An Iteratively Adaptive Multi-scale Finite Element Method for Elliptic PDEs With Rough Coefficients

Thomas Y. Hou[a], Feng-Nan Hwang[b,*], Pengfei Liu[a], Chien-Chou Yao[b]

[a]*Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, CA 91125, USA*
[b]*Department of Mathematics, National Central University, Jhongli District, Taoyuan City 32001, Taiwan*

Abstract

We propose an iteratively adaptive Multi-scale Finite Element Method (MsFEM) for elliptic PDEs with rough coefficients. The choice of the local boundary conditions for the multi-sale basis functions determines the accuracy of the MsFEM numerical solution, and one needs to incorporate the global information of the elliptic equation into the local boundary conditions of the multi-scale basis functions to recover the underlying fine-mesh solution of the equation. In our proposed iteratively adaptive method, we achieve this global-to-local information transfer through the combination of coarse-mesh solving using adaptive multi-scale basis functions and fine-mesh smoothing operations. In each iteration step, we first update the multi-scale basis functions based on the approximate numerical solutions of the previous iteration steps, and obtain the coarse-mesh approximate solution using a Galerkin projection. Then we apply several steps of smoothing operations to the coarse-mesh approximate solution on the underlying fine mesh to get the updated approximate numerical solution. The proposed algorithm can be viewed as a nonlinear two-level multi-grid method with the restriction and prolongation operators adapted to the approximate numerical solutions of the previous iteration steps. Convergence analysis of the proposed algorithm is carried out under the framework of two-level multi-grid method, and the harmonic coordinates are employed to establish the approximation property of the adaptive multi-scale basis functions. We demonstrate the efficiency of our proposed multi-scale methods through several numerical examples including a multi-scale coefficient problem, a high-contrast interface problem, and a convection-dominated diffusion problem.

Key words: Iteratively adaptive MsFEM, global-to-local information transfer, two-level multi-grid method, elliptic equation with rough coefficients, convection-dominated diffusion equation.

1. Introduction

Many problems of practical importance in science and engineering have multi-scale features. Porous media simulation and composite materials modeling are typical examples of such kind. The small-scale features of the problem can have significant impact on the large-scale properties of the solutions, and one needs a very fine discretization that resolves the small-scale variation of the problem to obtain faithful numerical results. For these problems, methods that allow people to effectively incorporate the small-scale features of the problem into the large-scale properties of the solutions are desired. In this work, we consider the following second order linear elliptic partial differential equation (PDE) with rough coefficient,

$$\begin{cases} -\mathrm{div}(a(x)\nabla u(x)) = f(x), & x \in D, \\ u(x)|_{\partial D} = 0, \end{cases} \tag{1}$$

and the proposed method will also be applied to a convection-dominated diffusion problem.

Here we assume that D is a convex polygon domain in R^2, and the forcing term $f(x) \in L^2(D)$. We also assume that the equation is uniformly elliptic, i.e., there exist $\lambda_{\min} > 0$ and $\lambda_{\max} > 0$ such that

$$a(x) \in [\lambda_{\min}, \lambda_{\max}]. \tag{2}$$

We do not assume any regularity of the coefficient $a(x) \in L^\infty(\Omega)$, which can be arbitrarily rough, thus equation (1) can be used to model diffusion process in strongly heterogeneous media. Classical finite element methods use piecewise linear (polynomial) functions to discretize the equation, and for smooth coefficient problems, their convergence depends on the following approximation property and regularity result

$$\|u(x) - Ju(x)\|_{H_0^1(D)} \le CH\|u(x)\|_{H^2(D)}, \quad \|u(x)\|_{H^2(D)} \le C\|f(x)\|_{L^2(D)}, \tag{3}$$

where $Ju(x)$ is the piecewise linear interpolation of $u(x)$, and H is the underlying mesh size. We can see that $O(H)$ accuracy in the $H_0^1(D)$ norm can be obtained if a discretization of size H is employed. However, classical finite element methods may fail if the coefficient $a(x)$ is rough, because the solution $u(x)$ loses regularity and $\|u(x)\|_{H^2(D)}$ cannot be bounded by $\|f(x)\|_{L^2(D)}$ in (3). It is actually shown in [9] that the polynomial finite element methods can perform arbitrarily badly in this setting. In practice, one needs a much finer mesh that resolves the small scale features of $a(x)$ to achieve the desired accuracy.

1.1. The multi-scale finite element method

The Multi-scale Finite Element Methods (MsFEM) [44, 45, 46, 28, 19] were developed to solve these multi-scale PDEs. Instead of piecewise polynomials in the classical finite element methods, MsFEM employs basis functions that incorporate properties of the elliptic operator to discretize the equation (1),

$$\phi_1(x), \phi_2(x), \ldots, \phi_n(x) \in H_0^1(\Omega). \tag{4}$$

Consider a coarse-mesh discretization of the domain with mesh size H, and an underlying fine mesh of size h that refines the coarse mesh and resolves the small-scale variations of the problem. This two-level discretization of the domain is illustrated in Figure 1. The MsFEM basis functions $\phi_j(x)$ are constructed by numerically solving local elliptic boundary value problems on each coarse element D_i,

$$-\mathrm{div}(a(x)\nabla\phi_{i_j}(x)) = 0, \quad x \in D_i, \quad \phi_{i_j}(x_{i_k}) = \delta_{kj}. \tag{5}$$

(a) A Coarse Element. (b) Two-level Discretization.
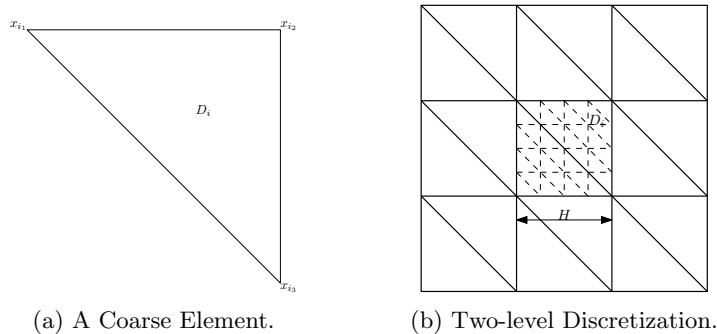
Figure 1: Discretization of the problem.

With the basis functions $\phi_i(x)$, $i = 1, \ldots, n$, MsFEM finds the numerical solution

$$u_H^{MS}(x) \in V_H = \mathrm{span}\{\phi_1(x), \phi_2(x), \ldots, \phi_n(x)\} \subset H_0^1(\Omega), \tag{6}$$

using the Galerkin projection. Namely, one finds $u_H^{MS}(x) \in V_H$, such that

$$a(u_H^{MS}(x), v(x)) = \langle f(x), v(x) \rangle, \quad \text{for all} \quad v(x) \in V_H, \tag{7}$$

where $a(u(x), v(x)) = \int_\Omega \nabla u(x)^t a(x) \nabla v(x) \mathrm{d}x$, $\langle f(x), v(x) \rangle = \int_\Omega f(x) v(x) \mathrm{d}x$.

The numerical solution defined above satisfies the following optimal property under the energy norm

$$\|u(x) - u_H^{MS}(x)\|_E = \inf_{v(x) \in V_H} \|u(x) - v(x)\|_E, \quad \|u(x)\|_E^2 = a(u(x), u(x)). \tag{8}$$

Convergence analysis of MsFEM in the periodic homogenization setting, where $a(x) = A(\frac{x}{\epsilon})$ with $\epsilon = o(1)$, was given in [45, 29, 17]. To construct the multi-scale basis function $\phi_{i_j}(x)$ on D_i in (5), appropriate local boundary conditions on ∂D_i are needed. We will review in section 2 that the local boundary conditions of the basis functions are crucial and determine the accuracy of the MsFEM numerical solution. With the optimal choice of local boundary conditions for the basis and some local bubble corrections, MsFEM can indeed recover the underlying fine-mesh solution. However, due to the nonlocal nature of elliptic equations, the local solution to (1) depends on the coefficient $a(x)$ and the forcing term $f(x)$ on the whole domain D. And to exactly recover the underlying fine-mesh solution, one needs to incorporate the global information of the elliptic equation into the local boundary conditions of the basis functions in (5).

1.2. Main contributions of this work

In this paper we propose an iteratively adaptive method, Algorithm 1, to construct the multi-scale basis. The global information of the elliptic equation is transferred to the local boundary conditions of the multi-scale basis through iterative coarse-mesh solving and fine-mesh smoothing operations.

In the s-th step of the iteration, we first construct basis functions $\phi_i^s(x)$, $i = 1, 2 \ldots, n$, based on the approximate numerical solutions from the previous several iteration steps, namely, $u^k(x)$, $k \leq s - 1$. We require that these multi-scale basis functions can exactly recover $u^{s-1}(x)$ on the boundary of the coarse-mesh elements, $\Gamma = \cup_i \partial D_i$. Then we solve the equation (1) using the basis functions $\phi_i^s(x)$ and a Galerkin projection, and correct the numerical solution $u_H^{(s)}(x)$ by solving a series of local bubble problems. Finally, we apply several steps of smoothing operations on the fine mesh to the numerical solution $u_H^{(s)}(x)$ to get the updated approximate solution $u^s(x)$. Throughout the iteration process, the approximation property of the basis functions $\phi_i^s(x)$ to the solution $u(x)$ gets improved. Our numerical results suggest that the numerical solutions $u^s(x)$ converge to the underlying fine-mesh solution exponentially fast. The coarse-mesh solving and the fine-mesh smoothing operations are the two key steps of the proposed algorithm.

Algorithm 2 is derived as a stabilized variant of Algorithm 1, and is essentially a two-level multi-grid method that employs multi-scale basis functions for the restriction and prolongation operations. By comparing Algorithm 1 with Algorithm 2, we show that even without explicit restriction operation, Algorithm 1 can indeed be viewed as a nonlinear version of the two-level multi-grid method where the restriction and prolongation operators are adapted to the approximate solutions of the previous iteration steps.

Convergence analysis of the proposed algorithms is carried out under the framework of the two-level multi-grid method. And the harmonic coordinates [59] are employed to establish the approximation property of the adaptively constructed multi-scale basis functions to the solution space of equation (1).

There exist some multi-scale methods in the literature that adaptively update the boundary conditions of the multi-scale basis. In [25, 54], the adaptive multi-scale basis functions in each iteration step are constructed using the local oversampling technique, and the local boundary conditions on the oversampling domain are constructed based on approximate solutions of the previous steps. There exist no smoothing operations in these methods, and they cannot obtain fine mesh accuracy as the present work. In [47], an adaptive bubble part is introduced in each iteration step, and the resulting method is similar to Algorithm 2 in this work. The method in [47] employs a finite volume formulation, and a dual mesh is needed to formulate the coarse-mesh linear system. In additin, the bubble correction terms in [47] need to be constructed explicitly by solving local boundary value problems. No convergence analysis has been provided for the previous adaptive multi-scale methods [47, 25, 54]. After suitable modifications, these methods may be analyzed under the same two-level multi-grid method framework presented in the current paper.

To demonstrate the applicability of our proposed methods for a variety of applications with multi-scale features, we consider three classes of PDE problems, including elliptic problems with rough coefficients [39, 40, 44, 45], elliptic interface problems with high-contrast discontinuous coefficients [18, 19, 49, 54], and convection-dominated diffusion equations [21, 31, 33, 34, 66]. In all the three types of problems that we consider, we observe significant improvement of the proposed methods over classical methods.

1.3. Review of multi-scale methods

Considerable amount of efforts has been devoted to analyzing and numerically solving the multi-scale PDE (1), and we list a few previous works related to the present paper below.

The classical homogenization theories, including the periodic homogenization where $a(x) = A(\frac{x}{\epsilon})$ [11, 48, 62, 20, 3, 2], and the H, G, Γ-convergence theories [56, 23, 22, 68, 67, 55, 35], consider the convergence (in an appropriate sense) of a sequence of operators parameterized by ϵ as $\epsilon \to 0$.

Numerical methods for the multi-scale elliptic PDE (1) can be divided into two categories. Methods in the first category can be viewed as fast solvers of the equation on the underlying fine mesh and alternatives of the multi-grid methods [32, 13, 36, 37]. Multi-grid methods are known as the fastest solvers for elliptic boundary value problems with linear computational complexity. However, due to the lack of regularity of the solutions, classical multi-grid method in general would fail for the multi-scale elliptic PDE (1) [30, 26]. Methods in the first category include the algebraic multi-grid methods (AMG) [1, 15, 14, 64, 69], the hierarchical matrix methods [38, 10], AMG based on energy minimization [51, 74, 76, 75, 72], stabilized hierarchical basis methods [71, 70, 73], and wavelet based methods [16, 12, 4, 24].

Methods in the second category, including the MsFEM, are essentially model reduction methods. They employ special basis functions adapted to the elliptic operator instead of piecewise polynomials to descretize the equation to reduce the total degrees of freedom. These methods generally employ a two-level discretization as shown in Figure 1, and are only able to achieve coarse-mesh accuracy. Methods in the second category include the generalized finite element methods [53, 8, 5], oversampling methods [7, 6, 43], localization methods [60, 50, 41, 63, 42], harmonic coordinates transformation [59], the polyharmonic splines [61, 57], and a recent method that depends on the multi-resolution decomposition of the solution space [58].

