MULTISCALE MODEL. SIMUL.
Vol. 21, No. 3, pp. 849–883

EXPONENTIALLY CONVERGENT MULTISCALE METHODS FOR 2D HIGH FREQUENCY HETEROGENEOUS HELMHOLTZ EQUATIONS*

YIFAN CHEN†, THOMAS Y. HOU†, AND YIXUAN WANG†

Abstract. In this paper, we present a multiscale framework for solving the Helmholtz equation in heterogeneous media without scale separation and in the high frequency regime where the wavenumber k can be large. The main innovation is that our methods achieve a nearly exponential rate of convergence with respect to the computational degrees of freedom, using a coarse grid of mesh size $O(1/k)$ without suffering from the well-known pollution effect. The key idea is a nonoverlapped domain decomposition and its associated coarse-fine scale decomposition of the solution space that adapts to the media property and wavenumber; this decomposition is inspired by the multiscale finite element method. We show that the coarse part is of *low complexity* in the sense that it can be approximated with a nearly exponential rate of convergence via local basis functions, due to the compactness of a restriction operator that maps Helmholtz-harmonic functions to their interpolation residues on edges, while the fine part is *local* such that it can be computed efficiently using the local information of the right-hand side. The combination of the two parts yields the overall nearly exponential rate of convergence of our multiscale method. Our method draws many connections to multiscale methods in the literature, which we will comment in detail. We demonstrate the effectiveness of our methods theoretically and numerically; an exponential rate of convergence is consistently observed and confirmed. In addition, we observe the robustness of our methods regarding the high contrast in the media numerically. We specifically focus on 2D problems in our exposition since the geometry of nonoverlapped domain decomposition is simplest to explain in such cases.

Key words. the Helmholtz equation, heterogeneous media, high frequency, exponential convergence, multiscale methods, high contrast

MSC codes. 65N12, 65N15, 65N30, 31A35

DOI. 10.1137/22M1507802

1. Introduction.

This paper focuses on solving the Helmholtz equation in heterogeneous media and highfrequency regimes. We consider the model problem in a bounded polygonal domain $\Omega \subset \mathbb{R}^d$ with a Lipschitz boundary Γ. For generality, the boundary can contain three disjoint parts $\Gamma = \Gamma_D \cup \Gamma_N \cup \Gamma_R$, where Γ_D, Γ_N, and Γ_R correspond to the Dirichlet-, Neumann-, and Robin-type conditions, respectively. Given positive constants $A_{\min}$, $A_{\max}$, $\beta_{\min}$, $\beta_{\max}$, $V_{\min}$, $V_{\max}$ and functions $A, \beta, V : \Omega \to \mathbb{R}$ that satisfy $A_{\min} \le A(x) \le A_{\max}$, $\beta_{\min} \le \beta(x) \le \beta_{\max}$, and $V_{\min} \le V(x) \le V_{\max}$, the Helmholtz equation with homogeneous boundary conditions[1] is formulated as follows:

$$\begin{cases} -\nabla \cdot (A \nabla u) - k^2 V^2 u = f \ \text{ in } \Omega, \\ u = 0 \ \text{ on } \Gamma_D, \\ A \nabla u \cdot \nu = T_k u \ \text{ on } \Gamma_N \cup \Gamma_R . \end{cases} \tag{1.1}$$

*Received by the editors July 7, 2022; accepted for publication (in revised form) March 6, 2023; published electronically July 13, 2023.
https://doi.org/10.1137/22M1507802
Funding: This research is in part supported by NSF grants DMS-1912654 and DMS 2205590. The first author is also grateful to the support from the Caltech Kortchak Scholar Program.
†Applied and Computational Mathematics, Caltech, Pasadena, CA 91125 USA (yifanc@caltech.edu, hou@cms.caltech.edu, roywang@caltech.edu).
[1]For simplicity of presentation, homogeneous boundary conditions are considered here. Generalization to nonhomogeneous data is straightforward; see section 5.3 or [13] (section 5.3).

Here, ν is the outer normal to the boundary. The boundary operator satisfies $T_k u = 0$ for $x \in \Gamma_N$ and $T_k u = ik\beta u$ for $x \in \Gamma_R$, where i denotes the imaginary number. The wavenumber k is real and positive, and functions u and f are complex-valued. The aim of this paper is to design a multiscale method for solving (1.1) that achieves a nearly exponential rate of convergence with respect to the computational degrees of freedom. This is a challenging problem due to combined difficulties of heterogeneity and high frequency. We review the related literature of this research field in section 1.1 and discuss our methodology as well as its motivations and related work in section 1.2.

1.1. Literature for solving Helmholtz equations. The Helmholtz equation has been widely used in studying wave propagation in complex media. Numerical simulation of this equation still remains a challenging task, especially in the regime where the wavenumber k is large. The main numerical difficulty lies in the highly oscillatory pattern of the solution. Furthermore, the operator in the equation is indefinite, which leads to severe instability issues for standard numerical solvers such as the finite element method (FEM). Indeed, a well-known preasymptotic effect called the pollution effect [5] can occur—that is, in order to get a reasonably accurate solution, the mesh size H in the FEM needs to be much smaller than $1/k$. For example, for a standard P1-FEM approach, the mesh size needs to satisfy $H = O(1/k^2)$ for quasi-optimality of the solution [1, 5]. This constraint on H is much stronger than the typical condition in the approximation theory for representing an oscillatory function with frequency k, where $H = O(1/k)$, i.e., a fixed number of grid points per wavelength, would suffice for an accurate approximate solution.

In the literature, there have been many attempts to overcome or alleviate the difficulty associated with the pollution effect so that a mesh size of $H = O(1/k)$ can be used. We highlight two classes of methods, namely, the *hp*-FEM and multiscale methods, which can theoretically deal with the pollution effect under their respective model assumptions. The *hp*-FEM is proposed in [40, 41], which is an FEM using local high order polynomials. It is shown that by choosing the degrees of local polynomials $p = O(\log k)$, the pollution effect can be suppressed in principle for the Helmholtz equation with constant A, V, and β. Nevertheless, to the best of our knowledge, there have been no theoretical results for this methodology when these coefficients become rough. There have been some recent developments for *hp*-FEM methods when piecewise regularity of the coefficients is assumed [6, 35]. In general, it is well known that polynomials might behave arbitrarily badly even for elliptic equations with rough coefficients [4].

Multiscale methods, on the other hand, have long been developed to address the difficulty associated with rough coefficients in elliptic equations. In particular, we mention the localized orthogonal decomposition (LOD)- and gamblets-related approaches [39, 29, 47, 45, 46, 11, 12], variants of the multiscale finite element method (MsFEM) [33, 15, 32, 14, 21, 13], and generalized FEMs based on partition of unity methods (PUMs) [2, 53, 9, 10, 3, 52, 37, 38], which are the ones most related to this paper. Recently, the LOD method has been generalized to the case of Helmholtz equations with high wavenumber and heterogeneous media [48, 24, 8, 49]. They show that with a coarse mesh of size $O(H)$ and localized multiscale basis functions of support size $O(H \log(1/H) \log k)$, the pollution effect can be overcome once the stability constant of the solution operator of the Helmholtz equation is of at most polynomial growth. An error of at most $O(H)$ is established. Very recently, there is also a superlocalized version of an LOD-type method for the Helmholtz equations, proposed in [20], where the support of basis functions is further reduced to $O(H \log^{(d-1)/d}(k/H))$.

From the perspective of MsFEM methodology, the authors in [23] introduce wavelet-based edge MsFEM to address the pollution effect successfully. Their basis functions are all of local support size $O(H)$. On the theoretical side, they require $O(k)$ number of basis functions in each element in order to achieve $O(H)$ accuracy. In contrast, our method in this paper, which can be viewed as a generalization of the MsFEM, only requires an $O(\log^{d+1} k)$ number of basis function of support size $O(H)$ in each element to handle the pollution effect and to achieve $O(H)$ accuracy. More importantly, our method yields an overall exponential rate of convergence regarding the number of basis functions, thanks to a systematic decomposition and treatment of coarse and fine scale parts of the solution.

In the literature, multiscale methods with exponential convergence for elliptic equations with rough media first appeared in [2], which is based on local optimal basis approximation combined with the PUM. There have been a number of recent papers that are actively working on improving the methodology [53, 9, 10, 3, 52, 37, 38], aiming for more refined continuous and discrete analysis, randomized computation, efficient implementation, and generalization beyond elliptic equations. Our initial work [13] on exponentially convergent multiscale methods for elliptic equations draws many motivations from these results, especially the Caccioppoli-type inequality that is essential for proving the exponential convergence. Different from the PUM-based approach, our method is based on nonoverlapped domain decomposition. More comparisons will be discussed in subsection 1.2. While revising this paper on solving the Helmholtz equations, we found that the authors in [36] also proposed an exponentially convergent method for the Helmholtz equations using the PUM-based optimal local approximation methodology.

In addition to those methods mentioned above, there have also been several algorithms based on the MsFEM methodology [43, 22] or the heterogenous multiscale methodology [44] with particular empirical success for solving the Helmholtz equation. It is also worth noting that, in conjunction with designing a good discretization scheme as above, one could also consider fast solvers for the discrete linear system. See, for example, the method of sweeping preconditioners [16, 17, 50], where a preconditioning matrix is constructed to compute approximations of the Schur complements successively. Very recently, the LOD approach has also been combined with the hierarchical approach of gamblets [28] to get a multiresolution solver for the discrete system.

1.2. Main contributions and motivations. In this paper, we propose a multiscale framework for solving the Helmholtz equation in rough media and high frequency regimes, specifically in dimension $d = 2$, where the mesh geometry of the nonoverlapped domain decomposition is simplest to describe. Our idea is based on a multiscale method in our previous work [13] for solving elliptic equations with rough coefficients in an exponentially convergent manner. This paper aims to extend this framework to the more challenging Helmholtz equation where the operator is non-Hermitian and indefinite. It is perhaps surprising that the techniques in multiscale methods for elliptic equations can be systematically adapted to the Helmholtz equation. Indeed, it has been proved in [18] that the Green function of the Helmholtz equations requires a polynomial in k number of degrees of freedom to approximate, where they consider basis functions independent of the right-hand side. Here, our results demonstrate that one can actually compress the solution operator exponentially efficiently by adding a number of local basis functions that depend on the local information of the right-hand side. This shows that one can still achieve significant

compression of the high frequency Helmholtz solution operator with rough coefficients by developing a data-driven compression operator adapted to the right-hand side.

We outline the main contributions of this paper below.

1. In studying the solution behavior of the Helmholtz equation (1.1), we introduce a coarse-fine scale decomposition of its solution space. This decomposition is adapted to the coarse mesh structure; a mesh size of $O(1/k)$ suffices to make this coarse-fine scale decomposition well defined. Moreover, the decomposition is adapted to the coefficients A, V, β and the wavenumber k.
2. Analytically, we show the fine scale part is of $O(H)$ in the energy norm, and it can be computed efficiently by solving the Helmholtz equations locally. Meanwhile, we prove that the space of the coarse scale part is of low complexity such that there exist local multiscale basis functions that can approximate this part in a nearly exponentially convergent manner. These serve as the cornerstone of our multiscale numerical method.
3. Numerically, we propose a multiscale framework that solves the two parts separately. The nearly exponential rate of convergence in the energy norm and L^2 norm is theoretically proved in this paper.
4. Experimentally, we conduct a number of numerical tests and observe consistently that our multiscale methods give a nearly exponential rate of convergence, even for problems with high-contrast media. Based on these numerical studies, several recommendations for efficient implementations of our methods are provided, especially on how to design the offline and online computation to handle multiple right-hand sides efficiently.

