DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS
Volume **36**, Number **8**, August **2016**

doi:10.3934/dcds.2016.36.4451

pp. **4451–4476**

OPTIMAL LOCAL MULTI-SCALE BASIS FUNCTIONS FOR LINEAR ELLIPTIC EQUATIONS WITH ROUGH COEFFICIENTS

Thomas Y. Hou and Pengfei Liu*

1200 E California Blvd, MC 9-94
California Institute of Technology
Pasadena, CA 91125, USA

Dedicated to Professor Peter Lax in the occasion of his 90th birthday with admiration and friendship.

Abstract. This paper addresses a multi-scale finite element method for second order linear elliptic equations with rough coefficients, which is based on the compactness of the solution operator, and does not depend on any scale-separation or periodicity assumption of the coefficient. We consider a special type of basis functions, the multi-scale basis, which are harmonic on each element and show that they have optimal approximation property for fixed local boundary conditions. To build the optimal local boundary conditions, we introduce a set of interpolation basis functions, and reduce our problem to approximating the interpolation residual of the solution space on each edge of the coarse mesh. And this is achieved through the singular value decompositions of some local oversampling operators. Rigorous error control can be obtained through thresholding in constructing the basis functions. The optimal interpolation basis functions are also identified and they can be constructed by solving some local least square problems. Numerical results for several problems with rough coefficients and high contrast inclusions are presented to demonstrate the capacity of our method in identifying and exploiting the compact structure of the local solution space to achieve computational savings.

1. Introduction.

Many problems of practical importance in science and engineering have multi-scale feature: composite materials and flows in porous media are typical examples of such kind. In some cases the quantities of interest (QoI) are only related to the large-scale properties of the solutions, but since the fine-scale features of the model can have significant impact on the large-scale properties of the solutions, one needs to use a very fine mesh to resolve the small-scale variations of the problem to get faithful numerical results. The computational cost can be prohibitive. For these so-called *multi-scale problems*, it is desirable to develop upscaling methods that allow us to efficiently incorporate the small-scale features of the problem into the large-scale properties of the solutions.

2010 *Mathematics Subject Classification.* Primary: 65N30; Secondary: 35J25.

Key words and phrases. Elliptic PDEs with rough coefficients, multi-scale finite element method, oversampling, optimal multi-scale basis, high-contrast.

The research is supported by Air force MURI Grant FA9550-09-1-0613, DOE Grant DE-FG02-06ER257, and NSF Grant No. DMS-1318377, DMS-1159138.

* Corresponding author: P. Liu.

In this work, we use the following second order linear elliptic equation with homogeneous Dirichlet boundary condition to illustrate our upscaling methodology,

$$\begin{cases} -\mathrm{div}(a(x)\nabla u(x)) = f(x), & x \in D, \\ u(x)|_{\partial D} = 0, \end{cases} \tag{1.1}$$

where D is a convex polygon domain in R^d with $d = 2, 3$. We assume that the equation is uniformly elliptic, i.e., there exist $\lambda_{\min} > 0$ and $\lambda_{\max} > 0$ such that

$$a(x) \in [\lambda_{\min}, \lambda_{\max}]. \tag{1.2}$$

We do not assume any regularity of the coefficient $a(x) \in L^\infty(D)$, which may have multiple spatial scales, thus the above equation (1.1) can be used to model diffusion process in strongly heterogeneous media. We also assume that in (1.1) the forcing function $f(x) \in L^2(D)$, not just in $H^{-1}(D)$. The existence of solution to (1.1), $u(x) \in H_0^1(D)$ follows from the Lax-Milgram theorem, and we have

$$c\|f\|_{H^{-1}(D)} \leq \|u(x)\|_{H_0^1(D)} \leq C\|f\|_{H^{-1}(D)}. \tag{1.3}$$

Classical finite element methods use piecewise linear (polynomial) functions to approximate the solution space, and the convergence of these methods generally depends on the following approximation property and regularity result

$$\|u(x) - Ju(x)\|_{H_0^1(D)} \leq CH\|u(x)\|_{H^2(D)}, \quad \|u(x)\|_{H^2(D)} \leq C\|f(x)\|_{L^2(D)}, \tag{1.4}$$

where Ju is the piecewise polynomial interpolation of $u(x)$, and H is the underlying mesh size. Thus $O(H)$ accuracy can be obtained if mesh of size $O(H)$ is employed in the discretization. Classical finite element methods may fail for these multi-scale problems, since for rough $a(x)$, $\|u(x)\|_{H^2}$ cannot be bounded by $\|f(x)\|_{L^2(D)}$ in (1.4). It is actually shown in [8] that the polynomial finite elements can perform arbitrarily badly in this setting. In practice, one needs a much finer mesh to get $O(H)$ accuracy, thus (1.1) can serve as a typical example of multi-scale problem.

One strategy to numerically solve the multi-scale problem (1.1) is using problem-dependent basis (instead of polynomials) that incorporates properties of the coefficient $a(x)$ to approximate the solution space. One first constructs basis functions

$$\phi_1(x), \phi_2(x), \ldots \phi_n(x) \in H_0^1(D), \tag{1.5}$$

that may depend on the elliptic operator in (1.1) and find the numerical solution

$$u_H(x) \in V_H(x) = \mathrm{span}\{\phi_1(x), \phi_2(x), \ldots \phi_n(x)\} \subset H_0^1(D), \tag{1.6}$$

using the Galerkin projection. Namely, we find $u_H(x) \in V_H$, such that

$$a(u_H(x), v(x)) = \langle f(x), v(x)\rangle, \quad \text{for all} \quad v(x) \in V_H, \tag{1.7}$$

where

$$a(u(x), v(x)) = \int_D \nabla u(x)^t a(x) \nabla v(x) \mathrm{d}x, \quad \langle f(x), v(x)\rangle = \int_D f(x)v(x)\mathrm{d}x. \tag{1.8}$$

The numerical solution satisfies the following optimal property

$$\|u(x) - u_H(x)\|_E = \inf_{v(x)\in V_H} \|u(x) - v(x)\|_E, \tag{1.9}$$

where the energy norm is equivalent to the $H_0^1(D)$ norm, and defined as

$$\|u(x)\|_E^2 = a(u(x), u(x)) = \int_D \nabla u(x)^t a(x) \nabla u(x) \mathrm{d}x. \tag{1.10}$$

In this work we will employ the above strategy to numerically solve (1.1). Note that to obtain the numerical solution $u_H(x)$ from the Galerkin projection (1.7), one needs to solve a linear system of size $n \times n$. Thus to make the computational cost small, we want the number of the basis functions used in (1.5) to be small. Besides, we want the basis functions in (1.5) to have compact support such that the stiffness matrix formed in (1.7) is sparse thus easy to compute and invert.

We propose an effective method to construct basis functions (1.5) with optimal local approximation property. Our method is based on the compactness of the solution operator to (1.1) restricted on local regions of the domain. To be specific, we introduce the following operator

$$T_i : \quad f(x) \to u_i(x) = u(x)|_{D_i}, \tag{1.11}$$

where D_i is a local subset of D with size $O(H)$, and H is chosen according to the desired order of accuracy. The compactness of the operator T_i will be demonstrated numerically in section 2. On each local region of the domain, D_i, we decompose the local solution $u_i(x)$ to two orthogonal parts with respect to the energy norm (1.8): an $a(x)$-harmonic part, and a local bubble part. We show that the bubble part of the solution is small and its compact structure can be easily identified by inverting the elliptic equation (1.1) locally on each region D_i. We consider approximating the solution space using a special type of basis functions that are $a(x)$-harmonic on each D_i, and call basis functions of such type *multi-scale basis*. Due to the smallness of the bubble part of the solution, we demonstrate that multi-scale basis functions are optimal in approximating the solution space for fixed local boundary conditions on ∂D_i if only $O(H)$ accuracy in the energy norm is desired.

The $a(x)$-harmonic part of the solution only depends on the restriction of the solution on the boundary of the local regions D_i, and we seek to identify the compact structure of the trace of the solution space on ∂D_i. Using a primary set of multi-scale interpolation basis functions (*nodal multi-scale basis*) , $\psi_i(x)$, we can reduce our problem to approximating the interpolation residual of the solution on each edge e of the coarse mesh, which we denote by $T_e f(x)$. We then introduce a local oversampling operator P_{OS} that maps the solution on an oversampling domain W to the interpolation residual $T_e f(x)$, and decompose T_e using P_{OS} and a global solution operator. We employ the compactness of the oversampling operator to construct the *edge multi-scale basis* for each edge e through singular value decomposition. The optimal choice of the nodal multi-scale basis functions $\psi_i(x)$ is identified as the solution to some local least square problems, which makes the singular values of P_{OS} have the fastest decay. Since the resulting basis functions (1.5) are $a(x)$-orthogonal to the bubble part of the solution space, we can add the bubble part back to our numerical solution by simply solving some local cell problems (independently from the Galerkin projection).

Our multiscale method consists of two stages: in the offline stage we identify the local compact structure of the solution space, and build multi-scale basis functions and the corresponding stiffness matrix; in the online stage, for any given forcing function $f(x) \in L^2(D)$, we solve the equation (1.1) efficiently using the multi-scale basis functions constructed offline with a very low computation cost. Our method can achieve significant computational savings in the multi-query setting where equation (1.1) needs to be solved for multiple times with different forcing.

Several numerical examples with rough coefficients and high-contrast channels are presented. Our method achieves high accuracy and significant computational

savings for these problems in the online stage. In this work we demonstrate our methodology through the second-order scalar elliptic equation, but it can be easily generalized to other linear elliptic problems such as the elasticity equations.

Below we review some related works in the literature. The classical homogenization theories, including the periodic homogenization [9, 33, 46, 15, 2, 1], and the H, G, Γ-convergence theories [41, 17, 16, 50, 49, 40, 24], consider the convergence of a sequence of operators parameterized by ϵ as $\epsilon \to 0$. In the Multi-scale Finite Element Method (MsFEM) [28, 29, 22, 11, 30, 21, 18, 19, 13, 20, 12], nodal basis functions that incorporate properties of the elliptic operator (1.5) are constructed by solving local elliptic boundary value problems. Convergence analysis of MsFEM in the periodic setting was given in [29, 22, 11]. An oversampling technique to reduce the resonance error introduced due to the artificial local boundary conditions in the basis functions was proposed in [29]. The MsFEM framework motivated a lot of interesting works and was further developed in [3, 31, 32, 35, 39]. In [38, 7, 4], the generalized finite element method was proposed, which provides a general framework to combine local approximation spaces together using a partition of unity formulation. In [6, 5], the local basis functions for this framework were constructed by solving some local spectral problems. In [36, 25, 47], the solution space is divided into two orthogonal parts, the coarse multi-scale space and the fine scale space. The coarse multi-scale space can approximate the solution space to (1.1) up to $O(H)$ accuracy in the energy norm. A set of basis functions were first identified as the solutions of some global elliptic problems, and then shown to decay exponentially fast. Thus their construction can be localized to regions of size $O(H \log(1/H))$ to retain the $O(H)$ approximation accuracy. Harmonic coordinates were introduced in [34], and employed for (quasi) one-dimensional elliptic problems in [7, 4]. In [44], the authors proved that the solutions to (1.1) gain an order of regularity with respect to the Harmonic coordinates and proposed upscaling methods based on this property. The polyharmonic spline functions were introduced in [45], where they were identified as solutions of some global optimization problems and shown to have super-localization property. These basis functions were later interpreted as conditional expectations in the Bayesian inference setting [42]. This novel point of view was further developed in [43], and a multi-resolution decomposition of the solution space was obtained based on a hierarchical information game formulation.