The proposed Algorithm 1 in this paper is able to achieve fine-mesh accuracy by iteratively incorporating the global information of the equation into the local boundary conditions of the basis functions. Thus it falls into the first category of multi-scale methods. Namely, it can be viewed as a fast fine-mesh solver and an alternative of the classical multi-grid methods. On the other hand, the proposed Algorithm 1 employs special basis to discretize the equations on the coarse mesh, and from this perspective it is similar to the model reduction method in the second category. In multi-scale methods of the second category, building basis with robust approximation property is in general expensive. The proposed method bypasses this difficulty by adaptively improving the approximation property of the multi-scale basis in the iteration.

The rest of the paper is organized as follows. In section 2, we briefly review multi-scale basis functions and the importance of their local boundary conditions. In section 3, we detail the proposed method to solve the multi-scale PDE (1) and derive a variant of this method. In section 4, we provide convergence analysis of the proposed algorithm under the two-level multi-grid method framework. In section 5, we present some numerical results employing the proposed method. Concluding remarks are made in section 6.

2. Motivation: Multi-scale Basis and the Importance of Their Local Boundary Conditions

Given a coarse-mesh discretization of the domain $D = \cup_i D_i$ as shown in Figure 1, we consider the following decomposition of the solution to (1), $u(x) = u^1(x) + u^2(x)$,

$$
\begin{cases} -\text{div}(a(x)\nabla u^1(x)) = 0, \ x \in D_i; \\ u^1(x)|_{\partial D_i} = u(x)|_{\partial D_i}; \end{cases} \qquad \begin{cases} -\text{div}(a(x)\nabla u^2(x)) = f(x), \ x \in D_i; \\ u^2(x)|_{\partial D_i} = 0. \end{cases} \tag{9}
$$

$u^1(x)$ is $a(x)$-harmonic in each coarse element and only depends on the trace of the solution on the edges of coarse elements, $\Gamma = \cup_i \partial D_i$. We call $u^1(x)$ the $a(x)$-harmonic part of the solution. $u^2(x)$ depends on

the forcing term $f(x)$ in each coarse element, and it can be constructed by solving a series of local bubble problems. We call $u^2(x)$ the bubble part of the solution. $u^1(x)$ and $u^2(x)$ are $a(x)$-orthogonal,

$$a(u^1(x), u^2(x)) = 0. \tag{10}$$

Besides, the local bubble part of the solution $u^2(x)$ is small on each coarse element,

$$\|u^2(x)|_{D_i}\|_{H_0^1(D_i)} \leq CH\|f(x)\|_{L^2(D_i)}, \quad \|u^2(x)|_{D_i}\|_{L^2(D_i)} \leq CH^2\|f(x)\|_{L^2(D_i)}, \tag{11}$$

which can be obtained using a scaling argument. Summing (11) up over each coarse element, we obtain

$$\|u^2(x)\|_{H_0^1(D)} \leq CH\|f(x)\|_{L^2(D)}, \quad \|u^2(x)\|_{L^2(D)} \leq CH^2\|f(x)\|_{L^2(D)}. \tag{12}$$

Motivated by the above orthogonal decomposition of the solution, a special type of basis functions, the *multi-scale basis functions* were introduced in [43], which generalize the basis functions employed in MsFEM. They are $a(x)$-harmonic within each coarse element, but are not necessarily nodal basis.

Definition [43] (Multi-scale basis) Consider basis functions $\psi_1(x), \psi_2(x), \ldots, \psi_n(x) \in H_0^1(D)$, if they are $a(x)$-harmonic on each coarse element D_j of the discretization (but not across the boundary), namely,

$$-\text{div}(a(x)\nabla\psi_i(x)) = 0, \quad x \in D_j,$$

then they are called multi-scale basis functions.

Due to the smallness of $u^2(x)$ (12), one can get the following optimal property of multi-scale basis.

Proposition [43] Consider a set of basis functions $\psi_i(x) \in H_0^1(D)$, $i = 1, 2 \ldots, m$, and a set of multi-scale basis $\phi_i(x)$, $i = 1, 2, \ldots, n$. Denote the Galerkin solutions using these basis functions as $u_H^\psi(x)$ and $u_H^{MS}(x)$ respectively, and let the union of coarse element boundaries be $\Gamma = \cup_i \partial D_i$. If

$$\text{span}\{\psi_1(x)|_\Gamma, \psi_2(x)|_\Gamma, \ldots, \psi_m(x)|_\Gamma\} = \text{span}\{\phi_1(x)|_\Gamma, \phi_2(x)|_\Gamma, \ldots, \phi_n(x)|_\Gamma\}, \tag{13}$$

then

$$\|u(x) - u_H^{MS}(x)\|_E^2 \leq \|u(x) - u_H^\psi(x)\|_E^2 + CH^2\|f(x)\|_{L^2(D)}^2. \tag{14}$$

The proposition says that for the same boundary conditions on Γ, the multi-scale basis functions perform at least as well as other basis functions, if only $O(H)$ accuracy in the energy norm is desired.

Moreover, the multi-scale basis functions $\phi_i(x)$, $i = 1, 2 \ldots, n$, are $a(x)$-orthogonal to the local bubble part of the solution $u^2(x)$, then according to the Galerkin projection formulation (7) and the optimal property (8), we have the following characterization of the numerical solution $u_H^{MS}(x)$,

$$\|u^1(x) - u_H^{MS}(x)\|_E = \inf_{v(x) \in V_H} \|u^1(x) - v(x)\|_E. \tag{15}$$

Recall that we can construct the local bubble part of the solution $u^2(x)$ by solving a series of local bubble problems (9) on each coarse element, independently from the Galerkin projection (7). We can then add $u^2(x)$ to $u_H^{MS}(x)$ to get the *corrected multi-scale solution* $u_H(x)$,

$$u_H(x) = u_H^{MS}(x) + u^2(x). \tag{16}$$

We have the following approximation property for the corrected multi-scale solution,

$$\|u(x) - u_H(x)\|_E = \|u^1(x) - u_H^{MS}(x)\|_E = \inf_{v(x) \in V_H} \|u^1(x) - v(x)\|_E. \tag{17}$$

The right hand side of (17) only depends on $u(x)|_\Gamma$ and the local boundary conditions of the multi-scale basis (13). With (17), one can focus on multi-scale basis functions, and their local boundary conditions

5

(13). If the multi-scale basis can exactly recover the solution on the boundaries of the coarse elements, Γ, the corrected multi-scale solution $u_H(x)$ can exactly recover the fine-mesh solution.

For a specific solution $u(x)$ to (1), one can choose local boundary conditions in (5) to make (17) vanish. However, these choices will depend on $u(x)|_\Gamma$, which is unknown in advance. Due to the nonlocal nature of elliptic operators, the local solution to (1) depends on the elliptic equation and the forcing term on the whole domain. Thus to exactly recover the underlying fine-mesh solution, one needs to incorporate the global information of the elliptic equation into the local boundary conditions of the multi-scale basis functions. In this paper, we propose an iteratively adaptive approach that achieves this global-to-local information transfer through the combination of coarse-mesh solving and fine-mesh smoothing operations.

3. An Iteratively Adaptive MsFEM for Multi-scale Elliptic PDEs

We propose an iteratively adaptive mulit-scale finite element method in this section, which iteratively updates the multi-scale basis functions based on the approximate numerical solutions in previous iteration steps. We also derive a stabilized variant of the proposed method, and it is essentially a two-level multi-grid method that employs a set of multi-scale basis for the restriction and prolongation operations.

3.1. An Iteratively Adaptive MsFEM

In this algorithm, we iteratively update the multi-scale basis functions $\phi_i^s(x)$ and the approximate numerical solution $u^s(x)$ simultaneously, where s is the iteration step.

Let the adaptive multi-scale finite element space at the s-th iteration step be

$$V_H^s = span\{\phi_i^s(x), i = 1, ..., n\}, \tag{18}$$

where n is the total number of interior coarse-mesh nodal points. At the s-th step, we first solve the equation (1) on the coarse mesh using trial space V_H^s (18) and the Galerkin projection (7) to get $u_H^{(s)}(x)$,

$$a(u_H^{(s)}(x), v_H(x)) = \langle f(x), v_H(x) \rangle, \quad \text{for all} \quad v_H(x) \in V_H^s. \tag{19}$$

Then we solve the local bubble problems (9) to get the local bubble part of the solution $u^2(x)$, and add it to $u_H^{(s)}(x)$. Recall that we can compute $u_H^{(s)}(x)$ and $u^2(x)$ independently because $\phi_i^s(x)$ are multi-scale basis functions, and they are $a(x)$-orthogonal to the local bubble part of the solution. We obtain

$$u^{(s)}(x) = u_H^{(s)}(x) + u^2(x). \tag{20}$$

These local bubble problems (9) need to be solved for only once in the iteration, and the obtained bubble part of the solution $u^2(x)$ will be used to correct $u_H^{(s)}(x)$ in each step of the iteration.

Recall that the multi-scale basis functions $\phi_i^s(x)$ and the local bubble part of the solution $u^2(x)$ are all constructed and stored using the piecewise linear basis functions on the underlying fine mesh. Then the approximate numerical solution $u^{(s)}(x)$ will also be stored as linear combination of the fine mesh piecewise linear basis. Discretizing the equation (1) using the fine mesh piecewise linear basis, we get the underlying fine mesh linear system. Next, we apply ν steps of smoothing operations, e.g., weighted Jacobi iteration, to the underlying fine mesh linear system using $u^{(s)}(x)$ as initial guess to get the updated $u^s(x)$,

$$u^s(x) = S^\nu u^{(s)}(x). \tag{21}$$

The purpose of this step is to eliminate the high frequency modes of the error in $u_H^{(s)}(x)$. Here the high-frequency modes correspond to the eigenfunctions of the elliptic operator with large eigenvalues.

Then we employ the approximate numerical solutions of previous iteration steps, $u^k(x)$, $k \leq s$, to construct the updated multi-scale basis functions for the next iteration step,

$$\phi_i^{s+1}(x), \quad i = 1, \ldots, n. \tag{22}$$

6

Here we require that the updated multi-scale basis functions (22) can exactly recover $u^s(x)|_\Gamma$,

$$u^s(x) - \sum_i u^s(x_i)\phi_i^{s+1}(x) = 0, \quad x \in \Gamma. \tag{23}$$

With the updated multi-scale basis $\phi_i^{s+1}(x)$, we can move to the next iteration step. There are several choices for the construction of multi-scale basis (22), and we describe two of them below.

1. One possible choice of the local boundary conditions for the multi-scale basis based on sampled solutions was proposed in [27] and later employed in [25, 54]. On the coarse element D_i as shown in Figure 1, we choose $\phi_{i_1}^{s+1}(x) = 0$ on the edge $x_{i_2}x_{i_3}$. And on the edges $x_{i_1}x_{i_j}$, $j = 2, 3$, we choose

$$\phi_{i_1}^{s+1}(x) = \begin{cases} \frac{u^s(x) - u^s(x_{i_j})}{u^s(x_{i_1}) - u^s(x_{i_j})}, & u^s(x_{i_1}) \neq u^s(x_{i_j}); \\ \text{linear}, & u^s(x_{i_1}) = u^s(x_{i_j}). \end{cases} \tag{24}$$

One can show that in the case that the values of $u(x)$ on any two adjacent node points are not equal,

$$u^s(x) = \sum_{i=1}^n u^s(x_i)\phi_i^{s+1}(x), \quad x \in \Gamma.$$

Namely, the basis functions can exactly recover the trace of the solution on Γ.

2. Another choice is employing the oversampling technique [44]. In the oversampling method, one solves two local boundary value problems (5) on an oversampling domain K_i that encloses the coarse element D_i, and uses the linear combination of these local sampled solutions to get the multi-sale basis on D_i. Now given two sampled solutions $u^s(x)$ and $u^{s-1}(x)$, we build the local boundary conditions of the multi-scale basis $\psi_{i_j}(x)$ on the D_i as the linear combination of constant, $u^s(x)$ and $u^{s-1}(x)$,

$$\bar{\phi}_{i_j}^{s+1}(x) = c_{i,j,0}^{s+1} + c_{i,j,1}^{s+1}u^s(x) + c_{i,j,2}^{s+1}u^{s-1}(x), \quad x \in \partial D_i, \tag{25}$$

where the coefficients $c_{i,j,k}^{s+1}$, $k = 0, 1, 2$ are determined by the condition that

$$\bar{\phi}_{i_j}^{s+1}(x_{i_l}) = \delta_{jl}. \tag{26}$$

The basis functions constructed as (25) are in general non-conforming and $\bar{\phi}_i(x)$ may have jump across the coarse-mesh edges. This is because the local boundary conditions on a coarse mesh edge e determined from its two neighbor coarse elements may be different. We can use a simple averaging technique to obtain conforming basis. For edge e, we denote its two neighbor coarse elements as D_{i_1}, D_{i_2}. We modify the local boundary conditions of $\bar{\phi}_j(x)$ on e, and denote the corresponding averaged basis as $\phi_j(x)$,

$$\phi_j(x)|_e = \frac{1}{2}(\bar{\phi}_{j,D_{i_1}}(x)|_e + \bar{\phi}_{j,D_{i_2}}(x)|_e), \tag{27}$$

where $\bar{\phi}_{j,D_{i_1}}(x) = \bar{\phi}_j(x)|_{D_{i_1}}$, $\bar{\phi}_{j,D_{i_2}}(x) = \bar{\phi}_j(x)|_{D_{i_2}}$. The price we pay for this averaging process is that the resulting basis $\phi_j(x)$ has enlarged support. This is shown in Figure 2. The same averaging technique to make the basis functions conforming has been employed in [25, 19, 54].