To the best of our knowledge, this multiscale framework is the first one that can be proved rigorously to achieve a nearly exponential rate of convergence in solving (1.1) with rough A, β, V and large k, especially for $d = 2$. It generalizes our previous work on exponential convergence for solving rough elliptic equations [13], which is motivated by the PUM approach using optimal local approximation spaces for elliptic equations [2].

Different from the PUM that uses an overlapped domain decomposition, our method relies on nonoverlapped domain decomposition and an edge coupling approach to combine local basis functions as in MsFEM. Our coarse-fine scale decomposition of the solution space is built on this nonoverlapped edge coupling. For elliptic equations, this decomposition is the same as the orthogonal decomposition in previous work on MsFEMs [32, 13] and approximate component mode synthesis [31, 30]. Under this line of methodology, this paper contributes a principled framework for obtaining nearly exponentially convergent basis functions for multiscale Helmholtz equations.

There are many differences between the multiscale methods based on PUMs and edge coupling. Basically, the support of basis functions in PUMs is usually larger than that of MsFEMs since nonoverlapped domain decomposition leads to smaller decomposed domains than its overlapped counterpart. There is no need to introduce additional freedom of partition of unity functions as well. On the other hand, in two dimensions, the number of local edges could be twice as many as the number of local domains, leading to more work in constructing the basis functions. Moreover, there will be an increasing design complexity for the nonoverlapped edge coupling approach for higher-dimensional problems since the boundaries of high-dimensional local domains become more complicated. This is why, in this paper, we are dedicated specifically to 2D Helmholtz equations for detailed analysis and numerical experiments.

We are not going to dive very deeply into the fundamental comparison between overlapped and nonoverlapped decomposition in multiscale methods. The aim of this paper is to demonstrate that one could achieve a nearly exponential convergence rate theoretically using the nonoverlapped edge coupling framework in a principled way and show that this method is very competitive numerically. A number of technical difficulties, such as the appropriate approximation space for the edge functions and the spectral analysis of the local restriction operator, are carefully addressed to lay out this framework. We believe this work could help future researchers understand and analyze multiscale methods that are built on different local decomposition and global coupling approaches.

Lastly, we remark that in principle, our multiscale algorithm can be applied to general Helmholtz equations numerically, while most of our theoretical results rely on analytical properties of the solution to (1.1), related to the well-posedness, stability and C^α estimates. Therefore, typical conditions (usually very mild) of these analytical properties will be assumed in this paper in order to get a rigorous theory. We will mention several references to these results in this paper. Some numerical examples in which these assumptions are violated will be also presented to illustrate the effectiveness of our algorithm in a general context.

1.3. Organization of the paper. The rest of this paper is organized as follows. In section 2, we review preliminary results for the Helmholtz equation, including the well-posedness, stability, adjoint problems, and Hölder C^α estimates. Section 3 is devoted to analyzing the solution space based on a coarse-fine scale decomposition. Moreover, the computational properties of the coarse and fine parts are rigorously studied in detail. Building upon these properties, in section 4 we develop the multiscale computational framework and prove the nearly exponential rate of convergence for our multiscale methods. The detailed numerical algorithms are discussed and implemented in section 5 for several Helmholtz equations. To improve the readability of our paper, some technical proofs of theorems and propositions will be deferred to section 6. Some concluding remarks are made in section 7.

2. Preliminaries on the Helmholtz equation.

Our multiscale algorithm relies on an in-depth understanding of the solution space of (1.1). To achieve this, we first present several analytic results for (1.1), which will serve as preliminaries for our subsequent discussions. We cover the weak formulation, the well-posedness of the equation, the stability estimates of the solution, and Hölder estimates.

2.1. Notations. We use $H^1(\Omega)$ to denote the standard complex Sobolev space in Ω, containing L^2 functions with L^2 first order derivatives. We write $(u,v)_D := \int_D u\bar{v}$ for any domain D. We use C as a generic constant, and its value can change from place to place; we will state explicitly the parameters that this constant may or may not depend on.

2.2. Analytic results. For the model problem (1.1), we consider the complex Sobolev space $\mathcal{H}(\Omega) := \{u \in H^1(\Omega) : u|_{\Gamma_D} = 0\}$ in which functions have zero trace on the Dirichlet boundary. This space is equipped with the norm $\|\cdot\|_{\mathcal{H}(\Omega)}$ such that

$$\|u\|_{\mathcal{H}(\Omega)} := \int_\Omega A|\nabla u|^2 + k^2V^2|u|^2\,.$$

The dual space of $\mathcal{H}(\Omega)$ is denoted by $\mathcal{H}^{-1}(\Omega)$ equipped with the norm $\|\cdot\|_{\mathcal{H}^{-1}(\Omega)}$; by definition, one has

$$\|f\|_{\mathcal{H}^{-1}(\Omega)} := \sup_{v\in\mathcal{H}(\Omega)} \frac{|(f,v)_\Omega|}{\|v\|_{\mathcal{H}(\Omega)}}.$$

Now, we present several analytic results pertaining to the Helmholtz equation (1.1).

Weak formulation. The weak formulation of (1.1) is given by

$$a(u,v) := (A\nabla u, \nabla v)_\Omega - k^2(V^2u, v)_\Omega - (T_k u, v)_{\Gamma_N\cup\Gamma_R} = (f,v)_\Omega \quad \forall\, v\in\mathcal{H}(\Omega). \tag{2.1}$$

Continuity estimate. By the Cauchy–Schwarz and trace inequalities (see Lemma 3.1 of [40]), the sesquilinear form $a(\cdot,\cdot)$ is bounded on $\mathcal{H}(\Omega)$ with a constant C_c independent of k; i.e., for any $u, v\in\mathcal{H}(\Omega)$, one has the continuity estimate

$$|a(u,v)| \le C_c\|u\|_{\mathcal{H}(\Omega)}\|v\|_{\mathcal{H}(\Omega)}. \tag{2.2}$$

Well-posedness and stability. If Γ_R has positive $d-1$ dimensional measure, then under some mild conditions (see Assumption 2.3 and Theorem 2.4 in [25]), problem (2.1) admits a unique solution given the right-hand side $f\in L^2(\Omega)$. We will assume these conditions. Let the solution operator be N_k, so that $u = N_k f$. Under the same conditions, this operator is stable (Theorem 2.4 in [25]) in the sense that

$$C_{\mathrm{stab}}(k) := \sup_{f\in L^2(\Omega)\setminus\{0\}} \frac{\|N_k f\|_{\mathcal{H}(\Omega)}}{\|f\|_{L^2(\Omega)}} < \infty. \tag{2.3}$$

To avoid getting into detailed discussions of these assumptions and for simplicity of presentation, we will base most of our arguments on assuming (2.3) holds.

The stability constant $C_{\mathrm{stab}}(k)$ will depend on k in general, and obtaining an explicit characterization of this dependence has been a hard task; see [7, 8, 26, 42, 51]. A prevalent and reasonable assumption on the constant is that of polynomial growth, namely, $C_{\mathrm{stab}}(k)\le C(1+k^\gamma)$ for some constants γ and C; see, for example, [34]. We are not going into detailed discussions on this assumption here, while we mention that the final error estimate of our numerical solution in this paper will depend on $C_{\mathrm{stab}}(k)$ explicitly; thus, those estimates on $C_{\mathrm{stab}}(k)$ in the literature can be readily applied to our context.

In addition, stability for $f\in L^2(\Omega)$ can yield well-posedness and stability for $f\in\mathcal{H}^{-1}(\Omega)$. According to Lemma 2.1 in [48] and also [19], one has

$$\sup_{f\in\mathcal{H}^{-1}(\Omega)\setminus\{0\}} \frac{\|N_k f\|_{\mathcal{H}(\Omega)}}{\|f\|_{\mathcal{H}^{-1}(\Omega)}} \le kC_{\mathrm{stab}}(k). \tag{2.4}$$

Adjoint problems. Due to the presence of the Robin boundary condition, $a(\cdot,\cdot)$ is not Hermitian. Its adjoint sesquilinear form is defined as $a^*(u,v) = \overline{a(v,u)}$. The adjoint problem for (2.1) is given by $a^*(u,v) = (f,v)_\Omega$ for any $v\in\mathcal{H}(\Omega)$. It also corresponds to the following PDE:

$$\begin{cases} -\nabla\cdot(A\nabla u) - k^2V^2u = f \ \text{ on } \Omega, \\ u = 0 \ \text{ in } \Gamma_D, \\ A\nabla u\cdot\nu = T_k^* u \ \text{ on } \Gamma_N\cup\Gamma_R, \end{cases}$$

where $T_k^* u := \overline{T_k \overline{u}} = -T_k u$. The adjoint solution operator is denoted by N_k^*. One can readily check that $N_k^* \overline{f} = \overline{N_k f}$. Therefore, the adjoint problem admits the same stability constant as the original problem; namely, it holds

$$C_{\text{stab}}(k) = \sup_{f \in L^2(\Omega)\setminus\{0\}} \frac{\|N_k^\star f\|_{\mathcal{H}(\Omega)}}{\|f\|_{L^2(\Omega)}} < \infty .$$

The adjoint problem will play a valuable role when we analyze the convergence property of our multiscale methods for the Helmholtz equation.

C^α Hölder regularity. We will need the C^α estimates of the solution in order to demonstrate the theoretical properties of our multiscale methods.

PROPOSITION 2.1. *Suppose $d \le 3$ and that (2.3) holds. If $f \in L^2(\Omega)$, then the solution $u \in C^\alpha(\Omega)$ for some $\alpha \in (0,1)$.*

We defer the proof of this proposition to subsection 6.1.

Remark 2.2. The global regularity estimate may depend on the wavenumber k. Nevertheless, we only use it to show qualitatively that our solution is continuous so that the nodal interpolation in subsection 3.4.2 is mathematically rigorous. Later, when we derive error estimates of our methods, we will only use the local version of the regularity estimate, where the constant is independent of the wavenumber; see Lemma 6.2.

We have presented several critical analytic results for the Helmholtz equation. Based on these results, we are now ready to study the solution space of (1.1) in the next section. The key is a coarse-fine scale decomposition of the solution space, which will play an essential role in designing our multiscale algorithms.

3. Coarse-fine scale decomposition. In this section, we develop a coarse-fine scale operator-adapted decomposition of the solution space. This decomposition is adaptive to the mesh structure, and a mesh of size $H = O(1/k)$ suffices to make this coarse-fine scale decomposition well defined. We discuss the setting of the mesh structure in subsection 3.1, followed by introducing the coarse-fine scale decomposition in subsection 3.2. In subsection 3.3, we show the finescale part is local and small up to $O(H)$ in the $\mathcal{H}(\Omega)$ norm. In subsection 3.4, we show the coarse scale component can be approximated via local edge basis functions in a nearly exponentially convergent manner.

3.1. Mesh structure. We begin by discussing related concepts of the mesh structure. The focus here is on $d = 2$. In the mesh structure, we discuss 2D elements in subsection 3.1.1, one-dimensional edges and zero-dimensional nodes, and their neighborship in subsection 3.1.2. See also Figure 1 for illustrations.

