RANDOMIZED SAMPLING FOR BASIS FUNCTIONS CONSTRUCTION IN GENERALIZED FINITE ELEMENT METHODS

KE CHEN, QIN LI, JIANFENG LU, AND STEPHEN J. WRIGHT

Abstract. In the context of generalized finite element methods for elliptic equations with rough coefficients $a(x)$, efficiency and accuracy of the numerical method depend critically on the use of appropriate basis functions. This work explores several random sampling strategies for construction of basis functions, and proposes a quantitative criterion to analyze and compare these sampling strategies. Numerical evidence shows that the optimal basis functions can be well approximated by a random projection of generalized eigenvalue problem onto subspace of a-harmonic functions.

1. Introduction

The focus of the paper is the construction of basis functions for elliptic equations with rough coefficients in the framework of generalized finite element methods. Consider an elliptic partial differential equation

$$\begin{cases} -\operatorname{div}\left(a(x)\nabla u(x)\right) = f(x)\,, & x \in \Omega; \\ u(x) = 0\,, & x \in \partial\Omega, \end{cases} \tag{1.1}$$

with $f \in L^2(\Omega)$ and the coefficient $a \in L^\infty(\Omega)$ uniformly elliptic, that is, there exists $\alpha_*, \beta_* > 0$ such that $a(x) \in [\alpha_*, \beta_*]$ for all $x \in \Omega$. Note that we assume only L^∞ regularity of a, so the coefficient could be rather rough, which poses challenges to conventional numerical methods, such as the standard finite element method with local polynomial basis functions.

Numerical methods can be designed to take advantage of certain analytical properties of the problem (1.1). A classical example is when a is two-scale, that is, $a(x) = a_0(x, \frac{x}{\varepsilon})$ where $a_0(x, y)$ is 1-periodic with respect to its second argument. (Thus, ε characterizes explicitly the small scale of the problem.) Using the theory of homogenization [4, 24], several numerical methods have been proposed over the past decades to capture the homogenized solution of the problem and possibly also provide some microscopic information. Approaches of this type include the multiscale finite element method [8, 13, 14, 12] and the heterogeneous multiscale method [6, 7, 20].

Date: January 23, 2018.

The work of Q.L. is supported in part by a start-up fund from UW-Madison and National Science Foundation under the grant DMS-1619778. The work of J.L. is supported in part by the National Science Foundation under grant DMS-1454939. The work of S.W. is supported in part by NSF Awards IIS-1447449, 1628384, and 1634597 and AFOSR Award FA9550-13-1-0138. Both Q.L. and S.W. are supported by TRIPODS: NSF 1740707.

While one could apply methods designed for numerical homogenization to the cases of rough media ($a \in L^\infty$), the lack of favorable structural properties often degrades the efficiency and convergence rates. Various numerical methods have been proposed for L^∞ media, including the generalized finite element method [2], upscaling based on harmonic coordinate [22], elliptic solvers based on $\mathcal{H}$-matrices [3, 10], and Bayesian numerical homogenization [21], to name just a few. Our work lies in the framework of the generalized finite element method (gFEM) originally proposed in [2]. The idea is to approximate local solution space by constructing good basis functions locally and to use either the partition-of-unity or discontinuous Galerkin method to obtain a global discretization.

According to the partition-of-unity finite-element theory, which we will recall briefly in Section 2, the global error is controlled by the accuracy of the numerical local solution spaces. Thus, global performance of the method depends critically on efficient preparation of accurate local solution spaces. Towards this end, in [1], Babuška and Lipton studied the Kolmogorov width of a finite dimensional approximation to the optimal solution space, and showed that the distance decays almost exponentially fast, as we recall in Section 2. The basis construction algorithm proposed in [1] follows the analysis closely: a full list of a-harmonic functions (up to discretization) is obtained in each patch, and local basis functions are obtained by a "post-processing" step of solving a generalized eigenvalue problem to select modes with highest "energy ratios". Since the roughness of a necessitates a fine discretization in each patch, and thus a large number of a-harmonic functions per patch, the overall computational cost of the above strategy is high.

Our work is based on the gFEM framework [2] together with the concept of optimal local solution space via Kolmogorov width studied in [1]. Our main contribution in this work is to accelerate the procedure in [1] by incorporating random sampling techniques. Randomized algorithms have been shown to be powerful in reducing computational complexities in looking for low rank structures of matrices. Since the generalized eigenvalues decay almost exponentially fast, the local solution space is of low rank, and we find that random sampling approaches can effectively capture the local solution space. The efficiency of the approach certainly depends on the particular random sampling strategy employed; we explore several strategies and identify the most successful ones.

The idea of random sampling or oversampling to construct local basis functions has been proposed before in the literature. In [9], the authors proposed to sample the a-harmonic function space by choosing random boundary conditions, and use SVD thresholding to obtain well-conditioned basis functions. The method was developed further in [5], where the randomized sample basis functions are selected by solving a local eigenvalue problem associated with the elliptic operator. This post-processing step is similar to the method developed in [17] in the context of solving elliptic eigenvalue problems with a rough coefficient.

There are several important differences between our approach and that of [9, 5]. First, we find that the best randomized sampling strategy is *not* based on random sampling of boundary conditions, but rather on projection of a Gaussian function randomly centered in the domain. Second, our postprocessing step is based on the Kolmogorov n-width analysis in [1], so a low rank approximation is guaranteed. Third, we provide a quantitative criterion for evaluating the efficiency and accuracy of the method.

Other basis construction approaches based on gFEM framework have been explored in the literature, mostly based on a similar offline and online strategy. In the offline step, one prepares the solution space (either local or global). In the online step, one assembles the basis through the Galerkin framework (see, for example, [19, 23, 15]). The random sampling strategy can be also explored in these contexts.

The organization of the rest of the paper is as follows. We review preliminaries in Section 2, including the basics of basis construction and error analysis . In Section 3, we describe the random sampling framework, and present a few particular sampling strategies. We will also connect and compare the framework with the randomized singular value decomposition (rSVD) in Section 3.3. To compare the various sampling approaches, we propose a criterion in Section 4, according to which random sampling in the interior of the subdomains is more effective than random sampling of boundary conditions, as we demonstrate through numerical examples in Section 5.

2. Preliminary Results

Here we provide some preliminary results about the generalized finite element method and review the construction of optimal basis set. Consider the following elliptic equation:

$$\begin{cases} \mathcal{L}u = -\mathrm{div}\left(a(x)\nabla u(x)\right) = f(x)\,, & x \in \Omega; \\ u(x) = 0\,, & x \in \partial\Omega, \end{cases} \tag{2.1}$$

with $0 < \alpha_* \leq a(x) \leq \beta_*$, where $\mathcal{L}$ denotes the elliptic operator. The corresponding weak formulation is given by

$$\langle a(x)\nabla u\,, \nabla v\rangle_{L^2(\Omega)} = \langle f\,, v\rangle_{L^2(\Omega)}$$

for all test functions v, where $\langle f, g\rangle_{L^2(\Omega)} := \int_\Omega f(x)g(x)\mathrm{d}x$.

In the Galerkin framework, one constructs the solution space first. Given the approximation space

$$\mathrm{Span}\{\phi_i\,, i = 1, 2, \ldots n\}\,, \tag{2.2}$$

we substitute the ansatz $U = \sum_{i=1}^n c_i\phi_i$ and test functions ϕ_i, $i = 1, 2, \ldots, n$ into (2.1) to obtain:

$$\sum_{j=1}^{n} \langle a(x)\nabla\phi_j\,, \nabla\phi_i\rangle_{L^2(\Omega)} c_j = \langle f\,, \phi_i\rangle_{L^2(\Omega)}\,.$$

We write this system in matrix form as follows:

$$\mathsf{A}\vec{c} = \vec{b}\,, \tag{2.3}$$

where A is a symmetric matrix with entries $\mathsf{A}_{mn} = \langle a\nabla\phi_m\,, \nabla\phi_n\rangle_{L^2(\Omega)}$, and $\vec{c}$ (with $\vec{c}_m = c_m$) is a list of coefficients to be determined. The right hand side is the load vector $\vec{b}$, with entries $\vec{b}_m = \langle f\,, \phi_m\rangle_{L^2(\Omega)}$.

It is well known that $u - U$ is a-perpendicular to the ansatz space (2.2), and the following quasi-optimality property holds:

$$\|u - U\|_{\mathcal{E}(\Omega)} \leq C\|u - \mathcal{P}u\|_{\mathcal{E}(\Omega)}\,,$$

where C is some constant depending on α_* and β_*, and $\mathcal{P}u$ is the projection of the true solution u onto the space (2.2). Here the energy norm on any subdomain $\omega \subset \Omega$ is defined by

$$\|v\|_{\mathcal{E}(\omega)} = \langle a\nabla v\,, \nabla v\rangle^{1/2}_{L^2(\omega)} := \left[\int_\omega a|\nabla v(x)|^2 \mathrm{d}x\right]^{1/2}. \tag{2.4}$$

Thus the existence of a small numerical error $\|u-U\|_{\mathcal{E}(\Omega)}$ requires a set of basis functions that form a space for which $\|u - \mathcal{P}u\|_{\mathcal{E}(\Omega)}$ is small.

The main difficulty of computing the elliptic equation with rough coefficient is the number of basis functions needed. When $a(x)$ is rough with ε as its smallest scale, for standard piecewise affine finite elements, the mesh size Δx needs to resolve the smallest scale, so that $\Delta x \ll \varepsilon$ in each dimension. It follows that the dimension n of the system (2.3) is $n = \mathcal{O}(1/\varepsilon^d) \gg 1$, where d is the spatial dimension. The large size of stiffness matrix A and its large condition number (usually on the order of $\mathcal{O}(1/\varepsilon^2)$) make the problem expensive to solve using this approach.

The question is then whether it is possible to design a Galerkin space for which n is independent of ε? As mentioned in Section 1, the offline-online procedure makes this possible, as we discuss next.

2.1. **Generalized Finite Element Method.** The generalized Finite Element Method was one of the earliest methods to utilize the offline-online procedure. This approach is based on the partition of unity. One first decomposes the domain Ω into many small patches ω_i, $i = 1, 2, \ldots, m$, that form an open cover of Ω. Each patch ω_i is assigned a partition-of-unity function ν_i that is zero outside ω_i and 1 over most of the set ω_i. Specifically, there are positive constants C_1 and C_2 such that

$$0 \leq \nu_i(x) \leq 1, \qquad \text{for all } x \in \Omega \text{ and all } i = 1, 2, \ldots, m, \tag{2.5a}$$

$$\nu_i(x) = 0, \qquad \text{for all } x \in \Omega \setminus \omega_i \text{ and all } i = 1, 2, \ldots, m, \tag{2.5b}$$

$$\max_{x\in\Omega} \nu_i(x) \leq C_1, \qquad \text{for all } i = 1, 2, \ldots, m, \tag{2.5c}$$

$$\max_{x\in\Omega} |\nabla\nu_i(x)| \leq \frac{C_2}{\operatorname{diam}(\omega_i)}, \quad \text{for all } i = 1, 2, \ldots, m. \tag{2.5d}$$

Moreover, we have

$$\sum_{i=1}^{m} \nu_i(x) = 1, \quad \text{for all } x \in \Omega\,. \tag{2.6}$$

In the offline step, basis functions are constructed for each patch ω_i, $i = 1, 2, \ldots, m$. We denote the numerical local solution space in patch ω_i by:

$$\Phi_{[i]} = \text{Span}\{\phi_{i,j}\,, j = 1, 2, \ldots, n_i\}\,, \tag{2.7}$$

where n_i is the number of basis functions for ω_i. In the online step, the Galerkin formulation is used, with the space in (2.2) replaced by:

$$\Phi := \bigoplus_{i=1,2,\ldots,m} \Phi_{[i]}\nu_i = \text{Span}\{\phi_{i,j}\nu_i\,, i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n_i\}\,. \tag{2.8}$$

Details can be found in [2].