Note that to determine the coefficients $c_{i,j,k}^{s+1}$ according to the condition (26), one needs to solve a 3×3 linear system. If the system is singular, we choose linear local boundary conditions for the multi-scale basis functions on element D_i. If the linear system is not singular, one can show that the resulting multi-scale basis functions (27) can recover $u^s(x)$ on $\Gamma = \cup \partial D_i$. Convergence analysis of Algorithm 1 with the choice of updated multi-scale basis (27) will be carried out in section 4.
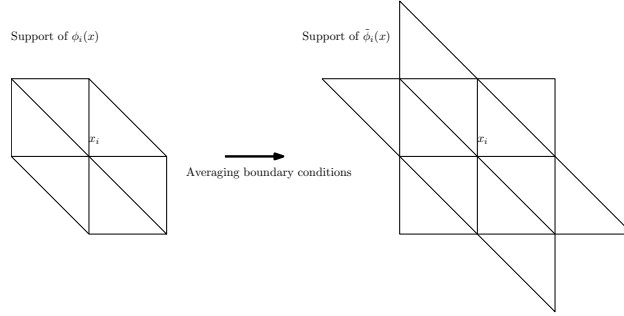
7

Figure 2: Averaging on the coarse mesh edge.

Algorithm 1 Iteratively adaptive multi-scale finite element algorithm

---

1: **for** $s = 1$ to $s_{\max}$ **do**
2:     **for** $i = 1$ to n **do**
3:         Construct the local boundary conditions for the multi-scale basis $\phi_i^s(x)$ based on the approximate solutions of the previous steps, $u^k(x), k \leq s - 1$. For example, employing formula (24) or (27).
4:         Construct the multi-scale basis $\phi_i^s(x)$ by solving local boundary value problems (5).
5:     **end for**
6:     Employ V_H^s (18) and the Galkerkin projection to solve (1) and get $u_H^{(s)}(x)$.
7:     Correct the coarse-mesh solution using the local bubble part $u^2(x)$,

$$u^{(s)}(x) = u_H^{(s)}(x) + u^2(x).$$

8:     Relax the fine mesh linear system $Au = b$ for ν times with the initial guess $u^{(s)}(x)$ to get $u^s(x)$.
9:     Check the user-defined stopping criteria.
10: **end for**

8

Iterative implementation of the above coarse mesh solving (19) and fine mesh smoothing (21) operations leads to the proposed iteratively adaptive multi-scale finite element method, Algorithm 1. We employ the Galerkin method to solve the local bubble problems (9) and construct multi-scale basis (5).

Remark 1 *There are similar multi-scale methods in the literature, for example [25, 54], which iteratively update the local boundary conditions of the multi-scale basis functions. In those works, the adaptive multi-scale basis functions are constructed using the oversampling technique. And the local boundary conditions of basis functions on the oversampling domain are constructed based on the approximate numerical solution of the previous iteration step, $u^s(x)$, and the rule (24). These works are different from our proposed method in the following aspects. In [25, 54], the multi-scale basis functions cannot recover $u^{s-1}(x)$ on the boundaries of coarse elements, Γ, because the local bubble part of $u^{s-1}(x)$ on K_i is not taken into account in the boundary conditions of the basis functions on ∂D_i. Thus these methods [25, 54] cannot obtain the underlying fine mesh solution even the iteration converges. Besides, there are no explicit smoothing operations in the previous methods [25, 54], and the transfer of the global information of the equation to the local boundary conditions of the multi-scale basis is achieved through oversampling and coarse-mesh solving.*

We will prove in section 4 under certain assumptions that the approximate numerical solutions $u^s(x)$ in Algorithm 1 converge to the underlying fine mesh solution exponentially fast, which means that the iteratively adaptive approach proposed in Algorithm 1 can effectively incorporate the global information of the equation into the local boundary conditions of the multi-scale basis (13).

Before we end this subsection, we make the following comments on the proposed Algorithm 1.

1. In the construction of the updated multi-scale basis functions, for certain $u^s(x)$, there may not exist nodal basis functions $\psi_i^{s+1}(x)$, $i = 1, \ldots, n$, which can recover the solution on Γ in (23).
   An extreme case is that $u^s(x_i) = 0$, $i = 1, 2 \ldots, n$, but $u^s(x)$ is not a zero function.
2. In each step of the iteration, one needs to construct the updated multi-scale basis functions by solving local boundary value problems on the fine mesh and assembling the updated stiffness matrix.
3. In constructing the local boundary conditions, one may encounter dividing by a small number in (24) or solving a nearly singular system in (26), thus suffer from some instability issue.

A stabilized variant of the proposed method, Algorithm 2, is derived in the next subsection, which introduces an adaptive correction term and keeps the multi-scale basis fixed throughout the iteration. Algorithm 2 is less expensive in each iteration step and avoids the potential instability issue in Algorithm 1.

3.2. A stabilized variant of the proposed algorithm

In Algorithm 1, the multi-scale basis functions $\phi_i^s(x)$ are constructed according to the approximate numerical solutions in the previous iteration steps, $u^k(x)$, $k < s$. Then we employ $\phi_i^s(x)$ and a Galerkin projection to get the approximate numerical solution $u^{(s)}(x)$ (20). We consider simplifying this step by using a set of multi-scale basis functions constructed before the iteration, which we denote as

$$\phi_i(x), \quad i = 1, 2, \ldots, n. \tag{28}$$

We have the following decomposition of $u^{s-1}(x)$ and $u^{(s)}(x)$ using $\phi_i(x)$,

$$u^{s-1}(x) = \sum_i u^{s-1}(x_i)\phi_i(x) + R^{s-1}(x), \tag{29}$$

$$u^{(s)}(x) = \sum_i u^{(s)}(x_i)\phi_i(x) + R^{(s)}(x), \tag{30}$$

where $R^{s-1}(x)$ and $R^{(s)}(x)$ are the interpolation residuals of $u^{s-1}(x)$ and $u^{(s)}(x)$ using basis $\phi_i(x)$.

Since both $u^{s-1}(x)$ and $u^{(s)}(x)$ approximate $u(x)$, we assume that the interpolation residuals $R^{s-1}(x)$ and $R^s(x)$ are close, $R^{s-1}(x) \approx R^{(s)}(x)$. Then $u^{(s)}(x)$ can be written as

$$u^{(s)}(x) \approx \sum_i u^{(s)}(x_i)\phi_i(x) + u^{s-1}(x) - \sum_i u^{s-1}(x_i)\phi_i(x) = \sum_i c_i\phi_i(x) + u^{s-1}(x). \tag{31}$$

9

And we construct an approximation of $u^{(s)}(x)$ (31) by solving the following optimization problem,

$$\min_{c_i} \|u(x) - \sum_i c_i \phi_i(x) - u^{s-1}(x)\|_E. \tag{32}$$

The resulting $e_H^s(x) = \sum_{i=1}^n c_i \phi_i(x)$ is the Galerkin finite element solution of equation

$$-\operatorname{div}(a(x)\nabla e(x)) = f(x) + \operatorname{div}(a(x)\nabla u^{s-1}(x)), \quad e(x)|_{\partial D} = 0, \tag{33}$$

using the multi-scale basis functions $\phi_i(x)$, $i = 1, \ldots, n$. And we get an approximation of $u^{(s)}(x)$,

$$u^{(s)}(x) \approx e_H^s(x) + u^{s-1}(x). \tag{34}$$

To reduce the error in the bubble part of $u^{(s)}(x)$ (34), we construct the bubble part of the solution to (33) by solving a series of local bubble problems as in (9). We denote the bubble part of the solution to (33) as $e^{s,2}(x)$. We use $e^{s,2}(x)$ to correct the approximate numerical solution in (34) and obtain

$$u^{(s)}(x) \approx u^{s-1}(x) + e_H^s(x) + e^{s,2}(x). \tag{35}$$

Replacing the construction of $u^{(s)}(x)$ in Algorithm 1 by the approximation (35), we get Algorithm 2. The interpolation residual $R^{s-1}(x)$ is introduced to motivate Algorithm 2 in (31), and in numerical implementation, we only need to solve equation (33) to get the approximate $u^{(s)}(x)$.

Algorithm 2 Iteratively Adaptive Multi-scale Finite Element Method

---

1: Construct multi-scale basis functions $\phi_i(x)$, $i = 1, \ldots, n$.
   Let $u^0(x) = 0$ be the initial guess of the solution.
2: **for** $s = 1$ to $s_{\max}$ **do**
3:   Solve equation (33) using multi-scale basis functions $\phi_i(x)$ to get $e_H^s(x)$
4:   Solve the local bubble part of the solution to (33), $e^{s,2}(x)$.
5:   Obtain the approximate numerical solution

$$u^{(s)}(x) = u^{s-1}(x) + e_H^s(x) + e^{s,2}(x).$$

6:   Relax the fine-mesh linear system for ν times with initial guess $u^{(s)}(x)$ to get $u^s(x)$.
7:   Check the user-defined stopping criteria.
8: **end for**

Because there is no need to construct updated multi-scale basis in each iteration step according to (24) or (25), the potential instability issue in the rules (24) and (26) is avoided, and this is is the reason that we call Algorithm 2 a stabilized variant of Algorithm 1. Moreover, the simplification procedure (35) also brings in computational savings in each iteration step of the Algorithm, since we do not need to construct the updated multi-scaled basis and the updated stiffness matrix in each iteration step.

The choice of multi-scale basis $\phi_i(x)$ (28) determines the convergence of Algorithm 2, and we will come back to the issue about the choice of the multi-scale basis in the convergence analysis in section 4.

Remark 2 *There exists a similar method in the literature [47], which also uses a bubble part to correct the coarse-mesh multi-scale solution. Algorithm 2 is different from the method in [47] in the following aspects. In [47], the finite element volume formulation is employed to construct the coarse-mesh linear system, and a set of dual coarse grids in addition to the primary coarse grids are required in the discretization. In [47], the correction term that enriches the finite element trial space is explicitly constructed in each step of the iteration by solving a series of local boundary value problems. Moreover, the local boundary conditions of the correction term on each coarse mesh edge are constructed by solving 1D problems in [47].*

4. Convergence of the Proposed Algorithms

In this section, we carry out the convergence analysis of the proposed Algorithm 1 and 2. We first show that Algorithm 2 is a two-level multi-grid method that employs a set of multi-scale basis for the restriction and prolongation operations. Then by comparing with Algorithm 2, we show that Algorithm 1 is a nonlinear version of the two-level multi-grid method with the restriction and prolongation operators adapted to the approximate solutions of the previous iteration steps. To establish the convergence of proposed Algorithm 1, one needs to analyze the approximation property of the constructed multi-scale basis to the solution space of (1) with $L^2(D)$ forcing $f(x)$, and this is achieved by employing the harmonic coordinates.

4.1. Reformulation of the two algorithms as two-level multi-grid methods

We denote the piecewise linear basis functions on the underlying fine mesh as

$$\psi_k(x), \quad k = 1, 2 \ldots, N. \tag{36}$$

The multi-scale basis functions $\phi_i^s(x)$ in Algorithm 1 (18) and $\phi_i(x)$ (28) in Algorithm 2 are all represented as linear combinations of $\psi_k(x)$. We denote these linear combinations as

$$\phi_i^s(x) = \sum_{k=1}^{N} D_{ki}^s \psi_k(x), \quad \phi_i(x) = \sum_{k=1}^{N} D_{ki} \psi_k(x), \tag{37}$$

then D^s and D are the matrix presentations of the multi-scale basis functions using $\psi_k(x)$.

Let $\phi_j^b(x)$, $j = 1, \ldots, N_b$ be the bubble basis functions, namely, the parts of fine mesh basis functions (36) that span the bubble part of the solution space. We have that for each $\phi_j^b(x)$, there exists k_j, such that

$$\phi_j^b(x) = \psi_{k_j}(x), \quad \psi_{k_j}(x)|_\Gamma = 0. \tag{38}$$

Then let D^b as the matrix representation of $\phi_j^b(x)$ using $\psi_k(x)$, and we have

$$D^b(i,j) = \delta_{i,k_j}. \tag{39}$$

We denote V_H^b as the space spanned by these bubble basis functions

$$V_H^b = \mathrm{span}\{\phi_1^b, \phi_2^b, \ldots, \phi_{N_b}^b\}. \tag{40}$$

When we solve equation (1) (in Algorithm 1) or (33) (in Algorithm 2) using the multi-scale basis (18) or (28), and add the corresponding local bubble parts back to the numerical solutions, we are essentially solving the equations using the multi-scale basis functions enriched with the bubble basis

$$\tilde{V}_H^s = V_H^s \oplus V_H^b, \quad \tilde{V}_H = V_H \oplus V_H^b. \tag{41}$$

This procedure can be separated into two steps, i.e., the coarse-mesh solving using V_H and local bubble correction using V_H^b, because the bubble basis is $a(x)$-orthogonal to the multi-scale basis.

We denote $\tilde{D}$ and $\tilde{D}^s$ as

$$\tilde{D} = [D, D^b], \quad \tilde{D}^s = [D^s, D^b], \tag{42}$$

which are the matrix representations of the multi-scale basis enriched with the bubble basis.

We denote the underlying fine-mesh linear system for equation (1) as

$$Au = f, \tag{43}$$

where A is the stiffness matrix, and f is the load vector,

$$A(i,j) = a(\psi_i(x), \psi_j(x)), \quad f(i) = \langle f(x), \psi_i(x) \rangle. \tag{44}$$

With the above notations, we derive the matrix representation of the iteration relation in Algorithm 1 and Algorithm 2. We consider Algorithm 2 first, which is a linear two-level multi-grid method.