3.1.1. Elements. We consider a shape regular and uniform partition of the domain Ω into finite elements, such as triangles and quadrilaterals. The collection of elements is denoted by $\mathcal{T}_H = \{T_1, T_2, \ldots, T_r\}$. For simplicity, we assume that each connected component of the domain is at least partitioned into two elements.

The mesh size is H, i.e., $\max_{T\in\mathcal{T}_H} \operatorname{diam}(T) = H$. The uniformity of the mesh implies $\min_{T\in\mathcal{T}_H} \operatorname{diam}(T) \ge c_0 H$ for some $0 < c_0 \le 1$ that is independent of H and T. The shape regularity property implies there is a constant $c_1 > 0$ independent of H and T such that $\max_{T\in\mathcal{T}_H} \operatorname{diam}(T)^d/|T| \le c_1$, where $|T|$ is the volume of T.

In this mesh, by using a scaling argument, the following Poincaré inequality will hold uniformly for $T \in \mathcal{T}_H$. This inequality will be used frequently later.

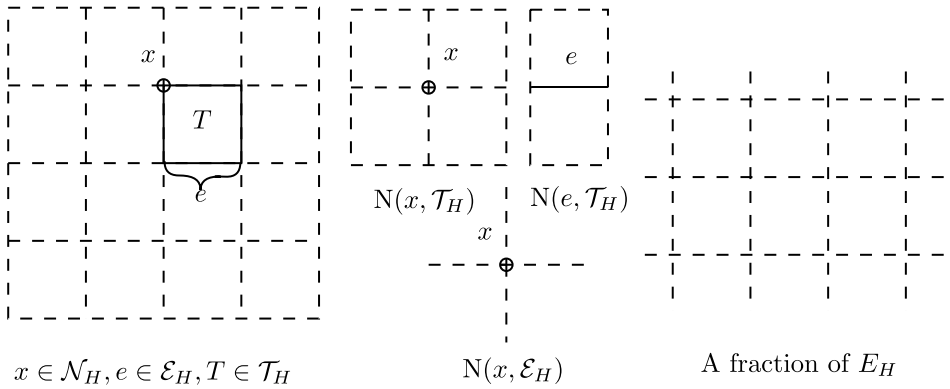

FIG. 1. *Geometry of the mesh.*

PROPOSITION 3.1 (the Poincaré inequality). *For any $T \in \mathcal{T}_H$ and a function $v \in H^1(T)$ that vanishes on one of the edges of T, it holds that*

$$\|v\|_{L^2(T)} \leq C_P H \|\nabla v\|_{L^2(T)}, \tag{3.1}$$

where C_P depends on c_0, c_1, and d.

3.1.2. Nodes, edges, and their neighbors. Let $\mathcal{N}_H = \{x_1, x_2, \ldots, x_p\}$ be the collection of interior nodes, and let $\mathcal{E}_H = \{e_1, e_2, \ldots, e_q\}$ be the collection of edges except those fully on the boundary of Ω. An edge $e \in \mathcal{E}_H$ is defined such that there exists two different elements T_i, T_j with $e = \overline{T}_i \bigcap \overline{T}_j$ that has co-dimension 1 in $\mathbb{R}^d$. We will use $E_H = \bigcup_{e \in \mathcal{E}_H} e \subset \Omega$ to denote the edges as a whole set.

We use the symbol $\sim$ to describe the neighbourship between nodes, edges, and elements. More precisely, if we consider a node $x \in \mathcal{N}_H$, an edge $e \in \mathcal{E}_H$, and an element $T \in \mathcal{T}_H$, then (1) $x \sim e$ denotes $x \in e$, (2) $e \sim T$ denotes $e \subset \overline{T}$, and (3) $x \sim T$ denotes $x \in \overline{T}$. The relationship $\sim$ is symmetric.

We use $\mathrm{N}(\cdot, \cdot)$ to describe the union of neighbors as a set. For example, $\mathrm{N}(x, \mathcal{E}_H) = \bigcup\{e \in \mathcal{E}_H : e \sim x\} \subset E_H$, $\mathrm{N}(x, \mathcal{T}_H) = \bigcup\{T \in \mathcal{T}_H : T \sim x\} \subset \Omega$, and $\mathrm{N}(e, \mathcal{T}_H) = \bigcup\{T \in \mathcal{T}_H : T \sim e\} \subset \Omega$.

3.2. Decomposition of solution space. With the mesh structure defined, we now discuss the coarse-fine scale decomposition of the solution space. We first discuss decomposition in the local element T in subsection 3.2.1 and then the global decomposition in subsection 3.2.2.

3.2.1. Local decomposition. A crucial requirement for the decomposition to be well defined is that the mesh size is order $O(1/k)$; see Assumption 1. As we will see later, this bound on H ensures that local Helmholtz problems in each element have properties that are similar to those of elliptic problems; thus, techniques in elliptic equations can then be applied.

Assumption 1. The mesh size satisfies $H \leq A_{\min}^{1/2}/(\sqrt{2} C_P V_{\max} k)$, where C_P is the constant in Proposition 3.1.

Given Assumption 1, we decompose[2] u into two parts $u = u_T^{\mathsf{h}} + u_T^{\mathsf{b}}$ in each element $T \in \mathcal{T}_H$. The two components satisfy

[2]This decomposition is inspired by that in the elliptic case [13].

$$
\begin{cases} -\nabla\cdot(A\nabla u_T^{\mathsf{h}}) - k^2V^2u_T^{\mathsf{h}} = 0 \ \text{ in } T, \\ \qquad\qquad\qquad\qquad u_T^{\mathsf{h}} = u \ \text{ on } \partial T\setminus(\Gamma_N\cup\Gamma_R), \\ \qquad\qquad\qquad\qquad A\nabla u_T^{\mathsf{h}}\cdot\nu = T_k u_T^{\mathsf{h}} \ \text{ on } \partial T\cap(\Gamma_N\cup\Gamma_R), \end{cases}
\tag{3.2}
$$

$$
\begin{cases} -\nabla\cdot(A\nabla u_T^{\mathsf{b}}) - k^2V^2u_T^{\mathsf{b}} = f \ \text{ in } T, \\ \qquad\qquad\qquad\qquad u_T^{\mathsf{b}} = 0 \ \text{ on } \partial T\setminus(\Gamma_N\cup\Gamma_R), \\ \qquad\qquad\qquad\qquad A\nabla u_T^{\mathsf{b}}\cdot\nu = T_k u_T^{\mathsf{b}} \ \text{ on } \partial T\cap(\Gamma_N\cup\Gamma_R). \end{cases}
$$

In short, the part u_T^{h} incorporates the boundary value of u, while u_T^{b} contains information of the right-hand side. Both equations in (3.2) should be understood in the standard weak sense using the following local sesquilinear form $a_T(\cdot,\cdot)$ in T:

$$
a_T(v,w) := (A\nabla v,\nabla w)_T - k^2(V^2v,w)_T - (T_kv,w)_{\partial T\cap(\Gamma_N\cup\Gamma_R)} \ \text{ for } v,w\in\mathcal{H}(T),
\tag{3.3}
$$

where $\mathcal{H}(T) := \mathcal{H}(\Omega)|_T$, the restriction of $\mathcal{H}(\Omega)$ in the domain T. The well-posedness of the two problems is due to the following proposition.

PROPOSITION 3.2. *Under Assumption* 1*, for $v\in\mathcal{H}(T)$ that vanishes on one of the edges of T, the corresponding sesquilinear form is coercive such that*

$$
\operatorname{Re} a_T(v,v) \geq \frac{1}{2}\|A^{1/2}\nabla v\|^2_{L^2(T)}.
$$

Proof. Using the Poincaré inequality (3.1) and Assumption 1, we get

$$
\begin{aligned}
\operatorname{Re} a_T(v,v) &= \|A^{1/2}\nabla v\|^2_{L^2(T)} - \|kVv\|^2_{L^2(T)} \\
&\geq (1 - C_P^2H^2k^2V_{\max}^2A_{\min}^{-1})\|A^{1/2}\nabla v\|^2_{L^2(T)} \geq \frac{1}{2}\|A^{1/2}\nabla v\|^2_{L^2(T)}.
\end{aligned}
\tag{3.4}
$$

∎

Since both equations in (3.2) contain Dirichlet's boundary condition on at least one of the edges of T, the coercivity implied by Proposition 3.2 suffices for the well-posedness. Consequently, the solutions u_T^{h} and u_T^{b} are well defined.

Remark 3.3. An important property is that u_T^{h} is "left-orthogonal" to u_T^{b} in T with respect to the local sesquilinear form $a_T(\cdot,\cdot)$ in T, in the sense of $a_T(u_T^{\mathsf{h}},u_T^{\mathsf{b}}) = 0$, according to the weak form of the equation. Note that we might not have $a_T(u_T^{\mathsf{b}},u_T^{\mathsf{h}}) = 0$ for T near the boundary (i.e., $\partial T\cap(\Gamma_N\cup\Gamma_R)\neq\emptyset$) due to the fact that $a_T(\cdot,\cdot)$ is not Hermitian here.

3.2.2. Global decomposition. In this subsection, we define a global decomposition $u = u^{\mathsf{b}} + u^{\mathsf{h}}$ such that for each T, it holds that $u^{\mathsf{h}}(x) = u_T^{\mathsf{h}}(x)$ and $u^{\mathsf{b}}(x) = u_T^{\mathsf{b}}(x)$ when $x\in T$. Both u^{h} and u^{b} are well defined and belong to $\mathcal{H}(\Omega)$ due to the continuity across edges. Here, the component u_T^{h} (resp., u^{h}) is called the local (resp., global) *Helmholtz-harmonic part* and u_T^{b} (resp., u^{b}) is the local (resp., global) *bubble part* of the solution u.

We further introduce the function space for the Helmholtz-harmonic part

$$
\begin{aligned}
V^{\mathsf{h}} := \{v\in\mathcal{H}(\Omega) : &-\nabla\cdot(A\nabla v) - k^2V^2v = 0 \text{ in each } T\in\mathcal{T}_H, \\
&A\nabla v\cdot\nu = T_kv, \text{ on } \Gamma_N\cup\Gamma_R\},
\end{aligned}
\tag{3.5}
$$

so that $u^{\mathsf{h}}\in V^{\mathsf{h}}$, and the space for the bubble part

$$
V^{\mathsf{b}} := \{v\in\mathcal{H}(\Omega) : v = 0 \text{ on } E_H\},
\tag{3.6}
$$

such that $u^{\mathsf{b}} \in V^{\mathsf{b}}$. In this way, the solution space of (1.1) can be decomposed to $V^{\mathsf{h}} + V^{\mathsf{b}}$. Furthermore, for any $v \in V^{\mathsf{h}}$ and $w \in V^{\mathsf{b}}$, it holds that $a(v, w) = 0$ by summing up local sesquilinear forms $a_T(\cdot, \cdot)$ and using Remark 3.3.

We will treat V^{b} as the fine scale or microscopic space and refer to V^{h} as the coarse scale or macroscopic space. The idea of our multiscale framework is to compute the two parts separately by exploring their own structures.

In the next two subsections, we will study the computational properties of $u^{\mathsf{h}} \in V^{\mathsf{h}}$ and $u^{\mathsf{b}} \in V^{\mathsf{b}}$, respectively. These properties serve as the cornerstone of designing our multiscale algorithm.

3.3. Local and small bubble part. In this subsection, we analyze the bubble part u^{b}. This part depends locally on f in each T. Thus, it can be computed efficiently in a parallel manner. Moreover, it is small and can be ignored if the target accuracy is $O(H)$; see Proposition 3.4.

