

Generalized multiscale finite element methods for wave propagation in heterogeneous media

Eric T. Chung*, Yalchin Efendiev† and Wing Tat Leung‡

May 1, 2018

Abstract

Numerical modeling of wave propagation in heterogeneous media is important in many applications. Due to the complex nature, direct numerical simulations on the fine grid are prohibitively expensive. It is therefore important to develop efficient and accurate methods that allow the use of coarse grids. In this paper, we present a multiscale finite element method for wave propagation on a coarse grid. The proposed method is based on the Generalized Multiscale Finite Element Method (GMsFEM) (see [13]). To construct multiscale basis functions, we start with two snapshot spaces in each coarse-grid block where one represents the degrees of freedom on the boundary and the other represents the degrees of freedom in the interior. We use local spectral problems to identify important modes in each snapshot space. These local spectral problems are different from each other and their formulations are based on the analysis. To our best knowledge, this is the first time where multiple snapshot spaces and multiple spectral problems are used and necessary for efficient computations. Using the dominant modes from local spectral problems, multiscale basis functions are constructed to represent the solution space locally within each coarse block. These multiscale basis functions are coupled via the symmetric interior penalty discontinuous Galerkin method which provides a block diagonal mass matrix, and, consequently, results in fast computations in an explicit time discretization. Our methods' stability and spectral convergence are rigorously analyzed. Numerical examples are presented to show our methods' performance. We also test oversampling strategies. In particular, we discuss how the modes from different snapshot spaces can affect the proposed methods' accuracy.

1 Introduction

Numerical modeling of wave propagation is important in many applications that include geophysics, material science, and so on. For example, in geophysics applications, wave propagation simulations play an important role in determining subsurface properties [32, 31, 29, 23, 30, 24]. These approaches include finite difference methods, finite element methods, and spectral methods that use polynomials basis [11, 19, 27, 18, 25, 20, 22, 21, 26, 33, 34, 35]. While these methods have different strengths and weaknesses, all of them tend to have limitations associated with discretization, especially in 3-D applications as frequency content of the simulated wavefield increases. Though the solutions to the wave equation have been shown to be accurate when the grid is fine enough [12], the practical limitations in discretization caused by limitations in computational power restrict this accuracy. An example of an application where this may be important is in the modeling of fractured media, where establishing reliable and accurate relationships between the properties of reflected seismic wavefields and variations in the density, orientation and compliance of fractures may help provide important constraints for hydrocarbon

*Department of Mathematics, The Chinese University of Hong Kong, Hong Kong SAR. Eric Chung's research is supported by the Hong Kong RGC General Research Fund project 400411.

†Department of Mathematics, Texas A&M University, College Station, TX 77843, USA & center for Numerical Porous Media (NumPor), King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Kingdom of Saudi Arabia

‡Department of Mathematics, The Chinese University of Hong Kong, Hong Kong SAR.

production. While more general finite element and spectral element methods may be able to address some problems by adapting grids to conform to heterogeneous structures, there are basic limitations associated with representing fine-scale features, and there is therefore a need to find approaches that reliably and accurately incorporate fine-scale features in a coarsely gridded model.

In this paper, we present a multiscale finite element method for wave propagation simulations on a coarse grid. The proposed method is based on the Generalized Multiscale Finite Element Method (GMsFEM) which was proposed in [13] and couples multiscale basis functions via a discontinuous Galerkin coupling (cf. [14]). To construct multiscale basis functions, we start with two snapshot spaces in each coarse-grid block where one represents the degrees of freedom on the coarse grid's boundary and the other represents the degrees of freedom in the interior. We use local spectral problems to identify important modes in each snapshot space. These local spectral problems are different from each other and their formulations are based on the analysis. Once local basis functions are identified, we couple these basis functions via Interior Penalty Discontinuous Galerkin method [17, 28].

Because these basis functions are discontinuous, the interior penalty discontinuous Galerkin (IPDG) method, for example [17, 10, 28], is an appropriate choice for solving the time-dependent partial differential equation. It generally yields a block diagonal mass matrix, hence the time stepping is very efficient. The staggered discontinuous Galerkin methods [3, 4] have been recently developed for the accurate wave simulations. By using a carefully chosen staggered grid, the resulting method is also energy conserving. Moreover, it is proved that (see [1, 6]) such method gives smaller dispersion errors, and therefore it is superior for the wave propagation. The staggered idea has also been extended to other problems, see for example [7, 8, 9, 5]. Recently, we have used standard MsFEM basis within staggered methods [2, 16]. These methods allow some limited upscaling and provide energy conserving numerical methods on staggered grids. In this paper, our goal is to construct a systematic enrichment by appropriately choosing snapshot spaces and corresponding local spectral problems.

We will focus our discussions on two-dimensional problems. The extension to three-dimensional problems is straightforward. Let $\Omega \subset \mathbb{R}^2$ be a bounded domain of two dimensions. The paper's aim is to develop a new multiscale method for the following wave equation

$$\frac{\partial^2 u}{\partial t^2} = \nabla \cdot (a \nabla u) + f \quad \text{in } [0, T] \times \Omega \quad (1)$$

with the homogeneous Dirichlet boundary condition $u = 0$ on $[0, T] \times \partial\Omega$. The extension to other boundary conditions will be reported in a forthcoming paper. The function $f(x, t)$ is a given source. The problem (1) is supplemented with the following initial conditions

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

We assume that the coefficient $a(x)$ is highly oscillatory, representing the complicated model in which the waves are simulated. It is well-known that solving (1) by standard methods requires a very fine mesh, which is computationally prohibited. Thus a coarse grid solution strategy is needed. Next we present our fine scale solver. The fine scale solution is considered as the exact solution when we discuss the convergence of our multiscale method in the following sections. We assume that the domain Ω is partitioned by a set of rectangles, called fine mesh, with maximum side length $h > 0$. We denote the resulting mesh by \mathcal{T}^h and the set of all edges and vertices by \mathcal{E}^h and \mathcal{N}^h respectively. We assume that the fine-mesh discretization of the wave equation provides an accurate approximation of the solution. The fine scale solver is the standard conforming bilinear finite element method. Let V_h be the standard conforming piecewise bilinear finite element space. We find $u_h \in V_h$ such that

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

where the bilinear form a is defined by

$$a(u, v) = \int_{\Omega} a \nabla u \cdot \nabla v, \quad \forall u, v \in V_h \quad (3)$$

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

The numerical results are presented for several representative examples. We investigate the GMS-FEM's accuracy and, in particular, how choosing modes from different snapshot spaces can affect the accuracy. Our numerical results show that choosing the basis functions from interior modes can improve the accuracy of GMSFEM substantially for wave equations. These results differ from those we observe for flow equations [13].

The paper is organized as follows. In Section 2, we will present the new multiscale method. Numerical results are shown in Section 3. Stability and spectral convergence of the semi-discrete scheme are proved in Section 4. In Section 5, the convergence of the fully-discrete scheme is also proved. Finally, conclusions are presented.

2 The generalized multiscale finite element method

In this section, we will give a detailed description of our new generalized multiscale finite element method. The method gives a numerical solver on a coarse grid, providing an efficient way to simulate waves in complicated media. As we will discuss next, the local basis functions are obtained via the solutions of some local spectral problems which are used to obtain the most dominant modes. These modes form the basis functions of our multiscale finite element method.

We introduce a coarse mesh that consists of union of connected fine-mesh grid blocks which is denoted by \mathcal{T}^H and the set of all edges by \mathcal{E}^H . We denote the size of the coarse mesh by H . Even though it is convenient to choose rectangular coarse grid blocks, the shapes of the coarse grid blocks can be quite general and our analysis can be applied without the assumption of rectangular coarse grid blocks.

For each coarse grid block K , we define $\partial\mathcal{T}^h(K)$ be the restriction of the conforming piecewise bilinear functions with respect to the fine mesh on ∂K . We remark that, for a coarse grid edge $e \in \mathcal{E}^H$ that is shared by two coarse grid blocks K_1 and K_2 , the values of the two functions in $\partial\mathcal{T}^h(K_1)$ and $\partial\mathcal{T}^h(K_2)$ on e are in general different. The union of all $\partial\mathcal{T}^h(K)$ is denoted by $\partial\mathcal{T}^h$. Moreover, we define $H^1(\mathcal{T}^H)$ as the space of functions whose restrictions on K belongs to $H^1(K)$.

2.1 Global IPDG solver

We will apply the standard symmetric IPDG approach to solve (1) on the coarse grid \mathcal{T}^H . The method follows the standard framework as discussed in [17, 28], but the finite element space will be replaced by the space spanned by our multiscale basis functions. We emphasize that the use of the IPDG approach is an example of the global coupling of our local multiscale basis functions, and other choices of coarse grid methods are equally good. The key to our proposed method's success of is the construction of our local multiscale basis functions.

First, we introduce some notations. For each interior coarse edge $e \in \mathcal{E}^H$, we let K^- and K^+ be the two coarse grid blocks having the common coarse edge e . Then we define the average and the jump operators respectively by

$$\begin{aligned} \{v\}_e &= \frac{v^+ + v^-}{2}, \\ [u]_e &= u^+ - u^-, \end{aligned}$$

where $u^\pm = u|_{K^\pm}$ and we have assumed that the normal vector on e is pointing from K^+ to K^- . For each coarse edge e that lies on the boundary of Ω , we define

$$\{v\}_e = v, \quad [u]_e = u$$

assuming the unit normal vector on e is pointing outside the domain. Let V_H be a finite dimensional function space which consists of functions that are smooth on each coarse grid blocks but are in general

discontinuous across coarse grid edges. We can then state the IPDG method as: find $u_H(t, \cdot) \in V_H$ such that

$$\left(\frac{\partial u_H}{\partial t^2}, v\right) + a_{DG}(u_H, v) = l(v), \quad \forall v \in V_H, \quad (4)$$

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

$$a_{DG}(u, v) = \sum_{K \in \mathcal{T}^H} \int_K a \nabla u \cdot \nabla v + \sum_{e \in \mathcal{E}^H} \left(- \int_e \{a \nabla u \cdot n\}_e [v]_e - \int_e \{a \nabla v \cdot n\}_e [u]_e + \frac{\gamma}{h} \int_e a [u]_e [v]_e \right)$$

$$l(v) = (f, v)$$

where $\gamma > 0$ is a penalty parameter and n denotes the unit normal vector on e . The initial conditions for the problem (4) are defined by $u_H(0) = P_H(g_0)$ and $(u_H)_t(0) = P_H(g_1)$, where P_H is the L^2 -projection operator into V_H .

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

$$(u_H^{n+1}, v) = 2(u_H^n, v) - (u_H^{n-1}, v) - \Delta t^2 \left(a_{DG}(u_H^n, v) - l(v) \right), \quad \forall v \in V_H \quad (5)$$

in each time step. Throughout the paper, the notation u^n represents the value of the function u at time t_n . The initial conditions are obtained as follows

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

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

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

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

2.2 Multiscale basis functions

In this section, we will give the definition of the space V_H . We will discuss the choice of our basis functions on one single coarse grid block K , as the definitions on other coarse grid blocks are similar. We recall that K is the union of a set of rectangular elements. We decompose the space V_H into two components, namely

$$V_H = V_H^1 + V_H^2.$$

The restrictions of V_H , V_H^1 and V_H^2 on K are denoted by $V_H(K)$, $V_H^1(K)$, and $V_H^2(K)$ respectively. Moreover, the restriction of the conforming space V_h on K is denoted by $V_h(K)$.