Let u^s be the fine-mesh discrete solution in the s-the step, namely,

$$u^s(x) = \sum_{i=1}^{N} u^s(i)\psi_i(x). \tag{45}$$

In Algorithm 2, we employ the enriched trial space $\tilde{V}_H$ in (41) to solve (33) and get

$$e_H^s(x) + e^{s,2}(x). \tag{46}$$

We first compute the corresponding load vector $\langle \phi_i(x), f(x) + \text{div}(a(x)\nabla u^s(x)) \rangle$, which is

$$\tilde{D}^T(f - Au^s). \tag{47}$$

The corresponding stiffness matrix employing the trial space (41) is

$$\tilde{D}^T A \tilde{D}, \tag{48}$$

and we can get the coarse-mesh solution based on (48) and (47),

$$(\tilde{D}^T A \tilde{D})^{-1} \tilde{D}^T(f - L_h u^s). \tag{49}$$

Next we represent the coarse-mesh solution (49) using the fine mesh basis $\psi_k(x)$, we get

$$e^s + e^{s,2} = \tilde{D}(\tilde{D}^T A \tilde{D})^{-1} \tilde{D}^T(f - L_h u^s). \tag{50}$$

We denote L_H^{-1} as the coarse-mesh solution operator,

$$L_H^{-1} = \tilde{D}(\tilde{D}^T A \tilde{D})^{-1} \tilde{D}^T, \tag{51}$$

then we have $e^s + e^{s,2} = L_H^{-1}(f - L_h u^s)$ in (50). We add it back to u^s in (35), and get

$$u^{(s)} = u^s + L_H^{-1}(f - L_h u^s). \tag{52}$$

Finally, we apply ν steps of smoothing operations to the approximate solution (52). Let S be iteration matrix in the smoothing operation, then we get the following iteration relation for Algorithm 2,

$$u^{s+1} = S^\nu(I - L_H^{-1} L_h)u^s + M, \tag{53}$$

where M is a constant vector that does not depend on u^s.

From (53), we can see that Algorithm 2 is a two-level multi-grid method that employs the multi-scale basis (28) and bubble basis (38) for the restriction and prolongation operators, which are $\tilde{D}^T$ and $\tilde{D}$ respectively.

Next we consider the proposed Algorithm 1. The only difference between Algorithm 1 and 2 is that in Algorithm 1, the multi-scale basis in the coarse-mesh solution operator L_H^{-1} is adapted to the approximate numerical solutions in previous iteration steps. To show this we consider replacing the multi-scale basis in the coarse-mesh correction (33) by $\phi_i^s(x)$, which are obtained according to (24) or (27).

Consider solving (33) using trial space $\tilde{V}_H^s$ (41) and a Galerkin projection to get $e_H^s(x) + e^{s,2}(x)$. Then

$$a(u(x) - u^{s-1}(x) - e_H^s(x) - e_H^{s,2}(x), v(x)) = 0, \quad \text{for} \quad v(x) \in V_H^s. \tag{54}$$

Recall that in our construction of the adaptive multi-scale basis $\phi_i^s(x)$, we require that they can recover the approximate solution $u^{s-1}(x)$ on Γ (23). Therefor we have $u^{s-1}(x) \in \tilde{V}_H^s$, and

$$u^{(s)}(x) = u^{s-1}(x) + e_H^s(x) + e_H^{s,2}(x) \in \tilde{V}_H^s. \tag{55}$$

The representation of $u^{(s)}(x)$ (55) and the orthogonality condition (54) together imply that $u^{(s)}(x)$ is just the Galerkin numerical solution to the original equation (1) employing the updated trial space $\tilde{V}_H^s$, which is the same as the approximate numerical solution $u^{(s)}(x)$ that we obtain in Algorithm 1, (20). Namely, if we replace the coarse mesh solution operator L_H^{-1} in Algorithm 2 using $(L_H^s)^{-1}$, we get Algorithm 1. So we conclude that Algorithm 1 is a nonlinear version of the two-level multi-grid method with the restriction and prolongation operators adapted to approximate solutions of the previous iteration steps.

Note that there is no explicit restriction operation in the proposed Algorithm 1, namely, we do not need to compute the fine mesh residual and pass it to coarse mesh as in (47). The reason that we can skip the restriction operation in Algorithm 1 is that we have required that the multi-scale basis $\phi_i^s(x)$ can recover the approximate numerical solution of the previous iteration step, $u^{s-1}(x)$, on Γ.

Similar to (53) in Algorithm 2, we get the following iteration relation for Algorithm 1,

$$u^s = S^{(\nu)}(I - (L_H^s)^{-1}L_h)u^{s-1} + M_s. \tag{56}$$

4.2. Estimates of the error decay in the two proposed algorithms

The exact fine-mesh solution $u^* = A^{-1}f$ is a fixed point of the iteration schemes (56) and (53). In this subsection, to simplify the notation, we will write L_H^{-1} in (53) as $(L_H^s)^{-1}$.

We have the following relation between the errors in two consecutive iteration steps,

$$u^{s+1} - u^* = S^{(\nu)}(I - (L_H^s)^{-1}L_h)(u^s - u^*). \tag{57}$$

We will provide convergence analysis of the proposed algorithms under the framework of the two-level multi-grid method, see [37], in which the iteration relation (57) is decomposed as (with $A = L_h$)

$$A(u^{s+1} - u^*) = (AS^\nu)[L_h^{-1} - (L_H^s)^{-1}][A(u^s - u^*)]. \tag{58}$$

Then we have the following estimate on the decay of the l^2 norm of the residual,

$$\|A(u^{s+1} - u^*)\| \le \|AS^\nu\|\|L_h^{-1} - (L_H^s)^{-1}\|\|A(u^s - u^*)\|, \tag{59}$$

where $\|\cdot\|$ is the discrete l^2 norm and the corresponding induced operator norm. We need

$$\|AS^\nu\|\|L_h^{-1} - (L_H^s)^{-1}\| < \rho < 1 \tag{60}$$

for the robust convergence of the iteration scheme (58). Classical convergence analysis of the two-level multi-grid methods estimates the two parts in (60) separately.

For the first part, one needs the following smoothing property,

$$\|AS^\nu\| = O(\frac{1}{\nu}), \quad \nu \to +\infty, \tag{61}$$

which requires that the smoothing operation can effectively damp the high frequency modes in the error.

For the second part in (60), classical convergence analysis requires the approximation property

$$\|L_h^{-1} - (L_H^s)^{-1}\| \le CH^2h^{-2}. \tag{62}$$

Note that $\|L_h^{-1}\| = O(h^{-2})$, and (62) requires $O(H^2)$ accuracy of the Galerkin solution (7) in $L^2(D)$,

$$\|u(x) - u_H^{MS}(x)\|_{L^2(D)} \le CH^2\|f(x)\|_{L^2(D)}. \tag{63}$$

For the multi-scale PDE with rough coefficient $a(x)$, the smoothing property (61) still holds for several types of smoothing operators. However, due to the lack of regularity of the solution, the approximation property (62) fails if piecewise linear basis functions are used in the coarse-mesh solving.

To recover the robust convergence (60) as for smooth coefficient problems, one needs the approximation property of the multi-scale basis functions $\phi_i^s(x)$ and $\phi_i(x)$ to the solution space of the equation, (63).

13

1. In Algorithm 2, if we have constructed a set of multi-scale basis functions $\phi_i(x)$ that can achieve the desired approximation property (63) to the solution space of (1), then there exists ν^* that depends on H/h, such that for $\nu > \nu^*$, one can obtain the robust convergence (60) for Algorithm 2.

   In the periodic homogenization setting, it was shown in [44] that the multi-scale basis functions (5) with linear local boundary conditions have good approximation property to the solution space of (1). We will see in the numerical results section that even for some rough coefficient problems without clear scale-separation, Algorithm 2 employing multi-scale basis functions $\phi_i(x)$ with linear local boundary conditions also performs significantly better than the classical two-level multi-grid method.

2. In Algorithm 1, the multi-scale basis functions $\phi_i^s(x)$ are constructed based on the approximate solutions in previous iteration steps, which can be viewed as sampled solutions to (1). Success of multi-scale methods that employ sampled solutions to construct local multi-scale basis can be explained using the harmonic coordinates. We will prove the approximation property (63) for multi-scale basis constructed according to (27) under certain assumptions of the harmonic coordinates in the next subsection.

   One key idea of Algorithm 1 is using approximate solutions of previous iteration steps as sampled solutions to construct the multi-scale basis, instead of generating sampled solutions by solving (1) with random forcing terms. The latter has the same computational complexity as the original problem.

3. Once we have generated multi-scale basis functions in Algorithm 1 that can achieve good approximation property to the solution space of the equation, (63), we can then switch to Algorithm 2 with this set of multi-scale basis, since Algorithm 2 is computationally more efficient for each iteration step.

Remark 3 *The adaptive multi-scale method in [47] can be directly put into the two-level multi-grid method framework as (53). The multi-scale methods in [25, 54] cannot obtain the fine mesh solution, thus these methods are not fine mesh solvers of the equation as two-level multi-grid method. However, the oversampling method employed there is similar to blockwise smoothing operations. With suitable modifications, these multi-scale methods may be analyzed as nonlinear two-level multi-grid methods as (56).*

Remark 4 *In (58), in order to effectively reduce the error $A(u^* - u^s)$, we would like to make $(L_h^{-1} - (L_H^s)^{-1})A(u^* - u^{s-1})$ as small as possible. This requires that the adaptively constructed multi-scale basis functions $\phi_i^s(x)$ approximate the error $u(x) - u^{s-1}(x)$. Recall that the multi-scale basis functions $\phi_i^s(x)$ are constructed based on approximate solutions in previous iteration steps $u^k(x), k < s$. It is not obvious that the multiscale basis functions based on $u^k(x), k < s$ would give a good approximation of the error $u(x) - u^{s-1}(x)$. In other words, this adaptively constructed multi-scale basis (27) may not have clear advantage in approximating $u(x) - u^{s-1}(x)$ over multi-scale basis functions constructed based on other sampled solutions of (1). However, in our numerical results in subsection (5.2), we will see that for the high contrast interface problem, Algorithm 1 actually performs better than Algorithm 2 that employs a set of multi-scale basis functions constructed from two randomly generated sampled solutions of (1).*

4.3. Approximation property of the multi-scale basis $\phi_i^s(x)$

We employ the harmonic coordinates to establish the approximation property of multi-scale basis functions constructed from (27). The $a(x)$-harmonic coordinates

$$F(x) = (F_1(x), F_2(x)),\tag{64}$$

are defined as the solutions to the following equations,

$$\begin{cases} -\mathrm{div}(a(x)\nabla F_i(x)) = 0; \\ F_i(x)|_{\partial D} = x_i. \end{cases}\tag{65}$$

A remarkable property proved in [59] is that under certain assumptions about $F(x)$, the solution to (1), $u(x)$, gains an order of regularity with respect to the harmonic coordinates,

$$\|u \circ F^{-1}(x)\|_{W^{2,p}(D)} \le C(p)\|f \circ F^{-1}(x)\|_{L^p(D)}, \quad p > 2.\tag{66}$$

We consider two sampled solutions of the equation (1), $v_1(x)$ and $v_2(x)$,

$$-\text{div}(a(x)\nabla v_1(x)) = g_1(x), \quad -\text{div}(a(x)\nabla v_2(x)) = g_2(x), \quad v_{1,2}(x)|_{\partial D} = 0, \tag{67}$$

where $v_1(x)$ and $v_2(x)$ are normalized based on

$$\|g_1(x)\|_{L^p(D)} = 1, \quad \|g_2(x)\|_{L^p(D)} = 1. \tag{68}$$

We build multi-scale basis based on $v_1(x)$, $v_2(x)$ following the procedure outlined in (25) and (27), and show the approximation property (63) of the resulting multi-scale basis $\phi_i(x)$, $i = 1, \ldots, n$. These basis functions are approximately local linear combinations of $v_1(x)$ and $v_2(x)$, (25). Our basic idea is the following: we consider approximating the solution $u(x)$ using the multi-scale basis $\phi_i(x)$ in the harmonic coordinates; on each $\tilde{D}_i = F(D_i)$, $\tilde{v}_1(x) = v_1 \circ F^{-1}(x)$ and $\tilde{v}_2(x) = v_2 \circ F^{-1}(x)$ has $W^{2,p}$ regularity, thus can be approximated by linear functions; the multi-scale basis functions in the harmonic coordinates, namely, $\tilde{\phi}_i(x) = \phi_i \circ F^{-1}(x)$, are approximately linear combinations of $\tilde{v}_1(x)$ and $\tilde{v}_2(x)$, thus are approximately local linear functions; $\tilde{u}(x) = u \circ F^{-1}(x)$ has $W^{2,p}$ regularity and can be approximated by the constructed $\tilde{\phi}_i(x)$. Then we pass the estimates in the harmonic coordinates to the physical space to get (63).

To facilitate the error estimation procedure, we make the following two relatively strong assumptions on the stability of the harmonic coordinates (64) and the sampled solutions (67).