PROPOSITION 3.4. *Under Assumption* 1*, it holds that*

$$\left\|u^{\mathsf{b}}\right\|_{\mathcal{H}(\Omega)} \leq \frac{3C_P}{A_{\min}^{1/2}} H \|f\|_{L^2(\Omega)}. \tag{3.7}$$

Proof. By definition, inside each patch T, it holds that $a_T(u^{\mathsf{b}}, u^{\mathsf{b}}) = (f, u^{\mathsf{b}})_T$. The coercivity estimate in (3.4) implies the inequality $\|kVu^{\mathsf{b}}\|^2_{L^2(T)} \leq \frac{1}{2}\|A^{1/2}\nabla u^{\mathsf{b}}\|^2_{L^2(T)}$. Using the estimate, we get

$$\begin{aligned}\operatorname{Re} a_T(u^{\mathsf{b}}, u^{\mathsf{b}}) &= \|A^{1/2}\nabla u^{\mathsf{b}}\|^2_{L^2(T)} - \|kVu^{\mathsf{b}}\|^2_{L^2(T)} \\ &\geq \frac{1}{3}(\|A^{1/2}\nabla u^{\mathsf{b}}\|^2_{L^2(T)} + \|kVu^{\mathsf{b}}\|^2_{L^2(T)}) = \frac{1}{3}\|u^{\mathsf{b}}\|^2_{\mathcal{H}(T)}.\end{aligned}$$

Combining the above estimate with the Cauchy–Schwarz inequality, we arrive at

$$\|u^{\mathsf{b}}\|^2_{\mathcal{H}(T)} \leq 3\operatorname{Re} a_T(u^{\mathsf{b}}, u^{\mathsf{b}}) = 3(f, u^{\mathsf{b}})_T \leq 3\|f\|_{L^2(T)}\|u^{\mathsf{b}}\|_{L^2(T)}.$$

Meanwhile, by the Poincaré inequality (3.1), we get

$$\|u^{\mathsf{b}}\|_{L^2(T)} \leq C_P H \|\nabla u^{\mathsf{b}}\|_{L^2(T)} \leq \frac{C_P H}{A_{\min}^{1/2}}\|u^{\mathsf{b}}\|_{\mathcal{H}(T)}.$$

Combining all the above inequalities gives $\|u^{\mathsf{b}}\|_{\mathcal{H}(T)} \leq 3(C_P H/A_{\min}^{1/2})\|f\|_{L^2(T)}$ for each element T. Summing them up for all elements T yields the desired conclusion. ∎

3.4. Low complexity of the Helmholtz-harmonic part. Now, we turn to the study of the Helmholtz-harmonic part u^{h}. The goal is to show that u^{h} can be approximated via local basis functions in an exponentially efficient manner. To achieve this, our approximation framework[3] contains three steps: (1) reducing the approximation of u^{h} to that of edge functions in subsection 3.4.1, (2) localizing the approximation to every single edge in subsection 3.4.2, and (3) realizing local approximation via oversampling and SVD in subsection 3.4.3. Combining all these three steps, we establish the low complexity in approximation of u^{h} in subsection 3.4.4.

3.4.1. Approximation via edge functions. We start with the first step of approximating u^{h}. By definition, u^{h} belongs to V^{h}. A key observation is that any function in V^{h} is determined entirely by its value on the edge set E_H. Thus, define

$$\tilde{V}^{\mathsf{h}} := \{\tilde{\psi} : E_H \to \mathbb{R}, \text{ there exists a function } \psi \in V^{\mathsf{h}}, \text{ such that } \tilde{\psi} = \psi|_{E_H}\};$$

[3]It is similar to that in our previous work for elliptic equations [13].

then, under Assumption 1, there is a one-to-one correspondence $\tilde{\psi} \in \tilde{V}^{\mathsf{h}} \leftrightarrow \psi \in V^{\mathsf{h}}$. More precisely, in each T, it holds that

$$\begin{cases} -\nabla \cdot (A\nabla\psi) - k^2 V^2 \psi = 0 & \text{in } T, \\ \psi = \tilde{\psi} & \text{on } \partial T \setminus (\Gamma_N \cup \Gamma_R), \\ A\nabla\psi \cdot \nu = T_k \psi & \text{on } \partial T \cap (\Gamma_N \cup \Gamma_R). \end{cases} \tag{3.8}$$

Indeed, we have $\tilde{V}^{\mathsf{h}} = H^{1/2}(E_H)$ by the trace theory since the local equation is elliptic. Using the above identification, approximating u^{h} corresponds to approximating $\tilde{u}^{\mathsf{h}}$, which is a function defined on edges and of lower complexity. We need to pay attention to the norm we use when approximating $\tilde{u}^{\mathsf{h}}$ so that we can use the error bound of the approximation to control the error of u^{h} in the energy norm. This will be the focus of the next section.

Remark 3.5. In the remaining part of the article, we will frequently use the correspondence between V^{h} and $\tilde{V}^{\mathsf{h}}$. Conventionally, when we write a tilde on the top of a function in V^{h}, it refers to its corresponding part in $\tilde{V}^{\mathsf{h}}$.

3.4.2. Localization of approximation. We discuss how to approximate the edge function $\tilde{u}^{\mathsf{h}}$, whose domain is E_H, which is nonlocal. Since it is often preferable to have localized basis functions for approximation and numerical algorithms, our second step is to localize the task of approximating $\tilde{u}^{\mathsf{h}}$ to every single edge.

To achieve localization, we study the geometry of the edge set E_H first. Observing that different edges only communicate with each other along their shared nodes, we can use nodal interpolation to localize the approximation. More precisely, we proceed with the following steps:

1. Interpolation: For each node $x_i \in \mathcal{N}_H$, choose $\tilde{\psi}_i$ to be the piecewise linear tent function on E_H, satisfying $\tilde{\psi}_i(x_j) = \delta_{ij}$ for each $x_j \in \mathcal{N}_H$. This defines an interpolation operator for $v \in V^{\mathsf{h}} \cap C(\overline{\Omega})$:
$$I_H v := \sum_{x_i \in \mathcal{N}_H} v(x_i)\psi_i(x).$$
Note that $\psi_i(x)$ is the same as the basis function constructed via the MsFEM [33]. The interpolation residual $v - I_H v$ vanishes on each $x_i \in \mathcal{N}_H$. Set[4] $v = \tilde{u}^{\mathsf{h}}$ and let $I_H \tilde{u}^{\mathsf{h}}$ be one part of the approximation for $\tilde{u}^{\mathsf{h}}$. Then, it remains to approximate the residue $\tilde{u}^{\mathsf{h}} - I_H \tilde{u}^{\mathsf{h}}$.
2. Localization: We wish to explore the fact that $\tilde{u}^{\mathsf{h}} - I_H \tilde{u}^{\mathsf{h}}$ vanishes on nodes to localize the subsequent approximation task. To achieve so, define $R_e \tilde{u}^{\mathsf{h}} = P_e(\tilde{u}^{\mathsf{h}} - I_H \tilde{u}^{\mathsf{h}}) := (\tilde{u}^{\mathsf{h}} - I_H \tilde{u}^{\mathsf{h}})|_e$. The goal is to find some basis functions on each e to approximate $R_e \tilde{u}^{\mathsf{h}}$. To make this problem precise, we need to specify the function space of $R_e \tilde{u}^{\mathsf{h}}$ and the norm for approximation. It turns out that the natural function space $R_e \tilde{u}^{\mathsf{h}}$ is the Lions–Magenes space; see the following Proposition 3.6.

PROPOSITION 3.6. *Let $d = 2$. Suppose $f \in L^2(\Omega)$ and that (2.3) holds. For each $e \in \mathcal{E}_H$, it holds that $R_e \tilde{u}^{\mathsf{h}} \in H_{00}^{1/2}(e)$, the Lions–Magenes space, which contains functions $v \in H^{1/2}(e)$ such that*

$$\frac{v(x)}{\operatorname{dist}(x, \partial e)} \in L^2(e).$$

Here, $\operatorname{dist}(x, \partial e)$ *is the Euclidean distance from x to the boundary of e.*

[4]Note that we can apply I_H to $\tilde{u}^{\mathsf{h}}$ due to the C^α estimate of u in Proposition 2.1.

It might seem unclear at this stage why we should consider such a complicated function space. In fact, this is related to the zero extension of functions. According to Chapter 33 of [54], $H_{00}^{1/2}(e)$ can also be characterized as the space of functions in $H^{1/2}(e)$ such that their zero extensions to E_H are still in $H^{1/2}(E_H)$. This is the key and in fact the only property that we will use for $H_{00}^{1/2}(e)$. The zero extension allows us to connect local approximation and global approximation. In the following, we will not distinguish $\tilde{\psi} \in H_{00}^{1/2}(e)$ and its zero extension to E_H that belongs to $H^{1/2}(E_H)$.

For any function in $H_{00}^{1/2}(e)$, we define a norm to measure approximation accuracy.

DEFINITION 3.7. *Let $d = 2$. The $\mathcal{H}^{1/2}(e)$ norm of a function $\tilde{\psi} \in H_{00}^{1/2}(e)$ is defined as*

$$\|\tilde{\psi}\|^2_{\mathcal{H}^{1/2}(e)} := \int_\Omega A|\nabla \psi|^2 + k^2|V\psi|^2\,, \tag{3.9}$$

where we have used the one-to-one correspondence $\tilde{\psi} \in \tilde{V}^{\mathsf{h}} \leftrightarrow \psi \in V^{\mathsf{h}}$. Here, we identify $\tilde{\psi}$ as the zero extension of its value on the edge e to E_H.

The $\mathcal{H}^{1/2}(e)$ norm in Definition 3.7 is the natural one to consider here since eventually, we aim for approximation accuracy in the energy norm.

The following theorem is the cornerstone for the above localization strategy. It states that a local accuracy guarantee can be seamlessly coupled to form a global accuracy guarantee.

THEOREM 3.8 (global error estimate). *Let $d = 2$. Suppose, for each edge e, there exists an edge function $\tilde{v}_e \in H_{00}^{1/2}(e)$ that satisfies*

$$\|R_e\tilde{u}^{\mathsf{h}} - \tilde{v}_e\|_{\mathcal{H}^{1/2}(e)} \le \epsilon_e\,. \tag{3.10}$$

Let $v_e \in V^{\mathsf{h}}$ be the corresponding part of $\tilde{v}_e \in \tilde{V}^{\mathsf{h}}$. Then, it holds that

$$\|u^{\mathsf{h}} - I_H u^{\mathsf{h}} - \sum_{e\in\mathcal{E}_H} v_e\|^2_{\mathcal{H}(\Omega)} \le C_{\mathrm{mesh}} \sum_{e\in\mathcal{E}_H} \epsilon_e^2\,, \tag{3.11}$$

where C_{mesh} is a constant depending on the number of edges for the elements only, e.g., for quadrilateral mesh $C_{\mathrm{mesh}} = 4$.

Given this theorem, to approximate u^{h} it suffices to find local edge basis functions that satisfy (3.10) for some desired ϵ_e. This is a localized task for each e.

The proofs for Propositions 3.6 and Theorem 3.8 are similar to those in the setting of elliptic equations [13]. However, for completeness, we will also present them here in subsections 6.2 and 6.3.