Definition of $V_H^1(K)$. To define $V_H^1(K)$, for each fine grid node x_i on the boundary of K , we find $w_{i,K} \in V_h(K)$ by solving

$$\int_K a \nabla w_{i,K} \cdot \nabla v = 0, \quad \forall v \in V_h(K) \quad (6)$$

with boundary condition $w_{i,K} = 1$ at x_i and $w_{i,K} = 0$ at the other grid points on the boundary of K . The functions $w_{i,K}$ defined above are the a -harmonic extensions of the unit basis functions of $\partial \mathcal{T}^h(K)$. We let n be the number of these a -harmonic extensions and define

$$V_H^1(K) = \text{span}\{w_{1,K}, \dots, w_{n,K}\}.$$

We remark that n is the number of boundary grid points on ∂K and its value changes with K . In our numerical simulations, we do not need to use all of these basis functions and use a local spectral problem in the space of snapshots to identify multiscale basis functions by choosing dominant modes. We use E to denote the sum of the reciprocal of the eigenvalues of the local spectral problem. We will choose

the eigenfunctions corresponding to small eigenvalues so that the sum of the reciprocals these small eigenvalues is a small percentage of E . The use of eigenfunctions corresponding to small eigenvalues means that we use the coarse component in $V_H^1(K)$ as the approximation space.

Local spectral problem on $V_H^1(K)$. The spectral problem we propose is

$$\int_K a \nabla w_\mu \cdot \nabla v = \frac{\mu}{H} \int_{\partial K} w_\mu v, \quad \forall v \in V_H^1(K). \quad (7)$$

We assume that the eigenvalues are ordered so that $0 = \mu_1 < \mu_2 \leq \mu_3 \leq \dots \leq \mu_n$. The corresponding eigenfunctions in the snapshot space are denoted by $\tilde{w}_{i,K}$, $i = 1, 2, \dots, n$, which are normalized with respect to the L^2 -norm on ∂K . We denote the total energy on the coarse grid block K by E_K which is defined by $E_K = \sum_{i=2}^n \mu_i^{-1}$. We can then choose the first p eigenvalues so that the sum $\sum_{i=2}^p \mu_i^{-1}$ is a portion of the total energy E_K . Note that we can take different p for different coarse grid blocks K . Finally we define

$$\tilde{V}_H^1(K) = \text{span}\{\tilde{w}_{1,K}, \dots, \tilde{w}_{p,K}\}.$$

Clearly, $\tilde{V}_H^1(K) \subset V_H^1(K)$.

Definition of $V_H^2(K)$. The space $V_H^2(K)$ contains functions in $V_h(K)$ that are zero on the boundary of K , which is denoted by $V_h^0(K)$.

Local spectral problem on $V_H^2(K)$. We will use a suitable (another) spectral problem to identify the important modes. The proposed eigenvalue problem has the following form: find $z_\lambda \in V_h^0(K)$ such that

$$\int_K a \nabla z_\lambda \cdot \nabla v = \frac{\lambda}{H^2} \int_K z_\lambda v, \quad \forall v \in V_h^0(K). \quad (8)$$

Assume that the eigenvalues are ordered so that $\lambda_{1,K} \leq \lambda_{2,K} \leq \dots$ and the corresponding eigenfunctions are denoted by $z_{i,K}$, which are normalized with respect to the L^2 -norm on K . In practice, for each coarse grid block, we can take the first m eigenfunctions, and the space $V_H^2(K)$ is spanned by these functions, that is

$$V_H^2(K) = \text{span}\{z_{1,K}, \dots, z_{m,K}\}.$$

In principle, one can choose different numbers of eigenfunctions for the space $V_H^2(K)$ for different coarse grid blocks. Nevertheless, our numerical results show that only the first few eigenfunctions are enough to obtain a reasonable accuracy.

Orthogonality of $V_H^1(K)$ and $V_H^2(K)$. Finally, we point out the following orthogonality condition which will be used in our analysis. For any $v \in V_H^1(K)$ and $u \in V_H^2(K)$, we conclude by (6) that

$$\int_K a \nabla v \cdot \nabla u = 0. \quad (9)$$

This means that the two spaces $V_H^1(K)$ and $V_H^2(K)$ are orthogonal.

3 Numerical Results

In this section, we will present some numerical examples to show the performance of our multiscale method. The media that we will consider is a heterogeneous field which is a modified Marmousi model (see the left plot of Figure 1). We have also considered more regular periodic highly heterogeneous fields and observed similar results. We will compare both the accuracy and efficiency of our method with the direct fine scale simulation defined in (2). To compare the accuracy, we will use the following error quantities

$$e_2 = \frac{\|u_H - u\|_{L^2(\Omega)}}{\|u\|_{L^2(\Omega)}}, \quad \bar{e}_2 = \frac{\sqrt{\sum_{K \in \mathcal{T}^H} |\int_K u_H - \int_K u|^2}}{\sqrt{\sum_{K \in \mathcal{T}^H} |\int_K u|^2}}, \quad e_{H^1} = \frac{\|\nabla(u_H - u)\|_{L^2(\Omega)}}{\|\nabla u\|_{L^2(\Omega)}}$$

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

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

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

The Ricker wavelet with frequency $f_0 = 20$

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

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

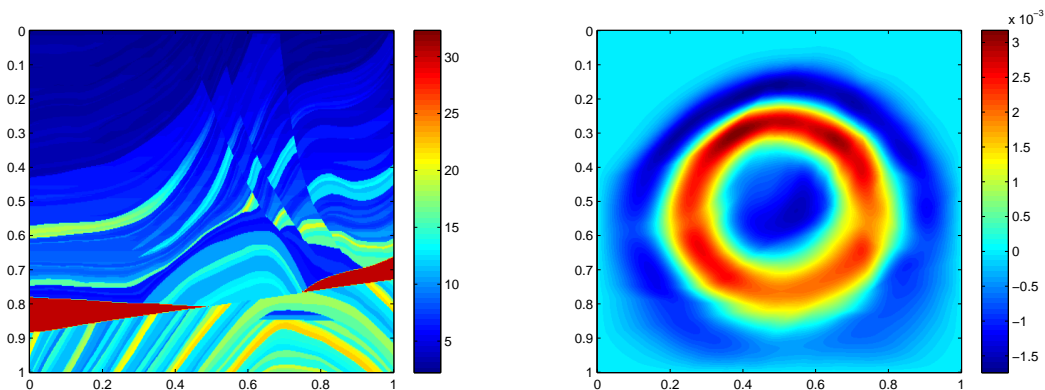


Figure 1: Left: a subset of the Marmousi model. Right: fine grid solution.

In Table 1 and Table 2, we present the errors and computational times for the case with $m = 1$, that is, we only use the first eigenfunction in the space V_H^2 . We see that if we use 80% of the total energy, the number of basis functions is between 33 and 40 on each coarse grid while the computational time for the offline procedure is 1019.06 and the time for online computations is 32.43. Note that the online computational time is about 59% of that of the online computational time of the direct fine grid simulation. The relative L^2 error and the relative error for cell averages are only 3.92% and 2.74% respectively. In addition, the relative error for the gradient is 14.86% and the jump error is 0.003. When 75% of the total energy is used, the number of basis functions is reduced to a number between 24 and 29 while the computational time for the offline procedures is 326.83 and the time for online computations is 18.21. The time for the online computation is 33% of the time required for direct fine grid simulation.

The relative L^2 error and the relative error for cell averages are increased slightly to 4.23% and 3.12%, respectively. In Table 1, we also present the values of μ_{min} for the space V_H^1 . Moreover, the eigenvalues are shown in Figure 3. The numerical solutions for these cases are shown in Figure 2. We note that the error decay is not fast mostly due to the error contribution because of the modes corresponding to the interior. Even though the error between the GMsFEM solution and the solution computed using the entire snapshot space V_H^1 is very small, the overall error between the GMsFEM solution and the fine-scale solution may not be small because we have only used one basis function in V_H^2 . Next, we will add more basis functions from V_H^2 and compare the errors.

Energy	Number of basis	e_2	\bar{e}_2	e_{H^1}	e_{Jump}	μ_{min}
75%	24-29	0.0423	0.0312	0.1542	4.7304e-04	1.9414
80%	33-40	0.0392	0.0274	0.1486	3.0671e-04	2.9992

Table 1: Errors for various choices of energy for the space V_H^1 .

Energy	t_{off}	t_{on}
75%	326.83	18.21
80%	1019.06	32.43

Table 2: Offline and online computational times.

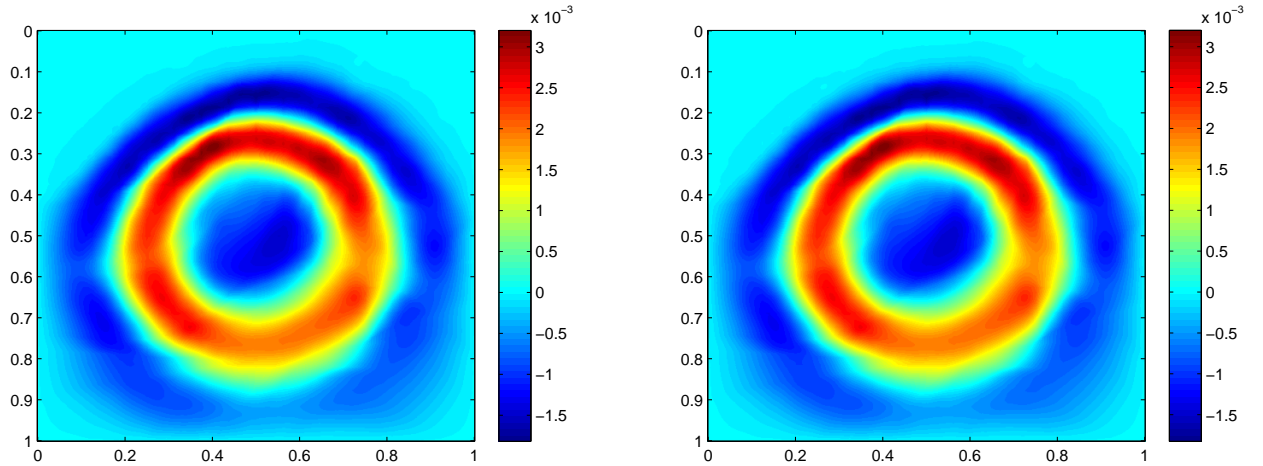


Figure 2: Left: 75% energy. Right: 80% energy.

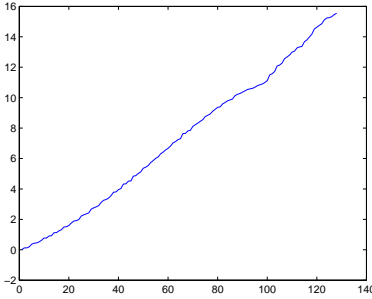


Figure 3: Eigenvalues for the space V_H^1 .

Next, we will investigate the use of more eigenfunctions in the space V_H^2 that will allow reducing the overall error. To do so, we consider the first case where 75% energy in the space V_H^1 is used and we consider using various number of eigenfunctions in V_H^2 . The errors and computational times are shown in Table 3 and Table 4. In general, we obtain better numerical approximations as more eigenfunctions are used. When two eigenfunctions are used (this corresponds to using less than 3% of the total local degrees of freedom in constructing all GMsFEM basis functions), the relative error is 3.52% and the online computational time is 18.64. When five eigenfunctions are used, the relative error is 1.93% and the online computational time is 18.21. Thus, we see that adding a few eigenfunctions in the space V_H^2 will improve the multiscale solution. This indicates that for the multiscale wave simulations, the modes that represent the interior nodes can improve the accuracy of the method and play an important role in obtaining an accurate solution. The numerical solution for these 4 cases are shown in Figure 4. We see that our method is able to capture the solution well. We also report the largest eigenvalue used in Table 3.

m	e_2	\bar{e}_2	e_{H^1}	e_{Jump}	λ_{min}
1	0.0423	0.0312	0.1542	4.7304e-04	3.4805e+04
2	0.0352	0.0259	0.1346	4.7030e-04	3.4873e+04
3	0.0227	0.0187	0.0945	4.5931e-04	5.5906e+04
5	0.0193	0.0163	0.0833	4.5910e-04	6.9650e+04

Table 3: Errors for various number of eigenfunctions in V_H^2 for using 75% energy in V_H^1 .

m	t_{off}	t_{on}
1	326.83	18.21
2	368.89	18.64
3	405.73	19.88
5	528.47	25.96

Table 4: Offline and online computational times for various number of eigenfunctions in V_H^2 for using 75% energy in V_H^1 .