1. The solution in the harmonic coordinates, $u \circ F^{-1}(x)$, satisfies a non-divergence form elliptic equation, and to obtain the regularity result (66), one needs the Cordes condition [52]. This requires estimates of $\nabla F(x)$. And to pass the estimates in the harmonic coordinates to the physical space, one also needs estimates of $\nabla F(x)$. Here for simplicity, we make a strong assumption that

$$\nabla F(x), \quad (\nabla F(x))^{-1} \in L^\infty(D). \tag{69}$$

For each coarse element D_i with node points x_{i_j}, $j = 1, 2, 3$, we denote the triangle of $\tilde{x}_{i_j} = F(x_{i_j})$ as $\triangle \tilde{x}_{i_1} \tilde{x}_{i_2} \tilde{x}_{i_3}$. To obtain the desired estimates in the harmonic coordinates, we assume that

$$\triangle \tilde{x}_{i_1} \tilde{x}_{i_2} \tilde{x}_{i_3} \text{ has bounded aspect ratio.} \tag{70}$$

2. On each coarse element D_i, the interpolation of $u(x)$ using the basis $\bar{\phi}_{i_j}(x)$ can be written as

$$c_i^0 + c_i^1 v_1(x) + c_i^2 v_2(x). \tag{71}$$

To obtain the estimates in the harmonic coordinates, we need the stability of the interpolation and some estimates on c_i^j. We denote the derivatives of the linear interpolation of $u \circ F^{-1}(x)$, $v_1 \circ F^{-1}(x)$ and $v_2 \circ F^{-1}(x)$ on the node points $\tilde{x}_{i_j}$, $j = 1, 2, 3$ as k_i^u, $k_i^{v_1}$ and $k_i^{v_2}$ respectively.

Then the coefficients c_i^j in (71) can be obtained as

$$(c_i^1, c_i^2)^T = [k_i^{v_1}, k_i^{v_2}]^{-1} k_i^u. \tag{72}$$

For the stability of the interpolation, we assume that

$$C_i = \|[k_i^{v_1}, k_i^{v_2}]^{-1}\| \le C. \tag{73}$$

With the two major assumptions made above, and some other technical assumptions, we can show that using the multi-scale basis obtained from $v_1(x)$, $v_2(x)$, the Galerkin numerical solution $u_H^{MS}(x)$ satisfies

$$\|u(x) - u_H^{MS}(x)\|_{L^2(D)} \le CH^2 \|f(x)\|_{L^2(D)}. \tag{74}$$

The proof of (74) is given in the appendix. Then if some linear combinations of the sampled solutions $u^s(x)$ and $u^{s-1}(x)$ satisfy the same stability condition as $v_1(x)$, $v_2(x)$ in (73), we can recover the approximation property (62) as for smooth coefficient problems. We have the following Theorem.

15

Remark 5 *For arbitrarily rough coefficient $a(x)$, one can only get that $F(x)$ is Hölder continuous and cannot get (69). However, under some integrability condition of $\nabla F(x)$, the same estimate procedure in this subsection still carries through, but we can only get weaker approximation property than (63).*

Theorem With the assumptions made above, the adaptive multi-scale basis $\phi_i^s(x)$ constructed following the procedure outlined in (25) and (27) in Algorithm 1 has the apprxomation property (63). And there exist ν^* that depends on H/h, such that for $\nu > \nu^*$, we can achieve the convergence (60) in Algorithm 1,

$$\|Au^s - f\| \leq \rho^s \|Au^0 - f\|, \quad \rho < 1. \tag{75}$$

If there exist a set of multi-scale basis $\phi_1(x), \ldots, \phi_n(x)$, which has the approximation property (63) to the solution space of (1). Then employing this set of basis in Algorithm 2, we have that there exists ν^* that depends on H/h, such that for $\nu > \nu^*$, we can achieve the robust convergence (60) for Algorithm 2.

$$\|Au^s - f\| \leq \rho^s \|Au^0 - f\|, \quad \rho < 1. \tag{76}$$

Remark 6 *Recall that $C_i = \|[k_i^{v_1}, k_i^{v_2}]^{-1}\|^{-1}$. C_i will be large if (1) $k_i^{v_1}$ and $k_i^{v_2}$ are close to being align to each other; or (2) $k_i^{v_j}$ is small. To avoid the first case, we want $v_1(x)$ are $v_2(x)$ to be independent sampled solutions to (1). To avoid the second case, we want $\|\nabla v_j(x)\|$ to be large.*

Remark 7 *Each iteration step in Algorithm 2 has the same computational cost as the classical two-level multi-grid method. On the other hand, Algorithm 1 is more expensive than Algorithm 2 since one needs to construct the adaptive multi-scale basis functions and the corresponding stiffness matrix in each iteration step. In the multi-query setting where the equation (1) needs to be solved for multiple times with different forcing functions, one can construct a set of multi-scale basis functions with good approximation property (63) by employing several steps of Algorithm 1 iterations in the first query, and then switch to Algorithm 2 with this set of multi-scale basis functions. Then the computational cost of the proposed algorithms will be the same as classical two-level multi-grid method for smooth coefficient problems.*

5. Numerical Results

In this section, to demonstrate the efficiency of the proposed iteratively adaptive multi-scale finite element method, we apply it to several problems with multi-scale features, including a rough coefficient problem, a high-contrast interface problem, and a convection-dominated diffusion equation. In particular, we will show its advantage over the classical two-level multi-grid method as fine mesh solvers.

We employ a two-level discretization of the domain $D = [0, 1] \times [0, 1]$ as shown in Figure 1. The coarse mesh size is chosen to be $H = \frac{1}{32}$ and the underlying fine mesh size is chosen to be $h = \frac{1}{1024}$.

5.1. An elliptic PDE with rough coefficient

In this subsection, we consider an example of the multi-scale PDE (1) where the coefficient $a(x, y)$ is rough and does not have clear scale-separation. $a(x, y)$ and $f(x, y)$ are given by

$$a(x, y) = |\tilde{a}(x, y)| + 0.5, \quad f(x, y) = 1. \tag{77}$$

The values of $\tilde{a}(x, y)$ on the node points of an intermediate mesh of size $\frac{1}{128}$ are iid Gaussian random variables. And the values of $a(x, y)$ on the fine mesh node points are obtained using piecewise linear interpolation based on its values on the intermediate mesh node points. A typical realization of the coefficient $a(x, y)$ is shown in Figure 3, and we will solve equation (1) with this realization of $a(x, y)$.

We first numerically verify our stability assumption on the harmonic coordinates (69). We solve the harmonic coordinates (65) on the underlying fine mesh and numerically compute $\|\nabla F(x)\|$ and $\|(\nabla F(x))^{-1}\|$. They are plotted in Figure 4. We can see that $\|\nabla F(x)\|$ and $\|(\nabla F(x))^{-1}\|$ are both bounded, namely, the stability assumption (69) holds for this rough coefficient elliptic problem.
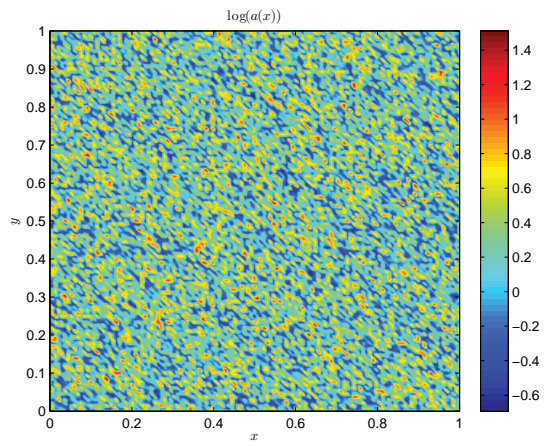
16

Figure 3: Rough coefficient $a(x, y)$ without scale-separation.
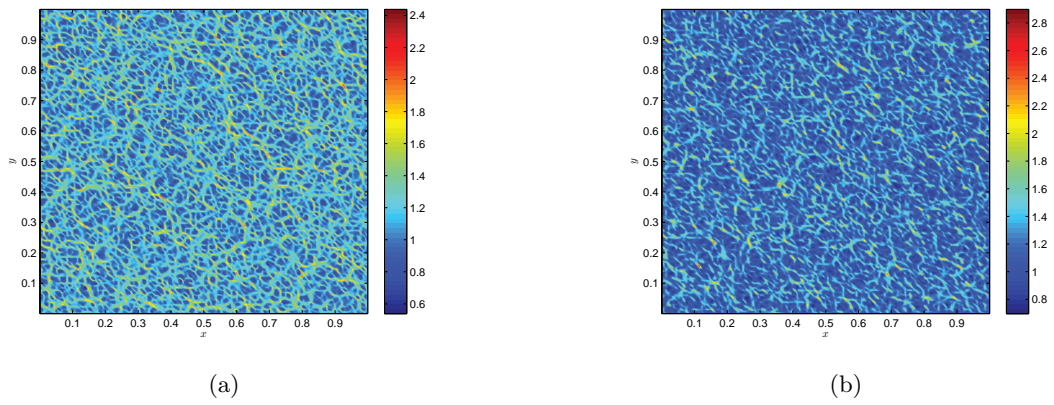

(a)

(b)

Figure 4: $\|\nabla F(x)\|$ and $\|(\nabla F(x))^{-1}\|$.

To start the iteration in Algorithm 1, we need to two sampled solutions to generate the multi-scale basis $\phi_i^1(x)$. We first solve the equation (1) using piecewise linear basis on the coarse mesh to get $u^{-1}(x)$. Then we apply 10 steps of Jacobi iteration on the fine mesh using $u^{-1}(x)$ as initial guess to obtain $u^0(x)$. With $u^{-1}(x)$ and $u^0(x)$, we build multi-scale $\phi_i^1(x)$ following the construction (25) and (27).

In the iteration of both Algorithm 1 and Algorithm 2, we choose $\nu = 10$ and the Gauss-Seidel iteration for the smoothing operations in step (21). There exist other smoothing methods that may perform better than the Gauss-Seidel iteration. However, for the purpose of comparing the proposed methods with the classical two-level multi-grid method, we will employ the simple Gauss-Seidel smoothing operation in implementation. We keep track of the l^2 norms of the error $\|u^s - u^*\|$ and the residual $\|Au^s - f\|$ in the iteration. We solve the multi-scale equation (1) using Algorithm 1, Algorithm 2, and the classical two-level multi-grid method under the same discretization. Note that the only difference between the three methods is the choice of the finite element basis functions in the coarse mesh solution operator $(L_H^s)^{-1}$. In Algorithm 1, we employ the adaptive multi-scale basis functions $\phi_i^s(x)$ and the bubble basis functions $\phi_j^b(x)$; in Algorithm 2, we employ the multi-scale basis functions with linear local boundary conditions $\phi_i(x)$ and $\phi_j^b(x)$; in the classical two-level multi-grid method we employ the coarse-mesh piecewise linear basis functions. The decay of $\|u^s - u^*\|$ and $\|Au^s - f\|$ for the three methods is shown in Figure 5.
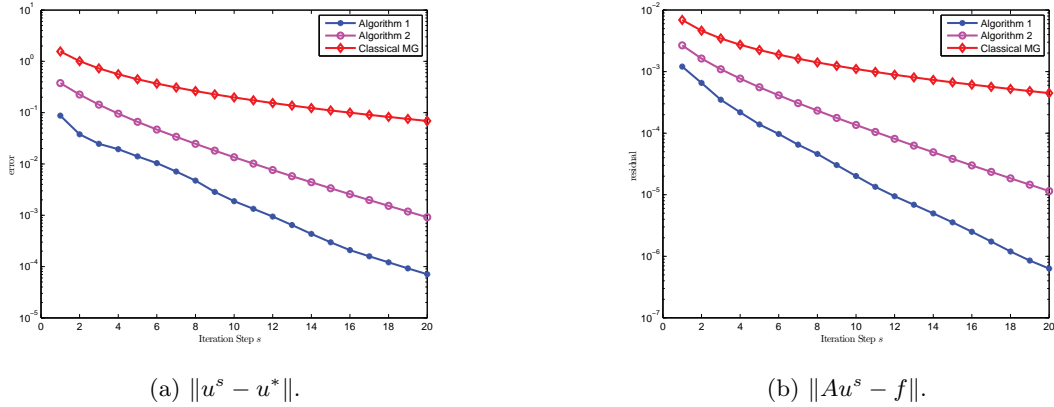

(a) $\|u^s - u^*\|$.   (b) $\|Au^s - f\|$.

Figure 5: Decay of the error $\|u^s - u^*\|$ and the residual $\|Au^s - f\|$.

According to the comparisons made in Figure 5, we can get the following conclusions.

1. The l^2 norms of the error and the residual both decay exponentially fast for Algorithm 1 and Algorithm 2. This agrees with our convergence analysis under the two-level multi-grid method framework.

2. The error and residual in Algorithm 1 and Algorithm 2 decay faster than the classical two-level multi-grid method. This is because the multi-scale basis functions employed in Algorithm 1 and 2 have better approximation property to the solution space of (1) than the coarse mesh piecewise linear basis.

In practical applications, even though the approximate solution $u^s(x)$ in Algorithm 1 and 2 converge to the underlying fine mesh solution $u_h(x)$ exponentially fast, we do not need to wait until

$$\|u^s(x) - u_h(x)\|_{L^2(D)} = O(h^2) \tag{78}$$

to stop the iteration. This is because the fine mesh size h is not chosen according to the desired order of accuracy, but according to the small-scale variations of the coefficient.

Even if we terminate the iteration based on the condition (78), we cannot get $O(h)$ accuracy in the numerical solution $u^s(x)$. Due to the oscillation in the coefficient $a(x)$, we have

$$\|u(x) - u_h(x)\|_{L^2(D)} \gg h^2. \tag{79}$$

18

If our desired order of accuracy is the coarse mesh size H, as in the multi-scale methods of the second category described in section 1.3, [7, 50, 43, 61], then we can terminate the iteration much earlier.