3.4.3. Local approximation via oversampling. The last step of approximation is to find local edge basis functions for each e so that (3.10) is satisfied. In this subsection, we discuss how to achieve this via oversampling and SVD, which can yield exponentially decaying ϵ_e. The general idea is to explore the fact that for a coarse scale function, its behavior on e can be controlled very well by that in an oversampling domain due to the compactness property of the restriction operator.

More precisely, for a given edge e, consider an oversampling domain ω_e associated with the edge. In general, any domain containing e in the interior can serve as a candidate. Here, for simplicity of presentation and as an illustrative example, we set

$$\omega_e = \overline{\bigcup\{T \in \mathcal{T}_H : \overline{T} \cap e \neq \emptyset\}}\,. \tag{3.12}$$

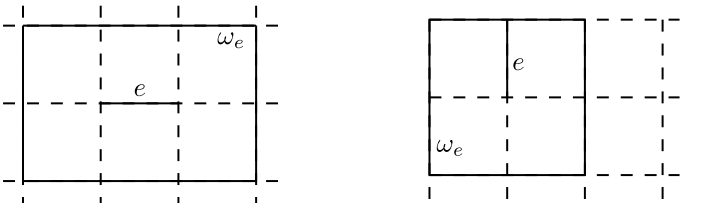

FIG. 2. *Illustration of oversampling domains. On the right, we use an edge connected to the upper boundary as an illustrating example.*

For interior edges and edges connected to the boundary, an illustration of this choice (3.12) for a quadrilateral mesh is given in Figure 2.

The key idea is to treat the residue $R_e\tilde{u}^{\mathsf{h}}$ as a restriction of a coarse scale function in ω_e and explore the compactness property of such restriction operators. By an abuse of notation via the correspondence of V^{h} and $\tilde{V}^{\mathsf{h}}$, for any $\mathcal{H}(\Omega)$ in V, we identify $R_e v$ as $R_e\tilde{v}^{\mathsf{h}}$. As a first step, we write

$$R_e\tilde{u}^{\mathsf{h}} = R_e u = R_e u^{\mathsf{h}}_{\omega_e} + R_e u^{\mathsf{b}}_{\omega_e}\,, \tag{3.13}$$

where we decompose u in ω_e into its coarse and fine scale components via (3.2) with T replaced by ω_e, and we shall use $u^{\mathsf{h}}_{\omega_e}$ and $u^{\mathsf{b}}_{\omega_e}$ to denote the corresponding local Helmholtz-harmonic and bubble part, respectively. Then, to approximate $R_e\tilde{u}^{\mathsf{h}}$, we could approximate the two terms in (3.13) separately. We will show that the first term can be approximated in an exponentially efficient manner due to a compactness property, and the second term can be computed locally and is very small.

Remark 3.9. One may ask whether the decomposition (3.13) in the oversampling domain is still well defined. Indeed, similar to (3.1), we have a uniform Poincaré inequality for every ω_e: for any edge e and $H^1(\omega_e)$ function v vanishing on any one of the edge boundaries of ω_e, it holds that

$$\|v\|_{L^2(\omega_e)} \le C'_P H \|\nabla v\|_{L^2(\omega_e)}\,, \tag{3.14}$$

where C'_P is a constant that only depends on c_0, c_1, d and our choice of oversampling domain. For the particular choice (3.12), $C'{}_P$ is a constant multiple of C_P; without loss of generality, we assume $C'{}_P \ge C_P$. Based on this observation, we will choose a small H so that Assumption 2 holds, which guarantees that local Helmholtz operators in the oversampling domain behave in a manner similar to that of elliptic case; this is similar to Proposition 3.2.

Assumption 2. The mesh size satisfies $H \le A_{\min}^{1/2}/(\sqrt{2}C'{}_P V_{\max} k)$, where C'_P is the constant in (3.14).

Note that Assumption 2 implies Assumption 1. Now, we discuss in detail how to deal with the two terms in (3.13).

1. For the first term, we consider the following function space in ω_e:

$$\begin{aligned} U(\omega_e) := \{v \in \mathcal{H}(\omega_e) : &-\nabla\cdot(A\nabla v) - k^2 V^2 v = 0 \quad \text{in } \omega_e, \\ &A\nabla v\cdot\nu = T_k v \quad \text{on } (\Gamma_N \cup \Gamma_R)\cap\partial\omega_e\}\,. \end{aligned} \tag{3.15}$$

Functions in this space are fully determined by their trace on $\partial\omega_e\backslash(\Gamma_N\cup\Gamma_R)$. By definition, $u^{\mathsf{h}}_{\omega_e}$ belongs to $U(\omega_e)$. Under Assumption 2, $(U(\omega_e), \|\cdot\|_{\mathcal{H}(\omega_e)})$

is a Hilbert space since the Helmholtz operator in ω_e is elliptic. Then, by abuse of notation, consider the operator

$$R_e : (U(\omega_e), \|\cdot\|_{\mathcal{H}(\omega_e)}) \to (H_{00}^{1/2}(e), \|\cdot\|_{\mathcal{H}^{1/2}(e)})$$

such that $R_e v = P_e(v - I_H v)$ for $v \in U(\omega_e)$. A critical property is that the singular values of R_e decay nearly exponentially fast; see Theorem 3.10. Its proof is deferred to subsection 6.4.

THEOREM 3.10. *Let $d = 2$. Under Assumption 2, the operator R_e is compact for each $e \in \mathcal{E}_H$. Denote the pairs of its left singular vectors and singular values by $\{\tilde{v}_{m,e}, \lambda_{m,e}\}_{m\in\mathbb{N}}$, where $\tilde{v}_{m,e} \in H_{00}^{1/2}(e)$ and the sequence $\{\lambda_{m,e}\}_{m\in\mathbb{N}}$ is in a descending order. Then, for any $\epsilon > 0$, it holds that*

$$\lambda_{m,e} \leq C_\epsilon \exp\left(-m^{(\frac{1}{d+1}-\epsilon)}\right), \tag{3.16}$$

where C_ϵ is a constant that is independent of k, H and may depend on ϵ, d, and the mesh parameters c_0, c_1.

Remark 3.11. As we can see from the proof, we actually show that (3.16) still holds by setting C_ϵ to be 1 and requiring for $m > N_\epsilon$ with N_ϵ depending on k and H. But we can also make the above inequality hold for all m by introducing the constant C_ϵ.

We discuss the implication of this theorem. By definition of singular values, if we set $W_{m,e} = \operatorname{span}\{\tilde{v}_{j,e}\}_{j=1}^{m-1}$, then Theorem 3.10 implies that

$$\min_{\tilde{v}_e \in W_{m,e}} \|R_e v - \tilde{v}_e\|_{\mathcal{H}^{1/2}(e)} \leq C_\epsilon \exp\left(-m^{(\frac{1}{d+1}-\epsilon)}\right) \|v\|_{\mathcal{H}(\omega_e)}. \tag{3.17}$$

Applying this result to $v = u^{\mathsf{h}}_{\omega_e} \in U(\omega_e)$ leads to

$$\min_{\tilde{v}_e \in W_{m,e}} \|R_e u^{\mathsf{h}}_{\omega_e} - \tilde{v}_e\|_{\mathcal{H}^{1/2}(e)} \leq C_\epsilon \exp\left(-m^{(\frac{1}{d+1}-\epsilon)}\right) \|u^{\mathsf{h}}_{\omega_e}\|_{\mathcal{H}(\omega_e)}. \tag{3.18}$$

Thus, there is a nearly exponential efficiency in approximating the first term $R_e u^{\mathsf{h}}_{\omega_e}$.

2. For the second term in (3.13), the oversampling bubble part $u^{\mathsf{b}}_{\omega_e}$ can be efficiently computed by solving local Helmholtz problems. Moreover, under Assumption 2, this term is small in the $\mathcal{H}(\Omega)$ norm, as shown in the following proposition.

PROPOSITION 3.12. *Under Assumption 2, for each $e \in \mathcal{E}_H$, the following estimate holds for the oversampling bubble part:*

$$\|R_e u^{\mathsf{b}}_{\omega_e}\|_{\mathcal{H}^{1/2}(e)} \leq CH\|f\|_{L^2(\omega_e)},$$

where C is a constant independent of k and H.

The proof is deferred to subsection 6.5.

We further define a special Helmholtz-harmonic function $u^{\mathsf{s}} \in V^{\mathsf{h}}$ such that that its restriction on each edge $e \in E_H$ equals $R_e u^{\mathsf{b}}_{\omega_e}$. Namely, this special Helmholtz-harmonic function accounts for the second term in (3.13) for each edge. By the previous proposition, we immediately have the estimate

$$\|u^{\mathsf{s}}\|_{\mathcal{H}(\Omega)} \leq CH\|f\|_{L^2(\Omega)},$$

where C is a constant independent of k and H. Along with Proposition 3.4, we conclude that there is a constant C_s independent of k and H such that

$$\|u^{\mathsf{s}}\|_{\mathcal{H}(\Omega)} + \|u^{\mathsf{b}}\|_{\mathcal{H}(\Omega)} \leq C_s H \|f\|_{L^2(\Omega)}\,. \tag{3.19}$$

Now, consider the following space of basis functions:

$$\tilde{V}^{(1)}_{H,m,e} := W_{m,e}\,.$$

In practice, this space can be computed locally by an SVD of R_e. Due to (3.13) and (3.18), we have the following error estimate on each e:

$$\min_{\tilde{v}_e \in \tilde{V}^{(1)}_{H,m,e}} \|R_e u^{\mathsf{h}} - u^{\mathsf{s}} - \tilde{v}_e\|_{\mathcal{H}^{1/2}(e)} \leq C_\epsilon \exp\left(-m^{(\frac{1}{d+1}-\epsilon)}\right) \|u^{\mathsf{h}}_{\omega_e}\|_{\mathcal{H}(\omega_e)}\,. \tag{3.20}$$

Remark 3.13. The operator R_e involves nodal interpolation, which is in general not stable for H^1 functions if the dimension is greater than 1. However, in Theorem 3.10, we take the domain of the operator to be $U(\omega_e)$, which contains Helmholtz-harmonic functions that are Hölder continuous due to the standard C^α estimates for elliptic equations. More specifically, Lemma 6.2 implies the stability of R_e in this space.

Remark 3.14. If we follow the proof of Lemma 3.13 in [38], it is be possible to remove the small parameter ϵ in Theorem 3.10 to get a better asymptotic bound $O(\exp(-m^{\frac{1}{d+1}}))$.

3.4.4. Low complexity in approximation. Finally, define the collection of edge basis functions

$$\tilde{V}^{(1)}_{H,m} = \text{span}\left\{\bigcup_e \tilde{V}^{(1)}_{H,m,e}\right\},$$

and denote by $\tilde{V}^{(0)}_H$ the span of the nodal interpolation basis used earlier, i.e., $\tilde{V}^{(0)}_H := \text{span}\,\{\tilde{\psi}_i\}$. Define the overall edge approximation $\tilde{V}_{H,m} = \text{span}\,\{\tilde{V}^{(0)}_H \bigcup \tilde{V}^{(1)}_{H,m}\}$. Let $V_{H,m} \subset V^{\mathsf{h}}$ be the corresponding part of $\tilde{V}_{H,m} \subset \tilde{V}^{\mathsf{h}}$, via (3.8). Then, using (3.20) and Theorem 3.8, we get a nearly exponentially decaying error estimate for approximating u^{h}; see Theorem 3.15.

