = \pi(I_h(\chi_i^{m,m-1}\eta)), \quad \|\rho\|_a^2 \leq D\|\pi(I_h(\chi_i^{m,m-1}\eta))\|_s^2, \quad \text{supp}(\rho) \subseteq K_{i,m} \setminus K_{i,m-1}. \tag{56}$$

Take $\tau = \rho - I_h(\chi_i^{m,m-1}\eta) \in V_h$. Again, $\pi(\tau) = 0$ and hence $\tau \in \widetilde{V}_h$. Taking $\psi = \tau - \zeta \in V_h(K_{i,m})$ in (54) and making use of the fact $\tau - \zeta \in \widetilde{V}_h$, we have

$$a_{DG}\left(\psi_j^{(i)} - \psi_{j,ms}^{(i)}, \tau - \zeta\right) = 0, \tag{57}$$

and therefore

$$\begin{aligned}
\left\|\psi_j^{(i)} - \psi_{j,ms}^{(i)}\right\|_a^2 &= a_{DG}\left(\psi_j^{(i)} - \psi_{j,ms}^{(i)}, \eta + \zeta\right) \\
&= a_{DG}\left(\psi_j^{(i)} - \psi_{j,ms}^{(i)}, \eta + \tau\right) \\
&\leq \left\|\psi_j^{(i)} - \psi_{j,ms}^{(i)}\right\|_a \|\eta + \tau\|_a,
\end{aligned} \tag{58}$$

which in turn implies

$$\begin{aligned}
\left\|\psi_j^{(i)} - \psi_{j,ms}^{(i)}\right\|_a^2 &\leq \|\eta + \tau\|_a^2 \\
&= \left\|I_h((1 - \chi_i^{m,m-1})\eta) + \rho\right\|_a^2 \\
&\leq 2\left(\left\|I_h((1 - \chi_i^{m,m-1})\eta)\right\|_a^2 + \|\rho\|_a^2\right) \\
&\leq 2\left(C_0^2\left\|I_h((1 - \chi_i^{m,m-1})\eta)\right\|_{DG}^2 + \|\rho\|_a^2\right) \\
&\leq 2\left(2C_0^2(1 + C_I^2)\left\|(1 - \chi_i^{m,m-1})\eta\right\|_{DG}^2 + \|\rho\|_a^2\right).
\end{aligned} \tag{59}$$

11

For the first term on the right hand side of (59), by using $\nabla\left((1-\chi_i^{m,m-1})\eta\right) = -\nabla\chi_i^{m,m-1}\eta + (1-\chi_i^{m,m-1})\nabla\eta$ and $0 \le 1 - \chi_i^{m,m-1} \le 1$, we have

$$\left\|(1-\chi_i^{m,m-1})\eta\right\|_a^2 \le 2\left(\|\eta\|_{DG(\Omega\setminus K_{i,m-1})}^2 + \|\eta\|_{s(\Omega\setminus K_{i,m-1})}^2\right). \tag{60}$$

For the second term on the right hand side of (59), using the definition of ρ in (56) and $0 \le 1 - \chi_i^{m,m-1} \le 1$, we obtain

$$\|\rho\|_a^2 \le D\|\pi(\chi_i^{m,m-1}\eta)\|_s^2 \le D\|\chi_i^{m,m-1}\eta\|_s^2 \le D\|\eta\|_{s(\Omega\setminus K_{i,m-1})}^2. \tag{61}$$

Moreover, since $\eta \in \widetilde{V}_h$, by the spectral problem (10), we have

$$\|\eta\|_{s(\Omega\setminus K_{i,m-1})}^2 \le \Lambda^{-1}\sum_{K_k \subset \Omega\setminus K_{i,m-1}} a_k(\eta,\eta). \tag{62}$$

Combining all these estimates, we obtain

$$\left\|\psi_j^{(i)} - \psi_{j,ms}^{(i)}\right\|_a^2 \le 10D(1+\Lambda^{-1})\|\eta\|_{DG(\Omega\setminus K_{i,m-1})}^2. \tag{63}$$

Next, we will provide a recursive estimate for η in the number of oversampling layers m. We take $\xi = 1 - \chi_i^{m-1,m-2}$. Then $0 \le \xi \le 1$ and $\xi = 1$ in $\Omega \setminus K_{i,m-1}$. Using $\nabla(\xi^2\eta) = \xi^2\nabla\eta + 2\xi\eta\nabla\xi$, for every $K \in \mathcal{T}^H$, we have

$$\int_K \kappa\nabla\eta\cdot\nabla(\xi^2\eta) = \int_K \kappa\nabla\eta\cdot\left(\xi^2\nabla\eta + 2\xi\eta\nabla\xi\right) = \int_K \kappa|\nabla(\xi\eta)|^2 - \int_K \kappa|\nabla\xi|^2\eta^2. \tag{64}$$

