GENERALIZED MULTISCALE FINITE ELEMENT METHOD FOR THE STEADY STATE LINEAR BOLTZMANN EQUATION[*]

ERIC CHUNG[†], YALCHIN EFENDIEV[‡], YANBO LI[§], AND QIN LI[¶]

Abstract. The Boltzmann equation, as a model equation in statistical mechanics, is used to describe the statistical behavior of a large number of particles driven by the same physics laws. Depending on the media and the particles to be modeled, the equation has slightly different forms. In this article, we investigate a model Boltzmann equation with highly oscillatory media in the small Knudsen number regime and study the numerical behavior of the generalized multiscale finite element method (GMsFEM) in the fluid regime when high oscillation in the media presents. The GMsFEM is a general approach [E. Chung, Y. Efendiev, and T. Y. Hou, *J. Comput. Phys.*, 320 (2016), pp. 69–95] to numerically treat equations with multiscale structures. The method is divided into the offline and online steps. In the offline step, basis functions are prepared from a snapshot space via a well-designed generalized eigenvalue problem (GEP), and these basis functions are then utilized to patch up for a solution through DG formulation in the online step to incorporate specific boundary and source information. We prove the well-posedness of the method on the Boltzmann equation and show that the GEP formulation provides a set of optimal basis functions that achieve spectral convergence. Such convergence is independent of the oscillation in the media, or the smallness of the Knudsen number, making it one of the few methods that simultaneously achieve numerical homogenization and asymptotic preserving properties across all scales of oscillations and the Knudsen number.

Key words. multiscale finite element method, linear Boltzmann equation, high contrast

AMS subject classifications. 65N12, 65N30

DOI. 10.1137/19M1256282

1. Introduction. The Boltzmann equation is a fundamental model in statistical mechanics. It traces the evolution of the distribution function on the phase space and describes the dynamics of a large number of particles that follow the same physics rules via a statistical manner. The equation encodes the particles' free transport and their interactions with the media and each other. Depending on the physics the particles follow, the interaction term may differ, but to a large extent, many particles, including neutrons, photons, and phonons, interact mainly with the media, making the collision term linear. The dynamics then can be described by the linear Boltzmann equation:

$$\partial_t u + v \cdot \nabla u = \sigma \mathcal{R} u(x, v) - \eta u, \quad (x, v) \in \Omega \times V. \tag{1}$$

In the equation, u is a function on the phase space $(x, v) \in \Omega \times V$. The evolution is governed by $v \cdot \nabla u$, a free transport term, and the terms on the right side of the

[†]Department of Mathematics, The Chinese University of Hong Kong, Hong Kong SAR (tschung@math.cuhk.edu.hk).

[‡]Department of Mathematics & Institute for Scientific Computation (ISC), Texas A&M University, College Station, TX 77840 (efendiev@math.tamu.edu).

[§]Department of Mathematics, Texas A&M University, College Station, TX 77840 (lyb@tamu.edu).

[¶]Department of Mathematics, University of Wisconsin, Madison, WI 53706 (qinli@math.wisc.edu).

equation that represent the "collision" and quantify the particles' interactions. These interactions include a pure absorption term ηu, where η is the absorption coefficient, and a scattering term $\sigma \mathcal{R} u$. The specific form of the operator $\mathcal{R}$ varies from particle to particle, and it is typically a functional independent of x. The strength of the interactions is governed by the size of σ and η. For photons specifically, the radiative transfer equation is used; these coefficients are termed the optical thickness. In this paper, for simplicity, we take $\mathcal{R}$ to be

$$(2) \qquad \mathcal{R}u(x,v) = \int_V u(x,v')dv' - u(x,v),$$

where dv is a normalized measure, i.e., $\int dv = 1$, and we set $\eta = 1$.

The equation demonstrates different behavior in different regimes. One particularly interesting regime is called the diffusion regime, in which the scattering coefficient is extremely strong and the pure absorption term is weak. Mathematically, considering the steady state case, we model the equation to

$$(3) \qquad v \cdot \nabla u + \epsilon\, u = \frac{1}{\epsilon}\sigma \mathcal{R}u(x,v)\,.$$

In this equation, ϵ is termed the Knudsen number, and it characterizes the ratio of the mean free path and the typical domain length. Physically it reflects the number of collisions a normal particle experiences inside Ω before emitting. When ϵ is small, the number of collisions per particle is large, meaning the particle gets scattered many times before emitting, and thus some kind of averaging effects take place, and the local equilibrium is achieved. In the case of (2), the equilibrium reads

$$(4) \qquad u(x,v) \sim \rho(x)\,,$$

and through asymptotic analysis, one could mathematically derive that ρ satisfies the diffusion equation:

$$(5) \qquad C\nabla \cdot \left(\frac{1}{\sigma}\nabla\rho\right) = \rho\,,$$

where C depends on the dimension.

The convergence from (3) to the asymptotic limit (5) was conjectured in [4] and was made rigorous in [3] for periodic boundary condition. In [29] the authors studied the boundary layer effect with geometric corrections, and the asymptotic convergence rate was shown to degrade [26, 27].

However, all the rigorous proofs are done assuming certain smoothness of σ. In particular, it is assumed that σ is sufficiently smooth. At the current stage, very limited work has been done when oscillations present in the media. Denoting δ the small scale in the media, we rewrite our equation as

$$(6) \qquad v \cdot \nabla u + \epsilon\, u = \frac{1}{\epsilon}\sigma^\delta \mathcal{R}u(x,v)\,,$$

where $\sigma^\delta(x) = \sigma(x, \frac{x}{\sigma})$ to explicitly reflect the fast variable $\frac{x}{\sigma}$ dependence. On the theoretical level, to our best knowledge, except a few cases [22], the theory is largely lacking, except a few cases [22], and to a large extent, we do not yet know the resonance of the two parameters and how they contribute in the asymptotic limits of the equation. And on the computation level, the only numerical study known to the authors is presented in [25], where the limits are taken in order: $\delta \ll \epsilon \ll 1$.

The problem is very challenging on the numerical level. The small ϵ makes the collision term extremely stiff, bringing ill-conditioning to the associated discrete system, and thus create a severe stability issue; the small δ brings wild oscillations to the media and the solution, and for accuracy of the numerical solution, high resolution is needed and small discretization is necessary, driving up the numerical cost.

This is certainly unaffordable, especially in the zero limit of ϵ and δ. The main goal of this paper is to develop a general numerical treatment that could deal with the equation with a wide range of ϵ and δ, and perform uniformly well, with the error term independent on the small parameters.

The approach we take is in the line of the generalized multiscale finite element method (GMsFEM). This is an offline-online framework that builds a good set of local basis functions during the offline step and patches local solutions up in the online step, similarly to the original multiscale finite element method. One main feature of GMsFEM is its basis selection procedure in the offline step where a special generalized eigenvalue problem (GEP) is designed. This special GEP encodes the oscillations and the ill-conditioning of the problem.

More explicitly, like many other multiscale methods, we build nested grids with coarse grid H and fine grid h satisfying $H \gg \epsilon \gg h$. In the offline step, local basis functions are constructed within coarse mesh H on fine mesh h that capture fine scale structure and preserve the heterogeneities in the media; in the online step, the basis functions are patched up through Galerkin framework [2, 6, 17, 13, 15, 16, 20, 10, 11, 9, 19, 1, 24]. The online step is rather standard and different methods give various algorithms in the offline step. What makes GMsFEM favorable is indeed its offline step, in which the full list of a-harmonic functions are collected, and then the most "representative" modes are selected through a specially designed GEP. The definition of the matrices in the GEP is associated with the final error term, which permits certain spectral error decay. We should mention GMsFEM was initially used for elliptic equations containing strong heterogeneous media, a topic about which the literature is extremely rich. For this particular problem, there is another category of method: upscaling-type methods. In upscaling methods, either locally or globally an effective media is numerically prepared so that equations can be computed on coarser grids with the effective media replacing the heterogeneous one [12, 30, 18, 21]. But this approach is not going to be pursued in this paper.

As a framework, the GMsFEM approach is rather easy to use, and the main mathematical challenge, when utilized to tackle different equations, is to develop the right GEP. For the linear Boltzmann equation with heterogeneous media, we frame the problem in the discontinuous Galerkin setting and are able to find two matrices that resemble the mass and stiffness matrices in the GEP of the elliptic equations, which allows us to show the optimality of the basis functions with respect to a physically meaningful norm. As the standard approach, these basis functions are then used in the online computation.

The paper will be organized in the following way. In section 2, we introduce some preliminaries. Both discrete ordinates, the standard kinetic solver and GMsFEM for the elliptic equation will be presented. Some properties will be presented. We present the algorithm in section 3, which is further divided into two subsections introducing offline and online procedures. Section 4 contains the analysis where we present the well-posedness and convergence results. The small ϵ limit of the method will also be discussed. Numerical results will be shown in section 5.

To end the introduction, we comment that the scaling problem studied in this paper is not mathematically artificial. In fact, as one redefines $x \to \frac{x}{\epsilon}$, $\sigma(x)$ should

have been automatically changed to highly oscillatory media $\sigma(x/\epsilon)$. Another practical example is to inject light into crystals, where the radiative transfer equation (one particular linear Boltzmann equation) is utilized. In this case, the periodic crystal structure should be encoded in the media and the period that corresponds to δ in our math formulation is expected to be small.

2. Preliminaries. In this section we prepare some basic important concepts. In particular, we will first present the discrete ordinate method for the linear Boltzmann equation, and then give a brief account of the GMsFEM. They are the building blocks for the algorithm designed in this paper.

2.1. Discrete ordinate method. The Boltzmann equation gives a statistical description of particle dynamics. Its extensive use in all kinds of engineering problems brought its great popularity, and literature on both theory and numerics has been very rich. Among all numerical methods developed for the Boltzmann equation, the discrete ordinate method stands out for its simplicity and intuitiveness and is the method we will use in our GMsFEM. Essentially it discretizes the velocity domain, and the semidiscrete system is a coupled PDE in the physical space.

We start the discussion with the following model equation:

$$(7) \qquad \begin{aligned} \mathbf{v}\cdot \nabla u(x,\mathbf{v}) + \epsilon u(x,\mathbf{v}) &= \frac{1}{\epsilon a^\delta}\mathcal{R}u(x,\mathbf{v}) &&\text{in } \Omega \times S^1, \\ u(x,\mathbf{v}) &= g(x,\mathbf{v}) &&\text{on } \Gamma^-, \end{aligned}$$

where $x \in \Omega \subset \mathbb{R}^2$ is a bounded domain with a Lipschitz boundary $\partial\Omega$. The velocity is $v \in S^1$, the unit circle. The media a^δ presents fine scale structure at δ order, and the stiffness of the collision operator $\mathcal{R}$ is determined by $\frac{1}{\epsilon} \gg 1$. We have the inflow boundary condition, with the inflow data $g(x,\mathbf{v})$ defined on Γ^-, a collection of coordinates on the boundary with velocity pointing into the domain:

$$\Gamma^- = \left\{ (x,v) \in \partial\Omega \times S^1 \,|\, \mathbf{v} \cdot \mathbf{n}_x < 0 \right\}.$$

Here $\mathbf{n}_x$ is the unit outer normal direction at $x \in \partial\Omega$. For simplicity, we use the model collision operator with homogeneous scattering coefficient:

$$\mathcal{R}u(x,\mathbf{v}) = \overline{u}(x) - u(x,\mathbf{v}) = \frac{1}{2\pi}\int_{S^1} u(x,\mathbf{v})d\mathbf{v} - u(x,\mathbf{v}).$$

The discrete ordinate method, denoted by S_N, is a standard method to discretize the velocity domain. One first sample m quadrature points on S^1 and each sample point is associated with a weight, denoted by $\{(\mathbf{v}_i,\alpha_i), i=1,\ldots,m\}$, where $\mathbf{v}_i$ are the quadrature points and α_i are the corresponding positive weights. These quadrature points and weights are chosen so that

$$(8) \qquad \sum_{i=1}^m \alpha_i = 1 \quad \text{and} \quad \frac{1}{2\pi}\int_{S^1} u(x,\mathbf{v})d\mathbf{v} \approx \sum_{i=1}^m \alpha_i u(x,\mathbf{v}_i).$$

The equation then will be discretized into a semidiscrete system. Let $u_i(x) = u(x,\mathbf{v}_i)$; the integrodifferential (7) is then transformed into a system of m coupled PDEs:

$$(9) \qquad \mathbf{v}_i \cdot \nabla u_i + \epsilon u_i + \frac{1}{\epsilon a^\delta} \left(u_i - \sum_{j=1}^m \alpha_j u_j \right) = 0 \qquad \text{in } \Omega \,,$$

$$u_i = g_i \qquad \text{on } \Gamma^- \,,$$

where $g_i = g(x, \mathbf{v}_i)$ is the inflow boundary data. Denote

$$(10) \qquad a_{ij} = \begin{cases} \alpha_i - \alpha_i^2, & i = j, \\ -\alpha_i \alpha_j, & i \neq j. \end{cases}$$

Then (9) is further simplified to

$$(11) \qquad \mathbf{v}_i \cdot \nabla u_i + \epsilon u_i + \frac{1}{\epsilon a^\delta \alpha_i} \sum_j a_{ij} u_j = 0 \,.$$

Since $\{a_{ij}\}$ is basically the discrete version of the collision operator $-\mathcal{R}$, it resembles the properties of $\mathcal{R}$. In particular, the matrix is positive semidefinite with a known kernel.