The total number of basis functions is $\sum_{i=1}^{m} n_i$. If all n_i, $i = 1, 2, \ldots, m$ are bounded by a modest constant, the dimension of the space Φ is of order m, so the computation in the online step is potentially inexpensive. It is proved in [2] that the total approximation error is governed by the summation of all local approximation errors.

Theorem 2.1. Denote by u the solution to (2.1). Suppose $\{\omega_i\}_{i=1,2,\ldots,m}$ forms an open cover of Ω and let $\{\nu_i\}_{i=1,2,\ldots,m}$ denote a set of partition-of-unity functions as defined in (2.5). Assume that in each patch ω_i the solution can be approximated well by $\zeta_i \in \Phi_{[i]}$. Then the global error is small too. More specifically, if we assume that

$$\|u - \zeta_i\|_{L^2(\omega_i)} \leq \varepsilon_1(i) \quad \text{and} \quad \|u - \zeta_i\|_{\mathcal{E}(\omega_i)} \leq \varepsilon_2(i)\,, \tag{2.9}$$

and define

$$\zeta(x) = \sum_{i=1}^{m} \zeta_i(x)\nu_i(x)\,,$$

then $\zeta(x) \in H^1(\Omega)$, and for the constants C_1 and C_2 defined in (2.5), we have

$$\|u - \zeta\|_{L^2(\Omega)} \leq C_1 \left(\sum_{i=1}^{m} \varepsilon_1^2(i)\right)^{1/2}$$

and

$$\|u - \zeta\|_{\mathcal{E}(\Omega)} \leq C_2 \left(\sum_{i=1}^{m} \frac{\varepsilon_1^2(i)}{\text{diam}^2(\omega_i)} + C_1^2 \sum_{i=1}^{m} \varepsilon_2^2(i)\right)^{1/2}.$$

This theorem shows that the approximation error of the Galerkin numerical solution for the gFEM depends directly on the accuracy of the local approximation spaces in each patch.

2.2. **Low-Rank Local Solution Space.** One reason for the success of gFEM is that the local numerical solution space is approximately low-rank, meaning that n_i can be made small for all i in (2.7); see [1]. We review the relevant results in this section, and show how to find $\Phi_{[i]}$.

Denote by ω_i^* an enlargement of the patch ω_i. To simplify notation, we suppress subscripts i from here on. We introduce a restriction operator:

$$P : H_a(\omega^*)/\mathbb{R} \to H_a(\omega)/\mathbb{R}\,,$$

where $H_a(\omega^*)$ is the collection of all a-harmonic functions in ω^* and $H_a(\omega^*)/\mathbb{R}$ represents the quotient space of $H_a(\omega^*)$ with respect to constant function. (This modification is needed to make $\|\cdot\|_{\mathcal{E}(\omega^*)}$ a norm, since an a-harmonic function is defined only up to an additive constant.) The operator P is determined uniquely by $a(x)$ restricted in ω^* and ω. We denote its adjoint operator by $P^* : H_a(\omega)/\mathbb{R} \to H_a(\omega^*)/\mathbb{R}$. It is shown in [1] that the operator P^*P is a compact, self-adjoint, nonnegative operator on $H_a(\omega^*)/\mathbb{R}$.

To derive an n-dimensional approximation of $H_a(\omega)/\mathbb{R}$, we first define the distance of an arbitrary n-dimensional function subspace $S_n \subset H_a(\omega)/\mathbb{R}$ to $H_a(\omega^*)/\mathbb{R}$, as follows:

$$d(S_n, \omega^*) = \sup_{u \in H_a(\omega^*)/\mathbb{R}} \inf_{\xi \in S_n} \frac{\|Pu - \xi\|_{\mathcal{E}(\omega)}}{\|u\|_{\mathcal{E}(\omega^*)}}.$$

By considering all possible S_n, we can identify the optimal approximation space Φ_n (called the optimal n-dimensional set) that achieves the infimum:

$$\Phi_n := \arg\inf_{S_n} d(S_n, \omega^*) = \arg\inf_{S_n} \sup_{u \in H_a(\omega^*)/\mathbb{R}} \inf_{\xi \in S_n} \frac{\|Pu - \xi\|_{\mathcal{E}(\omega)}}{\|u\|_{\mathcal{E}(\omega^*)}}. \tag{2.10}$$

We now define a distance measure between ω and ω^* as follows:

$$d_n(\omega, \omega^*) = d(\Phi_n, \omega^*) = \sup_{u \in H_a(\omega^*)/\mathbb{R}} \inf_{\xi \in \Phi_n} \frac{\|Pu - \xi\|_{\mathcal{E}(\omega)}}{\|u\|_{\mathcal{E}(\omega^*)}}.$$

The term $d_n(\omega, \omega^*)$ is the celebrated Kolmogorov n-width of the compact operator P. It reflects how quickly a-harmonic functions supported on ω^* lose their energies when confined on ω. According to [25], the optimal space Φ_n and Kolmogorov n-width can be found explicitly, in terms of the eigendecomposition of P^*P on ω^*, which is

$$P^*P\psi_i = \lambda_i \psi_i, \tag{2.11}$$

with λ_i arranged in descending order. We then have

$$\Phi_n := P\Psi_n = \mathrm{Span}\{\phi_1, \phi_2, \dots, \phi_n\},$$

where

$$\Psi_n = \mathrm{Span}\{\psi_1, \cdots, \psi_n\} \quad \text{and} \quad \phi_i = P\psi_i, \tag{2.12}$$

and

$$d_n(\omega, \omega^*) = \sqrt{\lambda_{n+1}}. \tag{2.13}$$

Note that ψ_i are all supported in the enlarged domain ω^*, while ϕ_i are their confinements in ω. Almost-exponential decay of the Kolmogorov width with respect to n was proved in [1], as follows.

Theorem 2.2. The accuracy $d_n(\omega, \omega^*)$ has nearly exponential decay for n sufficiently large, that is, for any small ε, we have

$$d_n(\omega, \omega^*) \le e^{-n^{\left(\frac{1}{d+1} - \varepsilon\right)}}.$$

It follows that for any function u that is a-harmonic function in the patch $\omega^* \subset \mathbb{R}^2$, we can find a function in space $v \in \Phi_n$ for which

$$\|u - v\|_{\mathcal{E}(\omega)} \leq d_n(\omega, \omega^*)\|u\|_{\mathcal{E}(\omega^*)} \sim e^{-n^{1/3}}.$$

Remark 1. Note that d_n is the $(n+1)$-th singular value $\sqrt{\lambda_{n+1}}$ of P. Because of the fast decay of d_n with respect to n indicated by Theorem 2.2, P is considered to be an approximately-low-rank operator. It follows that almost all a-harmonic functions supported on ω^*, when confined in ω, look almost alike, and can be represented by a relatively small number of "important" modes.

Remark 2. We note that enlarging the domain for over-sampling is not a new idea. In [12], the boundary layer behavior confined in ω^*/ω was studied and utilized for computation.

2.3. **Computing the Local Solution Space.** We describe here the computation of an approximation to Φ_n via a discretized implementation of the definitions in the previous subsection. The idea is to discretize the enlarged patch ω^* with a fine mesh, collect all a-harmonic functions in this discrete setting, and then solve (2.11) in a finite basis. To collect all a-harmonic functions, we would need to solve the system with elliptic operator (1.1) locally, with all possible Dirichlet boundary conditions on $\partial\omega^*$. For ease of presentation, here and in sequel, we assume that we choose a piecewise-affine finite-element discretization of the patch for computing the local a-harmonic functions. Then the discretized a-harmonic functions are determined by their values on grid points $\{y_1, y_2, \ldots, y_{N_y}\}$ on the boundary of $\partial\omega^*$. We proceed in three stages.

Stage A. Approximate the space of a-harmonic functions $H_a(\omega^*)$ via the functions χ_i obtained by solving the following system, for $i = 1, 2, \ldots, N_y$:

$$\begin{cases} \mathcal{L}\chi_i = -\mathrm{div}\,(a(x)\nabla\chi_i) = 0\,, & x \in \omega^* \\ \chi|_{\partial\omega^*} = \delta_i\,, & y_i \in \partial\omega^*, \end{cases} \tag{2.14}$$

where δ_i is the hat function that peaks at y_i and equals zero at other grid points y_j, $j \neq i$. Recall that we have assumed a piecewise-affine finite-element discretization of ω^*.

Stage B. Compute (2.11) in the space spanned by $\{\chi_i\,, i = 1, 2, \ldots, N_y\}$. To do this, we first note that:

$$\langle P^*P\psi_i\,, \delta\rangle_{\mathcal{E}(\omega^*)} = \langle P\psi_i\,, P\delta\rangle_{\mathcal{E}(\omega)} = \langle \psi_i\,, \delta\rangle_{\mathcal{E}(\omega)}\,, \quad \forall\ \delta \in H_a(\omega^*)/\mathbb{R}\,, \tag{2.15}$$

Thus the weak formulation of (2.11) is given by

$$\langle \psi_i\,, \chi\rangle_{\mathcal{E}(\omega)} = \lambda_i\langle \psi_i\,, \chi\rangle_{\mathcal{E}(\omega^*)}\,, \quad \forall \chi \in \mathrm{span}\{\chi_1, \ldots, \chi_{N_y}\}\,.$$

Expanding the eigenfunction ψ_i in terms of χ_j, $j = 1, 2, \ldots, N_y$:

$$\psi_i = \sum_j c_j^{(i)}\chi_j, \tag{2.16}$$

we obtain the following equation for the coefficient vector $\vec{c}^{(i)}$:

$$\sum_j c_j^{(i)} \langle \chi_j \, , \chi_k \rangle_{\mathcal{E}(\omega)} = \lambda_i \sum_j c_j^{(i)} \langle \chi_j \, , \chi_k \rangle_{\mathcal{E}(\omega^*)} \, .$$

This system can be written as

$$\begin{aligned} &\mathsf{S}\vec{c}^{(i)} = \lambda_i \mathsf{S}^* \vec{c}^{(i)} \, , \\ \text{with} \qquad &\mathsf{S}_{mn} = \langle \chi_m \, , \chi_n \rangle_{\mathcal{E}(\omega)} = \langle \chi_m \, , \chi_n \rangle_{\mathcal{E}(\omega)} \quad \text{and} \quad \mathsf{S}^*_{mn} = \langle \chi_m \, , \chi_n \rangle_{\mathcal{E}(\omega^*)} \, . \end{aligned} \tag{2.17}$$

This generalized eigenvalue problem can be solved for λ_i, $i = 1, 2, \ldots, N_y$, arranged in descending order, and their associated eigenfunctions ψ_i, $i = 1, 2, \ldots, N_y$ defined from (2.16) using the generalized eigenvectors $\vec{c}^{(i)}$ from (2.17). Index n is chosen such that $\lambda_{n+1} < \text{TOL} < \lambda_n$, where TOL is a given error tolerance.