In addition, using $\nabla(\xi\eta) = \xi\nabla\eta + \eta\nabla\xi$, for every $E \in \mathcal{E}^H$, we have

$$\begin{aligned}
&-\int_E \{\kappa\nabla\eta\cdot n_E\}[\![\xi^2\eta]\!] - \int_E \{\kappa\nabla(\xi^2\eta)\cdot n_E\}[\![\eta]\!] + \frac{\gamma}{h}\int_E \overline{\kappa}[\![\eta]\!][\![\xi^2\eta]\!] \\
&= -\int_E \{\kappa\nabla\eta\cdot n_E\}[\![\xi^2\eta]\!] - \int_E \{\kappa(\xi^2\nabla\eta + 2\xi\eta\nabla\xi)\cdot n_E\}[\![\eta]\!] + \frac{\gamma}{h}\int_E \overline{\kappa}[\![\eta]\!][\![\xi^2\eta]\!] \\
&= -2\left(\int_E \{\kappa\xi\nabla\eta\cdot n_E\}[\![\xi\eta]\!] + \int_E \{\kappa\eta\nabla\xi\cdot n_E\}[\![\xi\eta]\!]\right) + \frac{\gamma}{h}\int_E \overline{\kappa}[\![\xi\eta]\!]^2 \\
&= -2\int_E \{\kappa\nabla(\xi\eta)\cdot n_E\}[\![\xi\eta]\!] + \frac{\gamma}{h}\int_E \overline{\kappa}[\![\xi\eta]\!]^2.
\end{aligned} \tag{65}$$

Summing over $K \in \mathcal{T}^H$ and $E \in \mathcal{E}^H$, we obtain

$$\|\xi\eta\|_a^2 \le a_{DG}(\eta,\xi^2\eta) + \|\eta\|_{s(K_{i,m-1}\setminus K_{i,m-2})}^2, \tag{66}$$

where we make use of the fact that $\nabla\xi = 0$ outside $K_{i,m-1} \setminus K_{i,m-2}$. We start with estimating the first term on the right hand side of (66). For any coarse element $K_k \in \Omega \setminus K_{i,m-1}$, since $\xi = 1$ in K_k and $\eta \in \widetilde{V}_h$, we have

$$s\left(\xi^2\eta, \phi_j^{(k)}\right) = s\left(\eta, \phi_j^{(k)}\right) = 0 \text{ for all } j = 1, 2, \ldots, L_k. \tag{67}$$

On the other hand, for any coarse element $K_k \in K_{i,m-2}$, since $\xi = 0$ in K_k, we have

$$s\left(\xi^2\eta, \phi_j^{(k)}\right) = 0 \text{ for all } j = 1, 2, \ldots, L_k. \tag{68}$$

12

Therefore, $\text{supp}(\pi(I_h(\xi^2\eta))) \subset K_{i,m-1} \setminus K_{i,m-2}$. By Lemma 2, there exists $\sigma \in V_h$ such that

$$\pi(\sigma) = \pi(I_h(\xi^2\eta)), \quad \|\gamma\|_a^2 \le D\|\pi(I_h(\xi^2\eta))\|_s^2, \quad \text{supp}(\sigma) \subset K_{i,m-1} \setminus K_{i,m-2}. \tag{69}$$

For any coarse element $K_k \subset K_{i,m-1} \setminus K_{i,m-2}$, since $0 \le \xi \le 1$ and $\pi(\eta) = 0$, we have

$$\|\pi(I_h(\xi^2\eta))\|_{s(K_k)}^2 \le \|I_h(\xi^2\eta)\|_{s(K_k)}^2 \le \|I_h(\eta)\|_{s(K_k)}^2 = \|\eta\|_{s(K_k)}^2 \le \Lambda^{-1} a_k(\eta,\eta). \tag{70}$$

Summing over $K_k \subset K_{i,m-1} \setminus K_{i,m-2}$, we obtain

$$\|\pi(I_h(\xi^2\eta))\|_s^2 \le \Lambda^{-1} \sum_{K_k \subset K_{i,m-1} \setminus K_{i,m-2}} a_k(\eta,\eta). \tag{71}$$

We take $\theta = I_h(\xi^2\eta) - \sigma$. Again, $\pi(\theta) = 0$ and $\theta \in \widetilde{V}_h$, which yields

$$a_{DG}\left(\psi_j^{(i)}, \theta\right) = 0. \tag{72}$$

On the other hand, $\text{supp}(\theta) \subset \Omega \setminus K_{i,m-2}$ and $\text{supp}(\widetilde{\phi}_j^{(i)}) \subset K_i$. Since θ and $\widetilde{\phi}_j^{(i)}$ has disjoint supports, we have

$$a_{DG}\left(\widetilde{\phi}_j^{(i)}, \theta\right) = 0. \tag{73}$$

Therefore, we obtain

$$a_{DG}(\eta, \theta) = a_{DG}\left(\psi_j^{(i)} - \widetilde{\phi}_j^{(i)}, \theta\right) = 0. \tag{74}$$