PROPOSITION 1. *Define a matrix* A *so that* $\mathsf{A}_{ij} = a_{ij}$, *we claim*
- A *is positive semidefinite,*
- $\mathsf{u}^\top \mathsf{A} \mathsf{u} = 0$ *if and only if* $\mathsf{u} = (u_1, u_2, \ldots, u_m)$ *is isotropic, i.e.,* $u_1 = u_2 = \cdots = u_m$,
- $\mathsf{v}^\top \mathsf{A} \mathsf{u} = 0$ *if either* u *or* v *is isotropic.*

Proof. The computation is straightforward:

$$\mathsf{u}^\top \mathsf{A} \mathsf{u} = \sum_{i,j=1}^m a_{ij} u_j u_i = \sum_{i=1}^m (\alpha_i - \alpha_i^2) u_i^2 - 2 \sum_{i<j} \alpha_i \alpha_j u_i u_j$$

$$= \sum_{i=1}^m \alpha_i \sum_{j \neq i} \alpha_j u_i^2 - 2 \sum_{i<j} \alpha_i \alpha_j u_i u_j$$

$$= \sum_{i<j} \alpha_i \alpha_j (u_i - u_j)^2 \geq 0 \,.$$

The equal sign is achieved only when $u_i = u_j$ for all $i \neq j$.

To show the third bullet point, we note that the matrix A is symmetric; it suffices to assume that u is isotropic: $u_1 = u_2 = \cdots = u_m = \overline{u}$. Then,

$$\mathsf{v}^\top \mathsf{A} \mathsf{u} = \sum_{i,j=1}^m a_{ij} u_j v_i = \sum_{i,j=1}^m a_{ij} \overline{u} v_i \sum_{i=1}^m \overline{u} v_i \sum_{j=1}^m a_{ij}$$

$$= \sum_{i=1}^m \overline{u} v_i \left(\alpha_i - \alpha_i^2 + \sum_{j \neq i} -\alpha_i \alpha_j \right)$$

$$= \sum_{i=1}^m \overline{u} v_i \alpha_i \left(1 - \alpha_i - \sum_{j \neq i} \alpha_j \right) = 0,$$

where we have used the weight condition (8): $\sum_i \alpha_i = 1$. $\square$

The discrete ordinate method is a classical way of reformulating a Boltzmann-type equation supported on an (x, v)-domain to a coupled system of advection-diffusion equations that only have physical domain x-dependence. This reformulation quickly allows one to apply many different methods developed to treat equations on physical space. There are many choices. In this article, we confine ourselves to the GMsFEM framework.

2.2. Generalized multiscale finite element method. The discrete ordinate method is used to discretize the velocity domain, and for the spatial domain, we follow the GMsFEM approach, which, by choosing "optimal basis functions" via a special design of a GEP, we can obtain a reduced model that is robust for all values of ϵ and δ. For the completeness of the paper, we now present a general idea of GMsFEM, and its application to the heterogeneous Boltzmann equation will be discussed in detail in section 3.

The GMsFEM uses two stages: offline and online. In the offline stage, a small dimensional approximation space is constructed to solve the global problem for any external source on a coarse grid, whose grid size does not need to resolve any scales of the media and solution. The offline stage consists of two main concepts. The snapshot space, V_{snap}^i, is constructed for a generic coarse element K_i. The snapshot solutions are used to compute local multiscale basis functions. An appropriate snapshot space can

- provide a faster convergence,
- provide problem relevant restrictions on the coarse spaces (e.g., divergence-free solutions),
- reduce the cost associated with constructing the offline spaces.

Standard choices of snapshot spaces (see [7]) are (1) all fine grid functions; (2) snapshots of local solutions; (3) oversampling snapshots of local solutions; and (4) force-based snapshots. In this paper, we will use snapshots of local solutions.

More specifically, these are functions $\eta_l^{(i)}$ that satisfy

$$\mathcal{L}\left(\eta_l^{(i)}\right) = 0 \ \text{ in } K_i$$

subject to some boundary conditions, where $\mathcal{L}$ is the differential operator under consideration, and l is the index for the boundary condition. One can use all fine grid delta functions as boundary conditions or randomized boundary conditions [7, 5].

The offline space, V_H, is computed for each K_i (with elements of the space denoted $\psi_l^{(i)}$). We perform a spectral decomposition in the snapshot space and select the dominant eigenfunctions (corresponding to the smallest eigenvalues) to construct the offline (multiscale) space. The convergence rate of the resulting method is proportional to $1/\Lambda_*$, where Λ_* is the smallest eigenvalue that the corresponding eigenvector is not included in the multiscale space. We would like to select a local spectral problem such that we can remove many small eigenvalues with fewer multiscale basis functions. The choice of spectral problems is usually problem dependent and is based on convergence analysis. In general, the error is decomposed into coarse subdomains. If we take the elliptic operator $-\nabla \cdot (\kappa \nabla u)$ with heterogeneous coefficient κ as an example, the energy functional corresponding to the domain Ω is denoted by $a_\Omega(u, u)$, e.g., $a_\Omega(u, u) = \int_\Omega \kappa \nabla u \cdot \nabla u$. Then,

$$(12) \qquad a_\Omega(u - u_H, u - u_H) \preceq \sum_K a_K \left(u^K - u_H^K, u^K - u_H^K\right),$$

where K are coarse regions (K_i), and u^K is the localization of the solution. The local spectral problem is chosen to bound $a_K(u^K - u_H^K, u^K - u_H^K)$. We seek the subspace V_H^i such that for any $\eta \in V_{\text{snap}}^i$, there exists $\eta_0 \in V_H^i$ with

$$(13) \qquad a_{K_i}(\eta - \eta_0, \eta - \eta_0) \preceq \beta s_{K_i}(\eta - \eta_0, \eta - \eta_0),$$

where $s_{K_i}(\cdot, \cdot)$ is an auxiliary bilinear form, and β is an accuracy parameter. The auxiliary bilinear form needs to be chosen such that the solution is bounded in the corresponding norm.

Finally, in the online stage, the space V_H is used together with a suitable coarse grid discretization to solve the problem. The same space V_H is used for all input sources.

3. GMsFEM for heterogeneous Boltzmann equation. We now apply the GMsFEM approach to numerically study the heterogeneous Boltzmann equation, expressed in the discrete ordinate system (9).

The numerical difficulties in solving this equation are summarized as follows. First, the media a^δ is highly oscillatory, which implies that the solution u_i is also highly oscillatory and of multiscale nature. In order to capture these scales, the mesh size h has to be smaller than δ, which in turn brings prohibitive numerical cost. Second, the operator $\mathcal{L}$ is scaled by $\frac{1}{\epsilon}$, and in the zero limit of ϵ, the term is extremely stiff, and this brings concern in stability. It is our aim in this paper to develop a multiscale method that can address these issues. In particular, inspired by GMsFEM, we will design a numerical method that relies on offline basis construction and online basis patching procedure, and its numerical error has limited dependence on the two small parameters.

We will construct nested grids and call $\mathcal{T}^h$ the partition of Ω into fine finite elements and $\mathcal{T}^H$ the partition into coarse elements, where h and H are the fine and coarse mesh sizes, respectively. For simpler notation, we consider rectangular coarse elements as shown in Figure 1. The basis functions and discretization are based on the coarse grid, and the fine grid is used to numerically compute the basis functions. We also denote the collection of coarse edges $\mathcal{E}_H$, and $\mathcal{E}_H^0 = \mathcal{E}_H \backslash \partial\Omega$ the collection of coarse edges in the interior of the domain.

The discontinuous Galerkin method allows one to pick different values of the solution on different sides of the edges. Suppose two adjacent coarse blocks τ_i and τ_j share an edge and that τ_i is the upwind block; then we denote $w^+ = w|_{\tau_i}$ and
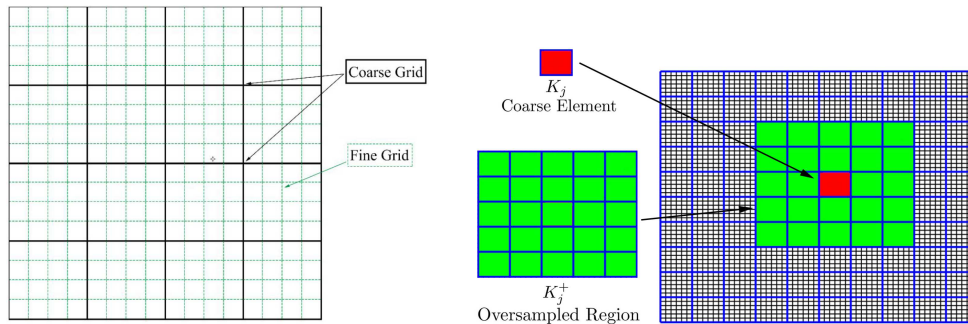

FIG. 1. *Left: an illustration of fine and coarse grids. Right: an illustration of a coarse neighborhood and a coarse element.*
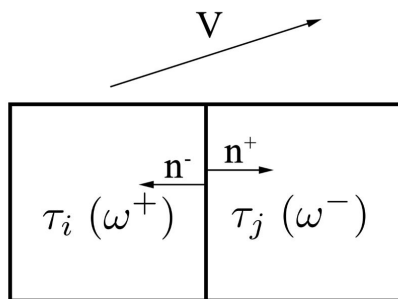
FIG. 2. *An illustration of upwind and downwind blocks.*

$w^- = w|_{\tau_j}$. Notice that depending on the direction of a specific $\mathbf{v}_i$, a different block could be picked as the upwind block, as shown in Figure 2.