Stage C. Obtain Φ_n by substituting the functions ψ_i, $i = 1, 2, \ldots, N_y$ from Stage B into (2.12).

3. Randomized Sampling Methods for Local Bases

In this section we propose a class of random sampling methods to construct local basis functions efficiently. As seen in Section 2.3, finding the optimal basis functions amounts to solving the generalized eigenvalue problem in (2.17). The main cost does not come from performing the eigenvalue decomposition, but rather from computing the a-harmonic functions χ_i, which are used to construct the matrices S and S^* in (2.17). As shown in Section 2.2, the eigenvalues decay almost exponentially fast, indicating that, locally, only a limited number of modes is needed to well represent the whole solution space. This low-rank structure motivates us to consider randomized sampling techniques.

Randomized algorithms have been highly successful in compressed sensing, where they are used to extract low rank structure efficiently from data. The Johnson-Lindenstrauss lemma [16] suggests that structure in high dimensional data points is largely preserved when projected onto random lower-dimensional spaces. The randomized SVD (rSVD) algorithm uses this idea to captures the principal components of a large matrix by random projection of its row and column spaces into smaller subspaces; see [11] for a review. In the current numerical PDE context, knowing that the local solution space is essentially low-rank, we seek to adopt the random sampling idea to generate local approximate solution spaces efficiently.

For our problem, unfortunately, randomized SVD cannot be applied directly, as we discuss further in Section 3.3. Instead, we propose a method based on Galerkin approximation of the generalized eigenvalue problem on a small subspace. One immediate problem is that an arbitrarily given random function is not necessarily a-harmonic. Thus, our method first generates a random collection of functions and projects them

onto the a-harmonic function space, and then solves the generalized eigenvalue problem (2.17) on the subspace to find the optimal basis functions. A detailed description of our procedure is shown in Algorithm 1.

Algorithm 1 Determining Optimal Local Bases

Stage 1: Randomly generate a collection of N_r a-harmonic functions.

Stage 1-A: Randomly pick functions $\{\xi_k : k = 1, 2, \ldots, N_r\}$ supported on ω^*;

Stage 1-B: For each $k = 1, 2, \ldots, N_r$, project ξ_k onto the a-harmonic function space to obtain γ_k;

Stage 2: Solve the generalized eigenvalue problem to determine leading modes.

Stage 2-A: Define:

$$\mathsf{S}_{\gamma,mn} := \langle \gamma_m \, , \gamma_n \rangle_{\mathcal{E}(\omega)} \, , \quad \mathsf{S}^*_{\gamma,mn} := \langle \gamma_m \, , \gamma_n \rangle_{\mathcal{E}(\omega^*)} \, , \tag{3.1}$$

and compute the associated generalized eigenvalue problem:

$$\mathsf{S}_\gamma \vec{v}^\gamma = \lambda^\gamma \mathsf{S}^*_\gamma \vec{v}^\gamma \, , \tag{3.2}$$

with $(\lambda^\gamma_j \, , \vec{v}^\gamma_j)$ denoting the j-th eigen-pairs, such that $\lambda^\gamma_1 \geq \lambda^\gamma_2 \geq \cdots$

Stage 2-B: Collect the first few eigenfucntions according to the preset error tolerance and use them as the local basis functions:

$$\Phi^r_n = \mathrm{Span}\{\phi^r_1, \phi^r_2, \ldots, \phi^r_n\} = \mathrm{Span}\{P\psi^r_j \, , j = 1, 2, \ldots, n\} \, , \tag{3.3}$$

where $\psi^r_j = \sum_k \vec{v}^\gamma_{j,k} \gamma_k$ and n is chosen such that $\lambda^\gamma_1 \geq \cdots \geq \lambda^\gamma_n \geq \mathrm{TOL} > \lambda^\gamma_{n+1}$.

Note that the steps in Stage 2 of Algorithm 1 are parallel to those of Section 2.3, but only a small number of functions γ_k is used in the generalized eigenvalue problem, rather that the whole list of a-harmonic functions (i.e., $N_r \ll N_y$). We therefore save significant computation in preparing the a-harmonic function space, in assembling the S and S^* matrices, and in solving the generalized eigenvalue decomposition.

The key is to use the random sampling strategy in Stage 1 of Algorithm 1 to generate an effective small subspace for the generalized eigenvalue problem. This aspect of the algorithm will be the focus of the rest of this section.

3.1. **a-Harmonic Projection.** Let us first discuss the a-harmonic projection of a given function ξ supported on ω^*. This problem can be formulated as a PDE-constrained optimization problem:

$$\min_\gamma \frac{1}{2} \|\gamma - \xi\|^2_{L^2(\omega^*)} \quad \text{s.t.} \quad \mathcal{L}\gamma = 0 \, , \tag{3.4}$$

where $\mathcal{L} = -\operatorname{div} a\nabla$ is the elliptic operator defined in (2.1). The Lagrangian function for (3.4) is as follows:

$$F(\gamma, \lambda) := \frac{1}{2} \|\gamma - \xi\|^2_{L^2(\omega^*)} - \langle \lambda \, , \mathcal{L}\gamma \rangle_{L^2(\omega^*)} \, , \tag{3.5}$$

where λ is a Lagrange multiplier. In the discrete setting, we form a grid $\{x_i\}$ over ω^* and denote by ζ_i the hat function centered at grid point x_i. (Recall that we have assumed piecewise affine finite element discretization.) The Lagrangian function for the corresponding discretized optimization problem is

$$F(\gamma,\lambda)=\frac{1}{2}(\gamma^i-\xi^i)^\top(\gamma^i-\xi^i)-\lambda^\top \mathsf{A}^{ii}\gamma^i-\lambda^\top \mathsf{A}^{ib}\gamma^b\,, \tag{3.6}$$

where the superindices i and b stand for interior and boundary grids, respectively, and A is the stiffness matrix:

$$\mathsf{A}_{mn}=\langle a\nabla\zeta_m\,,\nabla\zeta_n\rangle_{L^2(\omega^*)}\,.$$

In the discrete setting, λ is a vector of the same length as γ^i (the number of grid points in the interior). Note that in the translation to the discrete setting, we represent $\mathcal{L}\gamma=0$ by $\mathsf{A}\gamma=0$, which leads to

$$\mathsf{A}^{ii}\gamma^i+\mathsf{A}^{ib}\gamma^b=0\,.$$

Here A^{ii} is the stiffness matrix confined in the interior, and A^{ib} is the part of the stiffness matrix generated by taking the inner product of the interior basis functions and the boundary basis functions. To solve the minimization problem, we take the partial derivatives of (3.6) with respect to γ and λ and set them equal to zero, as follows:

$$\begin{aligned}\nabla_{\gamma^i}F&=\gamma^i-\xi^i-\mathsf{A}^{ii}\lambda=0\,,\\ \nabla_{\gamma^b}F&=\mathsf{A}^{ib\top}\lambda\qquad\quad\;\;=0\,,\\ \nabla_{\lambda}F&=\mathsf{A}^{ii}\gamma^i+\mathsf{A}^{ib}\gamma^b\;=0\,.\end{aligned}$$

Some manipulation yields the following systems for γ^b and γ^i:

$$\mathsf{A}^{ib\top}\left(\mathsf{A}^{ii}\right)^{-2}\mathsf{A}^{ib}\gamma^b=-\mathsf{A}^{ib\top}\left(\mathsf{A}^{ii}\right)^{-1}\xi^i,\quad \gamma^i=-\left(\mathsf{A}^{ii}\right)^{-1}\mathsf{A}^{ib}\gamma^b\,,$$

and the solution to this system gives the solution of (3.4) in the discrete setting.

3.2. **Random sampling strategies.** We have many possible choices for the random functions functions ξ_k, $k=1,2,\ldots,N_r$ in Stage 1-A of 1. Here we list a few natural ones.

1. *Interior δ-function.* Choose a random grid point in ω and set $\xi(x)=1$ at this grid point, and zero at all other grid points. That is, ξ is the hat function associated with the grid point x.
2. *Interior i.i.d. function.* Choose the value of ξ at each grid point in ω independently from a standard normal Gaussian distribution. The values of ξ at grid points in $\omega^*\backslash\omega$ are set to 0.
3. *Full-domain i.i.d. function.* The same as in 2, except that the values of ξ at the grid points in $\omega^*\backslash\omega$ are also chosen as Gaussian random variables.
4. *Random Gaussian.* Choose a random grid point $x_0\in\omega$ and set $\xi(x)=e^{-\frac{(x-x_0)^2}{2}}$ at all grid points in ω^*.

We aim to select basis functions (through Stage 2) that are associated with the largest eigenvalues, so that the Kolmogorov n-width can be small (2.13). Thus, we hope that in Stage 1, the chosen functions ξ_k provide large eigenvalues λ_i in (2.17). A large value of λ indicates that a large portion of the energy is maintained in ω, with only a small amount coming from the buffer region $\omega^*\backslash\omega$. It therefore suggests to choose functions ξ_k with most of their variations inside ω. However, the projection to a-harmonic space step makes the locality of the resulting functions hard to predict. In Section 4, we propose and analyze a criterion for the performance of the random sampling schemes. In particular, we compare the choices listed above.

We mention here two other approaches that have been proposed in the literature for producing bases of a-harmonic functions. In these previous works, the functions obtained are used directly as basis functions, without applying the post-processing step of the generalized eigenvalue problem (2.17) in Phase 2 above.

1. *Random boundary sampling.* In [9], the authors proposed to obtain a list of random a-harmonic functions by computing the local elliptic equation with i.i.d. random Dirichlet boundary conditions. Assuming there are N_y grid points on the boundary $\partial\omega^*$, we define g to be a vector of length N_y with i.i.d. random variables for each component. We then define γ by solving

$$\begin{cases} \mathcal{L}\gamma = 0\,, & x \in \omega^* \\ \gamma|_{\partial\omega^*} = g. \end{cases} \tag{3.7}$$

   This process is repeated for N_r times to obtain N_r random a-harmonic functions $\{\gamma_k : k = 1, 2, \ldots, N_r\}$.
2. *Randomized boundary sampling with exponential covariance.* A technique in which the Dirichlet boundary conditions are chosen to be random Gaussian variables with a specified covariance matrix is described in [18]. This matrix is assumed to be an exponential function, that is,

$$\mathrm{Cov}(y_i, y_j) = \exp(-|y_i - y_j|/\sigma)\,. \tag{3.8}$$

   The first few modes of a Karhunen-Loéve expansion are used to construct a boundary condition in (3.7), with which basis functions are computed. Although a justification for this approach is not provided, numerical computations show that it is more efficient than the i.i.d. random boundary sampling.