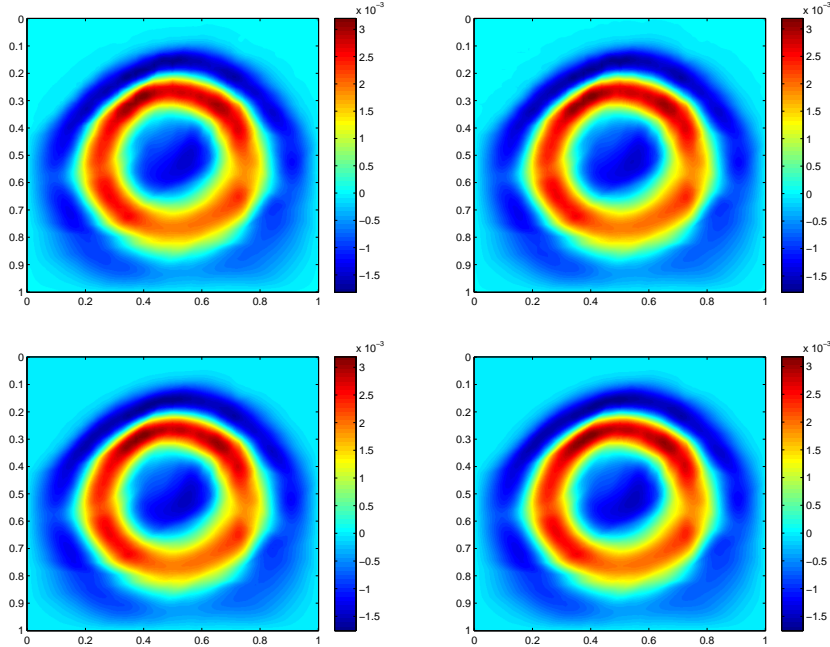


Figure 4: Numerical solutions with various number of eigenfunctions (m) in the space V_H^2 for using 75% energy in V_H^1 . Upper-left: $m = 1$. Upper-right: $m = 2$. Lower-left: $m = 3$. Lower-right: $m = 5$.

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

3.1 The use of oversampling

In this section, we present the performance of the method when the basis functions in the space V_H^1 are obtained by oversampling. We consider the previous example. The oversampling technique is used and the harmonic extension problems are solved on enlarged coarse grids, which are obtained by extending the original coarse grids by $H/16$ on each side. The results for using one basis functions in V_H^2 and various number of basis functions in V_H^1 are shown in Table 5. Moreover, we compute the errors using 73% energy for V_H^1 and various number of basis functions in V_H^2 . The results are presented in Table 6. We observe that there is no improvement in this case. This is due to the error from the modes representing internal nodes.

Energy	Number of basis	e_2	\bar{e}_2	e_{H^1}	e_{Jump}	μ_{min}
73%	24-30	0.0673	0.0583	0.1866	5.2038e-04	1.6755
79%	33-40	0.0640	0.0548	0.1827	3.4797e-04	2.5681
84%	45-55	0.0626	0.0534	0.1809	2.6388e-04	3.7918

Table 5: Simulation results with one basis function in V_H^2 .

4 Stability and convergence

In this section, we will prove the stability and convergence of the generalized multiscale finite element method constructed in Section 2. We will first state and prove some preliminary results, and then prove

m	e_2	$\overline{e_2}$	e_{H^1}	e_{Jump}	λ_{min}
1	0.0673	0.0583	0.1866	5.2038e-04	3.4805e+04
2	0.0596	0.0524	0.1666	5.1865e-04	3.4873e+04
3	0.0488	0.0449	0.1332	5.0929e-04	5.5906e+04
5	0.0449	0.0419	0.1220	5.0793e-04	6.9650e+04

Table 6: Errors and computational times for various number of eigenfunctions in V_H^2 for using $E = 73\%$.

the main convergence theorem for the semi-discrete scheme (4).

4.1 Preliminaries

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

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

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

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

and the a -semi-norm by

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

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

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

Assumption 1 *The function $a(x)$ is bounded, that is, there exist positive numbers a_0 and a_1 such that*

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

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

In the following, we will describe the consistency of the method (4). We define the consistency error by

$$R_{u_h}(v) = l(v) - \left(\frac{\partial^2 u_h}{\partial t^2}, v \right) - a_{DG}(u_h, v), \quad \forall v \in V_H, \quad (10)$$

where u_h is the fine grid finite element solution defined in (2). Clearly, we have

$$R_{u_h}(v) = 0, \quad \forall v \in V_H^2 \quad (11)$$

since $V_H^2 \subset V_h$. Thus, we only need to estimate $R_{u_h}(v)$ for $v \in V_H^1$. The following lemma states that the method (4) is consistent with the fine grid solution defined by (2). The proof will be presented in the Appendix.

Lemma 1 *Let u_h and u be the finite element solution defined in (2) and the exact solution of the wave propagation problem (1) respectively. If $u \in H^2(\Omega)$, then we have*

$$|R_{u_h}(v)| \leq C(u, f)h\|v\|_a, \quad v \in V_H^1 \quad (12)$$

where $C(u, f)$ is a constant which depends on the solution u and the source term f but independent of the fine mesh size h . This inequality gives the consistency of our method.

Next we will prove that the bilinear form a_{DG} satisfies the following coercivity and continuity conditions for suitably chosen penalty parameter $\gamma > 0$.

Lemma 2 *Let γ be sufficiently large. Then we have*

$$\frac{1}{2}\|v\|_a^2 \leq a_{DG}(v, v), \quad \forall v \in V_H \quad (13)$$

and

$$a_{DG}(u, v) \leq 2\|u\|_a\|v\|_a, \quad \forall u, v \in V_H. \quad (14)$$

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

$$a_{DG}(v, v) = \sum_{K \in \mathcal{T}^H} \int_K a \nabla v \cdot \nabla v + \sum_{e \in \mathcal{E}^H} \left(-2 \int_e \{a \nabla v \cdot n\}_e [v]_e + \frac{\gamma}{h} \int_e a [v]_e^2 \right),$$

and by the definition of the a -norm, we have

$$a_{DG}(v, v) = \|v\|_a^2 - 2 \sum_{e \in \mathcal{E}^H} \left(\int_e \{a \nabla v \cdot n\}_e [v]_e \right). \quad (15)$$

Using the Cauchy-Schwarz inequality, we obtain

$$2 \sum_{e \in \mathcal{E}^H} \left(\int_e \{a \nabla v \cdot n\}_e [v]_e \right) \leq \frac{2h}{\gamma a_0} \sum_{e \in \mathcal{E}^H} \int_e \{a \nabla v \cdot n\}_e^2 + \frac{a_0}{2} \sum_{e \in \mathcal{E}^H} \frac{\gamma}{h} \int_e [v]_e^2$$

which implies

$$2 \sum_{e \in \mathcal{E}^H} \left(\int_e \{a \nabla v \cdot n\}_e [v]_e \right) \leq \frac{h}{\gamma a_0} \sum_{K \in \mathcal{T}^H} \int_{\partial K} (a \nabla v \cdot n)^2 + \frac{1}{2} \sum_{e \in \mathcal{E}^H} \frac{\gamma}{h} \int_e a [v]_e^2.$$

Since v is a piecewise linear function, there is a uniform constant $\Lambda > 0$ such that

$$2h \sum_{K \in \mathcal{T}^H} \int_{\partial K} (a \nabla v \cdot n)^2 \leq \Lambda a_1 |v|_a^2. \quad (16)$$

So we have

$$2 \sum_{e \in \mathcal{E}^H} \left(\int_e \{a \nabla v \cdot n\}_e [v]_e \right) \leq \frac{\Lambda a_1}{2\gamma a_0} |v|_a^2 + \frac{1}{2} \sum_{e \in \mathcal{E}^H} \frac{\gamma}{h} \int_e a [v]_e^2.$$

Therefore, from (15), we obtain

$$\frac{1}{2}\|v\|_a^2 \leq a_{DG}(v, v), \quad \forall v \in V_H$$

if we take $\gamma \geq \Lambda a_1 a_0^{-1}$. Thus, we have proved (13).

We can prove (14) in the similar way. By the definition of a_{DG} , we have

$$\begin{aligned}
& |a_{DG}(u, v)| \\
&= \left| \sum_{K \in \mathcal{T}^H} \int_K a \nabla u \cdot \nabla v + \sum_{e \in \mathcal{E}^H} \left(- \int_e \{a \nabla u \cdot n\}_e \cdot [v]_e - \int_e \{a \nabla v \cdot n\}_e \cdot [u]_e + \frac{\gamma}{h} \int_e a [u]_e \cdot [v]_e \right) \right| \\
&\leq I_1 + I_2 + I_3 + I_4,
\end{aligned}$$

where

$$\begin{aligned}
I_1 &= \sum_{K \in \mathcal{T}^H} \left| \int_K a \nabla u \cdot \nabla v \right|, & I_2 &= \sum_{e \in \mathcal{E}^H} \left| \int_e \{a \nabla u \cdot n\}_e \cdot [v]_e \right|, \\
I_3 &= \sum_{e \in \mathcal{E}^H} \left| \int_e \{a \nabla v \cdot n\}_e \cdot [u]_e \right|, & I_4 &= \sum_{e \in \mathcal{E}^H} \frac{\gamma}{h} \int_e |a [u]_e \cdot [v]_e|.
\end{aligned}$$

First, we note that I_1 and I_4 can be estimated easily as follows:

$$\begin{aligned}
I_1 &\leq \left(\sum_{K \in \mathcal{T}^H} \int_K a |\nabla u|^2 \right)^{\frac{1}{2}} \left(\sum_{K \in \mathcal{T}^H} \int_K a |\nabla v|^2 \right)^{\frac{1}{2}}, \\
I_4 &\leq \left(\sum_{e \in \mathcal{E}^H} \frac{\gamma}{h} \int_e a [u]_e^2 \right)^{\frac{1}{2}} \left(\sum_{e \in \mathcal{E}^H} \frac{\gamma}{h} \int_e a [v]_e^2 \right)^{\frac{1}{2}}.
\end{aligned}$$

For I_2 , we can estimate as follows

$$\begin{aligned}
I_2 &\leq \left(\frac{h}{\gamma} \sum_{e \in \mathcal{E}^H} \int_e \{a \nabla u \cdot n\}_e^2 \right)^{\frac{1}{2}} \left(\sum_{e \in \mathcal{E}^H} \frac{\gamma}{h} \int_e [v]_e^2 \right)^{\frac{1}{2}} \\
&\leq \left(\frac{h}{\gamma a_0} \sum_{K \in \mathcal{T}^H} \int_{\partial K} (a \nabla u \cdot n)^2 \right)^{\frac{1}{2}} \left(\sum_{e \in \mathcal{E}^H} \frac{\gamma}{h} \int_e a [v]_e^2 \right)^{\frac{1}{2}} \\
&\leq \left(\frac{\Lambda a_1}{\gamma a_0} \right)^{\frac{1}{2}} |u|_a \left(\sum_{e \in \mathcal{E}^H} \frac{\gamma}{h} \int_e [v]_e^2 \right)^{\frac{1}{2}}.
\end{aligned}$$

The same idea can be applied to estimate I_3 to obtain

$$I_3 \leq \left(\frac{\Lambda a_1}{\gamma a_0} \right)^{\frac{1}{2}} |v|_a \left(\sum_{e \in \mathcal{E}^H} \frac{\gamma}{h} \int_e a [u]_e^2 \right)^{\frac{1}{2}}.$$

Finally, combining the above estimates, we have

$$|a_{DG}(u, v)| \leq 2 \|u\|_a \|v\|_a$$

when $\gamma \geq \Lambda a_1 a_0^{-1}$.