Recall from the definition that $I_h(\xi^2\eta) = \theta + \sigma$ and $\text{supp}(\sigma) \subset K_{i,m-1} \setminus K_{i,m-2}$. Hence we have

$$\begin{aligned}
a_{DG}(\eta, I_h(\xi^2\eta)) &= a_{DG}(\eta, \sigma) \\
&\le C_0 \|\eta\|_{DG(K_{i,m-1} \setminus K_{i,m-2})} \|\sigma\|_a \\
&\le C_0 D^{\frac{1}{2}} \|\eta\|_{DG(K_{i,m-1} \setminus K_{i,m-2})} \|\pi(I_h(\xi^2\eta))\|_s \\
&\le D\Lambda^{-\frac{1}{2}} \|\eta\|_{DG(K_{i,m-1} \setminus K_{i,m-2})}^2.
\end{aligned} \tag{75}$$

On the other hand, making use of the fact that $\xi^2 = 0$ in $K_{i,m-2}$ and $\xi^2 = 1$ in $\Omega \setminus K_{i,m-1}$, we observe that $\xi^2\eta = I_h(\xi^2\eta)$ outside $K_{i,m-1} \setminus K_{i,m-2}$. Moreover, $\xi^2\eta - I_h(\xi^2\eta)$ is globally continuous. Thus, we obtain

$$\begin{aligned}
a_{DG}(\eta, \xi^2\eta - I_h(\xi^2\eta)) &\le C_0^2 \|\eta\|_{DG(K_{i,m-1} \setminus K_{i,m-2})} \left\|\xi^2\eta - I_h(\xi^2\eta)\right\|_{DG(K_{i,m-1} \setminus K_{i,m-2})} \\
&\le C_0^2 C_I \|\eta\|_{DG(K_{i,m-1} \setminus K_{i,m-2})} \left\|\xi^2\eta\right\|_{DG(K_{i,m-1} \setminus K_{i,m-2})} \\
&\le \frac{D}{2} \left(\|\eta\|_{DG(K_{i,m-1} \setminus K_{i,m-2})}^2 + \left\|\xi^2\eta\right\|_{DG(K_{i,m-1} \setminus K_{i,m-2})}^2\right).
\end{aligned} \tag{76}$$

Again, using $\nabla(\xi^2\eta) = \xi^2\nabla\eta + 2\xi\eta\nabla\xi$, we have

$$\left\|\xi^2\eta\right\|_{DG(K_{i,m-1} \setminus K_{i,m-2})}^2 \le 2\|\eta\|_{DG(K_{i,m-1} \setminus K_{i,m-2})}^2 + 8\|\eta\|_{s(K_{i,m-1} \setminus K_{i,m-2})}^2. \tag{77}$$

Combining (66), (75), (76) and (77), we arrive at

$$\|\xi\eta\|_a^2 \le D\left(\left(\frac{3}{2} + \Lambda^{-\frac{1}{2}}\right) \|\eta\|_{DG(K_{i,m-1} \setminus K_{i,m-2})}^2 + 5\|\eta\|_{s(K_{i,m-1} \setminus K_{i,m-2})}^2\right). \tag{78}$$

Moreover, since $\pi(\eta) = 0$, we have

$$\|\eta\|_{s(K_{i,m-1} \setminus K_{i,m-2})} \le \Lambda^{-\frac{1}{2}} \|\eta\|_{DG(K_{i,m-1} \setminus K_{i,m-2})}, \tag{79}$$

13

which implies

$$\|\xi\eta\|_a^2 \le 6D\left(1 + \Lambda^{-\frac{1}{2}}\right)\|\eta\|_{DG(K_{i,m-1}\setminus K_{i,m-2})}^2. \tag{80}$$

By the equivalence of norms, we have

$$\|\eta\|_{DG(\Omega\setminus K_{i,m-1})}^2 \le C_0^2\|\xi\eta\|_a^2 \le 6D^2\left(1 + \Lambda^{-\frac{1}{2}}\right)\|\eta\|_{DG(K_{i,m-1}\setminus K_{i,m-2})}^2. \tag{81}$$

We obtain the recurrence estimate

$$\begin{aligned}
\|\eta\|_{DG(\Omega\setminus K_{i,m-2})}^2 &= \|\eta\|_{DG(\Omega\setminus K_{i,m-1})}^2 + \|\eta\|_{DG(K_{i,m-1}\setminus K_{i,m-2})}^2 \\
&\ge \left(1 + 6D^{-2}\left(1 + \Lambda^{-\frac{1}{2}}\right)^{-1}\right)\|\eta\|_{DG(\Omega\setminus K_{i,m-1})}^2.
\end{aligned} \tag{82}$$

Inductively, we have

$$\begin{aligned}
\|\eta\|_{DG(\Omega\setminus K_{i,m-1})}^2 &\le \left(1 + 6D^{-2}\left(1 + \Lambda^{-\frac{1}{2}}\right)^{-1}\right)^{1-m}\|\eta\|_{DG(\Omega\setminus K_{i,1})}^2 \\
&\le D\left(1 + 6D^{-2}\left(1 + \Lambda^{-\frac{1}{2}}\right)^{-1}\right)^{1-m}\|\eta\|_a^2.
\end{aligned} \tag{83}$$