3.3. **Connection with Randomized SVD.** We briefly address the connection of the random sampling method we propose in this paper with the well-known randomized SVD (rSVD) algorithm. Although rSVD cannot be used directly in our problem, it served as a motivation for our randomized sampling strategies.

The randomized SVD algorithm, studied thoroughly in [11], speeds up the computation of the SVD of a matrix when the matrix is large and approximately low rank. With high probability, the singular vector structure is largely preserved when the matrix is

projected onto a random subspace. Specifically, for a random matrix R with a small number of columns (the number depending on the rank of A), it is proved in [11] if we obtain Q from the QR factorization of AR, we have

$$\|\mathsf{A} - \mathsf{Q}\mathsf{Q}^\top \mathsf{A}\|_2 \ll \|\mathsf{A}\|_2 \,. \tag{3.9}$$

This bound implies that any vector in the range space of A can be well approximated by its projection into the space spanned by Q. For example, if $\vec{u} = \mathsf{A}\vec{v}$, we have from (3.9) that

$$\|\vec{u} - \mathsf{Q}\mathsf{Q}^\top \vec{u}\| \ll \|\vec{u}\| \,. \tag{3.10}$$

We note that Q and AR span the same column space, but Q is easier to work with and better conditioned, because its columns are orthonormal. Equivalent to (3.10), we can also say that any $\vec{u}$ in the image of A can be approximated well using a linear combination of the columns of AR.

To see the connection between rSVD and our problem, we first write the generalized eigenvalue problem (2.17) into a SVD form. Recall the definitions (3.1) of S and S^*:

$$\mathsf{S}_{mn} = \int_\omega a(x)\nabla\chi_m(x)\cdot\nabla\chi_n(x)\mathrm{d}x \,, \quad \mathsf{S}^*_{mn} = \int_{\omega^*} a(x)\nabla\chi_m(x)\cdot\nabla\chi_n(x)\mathrm{d}x \,,$$

and define

$$\Phi^* = \left[\sqrt{a}\nabla\chi_1, \sqrt{a}\nabla\chi_2, \ldots, \sqrt{a}\nabla\chi_{N_y}\right] \,, \quad \Phi = \Phi^*|_\omega \,. \tag{3.11}$$

Then one could denote $\mathsf{S} = \Phi^\top\Phi$ and $\mathsf{S}^* = \Phi^{*\top}\Phi^*$, and rewrite the generalized eigenvalue problem (2.17) as follows:

$$\Phi^\top\Phi\vec{c} = \lambda\Phi^{*\top}\Phi^*\vec{c}. \tag{3.12}$$

We write the QR factorization for Φ^* as follows:

$$\Phi^* = \mathsf{Q}_{\Phi^*}\mathsf{R}_{\Phi^*} \,,$$

and denote $\vec{d} = \mathsf{R}_{\Phi^*}\vec{c}$. By substituting into (3.12), we obtain

$$\left(\Phi\mathsf{R}^{-1}_{\Phi^*}\right)^\top \left(\Phi\mathsf{R}^{-1}_{\Phi^*}\right)\vec{d} = \lambda\vec{d},$$

meaning that $(\sqrt{\lambda}, \vec{d})$ forms a singular value pair of the matrix $\Phi R^{-1}_{\Phi^*}$.

According to the rSVD argument, the leading singular vectors of $\Phi\mathsf{R}^{-1}_{\Phi^*}$ are captured by those of

$$\Phi\mathsf{R}^{-1}_{\Phi^*}\mathsf{R} \,, \tag{3.13}$$

where R is a matrix whose entries are i.i.d Gaussian random variables. Specifically, with high probability, the leading singular values of $\Phi\mathsf{R}^{-1}_{\Phi^*}\mathsf{R}$ are almost the same as those of $\Phi\mathsf{R}^{-1}_{\Phi^*}$, and the column space spanned by (3.13) largely covers the image of $\Phi\mathsf{R}^{-1}_{\Phi^*}$, as in (3.11).

We now interpret $\Phi \mathsf{R}_{\Phi^*}^{-1}\mathsf{R}$ from the viewpoint of PDEs. Decomposing $\mathsf{R}_{\Phi^*}^{-1}\mathsf{R}$ into columns as follows:

$$\mathsf{R}_{\Phi^*}^{-1}\mathsf{R} = [r_1, r_2, \dots], \quad \text{with} \quad r_k = [r_{k1}, r_{k2}, \cdots]^\top, \tag{3.14}$$

and denoting $\gamma_k = \sum_j r_{kj}\chi_j$, we have from (3.11) that

$$\Phi r_k = \sqrt{a}\nabla\left(\sum_j r_{kj}\chi_j\right) \doteq \sqrt{a}\nabla\gamma_k\,.$$

Numerically, this corresponds to solving the following system for γ_k:

$$\begin{cases} \mathcal{L}\gamma_k = -\mathrm{div}\,(a(x)\nabla\gamma_k) = 0\,, & x \in \omega^* \\ \gamma_k|_{\partial\omega^*} = \sum_j r_{kj}\delta_{y_j}. \end{cases} \tag{3.15}$$

It is apparent from this equation that to obtain $\Phi\mathsf{R}_{\Phi^*}^{-1}\mathsf{R}$, we do not need to compute all functions χ_j, $j = 1, 2, \dots, N_y$ and use them to construct γ_k. Rather, we can compute γ_k directly by solving the elliptic equation with random boundary conditions given by r_{kj}, $j = 1, 2, \dots, N_y$. The cost of this procedure is now proportional to N_r, which is much less than N_y.

Unfortunately, this procedure is difficult to implement in a manner that accords with the rSVD theory. R is constructed using i.i.d. Gaussian random variables, but $\mathsf{R}_{\Phi^*}^{-1}$ is unknown ahead of time, so the distribution of r_k (3.14) is unknown. The theory here suggests that there is a random sampling strategy that achieves the accuracy and efficiency that characterize rSVD, but it does not provide such a strategy.

4. Efficiency of Various Random Sampling Methods

As discussed in Section 3, given multiple ways to choose the random samples in Stage 1 of Algorithm 1, it is natural to ask which one is better, and how do to predetermine the approximation accuracy? We answer these questions in this section.

The key requirement is that Algorithm 1 should capture the high-energy modes of (2.11) — the modes that correspond to the highest values of λ_i. We start with a simple example in Section 4.1 that finds the relationship between the energy captured by a certain single mode, and the angle that that mode makes with the highest energy mode (which corresponds to the Kolmogorov 1-width). We extend this discussion beyond a single mode in Section 4.2, again using the language of linear algebra. The relevance to local PDE basis construction is outlined in Section 4.3.

4.1. **A One-Mode Example.**

Suppose we are working in a three-dimensional space, with symmetric positive definite matrices A and B and generalized eigenvectors x_1, x_2, and x_3 such that

$$\langle x_i\,, x_j\rangle_{\mathsf{B}} = x_i^\top \mathsf{B} x_i = \delta_{ij}\,, \quad \langle x_i\,, x_j\rangle_{\mathsf{A}} = x_i^\top \mathsf{A} x_i = \delta_{ij}\lambda_i\,, \tag{4.1}$$

for generalized eigenvalues $\lambda_1 > \lambda_2 > \lambda_3$. We thus have

$$\mathsf{A} \cdot x_i = \lambda_i \mathsf{B} \cdot x_i, \quad i = 1, 2, 3.$$

Suppose we have some vector x that is intended to approximate the leading eigenvector x_1. The energy of x is

$$E(x) = \frac{x^\top \mathsf{A} x}{x^\top \mathsf{B} x}\,, \tag{4.2}$$

and the angle between the space spanned by x and that spanned by x_1 is defined by:

$$\angle(x, x_1) = \max_{\beta:|\beta|\le 1} \min_{\alpha} \frac{\|\alpha x - \beta x_1\|_{\mathsf{A}}}{\|x\|_{\mathsf{B}} \|x_1\|_{\mathsf{B}}} = \max_{\beta:|\beta|\le 1} \min_{\alpha} \frac{\|\alpha x - \beta x_1\|_{\mathsf{A}}}{\|x\|_{\mathsf{B}}}\,. \tag{4.3}$$

We have the following result (which generalizes easily to dimension greater than 3).

Proposition 1. The angle (4.3) is bounded in terms of the energy (4.2) as follows:

$$\angle(x, x_1) \le \sqrt{\frac{\lambda_1 \lambda_2 \left(\lambda_1 - E(x)\right)}{(\lambda_1 - \lambda_2) E(x)}}\,. \tag{4.4}$$

Proof. The proof is simple algebra. As $\{x_1, x_2, x_3\}$ span the entire space and are B-orthogonal, we have

$$x = w_1 x_1 + w_2 x_2 + w_3 x_3\,, \tag{4.5}$$

with $w_i = x^\top \mathsf{B} x_i$, $i = 1, 2, 3$. According to the definition of the angle, one can reduce the problem by setting $\beta = 1$ and $\sum_i w_i^2 = 1$, so that $\|x\|_{\mathsf{B}} = 1$ in (4.3). (With these normalizations, we have from (4.1) and (4.2) that $E(x) = x^T A x = \lambda_1 w_1^2 + \lambda_2 w_2^2 + \lambda_3 w_3^2$.) We thus have

$$\begin{aligned}
\left(\angle(x, x_1)\right)^2 &= \min_{\alpha} \|\alpha x - x_1\|_{\mathsf{A}}^2 \\
&= \min_{\alpha} \|(\alpha w_1 - 1) x_1 + \alpha w_2 x_2 + \alpha w_3 x_3\|_{\mathsf{A}}^2 \\
&= \min_{\alpha} \left((\alpha w_1 - 1)^2 \lambda_1 + \alpha^2 w_2^2 \lambda_2 + \alpha^2 w_3^2 \lambda_3\right).
\end{aligned}$$

The minimum is achieved at $\alpha = w_1 \lambda_1 / E(x)$, with the minimized angle being

$$\left(\angle(x, x_1)\right)^2 = \frac{E(x) - w_1^2 \lambda_1}{E(x)} \lambda_1\,. \tag{4.6}$$

To bound the numerator in (4.6) we observe that

$$E(x) - w_1^2 \lambda_1 = w_2^2 \lambda_2 + w_3^2 \lambda_3 \le w_2^2 \lambda_2 + w_3^2 \lambda_2 = (1 - w_1^2) \lambda_2,$$

and moreover

$$E(x) \le w_1^2 \lambda_1 + (1 - w_1^2) \lambda_2 \;\Rightarrow\; \lambda_1 - E(x) \ge (1 - w_1^2)(\lambda_1 - \lambda_2) \;\Rightarrow\; 1 - w_1^2 \le \frac{\lambda_1 - E(x)}{\lambda_1 - \lambda_2}\,.$$

By conbining these last two bounds, we obtain

$$E(x) - w_1^2 \lambda_1 \le \lambda_2 \frac{\lambda_1 - E(x)}{\lambda_1 - \lambda_2}.$$

By substituting this bound into (4.6), we obtain (4.4). □

Note that the bound (4.4) decreases to zero as $\lambda_1 - E(x) \downarrow 0$.