□

Next, we will prove the convergence of the semi-discrete scheme (4). First, we define the following error quantities. Let

$$\eta = u_h - w_H, \quad \xi = u_H - w_H, \quad \text{and} \quad \varepsilon = u_h - u_H, \tag{17}$$

where $w_H \in V_H$ is defined by solving the following elliptic projection problem

$$a_{DG}(w_H, v) = a_{DG}(u_h, v) + R_{u_h}(v), \quad \forall v \in V_H. \quad (18)$$

Notice that ε is the difference between the multiscale solution u_H and the fine grid finite element solution u_h . Moreover, η measures the difference between the fine grid solution u_h as its projection w_H . In the following, we will prove estimates for ε . First, we let

$$\|\varepsilon\|_{L^\infty([0,T];L^2(\Omega))} = \max_{0 \leq t \leq T} \|\varepsilon\|_{L^2(\Omega)} \quad \text{and} \quad \|\varepsilon\|_{L^\infty([0,T];a)} = \max_{0 \leq t \leq T} \|\varepsilon\|_a.$$

Then we will prove the following two inequalities, which estimate the error for the solution ε by the error for the projection η and the initial errors \mathcal{I}_1 and \mathcal{I}_2 , which are defined in the statements of the theorems.

Theorem 1 *Let ε, η and ξ be the error quantities defined in (17). Then we have the following error bound*

$$\begin{aligned} & \|\varepsilon_t\|_{L^\infty([0,T];L^2(\Omega))} + \|\varepsilon\|_{L^\infty([0,T];a)} \\ & \leq C \left(\|\eta_t\|_{L^\infty([0,T];L^2(\Omega))} + \|\eta\|_{L^\infty([0,T];H^1(\mathcal{T}^H))} + \|\eta_{tt}\|_{L^1([0,T];L^2(\Omega))} + \mathcal{I}_1 \right), \end{aligned} \quad (19)$$

where $\mathcal{I}_1 = \|\xi_t(0)\|_{L^2(\Omega)} + \|\xi(0)\|_{H^1(\mathcal{T}^H)}$.

Proof: First, using (4) and the definition of ξ , we have

$$(\xi_{tt}, v) + a_{DG}(\xi, v) = (f, v) - ((w_H)_{tt}, v) - a_{DG}(w_H, v).$$

Then by (12), we have

$$(\xi_{tt}, v) + a_{DG}(\xi, v) = (\eta_{tt}, v). \quad (20)$$

Taking $v = \xi_t$ in (20), we have

$$(\xi_{tt}, \xi_t) + a_{DG}(\xi, \xi_t) = (\eta_{tt}, \xi_t),$$

which implies

$$\frac{1}{2} \frac{d}{dt} \left(\|\xi_t\|_{L^2(\Omega)}^2 + a_{DG}(\xi, \xi) \right) \leq \|\eta_{tt}\|_{L^2(\Omega)} \|\xi_t\|_{L^2(\Omega)}.$$

Integrating from $t = 0$ to $t = \tau$, we have

$$\begin{aligned} \|\xi_t(\tau)\|_{L^2(\Omega)}^2 + \frac{1}{2} \|\xi(\tau)\|_a^2 & \leq \|\xi_t(0)\|_{L^2(\Omega)}^2 + 2\|\xi(0)\|_a^2 + 2 \int_0^\tau \|\eta_{tt}\|_{L^2(\Omega)} \|\xi_t\|_{L^2(\Omega)} \\ & \leq \|\xi_t(0)\|_{L^2(\Omega)}^2 + 2\|\xi(0)\|_a^2 + 2 \max_{0 \leq t \leq T} \|\xi_t\|_{L^2(\Omega)} \int_0^\tau \|\eta_{tt}\|_{L^2(\Omega)}. \end{aligned}$$

Therefore, we obtain

$$\|\xi_t\|_{L^\infty([0,T];L^2(\Omega))}^2 + \|\xi\|_{L^\infty([0,T];a)}^2 \leq C \left(\|\xi_t(0)\|_{L^2(\Omega)}^2 + \|\xi(0)\|_a^2 + \left(\int_0^T \|\eta_{tt}\|_{L^2(\Omega)} dt \right)^2 \right).$$

Finally, (19) is proved by noting that $\varepsilon = \eta - \xi$. □

Theorem 2 *Let ε, η and ξ be the error quantities defined in (17). Then we have the following error bound*

$$\|\varepsilon\|_{L^\infty([0,T];L^2(\Omega))} \leq C \left(\|\eta_t\|_{L^1([0,T],L^2(\Omega))} + \|\eta\|_{L^\infty([0,T];L^2(\Omega))} + \mathcal{I}_2 \right), \quad (21)$$

where $\mathcal{I}_2 = \|\xi(0)\|_{L^2(\Omega)}$.

Proof: Integrating by parts with respect to time in (20), we have

$$-(\xi_t, v_t) + \partial_t(\xi_t, v) + a_{DG}(\xi, v) = \partial_t(\eta_t, v) - (\eta_t, v_t).$$

Taking $v(x, t) = \int_t^\gamma \xi(x, \tau) d\tau$, we have $v_t = -\xi$ and $v(\gamma) = 0$. So,

$$(\xi_t, \xi) - \partial_t(\xi_t, v) - a_{DG}(v_t, v) = \partial_t(\eta_t, v) + (\eta_t, \xi),$$

which implies that

$$\frac{1}{2} \frac{d}{dt} \|\xi\|_{L^2(\Omega)}^2 - \partial_t(\xi_t, v) - \frac{1}{2} \frac{d}{dt} a_{DG}(v, v) = \partial_t(\eta_t, v) + (\eta_t, \xi).$$

Integrating from $t = 0$ to $t = \gamma$, we have

$$\frac{1}{2} \|\xi(\gamma)\|_{L^2(\Omega)}^2 - \frac{1}{2} \|\xi(0)\|_{L^2(\Omega)}^2 + (\xi_t(0), v(0)) + \frac{1}{2} a_{DG}(v(0), v(0)) = (\eta_t(0), v(0)) + \int_0^\gamma (\eta_t, \xi).$$

Since $\xi_t - \eta_t = (u_H - u_h)_t$, we obtain

$$(\xi_t(0) - \eta_t(0), v(0)) = ((u_H - u_h)_t(0), v(0)) = 0.$$

Using the coercivity of a_{DG} , we have

$$\begin{aligned} \|\xi(\gamma)\|_{L^2(\Omega)}^2 &\leq \|\xi(0)\|_{L^2(\Omega)}^2 + 2 \int_0^\gamma \|\eta_t\|_{L^2(\Omega)} \|\xi\|_{L^2(\Omega)} \\ &\leq \|\xi(0)\|_{L^2(\Omega)}^2 + 2 \max_{0 \leq t \leq T} \|\xi\|_{L^2(\Omega)} \int_0^T \|\eta_t\|_{L^2(\Omega)}. \end{aligned}$$

Hence (21) is proved by noting that $\varepsilon = \eta - \xi$. □

From Theorem 1 and Theorem 2, we see that, in order to estimate the error $\varepsilon = u_h - u_H$, we will need to find a bound for η given that the initial values $\xi_t(0)$ and $\xi(0)$ are sufficiently accurate.

4.2 Convergence analysis

In this section, we will derive an error bound for $\eta = u_h - w_H$. Notice that, on each coarse grid block K , we can express u_h as

$$u_h = \sum_{i=1}^n c_{i,K} \tilde{w}_{i,K} + \sum_{i=1}^{n_0} d_{i,K} z_{i,K} = u_{1,K} + u_{2,K}$$

for some suitable coefficients $c_{i,K}$ and $d_{i,K}$ determined by a L^2 -type projection, where n_0 is the dimension of $V_h^0(K)$. We write $u_h = u_1 + u_2$ with $u_i|_K = u_{i,K}$ for $i = 1, 2$. Moreover, we recall that $C(u, f)$, defined in (34), is the constant appearing in the consistency error estimate in Lemma 1. In the following theorem, we will give an estimate for the difference between the fine grid solution u_h and the projection of u_h into the coarse space V_H defined in (18). The theorem says that such difference is bounded by a best approximation error $\|u_h - v\|_a$ and a consistency error $hC(u, f)$. We emphasize that, even though the coarse mesh size H is fixed, but the fine mesh size h can be arbitrary small, and hence the consistency error is small compared with the best approximation error $\|u_h - v\|_a$.

Theorem 3 *Let $w_H \in V_H$ be the solution of (18) and u_h be the solution of (2). Then we have*

$$\|u_h - w_H\|_a \leq C(\|u_h - v\|_a + hC(u, f)), \quad \forall v \in V_H. \quad (22)$$

Proof: By the definition of w_H , we have

$$a_{DG}(w_H, v) = a_{DG}(u_h, v) + R_{u_h}(v), \quad \forall v \in V_H.$$

So, we have

$$a_{DG}(w_H - v, w_H - v) = a_{DG}(u_h - v, w_H - v) + R_{u_h}(w_H - v).$$

By (13), (14) and (12), we get

$$\begin{aligned} \|w_H - v\|_a^2 &\leq 2a_{DG}(w_H - v, w_H - v) \\ &= 2a_{DG}(u_h - v, w_H - v) + 2R_{u_h}(w_H - v) \\ &\leq C(\|u_h - v\|_a + hC(u, f))\|w_H - v\|_a. \end{aligned}$$

Finally, we obtain

$$\begin{aligned} \|u_h - w_H\|_a &\leq \|u_h - v\|_a + \|w_H - v\|_a \\ &\leq C(\|u_h - v\|_a + hC(u, f)). \end{aligned}$$

□

From the above theorem, we see that the error $\|u_h - w_H\|_a$ is controlled by the quantity $\|u_h - v\|_a$ for an arbitrary choice of the function $v \in V_H$. Thus, to obtain our final error bound, we only need to find a suitable function $v \in V_H$ to approximate the finite element solution u_h . In the following theorem, we will choose a specific v in Theorem 3 and prove the corresponding error estimate.

Theorem 4 *Let $u_h \in V_h$ be the finite element solution. Then we have*

$$\|u_h - \phi\|_a^2 \leq \sum_{K \in \mathcal{T}^H} \left(\frac{H}{\mu_{p+1,K}} \left(1 + \frac{2a_1\gamma H}{h\mu_{p+1,K}} \right) \int_{\partial K} \left(a \frac{\partial u_1}{\partial n} \right)^2 + \frac{H^2}{\lambda_{m+1,K}} \|f - u_{tt}\|_{L^2(K)}^2 \right), \quad (23)$$

where the function $\phi \in V_H$ is defined as

$$\phi|_K = \sum_{i=1}^p c_{i,K} \tilde{w}_{i,K} + \sum_{i=1}^m d_{i,K} z_{i,K} = \phi_{1,K} + \phi_{2,K}.$$

Proof: For a given coarse grid block K , using the orthogonality condition (9), we have

$$\int_K a |\nabla(u_h - \phi)|^2 = \int_K a |\nabla(u_1 - \phi_1)|^2 + \int_K a |\nabla(u_2 - \phi_2)|^2$$

which implies

$$\|u_h - \phi\|_a^2 = \|u_1 - \phi_1\|_a^2 + \|u_2 - \phi_2\|_a^2,$$

where we write $\phi = \phi_1 + \phi_2$ and $\phi_i|_K = \phi_{i,K}$, for $i = 1, 2$. We will first estimate $\|u_1 - \phi_1\|_a^2$. By the definition of a -norm, we have

$$\begin{aligned} \|u_1 - \phi_1\|_a^2 &= \sum_{K \in \mathcal{T}^H} \left(\int_K a |\nabla(u_1 - \phi_1)|^2 + \sum_{e \in \mathcal{E}^H} \frac{\gamma}{h} \int_e a |[(u_1 - \phi_1)]_e|^2 \right) \\ &\leq \sum_{K \in \mathcal{T}^H} \left(\int_K a |\nabla(u_1 - \phi_1)|^2 + \frac{2\gamma}{h} \int_{\partial K} a |(u_1 - \phi_1)|^2 \right) \\ &\leq \sum_{K \in \mathcal{T}^H} \left(\int_K a |\nabla(u_1 - \phi_1)|^2 + \frac{2a_1\gamma}{h} \int_{\partial K} |(u_1 - \phi_1)|^2 \right). \end{aligned} \quad (24)$$

Next, we will estimate the right hand side of (24) for each K .