For this specific problem, with the Gauss-Seidel iteration for the smoothing operation, we obtain $\frac{H^2}{10}$ in $L^2(D)$ accuracy after 1 step of iteration in Algorithm 1 and 4 steps of iteration in Algorithm 2.

In the multi-scale methods, for example, [7, 50, 43, 61], constructing multi-scale basis functions with robust approximation property could be very expensive. Our method, Algorithm 1, bypasses this difficulty by adaptively updating the multi-scale basis to improve the approximation property of the basis.

5.2. A high-contrast interface problem

In this subsection, we consider an elliptic problem with a high-contrast interface. Note that the high-contrast inclusions violate the uniform ellipticity assumption (2). However, the proposed Algorithm 1 still works well and performs much better than the classical two-level multi-grid method.

The high-contrast coefficient $a(x, y)$ that we consider in this example is illustrated in Figure 6, and the right hand side function is chosen as $f(x, y) = 1$. As in the previous numerical example, we solve this high-contrast interface problem using Algorithm 1, Algorithm 2 and the classical two-level multi-grid method under the same discretization. We choose $\nu = 10$ and the Gauss-Seidel iteration in the smoothing operation. The decay of the error and the residual in the iteration is plotted in Figure 7.
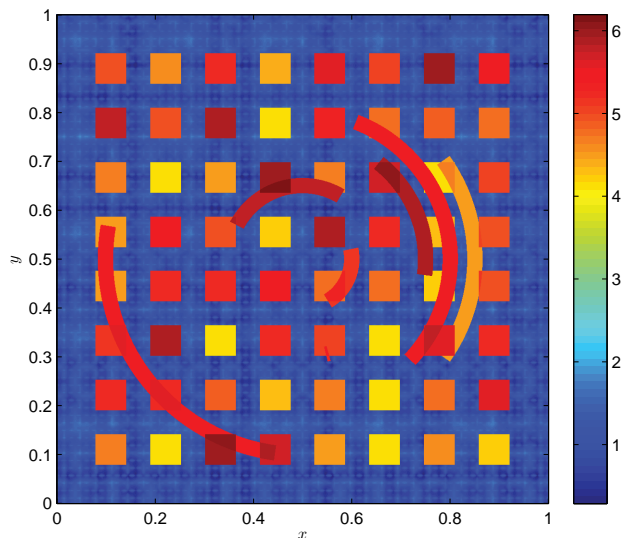

Figure 6: The high-contrast coefficient. $\log_{10} a(x)$.

For this high-contrast interface problem, no convergence is observed for the proposed Algorithm 2 employing multi-scale basis with linear local boundary conditions and the classical two-level multi-grid method. Due to the high contrast inclusions, the derivatives of the solution has large jump across the boundaries of the inclusions. Multi-scale basis with linear local boundary conditions cannot accurately capture these jumps, thus have very poor approximation property to the solution space of (1). This explains the failure of Algorithm 2 and the classical two-level multi-grid method observed in Figure 7.

For this high-contrast interface problem, the l^2 norms of the error $u^s - u^*$ and the residual $Au^s - f$ in the iteration of Algorithm 1 both converge exponentially fast. The adaptive multi-scale basis functions in Algorithm 1 are constructed based on sampled solutions of the equation, which can capture this jump of derivatives, thus have much better approximation property to the solution space of (1).

19

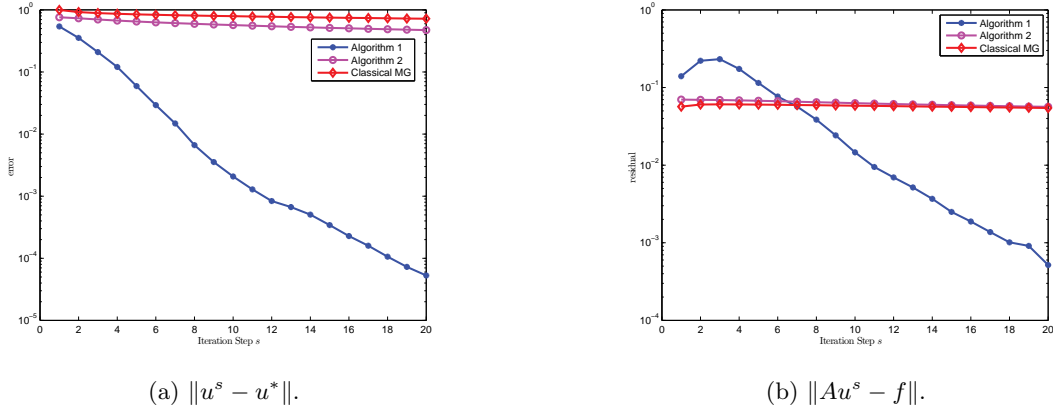(a) $\|u^s - u^*\|$.

(b) $\|Au^s - f\|$.

Figure 7: Decay of l^2 error and the residual for the high-contrast interface problem.

Next we consider constructing multi-scale basis functions $\phi_i(x), i = 1, 2, \ldots, n$ based on randomly generated sampled solutions of (1) following (25) and (27), and employing them in Algorithm 2 to solve the equation (1). For this purpose, we consider the following finite dimensional subspace of $L^p(D)$,

$$\{1, \sin(2\pi x), \cos(2\pi x), \ldots \sin(10\pi x), \cos(10\pi x)\} \otimes \{1, \sin(2\pi y), \cos(2\pi y), \ldots \sin(10\pi y), \cos(10\pi y)\}, \quad (80)$$

and generate random forcing functions $f(x)$ from this finite-dimensional subspace of $L^p(D)$.

We remark that solving (1) with randomly generated forcing has the same computational complexity as the original problem, and we design this example to compare the performance of Algorithm 1 and Algorithm 2. We consider two approaches for Algorithm 2. In the first approach, we generate a set of multi-scale basis functions before the iteration, and employ them in the coarse mesh solving step of Algorithm (2). In the second approach, we generate a new set of multi-scale basis functions for the coarse mesh solving operation in each iteration step of Algorithm 2. For this high-contrast interface problem, the convergence of the solutions is plotted in Figure 8. We observe that the iterative error corresponding to Algorithm 2 using the second approach indeed decays faster than that using the first approach. This shows that updating the multi-scale basis functions during the iteration indeed helps to reduce the iteration error. However, the iterative error in Algorithm 1 decays much faster than the iterative errors of Algorithm 2 even if we update multi-scale basis functions based on randomly generated sampled solutions (the second approach).

The complete understanding of this phenomenon requires further study. We would like to point out that that in our convergence analysis presented in section 4, we simply view the adaptive multi-scale basis functions $\phi_i^s(x)$ as generated from two sampled solutions. But there is an important factor that could potentially explain the difference in the performance between Algorithm 1 and Algorithm 2. It is important to note that the sample solutions that we use to construct the multiscale basis functions, $u^{s-2}(x)$ and $u^{s-1}(x)$, are tailored to the right hand side $Au^{s-1} - f$ in (58). It is possible that the adaptively constructed multiscale basis functions, $\phi_i^s(x)$, may give a better approximation of the solution corresponding to the specific right hand side function during the iteration process. The results in Figure 8 implies that the combination of fine-mesh smoothing operations with coarse mesh solving using the adaptive multi-scale basis constructed in (27) can more effectively reduce the error in $u^{s-1}(x)$, than the combination of fine mesh smoothing operations with multi-scale basis functions generated from random sampled solutions of the equation.
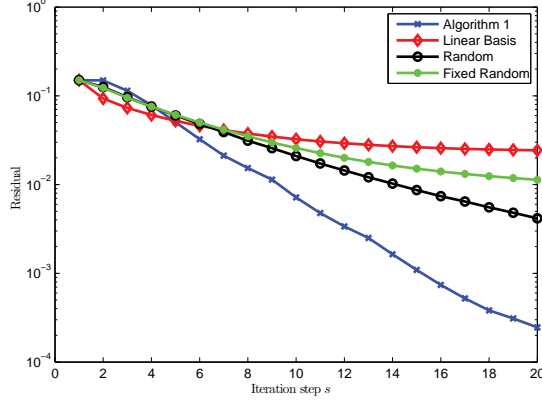
20

Figure 8: Comparison of Algorithm 1 with Algorithm (2) using multi-scale basis functions generated from random sampled solutions of (1). The line labeled "Fixed Random" corresponds the first approach described above for Algorithm 2, and the line labeled "Random" corresponds to the second approach for Algorithm 2. The line labeled "Linear Basis" corresponds to the classical two-level multi-scale method.

5.3. A convection-dominated diffusion equation

The proposed iteratively adaptive multi-scale finite element method, Algorithm 1, can also be naturally applied to solve convection dominated diffusion problems. In this subsection, we consider equation

$$\begin{cases} -\kappa\Delta u(x) + \mathbf{a}\cdot\nabla u(x) = f, & x\in\Omega, \\ u(x) = g(x), & x\in\partial\Omega, \end{cases} \tag{81}$$

where $\kappa = 10^{-6}$ is the diffusivity, $\mathbf{a} = (a_1, a_2)^T$ is the velocity field, $f(x)$ is a given source term, and g is given Dirichlet boundary condition. We consider the following four test cases as in [34]

1. Convective field with a fixed angle. $\Omega = [0,1]\times[0,1]$, $\mathbf{a} = (\sqrt{2}/2, \sqrt{2}/2)^T$, $f = 0$.
2. Thermal boundary layer problem. $\Omega = [0,1]\times[0,0.5]$, $\mathbf{a} = (2y, 0)^T$, $f = 0$.
3. Double ramp problem, $\Omega = [0,0.5]\times[0.5,1]\cup[0.5,1.5]\times[0,1]$, $\mathbf{a} = (1,0)^T$, and $f = 1$.
4. Smith-Hutton problem [69], $\Omega = [-1,1]\times[0,1]$, $\mathbf{a} = (2y(1-x^2), -2x(1-y^2))^T$, and $f = 0$.

The problem setting and boundary conditions for the four cases are illustrated in Figure 9.

In the discretization of the problem, the coarse mesh size is $H = \frac{1}{16}$ and fine mesh size is $h = \frac{1}{128}$. The streamline-crosswind diffusion method (SD/CD) [66] with a set of uniform rectangular meshes is used to construct the multi-scale basis with boundary conditions given by (24). And in the coarse mesh solving, we employ the following discretization of the equation: find $u^{(s)}(x) \in V_H^s$, such that

$$a_c(u^{(s)}(x), v(x)) = (f(x), v(x))_c, \quad \forall v(x)\in V_H^s, \tag{82}$$

where

$$a_c(u(x), v(x)) = a(u(x), v(x)) + \sum_i \tau_i(\mathbf{a}\cdot\nabla u, \mathbf{a}\cdot\nabla v)_{D_i} + \sum_i (\kappa_m - \kappa)(\mathbf{b}\cdot\nabla u, \mathbf{b}\cdot\nabla v)_{D_i} \tag{83}$$

and

$$(f, v)_c = (f, v) + \sum_i \tau_i(f, \mathbf{b}\cdot\nabla v)_{D_i}. \tag{84}$$

Here $a(u, v) = \kappa(\nabla u, \nabla v) + (\mathbf{a}\cdot\nabla u, \nabla v)$ is a bilinear form and (f, v) is a linear form.
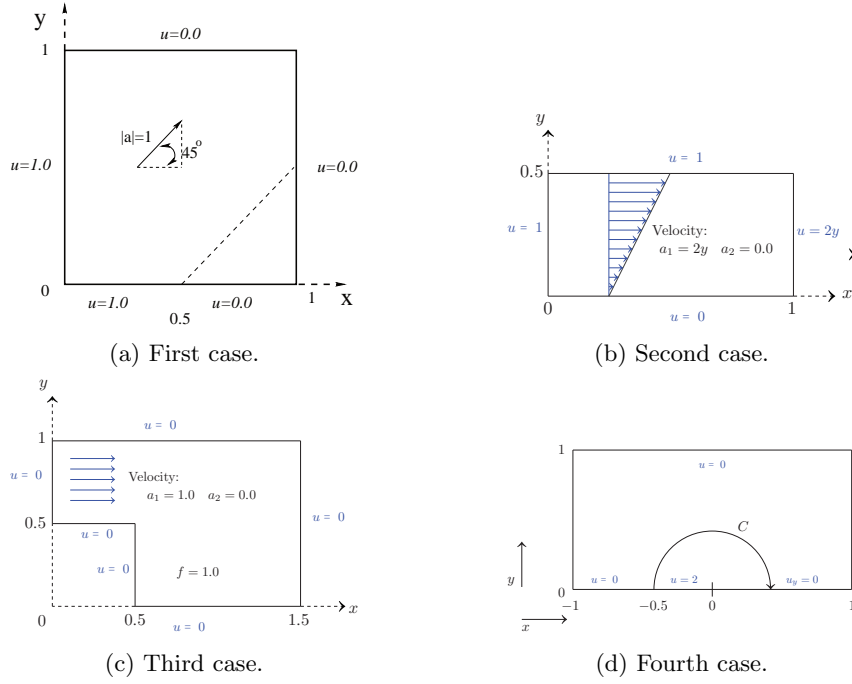
21

Figure 9: Boundary conditions for the four convection-diffusion problems.