According to (4.4), a larger gap in the spectrum between λ_1 and λ_2 yields a tighter bound, thus better control over the angle. The theorem indicates that the "energy" is the quantity that measures how well the randomly given vector x captures the first mode, and thus serves as the criterion for the quality of the approximation.

4.2. **Higher-Dimensional Criteria.** In this section, we seek the counterpart in higher dimensional space of the previous result. Suppose now that the two symmetric positive definite matrices A and B are $n \times n$, and their generalized eigenpairs (λ_i, x_i) satisfy the following:

$$\langle x_i \, , x_j \rangle_{\mathsf{B}} = \delta_{ij} \, , \quad \langle x_i \, , x_j \rangle_{\mathsf{A}} = \delta_{ij} \lambda_i \, , \tag{4.7}$$

so that

$$\mathsf{A} \cdot x_i = \lambda_i \mathsf{B} \cdot x_i \, , \quad \text{with} \quad \lambda_1 \geq \cdots \geq \lambda_k > \lambda_{k+1} \geq \cdots \geq \lambda_n > 0 \, ,$$

or

$$\mathsf{A} \cdot \mathsf{X} = \mathsf{B} \cdot \mathsf{X} \cdot \Lambda \, , \quad \text{with} \quad \Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \, . \tag{4.8}$$

Suppose we are trying to recover $\mathsf{X}^h = [x_1, x_2, \ldots, x_k]$, the matrix whose columns are the first k modes. Define $\mathsf{X}^l = [x_{k+1}, \ldots, x_n]$ to collect the remaining modes, and let $\mathsf{Y} = [y_1, y_2, \ldots, y_k]$ be our proposed approximation to X^h. We seek a quantity that measures how well Y approximates X^h in the energy norm.

To proceed, we assume first that Y is B-orthonormal, that is,

$$\mathsf{Y}^\top \mathsf{B} \mathsf{Y} = \mathbb{1} \, . \tag{4.9}$$

(If necessary, a QR procedure based on the B-norm can be used to ensure this property.) Extending (4.2), we define the energy for an $n \times k$ matrix Z as follows:

$$E(\mathsf{Z}) := \frac{\operatorname{Tr}(\mathsf{Z}^\top \mathsf{A} \mathsf{Z})}{\operatorname{Tr}(\mathsf{Z}^\top \mathsf{B} \mathsf{Z})} \tag{4.10}$$

The Kolmogorov width between space Y and the optimal subspace X^h is a generalization of the angle (4.3), defined as follows:

$$d(\mathsf{X}^h, \mathsf{Y}) = \sup_{z \in \text{Span } \mathsf{X}^h, \, \|z\|_{\mathsf{B}} \leq 1} \inf_{y \in \mathsf{Y}} \|y - z\|_{\mathsf{A}} \, , \tag{4.11}$$

where "Span X^h" denotes the space spanned by the columns of X^h. We show below how $E(\mathsf{Y})$ is related to $d(\mathsf{X}^h, \mathsf{Y})$.

Since X spans the entire space, we can express Y as follows:

$$\mathsf{Y} = \mathsf{X} \cdot \mathsf{C} = \mathsf{X}^h \cdot \mathsf{C}^h + \mathsf{X}^l \cdot \mathsf{C}^l \tag{4.12}$$

where $\mathsf{C} \in \mathbb{R}^{n \times k}$. We have that C has orthonormal columns, because from (4.7) and (4.9), we have

$$\mathsf{Y} = \mathsf{X} \cdot \mathsf{C} \; \Rightarrow \; \mathbb{1} = \mathsf{Y}^\top \cdot \mathsf{B} \cdot \mathsf{Y} = \mathsf{C}^\top \cdot \mathsf{X}^\top \cdot \mathsf{B} \cdot \mathsf{X} \cdot \mathsf{C} = \mathsf{C}^\top \cdot \mathsf{C}. \tag{4.13}$$

We denote by C^h the upper $\mathbb{R}^{k\times k}$ portion of C, and by C^l the lower $\mathbb{R}^{(n-k)\times k}$ portion. We denote the elements of C by c_{ji}, that is

$$\mathsf{C}^h = [c_{ji}]_{j=1,2,\ldots,k;\, i=1,2,\ldots,k}, \quad \mathsf{C}^l = [c_{ji}]_{j=k+1,k+2,\ldots,n;\, i=1,2,\ldots,k}. \tag{4.14}$$

By orthonormality of C, we have that

$$\sum_{j=1}^{k} c_{ji}^2 + \sum_{j=k+1}^{n} c_{ji}^2 = 1, \quad i = 1, 2, \ldots, k.$$

and thus

$$\left[C^{l\top} C^l\right]_{ii} = 1 - \left[C^{h\top} C^h\right]_{ii} = 1 - \sum_{j=1}^{k} c_{ji}^2, \quad i = 1, 2, \ldots, k. \tag{4.15}$$

Lemma 4.1. The trace of $\mathsf{C}^{l\top}\mathsf{C}^l$ is bounded by energy difference between optimal space X^h and the proposed space Y

$$\mathrm{Tr}(\mathsf{C}^{l\top}\mathsf{C}^l) \leq \frac{k\left(E(\mathsf{X}^h) - E(\mathsf{Y})\right)}{\lambda_k - \lambda_{k+1}}\,. \tag{4.16}$$

Furthermore, C^h is invertible if

$$E(\mathsf{X}^h) - E(\mathsf{Y}) < \frac{\lambda_k - \lambda_{k+1}}{k}\,. \tag{4.17}$$

Proof. Since Y is B-orthonormal, we have from (4.12) that

$$\mathsf{C}^{h\top}\mathsf{C}^h + \mathsf{C}^{l\top}\mathsf{C}^l = \mathbb{1}\,. \tag{4.18}$$

Since both X^h and Y are B-orthonormal and have k columns, we have:

$$\mathrm{Tr}\left(\mathsf{X}^{h\top}\mathsf{B}\mathsf{X}^h\right) = \mathrm{Tr}\left(\mathsf{Y}^\top\mathsf{B}\mathsf{Y}\right) = k\,.$$

By plugging X^h and Y into the definition of energy (4.10), we have

$$k\left(E(\mathsf{X}^h) - E(\mathsf{Y})\right) = \mathrm{Tr}\left(\mathsf{X}^{h\top}\mathsf{A}\mathsf{X}^h - \mathsf{Y}^\top\mathsf{A}\mathsf{Y}\right)$$

By substituting for Y from (4.12), and using (4.7), we have

$$\begin{aligned} k\left(E(\mathsf{X}^h) - E(\mathsf{Y})\right) &= \mathrm{Tr}\left(\mathsf{X}^{h\top}\mathsf{A}\mathsf{X}^h - \mathsf{C}^{h\top}\mathsf{X}^{h\top}\mathsf{A}\mathsf{X}^h\mathsf{C}^h - \mathsf{C}^{l\top}\mathsf{X}^{l\top}\mathsf{A}\mathsf{X}^l\mathsf{C}^l\right) \\ &= \mathrm{Tr}\left(\Lambda^h - \mathsf{C}^{h\top}\Lambda^h\mathsf{C}^h - \mathsf{C}^{l\top}\Lambda^l\mathsf{C}^l\right)\,. \end{aligned} \tag{4.19}$$

For the terms on the right-hand side of (4.19), we have

$$\mathrm{Tr}\left(\mathsf{C}^{l\top}\Lambda^l\mathsf{C}^l\right) \leq \lambda_{k+1}\,\mathrm{Tr}\left(\mathsf{C}^{l\top}\mathsf{C}^l\right)\,, \tag{4.20}$$

and that

$$\begin{aligned}\operatorname{Tr}\left(\Lambda^h - \mathsf{C}^{h\top}\Lambda^h\mathsf{C}^h\right) &= \sum_{j=1}^{k}\lambda_j - \sum_{i=1}^{k}\sum_{j=1}^{k}\lambda_j c_{ji}^2 \\ &= \sum_{j=1}^{k}\lambda_j\left(1-\sum_{i=1}^{k}c_{ji}^2\right) \\ &\geq \lambda_k\sum_{j=1}^{k}\left(1-\sum_{i=1}^{k}c_{ji}^2\right) \\ &= \lambda_k \operatorname{Tr}\left(\mathsf{C}^{l\top}\mathsf{C}^l\right) \qquad (4.21)\end{aligned}$$

where we used (4.15). By bubstituting (4.20) and (4.21) into (4.19), we obtain

$$k\left(E(\mathsf{X}^h) - E(\mathsf{Y})\right) \geq (\lambda_k - \lambda_{k+1})\operatorname{Tr}\left(\mathsf{C}^{l\top}\mathsf{C}^l\right),$$

which is equivalent to (4.16).

When condition (4.17) holds, we have from (4.16) that $\operatorname{Tr}(\mathsf{C}^{l\top}\mathsf{C}^l) < 1$. Thus since $\mathsf{C}^{h\top}\mathsf{C}^h = \mathbb{1} - \mathsf{C}^{l\top}\mathsf{C}^l$ and setting $\|\mathsf{C}^{l\top}\mathsf{C}\| \leq \operatorname{Tr}(\mathsf{C}^{l\top}\mathsf{C}^l) < 1$, we have that $\mathsf{C}^{h\top}\mathsf{C}^h$ is nonsingular, so that the $k \times k$ matrix C^h is nonsingular. □

We finally use energy distance $E(\mathsf{X}^h) - E(\mathsf{Y})$ to estimate the Kolmogorov width $d(\mathsf{X}^h, \mathsf{Y})$, as follows.