The remaining part of this paper is organized as follows. In section 2, we demonstrate the compactness of the solution operator restricted to local regions of the domain. In section 3, we decompose the solutions on each local region to different parts corresponding to the trace of the solution on the edges of the coarse mesh, and identify their compact structures separately. In section 4, numerical results are presented to demonstrate the capacity of our method in identifying and exploiting the compactness of the solution space to achieve computational savings. Concluding remarks are made in section 5.

2. Compactness of the solution space restricted to local regions of the domain.

The existence of a finite number of basis functions (1.5) that can approximate the solution space to (1.1) up to any accuracy is implied by the compactness of the solution operator, T, which maps from the forcing function $f(x) \in L^2(D)$ to the corresponding solution $u(x) \in H_0^1(D)$.

$$T: \quad f(x) \in L^2(D) \to u(x) \in H_0^1(D). \tag{2.1}$$

The compactness of T was analyzed in [37, 10], and employed for elliptic equations with random input data recently in [26, 27] for stochastic model reduction.

To be specific, the solution operator T can be decomposed as

$$T = L^{-1} I_{L^2(D)\to H^{-1}(D)}, \tag{2.2}$$

where L^{-1} maps $f(x) \in H^{-1}(D)$ to the solution $u(x) \in H_0^1(D)$, and $I_{L^2(D)\to H^{-1}(D)}$ is the embedding operator from $L^2(D)$ to $H^{-1}(D)$. From (1.3), we can see that L^{-1} is continuous and indeed a homomorphism, and the compactness of $I_{L^2(D)\to H^{-1}(D)}$ is well known based on the Sobolev space theory [23]. Thus the compactness of T follows from the decomposition (2.2). To quantify the approximability of T by a finite-rank operator, we consider its Kolmogorov-n width [48].

Definition 2.1 (Kolmogorov n-width)**.** For a compact linear operator T that maps between two Hilbert spaces, we define its Kolmogorov n-width as

$$d_n(T) = \inf_{T_n} \|T - T_n\|, \tag{2.3}$$

where T_n runs over all rank-n linear operators.

Due to the fact that L^{-1} is a homomorphism, one can easily see that the Kolmogorov-n width of T (mapping from $L^2(D)$ to $H_0^1(D)$) is only different from that of $I_{L^2(D)\to H^{-1}(D)}$ by a constant factor ,

$$c d_n(I_{L^2(D)\to H^{-1}(D)}) \le d_n(T) \le C d_n(I_{L^2(D)\to H^{-1}(D)}), \tag{2.4}$$

where C and c depend on $\lambda_{\min}$, $\lambda_{\max}$ (1.2) and D.

The Kolmogorov-n width of the embedding operator is well-known [37, 10],

$$d_n(I_{L^2(D)\to H_0^1(D)}) = n^{-1/d}(C + o(1)), \quad n \to \infty. \tag{2.5}$$

From (2.4) (2.5) and (2.3), we obtain that there exist n basis functions, (1.5), with the following approximation property to the solution space of (1.1),

$$\sup_{\|f(x)\|_{L^2(D)}=1} \inf_{c_i} \|\sum_{i=1}^{n} c_i \phi_i(x) - u(x)\|_{H_0^1(D)} \le C n^{-1/d}. \tag{2.6}$$

The approximation property (2.6) is optimal, and does not depend on the regularity of the coefficient $a(x)$. For practical applications in multi-scale problems, we want the basis functions $\phi_i(x)$ to have local support such that the corresponding stiffness matrix in (1.7) is sparse and easy to invert. However, the basis functions in (2.6) whose existence is implied by (2.4) and (2.5) may be nonlocal.

Since our objective is to find basis functions (1.5) with local support, we consider a local region of the domain, D_i with diameter $O(H)$, and a slightly larger local domain that contains D_i, W, which we call the oversampling region. We consider the restriction of the solutions to (1.1) on W,

$$u_W(x) = u(x)|_W. \tag{2.7}$$

The local solution $u_W(x)$ can be decomposed to two parts,

$$u_W(x) = u_W^1(x) + u_W^2(x), \tag{2.8}$$

where

$$\begin{cases} -\mathrm{div}(a(x)\nabla u_W^1(x)) = 0, & x \in W, \\ u_W^1(x) = u_W(x), & x \in \partial W, \end{cases} \tag{2.9}$$

and

$$\begin{cases} -\mathrm{div}(a(x)\nabla u_W^2(x)) = f(x), & x \in W, \\ u_W^2(x) = 0, & x \in \partial W. \end{cases} \tag{2.10}$$

We call the first part $u_W^1(x)$ the local $a(x)$-harmonic part, and the second part $u_W^2(x)$ the local bubble part. The two parts are orthogonal with respect to the local inner product, $a_W(\cdot,\cdot)$,

$$a_W(u_W^1(x), u_W^2(x)) = \int_W \nabla u_W^1(x)^t a(x) \nabla u_W^2(x) \mathrm{d}x = 0. \tag{2.11}$$

The local bubble part $u_W^2(x)$ is small in the sense that

$$\|u_W^2(x)\|^2_{H_0^1(W)} \leq CH^2 \|f(x)\|^2_{L^2(W)}, \tag{2.12}$$

which can be obtained from (1.3) and a scaling argument. Inequality (2.12) implies that if we only want to obtain $O(H)$ accuracy in our numerical solution, we can simply neglect the local bubble part.

Then we consider a local solution operator T_i that maps $f(x)$ to the local $a(x)$-harmonic part, $u_W^1(x)$ restricted on D_i,

$$T_i : \ f(x) \in L^2(D) \to u_W^1(x)|_{D_i} \in H^1(D_i), \tag{2.13}$$

and we want to construct local basis functions on D_i that can approximate the range of T_i. To demonstrate the compactness of T_D, we choose a set of orthonormal basis in the domain and range of T_i to discretize T_i as a matrix, and compute the decay of its singular values. We consider the following choice of coefficient in (1.1), which has multiple fine spatial scales and is illustrated in Figure 1a,

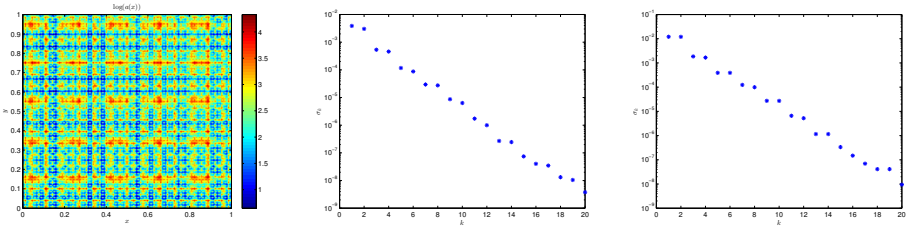
$$\begin{aligned} a(x) = \frac{1}{6}\Bigg(& \frac{1.1+\sin(2\pi x/\epsilon_1)}{1.1+\sin(2\pi y/\epsilon_1)} + \frac{1.1+\sin(2\pi y/\epsilon_2)}{1.1+\cos(2\pi x/\epsilon_2)} + \frac{1.1+\cos(2\pi x/\epsilon_3)}{1.1+\sin(2\pi y/\epsilon_3)} \\ & + \frac{1.1+\sin(2\pi y/\epsilon_4)}{1.1+\cos(2\pi x/\epsilon_4)} + \frac{1.1+\cos(2\pi x/\epsilon_5)}{1.1+\sin(2\pi y/\epsilon_5)} + \sin(4x^2y^2) + 1 \Bigg), \end{aligned} \tag{2.14}$$

where $\epsilon_1 = \frac{1}{5}$, $\epsilon_2 = \frac{1}{13}$, $\epsilon_3 = \frac{1}{17}$, $\epsilon_4 = \frac{1}{31}$, $\epsilon_5 = \frac{1}{65}$.

We choose $D = [0,1] \times [0,1]$, the oversampling region $W = [14H, 17H] \times [14H, 17H]$, and the local region $D_i = [15H, 16H] \times [15H, 16H]$, where $H = 1/32$. The decay of the singular values of the local solution operator (2.13) is plotted in Figure 1b. Then we compute the singular values for the local solution operator (2.13) to the Poisson equation in the same setting, the decay of which is plotted in Figure 1c. We can see that the singular values of the local solution operator decay very fast, and this fast decay does not deteriorate due to the roughness of the coefficient.

The fast decay of singular values of T_i implies that we can use a very small number of local basis functions (the first several left singular vectors of T_i) to get very good local approximation property. However, we cannot afford to construct T_i explicitly since it is a solution operator and its construction involves solving the equation (1.1) many times globally. It is known that for a low-rank operator, the main action of T_i can be captured in its image on some random vectors. This fact, to some degree, explains the success of some global upscaling methods [18, 14, 44] that use the linear combination of a small number of sampled global solutions to approximate the local solution space.

We will not pursue this perspective in this work. Instead, we introduce a local oversampling operator and construct optimal local multi-scale basis functions



(A) The rough coefficient. (B) Rough coefficient (1.1). (C) Poisson equation.

FIGURE 1. The fast decay of the singular values of the local solution operator (2.13), for rough coefficient (1.1) and possion equation.

employing the compactness of the oversampling operator through singular value decomposition. The resulting method does not involve any global solver of the equation (1.1). We will give the details of this method in the next section.

3. **Identifying the compact structure using oversampling.** In this section, we identify the compact structure of the local solution space through oversampling. In our numerical examples, the domain D is chosen to be $[0,1] \times [0,1]$, and we discretize D using a coarse square mesh of size H, which should be chosen according to the desired order of accuracy. With this discretization, we have

$$D = \cup_{i=1}^{N} D_i, \tag{3.1}$$

where D_i have disjoint interiors. Underlying this coarse mesh, we use a triangle fine mesh of size $O(h)$, which is a refinement of the coarse mesh. The fine mesh size h should be chosen such that it can resolve the small scale variation of the multi-scale coefficient in (1.1). In our method we solve the equation (1.1) on the coarse mesh, and the basis functions that we use are constructed and saved using linear basis functions on the fine mesh. The two level discretization is illustrated in Figure 2.