We note that the eigenvalue problem (7) is motivated by the right hand side of (24). In particular, based on the right hand side of (24), we consider

$$\int_K a \nabla w_\mu \cdot \nabla v + \frac{1}{H} \int_{\partial K} w_\mu v = \widehat{\mu} \int_K R(w_\mu) \cdot R(v), \quad \forall v \in V_H^1(K), \quad (25)$$

where the choice of R , e.g., $R = \sqrt{a} \nabla w_\mu$, depends on how we would like to bound the error. Indeed, choosing the eigenvectors that correspond to the largest L_K eigenvalues, one can guarantee that the best L_K dimensional space *in the space of snapshots* is given by the first L_K dominant eigenvectors. The choice of $R(\cdot)$ is important and can influence the eigenvalue behavior. For example, the use of oversampling domains both for the snapshot space and the eigenvalue can provide a faster convergence. In this paper, we take

$$R = \sqrt{a} \nabla w_\mu,$$

which allows estimating the right hand side of (24) by the energy norm. Note that, in (7), we use the smallest eigenvalues to determine the basis functions which is the same as choosing the largest eigenvectors that correspond to the largest eigenvalues of (25) because $\widehat{\mu} = 1 + \frac{1}{\mu}$.

Note that the eigenvalue problem (7) is equivalent to

$$a \frac{\partial w_\mu}{\partial n} = \frac{\mu}{H} w_\mu \quad \text{on} \quad \partial K.$$

So, for each K ,

$$\int_{\partial K} \left(a \frac{\partial u_{1,K}}{\partial n} \right)^2 = \int_{\partial K} \left(a \frac{\partial}{\partial n} \left(\sum_{i=1}^n c_{i,K} w_{i,K} \right) \right)^2 = \int_{\partial K} \left(\sum_{i=1}^n \frac{\mu_{i,K}}{H} c_{i,K} w_{i,K} \right)^2 = \sum_{i=1}^n \left(\frac{\mu_{i,K} c_{i,K}}{H} \right)^2, \quad (26)$$

where we have used the fact that $\int_{\partial K} w_{i,K} w_{j,K} = \delta_{ij}$. Then, by using the eigenvalue problem defined in (7), we have

$$\frac{1}{h} \int_{\partial K} |(u_{1,K} - \phi_{1,K})|^2 = \frac{1}{h} \sum_{i=p+1}^{4n} c_{i,K}^2 \leq \frac{H^2}{h \mu_{p+1,K}^2} \sum_{i=p+1}^{4n} \left(\frac{\mu_{i,K}}{H} \right)^2 c_{i,K}^2$$

and

$$\int_K a |\nabla (u_{1,K} - \phi_{1,K})|^2 = \sum_{i=p+1}^{4n} \frac{\mu_{i,K}}{H} c_{i,K}^2 \leq \frac{H}{\mu_{p+1,K}} \sum_{i=p+1}^{4n} \left(\frac{\mu_{i,K}}{H} \right)^2 c_{i,K}^2.$$

Note that, by using (26), we have,

$$\sum_{i=p+1}^{4n} \left(\frac{\mu_{i,K}}{H} \right)^2 c_{i,K}^2 \leq \sum_{i=1}^{4n} \left(\frac{\mu_{i,K}}{H} \right)^2 c_{i,K}^2 = \int_{\partial K} \left(a \frac{\partial u_{1,K}}{\partial n} \right)^2.$$

Therefore

$$\begin{aligned} \|u_1 - \phi_1\|_a^2 &\leq \sum_{K \in \mathcal{T}^H} \left(\frac{H}{\mu_{p+1,K}} \left(1 + \frac{2a_1 \gamma H}{h \mu_{p+1,K}} \right) \sum_{i=p+1}^{4n} \left(\frac{\mu_{i,K}}{H} \right)^2 c_{i,K}^2 \right) \\ &\leq \sum_{K \in \mathcal{T}^H} \left(\frac{H}{\mu_{p+1,K}} \left(1 + \frac{2a_1 \gamma H}{h \mu_{p+1,K}} \right) \int_{\partial K} \left(a \frac{\partial u_1}{\partial n} \right)^2 \right). \end{aligned} \quad (27)$$

Next, we will estimate $|u_2 - \phi_2|_a^2$. Since u_h satisfies

$$\int_K a \nabla u_h \cdot \nabla v = \int_K (f - (u_h)_{tt}) v, \quad \forall v \in V_h^0(K).$$

Putting $v = z_{i,K}$, we obtain

$$\frac{\lambda_{i,K}}{H^2} d_{i,K} = \int_K a \nabla u_h \cdot \nabla z_{i,K} = \int_K (f - (u_h)_{tt}) z_{i,K}.$$

We define $f_{i,K} = \int_K (f - (u_h)_{tt}) z_{i,K}$. Then we have $f_{i,K} = \frac{\lambda_{i,K}}{H^2} d_{i,K}$ and

$$\sum_{i=1}^{n_0} f_{i,K}^2 \leq \|f - (u_h)_{tt}\|_{L^2(K)}^2.$$

Hence,

$$\begin{aligned} |u_2 - \phi_2|_a^2 &= \sum_{K \in \mathcal{T}^H} \int_K a |\nabla(u_2 - \phi_2)|^2 \\ &= \sum_{K \in \mathcal{T}^H} \sum_{i \geq m+1} \frac{\lambda_{i,K}}{H^2} d_{i,K}^2 \\ &\leq \sum_{K \in \mathcal{T}^H} \frac{H^2}{\lambda_{m+1,K}} \sum_{i \geq m+1} \frac{\lambda_{i,K}^2}{H^4} d_{i,K}^2 \\ &= \sum_{K \in \mathcal{T}^H} \frac{H^2}{\lambda_{m+1,K}} \sum_{i \geq m+1} f_{i,K}^2 \\ &\leq \sum_{K \in \mathcal{T}^H} \frac{H^2}{\lambda_{m+1,K}} \|f - (u_h)_{tt}\|_{L^2(K)}^2. \end{aligned}$$

□

We note that, by the technique in [14], we can also derive a bound for $\|u_1 - \phi_1\|_a$ as follows

$$\|u_1 - \phi_1\|_a^2 \leq \sum_{K \in \mathcal{T}^H} \sum_{i \geq p+1} c_{i,K}^2.$$

This bound shows the decay of the error when more basis functions are used.

The bound in (23) gives the spectral convergence of our GMsFEM. Notice that, the term

$$H \sum_{K \in \mathcal{T}^H} \sum_{\partial K} \left(a \frac{\partial u_1}{\partial n} \right)^2$$

is uniformly bounded and can be considered as a norm for u_1 . Thus, (23) states that the error behaves as $O(\mu_{p+1,K}^{-1} + \lambda_{m+1,K}^{-1})$. We note that the eigenvalues increase (and go to the infinity as the fine mesh size decreases) and thus the error decreases as we increase the coarse space dimension.

Combining the results in Theorem 3 and Theorem 4, we obtain

$$\|\eta\|_a^2 \leq C \sum_{K \in \mathcal{T}^H} \left(\frac{H^2}{\lambda_{m+1,K}} \|f - u_{tt}\|_{L^2(K)}^2 + \frac{H}{\mu_{p+1,K}} \left(1 + \frac{2a_1 \gamma H}{h \mu_{p+1,K}} \right) \int_{\partial K} \left(a \frac{\partial u_1}{\partial n} \right)^2 \right) + h^2 C(u, f)^2.$$

Similarly, we obtain

$$\|\eta_t\|_a^2 \leq C \sum_{K \in \mathcal{T}^H} \left(\frac{H^2}{\lambda_{m+1,K}} \|f_t - u_{ttt}\|_{L^2(K)}^2 + \frac{H}{\mu_{p+1,K}} \left(1 + \frac{2a_1 \gamma H}{h \mu_{p+1,K}} \right) \int_{\partial K} \left(a \frac{\partial (u_1)_t}{\partial n} \right)^2 \right) + h^2 C(u_t, f_t)^2.$$

Finally, using these bounds for η , as well as the estimates proved in Theorem 1 and Theorem 2, we obtain estimates for the error ε .

5 Fully discretization

In this section, we will prove the convergence of the fully discrete scheme (5). To simplify the notations, we define the second order central difference operator δ^2 by

$$\delta^2(u^n) = \frac{u^{n+1} - 2u^n + u^{n-1}}{\Delta t^2}.$$

From the semi-discrete scheme (4), we have

$$((u_h)_{tt}^n, v) + a_{DG}(u_h^n, v) = (f^n, v) - R_{u_h^n}(v)$$

and the fully discrete scheme (5) can be written as

$$(\delta^2(u_H^n), v) + a_{DG}(u_H^n, v) = (f^n, v), \quad \text{for } n \geq 1.$$

Moreover, we define

$$r^n = \begin{cases} u_{tt}^n - \delta^2(u_H^n), & \text{for } n \geq 1, \\ \Delta t^{-2}(\xi^1 - \xi^0), & \text{for } n = 0, \end{cases} \quad (28)$$

and

$$R^n = \Delta t \sum_{i=0}^n r^i.$$

In order to prove the convergence for the fully discrete scheme, we first prove the following lemma. The result will be needed in the derivation of an upper bound for the time step size Δt .

Lemma 3 *There exists a positive constant $\beta(h)$ such that*

$$a_{DG}(v, v) \leq \beta(h)^{-1} \|v\|_{L^2(\Omega)}^2, \quad \forall v \in V_H.$$

Moreover, the constant $\beta(h)$ can be taken as $h^2 a_1^{-1} (24 + 32\sqrt{3\Lambda} + 16\gamma)^{-1}$.

Proof: We first note that, if p is a linear function defined on the interval $I = [x_1 - h/2, x_1 + h/2]$, then we have

$$\|p\|_{L^\infty(I)}^2 \leq \frac{4}{h} \|p\|_{L^2(I)}^2 \quad (29)$$

$$|p|_{H^1(I)}^2 \leq \frac{12}{h^2} \|p\|_{L^2(I)}^2. \quad (30)$$

Then by the definition of a_{DG} and the Cauchy-Schwarz inequality, we have

$$\begin{aligned} & a_{DG}(v, v) \\ & \leq \sum_{K \in \mathcal{T}^H} \int_K a |\nabla v|^2 - 2 \sum_{e \in \mathcal{E}^H} \left(\int_e \{a \nabla v \cdot n\}_e \cdot [v]_e + \frac{\gamma}{h} \int_e a [v]_e^2 \right) \\ & \leq \sum_{K \in \mathcal{T}^H} \int_K a |\nabla v|^2 + 2 \left(\sum_{K \in \mathcal{T}^H} h \|a \nabla v \cdot n_{\partial K}\|_{L^2(\partial K)}^2 \right)^{\frac{1}{2}} \left(\sum_{e \in \mathcal{E}^H} h^{-1} \|[v]\|_{L^2(e)}^2 \right)^{\frac{1}{2}} + \frac{\gamma}{h} \sum_{e \in \mathcal{E}^H} \int_e a [v]_e^2. \end{aligned}$$

Then by using (16), $a \leq a_1$ and estimating the jump terms by $L^2(\partial K)$ norms, we have

$$\begin{aligned} & a_{DG}(v, v) \\ & \leq a_1 \left(\sum_{K \in \mathcal{T}^H} \int_K |\nabla v|^2 + 2 \left(\sum_{K \in \mathcal{T}^H} \Lambda \int_K |\nabla v|^2 \right)^{\frac{1}{2}} \left(\sum_{e \in \mathcal{E}^H} h^{-1} \|[v]\|_{L^2(e)}^2 \right)^{\frac{1}{2}} + \frac{\gamma}{h} \sum_{e \in \mathcal{E}^H} \int_e [v]_e^2 \right) \\ & \leq a_1 \left(\sum_{K \in \mathcal{T}^H} \int_K |\nabla v|^2 + 4 \left(\sum_{K \in \mathcal{T}^H} \Lambda \int_K |\nabla v|^2 \right)^{\frac{1}{2}} \left(\sum_{K \in \mathcal{T}^H} h^{-1} \|v\|_{L^2(\partial K)}^2 \right)^{\frac{1}{2}} + \frac{2\gamma}{h} \sum_{K \in \mathcal{T}^H} \|v\|_{L^2(\partial K)}^2 \right). \end{aligned} \quad (31)$$

Thus, it remains to estimate $\|\nabla v\|_{L^2(K)}$ and $\|v\|_{L^2(\partial K)}$.