For the fine scale approximation, we choose the discrete function space to be

$$V_h = \left\{ v \in L^2(\Omega) \mid v|_\tau \in Q_1(\tau) \; \forall \tau \in \mathcal{T}^h \text{ and } v|_K \in C^0(K) \; \forall K \in \mathcal{T}^H \right\},$$

and we seek numerical solution such that

$$u_h = (u_{h,1}, u_{h,2}, \ldots, u_{h,m}) \in (V_h)^m.$$

This means the numerical solution for each $u_{h,i}$, when confined in each fine grid, is a linear function and is continuous function across coarse grids. In the variational formulation for all $i = 1, 2, \ldots, m$, we have

$$(14) \quad -\int_\Omega u_{h,i} \nabla w_i \cdot \mathbf{v}_i + \sum_{e \in \mathcal{E}_H^0} \int_e u_{h,i}^+ [w_i] \cdot \mathbf{v}_i + \sum_{e \in \Gamma^+} \int_e u_{h,i} w_i \mathbf{v}_i \cdot \mathbf{n} + \int_\Omega \epsilon u_{h,i} w_i$$

$$+ \int_\Omega \frac{1}{\epsilon a^\delta} \left(u_{h,i} - \sum_{j=1}^m \alpha_j u_{h,j} \right) w_i = -\sum_{e \in \Gamma^-} \int_e g_i w_i \mathbf{v}_i \cdot \mathbf{n} \quad \forall w_i \in V_h,$$

or with the definition of $\{a_{ij}\}$ in (10), they could be summed up to

$$(15)$$

$$\sum_{i=1}^m \alpha_i \left(-\int_\Omega u_{h,i} \nabla w_i \cdot \mathbf{v}_i + \sum_{e \in \mathcal{E}_H^0} \int_e u_{h,i}^+ [w_i] \cdot \mathbf{v}_i + \sum_{e \in \Gamma^+} \int_e u_{h,i} w_i \mathbf{v}_i \cdot \mathbf{n} + \int_\Omega \epsilon u_{h,i} w_i \right)$$

$$+ \int_\Omega \frac{1}{\epsilon a^\delta} \sum_{i,j=1}^m a_{ij} u_{h,j} w_i = -\sum_{i=1}^m \alpha_i \sum_{e \in \Gamma^-} \int_e g_i w_i \mathbf{v}_i \cdot \mathbf{n} \quad \forall w \in (V_h)^m.$$

In the equation, we have used upwind approximation for $\mathbf{v}_i \cdot \nabla u_i$ and the jump operator $[\cdot]$ is defined by

$$[w] = \begin{cases} w^- \mathbf{n}^- + w^+ \mathbf{n}^+ & \text{on } \mathcal{E}_H^0, \\ w^- \mathbf{n}^- & \text{on } \Gamma^-, \\ w^+ \mathbf{n}^+ & \text{on } \Gamma^+. \end{cases}$$

For notational simplicity, we define two bilinear operators

$$a(u, w) = \sum_{i=1}^{m} \alpha_i a_i(u_i, w_i) \quad \text{with}$$

$$a_i(u_i, w_i) = -\int_{\Omega} u_i \nabla w_i \cdot \mathbf{v}_i + \sum_{e \in \mathcal{E}_H^0} \int_e u_i^+ [w_i] \cdot \mathbf{v}_i + \sum_{e \in \Gamma^+} \int_e u_i w_i \mathbf{v}_i \cdot \mathbf{n},$$

$$l(u, w) = \sum_{i=1}^{m} \alpha_i \int_{\Omega} \frac{1}{\epsilon a^{\delta}} \left(u_i - \sum_{j=1}^{m} \alpha_j u_j \right) w_i + \sum_{i=1}^{m} \alpha_i \int_{\Omega} \epsilon u_i w_i$$

$$= \int_{\Omega} \frac{1}{\epsilon a^{\delta}} \sum_{i,j=1}^{m} a_{ij} u_j w_i + \sum_{i=1}^{m} \alpha_i \int_{\Omega} \epsilon u_i w_i,$$

and a linear operator

$$F(w) = \sum_{i=1}^{m} \alpha_i F_i(w_i) \quad \text{with} \quad F_i(w_i) = -\sum_{e \in \Gamma^-} \int_e g_i w_i \mathbf{v}_i \cdot \mathbf{n}.$$

With this notation, (15) is now written

(16) $$a(u_h, w) + l(u_h, w) = F(w) \quad \forall w \in (V_h)^m.$$

With $h \ll \min\{\epsilon, \delta\}$, it is a standard result that $u_h \approx u$, with an error term of size $\mathcal{O}(h^2/\min(\epsilon, \delta))$. For significantly small h, the function u_h is considered as a reference solution in accessing the performance of our method.

However, using small h that resolves ϵ and δ leads to a very big system that is numerically very costly. We would like to develop an algorithm that seeks solution only on the coarse grid H and the corresponding solution $u_H \approx u_h \approx u$. To do that, an offline-online procedure developed in [5] for the elliptic equation, that is, GMsFEM, will be pursued. In the offline step, an approximate space V_H is constructed to replace V_h. This newly constructed space would have much fewer degrees of freedom but preserves V_h's important factors. The final multiscale solution will be computed in the online step where the boundary condition $g(x, \mathbf{v})$ will be taken into account to determine the degrees of freedom in V_H.

We quickly review the online stage in section 3.1, and the complicated offline step will be discussed in detail in section 3.2.

3.1. Online computation. In the online stage, we will use the multiscale basis functions together with a coarse grid discretization to solve the given problem. The coarse grid discretization we used is a discontinuous Galerkin method with upwind flux. Assume that a multiscale finite element space $V_H = \text{span}\{\phi_p\}$ is determined, and this space, in some sense, approximates $(V_h)^m$. Then similar to the formulation as in (15), the solution will be sought in

$$u_H = (u_{H,1}, u_{H,2}, \ldots, u_{H,m}) \in V_H$$

so that

(17)
$$\sum_{i=1}^{m} \alpha_i \left(-\int_{\Omega} u_{H,i} \nabla w_i \cdot \mathbf{v}_i + \sum_{e \in \mathcal{E}_H^0} \int_e u_{H,i}^+ [w_i] \cdot \mathbf{v}_i + \sum_{e \in \Gamma^+} \int_e u_{H,i} w_i \mathbf{v}_i \cdot \mathbf{n} + \int_{\Omega} \epsilon u_{H,i} w_i \right)$$
$$+ \int_{\Omega} \frac{1}{\epsilon a^{\delta}} \sum_{i,j=1}^{m} a_{ij} u_{H,j} w_i = -\sum_{i=1}^{m} \alpha_i \sum_{e \in \Gamma^-} \int_e g_i w_i \mathbf{v}_i \cdot \mathbf{n} \quad \forall w \in V_H.$$

Similar to (16), we use a compact notation:

$$a(u_H, w) + l(u_H, w) = F(w) \quad \forall w \in V_H \,. \tag{18}$$

To implement the scheme above, we define the following matrices:

$$A_{pq} = a(\phi_p, \phi_q)\,, \quad L_{pq} = l(\phi_p, \phi_q)\,, \quad \text{and} \quad b_p = F(\phi_p)\,. \tag{19}$$

Then the multiscale solution u_H is formulated as

$$u_H = \sum_p U_p \phi_p \,, \tag{20}$$

where the coefficient vector U solves $(A + L)U = b$.

3.2. Construction of V_H. The key to the success of our method is the construction of the space V_H on the coarse mesh during the offline stage. We will give the details here.

As discussed in section 2.2, the offline step is further decomposed into two substages: constructing the snapshot space, and selecting modes associated with small eigenvalues. These two stages will be presented in sections 3.2.1 and 3.2.2, respectively.

In the snapshot space construction stage, in each coarse region, the Boltzmann solution will be solved multiple times together with all possible boundary conditions resolved by the fine grid. This give a high dimensional space. However, some modes in the snapshot space are more important than the others, and they dominate the numerical solution. To identify these basis functions, a specially designed local spectral problem (GEP) is formulated and solved. The modes that correspond to the smallest eigenvalues are selected to form V_H. The number of modes to be selected depends on the error tolerance and the eigenvalues of the GEP. The design of the local spectral problem is to encode the convergence error that is to be discussed in section 4.

3.2.1. Snapshot space. We present the construction of the snapshot spaces in this subsection. The procedure is the same in each coarse element, and we take the coarse element K_j as an example. The snapshot space for this particular element is denoted by V_{snap}^j. We use the notation $J^i(D)$ to denote the set of all nodes of the fine mesh $\mathcal{T}^h$ lying in the upwind part of ∂D associated with velocity $\mathbf{v}_i$. And we also use $J(D) = \bigoplus_{i=1}^m J^i(D)$ to denote the union. Then the snapshot space is simply the linear span of solutions to the local Boltzmann equation with delta function as boundary condition, namely,

$$V_{\text{snap}}^j = \left\{ {}_n\psi_l^j : n = 1, \ldots, m, \ x_l \in J(K_j) \right\} \,, \tag{21}$$

where ${}_n\psi_l^j = ({}_n\psi_{l,1}^j, {}_n\psi_{l,2}^j, \ldots, {}_n\psi_{l,m}^j)$ solves

$$\tag{22}$$
$$\begin{cases} \mathbf{v}_i \cdot \nabla \,{}_n\psi_{l,i}^j + \epsilon \,{}_n\psi_{l,i}^j + \frac{1}{\epsilon a^\delta} \left({}_n\psi_{l,i}^j - \sum_{q=1}^m \alpha_{qn} \psi_{l,q}^j \right) = 0 & \text{in} \quad K_j \ \forall \ i = 1, 2, \ldots, m\,, \\ {}_n\psi_l^j = \delta_l \mathbf{e}_n & \text{on} \quad J(K_j)\,. \end{cases}$$

Here we use multi-index Kronecker delta function $\delta_l \mathbf{e}_n$, where $\mathbf{e}_n$ is the standard basis in $\mathbb{R}^m$ and δ_l is the standard Kronecker delta function:

$$\delta_l(x_k) = \begin{cases} 1, & k = l, \\ 0, & k \neq l, \end{cases} \quad x_k \in J(K_j)\,.$$

This strategy is summarized in Algorithm DETLOCAL.

The full snapshot space is given by

$$\tag{23} V_{\text{snap}} = \bigoplus_j V_{\text{snap}}^j \,.$$

Remark 1. Numerically to prepare all snapshot basis functions is hard. It requires the computation of the local Boltzmann equation with a large number of possible incoming delta functions. To reduce the cost of computation, we use the idea of oversampling [15]. To do so, the local computational domain is slightly enlarged to K_j^+ (see Figure 1), and a collection of random boundary condition is imposed on K_j^+. The low rank structure of the solution space allows one to correctly capture the range, even with a limited sampling. In particular, we define the snapshot space

$$\tag{24} V_{\text{snap}}^j = \left\{ {}_n\psi_l^{j,+}|_{K_j} : n = 1, \ldots, m, \ l = 1, \ldots, k_j \right\} \,,$$

where k_j is the number of snapshot functions we could customize, and ${}_n\psi_l^{j,+} = ({}_n\psi_{l,1}^{j,+}, {}_n\psi_{l,2}^{j,+}, \ldots, {}_n\psi_{l,m}^{j,+})$ solves

(25)

$$\mathbf{v}_i \cdot \nabla_n \psi_{l,i}^{j,+} + \epsilon_n \psi_{l,i}^{j,+} + \frac{1}{\epsilon a^\delta}\left({}_n\psi_{l,i}^{j,+} - \sum_{q=1}^m \alpha_{qn}\psi_{l,q}^{j,+}\right) = 0 \quad \text{in} \quad K_j^+$$
$$\forall \ i = 1, 2, \ldots, m \,,$$
$$_n\psi_l^j = r_l \mathbf{e}_n \quad \text{on} \quad J(K_j^+)\,,$$

where r_l are random independent and identically distributed (i.i.d.) Gaussian sampling on $J(K_j^+)$. The solutions ${}_n\psi_l^{j,+}$ confined on K_j are then used to form the snapshot spaces. We remark that the use of randomized boundary conditions on oversampling domains is able to reduce the offline computational cost as there is no need to impose delta function boundary conditions as in (22).

This strategy is summarized in Algorithm RANLOCAL.

Similar to (16) and (18), we can solve the snapshot solution $u_{\text{snap}} \in V_{\text{snap}}$ by the following equation:

$$\tag{26} a(u_{\text{snap}}, w) + l(u_{\text{snap}}, w) = F(w) \quad \forall w \in V_{\text{snap}} \,.$$

We note that the snapshot solution can be considered as a reference solution. The error of the snapshot solution is related to the approximation property of the snapshot space in the fine scale space.

3.2.2. Offline space. Now, we will present the construction of the solution space V_H, with the property we mentioned in (13). In the end V_H, when confined on each coarse element, say, K_j, will be a subspace of V_{snap}^j, and the construction of V_H is the choice of the most appropriate basis functions in V_{snap}^j to be included. The procedure is further divided into two substeps: the energy minimizing oversampling, and a design of a GEP, as used in [8].

We first denote the local oversampled snapshot space $\bigoplus_{K_i \subset K_j^+} V_{\text{snap}}^i$ by $V_{\text{snap}}^{j,+}$. Notice that, for a given coarse element K_j and its corresponding oversampling region K_j^+, the space $V_{\text{snap}}^{j,+}$ is the union of all snapshot spaces V_{snap}^i with the condition that $K_i \subset K_j^+$.