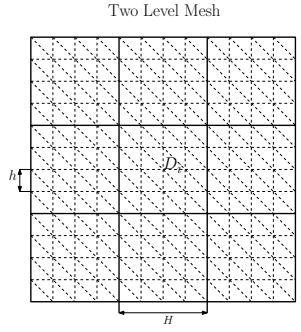


FIGURE 2

3.1. **The multi-scale basis.** We first introduce a special class of problem-dependent basis functions (1.5), which we call the *multi-scale basis*. These basis functions generalize the basis employed in MsFEM [28], but are not necessarily nodal.

Definition 3.1 (Multi-Scale basis)**.** For a discretization of D (3.1), we consider basis functions

$$\phi_1(x), \phi_2(x), \dots \phi_n(x) \in H_0^1(D). \tag{3.2}$$

If they are $a(x)$-harmonic on each element of the coarse discretization, D_j,

$$-\operatorname{div}(a(x)\nabla\phi_i(x)) = 0, \quad x \in D_j, \tag{3.3}$$

then we call them multi-scale basis functions.

Clearly, multi-scale basis functions are determined by their traces on the boundary of coarse elements ∂D_i since they are $a(x)$-harmonic in each D_i. We denote

$$\Gamma = \cup_{i=1}^N \partial D_i. \tag{3.4}$$

The following proposition implies that if the desired accuracy is $O(H)$, multi-scale basis functions are optimal for fixed local boundary conditions on Γ (3.4).

Proposition 1. *Consider a set of basis functions $\psi_i(x) \in H_0^1(D), i = 1, 2, \dots m$, and a set of multi-scale basis functions $\phi_i(x) \in H_0^1(D), i = 1, 2 \dots, n$ on a coarse mesh of size H, as shown in Figure 2. Denote the corresponding Galerkin numerical solution (1.7) to (1.1) using $\psi_i(x), i = 1, \dots m$ as $u^\psi(x)$, and the Galerkin solution using $\phi_i(x), i = 1, \dots n$ as $u_H^{MS}(x)$. If*

$$\operatorname{span}\{\phi_1(x)|_\Gamma, \dots, \phi_n(x)|_\Gamma\} = \operatorname{span}\{\psi_1(x)|_\Gamma, \dots, \psi_m(x)|_\Gamma\}. \tag{3.5}$$

Then we have

$$\|u(x) - u_H^{MS}(x)\|_E^2 \le \|u(x) - u^\psi(x)\|_E^2 + C\|f\|_{L^2(D)}^2 H^2. \tag{3.6}$$

Namely, if only $O(H)$ accuracy in the energy norm is desired, the multi-scale basis can perform as well as other set of basis functions, given that the local boundary conditions of the basis functions are the same.

To prove the above proposition, we first decompose the solution $u(x)$ to (1.1) to two parts. On each coarse mesh element D_i, we consider

$$u_i(x) = u(x)|_{D_i}, \tag{3.7}$$

and decompose it to an $a(x)$-harmonic part and a local bubble part, as in (2.8),

$$u_i(x) = u_i^1(x) + u_i^2(x), \quad x \in D_i, \tag{3.8}$$

where $u_i^1(x)$ is the local $a(x)$-harmonic part, and $u_i^2(x)$ is the local bubble part. Combining these local decompositions from all coarse elements D_i together we get,

$$u(x) = u^1(x) + u^2(x), \quad u^1(x) = \sum_{i=1}^N u_i^1(x), \quad u^2(x) = \sum_{i=1}^N u_i^2(x). \tag{3.9}$$

One can see that the two parts $u^1(x)$, $u^2(x)$ are orthogonal with respect to the $a(\cdot,\cdot)$ inner product (1.8),

$$a(u^1(x), u^2(x)) = 0. \tag{3.10}$$

Moreover, the combination of the local bubble parts is small according to (2.12). Specifically, we have

$$\|u^2(x)\|_E \le CH\|f\|_{L^2(D)}. \tag{3.11}$$

Next we prove the proposition 1.

Proof. Denote the numerical solution using $\psi_i(x)$ $(i = 1, \dots m)$ as

$$u^{\psi}(x) = \sum_{i=1}^{m} d_i \psi_i(x),$$

then, there exist c_i $(i = 1, \dots n)$ such that $u_H^{ms} = \sum_{i=1}^{n} c_i \phi_i(x)$, and

$$u_H^{ms}(x)|_{\Gamma} = u^{\psi}(x)|_{\Gamma}. \tag{3.12}$$

Then we consider

$$\|u_H^{ms}(x) - u(x)\|_E^2 = \|u^2(x) + u^1(x) - u_H^{ms}(x)\|_E^2. \tag{3.13}$$

Since $u_H^{ms}(x) \in H_0^1(D)$ is $a(x)$-harmonic on each coarse element D_i, we have

$$\begin{aligned} a(u^2(x), u_H^{ms}(x)) &= \sum_{i=1}^{N} \int_{D_i} \nabla u_i^2(x)^t a(x) \nabla u_H^{ms}(x) \mathrm{d}x \\ &= -\sum_{i=1}^{N} \int_{D_i} u_i^2(x) \mathrm{div}(a(x) \nabla u_H^{ms}(x)) \mathrm{d}x = 0. \end{aligned} \tag{3.14}$$

Thus $u^2(x)$ is a-orthogonal to $u^1(x) - u_H^{ms}(x)$, and according to (3.11) we have

$$\begin{aligned} \|u(x) - u_H^{ms}(x)\|_E^2 &= \|u^2(x)\|_E^2 + \|u^1(x) - u_H^{ms}(x)\|_E^2 \\ &\le \|u^1(x) - u_H^{ms}(x)\|_E^2 + C\|f\|_{L^2(D)}^2 H^2. \end{aligned} \tag{3.15}$$

Then we consider $u_e(x) = u(x) - u^{\psi}(x)$, and decompose it to two parts as we did for $u(x)$ in (3.9),

$$u_e(x) = u_e^1(x) + u_e^2(x), \quad a(u_e^1(x), u_e^2(x)) = 0. \tag{3.16}$$

Consequently, we have

$$\|u(x) - u^{\psi}(x)\|_E^2 = \|u_e^1(x)\|_E^2 + \|u_e^2(x)\|_E^2 \ge \|u_e^1(x)\|_E^2. \tag{3.17}$$

According to (3.12), we have

$$u_e^1(x) = u^1(x) - u_H^{ms}(x), \tag{3.18}$$

since they are equal on Γ and $a(x)$-harmonic on each D_i.

Finally based on (3.15), (3.17), and the optimal property (1.9), we have

$$\|u(x) - u_H^{MS}(x)\|_E^2 \le \|u(x) - u_H^{ms}(x)\|_E^2 \le \|u(x) - u^{\psi}(x)\|_E^2 + C\|f\|_{L^2(D)}^2 H^2. \tag{3.19}$$

This completes the proof. □

As we have shown in (3.11), the bubble part of the solution $u^2(x)$ is small and of $O(H)$ in the energy norm, thus can be neglected if the desired accuracy in the numerical solution is $O(H)$. In our method, we use multi-scale basis functions in (1.5) to approximate the solution space. The multi-scale basis functions are locally $a(x)$-harmonic functions, and are $a(x)$-orthogonal to the bubble part of solution. Due to this $a(x)$-orthogonality and the Galerkin projection formulation in (1.7), multi-scale basis functions only approximate the $a(x)$-harmonic part of the solution and will not bring in additional errors in the bubble part. Thus we can recover the bubble part of solution $u^2(x)$ independently by solving some local bubble problems (2.10). By adding $u^2(x)$ back to $u_H^{MS}(x)$, we can get numerical solution that is free of error in the bubble part. This is one of the advantages of using multi-scale basis in (1.5).

To construct local multi-scale basis functions, we introduce a set of nodal multi-scale basis and decompose the interpolation residual of the $a(x)$-harmonic part of the solution $u^1(x)$ to different parts corresponding to different edges of the coarse mesh, and approximate them separately. This will be detailed in the next subsection.

3.2. **Decomposition of the $a(x)$-harmonic part of the solution.** To identify the compact structure of the $a(x)$-harmonic part of the solution, we first introduce a set of primary interpolation multi-scale basis $\psi_i(x), i = 1, \dots n$ associated with the coarse mesh node points $x_1, x_2, \dots x_n$, which we call *nodal multi-scale basis*,

$$\psi_i(x_j) = \delta_{ij}; \quad -\mathrm{div}(a(x)\nabla\psi_i(x)) = 0, \quad x \in D_j. \tag{3.20}$$

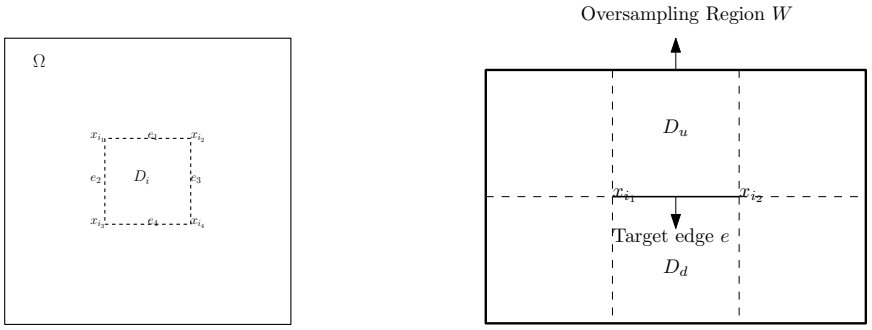
We also require that $\psi_i(x)$ is supported on the four coarse elements around x_i. For example, we can simply choose the multi-scale basis $\psi_i(x)$ to be linear on the boundaries of coarse elements. We will discuss about the optimal choice of these nodal multi-scale basis functions in subsection 3.4.

For $f(x) \in L^2(D)$, and the spatial dimension $d = 2, 3$, we have that $u(x)$ is Hölder continuous on D [23], so we can consider the interpolation of $u^1(x)$, namely the $a(x)$-harmonic part of the solution, using the nodal multi-scale basis functions $\psi_j(x)$, and get the residual,

$$v(x) = u^1(x) - \sum_i u(x_i)\psi_i(x). \tag{3.21}$$

For a coarse mesh element D_i, we denote its four nodes points as x_{i_1}, x_{i_2}, x_{i_3} and x_{i_4}, then we get the restriction of the residual (3.21) on D_i,

$$v(x)|_{D_i} = u_i^1(x) - u(x_i^1)\psi_{i_1}(x) - u(x_{i_2})\psi_{i_2}(x) - u(x_{i_3})\psi_{i_3}(x) - u(x_{i_4})\psi_{i_4}(x). \tag{3.22}$$



(A) A coarse element. (B) Oversampling region.