We will estimate the term $\|\nabla v\|_{L^2(K)}$ first. For a given coarse grid block K , we can write it as the union of fine grid blocks $K = \cup_{F \subset K} F$, where we use F to represent a generic fine grid block. Since the fine grid blocks are rectangles, we can write F as a tensor product of two intervals, namely, $F = I_x^F \times I_y^F$. For any $v \in V_H$ we can also write the restriction of v on F as $v(x, y) = v_{F,1}(x)v_{F,2}(y)$.

$$\begin{aligned} \int_K |\nabla v|^2 &= \sum_{F \subset K} \int_F |\nabla v|^2 \\ &= \sum_{F \subset K} \left(h(v'_{F,2})^2 \int_{I_x^F} (v_{F,1}(x))^2 + h(v'_{F,1})^2 \int_{I_y^F} (v_{F,2}(y))^2 \right) \\ &= \sum_{F \subset K} \left(\int_{I_y^F} (v'_{F,2}(y))^2 \int_{I_x^F} (v_{F,1}(x))^2 + \int_{I_x^F} (v'_{F,1}(x))^2 \int_{I_y^F} (v_{F,2}(y))^2 \right). \end{aligned}$$

Then, using (30), we have

$$\begin{aligned} \int_K |\nabla v|^2 &\leq 12h^{-2} \sum_{F \subset K} \left(\int_{I_y^F} (v_{F,2}(y))^2 \int_{I_x^F} (v_{F,1}(x))^2 + \int_{I_x^F} (v_{F,1}(x))^2 \int_{I_y^F} (v_{F,2}(y))^2 \right) \\ &= 24h^{-2} \sum_{F \subset K} \int_F |v|^2. \end{aligned}$$

Next, we estimate the term $\|v\|_{L^2(\partial K)}$. For a generic fine grid cell F , we write $I_x^F = [x_1, x_2]$ and $I_y^F = [y_1, y_2]$. Then, by using (29),

$$\begin{aligned} \|v\|_{L^2(\partial K)}^2 &= \sum_{F \subset K} \int_{\partial F \cap \partial K} (v_F)^2 \\ &= \sum_{F \subset K} \left(\int_{\partial F \cap (I_x \times \{y_1\})} (v_{F,2}(y_1)v_{F,1}(x))^2 + \int_{\partial F \cap (I_x \times \{y_2\})} (v_{F,2}(y_2)v_{F,1}(x))^2 \right) \\ &\quad + \sum_{F \subset K} \left(\int_{\partial F \cap (\{x_1\} \times I_y)} (v_{F,1}(x_1)v_{F,2}(y))^2 + \int_{\partial F \cap (\{x_2\} \times I_y)} (v_{F,1}(x_2)v_{F,2}(y))^2 \right) \\ &\leq \frac{4}{h} \sum_{F \subset K} \left(\int_{\partial F \cap (I_x \times \{y_1\})} \int_{[y_1, y_1+h]} v_{F,2}(y)^2 (v_{F,1}(x))^2 + \int_{\partial F \cap (I_x \times \{y_2\})} \int_{[y_2-h, y_2]} v_{F,2}(y)^2 (v_{F,1}(x))^2 \right) \\ &\quad + \frac{4}{h} \sum_{F \subset K} \left(\int_{\partial F \cap (\{x_1\} \times I_y)} \int_{[x_1, x_1+h]} v_{F,1}(x)^2 v_{F,2}(y)^2 + \int_{\partial F \cap (\{x_2\} \times I_y)} \int_{[x_2-h, x_2]} v_{F,1}(x)^2 v_{F,2}(y)^2 \right) \\ &\leq \frac{4}{h} \sum_{F \subset K} \left(\int_F (v_{F,2}(y)v_{F,1}(x))^2 + \int_F (v_{F,1}(x)v_{F,2}(y))^2 \right) \\ &= \frac{8}{h} \|v\|_{L^2(K)}^2. \end{aligned}$$

Consequently, combining the above results and using (31),

$$a_{DG}(v, v) \leq \frac{a_1}{h^2} \left(24 + 32\sqrt{3\Lambda} + 16\gamma \right) \|v\|_{L^2(\Omega)}^2.$$

□

Finally, we will state and prove the convergence of the fully discrete scheme (5).

Theorem 5 Assume that the time step size Δt satisfies the stability condition $\Delta t^2 < 4\beta(h)$. We have

$$\max_{0 \leq n \leq N} \|\varepsilon^n\|_{L^2(\Omega)} \leq C \left(\|\varepsilon^0\|_{L^2(\Omega)} + \max_{0 \leq n \leq N} \|\eta^n\|_{L^2(\Omega)} + \Delta t \sum_{n=0}^N \|R^n\|_{L^2(\Omega)} \right). \quad (32)$$

Proof: Notice that

$$(\delta^2(u_H^n - w_H^n + w_H^n - u_h^n), v) + a_{DG}(u_H^n - u_h^n, v) = ((u_h)_{tt}^n - \delta^2(u_h^n), v) + R_{u_h^n}(v).$$

Using the definitions of ξ and w_H , we have

$$(\delta^2(\xi^n), v) + a_{DG}(\xi^n, v) = (r^n, v), \quad \text{for } n \geq 1.$$

So we have

$$\left(\frac{\xi^{n+1} - \xi^n}{\Delta t}, v \right) - \left(\frac{\xi^n - \xi^{n-1}}{\Delta t}, v \right) + \Delta t a_{DG}(\xi^n, v) = \Delta t (r^n, v).$$

Summing up, for $n \geq 1$,

$$\left(\frac{\xi^{n+1} - \xi^n}{\Delta t}, v \right) - \left(\frac{\xi^1 - \xi^0}{\Delta t}, v \right) + \Delta t \sum_{i=1}^n a_{DG}(\xi^i, v) = \Delta t \sum_{i=1}^n (r^i, v).$$

To simplify the notations, we define

$$\Xi^n = \Delta t \sum_{i=1}^n \xi^i, \quad \text{for } n \geq 1; \quad \text{and } \Xi^0 = 0.$$

Then we get

$$\left(\frac{\xi^{n+1} - \xi^n}{\Delta t}, v \right) + a_{DG}(\Xi^n, v) = (R^n, v), \quad n \geq 1.$$

Substituting $v = \xi^{n+1} + \xi^n$, we have

$$\|\xi^{n+1}\|_{L^2(\Omega)}^2 - \|\xi^n\|_{L^2(\Omega)}^2 + \Delta t a_{DG}(\Xi^n, \xi^{n+1} + \xi^n) = \Delta t (R^n, \xi^{n+1} + \xi^n),$$

and summing for all $n \geq 1$, we have

$$\|\xi^{n+1}\|_{L^2(\Omega)}^2 - \|\xi^1\|_{L^2(\Omega)}^2 + \Delta t \sum_{i=1}^n a_{DG}(\Xi^i, \xi^{i+1} + \xi^i) = \Delta t \sum_{i=1}^n (R^i, \xi^{i+1} + \xi^i).$$

Notice that we have $\Xi^{n+1} - \Xi^{n-1} = \Delta t(\xi^{n+1} + \xi^n)$ for $n \geq 1$. So

$$\begin{aligned} \Delta t \sum_{i=1}^n \tilde{a}_{DG}(\Xi^i, \xi^{i+1} + \xi^i) &= \sum_{i=1}^n a_{DG}(\Xi^i, \Xi^{i+1} - \Xi^{i-1}) \\ &= \sum_{i=1}^n a_{DG}(\Xi^i, \Xi^{i+1}) - \sum_{i=0}^{n-1} a_{DG}(\Xi^i, \Xi^{i+1}) \\ &= a_{DG}(\Xi^n, \Xi^{n+1}). \end{aligned}$$

Moreover,

$$\begin{aligned} a_{DG}(\Xi^n, \Xi^{n+1}) &= a_{DG}\left(\frac{\Xi^n + \Xi^{n+1}}{2}, \frac{\Xi^n + \Xi^{n+1}}{2}\right) - a_{DG}\left(\frac{\Xi^n - \Xi^{n+1}}{2}, \frac{\Xi^n - \Xi^{n+1}}{2}\right) \\ &\geq -\frac{\Delta t^2}{4} a_{DG}(\xi^{n+1}, \xi^{n+1}). \end{aligned}$$

So we have

$$\|\xi^{n+1}\|_{L^2(\Omega)}^2 - \frac{\Delta t^2}{4} a_{DG}(\xi^{n+1}, \xi^{n+1}) \leq \|\xi^1\|_{L^2(\Omega)}^2 + \Delta t \sum_{i=1}^n (R^i, \xi^{i+1} + \xi^i), \quad n \geq 1.$$

Using the assumption $\Delta t^2 < 4\beta(h)$, we have $C_s = 1 - \frac{\Delta t^2}{4\beta(h)} > 0$. Therefore,

$$\begin{aligned} C_s \|\xi^{n+1}\|_{L^2(\Omega)}^2 &\leq \|\xi^1\|_{L^2(\Omega)}^2 + \Delta t \sum_{i=1}^n (R^i, \xi^{i+1} + \xi^i) \\ &\leq \|\xi^1\|_{L^2(\Omega)}^2 + 2\Delta t \max_{1 \leq i \leq n+1} \{\|\xi^i\|_{L^2(\Omega)}\} \sum_{i=1}^n \|R^i\|_{L^2(\Omega)} \\ &\leq \|\xi^1\|_{L^2(\Omega)}^2 + \frac{C_s}{2} \max_{1 \leq i \leq n+1} \{\|\xi^i\|_{L^2(\Omega)}\}^2 + \frac{2}{C_s} \left(\Delta t \sum_{i=1}^n \|R^i\|_{L^2(\Omega)} \right)^2. \end{aligned}$$

Then

$$\max_{1 \leq i \leq n+1} \{\|\xi^i\|_{L^2(\Omega)}\} \leq \sqrt{\frac{2}{C_s}} \|\xi^1\|_{L^2(\Omega)} + \frac{2}{C_s} \Delta t \sum_{i=1}^n \|R^i\|_{L^2(\Omega)}.$$

Since $\xi^1 = \xi^0 + \Delta t^2 r^0$, we have

$$\max_{1 \leq i \leq n+1} \{\|\xi^i\|_{L^2(\Omega)}\} \leq C \left(\|\xi^0\|_{L^2(\Omega)} + \Delta t^2 \|r^0\|_{L^2(\Omega)} + \Delta t \sum_{i=1}^n \|R^i\|_{L^2(\Omega)} \right)$$

and using the definition of R^0 ,

$$\max_{1 \leq i \leq n+1} \{\|\xi^i\|_{L^2(\Omega)}\} \leq C \left(\|\xi^0\|_{L^2(\Omega)} + \Delta t \sum_{i=0}^n \|R^i\|_{L^2(\Omega)} \right).$$

Thus,

$$\max_{0 \leq i \leq n+1} \{\|\xi^i\|_{L^2(\Omega)}\} \leq C \left(\|\xi^0\|_{L^2(\Omega)} + \Delta t \sum_{i=0}^n \|R^i\|_{L^2(\Omega)} \right).$$

Finally, by using the relation $\varepsilon = \eta - \xi$, we obtain (32). □

Next, we will estimate the right hand side of (32). To do so, we prove the following lemmas.