Then the energy minimizing snapshots are calculated. For any snapshot function $\psi \in V_{\text{snap}}^j$, its energy minimizing extension $\widetilde{\psi}$ has the smallest energy in some norm $a_{\text{Energy}}^j(\cdot, \cdot)$ and is sought in the local oversampled snapshot space $V_{\text{snap}}^{j,+}$ with the constraint $\widetilde{\psi}|_{K_j} = \psi|_{K_j}$. In mathematical expression, for any $\psi \in V_{\text{snap}}^j$, we seek $\widetilde{\psi} \in V_{\text{snap}}^{j,+}$ so that

$$
(27) \qquad \begin{aligned}
\widetilde{\psi} &= \operatorname{argmin}_{\widetilde{\phi} \in V_{\text{snap}}^{j,+}} a_{\text{Energy}}^j(\widetilde{\phi}, \widetilde{\phi}) \\
\text{s.t.} \quad \widetilde{\psi} &= \psi \quad \text{in } K_j
\end{aligned}
$$

in which

$$
(28)
$$

$$
a_{\text{Energy}}^j(\widetilde{\phi}, \widetilde{\phi}) = \sum_{i=1}^m \alpha_i \left(\int_{K_j^+} |\nabla \widetilde{\phi}_i|^2 + \frac{1}{H} \sum_{e \in \mathcal{E}_H^0(K_j^+)} \int_e [\widetilde{\phi}_i]^2 \right) + \int_{K_j^+} \frac{1}{\epsilon a^\delta} \sum_{i,l=1}^m a_{il} \widetilde{\phi}_l \widetilde{\phi}_i \, .
$$

We notice that this construction is well-defined and the strategy is summarized in Algorithm ENERGYMIN. As one can see, $\widetilde{\psi}$ is an extension of ψ onto the oversampling domain that achieves the minimum energy, defined in (28), and this extension is crucial, as will be seen in the later analysis.

Remark 2. This is about a stable decomposition property. It is important that the local basis functions satisfy a stable decomposition property. More precisely, the sum of local energies is bounded by the global energy.

Next, we define the two bilinear operators $a_{K_j}(\cdot, \cdot)$ and $s_{K_j}(\cdot, \cdot)$, mentioned in (13). For simplicity of notation, we use $a^j(\cdot, \cdot)$ and $s^j(\cdot, \cdot)$ instead. For the element K_j, define

$$
(29)
$$

$$
a^j(\phi, \eta) = \sum_{i=1}^m \alpha_i \left(\int_{K_j^+} \nabla \widetilde{\phi}_i \cdot \nabla \widetilde{\eta}_i + \frac{1}{H} \sum_{e \in \mathcal{E}_H^0(K_j^+)} \int_e [\widetilde{\phi}_i][\widetilde{\eta}_i] \right) + \int_{K_j^+} \frac{1}{\epsilon a^\delta} \sum_{i,l=1}^m a_{il} \widetilde{\phi}_l \widetilde{\eta}_i \, ,
$$

$$
s^j(\phi, \eta) = \sum_{i=1}^m \alpha_i \left(\frac{1}{2} \sum_{K \subset K_j^+} \int_{\partial K} |\mathbf{v}_i \cdot \mathbf{n}| \, \widetilde{\phi}_i \widetilde{\eta}_i + \int_{K_j^+} \epsilon \widetilde{\phi}_i \widetilde{\eta}_i \right) + \int_{K_j^+} \frac{1}{\epsilon a^\delta} \sum_{i,l=1}^m a_{il} \widetilde{\phi}_l \widetilde{\eta}_i \, .
$$

Using the above bilinear forms, a spectral problem is defined. On K_j, we look for $(\phi_k^j, \lambda_k^j) \in V_{\text{snap}}^j \times \mathbb{R}$ such that

$$
a^j \left(\phi_k^j, \eta \right) = \lambda s^j \left(\phi_k^j, \eta \right) \quad \forall \eta \in V_{\text{snap}}^j,
$$

where the eigenvalues are ordered in the ascending way:

$$
\lambda_{j,1} \le \lambda_{j,2} \le \cdots .
$$

For implementation, we define the following matrices:

$$
(30) \qquad A_{pq}^j = a^j \left(\psi_p^j, \psi_q^j \right) \quad \text{and} \quad S_{pq}^j = s^j \left(\psi_p^j, \psi_q^j \right) \, .
$$

Then the pair (ϕ_k^j, λ_k^j) is computed by solving

$$(31) \qquad A^j \mathbf{c}_k = \lambda_k S^j \mathbf{c}_k \quad \text{with} \quad \phi_k^j = \sum_p c_{k,p} \psi_p^j .$$

Suppose L_j modes are used for each K_j. This strategy is summarized in Algorithm LOCALGEP. The offline space V_H is given by

$$(32) \qquad V_H^j = \text{span} \left\{ \phi_k^j : k = 1 \cdots L_j \right\} \quad \text{and} \quad V_H = \bigoplus_j V_H^j .$$

This will be the approximation space for solving the system (9) in the scheme (18).

3.3. Algorithm summary. We finally summarize the algorithm. Largely speaking, we prepare the basis functions in the offline step and patch them up in the online step. The offline step is further divided into preparing a snapshot space in which either all Green's functions are accumulated or a good random selection is obtained, and the basis selection step, in which the local GEP is computed and eigenfunctions with highest energies are chosen. These basis functions are ultimately used in the online step via the weak formulation (20). We summarize the procedure in Algorithm 1.

4. Analysis of the GMsFEM. In this section, we will present some analysis of our GMsFEM. In section 4.1, we will prove the well-posedness of the discrete system resulting from the GMsFEM, and in section 4.2, we will prove the convergence of the method. Finally, in section 4.3, we will analyze the behavior of the method when ϵ is small.

4.1. Well-posedness. We first show the well-posedness of the GMsFEM (18).

THEOREM 1. *Problem* (18) *has a unique solution, and the solution u_H satisfies the following stability condition:*

$$\sum_{i=1}^m \alpha_i \left(\frac{1}{4} \sum_{e \in \mathcal{E}_H} \int_e |\mathbf{v}_i \cdot \mathbf{n}| \, [u_{H,i}]^2 + \int_\Omega \epsilon u_{H,i}^2 \right) + \int_\Omega \frac{1}{\epsilon a^\delta} \sum_{i,j=1}^m a_{ij} u_{H,j} u_{H,i}$$

$$(33) \qquad \leq \sum_{i=1}^m \alpha_i \sum_{e \in \Gamma^-} \int_e |\mathbf{v}_i \cdot \mathbf{n}| \, g_i^2 .$$

Proof. Since the system (18) is a square linear system, showing the existence and uniqueness amounts to proving that $a(\widehat{u}, w) + l(\widehat{u}, w) = 0$ for all $w \in V_H$ only for trivial solution $\widehat{u} = 0$.

We will first prove the following inequalities:

$$l(u, u) \geq 0 \quad \text{and} \quad a(u, u) = \sum_i a(u_i, u_i) \geq 0 \quad \forall u \in V_H .$$

First, l is nonnegative since the matrix (a_{ij}) is a positive semidefinite matrix, as discussed in Proposition 1. Next, the nonnegativity of $a(\cdot, \cdot)$ is shown below:

$$(34)$$

$$a(u_i, u_i) = -\int_\Omega u_i \nabla u_i \cdot \mathbf{v}_i + \sum_{e \in \mathcal{E}_H^0} \int_e u_i^+ [u_i] \cdot \mathbf{v}_i + \sum_{e \in \Gamma^+} \int_e u_i^2 \mathbf{v}_i \cdot \mathbf{n}$$

$$= -\frac{1}{2} \sum_{\tau \in \mathcal{T}^h} \int_{\partial \tau} u_i^2 \mathbf{v}_i \cdot \mathbf{n} + \sum_{e \in \mathcal{E}_H^0} \int_e u_i^+ [u_i] \cdot \mathbf{v}_i + \sum_{e \in \Gamma^+} \int_e u_i^2 \mathbf{v}_i \cdot \mathbf{n}$$

$$= -\frac{1}{2} \sum_{e \in \Gamma^+} \int_e |\mathbf{v}_i \cdot \mathbf{n}| \, u_i^2 + \frac{1}{2} \sum_{e \in \Gamma^-} \int_e |\mathbf{v}_i \cdot \mathbf{n}| \, u_i^2 + \frac{1}{2} \sum_{e \in \mathcal{E}_H^0} \int_e |\mathbf{v}_i \cdot \mathbf{n}| \left(u_i^{-2} - u_i^{+2} \right)$$

Algorithm 1. Multiscale solver for $\mathcal{L}u = 0$ over Ω with $u = g$ on Γ^-

---

1: **Domain decomposition**
2:     Partition domain into nonoverlapping patches $\Omega = \bigcup_j K_j$.
3: **Offline stage:**
4:     **Snapshot space**
5:         Form snapshot space by calling $V_{\text{snap}}^j = \text{DETLOCAL}(K_j)$ or $V_{\text{snap}}^j = \text{RANLOCAL}(K_j^+)$.
6:     **Offline space**
7:         Form offline space by calling $V_H^j = \text{LOCALGEP}(K_j)$.
8:         $V_H = \bigoplus_j V_H^j = \text{span}\{\phi_p\}$.
9: **Online stage:**
10:     Use global inflow boundary data g to determine U using (20).
11: **Return:** approximated global solution $u_H = \sum_p U_p \phi_p$.

1: **function** $\text{DETLOCAL}(K_j)$
2:     Prepare full list of multi-index Kronecker delta function $\delta_l \mathbf{e}_n$ on $J(K_j)$.
3:     Find $_n\psi_l^j$ using (22).
4:     Formulate $V_{\text{snap}}^j = \{_n\psi_l^j : n = 1, \ldots, m, \ x_l \in J(K_j)\}$ according to (21).
5:     **Return:** Local snapshot space V_{snap}^j.
6: **end function**

1: **function** $\text{RANLOCAL}(K_j)$
2:     Prepare k_j random i.i.d. Gaussian vector $r_l \mathbf{e}_n$ on $J(K_j^+)$.
3:     Find $_n\psi_l^{j,+}$ using (25).
4:     Formulate $V_{\text{snap}}^j = \{_n\psi_l^{j,+}|_{K_j} : n = 1, \ldots, m, \ l = 1, \ldots, k_j\}$ according to (21).
5:     **Return:** Local snapshot space V_{snap}^j.
6: **end function**

1: **function** $\text{LOCALGEP}(K_j)$
2:     $V_{\text{snap}}^{j,+} = \bigoplus_{K_i \subset K_j^+} V_{\text{snap}}^i$.
3:     Compute $\widetilde{\psi} \in V_{\text{snap}}^{j,+}$ using (27).
4:     Solve the GEP (31) for $V_H^j = \text{span}\{\phi_k^j : k = 1 \cdots L_j\}$.
5:     **Return:** Offline space V_H^j.
6: **end function**

$$+ \sum_{e \in \mathcal{E}_H^0} \int_e |\mathbf{v}_i \cdot \mathbf{n}| \, u_i^+ \left(u_i^+ - u_i^-\right) + \sum_{e \in \Gamma^+} \int_e |\mathbf{v}_i \cdot \mathbf{n}| \, u_i^2$$

$$= \frac{1}{2} \sum_{e \in \Gamma^+} \int_e |\mathbf{v}_i \cdot \mathbf{n}| \, u_i^2 + \frac{1}{2} \sum_{e \in \Gamma^-} \int_e |\mathbf{v}_i \cdot \mathbf{n}| \, u_i^2 + \frac{1}{2} \sum_{e \in \mathcal{E}_H^0} \int_e |\mathbf{v}_i \cdot \mathbf{n}| \left(u_i^+ - u_i^-\right)^2$$

$$= \frac{1}{2} \sum_{e \in \mathcal{E}_H} \int_e |\mathbf{v}_i \cdot \mathbf{n}| \, [u_i]^2 \geq 0.$$

Assuming $a(\widehat{u}, w) + l(\widehat{u}, w) = 0$ for any $w \in (V_h)^m$, then setting $w = \widehat{u}$, we have

$$\text{(35)} \qquad \sum_{i=1}^{m} \alpha_i \left(\frac{1}{2} \sum_{e \in \mathcal{E}_H} \int_e |\mathbf{v}_i \cdot \mathbf{n}| \, [\widehat{u}_i]^2 + \int_\Omega \epsilon \widehat{u}_i^2 \right) + \int_\Omega \frac{1}{\epsilon a^\delta} \sum_{i,j=1}^{m} a_{ij} \widehat{u}_j \widehat{u}_i = 0.$$

According to Proposition 1, we have

$$\text{(36)} \qquad \widehat{u_1} = \widehat{u_2} = \cdots = \widehat{u_m} = 0 \,,$$

meaning $\widehat{u} = 0$, and the solution to (18) is thus unique. For stability, we start with

$$F(u_H) = -\sum_{i=1}^{m} \alpha_i \sum_{e \in \Gamma^-} \int_e g_i u_{H,i} \mathbf{v}_i \cdot \mathbf{n}$$
$$\leq \sum_{i=1}^{m} \alpha_i \sum_{e \in \Gamma^-} \int_e |\mathbf{v}_i \cdot \mathbf{n}| \, g_i^2 + \frac{1}{4} \sum_{i=1}^{m} \alpha_i \sum_{e \in \mathcal{E}_H} \int_e |\mathbf{v}_i \cdot \mathbf{n}| \, [u_{H,i}]^2 \,.$$

Considering $a(u_H, u_H) + l(u_H, u_H) = F(u_H)$, we conclude with the stability inequality (33). □

We notice that the snapshot equation (26) has the same structure, and the well-posedness is proved in the same way.