FIGURE 3. Decomposition of the interpolation residual.

Since the residual $v(x)|_{D_i}$ vanishes on the node points $x_{i_j}, j = 1, 2, 3, 4$, we can decompose the trace of $v(x)$ on ∂D_i to four parts, corresponding to the four edges of D_i, e_i^1, e_i^2, e_i^3, e_i^4, respectively,

$$\begin{aligned} v(x)|_{\partial D_i} &= v(x)|_{e_i^1} + v(x)|_{e_i^2} + v(x)|_{e_i^3} + v(x)|_{e_i^4} \\ &:= v_{e_i^1}(x) + v_{e_i^2}(x) + v_{e_i^3}(x) + v_{e_i^4}(x). \end{aligned} \tag{3.23}$$

This decomposition is illustrated in Figure 3a. Each part in the above decomposition (3.23) belongs to $H^{1/2}(\partial D_i)$ since they vanish on the node points, and we can extend

them to D_i to get four $a(x)$-harmonic components of $v(x)$. We still denote them as $v_{e_i^1}(x)$, $v_{e_i^2}(x)$, $v_{e_i^3}(x)$, $v_{e_i^4}(x)$ and get

$$v(x)|_{D_i} = v_{e_i^1}(x) + v_{e_i^2}(x) + v_{e_i^3}(x) + v_{e_i^4}(x). \tag{3.24}$$

Combining these local decompositions together, we have

$$v(x) = \sum_e v_e(x), \tag{3.25}$$

where $v_e(x)$ is the $a(x)$-harmonic extension of the interpolation error on the edge e to its two neighbor elements. In (3.25), we are actually dividing the interpolation error $v(x)$ in the $a(x)$-harmonic part of solution to different parts corresponding to errors on different edges e. This is possible since $v(x)$ vanishes on the node points, thanks to the interpolation operation using the nodal multi-scale basis $\psi_i(x)$ (3.21).

We seek to construct *edge multi-scale basis functions* on each edge e that approximate $v_e(x)$, and combine them with the nodal multi-scale basis (3.20) to get the whole trial space. We introduce the following operator for the edge e with endpoints x_{i_1} and x_{i_2}, which maps $f(x) \in L^2(D)$ to the interpolation residual,

$$T_e:\ f(x) \in L^2(D) \to v_e(x) = u(x) - u(x_{i_1})\phi_{i_1}(x) - u(x_{i_2})\phi_{i_2}(x) \in H^{1/2}(e). \tag{3.26}$$

The left singular vectors of T_e form the optimal edge multi-scale basis functions. However, T_e is a global operator and its construction involves solving the equation (1.1) globally. In the next section, we decompose T_e as a global solution operator and a local oversampling operator, and construct edge multi-scale basis functions that approximate the range of T_e through the oversampling operator.

3.3. **The oversampling operator.** To identify the compact structure of the solution space restricted on the edge e, we put it in an oversampling region that we denote by W. In our numerical examples, we use the square mesh of size H for the coarse discretization, and the oversampling region W is chosen as the union of the six elements around the edge e. It is illustrated in Figure 3b. Our method is also applicable to other types of discretizations like triangular mesh.

We remark that the idea of identifying the local structure of the solution space by putting it in a larger region, namely *oversampling*, was first proposed in [29] to reduce the resonance error due to artificial local boundary conditions of the multi-scale basis, and this strategy was later employed in [3, 6, 13].

We denote T_W as the operator that maps $f(x) \in L^2(D)$ to the oversampling solution $u_W(x) = u(x)|_W$, and $T_{W\to e}$ as the operator that maps $u_W(x)$ to the solution restricted on the edge e:

$$T_W:\ f(x) \to u_W(x) = u(x)|_W, \quad T_{W\to e}:\ u_W(x) \to u_e(x) = u_W(x)|_e. \tag{3.27}$$

We also introduce the interpolation residual operator using (3.20), P_e,

$$P_e:\ u(x)|_e \to u(x)|_e - u(x_{i_1})\psi_{i_1}(x) - u(x_{i_2})\psi_{i_2}(x). \tag{3.28}$$

With the above definitions, the operator T_e (3.26) can be decomposed as

$$T_e = P_e T_{W\to e} T_W. \tag{3.29}$$

We call the operator $P_e T_{W\to e}$ in the above decomposition (3.29) the oversampling operator, which maps the solution on W, $u_W(x)$ to the interpolation residual,

$$P_{OS} = P_e T_{W\to e}:\ u_W(x) \to v_e(x) = u_W(x) - u_W(x_{i_1})\psi_{i_1}(x) - u_W(x_{i_2})\psi_{i_2}(x), \tag{3.30}$$

where x_{i_1} and x_{i_2} are the two endpoints of e, and $\psi_{i_j}(x)$ are the nodal multi-scale basis functions (3.20).

We employ the compactness of the oversampling operator (3.30) to construct basis functions in $H^{1/2}(e)$ that vanish at x_{i_1} and x_{i_2}, and approximate the range of (3.26). To be specific, we use the first several left singular vectors of P_{OS} as the edge multi-scale basis functions associated with e. We first introduce appropriate inner products for the domain and range space of P_{OS}.

On the edge e, the image of T_e, $v_e(x) \in H^{1/2}(e)$ and vanishes on the two endpoints. We consider its $a(x)$-harmonic extension to the upper and lower coarse elements respectively, as shown in Figure 3b, and denote them as $v_e^u(x)$ and $v_e^d(x)$. Then we define

$$\|v_e(x)\|^2_{H^{1/2}(e)} = \frac{1}{2}\int_{D_u} (\nabla v_e^u)^t a(x) \nabla v_e^u \mathrm{d}x + \frac{1}{2}\int_{D_d} (\nabla v_e^d)^t a(x) \nabla v_e^d \mathrm{d}x. \tag{3.31}$$

In the domain of the operator $P_e T_{W\to e}$, namely, $u_W(x)$, we define its inner product as

$$\|u_W(x)\|^2_{V_W} = \int_W \nabla (u_W^1)^t a(x) \nabla u_W^1 + (u_W^1)^2 \mathrm{d}x + \int_W [\mathrm{div}(a(x)\nabla u_W)]^2, \tag{3.32}$$

where $u_W^1(x)$ is the local $a(x)$ harmonic part of the solution u_W. And we denote V_W as the Hilbert space of functions on W, which have bounded norm (3.32).

With the above inner products, we compute the singular value decomposition of the oversampling operator P_{OS}. To discretize the domain of P_{OS}, $u_W(x)$, we consider its two parts, the $a(x)$-harmonic part, and the bubble part. The $a(x)$-harmonic part only depends on the trace of $u_W(x)$ on ∂W, and we discretize $H^{1/2}(\partial W)$ using all the fine mesh piecewise linear functions. If ∂W intersects with ∂D, then we will only use fine mesh basis functions that vanish on ∂D. The bubble part of the solution $u_W^2(x)$ only depends on $f_W(x) = f(x)|_W$, and we discretize $f(x)$ using piecewise constant functions on the coarse mesh, which can justified by assuming certain regularity of $f(x)$.

With the above discretization of $u_W(x)$, we truncate the singular values of P_{OS} to $O(H)$ and select the corresponding left singular vectors as the boundary basis functions. We denote them as

$$v_e^1(x), \dots v_e^{k_e}(x) \in H^{1/2}(e), \tag{3.33}$$

which vanish on the two endpoints of e. According to the oversampling operator in (3.29), and our truncation criteria, we have the following approximation property

$$\inf_{c_i^e} \|u(x)|_e - u(x_{i_1})\phi_{i_1}(x) - u(x_{i_2})\phi_{i_2}(x) - \sum_{i=1}^{e_k} c_i^e v_e^i(x)\|_{H^{1/2}(e)} \le CH\|u_W^1(x)\|_{V_W}. \tag{3.34}$$

Then we extend these boundary basis functions to the two neighbourhood coarse elements D_u and D_d as $a(x)$-harmonic functions to get the edge multi-scale basis, by solving

$$\begin{cases} -\mathrm{div}(a(x)\nabla \phi_e^k(x)) = 0, x \in D_u, D_d, \\ \phi_e^k(x)|_e = v_e^k(x). \end{cases} \tag{3.35}$$

Finally, we combine the edge multi-scale basis functions for each edge e,

$$\phi_e^i(x), \quad i = 1, 2 \dots k_e, \tag{3.36}$$

with the nodal multi-scale basis functions $\psi_i(x)$ (3.20), and get the trial space,

$$V_H = \text{span}\{\phi_i(x), \quad i = 1, \dots n\}. \tag{3.37}$$

We have the following error estimate using the trial space (3.37).

Proposition 2. *Using the trial space consisting of the nodal multi-scale basis functions* (3.20) *and the edge multi-scale basis functions* (3.35)*, we obtain numerical solution to* (1.1) *using the Galerkin projection* (1.7)*. Then we have the following convergence property,*

$$\|u(x) - u_H^{MS}(x)\|_E \leq CH\|f(x)\|_{L^2}. \tag{3.38}$$

Remark 1. Using a simple Aubin-Nitsche duality argument and the convergence result in the energy norm (3.38), we can get the onvergence in L^2.

$$\|u(x) - u_H^{MS}(x)\|_{L^2(D)} \leq CH^2\|f(x)\|_{L^2}. \tag{3.39}$$

The proof of the convergence result (3.38) follows directly from the decomposition of the solution operator (3.29) and the truncation in the singular value decomposition of the oversampling operator.

Proof. We choose c_e^j as the ones in (3.34), and denote

$$u_H^{ms}(x) = \sum_{i=1}^{n} u(x_i)\psi_i(x) + \sum_e \sum_{j=1}^{k_e} c_e^j \phi_e^j(x) \in V_H. \tag{3.40}$$

We consider $\|u_H^{ms}(x) - u(x)\|_E$. Since the basis functions in (3.37) are multi-scale basis, namely, they are $a(x)$-harmonic on each D_i, we have

$$\|u_H^{ms} - u\|_E^2 = \|u_H^{ms} - u_1 - u_2\|_E^2 \leq \|u_H^{ms}(x) - u_1(x)\|_E^2 + CH^2\|f\|_{L^2(D)}^2, \tag{3.41}$$

where $u_1(x)$ and $u_2(x)$ are the $a(x)$-harmonic part and bubble parts of the solution.