THEOREM 3.15. *Let $d = 2$. Under Assumption* 2 *and* (2.3)*, it holds that*

$$\min_{v \in V_{H,m}} \|u^{\mathsf{h}} - u^{\mathsf{s}} - v\|_{\mathcal{H}(\Omega)} \leq C_d (C_{\text{stab}}(k) + H) \exp\left(-m^{(\frac{1}{d+1}-\epsilon)}\right) \|f\|_{L^2(\Omega)}\,,$$

where C_d is a generic constant independent of k, m, H.

Proof. By Theorem 3.10 and the global error estimate in Theorem 3.8, we get

$$\min_{v \in V_{H,m}} \|u^{\mathsf{h}} - u^{\mathsf{s}} - v\|^2_{\mathcal{H}(\Omega)} \leq C_{\text{mesh}} C_\epsilon^2 \exp\left(-2m^{(\frac{1}{d+1}-\epsilon)}\right) \sum_{e \in \mathcal{E}_H} \|u^{\mathsf{h}}_{\omega_e}\|^2_{\mathcal{H}(\omega_e)}\,. \tag{3.21}$$

Due to Assumption 2, we have the elliptic estimate for the oversampling bubble part

$$\|u^{\mathsf{b}}_{\omega_e}\|_{\mathcal{H}(\omega_e)} \leq \frac{3C'_P}{A^{1/2}_{\min}} H \|f\|_{L^2(\omega_e)}\,. \tag{3.22}$$

This is similar to Proposition 3.4, which is a consequence of Assumption 1. Then, using $u^{\mathsf{h}}_{\omega_e} = u - u^{\mathsf{b}}_{\omega_e}$, it follows that

$$\|u^{\mathsf{h}}_{\omega_e}\|^2_{\mathcal{H}(\omega_e)} \leq 2(\|u\|^2_{\mathcal{H}(\omega_e)} + \|u^{\mathsf{b}}_{\omega_e}\|^2_{\mathcal{H}(\omega_e)}) \leq \frac{18{C'}_P^2}{A_{\min}} H^2 \|f\|^2_{L^2(\omega_e)} + 2\|u\|^2_{\mathcal{H}(\omega_e)}\,. \tag{3.23}$$

Note that by our choice of oversampling domains, every element T can only be covered by $\{\omega_e\}_{e\in\mathcal{E}_H}$ at most C_1 times for a fixed C_1. Therefore, it holds that

$$\sum_{e\in\mathcal{E}_H} \|f\|^2_{L^2(\omega_e)} \leq C_1 \|f\|^2_{L^2(\Omega)}\,, \tag{3.24}$$

as well as

$$\sum_{e\in\mathcal{E}_H} \|u\|^2_{\mathcal{H}(\omega_e)} \leq C_1 \|u\|^2_{\mathcal{H}(\Omega)} \leq C_1 C^2_{\mathrm{stab}}(k) \|f\|^2_{L^2(\Omega)}\,, \tag{3.25}$$

where the last inequality is due to the a priori estimate (2.3). Combining (3.21), (3.23), (3.24), and (3.25) completes the proof. ∎

Clearly, Theorem 3.15 implies the low complexity property of the part $u^{\mathsf{h}} - u^{\mathsf{s}}$. Each edge contains at most m basis functions, so the space $V_{H,m}$ is of dimension $O(m/H^d)$, while the approximation accuracy is of order $\exp(-m^{(\frac{1}{d+1}-\epsilon)})$. We will use the space $V_{H,m}$ in our multiscale framework for approximating $u^{\mathsf{h}} - u^{\mathsf{s}}$.

Remark 3.16. $V_{H,m}$ does not depend on the right-hand side f or the solution u. Therefore, we can use the same $V_{H,m}$ for different right-hand sides.

4. The multiscale methods. In this section, we discuss the multiscale methods for solving (1.1), based on the coarse-fine scale decomposition established in the last section.

By the nature of a multiscale algorithm, we will handle the "coarse part" $u^{\mathsf{h}} - u^{\mathsf{s}}$ and the "fine part" $u^{\mathsf{b}} + u^{\mathsf{s}}$ separately. Conceptually, the locality and small magnitude of $u^{\mathsf{b}} + u^{\mathsf{s}}$ imply that it can be computed efficiently or ignored without affecting the accuracy much, and the low complexity of $u^{\mathsf{h}} - u^{\mathsf{s}}$ indicates that we can use a Galerkin method with a small number of basis functions to approximate it accurately.

In subsection 4.1, we outline our general multiscale computational framework. Depending on how the trial and test spaces in the Galerkin method are selected, we get two categories of algorithms, namely, the Ritz–Galerkin approach and Petrov–Galerkin approach, that we will make precise in subsections 4.2 and 4.3, respectively.

4.1. The multiscale framework. The bubble part u^{b} and the special function u^{s} are first computed locally. Given these parts, we form an effective equation for $u^{\mathsf{h}} - u^{\mathsf{s}}$ as

$$a(u^{\mathsf{h}} - u^{\mathsf{s}}, v) = (f, v)_\Omega - a(u^{\mathsf{b}} + u^{\mathsf{s}}, v) \tag{4.1}$$

for any $v \in \mathcal{H}(\Omega)$.

Remark 4.1. The right-hand side in (4.1) can be seen as a bounded linear functional on $v \in \mathcal{H}(\Omega)$. By the estimate in (2.4), this equation for $u^{\mathsf{h}} - u^{\mathsf{s}}$ (given fixed $u^{\mathsf{b}} + u^{\mathsf{s}}$) is well-posed.

Numerically, we solve (4.1) for $u^{\mathsf{h}} - u^{\mathsf{s}}$ using a Galerkin method. That is, we choose a trial space S and a test space S_{test} to find a numerical solution $u_S \in S$ that satisfies

$$a(u_S, v) = (f, v)_\Omega - a(u^{\mathsf{b}} + u^{\mathsf{s}}, v) \tag{4.2}$$

for any $v \in S_{\text{test}}$. If $S_{\text{test}} = S$, then it is called a Ritz–Galerkin method; otherwise, it is a Petrov–Galerkin method. Here, since the equation is formulated in the complex domain, we specifically refer to the choice $S_{\text{test}} = \overline{S}$ as the Petrov–Galerkin method.

In subsection 4.2, we formulate our Ritz–Galerkin method and present theories for the well-posedness of the discrete problem, as well as the error estimate in both the energy norm and the L^2 norm. In subsection 4.3, we discuss the Petrov–Galerkin method, which is more straightforward and appears more convenient in practical computation.

4.2. The Ritz–Galerkin method. First, we establish a general strategy for analyzing the Ritz–Galerkin method in solving (4.1). We start with a definition of the approximation accuracy of S.

DEFINITION 4.2. *For $S \subset V^{\mathsf{h}}$, the approximation accuracy of S is defined as*

$$\eta(S) := \sup_{f \in L^2(\Omega)\setminus\{0\}} \inf_{v \in S} \frac{\|u - v\|_{\mathcal{H}(\Omega)}}{\|f\|_{L^2(\Omega)}}, \tag{4.3}$$

where u and f are related via the Helmholtz equation in (1.1).

For the Ritz–Galerkin method, it turns out that $\eta(S)$ is critical in analyzing the solution errors of u_S.

THEOREM 4.3. *Suppose* (2.3) *holds and that $k\eta(S) \le 1/(4C_cV_{\max})$ as well as $\overline{S} = S$. Then, the following statements hold for the Ritz–Galerkin method:*

1. *The Galerkin solution u_S is a quasi-optimal approximation in the sense that*

$$\|u^{\mathsf{h}} - u^{\mathsf{s}} - u_S\|_{\mathcal{H}(\Omega)} \le 2C_c \inf_{v \in S} \|u^{\mathsf{h}} - u^{\mathsf{s}} - v\|_{\mathcal{H}(\Omega)},$$
$$\|u^{\mathsf{h}} - u^{\mathsf{s}} - u_S\|_{L^2(\Omega)} \le C_c\eta(S)\|u^{\mathsf{h}} - u^{\mathsf{s}} - u_S\|_{\mathcal{H}(\Omega)}.$$

2. *If we further assume $Hk \le 1/(8C_sC_cV_{\max})$ for constant C_s defined in* (3.19)*, the discrete problem satisfies the discrete inf-sup stability condition:*

$$\inf_{v \in S} \sup_{v' \in S\setminus\{0\}} \frac{|a(v, v')|}{\|v\|_{\mathcal{H}(\Omega)}\|v'\|_{\mathcal{H}(\Omega)}} \ge \frac{1}{4 + 3C_c^{-1} + 8kV_{\max}C_{\text{stab}}(k)}.$$

The proof of this theorem is deferred to subsection 6.6. It is inspired by the standard Gårding-type inequality for a posteriori estimates; see, for example, [40]. However, our proofs are slightly different since only the part $u^{\mathsf{h}} - u^{\mathsf{s}}$ is approximated via the basis functions.

The above theorem implies that once $\eta(S)$ is small, the discrete problem is well-posed, and the Galerkin solution approximates the exact solution accurately.

Given Theorem 4.3, we can choose $S = V_{H,m} + \overline{V_{H,m}}$ where $V_{H,m}$ is defined in Theorem 3.15 independent of the right-hand side. For the quantity $\eta(S)$, we have the following estimate using its subspace $V_{H,m}$:

$$\eta(S) \le \eta(V_{H,m}) \le \max(C_d, C_s)\left((C_{\text{stab}}(k) + H)\exp\left(-m^{(\frac{1}{d+1}-\epsilon)}\right) + H\right). \tag{4.4}$$

Here, we have used (3.19) for the small parts u^{b} and u^{s} of size $O(H)$ and Theorem 3.15 for the approximation error for $u^{\mathsf{h}} - u^{\mathsf{s}}$. Invoking Theorems 4.3 and 3.15, we get the following error analysis for the Galerkin solution.

THEOREM 4.4. *Let $d=2$. Suppose Assumption* 2 *and* (2.3) *hold, and*

$$\max(C_d, C_s)k\left((C_{\text{stab}}(k)+H)\exp\left(-m^{(\frac{1}{d+1}-\epsilon)}\right)+H\right) \le 1/(4C_cV_{\max}),$$

where C_s, C_d are generic constants defined in (3.19) *and Theorem* 3.15*, respectively. Then, using $S = V_{H,m} + \overline{V_{H,m}}$ in the Ritz–Galerkin method leads to a solution u_S that satisfies*

$$\|u^{\mathsf{h}} - u^{\mathsf{s}} - u_S\|_{\mathcal{H}(\Omega)} \le 2C_cC_d(C_{\text{stab}}(k)+H)\exp\left(-m^{(\frac{1}{d+1}-\epsilon)}\right)\|f\|_{L^2(\Omega)}. \tag{4.5}$$

For the ϵ that satisfies $\frac{1}{d+1} - \epsilon = \frac{1}{d+2}$, we can take $m \sim O(\log^{d+2}(kC_{\text{stab}}(k))$. Then, the condition in Theorem 4.4 holds, provided that the mesh size H satisfies the following Assumption 3.

Assumption 3. The mesh size satisfies $H \le 1/(8\max(C_d, C_s)C_cV_{\max}k)$.

Furthermore, if $C_{\text{stab}}(k) \le C(1+k^\gamma)$ for some constants γ and C, then the condition $m \sim O(\log^{d+2}(kC_{\text{stab}}(k))$ reduces to $m \sim \log^{d+2}(k)$. This implies that once m is moderately large, i.e., logarithmic in k, the nearly exponential convergence of the Galerkin solution shown in Theorem 4.4 will become effective. As in Remark 3.14, we can improve the index $d+2$ to $d+1$.