4.2. Convergence analysis. We now analyze the convergence of the proposed method. The goal of this section is to estimate the difference between the snapshot solution, u_{snap}, computed in (26), and the multiscale coarse solution, u_H, computed in (18). To do so, we first define the following norms. We define the V-norm as

$$\text{(37)} \qquad \|u\|_V^2 = \sum_{i=1}^{m} \alpha_i \, \|u_i\|_{V^i}^2 \quad \text{with} \quad \|u_i\|_{V^i}^2 = \frac{1}{2} \sum_{e \in \mathcal{E}_H} \int_e |\mathbf{v}_i \cdot \mathbf{n}| \, [u_i]^2$$

and the W-norm as

$$\text{(38)} \qquad \|u\|_W^2 = \sum_{i=1}^{m} \alpha_i \, \|u_i\|_{W^i}^2 \quad \text{with} \quad \|u_i\|_{W^i}^2 = \frac{1}{2} \sum_{K_j} \int_{\partial K_j} |\mathbf{v}_i \cdot \mathbf{n}| \, u_i^2 \,.$$

We also extend them by incorporating the collision term:

$$\text{(39)} \qquad \|u\|_{\widetilde{V}}^2 = \|u\|_V^2 + l(u, u) \quad \text{and} \quad \|u\|_{\widetilde{W}}^2 = \|u\|_W^2 + l(u, u) \,.$$

The total energy is now defined by

$$\|u\|_{\text{Energy}}^2 = \sum_{i=1}^{m} \alpha_i \left(\int_\Omega |\nabla u_i|^2 + \frac{1}{H} \sum_{e \in \mathcal{E}_H^0} \int_e [u_i]^2 \right) + \int_\Omega \frac{1}{\epsilon a^\delta} \sum_{i,j=1}^{m} a_{ij} u_j u_i \,.$$

Note that we have following propositions.

PROPOSITION 2. $a(u, u) = \|u\|_V^2$, and $a(u, u) + l(u, u) = \|u\|_{\widetilde{V}}^2$.

Proof. This proposition simply comes from the calculations in (34). □

PROPOSITION 3. *If $u \in V_{snap}$, we have*

(40) $$1. \quad \|u\|_{\widetilde{W}}^2 \leq \sum_j s^j(u|_{K_j}, u|_{K_j}),$$

(41) $$2. \quad \sum_j a^j(u|_{K_j}, u|_{K_j}) \leq M\|u\|_{Energy}^2.$$

Here a^j and s^j are bilinear operator defined in (29). $M = \max_{K,E}\{M_K, M_E\}$, where M_K is the number of oversampled regions K_j^+'s which have nonempty intersection with coarse block K, and M_E is the number of oversampled regions K_j^+'s whose interior coarse edges $\mathcal{E}_H^0(K_j^+)$ contain coarse edge E. They are both small numbers.

Proof. We denote $u|_{K_j}$ by u^j. So $u^j \in V_{\text{snap}}^j$. According to (27), u^j has an energy minimizing extension $\widetilde{u}^j \in V_{\text{snap}}^{j,+}$ that satisfies $\widetilde{u}^j = u^j$ in K_j. Then we have

$$\sum_{i=1}^m \alpha_i \left(\frac{1}{2} \sum_{K_j} \int_{\partial K_j} |\mathbf{v}_i \cdot \mathbf{n}| \, (u_i^j)^2 + \int_{K_j} \epsilon(u_i^j)^2 \right) + \int_{K_j} \frac{1}{\epsilon a^\delta} \sum_{i,l=1}^m a_{il} u_l^j u_i^j$$

$$= \sum_{i=1}^m \alpha_i \left(\frac{1}{2} \sum_{K_j} \int_{\partial K_j} |\mathbf{v}_i \cdot \mathbf{n}| \, (\widetilde{u}_i^j)^2 + \int_{K_j} \epsilon(\widetilde{u}_i^j)^2 \right) + \int_{K_j} \frac{1}{\epsilon a^\delta} \sum_{i,l=1}^m a_{il} \widetilde{u}_l^j \widetilde{u}_i^j$$

$$\leq \sum_{i=1}^m \alpha_i \left(\frac{1}{2} \sum_{K \subset K_j^+} \int_{\partial K} |\mathbf{v}_i \cdot \mathbf{n}| \, (\widetilde{u}_i^j)^2 + \int_{K_j^+} \epsilon(\widetilde{u}_i^j)^2 \right) + \int_{K_j^+} \frac{1}{\epsilon a^\delta} \sum_{i,l=1}^m a_{il} \widetilde{u}_l^j \widetilde{u}_i^j$$

$$= s^j\left(u^j, u^j\right).$$

Combining with the definition of $\|\cdot\|_{\widetilde{W}}^2$, we proved (40).

Next, we denote $u^{j,+} = u|_{K_j^+} \in V_{\text{snap}}^{j,+}$. By the definition of the energy minimizing extension in (27), we have

$$a^j(u^j, u^j)$$

$$= \sum_{i=1}^m \alpha_i \left(\int_{K_j^+} \left| \nabla \widetilde{u}_i^j \right|^2 + \frac{1}{H} \sum_{e \in \mathcal{E}_H^0(K_j^+)} \int_e \left[\widetilde{u}_i^j \right]^2 \right) + \int_{K_j^+} \frac{1}{\epsilon a^\delta} \sum_{i,l=1}^m a_{il} \widetilde{u}_l^j \widetilde{u}_i^j$$

$$\leq \sum_{i=1}^m \alpha_i \left(\int_{K_j^+} \left| \nabla u_i^{j,+} \right|^2 + \frac{1}{H} \sum_{e \in \mathcal{E}_H^0(K_j^+)} \int_e [u_i^{j,+}]^2 \right) + \int_{K_j^+} \frac{1}{\epsilon a^\delta} \sum_{i,l=1}^m a_{il} u_l^{j,+} u_i^{j,+}.$$

Hence,

$$\sum_j a^j(u^j, u^j) \leq \sum_j \sum_{i=1}^m \alpha_i \left(\int_{K_j^+} \left| \nabla u_i^{j,+} \right|^2 + \frac{1}{H} \sum_{e \in \mathcal{E}_H^0(K_j^+)} \int_e [u_i^{j,+}]^2 \right)$$

$$+ \sum_j \int_{K_j^+} \frac{1}{\epsilon a^\delta} \sum_{i,l=1}^m a_{il} u_l^{j,+} u_i^{j,+}$$

$$\leq \sum_{i=1}^m \alpha_i \left(\sum_j M_{K_j} \int_{K_j} \left| \nabla u_i^j \right|^2 + \frac{1}{H} \sum_{e \in \mathcal{E}_H^0} M_e \int_e [u_i^j]^2 \right)$$

$$+ \sum_j M_{K_j} \int_{K_j} \frac{1}{\epsilon a^\delta} \sum_{i,l=1}^m a_{il} u_l^j u_i^j$$

$$\leq M \|u\|_{\text{Energy}}^2,$$

and thus we have (41). □

For the convergence analysis, we first examine the best approximation property. For that, we have the following.

LEMMA 1. *Let u_{snap} be the snapshot solution to* (26) *and let u_H be the multiscale solution to* (18). *Then*

$$\text{(42)} \qquad \|u_{snap} - u_H\|_{\widetilde{V}}^2 \leq C \inf_{w \in V_H} \|u_{snap} - w\|_{\widetilde{W}}^2,$$

where C is a constant independent of ϵ, a^δ, and the mesh size.

Proof. Using (26) and (18), and the fact that $V_H \subset V_{\text{snap}}$, we have

$$\text{(43)} \qquad a(u_{\text{snap}} - u_H, w) + l(u_{\text{snap}} - u_H, w) = F(w) - F(w) = 0 \quad \forall w \in V_{\text{snap}}.$$

Then for all $w \in V_H$,

$$a(u_{\text{snap}} - u_H, u_{\text{snap}} - u_H) + l(u_{\text{snap}} - u_H, u_{\text{snap}} - u_H)$$
$$= a(u_{\text{snap}} - u_H, u_{\text{snap}} - w) + l(u_{\text{snap}} - u_H, u_{\text{snap}} - w).$$

Using Proposition 2, we have

$$\text{(44)} \qquad \|u_{\text{snap}} - u_H\|_{\widetilde{V}}^2 = a(u_{\text{snap}} - u_H, u_{\text{snap}} - w) + l(u_{\text{snap}} - u_H, u_{\text{snap}} - w).$$

To obtain (42), noticing that $u_{\text{snap}} - u_H$ and $u_{\text{snap}} - w$ are both in V_{snap}, it amounts to showing that

$$\text{(45)} \qquad a(u, w) + l(u, w) \leq C \|u\|_{\widetilde{V}} \|w\|_{\widetilde{W}} \quad \forall u, w \in V_{\text{snap}}.$$

In fact it suffices to show that

$$\text{(46)} \qquad a(u, w) + l(u, w) \leq \sqrt{2} \|u\|_V \|w\|_{\widetilde{W}} \quad \forall u, w \in V_{\text{snap}},$$

since it is obvious that $\|u\|_V \leq \|u\|_{\widetilde{V}}$.

To show (46), we first use integration by parts to obtain

$$\sum_{i=1}^m \alpha_i \left(-\int_\Omega u_i \nabla w_i \cdot \mathbf{v}_i \right) = \sum_{i=1}^m \alpha_i \left(\int_\Omega w_i \nabla u_i \cdot \mathbf{v}_i - \sum_{\tau \in \mathcal{T}^h} \int_{\partial \tau} \mathbf{v}_i \cdot \mathbf{n} u_i w_i \right)$$

$$= \sum_{i=1}^m \alpha_i \left(\sum_{K_j} \int_{K_j} \nabla u_i \cdot \mathbf{v}_i w_i - \sum_{K_j} \int_{\partial K_j} \mathbf{v}_i \cdot \mathbf{n} u_i w_i \right)$$

$$= -l(u, w) - \sum_{i=1}^m \alpha_i \sum_{K_j} \int_{\partial K_j} \mathbf{v}_i \cdot \mathbf{n} u_i w_i,$$

where we have used the continuity accross fine scales $\partial \tau$ and the assumption that, in each K_j, u satisfies the following equation:

(47)

$$-\int_{K_j} u_i \nabla w_i \cdot \mathbf{v}_i + \int_{\partial K_j} u_i w_i \mathbf{v}_i \cdot \mathbf{n} + \int_{K_j} \epsilon u_i w_i + \int_{K_j} \frac{1}{\epsilon a^\delta} \left(u_i - \sum_{q=1}^m \alpha_q u_q \right) w_i = 0$$

for all $i = 1, 2, \ldots, m$. Here w could be any function in $(V_h)^m$ restricted on K_j. In particular, (47) works for $w \in V_{\text{snap}}^j$.