Combining (63) and (83), we see that

$$\left\|\psi_j^{(i)} - \psi_{j,ms}^{(i)}\right\|_a^2 \le 10D^2\left(1 + \Lambda^{-1}\right)\left(1 + 6D^{-2}\left(1 + \Lambda^{-\frac{1}{2}}\right)^{-1}\right)^{1-m}\|\eta\|_a^2 \tag{84}$$

By the energy minimizing property of $\psi_j^{(i)}$, we have

$$\|\eta\|_a \le \|\psi_j^{(i)}\|_a + \|\widetilde{\phi}_j^{(i)}\|_a \le 2\|\widetilde{\phi}_j^{(i)}\|_a \le 2D^{\frac{1}{2}}\|\phi_j^{(i)}\|_{s(K_i)}. \tag{85}$$

We obtain the desired result. $\qquad\square$

Now, we are ready to establish our main theorem, which estimates the error between the solution u_h and the multiscale solution u_{ms}.

Theorem 4. *Let $u_h \in V_h$ be the solution of (4), $u_{glo} \in V_{glo}$ be the solution of (26) with the global multiscale basis functions defined by (17), and $u_{ms} \in V_{ms}$ be the multiscale solution of (22) with the localized multiscale basis functions defined on an oversampled domain with $m > 2$ coarse grid layers by (19). Then we have*

$$\|u_h - u_{ms}\|_a \le C\Lambda^{-\frac{1}{2}}\|\widetilde{\kappa}^{-\frac{1}{2}}f\|_{L^2(\Omega)} + Cm^d E^{\frac{1}{2}}\|u_{glo}\|_s, \tag{86}$$

Moreover, if we let $k = O\left(\log\left(\frac{\kappa_1}{H}\right)\right)$ and replace the multiscale partition of unity $\{\chi_j^{ms}\}$ by the bilinear partition of unity, we have

$$\|u_h - u_{ms}\|_a \le CH\Lambda^{-\frac{1}{2}}\|\kappa^{-\frac{1}{2}}f\|_{L^2(\Omega)}. \tag{87}$$

Proof. First, we write u_{glo} in the linear combination of the basis $\{\psi_k^{(j)}\}$

$$u_{glo} = \sum_{i=1}^{N}\sum_{j=1}^{L_i}\alpha_j^{(i)}\psi_j^{(i)}. \tag{88}$$

and define $\widehat{u}_{ms} \in V_{ms}$ by

$$\widehat{u}_{ms} = \sum_{i=1}^{N}\sum_{j=1}^{L_i}\alpha_j^{(i)}\psi_{j,ms}^{(i)}. \tag{89}$$

From (4) and (22), we obtain the Galerkin orthogonality

$$a_{DG}(u_h - u_{ms}, w) = 0 \text{ for all } w \in V_{ms}, \tag{90}$$

which gives

$$\|u_h - u_{ms}\|_a \leq \|u_h - \widehat{u}_{ms}\|_a \leq \|u_h - u_{glo}\|_a + \|u_{glo} - \widehat{u}_{ms}\|_a. \tag{91}$$

Using Lemma 3, we see that

$$
\begin{aligned}
\|u_{glo} - \widehat{u}_{ms}\|_a^2 &= \left\| \sum_{i=1}^{N} \sum_{j=1}^{L_i} \alpha_j^{(i)} (\psi_j^{(i)} - \psi_{j,ms}^{(i)}) \right\|_a^2 \\
&\leq C m^d \sum_{i=1}^{N} \left\| \sum_{j=1}^{L_i} \alpha_j^{(i)} (\psi_j^{(i)} - \psi_{j,ms}^{(i)}) \right\|_a^2 \\
&\leq C m^d E \sum_{i=1}^{N} \left\| \sum_{j=1}^{L_i} \alpha_j^{(i)} \phi_j^{(i)} \right\|_s^2 \\
&= C m^d E \|u_{glo}\|_s^2,
\end{aligned}
\tag{92}
$$

where the last equality follows from the orthogonality of the eigenfunctions in (10). Using the estimates (28) and (92) in (91), we have

$$\|u_h - u_{ms}\|_a \leq \Lambda^{-\frac{1}{2}} \|\widetilde{\kappa}^{-\frac{1}{2}} f\|_{L^2(\Omega)} + C m^{\frac{d}{2}} E^{\frac{1}{2}} \|u_{glo}\|_s. \tag{93}$$