The stabilization parameter τ_i is defined as

$$\tau_i = \frac{H_{D_i}^2}{6\kappa \max\{1, Pe_i\} + 6\kappa}, \tag{85}$$

with

$$Pe_i = \frac{|\mathbf{a}|_a H_{D_i}}{3\kappa}. \tag{86}$$

The artificial crosswind diffusion coefficient κ_m is given as

$$\kappa_m = \begin{cases} \kappa, & \text{if} \quad \kappa \geq H_{D_i}^{3/2}, \\ H_{D_i}^{3/2}, & \text{if} \quad \kappa < H_{D_i}^{3/2}. \end{cases} \tag{87}$$

And $\mathbf{b} = (-a_2, a_1)$ is the crosswind vector.

The initial and final MsFEM solutions for the 4 cases are plotted in Figure 10.

The initial MsFEM solutions are obtained using multi-scale basis with linear local boundary conditions. We observe that the accuracies of the MsFEM numerical solutions are significantly improved in the end of the iteration: the overshooting problem within the boundary layer regions is completely eliminated.

The smoothing operation (21) is chosen as the incomplete LU factorization without fill-in, namely, the ILU(0) method [65]. We choose $\nu = 1, 2, \ldots, 5$, and the stopping criteria for the Algorithm is $\|u^{s-1}(x) - u^s(x)\|_{L^2(\Omega)} \leq 10^{-4}$. To obtain the convergence, the steps of iterations required for the four cases are summarized in Table 1. For $\nu \geq 4$, the iteration converges within 6 steps for all the 4 cases.

6. Concluding Remarks

An iteratively adaptive multi-scale finite element is proposed, in which the multi-scale basis functions and the approximate solutions are updated simultaneously in each iteration step. The global information

22

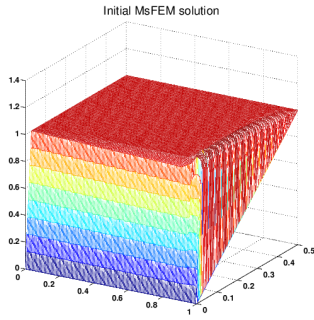| ν	Fixed angle	Thermal	Double ramp	Smith-Hutton
1	5	11	10	7
2	4	7	8	6
3	4	6	8	5
4	3	5	6	5
5	3	5	5	5

Table 1: Number of iteration steps to obtain convergence for the four tested cases.
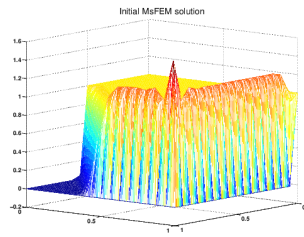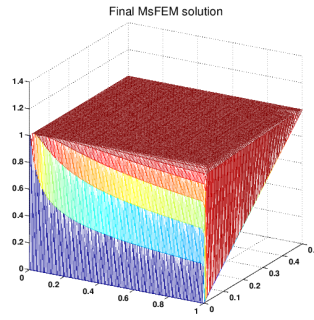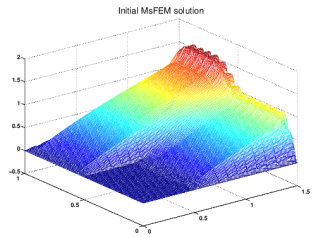

(a) Initial MsFEM solution. First case.
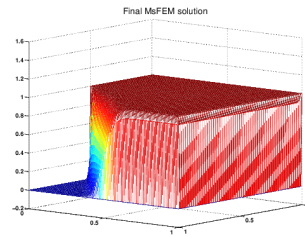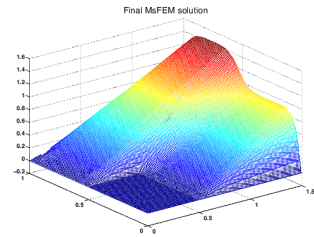
(b) Final MsFEM solution. First case.
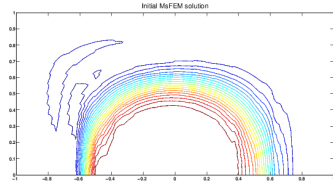
(c) Initial MsFEM solution. Second case.

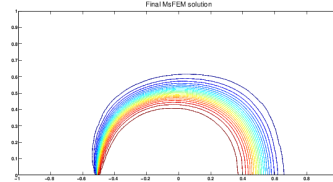(d) Final MsFEM solution. Second case.

(e) Initial MsFEM solution. Third case.

(f) Final MsFEM solution. Third case.

(g) Initial MsFEM solution. Fourth case.

(h) Final MsFEM solution. Fourth case.

Figure 10: Initial and final MsFEM solutions for the convection-dominated diffusion problems.

of the elliptic equation is transferred to the local boundary conditions of the multi-scale basis functions through coarse mesh solving using adaptive multi-scale basis and fine mesh smoothing operations.

A stabilized variant of the proposed algorithm is also derived, which is a two-level multi-grid method that employs a set of multi-scale basis functions for the restriction and prolongation operations. The proposed iteratively adaptive MsFEM can be viewed as a nonlinear version of the two-level multi-grid method.

Convergence analysis of the proposed algorithm is carried out in the two-level multi-grid method framework. And we employ the harmonic coordinates to establish the approximation property of the adaptively constructed multi-scale basis functions to the solution space of the equation.

Numerical results are presented for a rough coefficient problem, an elliptic interface problem and several convection dominated diffusion problems. Significant improvements over classical methods are observed for all these three types of PDE problems. In particular, the proposed Algorithm 1 has significant advantage over the classical two-level multi-grid method as a fine mesh solver of the multi-scale equation.

Acknowledgements

APPENDIX

Let $\phi_i(x)$, $i = 1, 2, \ldots, n$ be the multi-scale basis functions constructed using the sampled solutions $v_1(x)$, $v_2(x)$ obtained from (67) and following the procedure outlined in (25) and (27). Further, we denote $u_H^{MS}(x)$ as the Galerkin finite element solution of (1) using the resulting trial space

$$V_h = \text{span}\{\phi_1(x), \phi_2(x), \ldots, \phi_n(x)\}.$$

We will prove that

$$\|u(x) - u_H^{MS}(x)\|_{L^2(D)} \leq CH^2 \|f(x)\|_{L^2(D)}, \tag{88}$$

under the two assumptions in subsection 4.3, and some additional technical assumptions (93), (101), (103).

Proof:

We first consider the interpolation of the solution $u(x)$ using the non-conforming basis $\bar{\phi}_i(x)$. On the boundary of the coarse element D_i, the interpolation residual $e_i(x) = u(x) - \sum_{j=1}^{3} u(x_{i_j}) \bar{\phi}_{i_j}(x)$ is

$$e_i(x) = u(x) - c_i^0 - c_i^1 v_1(x) - c_i^2 v_2(x), \quad x \in \partial D_i. \tag{89}$$

Then employing the optimality property of multi-scale basis functions, (14), we have

$$\|e_i(x)\|_E^2 \leq \|u(x) - c_i^0 - c_i^1 v_1(x) - c_i^2 v_2(x)\|_E^2 + CH^2 \|f(x)\|_{L^2(D_i)}^2, \tag{90}$$

where $\|\cdot\|_E$ is the energy norm (8). To bound the energy norm of $E_i(x) = u(x) - c_i^0 - c_i^1 v_1(x) - c_i^2 v_2(x)$ in the above estimate, (90), we express this term in the harmonic coordinates, $\tilde{E}_i(x) = E_i \circ F^{-1}(x)$,

$$\tilde{E}_i(x) = \tilde{u}(x) - c_i^0 - c_i^1 \tilde{v}_1(x) - c_i^2 \tilde{v}_2(x), \quad x \in \tilde{D}_i. \tag{91}$$

We denote $\hat{v}_{i,1}(x)$, $\hat{v}_{i,2}(x)$ and $\hat{u}_i(x)$ as the linear interpolation of $\tilde{v}_1(x)$, $\tilde{v}_2(x)$ and $\tilde{u}(x)$ on the three node points, $\tilde{x}_{i_1}, \tilde{x}_{i_2}, \tilde{x}_{i_3}$. Then we can decompose (91) as

$$\tilde{E}_i(x) = [\tilde{u}(x) - c_i^0 - c_i^1 \hat{v}_{i,1}(x) - c_i^2 v_{i,2}] + [c_i^1(\hat{v}_{i,1}(x) - \tilde{v}_1(x)) + c_i^2(\hat{v}_{i,2}(x) - \tilde{v}_2(x))]. \tag{92}$$

The first term in (92) is simply $\tilde{u}(x) - \hat{u}_i(x)$, and to get an estimate of it on $\tilde{D}_i$, we need to put $\tilde{D}_i$ in a regular domain. Let $\hat{D}_i$ be the triangle which is the scaling of $\Delta \tilde{x}_{i_1} \tilde{x}_{i_2} \tilde{x}_{i_3}$ by a factor of 2, and has the same center as $\Delta \tilde{x}_{i_1} \tilde{x}_{i_2} \tilde{x}_{i_3}$. Then we assume that $\tilde{D}_i \subset \hat{D}_i$ and each $x \in D$ is covered by finite $\hat{D}_i$,

$$\tilde{D}_i \subset \hat{D}_i, \quad \max_{x \in D} \#\{i : x \in \hat{D}_i\} \leq C. \tag{93}$$

The element D_i, its image $\tilde{D}_i$, and the triangle $\hat{D}_i$ are illustrated in Figure 11.
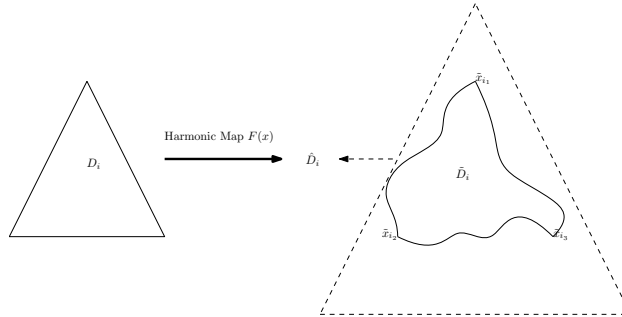

Figure 11: Coarse-mesh element in Harmonic Coordinate.

Then using the regularity result (66) and the assumption that $\hat{D}_i$ has bounded aspect ratio, we have

$$\|\tilde{u}(x) - c_i^0 - c_i^1 \hat{v}_{i,1}(x) - c_i^2 v_{i,2}(x)\|_{H^1(\tilde{D}_i)} \leq \|\tilde{u}(x) - c_i^0 - c_i^1 \hat{v}_{i,1}(x) - c_i^2 v_{i,2}(x)\|_{H^1(\hat{D}_i)} \leq CH\|\tilde{u}(x)\|_{H^2(\hat{D}_i)}. \quad (94)$$

Using the same assumptions as in (94), we can bound the second term in (92) by

$$\|c_i^1(\hat{v}_{i,1}(x) - \tilde{v}_1(x)) + c_i^2(\hat{v}_{i,2}(x) - \tilde{v}_2(x))\|_{H^1(\tilde{D}_i)} \leq |c_i^1| H \|\tilde{v}_1(x)\|_{H^2(\hat{D}_i)} + |c_i^2| H \|\tilde{v}_2(x)\|_{H^2(\hat{D}_i)}. \quad (95)$$

Putting estimates (94) and (95) together in the decomposition (92), we have

$$\|\tilde{E}_i(x)\|_{H^1(\tilde{D}_i)}^2 \leq CH^2\|\tilde{u}(x)\|_{H^2(\hat{D}_i)}^2 + C(c_i^1)^2 H^2 \|\tilde{v}_1(x)\|_{H^2(\hat{D}_i)}^2 + C(c_i^2)^2 H^2 \|\tilde{v}_2(x)\|_{H^2(\hat{D}_i)}^2. \quad (96)$$

Adding up the estimate (96) over each coarse element, we have

$$\sum_i \|E_i(x)\|_{H^1(D_i)}^2 \leq CH^2\|\tilde{u}(x)\|_{H^2(D)}^2 + C\max_i(c_i^1)^2 H^2 \|\tilde{v}_1(x)\|_{H^2(D)}^2 + C\max_i(c_i^2)^2 H^2 \|\tilde{v}_2(x)\|_{H^2(D)}^2. \quad (97)$$

Using the regularity result (66) and the normalization condition (68), we have

$$\|\tilde{u}(x)\|_{H^2(D)} \leq C\|f(x)\|_{L^p(D)}; \quad \|\tilde{v}_1(x)\|_{H^2(D)}, \|\tilde{v}_2(x)\|_{H^2(D)} \leq C. \quad (98)$$

For the coefficients c_i^1 and c_i^2, according to (72), we have

$$c_i^1, c_i^2 \leq C\|\tilde{u}(x)\|_{C^1(D)} \leq C\|f(x)\|_{L^p(D)}. \quad (99)$$

Putting the estimates (98), (99) and (97) in (90), we have

$$\sum_i \|e_i(x)\|_{H^1(D_i)}^2 \leq CH^2\|f(x)\|_{L^p(D)}^2. \quad (100)$$

To obtain the approximation property of the conforming basis (27), we simply assume that the error will not be amplified due to the averaging,

$$\|u(x) - \sum_i u(x_i)\phi_i(x)\|_{H^1(D)}^2 \leq C\sum_i \|e_i(x)\|_{H^1(D_i)}^2. \quad (101)$$

Then using the optimal property (8), (100) and (101), we have

$$\|u_H^{MS}(x) - u(x)\|_{H^1(D)} \leq CH\|f(x)\|_{L^p(D)}. \quad (102)$$

Finally, we assume that for all the $f(x)$ that occur in the iteration of Algorithm 1, we have

$$\|f(x)\|_{L^p(D)} \leq C\|f(x)\|_{L^2(D)}, \quad (103)$$

which requires that the $f(x)$ is wide spread. With (103), we get

$$\|u(x) - u_H^{MS}\|_{H^1(D)} \leq CH\|f(x)\|_{L^2(D)}. \quad (104)$$

Then using a Aubin-Nitsche duality argument, we get (88). The assumption (103) requires

$$\|u(x) - u_H^{MS}(x)\|_{L^p(D)} \leq C\|u(x) - u_H^{MS}(x)\|_{L^3(D)} \quad (105)$$

in the duality argument. We keep track of the constants C in (103) and (105) in the iteration of Algorithm 1 for the rough coefficient problem in subsection (5.1), and they are plotted in Figure 12.
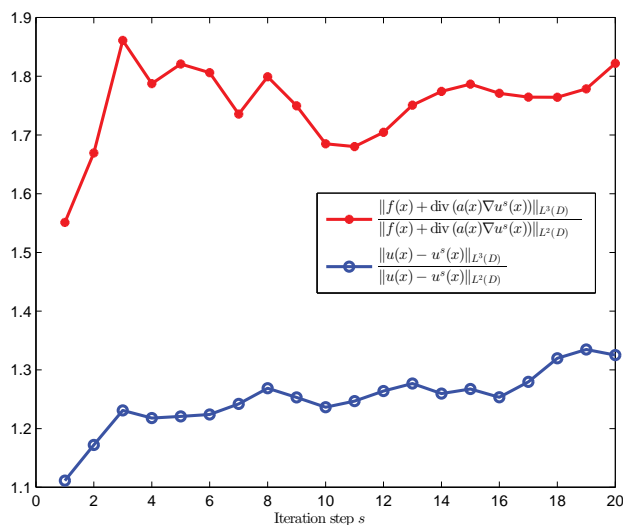
Figure 12: The constants in (103) and (105).