Using the definition of $a(\cdot, \cdot)$ and direct calculations, we have

$$a(u, w) + l(u, w) = \sum_{i=1}^m \alpha_i \left(-\sum_{K_j} \int_{\partial K_j} \mathbf{v}_i \cdot \mathbf{n} u_i w_i + \sum_{e \in \mathcal{E}_H^0} \int_e u_i^+ [w_i] \cdot \mathbf{v}_i \right.$$

$$\left. + \sum_{e \in \Gamma^+} \int_e u_i w_i \mathbf{v}_i \cdot \mathbf{n} \right)$$

$$+ \sum_{i=1}^m \alpha_i \sum_{K_j} \left(\int_{K_j} \epsilon u_i w_i + \frac{1}{\epsilon a^\delta} \left(u_i - \sum_{q=1}^m \alpha_q u_q \right) w_i \right)$$

$$= \sum_{i=1}^m \alpha_i \left(-\sum_{e \in \mathcal{E}_H^0} \int_e w_i^- [u_i] \cdot \mathbf{v}_i - \sum_{e \in \Gamma^-} \int_e u_i w_i \mathbf{v}_i \cdot \mathbf{n} \right).$$

Then, applying the Cauchy–Schwarz inequality, we have

$$a(u, w) + l(u, w) \le \left(\sum_{i=1}^m \alpha_i \left(\sum_{e \in \mathcal{E}_H^0} \int_e |\mathbf{v}_i \cdot \mathbf{n}| [u_i]^2 + \sum_{e \in \Gamma^-} \int_e |\mathbf{v}_i \cdot \mathbf{n}| u_i^2 \right) \right)^{1/2}$$

(48)

$$+ \left(\sum_{i=1}^m \alpha_i \left(\sum_{e \in \mathcal{E}_H^0} \int_e |\mathbf{v}_i \cdot \mathbf{n}| w_i^{-2} + \sum_{e \in \Gamma^-} \int_e |\mathbf{v}_i \cdot \mathbf{n}| w_i^2 \right) \right)^{1/2}.$$

The two terms on the right-hand side are taken care of separately. To handle the first term, recalling the definition of V-norm in (37), we have

$$\sum_{i=1}^m \alpha_i \left(\sum_{e \in \mathcal{E}_H^0} \int_e |\mathbf{v}_i \cdot \mathbf{n}| [u_i]^2 + \sum_{e \in \Gamma^-} \int_e |\mathbf{v}_i \cdot \mathbf{n}| u_i^2 \right) \le \sum_{i=1}^m \alpha_i \sum_{e \in \mathcal{E}_H} \int_e |\mathbf{v}_i \cdot \mathbf{n}| [u_i]^2$$

(49)

$$= \sqrt{2} \|u\|_V^2.$$

And to compute the second term, we notice that

$$0 = \sum_{i=1}^m \alpha_i \sum_{K_j} \left(-\int_{K_j} w_i \nabla w_i \cdot \mathbf{v}_i + \int_{\partial K_j} w_i^2 \mathbf{v}_i \cdot \mathbf{n} \right) + l(w, w)$$

$$= \sum_{i=1}^m \alpha_i \sum_{K_j} \left(-\frac{1}{2} \int_{\partial K_j} w_i^2 \mathbf{v}_i \cdot \mathbf{n} + \int_{\partial K_j} w_i^2 \mathbf{v}_i \cdot \mathbf{n} \right) + l(w, w)$$

$$= \frac{1}{2} \sum_{i=1}^m \alpha_i \sum_{K_j} \int_{\partial K_j} \mathbf{v}_i \cdot \mathbf{n} w_i^2 + l(w, w)$$

$$= \frac{1}{2}\sum_{i=1}^{m}\alpha_i\left(-\sum_{e\in\Gamma^-}\int_e|\mathbf{v}_i\cdot\mathbf{n}|\,w_i^2 + \sum_{e\in\Gamma^+}\int_e|\mathbf{v}_i\cdot\mathbf{n}|\,w_i^2\right.$$
$$\left. + \sum_{e\in\mathcal{E}_H^0}\int_e|\mathbf{v}_i\cdot\mathbf{n}|\left(w_i^{+\,2}-w_i^{-\,2}\right)\right) + l(w,w)\,,$$

which in turn gives

$$(50)\qquad \sum_{i=1}^{m}\alpha_i\left(\sum_{e\in\mathcal{E}_H^0}\int_e|\mathbf{v}_i\cdot\mathbf{n}|\,w_i^{-\,2} + \sum_{e\in\Gamma^-}\int_e|\mathbf{v}_i\cdot\mathbf{n}|\,w_i^2\right)$$
$$= \frac{1}{2}\sum_{i=1}^{m}\alpha_i\left(\sum_{e\in\Gamma^-}\int_e|\mathbf{v}_i\cdot\mathbf{n}|\,w_i^2 + \sum_{e\in\Gamma^+}\int_e|\mathbf{v}_i\cdot\mathbf{n}|\,w_i^2\right.$$
$$\left. + \sum_{e\in\mathcal{E}_H^0}\int_e|\mathbf{v}_i\cdot\mathbf{n}|\left(w_i^{+\,2}+w_i^{-\,2}\right)\right) + l(w,w)$$
$$= \frac{1}{2}\sum_{i=1}^{m}\alpha_i\sum_{K_j}\int_{\partial K_j}|\mathbf{v}_i\cdot\mathbf{n}|\,w_i^2 + l(w,w)$$
$$= \|w\|_{\widetilde{W}}^2\,.$$

Plugging (49) and (50) into (48), we have proved the desired boundedness condition (46), which concludes the proof of (42). □

Now, we are ready to prove our main convergence result in this section.

THEOREM 2. *Let u_{snap} be the snapshot solution to problem* (26) *and let u_H be the multiscale solution to problem* (18). *Then*

$$\|u_{snap}-u_H\|_{\widetilde{V}}^2 \le \frac{CM}{\Lambda_*}\|u_{snap}\|_{Energy}^2,$$

where $\Lambda_* = \min_j \lambda_{j,L_j+1}$, C is the same constant from Lemma 1, *and M is the same constant from Proposition* 3.

Proof. We first denote

$$u_{\text{snap}} = \sum_j u_{\text{snap}}|_{K_j} = \sum_j u_{\text{snap}}^j = \sum_{j,l} d_{j,l}\phi_l^j\,,$$

where ϕ_l^j is the lth multiscale basis function for the coarse element K_j (31). Note that $\text{span}\{\phi_l^j\}$ covers the entire snapshot space. We then define a projection of u_{snap}^j into V_H^j, as well as a projection of u_{snap} into V_H:

$$P^j(u_{\text{snap}}^j) = \sum_{l\le L_j} d_{j,l}\phi_l^j\,,\quad P(u_{\text{snap}}) = \sum_j\sum_{l\le L_j} d_{j,l}\phi_l^j.$$

It is easy to see that $P^j(u_{\text{snap}}^j) = P(u_{\text{snap}})|_{K_j}$. Combining with Proposition 3, we have

$$\inf_{w\in V_H}\|u_{\text{snap}}-w\|_{\widetilde{W}}^2 \le \|u_{\text{snap}}-P(u_{\text{snap}})\|_{\widetilde{W}}^2$$
$$\le \sum_j s^j\left(u_{\text{snap}}^j - P^j\left(u_{\text{snap}}^j\right), u_{\text{snap}}^j - P^j\left(u_{\text{snap}}^j\right)\right)$$

$$\leq \sum_j \frac{1}{\lambda_{L_j+1}^{(j)}} a^j \left(u_{\mathrm{snap},j}, u_{\mathrm{snap},j}\right)$$

$$\leq \frac{1}{\Lambda_*} \sum_j a^j \left(u_{\mathrm{snap}}^j, u_{\mathrm{snap}}^j\right)$$

$$\leq \frac{M}{\Lambda_*} \|u_{\mathrm{snap}}\|_{\mathrm{Energy}}^2.$$

Combining with Lemma 1, we proved the theorem.   □

In the above theorem, we estimate the error between the snapshot solution u_{snap} and the multiscale solution u_H. We see that the error is inversely proportional to the eigenvalues. This shows that the multiscale space gives a good approximation property in the snapshot space. In our analysis, we assume that the snapshot functions satisfy the PDE in the strong sense, that is, (47). On the other hand, there is an error between the snapshot solution u_{snap} and the fine scale solution u_h if we use Algorithm RanLocal in section 3.2. This amounts to an irreducible error, and the analysis of this is beyond the scope of this paper.

We should emphasize that the difficulty brought by small δ is encoded in the quality of Λ_* and thus is not explicitly expressed in the error analysis.

4.3. Small ϵ regime. An important property the algorithm satisfies is that it is robust with respect to the parameters. In the limiting regime of $\epsilon \to 0$, Λ_* has a positive lower bound, and this serves as the stability argument that allows the algorithm to be effective across regimes. In particular we will show the following.

THEOREM 3. *Denote $\lambda_{j,k}$ the kth eigenvalue of the GEP defined in (31) for coarse element K_j. It has an asymptotic limit in the zero limit of ϵ, meaning there is a constant $\lambda_{j,k}^0$ so that*

$$\left|\lambda_{j,k} - \lambda_{j,k}^0\right| = \mathcal{O}(\epsilon).$$

This theorem, when combined with our main Theorem 2, indicates that the error bound, which is controlled by $\Lambda_* = \frac{1}{\min_j\{\lambda_{j,L_j+1}\}}$, will not grow in ϵ and thus the error is uniformly bounded.

To show the theorem, we first start with a lemma.

LEMMA 2. *For every coarse element K_j, we have*

$$A^j = A^{j,0} + \mathcal{O}(\epsilon), \quad and \quad S^j = S^{j,0} + \mathcal{O}(\epsilon),$$

where entries in $A^{j,0}$ and $S^{j,0}$ are defined by

$$(51) \qquad A_{pq}^{j,0} = \sum_{i=1}^m \alpha_i \left(\int_{K_j^+} \nabla \widetilde{\psi}_{p,i}^{j,0} \cdot \nabla \widetilde{\psi}_{q,i}^{j,0} + \frac{1}{H} \sum_{e \in \mathcal{E}_H^0(K_j^+)} \int_e \left[\widetilde{\psi}_{p,i}^{j,0}\right] \left[\widetilde{\psi}_{q,i}^{j,0}\right] \right),$$

and

$$(52) \qquad S_{pq}^{j,0} = \sum_{i=1}^m \alpha_i \left(\frac{1}{2} \sum_{K \subset K_j^+} \int_{\partial K} |\mathbf{v}_i \cdot \mathbf{n}| \, \widetilde{\psi}_{p,i}^{j,0} \widetilde{\psi}_{q,i}^{j,0} \right),$$

where $V_{snap}^j = span\{\psi_p^j\}$, and $\widetilde{\psi}_p^j$ is the basis functions' energy minimizing extension. We further denote $\widetilde{\psi}_p^{j,0}$ the leading order asymptotic expansion of $\widetilde{\psi}_p^j$.

Proof. To proceed we notice that $\widetilde{\psi}_p^j \in V_{\text{snap}}^{j,+}$ can be written as the sum of some ψ^s's, where $\psi^s \in V_{\text{snap}}^s$ and $K_s \subset K_j^+$. Recall the assumption of (47). Then in each K_s and for all $i = 1, 2, \ldots, m$, we have

(53)

$$-\int_{K_j} \widetilde{\psi}_{p,i}^j \nabla w_i \cdot \mathbf{v}_i + \int_{\partial K_j} \widetilde{\psi}_{p,i}^j w_i \mathbf{v}_i \cdot \mathbf{n} + \int_{K_j} \epsilon \widetilde{\psi}_{p,i}^j w_i + \int_{K_j} \frac{1}{\epsilon a^\delta} \left(\widetilde{\psi}_{p,i}^j - \sum_{l=1}^m \alpha_l \widetilde{\psi}_{p,l}^j \right) w_i = 0.$$

Here w could be any function in $(V_h)^m$ restricted on K_j.

Therefore we have

(54)

$$\sum_{i=1}^m \alpha_i \left(-\int_{K_j} \widetilde{\psi}_{p,i}^j \nabla w_i \cdot \mathbf{v}_i + \int_{\partial K_j} \widetilde{\psi}_{p,i}^j w_i \mathbf{v}_i \cdot \mathbf{n} + \int_{K_j} \epsilon \widetilde{\psi}_{p,i}^j w_i \right) + \int_{K_j} \frac{1}{\epsilon a^\delta} \sum_{i,l=1}^m a_{il} \widetilde{\psi}_{p,l}^{j,0} w_i = 0.$$

Take the asymptotic expansion for $\widetilde{\psi}_p^j$,

(55)
$$\widetilde{\psi}_p^j = \widetilde{\psi}_p^{j,0} + \epsilon \widetilde{\psi}_p^{j,1} + \mathcal{O}\left(\epsilon^2\right),$$

set $w = \widetilde{\psi}_p^{j,0}$, and plug them back into (54). We have, in the leading order of $\frac{1}{\epsilon}$,

$$\sum_{i,l=1}^m a_{il} \widetilde{\psi}_{p,l}^{j,0} \widetilde{\psi}_{p,i}^{j,0} = 0,$$

meaning $\widetilde{\psi}_p^{j,0}$ is isotropic in each K_s due to Proposition 1. Therefore $\widetilde{\psi}_p^{j,0}$ is isotropic. The same analysis is applied to $\widetilde{\psi}_q^{j,0}$.