Lemma 4 *We have*

$$\|r^0\|_{L^2(\Omega)} \leq C(\Delta t^{-1} \|\eta_t\|_{L^\infty([0,T];L^2(\Omega))} + \Delta t \|(u_h)_{ttt}\|_{C([0,T];L^2(\Omega))}).$$

Proof: By (28), we have $r^0 = \Delta t^{-2}(\xi^1 - \xi^0)$ and by the definition of u_H^0 , we have

$$(u_H^0 - u^0, v) = 0, \quad \forall v \in V_H.$$

Then using the definitions of ξ^1 and ξ^0 , we have

$$\begin{aligned} (\xi^1 - \xi^0, v) &= (u_H^1 - w_H^1, v) - (u_H^0 - w_H^0, v) \\ &= (u_h^1 - w_H^1, v) + (u_H^1 - u_h^1, v) - (u_h^0 - w_H^0, v) \\ &= ((u_h^1 - u_h^0) - (w_H^1 - w_H^0), v) + (u_H^1 - u_h^1, v). \end{aligned}$$

The first term can be estimated in the following way

$$\begin{aligned} |((u_h^1 - u_h^0) - (w_H^1 - w_H^0), v)| &\leq \left| \left(\int_0^{t^1} \partial_t (u_h - w_H), v \right) \right| \\ &\leq \Delta t \|\eta_t\|_{L^\infty([0, T]; L^2(\Omega))} \|v\|_{L^2(\Omega)}. \end{aligned}$$

To estimate the second term, by the Taylor's expansion, we get

$$u_h^1 = u_h^0 + \Delta t (u_h)_t^0 + \frac{\Delta t^2}{2} (u_h)_{tt}^0 + \frac{\Delta t^3}{6} (u_h)_{ttt}(\cdot, s), \quad \text{where } 0 < s < t^1.$$

By the definition of u_H^1 ,

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

Thus,

$$\begin{aligned} (u_H^1 - u_h^1, v) &= \frac{\Delta t^2}{2} (\tilde{v} - (u_h)_{tt}^0, v) - \frac{\Delta t^3}{6} ((u_h)_{ttt}(\cdot, s), v) \\ &= \frac{\Delta t^2}{2} [(f^0, v) - a(u_h^0, v) + ((u_h)_{tt}^0, v)] - \frac{\Delta t^3}{6} ((u_h)_{ttt}(\cdot, s), v) \\ &= -\frac{\Delta t^3}{6} ((u_h)_{ttt}(\cdot, s), v) \end{aligned}$$

which proves the Lemma. □

Lemma 5 For $n \geq 1$, we have

$$\|r^n\|_{L^2(\Omega)} \leq C(\Delta t)^{-1} \int_{t_{n-1}}^{t_{n+1}} \|\eta_{tt}(\cdot, \tau)\|_{L^2(\Omega)} + \Delta t \int_{t_{n-1}}^{t_{n+1}} \|(u_h)_{tttt}(\cdot, \tau)\|_{L^2(\Omega)} d\tau.$$

Proof: By the definition of r^n ,

$$\begin{aligned} \|r^n\|_{L^2(\Omega)} &= \|(u_h)_{tt}^n - \delta^2 w_H^n\|_{L^2(\Omega)} \\ &\leq \|\delta^2 (w_H^n - u_h^n)\|_{L^2(\Omega)} + \|(u_h)_{tt}^n - \delta^2 u_h^n\|_{L^2(\Omega)}. \end{aligned}$$

Using the identity

$$v^{n+1} - 2v^n + v^{n-1} = \Delta t \int_{t_{n-1}}^{t_{n+1}} \left(1 - \frac{|\tau - t_n|}{\Delta t} \right) v_{tt}(\tau) d\tau,$$

the first term can be estimated as follows

$$\begin{aligned} (\delta^2 (w_H^n - u_h^n), v) &= \frac{1}{\Delta t} \int_{t_{n-1}}^{t_{n+1}} \left(1 - \frac{|\tau - t_n|}{\Delta t} \right) ((w_H)_{tt} - (u_h)_{tt}, v)(\tau) d\tau \\ &\leq \frac{1}{\Delta t} \int_{t_{n-1}}^{t_{n+1}} \|\eta_{tt}(\cdot, \tau)\|_{\tilde{L}^2(\Omega)} \|v\|_{\tilde{L}^2(\Omega)} d\tau. \end{aligned}$$

To estimate the term $\|(u_h)_{tt}^n - \delta^2 u_h^n\|_{L^2(\Omega)}$, we use

$$\delta^2 u_h^n = (u_h)_{tt}^n + \frac{1}{6\Delta t^2} \int_{t_{n-1}}^{t_{n+1}} (\Delta t - |\tau - t_n|)^3 (u_h)_{tttt}(\cdot, \tau) d\tau.$$

This implies

$$\|(u_h)_{tt}^n - \delta^2 u_h^n\|_{L^2(\Omega)} \leq \frac{\Delta t}{6} \int_{t_{n-1}}^{t_{n+1}} \|(u_h)_{tttt}(\cdot, \tau)\|_{L^2(\Omega)} d\tau.$$

□

Using the definition of R^n and the above two lemma, we get

$$\begin{aligned} & \|R^n\|_{L^2(\Omega)} \\ \leq & C \left(\int_0^{t_n} \|\eta_{tt}(\cdot, \tau)\|_{L^2(\Omega)} + \|\eta_t\|_{L^\infty([0, T]; L^2(\Omega))} + \Delta t^2 \int_0^{t_n} \|(u_h)_{tttt}(\cdot, \tau)\|_{L^2(\Omega)} + \Delta t^2 \|(u_h)_{ttt}\|_{C([0, T]; L^2(\Omega))} \right). \end{aligned}$$

Hence we obtain

$$\begin{aligned} & \Delta t \sum_{n=0}^N \|R^n\|_{L^2(\Omega)} \\ \leq & 2T \max_{0 \leq n \leq N} \|R^n\|_{L^2(\Omega)} \\ \leq & C \left(\int_0^T \|\eta_{tt}(\cdot, \tau)\|_{L^2(\Omega)} + \|\eta_t\|_{L^\infty([0, T]; L^2(\Omega))} + \Delta t^2 \int_0^T \|(u_h)_{tttt}(\cdot, \tau)\|_{L^2(\Omega)} + \Delta t^2 \|(u_h)_{ttt}\|_{C([0, T]; L^2(\Omega))} \right). \end{aligned}$$

Combining the estimates of η proved in Section 3 and (32), we obtain the error estimate for the fully discrete scheme (5).

6 Conclusions

In this paper, we present a multiscale simulation method based on Generalized Multiscale Finite Element Method for solving the wave equation in heterogeneous media. For the construction of multiscale basis functions, we divide the snapshot space into two spaces. The first snapshot space represents the degrees of the freedom associated with boundary nodes and consists of a -harmonic functions. The second snapshot space represents the interior degrees of the freedom and consists of all zero Dirichlet vectors. For each snapshot space, we introduce local spectral problems motivated by the analysis presented in the paper. We use these local spectral problems to identify important modes in each of the snapshot spaces. The local spectral problems are designed to achieve a high accuracy and motivated by the global coupling formulation. The use of multiple snapshot spaces and multiple spectral problems is one of the novelties of this work. Using the dominant modes from local spectral problems, multiscale basis functions are constructed to represent the solution space locally within each coarse block. These multiscale basis functions are coupled via the symmetric interior penalty discontinuous Galerkin method which provides a block diagonal mass matrix, and, consequently, results in fast computations in an explicit time discretization. Numerical examples are presented. In particular, we discuss how the modes from our snapshot spaces can affect the accuracy of the method. Our numerical results show that one can obtain an accurate approximation of the solution with GMsFEM using less than 3% of the total local degrees of freedom. We also test oversampling strategies following [15]. Analysis of the method is presented.

7 Appendix

In this Appendix, we will prove Lemma 1. Let $v \in V_H^1$. By assumption, $u \in H^2(\Omega)$, thus $a_{DG}(u, v)$ is well-defined and we have

$$\sum_{K \in \mathcal{T}^H} \left(\frac{\partial^2 u}{\partial t^2}, v \right)_{L^2(K)} + a_{DG}(u, v) = \sum_{K \in \mathcal{T}^H} (f, v)_{L^2(K)}, \quad \forall v \in H^1(\mathcal{T}^H).$$

Moreover, the following standard finite element error estimate holds

$$|u - u_h|_{H^1(\Omega)} \leq Ch|u|_{H^2(\Omega)}.$$

By the definition of the consistency error, we have

$$R_{u_h}(v) = \left(\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u_h}{\partial t^2}, v \right) + a(u, v) - a_{DG}(u_h, v), \quad \forall v \in H^1(\mathcal{T}^H). \quad (33)$$

Next, we define $v_c \in V_h$ in the following way. For each vertex in the triangulation, the value of v_c is defined as the average value of v at this vertex. Then by direct calculations, we have

$$\sum_{K \in \mathcal{T}^H} |v - v_c|_{H^1(K)}^2 \leq C \frac{1}{h} \sum_{e \in \mathcal{E}^H} \|[v]\|_{L^2(e)}^2$$

and

$$\sum_{K \in \mathcal{T}^H} \|v - v_c\|_{L^2(K)}^2 \leq Ch \sum_{e \in \mathcal{E}^H} \|[v]\|_{L^2(e)}^2.$$

Clearly, we have $[v - v_c]_e = [v]_e$ for all $e \in \mathcal{E}^H$ since $v_c \in C^0(\Omega)$. Therefore we get

$$\|v - v_c\|_{H^1(\mathcal{T}^H)}^2 \leq C \frac{1}{h} \sum_{e \in \mathcal{E}^H} \|[v]\|_{L^2(e)}^2.$$

By (33) and (2) as well as the fact that $a_{DG}(u_h, v_c) = a(u_h, v_c)$, we have

$$R_{u_h}(v) = \sum_{K \in \mathcal{T}^H} \left(\frac{\partial^2(u - u_h)}{\partial t^2}, v - v_c \right)_{L^2(K)} + a_{DG}(u - u_h, v - v_c).$$

Next, we will estimate the two terms on the right hand side. For the first term, we have

$$\begin{aligned} \sum_{K \in \mathcal{T}^H} \left(\frac{\partial^2(u - u_h)}{\partial t^2}, v - v_c \right)_{L^2(K)} &\leq \left\| \frac{\partial^2(u - u_h)}{\partial t^2} \right\|_{L^2(\Omega)} \left(\sum_{K \in \mathcal{T}^H} \|v - v_c\|_{L^2(K)}^2 \right)^{\frac{1}{2}} \\ &\leq Ch \left\| \frac{\partial^2(u - u_h)}{\partial t^2} \right\|_{L^2(\Omega)} \left(\frac{1}{h} \sum_{e \in \mathcal{E}^H} \|[v]\|_{L^2(e)}^2 \right)^{\frac{1}{2}}. \end{aligned}$$

For the second term, by the definition of a_{DG} and the Cauchy-Schwarz inequality, we have

$$\begin{aligned} &a_{DG}(u - u_h, v - v_c) \\ &= \sum_{K \in \mathcal{T}^H} \int_K a \nabla(u - u_h) \cdot \nabla(v - v_c) - \sum_{e \in \mathcal{E}^H} \int_e \{a \nabla(u - u_h) \cdot n\} [v] \\ &\leq \sum_{K \in \mathcal{T}^H} a_1 |u - u_h|_{H^1(K)} |v - v_c|_{H^1(K)} + \left(\sum_{K \in \mathcal{T}^H} \int_{\partial K} (a \nabla(u - u_h) \cdot n)^2 \right)^{\frac{1}{2}} \left(\sum_{e \in \mathcal{E}^H} \int_e [v]^2 \right)^{\frac{1}{2}}. \end{aligned}$$