References

[1] RE Alcouffe, A Brandt, JE Dendy, Jr, and JW Painter. The multi-grid method for the diffusion equation with strongly discontinuous coefficients. *SIAM Journal on Scientific and Statistical Computing*, 2(4):430–454, 1981.

[2] G Allaire. Homogenization and two-scale convergence. *SIAM Journal on Mathematical Analysis*, 23(6):1482–1518, 1992.

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

[4] A Averbuch, G Beylkin, R Coifman, and M Israeli. Multiscale inversion of elliptic operators. *Wavelet Analysis and Its Applications*, 7:341–359, 1998.

[5] 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(4):945–981, 1994.

[6] 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(02):493–515, 2014.

[7] I Babuska and R Lipton. Optimal local approximation spaces for generalized finite element methods with application to multiscale problems. *Multiscale Modeling & Simulation*, 9(1):373–406, 2011.

[8] I Babuška and J Osborn. Generalized finite element methods: their performance and their relation to mixed methods. *SIAM Journal on Numerical Analysis*, 20(3):510–536, 1983.

[9] I Babuška and J Osborn. Can a finite element method perform arbitrarily badly? *Mathematics of Computation of the American Mathematical Society*, 69(230):443–462, 2000.

[10] M Bebendorf. *Hierarchical matrices*. Springer, 2008.

[11] A Bensoussan, JL Lions, and G Papanicolaou. *Asymptotic analysis for periodic structures*, volume 374. American Mathematical Soc., 2011.

[12] G Beylkin and N Coult. A multiresolution strategy for reduction of elliptic pdes and eigenvalue problems. *Applied and Computational Harmonic Analysis*, 5(2):129–155, 1998.

[13] A Brandt. Multi-level adaptive technique (mlat) for fast numerical solution to boundary value problems. In *Proceedings of the Third International Conference on Numerical Methods in Fluid Mechanics*, pages 82–89. Springer, 1973.

[14] A Brandt. Algebraic multigrid theory: The symmetric case. *Applied mathematics and computation*, 19(1):23–56, 1986.

[15] A Brandt, S McCoruick, and J Huge. Algebraic multigrid (amg) for sparse matrix equations. *Sparsity and its Applications*, page 257, 1985.

[16] ME Brewster and G Beylkin. A multiresolution strategy for numerical homogenization. *Applied and Computational Harmonic Analysis*, 2(4):327–349, 1995.

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

[18] IL Chern and YC Shu. A coupling interface method for elliptic interface problems. *Journal of computational physics*, 225(2):2138–2174, 2007.

[19] CC Chu, I Graham, and TY Hou. A new multiscale finite element method for high-contrast elliptic interface problems. *Mathematics of Computation*, 79(272):1915–1955, 2010.

[20] D Cioranescu and P Donato. Introduction to homogenization. 2000.

27

[21] R. Codina. A discontinuity-capturing crosswind-dissipation for the finite element solution of the convection-diffusion equation. *Computer methods in applied mechanics and engineering*, 110:325–342, 1993.

[22] E De Giorgi. New problems in γ-convergence and g-convergence. *Free boundary problems*, 2:183–194, 1980.

[23] E De Giorgi. Sulla convergenza di alcune successioni dintegrali del tipo dellarea. *Ennio De Giorgi*, page 414, 2006.

[24] M Dorobantu and B Engquist. Wavelet-based numerical homogenization. *SIAM Journal on Numerical Analysis*, 35(2):540–559, 1998.

[25] LJ Durlofsky, Y Efendiev, and V Ginting. An adaptive local–global multiscale finite volume element method for two-phase flow simulations. *Advances in Water Resources*, 30(3):576–588, 2007.

[26] Y Efendiev, J Galvis, and PS Vassilevski. Spectral element agglomerate algebraic multigrid methods for elliptic problems with high-contrast coefficients. In *Domain decomposition methods in science and engineering XIX*, pages 407–414. Springer, 2011.

[27] Y Efendiev, V Ginting, TY Hou, and R Ewing. Accurate multiscale finite element methods for two-phase flow simulations. *Journal of Computational Physics*, 220(1):155–174, 2006.

[28] Y Efendiev, TY Hou, and V Ginting. Multiscale finite element methods for nonlinear problems and their applications. *Communications in Mathematical Sciences*, 2(4):553–589, 2004.

[29] Y Efendiev, TY Hou, and XH Wu. Convergence of a nonconforming multiscale finite element method. *SIAM Journal on Numerical Analysis*, 37(3):888–910, 2000.

[30] B Engquist and E Luo. Convergence of a multigrid method for elliptic equations with highly oscillatory coefficients. *SIAM journal on numerical analysis*, 34(6):2254–2273, 1997.

[31] C Farhat, I Kalashnikova, and R Tezaur. A higher-order discontinuous enrichment method for the solution of high péclet advection–diffusion problems on unstructured meshes. *International Journal for Numerical Methods in Engineering*, 81:604–636, 2010.

[32] RP Fedorenko. A relaxation method for solving elliptic difference equations. *USSR Computational Mathematics and Mathematical Physics*, 1(4):1092–1096, 1962.

[33] LP Franca, SL Frey, and TJR Hughes. Stabilized finite element methods: I. Application to the advective-diffusive model. *Computer Methods in Applied Mechanics and Engineering*, 95:253–276, 1992.

[34] LP Franca and FN Hwang. Refining the submesh strategy in the two-level finite element method: application to the advection–diffusion equation. *International Journal for Numerical Methods in Fluids*, 39:161–187, 2002.

[35] A Gloria. An analytical framework for the numerical homogenization of monotone elliptic operators and quasiconvex energies. *Multiscale Modeling & Simulation*, 5(3):996–1043, 2006.

[36] W Hackbusch. A fast iterative method for solving poissons equation in a general region. In *Numerical treatment of differential equations*, pages 51–62. Springer, 1978.

[37] W Hackbusch. *Multi-grid methods and applications*, volume 4. Springer Science & Business Media, 2013.

[38] W Hackbusch and L Grasedyck. An introduction to hierarchical matrices. *Mathematica Bohemica*, 127(2):229–241, 2002.

[39] H Hajibeygi, G Bonfigli, MA Hesse, and P Jenny. Iterative multiscale finite-volume method. *Journal of Computational Physics*, 227(19):8604–8621, 2008.

[40] H Hajibeygi and P Jenny. Adaptive iterative multiscale finite volume method. *Journal of Computational Physics*, 230:628–643, 2011.

[41] P Henning and D Peterseim. Oversampling for the multiscale finite element method. *Multiscale Modeling & Simulation*, 11(4):1149–1175, 2013.

[42] TY Hou and P Liu. A heterogeneous stochastic fem framework for elliptic pdes. *Journal of Computational Physics*, 281:942–969, 2015.

[43] TY Hou and P Liu. Optimal local multi-scale basis functions for linear elliptic equations with rough coefficient. *arXiv preprint arXiv:1508.00346*, 2015.

[44] TY Hou and XH Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *Journal of computational physics*, 134(1):169–189, 1997.

[45] TY Hou, XH 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(227):913–943, 1999.

[46] TY Hou, XH 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(2):185–205, 2004.

[47] P Jenny, SH Lee, and HA Tchelepi. Adaptive multiscale finite-volume method for multiphase flow and transport in porous media. *Multiscale Modeling & Simulation*, 3(1):50–64, 2005.

[48] VV Jikov, SM Kozlov, and OA Oleinik. *Homogenization of differential operators and integral functionals*. Springer Science & Business Media, 2012.

[49] Z Li. The immersed interface method using a finite element formulation. *Applied Numerical Mathematics*, 27:253–267, 1998.

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

[51] J Mandel, M Brezina, and P Vaněk. Energy optimization of algebraic multigrid bases. *Computing*, 62(3):205–228, 1999.

[52] A Maugeri, K Palagachev, and LG Softova. *Elliptic and Parabolic Equations with Discontinuous Coefficients*, volume 109. wiley, 2000.

[53] JM Melenk and I Babuška. The partition of unity finite element method: basic theory and applications. *Computer methods in applied mechanics and engineering*, 139(1):289–314, 1996.

[54] R Millward. *A new adaptive multiscale finite element method with applications to high contrast interface problems*. PhD thesis, University of Bath, 2011.

[55] F Murat. Compacité par compensation. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 5(3):489–507, 1978.

[56] F Murat and L Tartar. *H-convergence*. Springer, 1997.

[57] H Owhadi. Bayesian numerical homogenization. *Multi-scale Modeling & Simulation*, 13(13):812–828, 2015.

[58] H Owhadi. Mult-grid with rough coefficients and multiresolution operator decomposition from hierarchical information games. *arXiv preprint arXiv:1503.03467*, 2015.

[59] H Owhadi and L Zhang. Metric-based upscaling. *Communications on Pure and Applied Mathematics*, 60(5):675–723, 2007.

[60] H Owhadi and L Zhang. Localized bases for finite-dimensional homogenization approximations with nonseparated scales and high contrast. *Multiscale Modeling & Simulation*, 9(4):1373–1398, 2011.

[61] H Owhadi, L Zhang, and L Berlyand. Polyharmonic homogenization, rough polyharmonic splines and sparse super-localization. *ESAIM: Mathematical Modelling and Numerical Analysis*, 48(02):517–552, 2014.

[62] N Panasenko and NS Bakhvalov. Homogenization: Averaging processes in periodic media: Mathematical problems in the mechanics of composite materials, 1989.

[63] D Peterseim. Variational multiscale stabilization and the exponential decay of fine-scale correctors. *arXiv preprint arXiv:1505.07611*, 2015.

[64] JW Ruge and K Stüben. Algebraic multigrid. *Multigrid methods*, 3:73–130, 1987.

[65] Y Saad. *Iterative methods for sparse linear systems*. Siam, 2003.

[66] YT Shih and HC Elman. Iterative methods for stabilized discrete convection-diffusion problems. *IMA Journal of Numerical Analysis*, 20:333–358, 2000.

[67] S Spagnolo. Sulla convergenza di soluzioni di equazioni paraboliche ed ellittiche. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 22(4):571–597, 1968.

[68] S Spagnolo. Convergence in energy for elliptic operators. *Numerical Solutions of Partial Differential Equations III, Acad. Press, New York*, 1976.

[69] U Trottenberg, C Oosterlee, and A Schűller. *Multigrid*. Academic Press, 2000.

[70] P Vasilevski and J Wang. *Stabilizing the hierarchical basis by approximate wavelets*. Department of Mathematics, University of California, Los Angeles, 1995.

[71] PS Vassilevski. On two ways of stabilizing the hierarchical basis multilevel methods. *SIAM review*, 39(1):18–53, 1997.

[72] PS Vassilevski. General constrained energy minimization interpolation mappings for amg. *SIAM Journal on Scientific Computing*, 32(1):1–13, 2010.

[73] PS Vassilevski and J Wang. Stabilizing the hierarchical basis by approximate wavelets ii: Implementation and numerical results. *SIAM Journal on Scientific Computing*, 20(2):490–514, 1998.

[74] WL Wan, TF Chan, and B Smith. An energy-minimizing interpolation for robust multigrid methods. *SIAM Journal on Scientific Computing*, 21(4):1632–1649, 1999.

[75] J Xu and Y Zhu. Uniform convergent multigrid methods for elliptic problems with strongly discontinuous coefficients. *Mathematical Models and Methods in Applied Sciences*, 18(01):77–105, 2008.

[76] J Xu and L Zikatanov. On an energy minimizing basis for algebraic multigrid methods. *Computing and Visualization in Science*, 7(3-4):121–127, 2004.