Recalling the definition of A^j and a^n in (29), we have

$$a^j(\psi_p^j, \psi_q^j) = \sum_{i=1}^m \alpha_i \left(\int_{K_j^+} \nabla \widetilde{\psi}_{p,i}^j \cdot \nabla \widetilde{\psi}_{q,i}^j + \frac{1}{H} \sum_{e \in \mathcal{E}_H^0(K_j^+)} \int_e \left[\widetilde{\psi}_{p,i}^j \right] \left[\widetilde{\psi}_{q,i}^j \right] \right)$$

$$+ \int_{K_j^+} \frac{1}{\epsilon a^\delta} \sum_{i,l=1}^m a_{il} \widetilde{\psi}_{p,l}^j \widetilde{\psi}_{q,i}^j$$

$$= \frac{1}{\epsilon} \left[\int_{K_j^+} \frac{1}{a^\delta} \sum_{i,l=1}^m a_{il} \widetilde{\psi}_{p,l}^{j,0} \widetilde{\psi}_{q,i}^{j,0} \right]$$

$$+ 1 \left[\sum_{i=1}^m \alpha_i \left(\int_{K_j^+} \nabla \widetilde{\psi}_{p,i}^{j,0} \cdot \nabla \widetilde{\psi}_{q,i}^{j,0} + \frac{1}{H} \sum_{e \in \mathcal{E}_H^0(K_j^+)} \int_e \left[\widetilde{\psi}_{p,i}^{j,0} \right] \left[\widetilde{\psi}_{q,i}^{j,0} \right] \right) \right.$$

$$\left. + \int_{K_j^+} \frac{1}{a^\delta} \sum_{i,l=1}^m a_{il} \widetilde{\psi}_{p,l}^{j,0} \widetilde{\psi}_{q,i}^{j,1} + \int_{K_j^+} \frac{1}{a^\delta} \sum_{i,l=1}^m a_{il} \widetilde{\psi}_{p,l}^{j,1} \widetilde{\psi}_{q,i}^{j,0} \right]$$

$$+ \mathcal{O}(\epsilon).$$

Due to Proposition 1, we have

$$\int_{K_j^+} \frac{1}{a^\delta} \sum_{i,l=1}^m a_{il} \widetilde{\psi}_{p,l}^{j,0} \widetilde{\psi}_{q,i}^{j,0} = \int_{K_j^+} \frac{1}{a^\delta} \sum_{i,l=1}^m a_{il} \widetilde{\psi}_{p,l}^{j,0} \widetilde{\psi}_{q,i}^{j,1} = \int_{K_j^+} \frac{1}{a^\delta} \sum_{i,l=1}^m a_{il} \widetilde{\psi}_{p,l}^{j,1} \widetilde{\psi}_{q,i}^{j,0} = 0,$$

and thus

$$a^j(\psi_p^j, \psi_q^j) = \sum_{i=1}^m \alpha_i \left(\int_{K_j^+} \nabla \widetilde{\psi}_{p,i}^{j,0} \cdot \nabla \widetilde{\psi}_{q,i}^{j,0} + \frac{1}{H} \sum_{e \in \mathcal{E}_H^0(K_j^+)} \int_e \left[\widetilde{\psi}_{p,i}^{j,0} \right] \left[\widetilde{\psi}_{q,i}^{j,0} \right] \right) + \mathcal{O}(\epsilon)$$

$$\to \sum_{i=1}^m \alpha_i \left(\int_{K_j^+} \nabla \widetilde{\psi}_{p,i}^{j,0} \cdot \nabla \widetilde{\psi}_{q,i}^{j,0} + \frac{1}{H} \sum_{e \in \mathcal{E}_H^0(K_j^+)} \int_e \left[\widetilde{\psi}_{p,i}^{j,0} \right] \left[\widetilde{\psi}_{q,i}^{j,0} \right] \right).$$

The proof for s^j is the same and is omitted here. □

Theorem 3 is a straightforward consequence of the following perturbation theorem.

Proof for Theorem 3. According to Lemma 2, A^j and S^j have expansions $A^j = A^{j,0} + \mathcal{O}(\epsilon) =: A^{j,0} + \epsilon A^{j,1}$ and $S^j = S^{j,0} + \mathcal{O}(\epsilon) =: S^{j,0} + \epsilon S^{j,1}$. We also define $\mathbf{x}_k^{j,0}$ as the kth generalized eigenvector of the two matrices $A^{j,0}$ and $S^{j,0}$, i.e.,

(56) $$A^{j,0} \mathbf{x}_k^{j,0} = \lambda_{j,k}^0 S^{j,0} \mathbf{x}_k^{j,0}.$$

Using the absolute Weyl theorem for GEPs in [28], when ϵ is small enough such that $\epsilon \left\| S^{j,1} \right\|_2 < \lambda_{\min}(S^{j,0})$, we have

$$\begin{aligned}
\left| \lambda_{j,k} - \lambda_{j,k}^0 \right| &\leq \frac{\left\| \epsilon A^{j,1} \right\|_2}{\lambda_{\min}(S^{j,0})} + \frac{\left\| A^{j,0} \right\|_2 + \left\| \epsilon A^{j,1} \right\|_2}{\lambda_{\min}(S^{j,0})(\lambda_{\min}(S^{j,0}) - \left\| \epsilon S^{j,1} \right\|_2)} \left\| \epsilon S^{j,1} \right\|_2 \\
&= \frac{\epsilon \left\| A^{j,1} \right\|_2}{\lambda_{\min}(S^{j,0})} + \frac{\left\| A^{j,0} \right\|_2 + \epsilon \left\| A^{j,1} \right\|_2}{\lambda_{\min}(S^{j,0})(\lambda_{\min}(S^{j,0}) - \epsilon \left\| S^{j,1} \right\|_2)} \epsilon \left\| S^{j,1} \right\|_2 \\
&= \frac{\epsilon \mathcal{O}(1)}{\mathcal{O}(1)} + \frac{\mathcal{O}(1) + \epsilon \mathcal{O}(1)}{\mathcal{O}(1)(\mathcal{O}(1) - \epsilon \mathcal{O}(1))} \epsilon \mathcal{O}(1) \\
&= \mathcal{O}(\epsilon),
\end{aligned}$$

where $\|\cdot\|_2$ is the spectral norm of a matrix. □

According to the formula for $A^{j,0}$ and $S^{j,0}$ in (51) and (52), the eigenvalues are positive except that the smallest one is 0 with constant as corresponding eigenvector. So Λ_* has positive limit in the limiting regime of $\epsilon \to 0$.

5. Numerical results. We take boundary condition $g(x, \mathbf{v}) = \cos(2\pi(x_1 + x_2)) + 1$. And we set $m = 6$, and use the Gaussian quadrature rule to define $\{(\mathbf{v}_i, \alpha_i), i = 1, \ldots, m\}$. As for a^δ, we give two examples. In the first example, we will choose a^δ to be based on a high contrast media shown in Figure 3 (left) and choose a^δ to be an oscillatory function for the second example used in [25, 14, 23], shown in Figure 3 (right), with the expression

$$a^\delta = \frac{2 + 1.8 \sin(10\pi x_1)}{2 + 1.8 \cos(10\pi x_2)} + \frac{2 + \sin(10\pi x_2)}{2 + 1.8 \sin(10\pi x_1)}.$$

The space domain Ω is taken as the unit square $[0, 1] \times [0, 1]$ and is divided into 10×10 coarse blocks consisting of uniform squares. Each coarse element is then divided into 10×10 fine elements consisting of uniform squares. That is, Ω is partitioned by 100×100 square fine elements. And we use the oversampling technique in (24)–(25) to obtain the snapshot space. We define an oversampling region K_j^+ by enlarging K_j by one coarse grid layer.
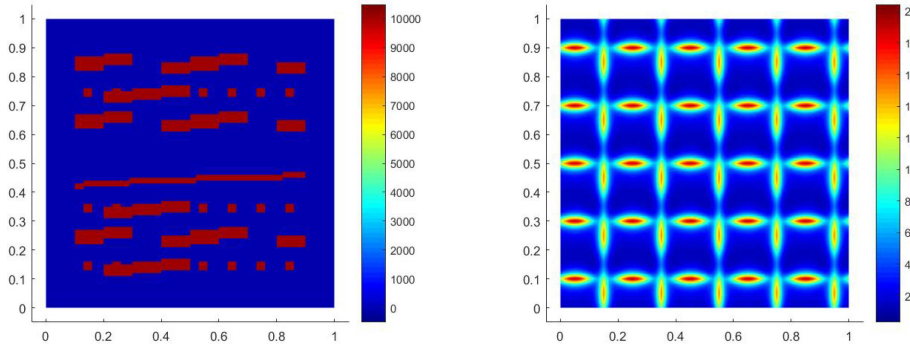
FIG. 3. *Left:* a^δ *for Example* 1. *Right:* a^δ *for Example* 2.

TABLE 1
Errors for Example 1 *with* $\epsilon = 10^{-1}$ *and* $a^\delta = \kappa^4$.

L	Snapshot ratio	e_1	e_2
1	0.79%	19.64%	9.41%
2	1.59%	17.68%	8.53%
3	2.38%	14.41%	7.40%
5	3.97%	8.11%	4.92%
7	5.56%	6.16%	3.62%
10	7.94%	3.44%	1.62%
15	11.90%	2.24%	1.04%
20	15.87%	1.64%	0.68%

To compare the accuracy, we will use the following error quantities:

$$e_1 = \left(\frac{\sum_{i=1}^m \alpha_i \int_\Omega |u_{h_i} - u_{H,i}|^2}{\sum_{i=1}^m \alpha_i \int_\Omega |u_{h_i}|^2} \right)^{1/2}, \quad e_2 = \left(\frac{\int_\Omega |\overline{u_h} - \overline{u_H}|^2}{\int_\Omega |\overline{u_h}|^2} \right)^{1/2},$$

where $\overline{u}$ is defined as $\overline{u} = \sum_{i=1}^m \alpha_i u_i$.

For Example 1, we first fix $a^\delta = \kappa^4$ and give the error tables for Knudsen number $\epsilon = 10^{-1}, 10^{-2}, 10^{-3}$, respectively. And L is the number of multiscale bases chosen from each coarse element, and snapshot ratio is defined by

$$\text{snapshot ratio} = \frac{\dim(V_H)}{\dim(V_{\text{snap}})}.$$

From Tables 1, 2, and 3, we can see this framework works for all Knudsen number ϵ, which verifies our proved conclusion. In addition, we see clearly the reduction of error when more basis functions are used, and the reduction of error is more rapid when fewer basis functions are used. We also observe that the method gives reasonable error levels with small snapshot ratios. On the other hand, Figure 4 shows the fine and multiscale solutions with $\epsilon = 10^{-2}$ and $L = 5$. From these figures, we observe very good agreements between the fine scale and multiscale solutions

Next, we fix $\epsilon = 10^{-2}$ and change the high contrast value of a^δ. We set $a^\delta = \kappa^2$, κ^4, κ^6, respectively. From Table 4, we can see that contrast values do not affect the error.

For Example 2, we give the error tables for $\epsilon = 5 \times 10^{-2}, 5 \times 10^{-3}, 5 \times 10^{-4}$, respectively. We present the errors for using various choices of number of basis functions in Tables 5, 6, and 7. We clearly see that, with a very small snapshot ratio, our method

TABLE 2
Errors for Example 1 with $\epsilon = 10^{-2}$ and $a^\delta = \kappa^4$.

L	Snapshot ratio	e_1	e_2
1	0.79%	12.05%	11.69%
2	1.59%	15.35%	15.17%
3	2.38%	3.73%	3.44%
5	3.97%	2.90%	2.64%
7	5.56%	2.61%	2.41%
10	7.94%	1.86%	1.67%
15	11.90%	1.20%	0.98%
20	15.87%	1.04%	0.83%

TABLE 3
Errors for Example 1 with $\epsilon = 10^{-3}$ and $a^\delta = \kappa^4$.