Theorem 4.1. Consider the optimal space spaned by the columns of X^h and the proposed space spanned by the columns of Y. If

$$E(\mathsf{X}^h) - E(\mathsf{Y}) \leq \frac{\lambda_k - \lambda_{k+1}}{2k}, \qquad (4.22)$$

then we have

$$d(\mathsf{X}^h, \mathsf{Y}) \leq \sqrt{\lambda_{k+1}\frac{\|\mathsf{C}^{l\top}\mathsf{C}^l\|}{1-\|\mathsf{C}^{l\top}\mathsf{C}^l\|}}. \qquad (4.23)$$

and furthermore

$$d(\mathsf{X}^h, \mathsf{Y}) \leq \sqrt{2\lambda_{k+1}\frac{k\left(E(\mathsf{X}^h) - E(\mathsf{Y})\right)}{\lambda_k - \lambda_{k+1}}}. \qquad (4.24)$$

Proof. Choosing an arbitrary $z = \mathsf{X}^h\alpha$ with $\|\alpha\| \leq 1$, we look for β such that $y = \mathsf{Y}\beta$ is closest to z in A-norm. The solution, obtained from the minimization problem

$$\min_{\beta} f(\beta) := \|y - z\|_{\mathsf{A}}^2 = (\mathsf{Y}\beta - \mathsf{X}^h\alpha)^\top\mathsf{A}(\mathsf{Y}\beta - \mathsf{X}^h\alpha) \qquad (4.25)$$

is

$$\beta^* = (\mathsf{Y}^\top\mathsf{A}\mathsf{Y})^{-1}\mathsf{Y}^\top\mathsf{A}\mathsf{X}^h\alpha. \qquad (4.26)$$

From (4.7) and (4.12), we have

$$\mathsf{Y}^\top\mathsf{A}\mathsf{Y} = \mathsf{C}^\top\Lambda\mathsf{C} = \mathsf{C}^{h\top}\Lambda^h\mathsf{C}^h + \mathsf{C}^{l\top}\Lambda^l\mathsf{C}^l, \qquad (4.27)$$

which is invertible, since C has orthonormal columns and Λ is diagonal and positive definite. Thus β^* is well defined by (4.26). By substituting in (4.25), we obtain

$$f(\beta^*) = -\alpha^\top(\mathsf{Y}^\top\mathsf{A}\mathsf{X}^h)^\top(\mathsf{Y}^\top\mathsf{A}\mathsf{Y})^{-1}(\mathsf{Y}^\top\mathsf{A}\mathsf{X}^h)\alpha + \alpha^\top\mathsf{X}^{h\top}\mathsf{A}\mathsf{X}^h\alpha. \tag{4.28}$$

Note from (4.12) and (4.8) that

$$\mathsf{A}\mathsf{Y} = \mathsf{A}\mathsf{X}^h\mathsf{C}^h + \mathsf{A}\mathsf{X}^l\mathsf{C}^l = \mathsf{B}\mathsf{X}^h\Lambda^h\mathsf{C}^h + \mathsf{B}\mathsf{X}^l\Lambda^l\mathsf{C}^l,$$

so from (4.7), we have

$$\mathsf{X}^{h\top}\mathsf{A}\mathsf{Y} = (\mathsf{X}^{h\top}\mathsf{B}\mathsf{X}^h)\Lambda^h\mathsf{C}^h + (\mathsf{X}^{h\top}\mathsf{B}\mathsf{X}^l)\Lambda^l\mathsf{C}^l = \Lambda^h\mathsf{C}^h.$$

By substituting this equality together with (4.27) into (4.28), we have

$$f(\beta^*) = -\alpha^\top(\Lambda^h\mathsf{C}^h)(\mathsf{C}^{h\top}\Lambda^h\mathsf{C}^h + \mathsf{C}^{l\top}\Lambda^l\mathsf{C}^l)^{-1}(\Lambda^h\mathsf{C}^h)^\top\alpha + \alpha^\top\Lambda^h\alpha\,. \tag{4.29}$$

Invertibility of C^h follows from Lemma 4.1 and the condition (4.22), so that $(\Lambda^h)^{1/2}\mathsf{C}^h$ is invertible, and we can transform (4.29) to

$$f(\beta^*) = -\alpha^\top(\Lambda^h)^{1/2}\left[\mathbb{1} + (\mathsf{C}^{h\top}(\Lambda^h)^{1/2})^{-1}\mathsf{C}^{l\top}\Lambda^l\mathsf{C}^l((\Lambda^h)^{1/2}\mathsf{C}^h)^{-1}\right]^{-1}(\lambda^h)^{1/2}\alpha + \alpha^T\Lambda^h\alpha. \tag{4.30}$$

For any matrix A with $(\mathbb{1}+\mathsf{A})$ nonsingular, we have that

$$(\mathbb{1}+\mathsf{A})^{-1} = \mathbb{1} - \mathsf{A} + (\mathbb{1}+\mathsf{A})^{-1}\mathsf{A}^2. \tag{4.31}$$

Moreover, if A is symmetric positive semidefinite, the last term $(\mathbb{1}+\mathsf{A})^{-1}\mathsf{A}^2$ is symmetric positive semidefinite, since if we write the eigenvalue decomposition of A as $\mathsf{A} = \mathsf{U}\mathsf{S}\mathsf{U}^\top$ where U is orthogonal and S is nonnegative diagonal, we have that $(\mathbb{1}+\mathsf{A})^{-1}\mathsf{A}^2 = \mathsf{A}^2(\mathbb{1}+\mathsf{A})^{-1} = \mathsf{U}(\mathbb{1}+\mathsf{S})^{-1}\mathsf{S}^2\mathsf{U}^\top$. Thus for any vector z, we have from (4.31) that

$$-z^\top(\mathbb{1}+\mathsf{A})^{-1}z + z^Tz \le -z^T(\mathbb{1}-\mathsf{A})z + z^Tz = z^T\mathsf{A}z.$$

By substituting $\mathsf{A} = (\mathsf{C}^{h\top}(\Lambda^h)^{1/2})^{-1}\mathsf{C}^{l\top}\Lambda^l\mathsf{C}^l((\Lambda^h)^{1/2}\mathsf{C}^h)^{-1}$ and $z = (\lambda^h)^{1/2}\alpha$ into this expression, we have from (4.30) that

$$\begin{aligned} f(\beta^*) &= \alpha^\top(\Lambda^h)^{1/2}(\mathsf{C}^{h\top}(\Lambda^h)^{1/2})^{-1}\mathsf{C}^{l\top}\Lambda^l\mathsf{C}^l((\Lambda^h)^{1/2}\mathsf{C}^h)^{-1}(\Lambda^h)^{1/2}\alpha \\ &= \alpha^\top(\mathsf{C}^h)^{-\top}\mathsf{C}^{l\top}\Lambda^l\mathsf{C}^l(\mathsf{C}^h)^{-1}\alpha \\ &\le \|\alpha\|^2\|(\mathsf{C}^{h\top}\mathsf{C}^h)^{-1}\|\|\mathsf{C}^{l\top}\Lambda^l\mathsf{C}^l\|. \end{aligned} \tag{4.32}$$

Note that $\|\alpha\| \le 1$, $\|\mathsf{C}^{l\top}\Lambda^l\mathsf{C}^l\| \le \lambda_{k+1}\|\mathsf{C}^{l\top}\mathsf{C}^l\|$ and

$$\left\|(\mathsf{C}^{h\top}\mathsf{C}^h)^{-1}\right\| = \left\|(\mathbb{1}-\mathsf{C}^{l\top}\mathsf{C}^l)^{-1}\right\| \le \sum_{i=0}^{\infty}\left\|\mathsf{C}^{l\top}\mathsf{C}^l\right\|^i = \frac{1}{1-\|\mathsf{C}^{l\top}\mathsf{C}^l\|}\,,$$

so by substituting into (4.32), we have

$$f(\beta^*) \le \lambda_{k+1}\frac{\|\mathsf{C}^{l\top}\mathsf{C}^l\|}{1-\|\mathsf{C}^{l\top}\mathsf{C}^l\|}\,. \tag{4.33}$$

Under condition (4.22) we have from Lemma 4.1 that

$$\|\mathsf{C}^{l\top}\mathsf{C}\| \leq \mathrm{Tr}(\mathsf{C}^{l\top}\mathsf{C}) \leq \frac{k\left(E(\mathsf{X}^h) - E(\mathsf{Y})\right)}{\lambda_k - \lambda_{k+1}} \leq \frac{1}{2},$$

so that

$$\frac{\|\mathsf{C}^{l\top}\mathsf{C}^l\|}{1 - \|\mathsf{C}^{l\top}\mathsf{C}^l\|} \leq 2\|\mathsf{C}^{l\top}\mathsf{C}^l\| \leq 2\frac{k\left(E(\mathsf{X}^h) - E(\mathsf{Y})\right)}{\lambda_k - \lambda_{k+1}},$$

yielding (4.24). □

4.3. **Criteria Used in Random Sampling for Local Basis Functions.** In our local basis construction problem, we identify A and B in Section 4.2 with S and S^*, respectively, from (2.17). After obtaining a number of random a-harmonic functions $\Gamma_n = \{\gamma_1, \gamma_2, \ldots, \gamma_n\}$, we compute the energy according to the definition in (4.10):

$$E(\Gamma_n) = \frac{\sum_{i=1}^n \langle a(x)|\nabla_x \gamma_i|^2\rangle_{\mathcal{E}(\omega)}}{\sum_{i=1}^n \langle a(x)|\nabla_x \gamma_i|^2\rangle_{\mathcal{E}(\omega^*)}}\,. \tag{4.34}$$

By Theorem 4.1, larger values of E indicate better sampling strategies.

Remark 3. The criterion in Theorem 4.1 indicates that a sampling strategy that results in energy $E(\mathsf{Y})$ close to the optimal value of $E(\mathsf{X}^h)$ will result in a small Kolmogorov width. Larger energy is achieved when the samples have their energies largely supported in the interior ω. Further, it suggests that construction of a-harmonic functions through random sampling of boundary conditions may not be the best strategy, because the boundary layer close to $\partial\omega^*$ quickly damps out the solution and the energies concentrate in the margin $\omega^*\backslash\omega$, leading to relatively small energy in ω, thus a small value of the energy (4.34). This would give a weak bound for the Kolmogorov width. These observations are borne out by the numerical experiments reported in the next section.

Theorem 4.2. In two-dimensional space, given any $u \in H_a(\omega^*)/\mathbb{R}$ and a subspace $\Gamma_n = \mathrm{Span}\{\gamma_1, \gamma_2, \ldots, \gamma_n\}$, a list of random samples of a-harmonic functions, the accuracy of approximating u with a function γ from Γ_n is bounded by the following estimate:

$$\min_{\gamma\in\Gamma_n} \|u - \gamma\|_{\mathcal{E}(\omega)} \leq \|u\|_{\mathcal{E}(\omega^*)} \left(e^{-n^{1/3}} + d(\Psi_n, \Gamma_n)\right)\,,$$

where Ψ_n is defined in (2.12) and $d(\Psi_n, \Gamma_n)$ is defined in equation (4.11).