Then we decompose $\|u_H^{ms}(x) - u_1(x)\|_E^2$ to different parts on D_i. For each part, according to the approximation property (3.34), and the definition (3.31), we have

$$\int_{D_i} \nabla(u_H^{ms}(x) - u_1(x))^t a(x) \nabla(u_H^{ms}(x) - u_1(x)) \mathrm{d}x \leq C \sum_W H^2 \|u_W(x)\|_{V_W}. \tag{3.42}$$

The sum over W corresponds to the oversampling regions for edges of D_i. There are four of them for each D_i. Summing up (3.42) for all the coarse elements of D, we have

$$\|u_H^{ms}(x) - u_1(x)\|_E^2 \leq CC_1H^2(\|u(x)\|_E^2 + \|u(x)\|_{L^2(D)^2} + \|f(x)\|_{L^2(D)}), \tag{3.43}$$

where C_1 depends on the size the oversampling region W.

Substituting (3.43) into (3.41), and using (1.3) we have

$$\|u(x) - u_H^{ms}(x)\|_E \leq CH\|f(x)\|_{L^2(D)}. \tag{3.44}$$

Then using the optimal approximation property (1.9), we finish the proof. □

To make the number of multi-scale basis functions in (1.5) small, we want the singular values of $P_{OS} = P_e T_{W \to e}$ decay fast. P_{OS} can be decomposed to two parts: the first part acts on the $a(x)$-harmonic part of $u_W(x)$, and we denote it as P_{OS}^1; the second part acts on the bubble that only depends on $f_W(x)$, and we denote it as P_{OS}^2. For the first part, similar analysis has been done in [6] in a slightly different setting, and the method there also applies to our problem. We have the following result.

Proposition 3. *Denote the singular values of P_{OS}^1 that acts on $a(x)$-harmonic functions on W as σ_k, then for our choice of the oversampling domain, we have the following upper bound on the decay of σ_k: for any $\epsilon > 0$, there exist C such that*

$$\sigma_k \leq C \exp\{-k^{1/(d+1)-\epsilon}\}, \tag{3.45}$$

where d is the dimension of the domain D.

The second part P_{OS}^2 is small according to (2.12), and the decay rate of its singular values can be obtained from (2.5) using a simple scaling argument.

Remark 2. In our definition of (3.32), we take into account the fact that $f(x) \in L^2(D)$. If we choose $\|\cdot\|_{V_W}$ as the $H^1(W)$ norm, then P_{OS} will not be well-defined due to the lack of continuity of $u_W(x)$ for $d \geq 2$.

Remark 3. We will numerically investigate how does the decay of the singular values of P_{OS}^1 depend on the size of the oversampling domain in section 4.1. We will see that for larger oversampling region, the singular values of P_{OS}^2 will be smaller. However, in our error estimate from (3.42) to (3.43), we sum up the errors on different coarse meshes, and since the oversampling regions for different edges have overlapping, the constant C_1 in (3.43) will not be 1. For larger oversampling region, C_1 will be larger, and the error estimates deteriorates. Thus there exists a trade-off in choosing the oversampling size.

We will see in our numerical results section that the singular values of P_{OS} decay very fast, and a very small number of edge basis functions can achieve high local approximation accuracy.

As we have shown previously, the basis functions we obtain are multi-scale basis functions, thus are $a(x)$-orthogonal to the bubble part of the solution space. This gives us the flexibility to add the bubble parts back to the numerical solutions at local regions where higher accuracy is desired by simply solving some local bubble problems (2.10). In our truncation of the singular values of the local compact operator $P_e T_{W\to e}$, we choose the threshold to be $O(H)$, since $O(H)$ accuracy is required in (3.38). If we need higher accuracy than H, for example, $O(\epsilon)$ with $h \ll \epsilon \ll H$, then we can truncate the singular values of $P_e T_{W\to e}$ by ϵ. By doing so, the resulting multi-scale basis functions are able to approximate the $a(x)$-harmonic part of the solution space up to $O(\epsilon)$ accuracy. Then by adding back the bubble part of the solution $u^2(x)$ to the numerical solution $u_H^{MS}(x)$,

$$u_H(x) = u_H^{MS}(x) + u_2(x), \tag{3.46}$$

we can get $O(\epsilon)$ accuracy in our final numerical solutions,

$$\|u(x) - u_H(x)\|_{H_0^1(D)} \leq C\epsilon \|f(x)\|_{L^2(D)}, \quad \|u(x) - u_H(x)\|_{L^2(D)} \leq C\epsilon^2 \|f(x)\|_{L^2(D)}. \tag{3.47}$$

Namely, our upscaling strategy allows us to get very high accuracy that is permitted by the fine mesh discretization.

3.4. **Optimal nodal interpolation basis functions.** In constructing the multi-scale basis functions in the previous section, we need to choose a set of nodal basis functions (3.20) first, which allows us to reduce the problem to approximating the solution space restricted on each edge e, (3.29). The choice of these nodal basis functions will affect the oversampling operator $P_{OS} = P_e T_{W\to e}$. In this subsection, we identify the optimal nodal basis functions by solving local under-determined least square problems.

The oversampling operator for edge e, P_{OS}, depends on the nodal multi-scale basis functions $\psi_{i_1}(x)$ and $\psi_{i_2}(x)$ associated with the two endpoints of e, and we seek optimal nodal multi-scale basis functions $\psi_{i_1}(x)$ and $\psi_{i_2}(x)$, such that the singular values of P_{OS} have the fastest decay.

We consider the following optimization problem

$$\min_{\phi_{i_1}(x)} \|\psi_{i_1}(x)\|_{V_W}, \quad \min_{\phi_{i_2}(x)} \|\psi_{i_2}(x)\|_{V_W}, \quad \text{subject to} \tag{3.48a}$$

$$-\mathrm{div}(a(x)\nabla\psi_{i_j}(x)) \in L^2(W),\ x \in W, \quad \psi_{i_j}(x_{i_k}) = \delta_{jk}. \tag{3.48b}$$

where the norm $\|\cdot\|_{V_W}$ is defined in (3.32). We use the solution to (3.48), $\psi_{i_1}(x)|_e$ and $\psi_{i_2}(x)|_e$ as boundary conditions to construct the nodal multi-scale basis. The resulting nodal basis functions are optimal and we have the following theorem.

Theorem 3.2. *The optimization problem (3.48) has unique solutions. Let $\phi^*_{i_1}(x)$, $\phi^*_{i_2}(x)$ be the solution to (3.48), and $\phi_{i_1}(x)$ and $\phi_{i_2}(x)$ be two other nodal multi-scale basis functions. Denote the corresponding oversampling operators using these nodal basis functions as (3.30) as P^*_{OS} and P_{OS}, and let $\sigma^*_i, i = 1, 2, \dots$ and $\sigma_i, i = 1, 2, \dots$ be their singular values in descending order. Then we have*

$$\sigma^*_i \le \sigma_i. \tag{3.49}$$

Namely, using the nodal multi-scale basis obtained from (3.48), the singular values of the oversampling operator (3.30) have the fastest decay.

The proof of Theorem 3.2 is given in the appendix.

Note that in the minimization problem (3.48), the two nodal basis functions can be constructed independently by solving under-determined least square problems. Going over the oversample regions for each edge of the coarse mesh, we can construct the optimal nodal multi-scale basis functions (3.20) on all the boundaries of local regions. Then we extend them locally to $a(x)$-harmonic functions by solving some local boundary value problems as in (3.35) to get the nodal basis (3.20).

We will see in our numerical results section that, in some cases, the optimal nodal multi-scale basis functions (3.48) are enough to obtain good approximation property to the solution space of (1.1). Namely, there is no need to construct the edge multi-scale basis functions (3.35) and add them to the trial space.

3.5. **Implementation of the whole method.** The proposed multi-scale finite element method consists of two stages, the offline stage and the online stage. In the offline stage, we identify the compact structure of the solution space. In the online stage, for a given forcing function, we compute the numerical solution using the offline basis functions. The offline stage involves the following procedures.

1. Build the oversampling operator for each edge e of the coarse mesh.

   On each oversampling region W, we build the local $a(x)$-harmonic extension operator, which maps the boundary condition which belongs to $H^{1/2}(\partial W)$ to $a(x)$-harmonic functions on W. This step requires solving a series of boundary value problems on W. Then we discretize the local forcing function using piecewise constant functions on the coarse mesh. The above discretizations correspond to the two parts of V_W (3.32), which is the domain of P_{OS}. With this we construct the optimal nodal multi-scale basis by solving (3.48), and build the oversampling operator (3.3).
2. Compute the edge multi-scale basis using the oversampling operator.

Using the inner products (3.32) and (3.31) in the oversampling operator to compute its singular value decomposition. We truncate the singular values to ϵ, and save the corresponding left singular vectors, which are basis functions of $H^{1/2}(e)$, $v_e^1(x), v_e^2(x), \dots v_e^{k_e}(x)$. Then we extend these boundary basis functions to the two neighborhood coarse elements of e by solving local boundary value problems (3.35). Combining the nodal multi-scale basis functions with the edge multi-scale basis functions, we get the trial space

$$V_H = \text{span}\{\phi_1(x), \phi_2(x), \dots \phi_n(x)\}. \tag{3.50}$$

3. Compute the stiffness matrix.

   We save the multi-scale basis in (3.50), and compute the stiffness matrix,

$$M(i,j) = a(\phi_i(x), \phi_j(x)). \tag{3.51}$$

The online stage involves the following procedures.

1. Compute the load vector.

   For a given forcing function $f(x) \in L^2(D)$, we compute the corresponding load vector

$$b(i) = \int_D \phi_i(x) f(x) \mathrm{d}x, \quad i = 1, \dots n. \tag{3.52}$$

2. Compute the online numerical solution.

   Using load vector (3.52) and stiffness matrix (3.51), we solve linear system

$$Mc = b, \tag{3.53}$$

   and get the online numerical solution on the fine mesh

$$u_H^{MS}(x) = \sum_{i=1}^{n} c_i \phi_i(x). \tag{3.54}$$

   Recall that the edge multi-scale basis functions vanish on the coarse grid node points. So if we only want the coarse-mesh solution, we can simply select the coefficients in c corresponding to the nodal multi-scale basis functions (3.20).

3. Recover the bubble part of the solution.

   Solve the local boundary value problem (2.10) on each coarse mesh element D_i, and get $u_i^2(x)$. Combining the local bubbles together and adding them to the Galerkin solution (3.54), we get

$$u_H(x) = u_H^{MS}(x) + \sum_{i=1}^{N} u_i^2(x). \tag{3.55}$$

   If $O(H)$ accuracy is required in the numerical solution, this step is unnecessary since the bubble part does not impact the large scale properties of the solution.