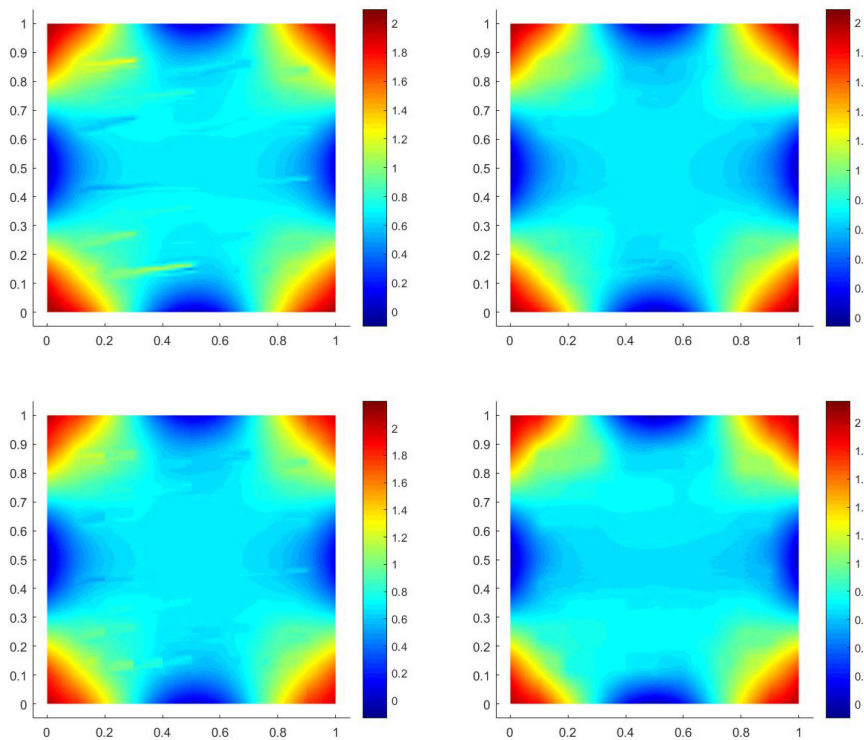| L	Snapshot ratio	e_1	e_2
1	0.79%	12.80%	12.80%
2	1.59%	26.43%	26.42%
3	2.38%	17.86%	17.85%
5	3.97%	4.45%	4.43%
7	5.56%	3.60%	3.59%
10	7.94%	3.55%	3.53%
15	11.90%	3.20%	3.18%
20	15.87%	3.19%	3.17%

FIG. 4. *Fine solution and multiscale solution for Example* 1. *Top left:* $u_{h,1}$. *Top right:* $\overline{u_h}$. *Bottom left:* $u_{H,1}$. *Bottom right:* $\overline{u_H}$.

TABLE 4
e_2 for Example 1 with different high contrast values of a^δ.

L	κ^2	κ^4	κ^6
1	11.70%	11.69%	11.68%
2	15.19%	15.17%	15.17%
3	3.41%	3.44%	3.45%
5	2.60%	2.64%	2.64%
7	2.37%	2.41%	2.41%
10	1.64%	1.67%	1.67%
15	0.96%	0.98%	0.98%
20	0.81%	0.83%	0.83%

TABLE 5
Errors for Example 2 with $\epsilon = 5 \times 10^{-2}$.

L	Snapshot ratio	e_1	e_2
1	0.79%	22.70%	9.73%
2	1.59%	20.36%	8.43%
3	2.38%	16.97%	8.13%
5	3.97%	11.94%	6.86%
7	5.56%	8.09%	4.64%
10	7.94%	4.70%	1.99%
15	11.90%	2.48%	1.22%
20	15.87%	1.86%	0.91%

TABLE 6
Errors for Example 2 with $\epsilon = 5 \times 10^{-3}$.

L	Snapshot ratio	e_1	e_2
1	0.79%	12.76%	11.98%
2	1.59%	11.02%	10.64%
3	2.38%	3.40%	2.97%
5	3.97%	2.04%	1.67%
7	5.56%	1.77%	1.43%
10	7.94%	1.50%	1.21%
15	11.90%	1.38%	1.15%
20	15.87%	1.17%	0.95%

TABLE 7
Errors for Example 2 with $\epsilon = 5 \times 10^{-4}$.

L	Snapshot ratio	e_1	e_2
1	0.79%	14.12%	14.11%
2	1.59%	20.87%	20.86%
3	2.38%	11.69%	11.69%
5	3.97%	2.95%	2.95%
7	5.56%	2.71%	2.71%
10	7.94%	2.71%	2.71%
15	11.90%	2.88%	2.88%
20	15.87%	2.93%	2.92%

is able to obtain solutions with very good accuracy. Furthermore, we observe a faster decay of the error when smaller numbers of basis functions are used. In Figure 5, we present the fine and multiscale solutions with $\epsilon = 5 \times 10^{-3}$ and $L = 5$. We observe very good agreement of both solutions.
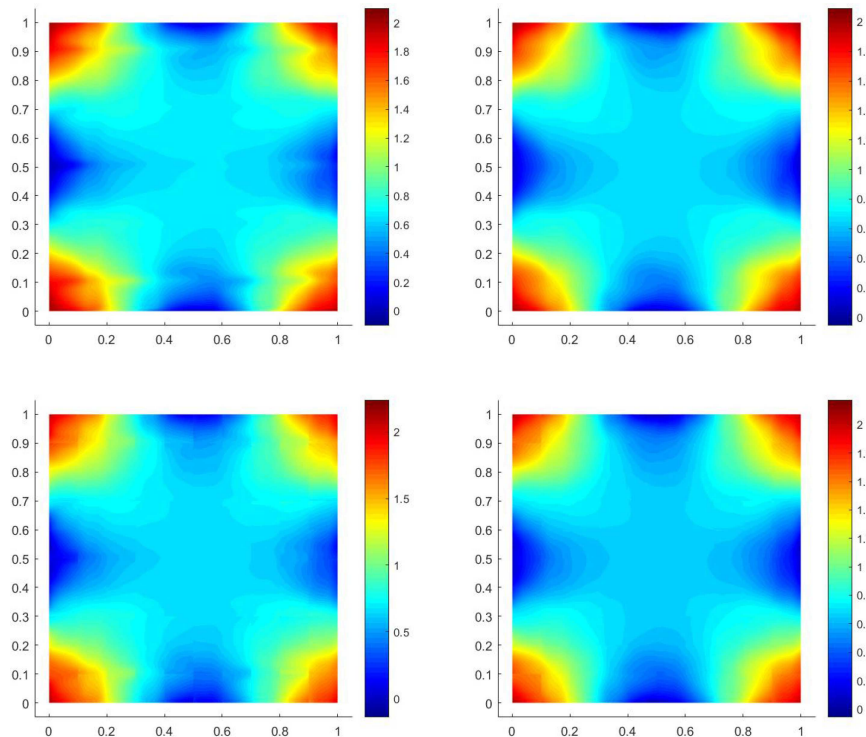
FIG. 5. *Fine solution and multisciale solution for Example* 2. *Top left:* $u_{h,1}$. *Top right:* $\overline{u_h}$. *Bottom left:* $u_{H,1}$. *Bottom right:* $\overline{u_H}$.

REFERENCES

[1] J. E. AARNES, *On the use of a mixed multiscale finite element method for greater flexibility and increased speed or improved accuracy in reservoir simulation*, SIAM J. Multiscale Model. Simul., 2 (2004), pp. 421–439.

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

[3] C. BARDOS, R. SANTOS, AND R. SENTIS, *Diffusion approximation and computation of the critical size*, Trans. Amer. Math. Soc., 284 (1984), pp. 617–649.

[4] A. BENSOUSSAN, J.-L. LIONS, AND G. PAPANICOLAOU, *Asymptotic Analysis for Periodic Structures*, AMS, Providence, RI, 2011.

[5] V. CALO, Y. EFENDIEV, J. GALVIS, AND G. LI, *Randomized Oversampling for Generalized Multiscale Finite Element Methods*, preprint, arXiv:1409.7114, 2014.

[6] C.-C. CHU, I. GRAHAM, AND T.-Y. HOU, *A new multiscale finite element method for high-contrast elliptic interface problems*, Math. Comp., 79 (2010), pp. 1915–1955.

[7] E. CHUNG, Y. EFENDIEV, AND T. Y. HOU, *Adaptive multiscale model reduction with generalized multiscale finite element methods*, J. Comput. Phys., 320 (2016), pp. 69–95.

[8] E. CHUNG, Y. EFENDIEV, AND W. T. LEUNG, *Generalized multiscale finite element methods with energy minimizing oversampling*, Internat. J. Numer. Methods Engrg., 117 (2019), pp. 316–343.

[9] E. T. CHUNG, Y. EFENDIEV, AND S. FU, *Generalized multiscale finite element method for elasticity equations*, GEM Int. J. Geomath., 5 (2014), pp. 225–254.

[10] E. T. CHUNG, Y. EFENDIEV, AND W. T. LEUNG, *Generalized multiscale finite element methods for wave propagation in heterogeneous media*, Multiscale Model. Simul., 12 (2014), pp. 1691–1721.

[11] E. T. CHUNG AND W. T. LEUNG, *A sub-grid structure enhanced discontinuous galerkin method for multiscale diffusion and convection-diffusion problems*, Commun. Comput. Phys., 14 (2013), pp. 370–392.

[12] L. J. Durlofsky, *Numerical calculation of equivalent grid block permeability tensors for heterogeneous porous media*, Water Resources Research, 27 (1991), pp. 699–708.

[13] Y. Efendiev, J. Galvis, and X.-H. Wu, *Multiscale finite element methods for high-contrast problems using local spectral basis functions*, J. Comput. Phys., 230 (2011), pp. 937–955.

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

[15] Y. Efendiev and T. Y. Hou, *Multiscale Finite Element Methods: Theory and Applications*, Surv. Tutor. Appl. Math. Sci. 4, Springer, New York, 2009.

[16] Y. Efendiev, T. Y. Hou, and V. Ginting, *Multiscale finite element methods for nonlinear problems and their applications*, Commun. Math. Sci., 2 (2004), pp. 553–589.

[17] B. Engquist and O. Runborg, *Heterogeneous multiscale methods*, Commun. Math. Sci., 1 (2013), p. 132.

[18] K. Gao, E. T. Chung, R. L. Gibson, Jr., S. Fu, and Y. Efendiev, *A numerical homogenization method for heterogeneous, anisotropic elastic media based on multiscale theory*, Geophysics, 80 (2015), pp. D385–D401.

[19] K. Gao, S. Fu, R. L. Gibson, E. T. Chung, and Y. Efendiev, *Generalized multiscale finite-element method (GMsFEM) for elastic wave propagation in heterogeneous, anisotropic media*, J. Comput. Phys., 295 (2015), pp. 161–188.

[20] M. Ghommem, M. Presho, V. M. Calo, and Y. Efendiev, *Mode decomposition methods for flows in high-contrast porous media. Global–local approach*, J. Comput. Phys., 253 (2013), pp. 226–238.

[21] R. L. Gibson, Jr., K. Gao, E. Chung, and Y. Efendiev, *Multiscale modeling of acoustic wave propagation in 2D media*, Geophysics, 79 (2014), pp. T61–T75.

[22] T. Goudon and A. Mellet, *Homogenization and diffusion asymptotics of the linear Boltzmann equation*, ESAIM Control Optim. Calc. Var., 9 (2003), pp. 371–398.

[23] T. Y. Hou and X. H. Wu, *A multiscale finite element method for PDEs with oscillatory coefficients*, in Numerical Treatment of Multi-Scale Problems, W. Hackbusch and G. Wittum, eds., Notes Numer. Fluid Mech. 70, Springer, Berlin, 1999, pp. 58–69.

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

[25] Q. Li and J. Lu, *An asymptotic preserving method for transport equations with oscillatory scattering coefficients*, Multiscale Model. Simul., 15 (2017), pp. 1694–1718.

[26] Q. Li, J. Lu, and W. Sun, *Diffusion approximations and domain decomposition method of linear transport equations: Asymptotics and numerics*, J. Comput. Phys., 292 (2015), pp. 141–167.

[27] Q. Li, J. Lu, and W. Sun, *Validity and regularization of classical half-space equations*, J. Statist. Phys., 166 (2017), pp. 398–433.

[28] Y. Nakatsukasa, *Absolute and relative weyl theorems for generalized eigenvalue problems*, Linear Algebra Appl., 432 (2010), pp. 242–248.

[29] L. Wu and Y. Guo, *Geometric correction for diffusive expansion of steady neutron transport equation*, Commun. Math. Phys., 336 (2015), pp. 1473–1553.

[30] X.-H. Wu, Y. Efendiev, and T. Y. Hou, *Analysis of upscaling absolute permeability*, Discrete Contin. Dyn. Syst. Ser. B, 2 (2002), pp. 185–204.