Proof. Without loss of generality, we assume $\|u\|_{\mathcal{E}(\omega^*)} = 1$. Consider the optimal basis $\{\psi_i\}_{i=1}^\infty \subset H_a(\omega^*)/\mathbb{R}$ computed in (2.15), for which we have

$$\langle \psi_i, \psi_j\rangle_{\mathcal{E}(\omega^*)} = \delta_{ij}\,, \quad \langle \psi_i, \psi_j\rangle_{\mathcal{E}(\omega)} = \lambda_i \delta_{ij}.$$

We therefore have scalars $u_1, u_2, \ldots$ such that

$$u = \sum_{i=1}^\infty u_i \psi_i, \quad \sum_{i=1}^\infty u_i^2 = 1\,.$$

Define $\tilde{u} \in H_a(\omega^*)/\mathbb{R}$ by

$$\tilde{u} = \sum_{i=1}^{n} u_i \psi_i \,,$$

the restriction $v = P\tilde{u} \in H_a(\omega)/\mathbb{R}$ has

$$\|u - v\|_{\mathcal{E}(\omega)} = \|u - P\tilde{u}\|_{\mathcal{E}(\omega)} = \left(\sum_{i=n+1}^{\infty} u_i^2 \lambda_i \right)^{1/2} .$$

By the definition (4.11) of Kolmogorov width, there exists $\gamma \in \Gamma_n$ such that

$$\|v - \gamma\|_{\mathcal{E}(\omega)} \leq \|\tilde{u}\|_{\mathcal{E}(\omega^*)} d(\Psi_n, \Gamma_n) \,,$$

where $\Psi_n := \mathrm{Span}\{\psi_1, \ldots, \psi_n\}$. We further note that

$$\|\tilde{u}\|_{\mathcal{E}(\omega^*)} = \left(\sum_{i=1}^{n} u_i^2 \right)^{1/2} ,$$

and therefore

$$\begin{aligned} \|u - \gamma\|_{\mathcal{E}(\omega)} &\leq \|u - v\|_{\mathcal{E}(\omega)} + \|v - \gamma\|_{\mathcal{E}(\omega)} \\ &\leq \left(\sum_{i=n+1}^{\infty} u_i^2 \lambda_i \right)^{1/2} + \left(\sum_{i=1}^{n} u_i^2 \right)^{1/2} d(\Psi_n, \Gamma_n) \\ &\leq (\lambda_{n+1})^{1/2} + d(\Psi_n, \Gamma_n) \\ &\leq e^{-n^{1/3}} + d(\Psi_n, \Gamma_n) \,, \end{aligned}$$

where the last inequality comes from Theorem 2.2. □

5. Computational Results

We present numerical results in this section that show how the Kolmogorov width between the optimal space and the solution space constructed via random sampling decreases with the number of basis functions, for different sampling strategies. Throughout the section, the domain ω and enlarged domain ω^* are defined by

$$\omega = [-1, 1] \times [-1, 1], \quad \omega^* = [-1.4, 1.4] \times [-1.4, 1.4].$$

The media $a(x, y)$ is defined to be

$$\begin{aligned} a(x,y) = \frac{1}{5} \Bigg(&\frac{1.1 + \sin(7\pi x)}{1.1 + \sin(7\pi y)} + \frac{1.1 + \sin(9\pi y)}{1.1 + \cos(9\pi x)} + \frac{1.1 + \cos(13\pi y)}{1.1 + \cos(13\pi x)} \\ &+ \frac{1.1 + \cos(9\pi x)}{1.1 + \sin(9\pi y)} + \frac{1.1 + \sin(7\pi y)}{1.1 + \sin(7\pi x)} \Bigg), \quad (x, y) \in \omega^* \,. \end{aligned} \tag{5.1}$$

Numerical results will be shown for discretization parameters $\mathrm{d}x = \mathrm{d}y = 1/40$.

The reference solution is obtained from the procedure summarized in Section 2.3. To find the optimal solution space, we prepare the entire a-harmonic function space by going through all possible boundary condition configurations, before computing the general eigenvalue problem (2.17) for basis selection. This process requires computation

of the elliptic equation (2.14) 444 times (each time with a hat function on the boundary of $\partial\omega^*$ concentrated at a specific grid point) followed by computation of the generalized eigenpairs of two matrices of size 444×444.

We also implement the random sampling method proposed in Section 3 with four different sampling strategies, together with a "boundary i.i.d. function" method adopted from [5]:

1. Interior delta: Perform harmonic projection described in Section 3.1 of a given delta function supported in ω;
2. Interior i.i.d. function: Perform harmonic projection described in Section 3.1 of a function given by i.i.d. Gaussian supported in ω;
3. Full-domain i.i.d. function: Same as the previous one, except that the support is enlarged to ω^*;
4. Random Gaussian: Perform harmonic projection of Gaussian distribution centered at randomly chosen point in ω.
5. Boundary i.i.d. function: This is the strategy described in (3.7) where the boundary condition on $\partial\omega^*$ is defined by i.i.d. Gaussian variables.

As we see below, the five strategies have varying degrees of efficiency, but all capture the low rank structure of the optimal space.

High-Energy Modes. The first four modes $\{\phi_{1,2,3,4}\}$ of the reference solution are shown in Figure 5.1. These are obtained by following the procedure described in Section 2.3. We note here the presence of boundary layers in ω^*, as the functions exhibit fine scale oscillations near the boundary $\partial\omega^*$; moreover, these oscillations in the boundary layer are unseen when solutions are confined to the patch ω.

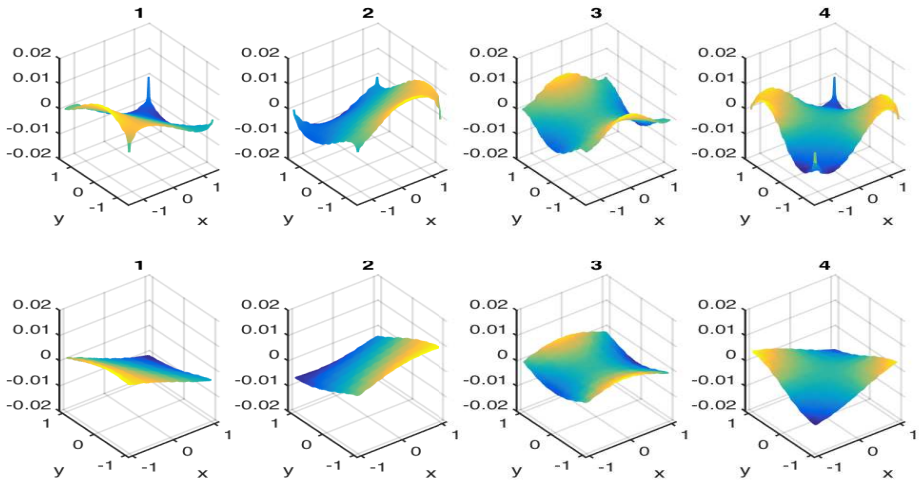

FIGURE 5.1. The first row shows the four modes $\phi_{1,2,3,4}$ supported on ω^*, and the second row shows the same modes confined in ω.

Recovery of General Eigenvalues. We now describe results obtained by random sampling method with the five sampling strategies. For each strategy, we sample only 20 a-harmonic functions for the computation in equation (3.2), hoping that these 20 random samples still capture the highest energy modes. In Figure 5.2, we plot (in log scale) the 20 generalized eigenvalues obtained from each of the five sampling strategies, together with the leading 20 eigenvalues from the optimal reference solution. All methods give almost exponential decay of the eigenvalues, but the random Gaussian method does by far the best at tracking the eigenvalues of the reference solution.

It is expected that as the number of random samples increases, all random methods should do better at capturing the eigenvalues of the reference solution. We see that this is true in Figure 5.3, where we use 300 random samples for all five sampling strategies. All except the strategies involving the full-domain i.i.d. function and possibly the boundary i.i.d. function do well at matching the reference eigenvalues.

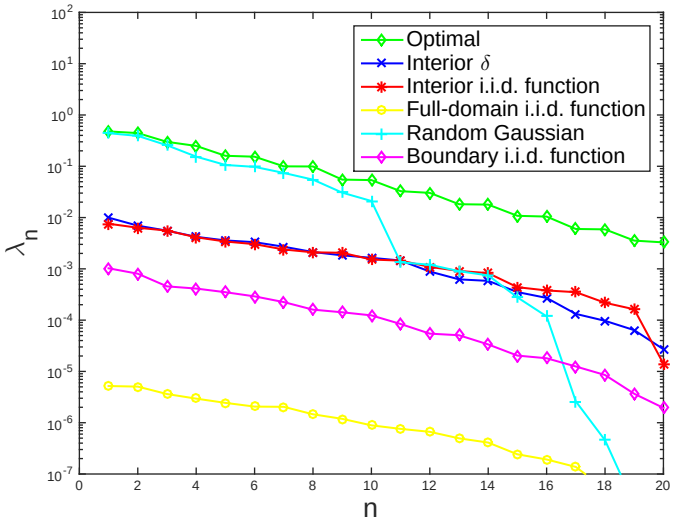

FIGURE 5.2. Eigenvalues obtained from the five different random sampling strategies, using 20 samples each, and the leading 20 eigenvalues from the reference solution. Eigenvalues are computed from (3.2) for the random sampling strategies and (2.17) for the reference solution. All methods show almost exponential decay of the eigenvalues, with the Random Gaussian strategy doing best at matching the reference eigenvalues.

Figure 5.4 shows the recovery of eigenspace by random sampling procedures. We regard Φ_5, defined in (2.10), as the optimal space (the space expanded by the five modes with highest energies), and use Φ_m^r defined in (3.3) to approximate it, for $m = 1, 2, \ldots, 20$. The vertical axis shows Kolmogorov width, whose computation is described in the appendix A. As expected, using more random modes leads to better recovery and thus smaller Kolmogorov width (4.11). The plots show Kolmogorov width decays roughly exponentially fast with respect to m, for all five sampling strategies. Once

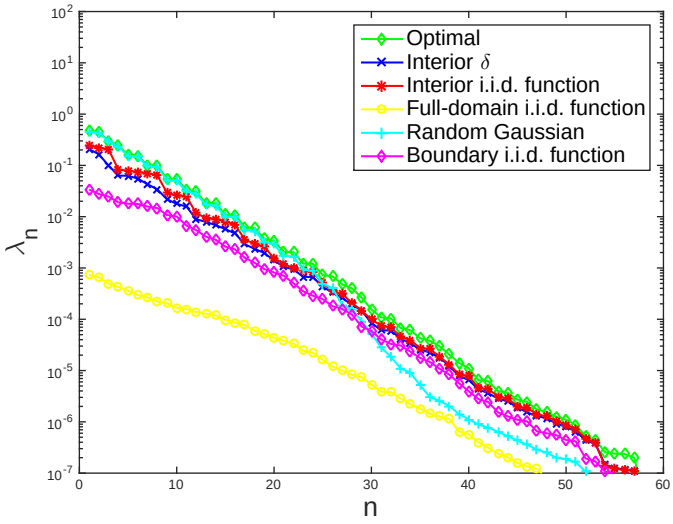

FIGURE 5.3. Same as Figure 5.2, but with 300 random samples instead of 20. Since $N_r \sim N_y$, the eigenvalues obtained from ransom sampling tend to match the reference eigenvalues more closely.

again, the Random Gaussian strategy is by far the best: Φ_{20}^r approximates Φ_5 with accuracy near 10^{-5}. The other four strategies attain accuracies of around 10^{-1} to 10^{-2} for $m = 20$.