We provide several additional remarks of the Ritz–Galerkin method below.

Remark 4.5. In the Ritz–Galerkin method, the trial and test spaces are $S = V_{H,m} + \overline{V_{H,m}}$. One can intuitively understand that $V_{H,m}$ is needed to represent the desired solution, and $\overline{V_{H,m}}$ is used for the approximation of the adjoint problem, which is required in the numerical analysis of the Helmholtz equation. There can be a lot of overlap between $V_{H,m}$ and $\overline{V_{H,m}}$: on each interior edge, since the singular vectors of R_e are real, these edge basis functions are real-valued. Thus, $V_{H,m}$ and $\overline{V_{H,m}}$ can only differ on the edges connected to the boundary, where the presence of the Robin boundary condition makes the operator non-Hermitian.

Remark 4.6. Combining (4.5) with the local computation of the fine parts will yield the overall error estimate for u, which is nearly exponentially convergent.

4.3. The Petrov–Galerkin method. In this subsection, we introduce the Petrov–Galerkin method. We choose $S = V_{H,m}$ and $S_{\text{test}} = \overline{V_{H,m}}$. We give the following remarks on this method.

Remark 4.7. The trial and test spaces in the Petrov–Galerkin method often have smaller dimensions than their Ritz–Galerkin counterpart since we do not put the complex conjugate $\overline{V_{H,m}}$ in S. This can save computational effort.

Remark 4.8. Our current theory does not address the stability of the discrete system and the $\mathcal{H}(\Omega)$ error estimate for the Petrov–Galerkin method. This is left for our future work. We note that our numerical experiments in the next section imply that these properties also hold for the Petrov–Galerkin method.

5. Numerical experiments. In this section, we will outline and discuss our numerical algorithms in detail based on the established theoretical analysis. Several Helmholtz equations are solved using our algorithm, which confirm our theoretical results. We also consider some examples in which our theoretical assumptions are not satisfied. Even for these examples, our methods still give a nearly exponential rate of convergence. This provides further evidence for the robustness of our methods.

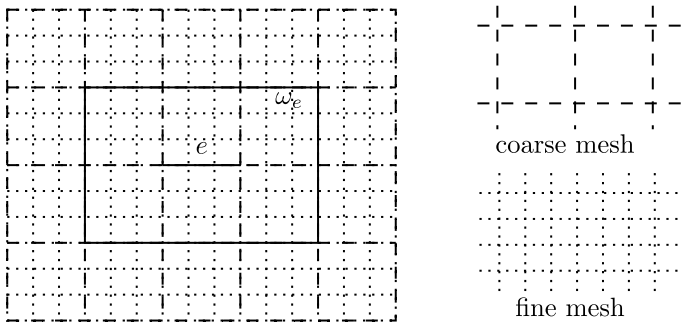

FIG. 3. *Two-level mesh: a fraction.*

5.1. Set-up. We consider the domain $\Omega = [0,1] \times [0,1]$ and discretize it by a uniform two-level quadrilateral mesh; see a fraction of this mesh in Figure 3, where we also show an edge e and its oversampling domain ω_e in solid lines. The coarse and fine mesh sizes are denoted by H and h, respectively.

For a given Helmholtz equation, we compute the reference solution u_{ref} using the classical FEM on the fine mesh; with a sufficiently small h, it is reasonable to treat u_{ref} as the ground truth u. We remark that via a posteriori estimates, we can check that the fine mesh indeed resolve the corresponding problems; thus, the associated fine mesh solutions could serve as good reference solutions for all of our numerical examples. To be precise, we check that the relative error between the solutions using fine mesh of size h and $h/2$ are small such that it is of order 10^{-2} in the energy norm and 10^{-4} in the L^2 norm.

The accuracy of a numerical solution u_{sol} is computed by comparing it with the reference solution u_{ref} on the fine mesh. The accuracy will be measured both in the L^2 norm and the energy norm:

$$e_{L^2} = \frac{\|u_{\text{ref}} - u_{\text{sol}}\|_{L^2(\Omega)}}{\|u_{\text{ref}}\|_{L^2(\Omega)}}, \qquad e_{\mathcal{H}} = \frac{\|u_{\text{ref}} - u_{\text{sol}}\|_{\mathcal{H}(\Omega)}}{\|u_{\text{ref}}\|_{\mathcal{H}(\Omega)}}. \tag{5.1}$$

5.2. Multiscale algorithms. We outline our numerical algorithms for obtaining u_{sol}. There are offline and online stages, depending on whether the steps involve the information of the right-hand side.

5.2.1. Offline stage. For each edge $e \in \mathcal{E}_H$ and its associated oversampling domain ω_e, the key step in the offline stage is to construct the discretized version of the operator

$$R_e : (U(\omega_e), \|\cdot\|_{\mathcal{H}(\omega_e)}) \to (H_{00}^{1/2}(e), \|\cdot\|_{\mathcal{H}^{1/2}(e)}),$$

which is defined by $R_e v = (v - I_H v)|_e$. Here, $U(\omega_e)$ is defined in (3.15), $\|\cdot\|_{\mathcal{H}(\omega_e)}$ is the energy norm in ω_e while $H_{00}^{1/2}(e)$ is the Lions–Magenes space, and $\|\cdot\|_{\mathcal{H}^{1/2}(e)}$ is defined in (3.9).

We note that functions in $U(\omega_e)$ are fully determined by their traces on $\partial\omega_e \backslash (\Gamma_N \cup \Gamma_R)$. Thus, we can take the discretized matrix version of R_e as a linear mapping from Dirichlet's data on $\partial\omega_e \backslash (\Gamma_N \cup \Gamma_R)$ to the image of R_e, which contains functions on the edge e. The discretization of the $\|\cdot\|_{\mathcal{H}(\omega_e)}$ and $\|\cdot\|_{\mathcal{H}^{1/2}(e)}$ norms leads to positive

definite matrices on the discretized domains $\partial\omega_e\backslash(\Gamma_N\cup\Gamma_R)$ and e. To obtain these positive definite matrices, we construct the Helmholtz-harmonic extension operators both on e and $\partial\omega_e\backslash(\Gamma_N\cup\Gamma_R)$, which maps boundary data to the Hemholtz-harmonic function in the domain. Based on this operator, we can calculate the energy norms of the extended Hemholtz-harmonic function. This leads to the required norms as well as the positive definite matrices defining these norms.[5]

With the discretized matrices constructed, the next step is to compute the top m left singular vectors of R_e for some selected $m\in\mathbb{N}$. This SVD problem turns out to be a generalized eigenvalue problem for these discrete matrices. For each e, denote the singular vectors by $\tilde{v}_{1,e},\dots,\tilde{v}_{m,e}\in H_{00}^{1/2}(e)$. Their Helmholtz-harmonic extensions to the domain are denoted by $v_{1,e},\dots,v_{m,e}\in\mathcal{H}(\Omega)$, obtained via the correspondence (3.8). The basis function space formed by the collection of all $v_{j,e}, 1\le j\le m$ and $e\in\mathcal{E}_H$, together with the interpolation part $\{\psi_i\}_{x_i\in\mathcal{N}_H}$, are denoted by $V_{H,m}$ and will constitute the Galerkin basis as defined in subsection 3.4.4. Note that here, $\{\psi_i\}_{x_i\in\mathcal{N}_H}$ are the same as the basis functions in the MsFEM.

We are now in a position to construct our Galerkin basis and the associated stiffness matrix. The construction depends on how to choose the trial and test spaces in the Galerkin method. We will outline two possible choices below:

- Ritz–Galerkin: $S=V_{H,m}+\overline{V_{H,m}}$ and $S_{\text{test}}=S$.
- Petrov–Galerkin: $S=V_{H,m}$ and $S_{\text{test}}=\overline{V_{H,m}}$.

5.2.2. Online stage. In the online stage, we solve the coarse and fine scales separately. First, we solve for u^{b} and u^{s}, and then we use the effective equation (4.1) to solve for $u^{\mathsf{h}}-u^{\mathsf{s}}$.

For the bubble part u^{b}, we solve the local Hemholtz problem in each element $T\in\mathcal{T}_H$, which leads to u_T^{b} defined in (3.2). Gluing them together leads to u^{b}.

For u^{s}, on each $e\in\mathcal{E}_H$ and ω_e, we construct the oversampling bubble part $u_{\omega_e}^{\mathsf{b}}$ via solving a local Helmholtz equation. Then, we get an edge function $R_e u_{\omega_e}^{\mathsf{b}}$ for each edge. We solve locally the Helmholtz-harmonic extension of these edge functions and add them together to obtain u^{s}.

Now, we can form the right-hand-side vector in our effective equation (4.1) and use the offline-assembled stiffness matrix to obtain the Galerkin solution for the part $u^{\mathsf{h}}-u^{\mathsf{s}}$.

This construction yields a practical numerical algorithm that efficiently handles multiple right-hand sides.

We note that all the above algorithms consider a uniform number of basis functions, namely, m, for each edge $e\in\mathcal{E}_H$. It is also possible to make this number vary with edges and thus fully adaptive to the problem's local properties such as the approach in [32]. Consequently, this will lead to an adaptive algorithm where the truncated singular values serve as local error indicators. We do not pursue this in detail here and will leave this to our future work.

In the following, we will test our algorithms for different model problems. Our general set-up is to fix a reasonable coarse scale H and then study how the errors behave as m changes for the two choices outlined above.

Remark 5.1. Our numerical experience implies that in the Ritz–Galerkin method, one does not need to add the conjugate space $\overline{V_{H,m}}$ into S while still obtaining an exponential rate of convergence.

[5]See also the implementation in subsection 4.2 of [13] on how these matrices are constructed for elliptic problems.

5.3. A high wavenumber example: Planar wave. We start with an example of planar wave where the coefficients are constant and the wavenumber is high. More precisely, we set $A = V = \beta = 1$ and $f = 0$. The wavenumber $k = 2^7$. We take the exact solution to be

$$u(x_1, x_2) = \exp(-ik(0.6x_1 + 0.8x_2)).$$

Using this solution, we are able to specify the Robin boundary condition on $\partial\Omega$. Note that this is an inhomogeneous boundary condition, so it is beyond our previous discussion. In this case, the inhomogeneous data are incorporated to the equation of the bubble part u^{b}, while the treatment for the Helmholtz-harmonic part remains the same as that in the homogeneous case. To be specific, now our decomposition on each element T is $u = u_T^{\mathsf{h}} + u_T^{\mathsf{b}} + u_T^{\mathsf{p}}$, where u_T^{p} stands for a particular solution. The part $u_T^{\mathsf{b}} + u_T^{\mathsf{p}}$ satisfies

$$\begin{aligned} -\nabla\cdot(A\nabla(u_T^{\mathsf{b}} + u_T^{\mathsf{p}})) - k^2V^2(u_T^{\mathsf{b}} + u_T^{\mathsf{p}}) &= f \quad \text{in } T, \\ u_T^{\mathsf{b}} + u_T^{\mathsf{p}} &= 0 \quad \text{on } \partial T \setminus (\Gamma_N \cup \Gamma_R), \\ A\nabla(u_T^{\mathsf{b}} + u_T^{\mathsf{p}})\cdot\nu &= T_k(u_T^{\mathsf{b}} + u_T^{\mathsf{p}}) + g \quad \text{on } \partial T \cap (\Gamma_N \cup \Gamma_R). \end{aligned}$$

We will use $u^{\mathsf{b}} + u^{\mathsf{p}}$ to replace u^{b} on the right-hand side of the effective equation for Galerkin solution (4.1). Similarly, when we compute the special Helmholtz-harmonic function u^{s} to account for the oversampling bubble part, its restriction on each edge equals $R_e(u^{\mathsf{b}}_{\omega_e} + u^{\mathsf{p}}_{\omega_e})$ instead of $R_e u^{\mathsf{b}}_{\omega_e}$. In this way we can take care of the boundary data via local particular problems and still obtain the desired accuracy. The error analysis in such case remains the same once we replace u_T^{b} in the homogeneous data case by $u_T^{\mathsf{b}} + u_T^{\mathsf{p}}$; in the bound, we will also have the norm of g.