This completes the first part of the theorem. Next, we assume the partition of unity functions are bilinear, and we are going to estimate $\|u_{glo}\|_s$. Using the fact that $|\nabla \chi_k| = O(H^{-1})$, we have

$$\|u_{glo}\|_s^2 \leq C H^{-2} \kappa_1 \|u_{glo}\|_{L^2(\Omega)}^2. \tag{94}$$

Then, by Poincaré inequality, we have

$$\|u_{glo}\|_{L^2(\Omega)}^2 \leq C \kappa_0^{-1} \|u_{glo}\|_a^2. \tag{95}$$

By taking $w = u_{glo} \in V_{glo}$ in (26), we obtain

$$\|u_{glo}\|_a^2 = (f, u_{glo})_{0,\Omega} \leq \|\widetilde{\kappa}^{-\frac{1}{2}} f\|_{L^2(\Omega)} \|u_{glo}\|_s. \tag{96}$$

Combining these estimates, we have

$$\|u_{glo}\|_s \leq C H^{-2} \kappa_0^{-1} \kappa_1 \|\widetilde{\kappa}^{-\frac{1}{2}} f\|_{L^2(\Omega)}. \tag{97}$$

To obtain our desired result, we need

$$H^{-2} \kappa_1 m^{\frac{d}{2}} E^{\frac{1}{2}} = O(1). \tag{98}$$

Taking logarithm, we have

$$\log(H^{-2}) + \log(\kappa_1) + \frac{d}{2} \log(m) + \frac{1-m}{2} \log\left(1 + \Lambda^{-\frac{1}{2}}\right) = O(1). \tag{99}$$

Thus, taking $m = O\left(\log\left(\frac{\kappa_1}{H}\right)\right)$ completes the proof of the second result. $\qquad\square$

5 Numerical results

In this section, we will present numerical examples with high contrast media to demonstrate the convergence of our proposed method with respect to the coarse mesh size H and the number of oversampling layers m, and illustrate possible improvements in error robustness with respect to contrast by employing the idea of constructing multiscale basis function by relaxation method introduced in [13]. Lastly, we examine the performance of applying the method to the wave equation. In all the experiments, the IPDG penalty parameter in (5) is set to be $\gamma = 4$, so as to ensure the coercivity of the bilinear form a_{DG}.

5.1 Experiment 1: flow problem

In the first experiment, we consider a highly heterogeneous permeability field κ in $\Omega = [0.1]^2$ as shown in Figure 3, with the background value is $\kappa = 1$ and the value in the channels and inclusions is 10^4. and the resolution is 400×400, i.e. κ is piecewise constant on a fine grid with mesh size $h = 1/400$. The coarse mesh size varies from $H = 1/80$ to $H = 1/10$, and the number of oversampling layers varies from $m = 3$ to $m = 6$. In all these combinations, there are no more than 3 high conductivity channels in a coarse block $K \in \mathcal{T}^H$. As a result, we have 3 small eigenvalues in a local spectral problem (10), and it suffices to use 3 auxiliary basis functions per coarse block to construct the correspoding localized multiscale basis functions. The source function is taken as

$$f(x, y) = 2\pi^2 \sin(\pi x) \sin(\pi y) \text{ for all } (x, y) \in \Omega. \tag{100}$$

Table 1 records the error when we take the number of oversampling layer to be approximately $m \approx 4 \log(1/H)/\log(1/10)$. The results show that the method provides optimal convergence in energy norm, which agrees with our theoretical finding in Section 4, and the L^2 error converges with second order. Table 2 records the error with various number of oversampling layers and a fixed coarse mesh sizes $H = 1/40$. It can be observed that increasing the number of oversampling layers improves the quality of approximations, but the decay in error is limited when the oversampling region is sufficiently large. This numerically verifies that the multiscale basis functions can indeed be localized.
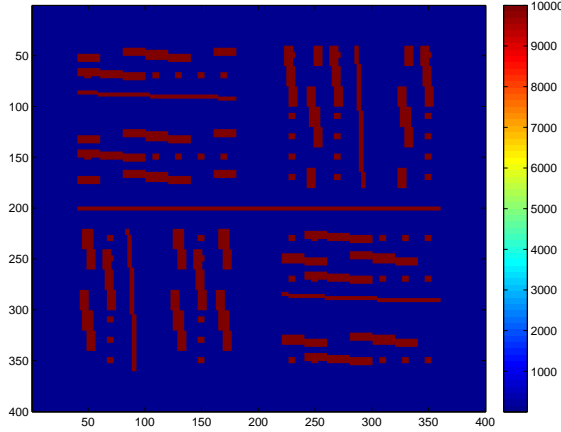

Figure 3: The permeability field κ for Experiment 1.