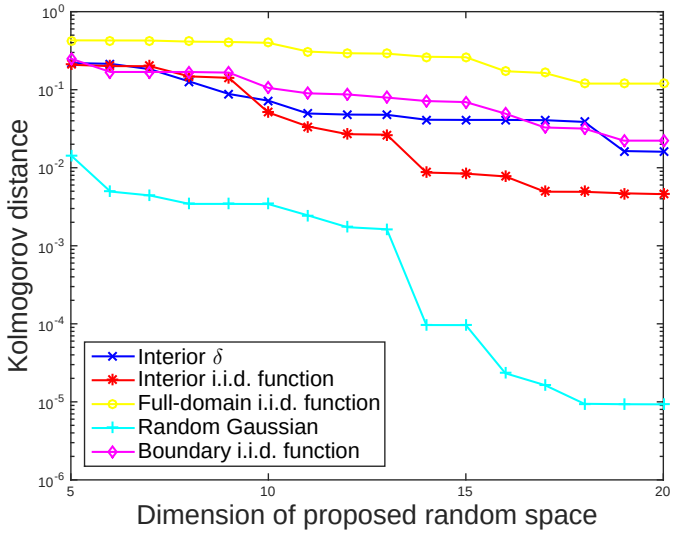

FIGURE 5.4. Kolmogorov distance between Φ_5 and Φ_m^r, for $m = 1, 2, \ldots, 20$ sampling modes.

Eigenspace Recovery for the Random Gaussian Strategy. Finally, we focus on the the Random Gaussian sampling strategy, which is clearly the most successful one. In Figure 5.5, we plot in the first row the high energy modes $\{\phi_{1,2,3,4}\}$ for the reference

solution, and in the second row we plot the high energy modes $\{\phi^r_{1,2,3,4}\}$ obtained from the Random Gaussian strategy with 20 samples. The similarity is evident.

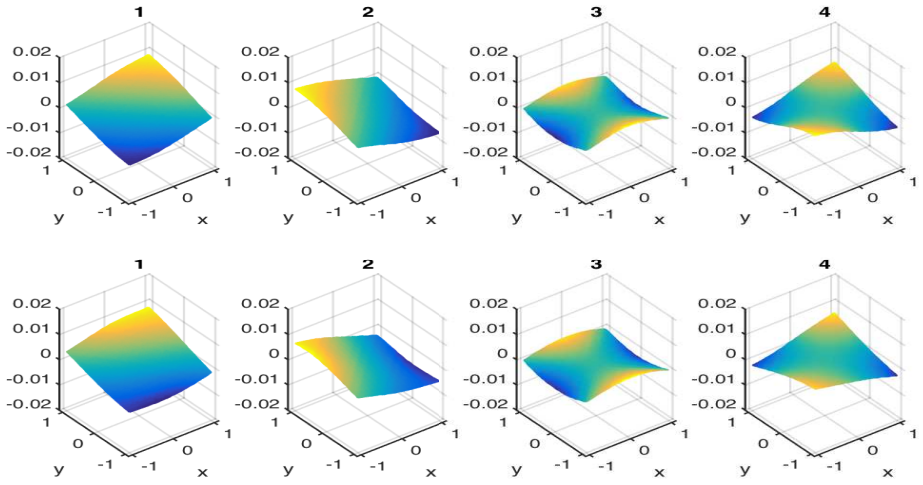

FIGURE 5.5. Recovery of high energy modes using Random Gaussian sampling strategy, with 20 samples. First row shows the first four modes of the reference solution; second row shows the first four modes obtained from the Random Gaussian sampling strategy.

6. Conclusion

In this paper we have proposed to compute elliptic equations with rough coefficients using random sampling strategies in the context of generalized finite element methods. It is shown that the random sampling methods capture the main part of the local solution spaces with high accuracy, and that efficiency can be evaluated by the energy contained in the proposed random space. Numerical comparisons of five different sampling strategies show that the use of Random Gaussian functions sampled in the interior of each patch, in conjunction with a-harmonic projection, yield the smallest Kolmogorov width, and thus perform better in approximating the solution space.

Appendix A. Calculation of the Kolmogorov Distance

Suppose we are given the optimal space

$$\Phi_m = \operatorname{Span} X, \text{ where } X = [\phi_1, \ldots, \phi_m]\,,$$

and a proposed space

$$\Phi^r_n = \operatorname{Span} Y, \text{ where } Y = [\phi^r_1, \ldots, \phi^r_n]\,, \quad \text{where } n \geq m,$$

such that

$$\langle \phi_i, \phi_j \rangle_{\mathcal{E}(\omega^*)} = \delta_{ij} \quad \text{and} \quad \langle \phi^r_i, \phi^r_j \rangle_{\mathcal{E}(\omega^*)} = \delta_{ij}\,.$$

Recall the following definition of Kolmogorov width from (4.11):

$$d(\Phi_m, \Phi_n^r) = \max_{x\in\Phi_m, \|x\|_{\mathcal{E}(\omega^*)}=1} \min_{y\in\Phi_n^r} \|x-y\|_{\mathcal{E}(\omega)}\,.$$

To calculate $d(\Phi_m, \Phi_n^r)$ explicitly, we write $x = X\alpha$ for some $\alpha \in \mathbb{R}^m$ and $y = Y\beta$ for some $\beta \in \mathbb{R}^n$. The Kolmogorov width is achieved when $\|x\|_{\mathcal{E}(\omega^*)} = 1$, which implies that $\|\alpha\|_{l^2} = 1$. By expanding the objective, we have

$$\begin{aligned}\frac{1}{2}\|x-y\|^2_{\mathcal{E}(\omega)} &= \frac{1}{2}\langle X\alpha - Y\beta, X\alpha - Y\beta\rangle_{\mathcal{E}(\omega)} \\ &= \frac{1}{2}\beta^\top Y_A\beta - \alpha^\top C_A\beta + \frac{1}{2}\alpha^\top X_A\alpha\,,\end{aligned}$$

where $(Y_A)_{ij} = \langle \phi_i^r, \phi_j^r\rangle_{\mathcal{E}(\omega)}$, $(X_A)_{ij} = \langle \phi_i, \phi_j\rangle_{\mathcal{E}(\omega)}$ and $(C_A)_{ij} = \langle \phi_j, \phi_i^r\rangle_{\mathcal{E}(\omega)}$. The minimizing value of β is given explicitly by

$$\beta = Y_A^{-1}C_A\alpha\,,$$

for which we have

$$\frac{1}{2}d(\Phi_m, \Phi_n^r)^2 = \max_{\|\alpha\|_{l^2}=1} \alpha^\top(X_A - C_A^\top Y_A^{-1}C_A)\alpha = \|X_A - C_A^\top Y_A^{-1}C_A\|_2^2\,.$$

The Kolmogorov width is therefore

$$d(\Phi_m, \Phi_n^r) = \sqrt{2}\|X_A - C_A^\top Y_A^{-1}C_A\|_2\,.$$

References

[1] I. Babuška and R. Lipton, *Optimal local approximation spaces for generalized finite element methods with application to multiscale problems*, Multiscale Model. Simul., 9 (2011), pp. 373–406.

[2] I. Babuška and J. Melenk, *The partition of unity method*, International Journal for Numerical Methods in Engineering, 40 (1997), p. 727758.

[3] M. Bebendorf, *Why finite element discretizations can be factored by triangular hierarchical matrices*, SIAM Journal on Numerical Analysis, 45 (2007), pp. 1472–1494.

[4] A. Bensoussan, J. Lions, and G. Papanicolaou, *Asymptotic Analysis for Periodic Structures*, AMS Chelsea Publishing Series, American Mathematical Society, 2011.

[5] V. M. Calo, Y. Efendiev, J. Galvis, and G. Li, *Randomized oversampling for generalized multiscale finite element methods*, Multiscale Modeling & Simulation, 14 (2016), pp. 482–501.

[6] W. E and B. Engquist, *The heterogeneous multi-scale methods*, Commun. Math. Sci., 1 (2003), pp. 87–133.

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

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

[9] J. Galvis and Y. Efendiev, *Domain decomposition preconditioners for multiscale flows in high contrast media: Reduced dimension coarse spaces*, Multiscale Model. Simul., 8 (2010), pp. 1621–1644.

[10] W. Hackbusch, *Hierarchical Matrices: Algorithms and Analysis*, Springer Series in Computational Mathematics, Springer, 2015.

[11] N. Halko, P. G. Martinsson, and J. A. Tropp, *Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions*, SIAM Review, 53 (2011), pp. 217–288.

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

[13] ———, *A multiscale finite element method for pdes with oscillatory coefficients*, Notes on Numerical Fluid Mechanics, 70 (1999), pp. 58–69.

[14] T. Y. Hou, X.-H. Wu, and Z. Cai, *Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients*, Math. Comp., 68 (1999), pp. 913–943.

[15] T. Y. Hou and P. Zhang, *Sparse operator compression of higher-order elliptic operators with rough coefficients*, Research in Mathematical Sciences, (in press).

[16] W. B. Johnson and J. Lindenstrauss, *Extensions of Lipshitz mapping into Hilbert space*, Contemporary Mathematics, 26 (1984), pp. 189–206.

[17] L. Lin, J. Lu, L. Ying, and W. E, *Adaptive local basis set for KohnSham density functional theory in a discontinuous Galerkin framework I: Total energy calculation*, J. Comput. Phys., 231 (2012), pp. 2140–2154.

[18] R. Lipton, P. Sinz, and M. Stuebner, *Uncertain loading and quantifying maximum energy concentration within composite structures*, Journal of Computational Physics, 325 (2016), pp. 38 – 52.

[19] A. Målqvist and D. Peterseim, *Localization of elliptic multiscale problems*, Math. Comp., 83 (2014), pp. 2583–2603.

[20] P. Ming and X. Yue, *Numerical methods for multiscale elliptic problems*, J. Comput. Phys., 214 (2006), pp. 421–445.

[21] H. Owhadi, *Bayesian numerical homogenization*, Multiscale Modeling & Simulation, 13 (2015), pp. 812–828.

[22] H. Owhadi and L. Zhang, *Metric-based upscaling*, Comm. Pure Appl. Math., 60 (2007), pp. 675–723.

[23] H. Owhadi, L. Zhang, and L. Berlyand, *Polyharmonic homogenization, rough polyharmonic splines and sparse super-localization*, M2AN Math. Model. Numer. Anal., 48 (2014), pp. 517–552.

[24] G. Papanicolaou, *Asymptotic analysis of transport process*, Bull. American Math. Soc., 81 (1975), pp. 330–392.

[25] A. Pinkus, *N-widths in approximation theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Springer, 1985.

Mathematics Department, University of Wisconsin-Madison, 480 Lincoln Dr., Madison, WI 53706 USA.

E-mail address: ke@math.wisc.edu

Mathematics Department, University of Wisconsin-Madison, 480 Lincoln Dr., Madison, WI 53706 USA.

E-mail address: qinli@math.wisc.edu

Department of Mathematics, Department of Physics and Department of Chemistry, Duke University, Box 90320, Durham, NC 27708 USA.

E-mail address: jianfeng@math.duke.edu

Computer Sciences Department, University of Wisconsin-Madison, 1210 W Dayton St, Madison, WI 53706 USA.

E-mail address: swright@cs.wisc.edu