To estimate the flux term above, we let I_K be the standard finite element interpolant. Then we have

$$\begin{aligned} &\sum_{K \in \mathcal{T}^H} \int_{\partial K} (a \nabla(u - u_h) \cdot n)^2 \\ &\leq 2 \left(\sum_{K \in \mathcal{T}^H} \int_{\partial K} (a \nabla(u - I_K(u)) \cdot n)^2 \right) + 2 \left(\sum_{K \in \mathcal{T}^H} \int_{\partial K} (a \nabla(I_K(u) - u_h) \cdot n)^2 \right) \\ &\leq Ca_1 \left(h |u|_{H^2(K)}^2 + \frac{1}{h} |I_K(u) - u_h|_{H^1(K)}^2 \right) \\ &\leq Ca_1 \left(h |u|_{H^2(K)}^2 + \frac{1}{h} |I_K(u) - u|_{H^1(K)}^2 + \frac{1}{h} |u - u_h|_{H^1(K)}^2 \right) \\ &\leq Ca_1 h |u|_{H^2(K)}^2. \end{aligned}$$

Next we will estimate the term $\sum_{K \in \mathcal{T}^H} a_1 |u - u_h|_{H^1(K)} |v - v_c|_{H^1(K)}$. We have

$$\begin{aligned}
& \sum_{K \in \mathcal{T}^H} a_1 |u - u_h|_{H^1(K)} \cdot |v - v_c|_{H^1(K)} \\
& \leq C a_1 \left(\sum_{K \in \mathcal{T}^H} |u - u_h|_{H^1(K)}^2 \right)^{\frac{1}{2}} \left(\sum_{K \in \mathcal{T}^H} |v - v_c|_{H^1(K)}^2 \right)^{\frac{1}{2}} \\
& \leq C a_1 \left(\frac{1}{h} \sum_{K \in \mathcal{T}^H} |u - u_h|_{H^1(K)}^2 \right)^{\frac{1}{2}} \left(\sum_{e \in \mathcal{E}^H} \|[v]\|_{L^2(e)}^2 \right)^{\frac{1}{2}} \\
& \leq C a_1 h |u|_{H^2} \left(\frac{1}{h} \sum_{e \in \mathcal{E}^H} \|[v]\|_{L^2(e)}^2 \right)^{\frac{1}{2}}.
\end{aligned}$$

Combining the above estimates, we get

$$|R_{u_h}(v)| \leq \frac{C}{a_0} h \left(\|(u - u_h)_{tt}\|_{L^2(K)} + a_1 |u|_{H^2(\Omega)} \right) \left(\frac{1}{h} \sum_{e \in \mathcal{E}^H} \|a[v]\|_{L^2(e)}^2 \right)^{\frac{1}{2}}.$$

Finally, we assume that the second time derivatives of u and u_h are smooth functions. Then we have

$$(u_{ttt}, v) + \int_{\Omega} a \nabla u_t \cdot \nabla v = \int_{\Omega} f_t v.$$

Letting $v = u_{tt}$, we have

$$\frac{d}{dt} (\|u_{tt}\|_{L^2(\Omega)}^2 + \int_{\Omega} a |\nabla u_t|^2) = \int_{\Omega} f_t u_{tt}$$

which leads to

$$\|u_{tt}\|_{L^\infty([0, T]; L^2(\Omega))} + \|u_t\|_{L^\infty([0, T]; a)} \leq C (\|f_t\|_{L^1([0, T]; L^2(\Omega))} + \|u_{tt}(\cdot, 0)\|_{L^2(\Omega)} + |u_t(\cdot, 0)|_a).$$

Similarly, for the finite element solution u_h , we have

$$\begin{aligned}
& \|(u_h)_{tt}\|_{L^\infty([0, T]; L^2(\Omega))} + \|(u_h)_t\|_{L^\infty([0, T]; a)} \\
& \leq C (\|f_t\|_{L^1([0, T]; L^2(\Omega))} + \|(u_h)_{tt}(\cdot, 0)\|_{L^2(\Omega)} + |(u_h)_t(\cdot, 0)|_a).
\end{aligned}$$

Consequently, we get

$$|R_{u_h}(v)| \leq C h (\|f_t\|_{L^1([0, T]; L^2(\Omega))} + \|u_{tt}(\cdot, 0)\|_{L^2(\Omega)} + a_1 |u|_{H^2(\Omega)}) \|v\|_{H^1(\mathcal{T}^H)}.$$

Finally, the constant $C(u, f)$ in the lemma can be chosen as

$$C(u, f) \approx \|f_t\|_{L^1([0, T]; L^2(\Omega))} + \|u_{tt}(\cdot, 0)\|_{L^2(\Omega)} + a_1 |u|_{H^2(\Omega)}. \quad (34)$$

References

- [1] H. Chan, E. Chung, and G. Cohen. Stability and dispersion analysis of staggered discontinuous Galerkin method for wave propagation. *Int. J. Numer. Anal. Model.*, 10:233–256, 2013.
- [2] E. Chung, Y. Efendiev, and R. Gibson. An energy-conserving discontinuous multiscale finite element method for the wave equation in heterogeneous media. *Advances in Adaptive Data Analysis*, 3:251–268, 2011.
- [3] E. Chung and B. Engquist. Optimal discontinuous Galerkin methods for wave propagation. *SIAM J. Numer. Anal.*, 44:2131–2158, 2006.

- [4] E. Chung and B. Engquist. Optimal discontinuous Galerkin methods for the acoustic wave equation in higher dimensions. *SIAM J. Numer. Anal.*, 47:3820–3848, 2009.
- [5] E. Chung and P. Ciarlet Jr. A staggered discontinuous Galerkin method for wave propagation in media with dielectrics and meta-materials. *J. Comput. Appl. Math.*, 239:189–207, 2013.
- [6] E. Chung, P. Ciarlet Jr., and T. Yu. Convergence and superconvergence of staggered discontinuous Galerkin methods for the three-dimensional maxwell’s equations on cartesian grids. *J. Comput. Phys.*, 235:14–31, 2013.
- [7] E. Chung, H. Kim, and O. Widlund. Two-level overlapping schwarz algorithms for a staggered discontinuous Galerkin method. *SIAM J. Numer. Anal.*, 51:47–67, 2013.
- [8] E. Chung and C. Lee. A staggered discontinuous Galerkin method for the convection-diffusion equation. *J. Numer. Math.*, 20:1–31, 2012.
- [9] E. Chung and C. Lee. A staggered discontinuous Galerkin method for the curl-curl operator. *IMA J. Numer. Anal.*, 32:1241–1265, 2012.
- [10] E. Chung and W. Leung. A sub-grid structure enhanced discontinuous Galerkin method for multiscale diffusion and convection-diffusion problems. *Commun. Comput. Phys.*, 14:370–392, 2013.
- [11] Jonas D. De Basabe and Mrinal K. Sen. New developments in the finite-element method for seismic modeling. *The Leading Edge*, 28(5):562–567, 2009.
- [12] Florence Delprat-Jannaud and Patrick Lailly. Wave propagation in heterogeneous media: Effects of fine-scale heterogeneity. *Geophysics*, 73(3):T37–T49, 2008.
- [13] Y. Efendiev, J. Galvis, and T. Hou. Generalized multiscale finite element method. 2012.
- [14] Y. Efendiev, J. Galvis, R. Lazarov, M. Moon, and M. Sarkis. Generalized multiscale finite element method. IPDG. Submitted, 2013.
- [15] Y. Efendiev, J. Galvis, G. Li, and M. Presho. Generalized multiscale finite element methods. oversampling strategies. *International Journal for Multiscale Computational Engineering*, to appear, 2013.
- [16] R. Gibson, K. Gao, E. Chung, and Y. Efendiev. Multiscale modeling of acoustic wave propagation in two-dimensional media. *Under Review*.
- [17] M. Grote and D. Schotzau. Optimal error estimates for the fully discrete interior penalty DG method for the wave equation. *J. Sci. Comput.*, 40:257–272, 2009.
- [18] Verena Hermann, Martin Käser, and Cristóbal E. Castro. Non-conforming hybrid meshes for efficient 2-D wave propagation using the discontinuous galerkin method. *Geophysical Journal International*, 184(2):746–758, 2011.
- [19] Martin Käser, Christian Pelties, Cristobal E. Castro, Hugues Djikpesse, and Michael Prange. Wave-field modeling in exploration seismology using the discontinuous Galerkin finite-element method on hpc infrastructure. *The Leading Edge*, 29(1):76–85, 2010.
- [20] Dimitri Komatitsch, Dominik GÖddeke, Gordon Erlebacher, and David Michéa. Modeling the propagation of elastic waves using spectral elements on a cluster of 192 gpus. *Computer Science - Research and Development*, 25:75–82, 2010. 10.1007/s00450-010-0109-1.
- [21] Dimitri Komatitsch and Jeroen Tromp. Introduction to the spectral element method for three-dimensional seismic wave propagation. *Geophysical Journal International*, 139(3):806–822, 1999.

- [22] Dimitri Komatitsch and Jeroen Tromp. Spectral-element simulations of global seismic wave propagation—i. validation. *Geophysical Journal International*, 149(2):390–412, 2002.
- [23] B. Lombard, J. Piraux, C. Gélis, and J. Virieux. Free and smooth boundaries in 2-D finite-difference schemes for transient elastic waves. *Geophysical Journal International*, 172(1):252–261, 2008.
- [24] Y. J. Masson and S. R. Pride. Finite-difference modeling of Biot’s poroelastic equations across all frequencies. *Geophysics*, 75(2):N33–N41, 2010.
- [25] Peter Moczo, Jozef Kristek, Martin Galis, Emmanuel Chaljub, and Vincent Etienne. 3-D finite-difference, finite-element, discontinuous-Galerkin and spectral-element schemes analysed for their accuracy with respect to p-wave to s-wave speed ratio. *Geophysical Journal International*, 187(3):1645–1667, 2011.
- [26] Christina Morency, Yang Luo, and Jeroen Tromp. Acoustic, elastic and poroelastic simulations of CO₂ sequestration crosswell monitoring based on spectral-element and adjoint methods. *Geophysical Journal International*, pages no–no, 2011.
- [27] C. Pelties, M. Käser, V. Hermann, and C. E. Castro. Regular versus irregular meshing for complicated models and their effect on synthetic seismograms. *Geophysical Journal International*, pages no–no, 2010.
- [28] Beatrice M. Riviere. *Discontinuous Galerkin Methods For Solving Elliptic And parabolic Equations: Theory and Implementation*. SIAM, 2008.
- [29] Erik H. Saenger, Radim Ciz, Oliver S. Krüger, Stefan M. Schmalholz, Boris Gurevich, and Serge A. Shapiro. Finite-difference modeling of wave propagation on microscale: A snapshot of the work in progress. *Geophysics*, 72(5):SM293–SM300, 2007.
- [30] William Symes, Igor S. Terentyev, and Tetyana Vdovina. Getting it right without knowing the answer: Quality control in a large seismic modeling project. *SEG Technical Program Expanded Abstracts*, 28(1):2602–2606, 2009.
- [31] Jean Virieux. SH-wave propagation in heterogeneous media: Velocity-stress finite-difference method. *Geophysics*, 49(11):1933–1942, 1984.
- [32] Jean Virieux. P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method. *Geophysics*, 51(4):889–901, 1986.
- [33] W. Zhang, L. Tong, and E. Chung. Exact nonreflecting boundary conditions for three dimensional poroelastic wave equations. *Comm. Math. Sci.*, To appear.
- [34] W. Zhang, L. Tong, and E. Chung. A new high accuracy locally one-dimensional scheme for the wave equation. *J. Comput. Appl. Math.*, 236:1343–1353, 2011.
- [35] W. Zhang, L. Tong, and E. Chung. Efficient simulation of wave propagation with implicit finite difference schemes. *Numer. Math. Theor. Meth. Appl.*, 5:205–228, 2012.