Next, we present the idea of the relaxed formulation of (19). Instead of using the method of Lagrange multiplier as in (20), the ϕ-orthogonality is imposed weakly by a penalty formula-

m	H	Energy error	L^2 error
4	1/10	7.4625%	0.7653%
6	1/20	1.5392%	0.0625%
7	1/40	0.7266%	0.0160%
8	1/80	0.3433%	0.0035%

Table 1: History of convergence with number of oversampling layers $m \approx 4\log(1/H)/\log(1/10)$ for Experiment 1.

m	Energy error	L^2 error
3	84.7517%	72.3079%
4	19.0936%	3.6716%
5	2.6687%	0.0720%
6	0.7836%	0.0161%
7	0.7266%	0.0160%
8	0.7259%	0.0160%

Table 2: Error table with different number of oversampling layers m and a fixed coarse mesh size $H = 1/40$ for Experiment 1.

tion. The localized multiscale basis function $\psi_{j,ms}^{(i)} \in V_h(K_{i,m})$ is defined as the solution of the following relaxed constrained energy minimization problem

$$\psi_{j,ms}^{(i)} = \operatorname{argmin}\left\{ a_{DG}(\psi, \psi) + s\left(\pi(\psi) - \phi_j^{(i)}, \pi(\psi) - \phi_j^{(i)}\right) : \psi \in V_h(K_{i,m})\right\}. \tag{101}$$

The minimization problem (101) is equivalent to the following variational problem: find $\psi_{j,ms}^{(i)} \in V_h(K_{i,m})$ such that

$$a_{DG}\left(\psi_{j,ms}^{(i)}, \psi\right) + s\left(\pi\left(\psi_{j,ms}^{(i)}\right), \pi(\psi)\right) = s\left(\phi_j^{(i)}, \pi(\psi)\right) \text{ for all } \psi \in V_h(K_{i,m}). \tag{102}$$

The construction of multiscale finite element space and coarse-scale model then follow (21) and (22) respectively. We compare the performance of the multiscale method with multiscale basis functions constructed by method of Lagrange multiplier (20) and the relaxation method (102) at different contrast values, where the coarse mesh size is taken as $H = 1/10$ and the number of oversampling layers as $m = 4$. In Table 3, we record the energy error and L^2 error with different contrast κ_1, where $\kappa_1 \gg 1$ is the value of κ in the high conductivity channels. It can be seen that the relaxation method is more robust with respect to contrast.

κ_1	Lagrange multiplier		Relaxation	
	Energy error	L^2 error	Energy error	L^2 error
10^4	7.4625%	0.7653%	6.3757%	0.6395%
10^5	12.6299%	1.6977%	6.3986%	0.6467%
10^6	32.1465%	10.5146%	6.4020%	0.6478%
10^7	64.1190%	41.8127%	6.4049%	0.6481%
10^8	77.1229%	60.4947%	6.4301%	0.6503%

Table 3: Comparison of the method of Lagrange multiplier and the relaxation method with different contrast values for Experiment 1.

5.2 Experiment 2: wave propagation

In the second experiment, we consider a simple wave equation in the space-time domain $\Omega_T = [0, T] \times \Omega$, where $[0, T]$ is the temporal domain and Ω is the spatial domain:

$$\frac{\partial^2 u}{\partial t^2} - \operatorname{div}(\kappa \nabla u) = f, \tag{103}$$

with homogeneous Dirichlet boundary condition $u = 0$ on $[0, T] \times \partial\Omega$. Here div and ∇ correspond to the divergence and gradient with respect to the spatial variable x, and κ is the bulk modulus which is assumed to be stationary and highly oscillatory. Using second-order central difference for temporal discretization on a uniform grid points $t_k = k\Delta t$ and IPDG formulation on the space V_h, the fully-discrete DG finite element scheme reads: find $u_h^{n+1} \in V_h$ such that

$$\frac{u_h^{n+1} - 2u_h^n + u_h^{n-1}}{\Delta t^2} + a_{DG}(u_h^n, w) = \int_\Omega f^n w \text{ for all } w \in V_h, \tag{104}$$

where the superscript k stands for the evaluation of a function at the time instant t_k, and the initial conditions u_h^0, u_h^1 are given. Next, we use the CEM method in Section 3 to construct a multiscale finite element space V_{ms}. The coarse-scale model then reads: find $u_{ms}^{n+1} \in V_{ms}$ such that

$$\frac{u_{ms}^{n+1} - 2u_{ms}^n + u_{ms}^{n-1}}{\Delta t^2} + a_{DG}(u_{ms}^n, w) = \int_\Omega f^n w \text{ for all } w \in V_{ms}. \tag{105}$$

In this experiment, we take the bulk modulus on the spatial domain $\Omega = [0, 1]^2$ as part of the Marmousi model as shown in Figure 4. The fine mesh size is taken as $h = 1/256$. The coarse mesh size varies from $H = 1/32$ to $H = 1/8$, and the number of oversampling layers varies from $m = 4$ to $m = 6$. In all these combinations, we use 4 auxiliary basis functions per coarse block to construct the correspoding localized multiscale basis functions. The source function f is taken as the Ricker wavelet

$$f(t, x, y) = \frac{t - 2/f_0}{4h^2} \exp\left(-\pi^2 f_0^2 (t - 2/f_0)^2\right) \exp\left(\frac{(x - 0.5)^2 + (y - 0.5)^2}{4h^2}\right) \text{ for all } (t, x, y) \in \Omega_T. \tag{106}$$

where the central frequency is chosen as $f_0 = 20$. We iteratively solve for the numerical solution at the final time $T = 0.2$ with time step size $\Delta t = 10^{-4}$. Table 4 records the error of the final solution when we take the number of oversampling layer to be approximately $m \approx 4\log(1/H)/\log(1/8)$. It can been observed that the method results in good accuracy and desired convergence in error. Figure 5 depicts the numerical solutions by the fine-scale formulation and the coarse-scale formulation at the final time $T = 2$. The comparison suggests that the CEM method provides very good accuracy at a reduced computational expense.

m	H	Energy error	L^2 error
4	1/8	67.1831%	40.4876%
5	1/16	24.1360%	9.7541%
6	1/32	4.8839%	1.1559%

Table 4: History of convergence for Experiment 2.