Note that in the offline stage, we need to solve a series of boundary value problems for each edge of the coarse mesh to construct the oversampling operator (3.30), and then compute its singular value decomposition, which is relatively expensive. However, the constructions of edge multi-scale basis functions on different edges are independent, thus the offline stage can be implemented on a parallel machine to accelerate the computation. In the online stage, the main computational cost comes from solving the linear system (3.53). Our numerical results in the next section suggest that a very small number of multi-scale basis functions are enough to obtain the coarse mesh accuracy, $O(H)$, thus the linear system (3.53) is small and

sparse. This implies that the online computation in our method is efficient, and our method can bring in significant computational savings in the multi-query setting, where the equation (1.1) needs to be solved for multiple times using different forcing functions.

4. **Numerical results.** In this section, we present numerical examples that have multiple-scale features and high-contrast channels to demonstrate the capacity of our method in identifying and exploiting the compact structure of the local solution space to achieve computational savings in the online stage. We discretize the domain of the problems $D = [0,1] \times [0,1]$ using a two-level mesh as shown in Figure 2. The coarse mesh is of size $H = 1/32$, and the fine mesh is of size $h = 1/1024$.

4.1. **An example with multiple spatial scales.** The first example we consider is one that has multiple spatial scales. The coefficient is given by (2.14), and it is visualized in Figure 1a. For each edge of the coarse mesh, we compute the singular value decomposition of the oversampling operator, and truncate the singular values at $\epsilon = H$, which guarantees $O(H)$ accuracy in the online numerical solution. After the offline stage, multi-scale basis functions (1.5) are constructed, and the average number of edge multi-scale basis functions associated with each edge is

$$\bar{k}_e = \frac{\sum_e k_e}{\#(e)} \approx 1.00. \tag{4.1}$$

We see that $\bar{k}_e$ is very small. Actually only 1 or 2 edge multi-scale basis functions are constructed for each edge of the coarse mesh, in addition to the nodal multi-scale basis functions. And this implies the efficiency of our method in the online stage since the resulting stiffness matrix is small and sparse.

To measure the error in our online numerical solution, we need to choose a reference solution. Since the multi-scale basis functions are constructed and saved on the fine mesh of size h, we will use the piecewise linear finite element solution on the fine mesh as the reference.

In the online stage, we choose the forcing function $f(x, y)$ to be

$$f(x, y) = 1, \quad (x, y) \in D. \tag{4.2}$$

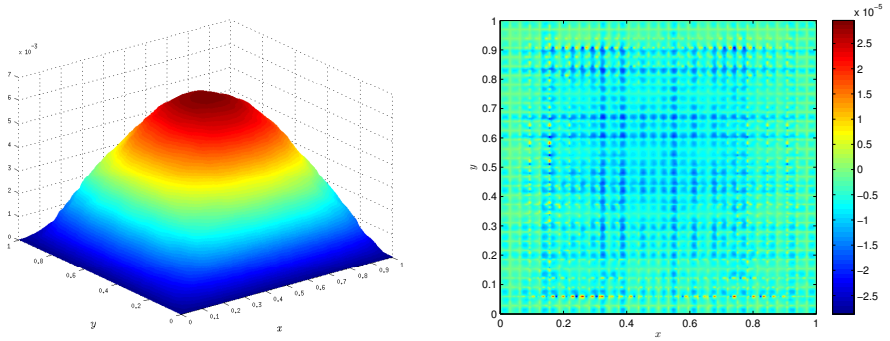
Recall that the basis functions that we use are $a(x)$-harmonic in each D_i, and we can add back the bubble part of the solution by simply solving some local cell problems. Our online numerical solution $u_H^{MS}(x)$, and the error $u(x) - u_H^{MS}(x)$ are plotted in Figure 4. We can see that our method achieves high accuracy and the numerical error in the online numerical solution is very small.

We measure the error of the numerical solution in the energy norm and L^2 norm. We denote the numerical solution as $u_H^{MS}(x)$, and the corrected solution using the bubble as $u_H(x)$ (3.55). We compute

$$E_E^{MS} = \frac{\|u(x) - u_H^{MS}(x)\|_E}{\|u(x)\|_E}, \qquad E_E = \frac{\|u(x) - u_H(x)\|_E}{\|u(x)\|_E}, \tag{4.3a}$$

$$E_{L^2}^{MS} = \frac{\|u(x) - u_H^{MS}(x)\|_{L^2(D)}}{\|u(x)\|_{L^2(D)}}, \qquad E_{L^2} = \frac{\|u(x) - u_H(x)\|_{L^2(D)}}{\|u(x)\|_{L^2(D)}}. \tag{4.3b}$$

The results are listed in Table 1. We can see that by adding the bubble part of solution back to the numerical solution, the error in L^2 norm and energy norm are both reduced by about one half. This implies the numerical error in the $a(x)$-harmonic part of the solution is about the same as that in the bubble part. Since



(A) The fine mesh solution. (B) Error in our online numerical solution.

FIGURE 4. Online numerical solutions.

	Energy Norm	L^2 Norm	Coarse-space Dimension
$u_H^{MS}(x)$	4.16×10^{-2}	1.73×10^{-3}	2945
$u^H(x)$	2.67×10^{-2}	8.75×10^{-4}	2945 + 1024 local solvers
Interpolation basis	7.57×10^{-2}	5.95×10^{-3}	961

TABLE 1. $u_H^{MS}(x)$ denotes the online numerical solution (3.54); $u_H(x) = u_H^{MS}(x) + u^2(x)$ denotes the corrected numerical solution (3.55); 'Interpolation basis' denotes the numerical solution obtained using only the nodal multi-scale basis (3.20) in the trial space.

the latter is of order $O(H)$ in the energy norm and of order $O(H^2)$ in the L^2 norm, this result confirms our error estimates (3.38) and (3.39).

Then we consider only using the optimal nodal multi-scale basis functions $\psi_i(x), i = 1, \dots N$ in the trial space (3.50), namely, we do not use the edge multi-scale basis functions. The error in the corresponding numerical solution is also listed in Table 1. We can see that the relative error in L^2 is also small, which means the numerical solution can capture the large-scale property of the solution. However, the errors in the energy norm and L^2 norm are both significantly larger than that in (3.54), which implies the necessity of enriching the trial space using the edge multi-scale basis functions when higher accuracy is required.

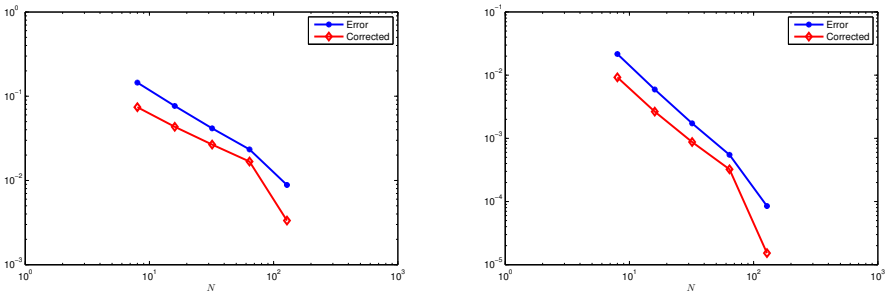
To further demonstrate the convergence rate in (3.38) and (3.39), we consider a sequence of coarse meshes with $H = 2^{-k}, k = 3, 4, 5, 6, 7$. For each H, we compute the error in the online numerical solution, and the decay of the numerical error with respect to H is plotted in Figure 5.

A simple linear regression reveals that

$$\log E_E^{MS} \approx -0.97 \log N + 1.59 \times 10^{-1}, \quad \log E_{L^2}^{MS} \approx -1.94 \log N + 2.93 \times 10^{-1}. \tag{4.4}$$

which agree with the error estimates (3.38) and (3.39).

To demonstrate that our method can achieve higher accuracy than $O(H)$, we consider truncating the singular values of the oversampling operator (3.30) using different ϵ, ranging from H to h, and adding the bubble part of the solution back to our numerical solution. The numerical errors in $u_H(x)$ (3.55) decay with ϵ, and it is



(A) Energy norm. (B) L^2 norm.

FIGURE 5. Convergence of the online numerical solution with $N = 1/H$. The line labeled 'Error' corresponds to the multi-scale numerical solution $u_H^{MS}(x)$ (3.54). The line label 'corrected' corresponds to the numerical solution corrected by the bubble part, $u_H(x) = u_H^{MS}(x) + u^2(x)$ (3.55). The slope of the lines in the left plot is approximately -0.97, while the slope of the lines in the right is approximately -1.94

plotted in Figure 6. We can see that the error in energy norm decays linearly with ϵ, and the error in L^2 norm decays quadratically with ϵ, which agrees with (3.47). We comment that the numerical error is measured using the fine mesh solution as the reference, so for ϵ close to h, the error in our numerical solution is actually dominated by the fine mesh discretization error.

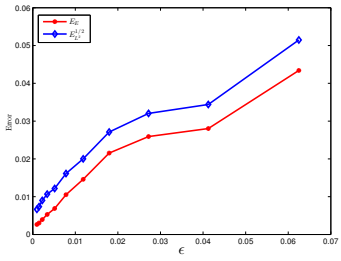
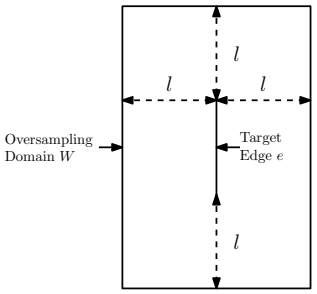


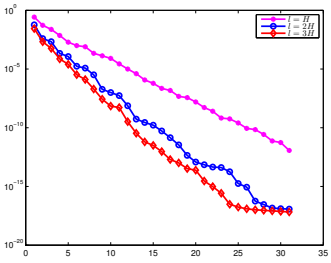
FIGURE 6. Numerical errors in energy norm and L^2 norm of the numerical solution (3.54) for different ϵ. The line labeled E_E corresponds to the error in the energy norm, while the line labeled $E_{L^2}^{1/2}$ corresponds to the square root of the error in L^2 norm. We can see that the energy norm of the error decays linearly with ϵ, and the L^2 norm of the error decays quadratically with ϵ.

Then we numerically investigate how does the size of oversampling domain W affect the decay of the singular values of the oversampling operator P_{OS}. We only consider the first part of P_{OS}, namely, P_{OS}^1 which maps the $a(x)$-harmonic part of the solution $u_W^1(x)$ to the interpolation residual, since the singular values of P_{OS}^2 is well-known according to (2.5). We consider a vertical edge e of the coarse mesh

which lies in the center of the domain D, and let l be the distance between ∂W and e. The choice of W based on l is illustrated in Figure 7a. For different choices of l, the decay of the singular values of P_{OS} is shown in Figure 7b. We can see that the singular values of P_{OS}^1 decay exponentially fast, and for larger oversampling domain W, the singular values of P_{OS}^1 is smaller. However, as we argued in Remark 3, for larger W, the constant C in the error estimates (3.38), (3.39) will be larger, and there is a tradeoff in choosing the oversampling size l. We simply choose $l = H$ in our numerical results.