6 Conclusion

In this paper, we present CEM-GMsDGM, a local multiscale model reduction approach in the discontinuous Galerkin framework. The multiscale basis functions are defined in coarse oversampled regions by a constraint energy minimization problem, which are in general discontinuous
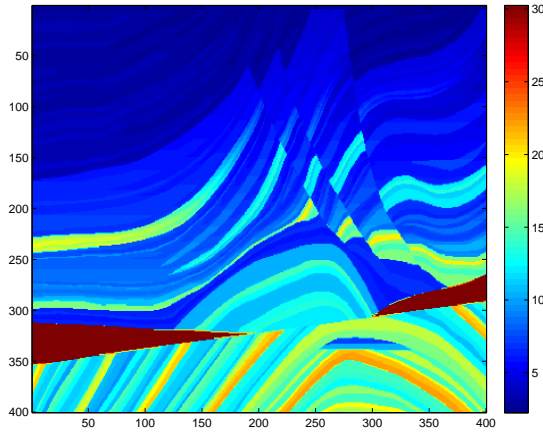
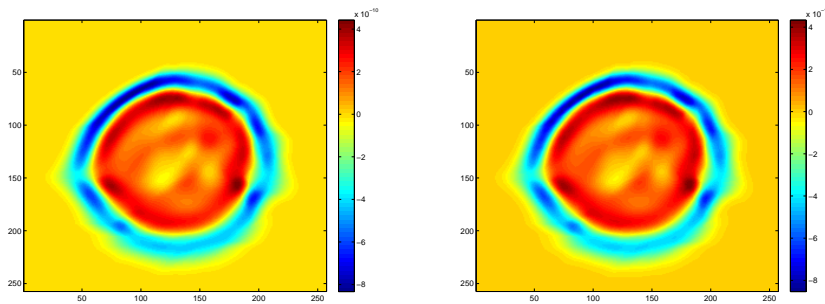Figure 4: Marmousi model for Experiment 2.

Figure 5: Plots of numerical solution for Experiment 2. Fine solution (left) and multiscale solution (right).

on the coarse grid, and coupled by the IPDG formulation. Thanks to the definition of local spectral problems, the dimension of auxiliary space is minimal for sufficiently representing the high conductivity regions, and provides the most locally compressed multiscale space. In our analysis for the Darcy flow problem, we show that the method provides optimal convergence in the coarse mesh size, which is independent of the contrast, provided that the oversampling size is appropriately chosen. The convergence of the method for solving Darcy flow is theoretically analyzed and numerically verified. Numerical results for applying the method on a simple wave equation are also presented.

References

[1] A. Abdulle. On a priori error analysis of fully discrete heterogeneous multiscale fem. *SIAM J. Multiscale Modeling and Simulation*, 4(2):447–459, 2005.

[2] A Buffa, TJR Hughes, and G Sangalli. Analysis of a multiscale discontinuous Galerkin method for convection-diffusion problems. *SIAM Journal on Numerical Analysis*, 44(4):1420–1440, 2006.

[3] Victor Calo, Yalchin Efendiev, and Juan Galvis. A note on variational multiscale methods for high-contrast heterogeneous porous media flows with rough source terms. *Advances in water resources*, 34(9):1177–1185, 2011.

[4] Z. Chen and T. Y. Hou. A mixed multiscale finite element method for elliptic problems with oscillating coefficients. *Mathematics of Computation*, 72(242):541–576, 2003.

[5] Siu Wun Cheung, Eric T Chung, Yalchin Efendiev, Wing Tat Leung, and Maria Vasilyeva. Constraint energy minimizing generalized multiscale finite element method for dual continuum model. *arXiv preprint arXiv:1807.10955*, 2018.

[6] Siu Wun Cheung and Nilabja Guha. Dynamic data-driven bayesian gmsfem. *Journal of Computational and Applied Mathematics*, 353:72 – 85, 2019.

[7] C.C. Chu, I.G. Graham, and T. Hou. A new multiscale finite element methods for high-contrast elliptic interface problem. *Mathematics of Computation*, 79:1915–1955, 2010.

[8] Eric Chung, Yalchin Efendiev, and Thomas Y Hou. Adaptive multiscale model reduction with generalized multiscale finite element methods. *Journal of Computational Physics*, 320:69–95, 2016.