(A) The oversampling region.



(B) Decay of the singular value of P_{OS}^1.

4.2. **An example without scale-separation.** In this subsection, we consider an example where the coefficient $a(x)$ has no scale-separation,

$$a(x, y) = |\tilde{a}| + 0.5. \tag{4.5}$$

The values of $\tilde{a}$ on the node points of an intermediate mesh of size $\frac{1}{128}$ are independent standard Gaussian random variables. And $a(x, y)$ is piecewise linear on the same mesh. For a typical realization of $\tilde{a}$, the coefficient $a(x, y)$ is rough and has no clear scale-separation. One realization of the coefficient (4.5) is illustrated in Figure 8.

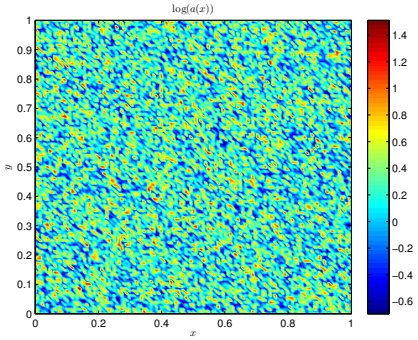


FIGURE 8. The rough coefficient $a(x)$ without scale-separation.

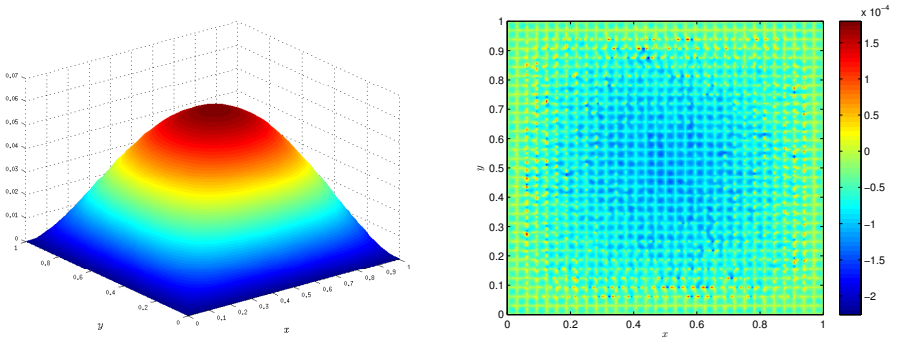
We discretize the spatial domain D using a two-level mesh as shown in Figure 2, and then solve the optimization problem (3.48) and build the oversampling operator (3.30) on the fine mesh. We truncate the singular value decomposition of the

	Energy Norm	L^2 Norm	Coarse-space Dimension
$u_H^{MS}(x)$	4.44×10^{-2}	1.98×10^{-3}	2945
$u_H(x)$	3.17×10^{-2}	1.18×10^{-3}	2945 + 1024 local solvers

TABLE 2. Errors in the online numerical solution $u_H^{MS}(x)$ (3.54), and the numerical solution corrected by the local bubble parts, $u_H(x) = u_H^{MS}(x) + u^2(x)$ (3.55).

oversampling operator to $\epsilon = H$. After the offline stage, the average number of edge multi-scale basis functions associated with each edge of the coarse mesh, (4.1), is $\bar{k}_e \approx 1.00$. The smallness of $\bar{k}_e$ reflects the compactness of the local solution space.

In the online stage, we choose the forcing function $f(x)$ to be same as (4.2), and measure the error of the online numerical solution using the fine mesh solution as reference. The numerical solutions are plotted in Figure 9. We can see that the errors in the online numerical solutions are small. We measure the error (4.3) in the energy norm and the L^2 norm. The results are summarized in Table 2. Again we see that our method achieves very high accuracy in the online stage, which reflects that the good performance of our method does not depend on the scale-separation of the coefficient.



(A) The fine mesh solution. (B) Error in the numerical solution $u_H^{MS}(x)$.

FIGURE 9. Online numerical solutions.

4.3. **An example with high-contrast channels.** In this subsection, we consider an example with high-contrast channels. The high contrast in the coefficient violates our uniform ellipticity assumption (1.2), and brings in additional difficulty. The coefficient that we consider here is the one with multiple scales (2.14) added with some high conductivity patches and channels. $\log_{10} a(x)$ is plotted in Figure 10, which has very strong heterogeneity.

We discretize the problem in the spatial direction as the previous two examples, and build the oversampling operator for each edge e of the coarse mesh. We truncate the singular value decomposition of the oversampling operators to $\epsilon = H$. The average number of edge multi-scale basis functions associated with each edge is $\bar{k}_e = 0.89$ (4.1). Namely, on average, we use less than one edge multi-scale basis function for each edge of the coarse mesh, which reflects the compactness of the solution space on local regions of the domain for this problem with high contrast channels. In the online stage, we choose the forcing function to be (4.2). The

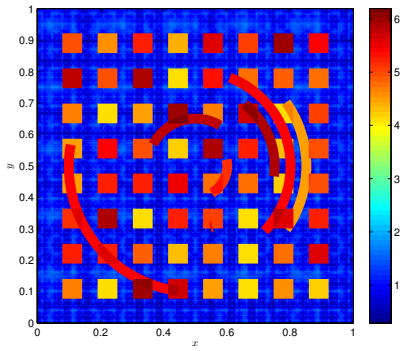
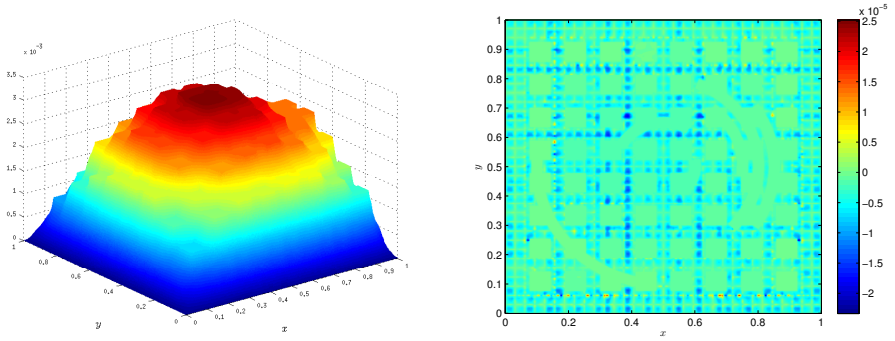


FIGURE 10. $\log_{10} a(x)$. The coefficient with high-contrast channels.

	Energy Norm	L^2 Norm	Coarse-space Dimension
$u_H^{MS}(x)$	3.67×10^{-2}	1.64×10^{-3}	2727
$u_H(x)$	2.02×10^{-2}	6.13×10^{-4}	2727 + 1024 local solvers

TABLE 3. Errors in the numerical solution $u_H^{MS}(x)$ (3.54), and the corrected numerical solution $u_H(x) = u_H^{MS}(x) + u^2(x)$ (3.55).

numerical errors are plotted in Figure 11. We can see that our numerical solutions have high accuracy and can capture the large-scale properties of the solution.



(A) The fine mesh solution. (B) Error in the numerical solution $u_H^{MS}(x)$.

FIGURE 11. Online numerical solutions.

The numerical errors of the solutions (4.3) are listed in Table 3. Again, we see that we obtain high accuracy for our online numerical solution.

5. **Concluding remarks.** In this paper, a novel multi-scale finite element method is proposed, which is based on the compactness of the solution space restricted on local regions of the domain, and does not depend on any scale-separation or periodicity assumption of the coefficient. We introduced a special type of basis functions, namely, the *multi-scale* basis, which are harmonic on each coarse element, and

showed that multi-scale basis is optimal in approximating the solution for fixed local boundary conditions. By introducing a primary set of nodal multi-scale basis functions, we reduce our problem to approximating the interpolation residual of solution space on each edge of the coarse mesh. Then we construct edge multi-scale basis functions for each edge of the coarse mesh separately employing an oversampling operator which is local and compact. The optimal nodal multi-scale basis is also identified as the solution of some under-determined least square problems. Numerical results suggest that our method can achieve high efficiency and accuracy for the challenging problems without scale-separation, or having high-contrast inclusions.

Appendix: Proof of Theorem 3.2. Recall that the restriction operator $P_{W\to e}$ (3.27) maps $u(x)|_W$ to $H^{1/2}(e)\cap C^{\alpha}(e)$,

$$P_{W\to e}:\ u_W(x)\to u_e(x)\in H^{1/2}(e)\cap C^{\alpha}(e). \tag{5.1}$$

We introduced in (3.32) the following inner product on the domain of $T_{W\to e}$, which is denoted by V_W,

$$\|u_W(x)\|^2_{V_W}=a(u^1_W(x),u^1_W(x))+\|u^1_W(x)\|^2_{L^2(W)}+\|\mathrm{div}(a(x)\nabla u_W(x))\|^2_{L^2(W)}. \tag{5.2}$$

We consider P_1 and P_2, which are the bounded linear functionals that map $u_W(x)\in V_W$ to its values on x_{i_1} and x_{i_2} respectively,

$$P_1:\ u_W(x)\to u_W(x_{i_1}),\quad P_2:\ u_W(x)\to u_W(x_{i_2}). \tag{5.3}$$

The two operators P_1 and P_2 are linearly independent, thus there exist $\phi_{i_1}(x),\phi_{i_2}(x)\in V_W$, s.t.

$$P_j(\phi_{x_{i_k}}(x))=\delta_{jk},\quad j,k=1,2. \tag{5.4}$$

We denote the intersection of the kernels of P_1 and P_2 as V^0_W, which is a closed subspace of V_W. And we denote the projection of $\psi_{i_1}(x)$ and $\psi_{i_2}(x)$ to the orthogonal complement of V^0_W, $(V^0_W)^{\perp}$, by $\psi^*_{i_1}(x)$ and $\psi^*_{i_2}(x)$, namely,

$$\psi_{i_1}(x)-\psi^*_{i_1}(x),\ \psi_{i_2}(x)-\psi^*_{i_2}(x)\in V^0_W,\quad \psi^*_{i_1}(x),\psi^*_{i_2}(x)\perp V^0_W, \tag{5.5}$$

where the orthogonality is in the sense of (5.2). Then due to the definition of V^0_W, we have

$$\psi^*_{i_j}(x_{i_k})=\delta_{jk},\ \|\psi_{i_j}(x)\|^2_{V_W}=\|\psi^*_{i_j}(x)\|^2_{V_W}+\|\psi_{i_j}(x)-\psi^*_{i_j}(x)\|^2_{V_W},\ j,k=1,2. \tag{5.6}$$

Using (5.6), we can prove that the optimization problem (3.48) has unique solutions $\psi^*_{i_1}(x)$ and $\psi^*_{i_2}(x)$.

Next we prove property (3.49). We choose $\psi^*_{i_1}(x)|_e$ and $\psi^*_{i_2}(x)|_e$ as the interpolation basis in (3.3), and get the oversampling operator P^*_{OS}. For any other two interpolation basis functions, $\psi_{i_1}(x)$ and $\psi_{i_2}(x)$, we denote the corresponding oversampling operator as P_{OS}. Denote the singular values of P^*_{OS} and P_{OS} as σ^*_k and σ_k respectively, and we will prove that $\sigma^*_k\le\sigma_k$. We achieve this by showing that P^*_{OS} and P_{OS} are equal on the subspace V^0_W, while P^*_{OS} vanishes on the orthogonal complement of V^0_W.

We use the following characterization of singular values,

$$\sigma_k=\sup_{V_k\subset V_W}\ \inf_{\substack{u(x)\in V_k\\ \|u(x)\|_{V_W}=1}}\|P_{OS}u(x)\|_{H^{1/2}(e)}. \tag{5.7}$$

where V_k runs over k-dimensional subspace of V_W.

For any $u(x) \in V_W$, we consider its projection to V_W^0 and $(V_W^0)^\perp$, and denote them as $u_1(x)$, $u_2(x)$ respectively,

$$u(x) = u_1(x) + u_2(x). \tag{5.8}$$

Then according to the definition of V_W^0 and our choice of the interpolation basis, we have

$$P_{OS}^*(u_1(x)) = P_{W\to e}(u_1(x)), \quad P_{OS}^* u_2(x) = 0. \tag{5.9}$$

Then according to (5.9), we get

$$\begin{aligned}\sigma_k^* &= \sup_{V_k \subset V_W^0} \inf_{\substack{u(x)\in V_k \\ \|u(x)\|_{V_W}=1}} \|P_{OS}^* u(x)\|_{H^{1/2}(e)} \\ &= \sup_{V_k \subset V_W^0} \inf_{\substack{u(x)\in V_k \\ \|u(x)\|_{V_W}=1}} \|P_{W\to e}(u(x))\|_{H^{1/2}(e)}, \end{aligned} \tag{5.10}$$

where V_k runs over k-dimensional subspace of V_W^0. In the first equality, we have used $P_{OS}^* u_2(x) = 0$, while in the second equality we have used $P_{OS}^* u_1(x) = P_{W\to e}(u_1(x))$.

And for σ_k, we have

$$\begin{aligned}\sigma_k &\geq \sup_{V_k \subset V_W^0} \inf_{\substack{u(x)\in V_k \\ \|u(x)\|_{V_W}=1}} \|P_{OS} u(x)\|_{H^{1/2}(e)} \\ &= \sup_{V_k \subset V_W^0} \inf_{\substack{u(x)\in V_k \\ \|u(x)\|_{V_W}=1}} \|P_{W\to e}(u(x))\|_{H^{1/2}(e)}, \end{aligned} \tag{5.11}$$

where V_k runs over k-dimensional subspace of V_W^0. The inequality is due to that we restrict V_k to be a subspace of $V_W^0 \subset V_W$.

Using (5.10) and (5.11), we finish the proof of Theorem 3.2.

REFERENCES

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

[2] G. Allaire, *Shape optimization by the homogenization method*, vol. 146, Springer Science & Business Media, 2012.

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

[4] I. Babuška, G. Caloz and J. Osborn, Special finite element methods for a class of second order elliptic problems with rough coefficients, *SIAM Journal on Numerical Analysis*, **31** (1994), 945–981.

[5] I. Babuška, X. Huang and R. Lipton, Machine computation using the exponentially convergent multiscale spectral generalized finite element method, *ESAIM: Mathematical Modelling and Numerical Analysis*, **48** (2014), 493–515.

[6] I. Babuska and R. Lipton, Optimal local approximation spaces for generalized finite element methods with application to multiscale problems, *Multiscale Modeling & Simulation*, **9** (2011), 373–406.

[7] I. Babuška and J. Osborn, Generalized finite element methods: Their performance and their relation to mixed methods, *SIAM Journal on Numerical Analysis*, **20** (1983), 510–536.

[8] I. Babuška and J. Osborn, Can a finite element method perform arbitrarily badly?, *Mathematics of Computation of the American Mathematical Society*, **69** (2000), 443–462.

[9] A. Bensoussan, J. Lions and G. Papanicolaou, *Asymptotic Analysis for Periodic Structures*, vol. 374, American Mathematical Soc., 2011.

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

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

[12] C. Chu, I. Graham and T. Y. Hou, A new multiscale finite element method for high-contrast elliptic interface problems, *Mathematics of Computation*, **79** (2010), 1915–1955.
[13] J. Chu, Y. Efendiev, V. Ginting and T. Y. Hou, Flow based oversampling technique for multiscale finite element methods, *Advances in Water Resources*, **31** (2008), 599–608.
[14] M. Ci, T. Y. Hou and Z. Shi, A multiscale model reduction method for partial differential equations, *ESAIM-Mathematical Modelling and Numerical Analysis*, **48** (2014), 449–474.
[15] D. Cioranescu and P. Donato, Introduction to homogenization.
[16] E. De Giorgi, New problems in γ-convergence and g-convergence, *Free boundary problems*, **2** (1980), 183–194.
[17] E. De Giorgi, Sulla convergenza di alcune successioni d'integrali del tipo dell'area, Ennio De Giorgi, 414.
[18] Y. Efendiev, V. Ginting, T. Y. Hou and R. Ewing, Accurate multiscale finite element methods for two-phase flow simulations, *Journal of Computational Physics*, **220** (2006), 155–174.
[19] Y. Efendiev and T. Y. Hou, Multiscale finite element methods for porous media flows and their applications, *Applied Numerical Mathematics*, **57** (2007), 577–596.
[20] Y. Efendiev and T. Y. Hou, *Multiscale Finite Element Methods: Theory and Applications*, vol. 4, Springer Science & Business Media, 2009.
[21] Y. Efendiev, T. Y. Hou and V. Ginting, Multiscale finite element methods for nonlinear problems and their applications, *Communications in Mathematical Sciences*, **2** (2004), 553–589.
[22] Y. Efendiev, T. Y. Hou and X. Wu, Convergence of a nonconforming multiscale finite element method, *SIAM Journal on Numerical Analysis*, **37** (2000), 888–910.
[23] D. Gilbarg and N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, vol. 224, Springer Science & Business Media, 2001.
[24] A. Gloria, An analytical framework for the numerical homogenization of monotone elliptic operators and quasiconvex energies, *Multiscale Modeling & Simulation*, **5** (2006), 996–1043.
[25] P. Henning and D. Peterseim, Oversampling for the multiscale finite element method, *Multiscale Modeling & Simulation*, **11** (2013), 1149–1175.
[26] T. Y. Hou and P. Liu, A heterogeneous stochastic fem framework for elliptic pdes, *Journal of Computational Physics*, **281** (2015), 942–969.
[27] T. Y. Hou, P. Liu and Z. Zhang, A model reduction method for elliptic pdes with random input using the heterogeneous stochastic fem framework, *Bulletin of the Institute of Mathematics*, **11** (2016), 179–216.
[28] T. Y. Hou and X. Wu, A multiscale finite element method for elliptic problems in composite materials and porous media, *Journal of computational physics*, **134** (1997), 169–189.
[29] T. Y. Hou, X. Wu and Z. Cai, Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients, *Mathematics of Computation of the American Mathematical Society*, **68** (1999), 913–943.
[30] T. Y. Hou, X. Wu and Y. Zhang, Removing the cell resonance error in the multiscale finite element method via a petrov-galerkin formulation, *Communications in Mathematical Sciences*, **2** (2004), 185–205.
[31] P. Jenny, S. Lee and H. Tchelepi, Multi-scale finite-volume method for elliptic problems in subsurface flow simulation, *Journal of Computational Physics*, **187** (2003), 47–67.
[32] P. Jenny, S. Lee and H. Tchelepi, Adaptive multiscale finite-volume method for multiphase flow and transport in porous media, *Multiscale Modeling & Simulation*, **3** (2005), 50–64.
[33] V. Jikov, S. Kozlov and O. Oleinik, *Homogenization of Differential Operators and Integral Functionals*, Springer Science & Business Media, 2012.
[34] S. Kozlov, Averaging of random operators, *Matematicheskii Sbornik*, **151** (1979), 188–202.
[35] I. Lunati and P. Jenny, Multiscale finite-volume method for compressible multiphase flow in porous media, *Journal of Computational Physics*, **216** (2006), 616–636.
[36] A. Målqvist and D. Peterseim, Localization of elliptic multiscale problems, *Mathematics of Computation*, **83** (2014), 2583–2603.
[37] J. Melenk, On n-widths for elliptic problems, *Journal of mathematical analysis and applications*, **247** (2000), 272–289.
[38] J. Melenk and I. Babuška, The partition of unity finite element method: Basic theory and applications, *Computer methods in applied mechanics and engineering*, **139** (1996), 289–314.
[39] R. Millward, *A New Adaptive Multiscale Finite Element Method with Applications to High Contrast Interface Problems*, PhD thesis, University of Bath, 2011.

[40] F. Murat, Compacité par compensation, *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, **5** (1978), 489–507.
[41] F. Murat and L. Tartar, *H-convergence*, Springer, 1997.
[42] H. Owhadi, Bayesian numerical homogenization, *Multiscale Modeling & Simulation*, **13** (2015), 812–828.
[43] H. Owhadi, Mult-grid with rough coefficients and multiresolution operator decomposition from hierarchical information games, arXiv preprint arXiv:1503.03467.
[44] H. Owhadi and L. Zhang, Metric-based upscaling, *Communications on Pure and Applied Mathematics*, **60** (2007), 675–723.
[45] H. Owhadi, L. Zhang and L. Berlyand, Polyharmonic homogenization, rough polyharmonic splines and sparse super-localization, *ESAIM: Mathematical Modelling and Numerical Analysis*, **48** (2014), 517–552.
[46] N. Panasenko and N. Bakhvalov, Homogenization: Averaging processes in periodic media: Mathematical problems in the mechanics of composite materials, 1989.
[47] D. Peterseim, Variational multiscale stabilization and the exponential decay of fine-scale correctors, arXiv:1505.07611.
[48] A. Pinkus, *n-Width in Approximation Theory*, Springer, 1985.
[49] S. Spagnolo, Sulla convergenza di soluzioni di equazioni paraboliche ed ellittiche, *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, **22** (1968), 571–597.
[50] S. Spagnolo, Convergence in energy for elliptic operators, Numerical Solutions of Partial Differential Equations III, Acad. Press, New York.

Received June 2015; revised October 2015.

E-mail address: hou@cms.caltech.edu
E-mail address: plliu@caltech.edu