[9] Eric Chung, Yalchin Efendiev, and Wing Tat Leung. Constraint energy minimizing generalized multiscale finite element method in the mixed formulation. *Computational Geosciences*, 22(3):677–693, 2018.

[10] Eric T Chung, Yalchin Efendiev, Richard L Gibson Jr, and Maria Vasilyeva. A generalized multiscale finite element method for elastic wave propagation in fractured media. *GEM-International Journal on Geomathematics*, pages 1–20, 2015.

[11] Eric T Chung, Yalchin Efendiev, and Wing Tat Leung. Residual-driven online generalized multiscale finite element methods. *Journal of Computational Physics*, 302:176–190, 2015.

[12] Eric T. Chung, Yalchin Efendiev, and Wing Tat Leung. An online generalized multiscale discontinuous galerkin method (gmsdgm) for flows in heterogeneous media. *Communications in Computational Physics*, 21(2):401422, 2017.

[13] Eric T Chung, Yalchin Efendiev, and Wing Tat Leung. Constraint energy minimizing generalized multiscale finite element method. *Computer Methods in Applied Mechanics and Engineering*, 339:298–319, 2018.

[14] E.T. Chung, Y. Efendiev, and W.T. Leung. An adaptive generalized multiscale discontinuous galerkin method for high-contrast flow problems. *SIAM Multiscale Modeling and Simulation*, 16(3):1227–1257, 2018.

[15] ET Chung, Y Efendiev, and G Li. An adaptive GMsFEM for high-contrast flow problems. *Journal of Computational Physics*, 273:54–76, 2014.

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

[17] W. E, P. Ming, and P. Zhang. Analysis of the heterogeneous multiscale method for elliptic homogenization problems. *J. Amer. Math. Soc.*, 18(1):121–156, 2005.

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

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

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

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

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

[23] Y. Efendiev, T. Hou, and X.H. Wu. Convergence of a nonconforming multiscale finite element method. *SIAM J. Numer. Anal.*, 37:888–910, 2000.

[24] Yalchin Efendiev, Raytcho Lazarov, Minam Moon, and Ke Shi. A spectral multiscale hybridizable discontinuous Galerkin method for second order elliptic problems. *Computer Methods in Applied Mechanics and Engineering*, 292:243–256, 2015.

[25] Yalchin Efendiev, Wing Tat Leung, S. W. Cheung, N. Guha, V. H. Hoang, and B. Mallick. Bayesian multiscale finite element methods. modeling missing subgrid information probabilistically. *International Journal for Multiscale Computational Engineering*, 15(2):175–197, 2017.

[26] D. Elfverson, E. Georgoulis, A. Mlqvist, and D. Peterseim. Convergence of a discontinuous galerkin multiscale method. *SIAM Journal on Numerical Analysis*, 51(6):3351–3372, 2013.

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

[28] Thomas Y Hou and Pengchuan Zhang. Sparse operator compression of higher-order elliptic operators with rough coefficients. *Research in the Mathematical Sciences*, 4(1):24, 2017.

[29] T.J.R. Hughes, G.R. Feijóo, L. Mazzei, and J.-B. Quincy. The variational multiscale method - a paradigm for computational mechanics. *Comput. Methods Appl. Mech Engrg.*, 127:3–24, 1998.

[30] TJR Hughes and G Sangalli. Variational multiscale analysis: the fine-scale Green's function, projection, optimization, localization, and stabilized methods. *SIAM Journal on Numerical Analysis*, 45(2):539–557, 2007.

[31] O. Iliev, R. Lazarov, and J. Willems. Variational multiscale finite element method for flows in highly porous media. *Multiscale Model. Simul.*, 9(4):1350–1372, 2011.

[32] Axel Målqvist and Daniel Peterseim. Localization of elliptic multiscale problems. *Mathematics of Computation*, 83(290):2583–2603, 2014.

[33] Houman Owhadi. Multigrid with rough coefficients and multiresolution operator decomposition from hierarchical information games. *SIAM Review*, 59(1):99–149, 2017.

[34] Houman Owhadi, Lei Zhang, and Leonid Berlyand. Polyharmonic homogenization, rough polyharmonic splines and sparse super-localization. *ESAIM: Mathematical Modelling and Numerical Analysis*, 48(2):517–552, 2014.

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

[36] Jun Sur Richard Park, Siu Wun Cheung, Tina Mai, and Viet Ha Hoang. Multiscale simulations for upscaled multi-continuum flows. *arXiv preprint arXiv:1909.04722*, 2019.

[37] Béatrice Rivière. *Discontinuous Galerkin methods for solving elliptic and parabolic equations: theory and implementation.* Society for Industrial and Applied Mathematics, 2008.

[38] Min Wang, Siu Wun Cheung, Eric T. Chung, Maria Vasilyeva, and Yuhe Wang. Generalized multiscale multicontinuum model for fractured vuggy carbonate reservoirs. *Journal of Computational and Applied Mathematics*, 366:112370, 2020.

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