We set the fine mesh $h = 2^{-10}$ and coarse mesh $H = 2^{-5}$. We vary the number of edge basis functions in each $e \in \mathcal{E}_H$, choosing $m = 1, 2, \ldots, 7$ and implementing the two algorithms outlined in subsection 5.2.2. The results are shown in Figure 4. We observe that the online basis approaches achieve nearly exponential decaying errors with respect to m. The difference between the Ritz–Galerkin and Petrov–Galerkin approaches is almost negligible. We can see that a few bases per edge suffice for very high accuracy.

Furthermore, we make some comparison between our edge coupling approach (the Ritz–Galerkin version) and the PUM approach reported in [36]. We adopt the same

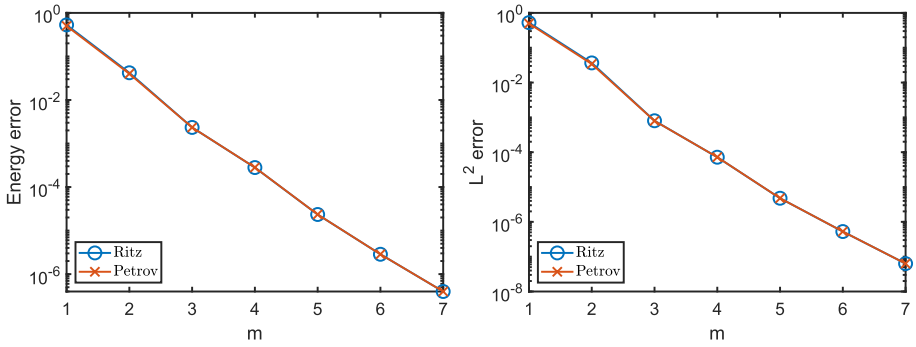

FIG. 4. *Numerical results for the high wavenumber example. Left: $e_{\mathcal{H}}$ versus m; right: e_{L^2} versus m.*

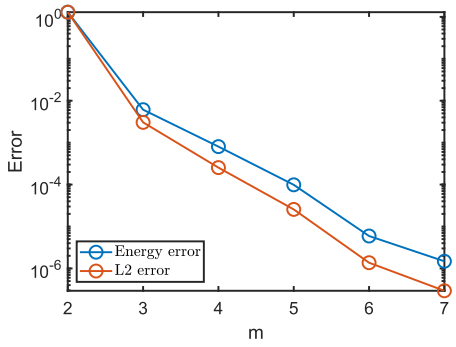

FIG. 5. *Numerical results for the high wavenumber example with* $k = 100$, $H = 1/20$, $h = 1/1000$.

setting there with $k = 100$, $H = 1/20$, $h = 1/1000$ and vary the number of edge basis functions in each $e \in \mathcal{E}_H$, choosing $m = 2, 3, \ldots, 7$. We present the results in Figure 5. We see that both errors decay very fast, and in particular, the error in our method for $m = 7$ is smaller than the error in [36], with oversampling ratio $H_*/H = 2$ and 35 local bases per patch. With the same wavenumber and number of coarse patches, our method uses a slightly larger oversampling domain while reducing the number of multiscale bases by a factor of around $35/(2 \times 7) = 2.5$. Here, we have used the fact that the number of edges is twice as many as domains in two dimensions. Nevertheless, the support of basis functions in our approach and PUM approach could be different by a factor of two, and the size of the overlapped domain decomposition in the PUM approach could also influence the result, leading to additional complexities for comparison. More detailed numerical study of the two approaches could be of future interest.

5.4. A high contrast example: Mie resonances. In this example, we consider an $A(x)$ with high contrast channels. More precisely, define the domain

$$\Omega_\varepsilon = (0.25, 0.75)^2 \cap \bigcup_{j \in \mathbb{Z}^2} \varepsilon \left(j + (0.25, 0.75)^2 \right), \tag{5.2}$$

and the coefficient is defined as

$$A(x) = \begin{cases} 1, & x \notin \Omega_\varepsilon \\ \varepsilon^2, & x \in \Omega_\varepsilon . \end{cases}$$

Here, ε is a parameter controlling the contrast. We choose $\varepsilon = 2^{-4}$ and visualize $\log_{10} A(x)$ in the left plot of Figure 6.

We take $\beta = 1, V = 1, k = 9$. For such a choice of k, the model exhibits an unusual behavior induced by Mie resonances in the small inclusions; see [44, 49]. An accurate numerical solution for this model would be hard to compute, and it serves as a proper benchmark for our method. The right-hand side is

$$f(x_1, x_2) = \begin{cases} 10000 \exp\left(-\dfrac{1}{1 - 400 \times \operatorname{dist}(x, z)^2} \right), & \operatorname{dist}(x, z)^2 < \dfrac{1}{400} \\ 0 \quad \text{otherwise}, \end{cases}$$

where $z = (0.125, 0.5)$ and $\operatorname{dist}(x, z)^2 = (x_1 - 0.125)^2 + (x_2 - 0.5)^2$. We impose the homogeneous Robin boundary condition on $\partial\Omega$. We take the fine mesh $h = 2^{-9}$ and

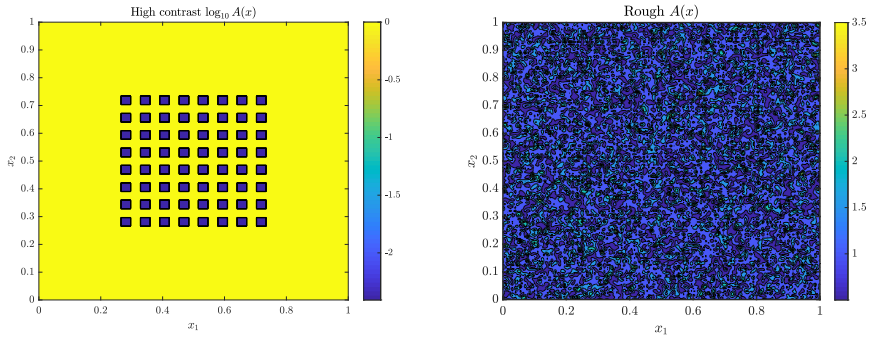

FIG. 6. *Left: the contour of* $\log_{10} A$ *for the high contrast example; right: the contour of* A *for the rough media example.*

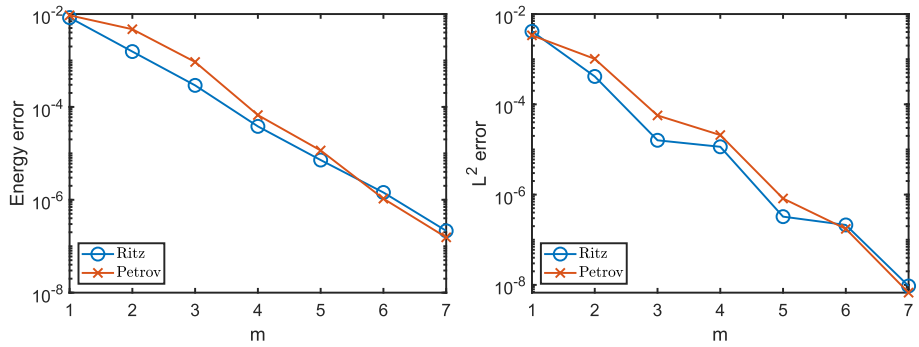

FIG. 7. *Numerical results for the high contrast example. Left:* $e_{\mathcal{H}}$ *versus* m*; right:* e_{L^2} *versus* m.

the coarse mesh $H = 2^{-5}$. As before, we take $m = 1, 2, \ldots, 7$, and the numerical results are shown in Figure 7. A nearly exponential rate of convergence is observed consistently, and in this particular example, the Ritz method slightly outperforms the Petrov method.

5.5. A numerical example with mixed boundary and rough field. In the last example, we consider a mixed boundary problem. We impose the homogeneous Dirichlet boundary condition on $(x_1, 0), x_1 \in [0, 1]$, the homogeneous Neumann boundary condition on $(x_1, 1), x_1 \in [0, 1]$, and the homogeneous Robin boundary condition on the other two parts of $\partial\Omega$. We choose $A(x)$ to be a realization of some random field; more precisely,

$$A(x) = |\xi(x)| + 0.5\,, \tag{5.3}$$

where the field $\xi(x)$ satisfies

$$\xi(x) = a_{11}\xi_{i,j} + a_{21}\xi_{i+1,j} + a_{12}\xi_{i,j+1} + a_{22}\xi_{i+1,j+1}, \text{ if } x \in \left[\frac{i}{2^7}, \frac{i+1}{2^7}\right) \times \left[\frac{j}{2^7}, \frac{j+1}{2^7}\right).$$

Here, $\{\xi_{i,j}, 0 \le i, j \le 2^7\}$ are independent and identically distributed unit Gaussian random variables. In addition, $a_{11} = (i + 1 - 2^7 x_1)(j + 1 - 2^7 x_2)$, $a_{21} = (2^7 x_1 - i)(j + 1 - 2^7 x_2)$, $a_{12} = (i + 1 - 2^7 x_1)(2^7 x_2 - j)$, $a_{22} = (2^7 x_1 - i)(2^7 x_2 - j)$ are interpolating coefficients to make $\xi(x)$ piecewise linear. A sample from this field is displayed in the right plot of Figure 6.

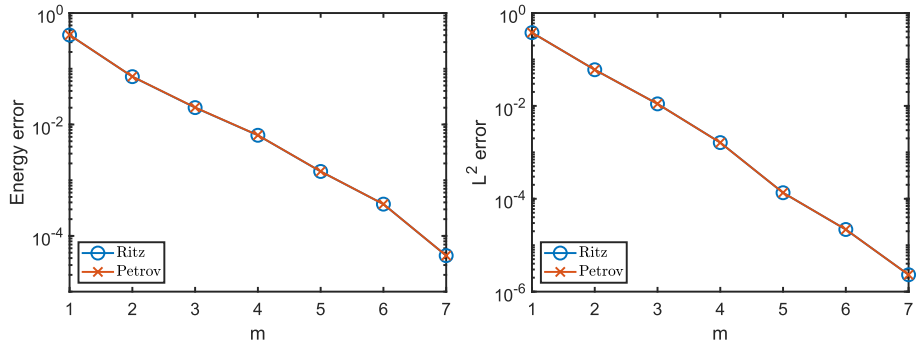

FIG. 8. *Numerical results for the mixed boundary and rough field example. Left: $e_{\mathcal{H}}$ versus m; right: e_{L^2} versus m.*

Moreover, we also take $V(x)$ and $\beta(x)$ as independent samples drawn from this random field. We choose the wavenumber $k = 2^5$, the right-hand side $f(x_1, x_2) = x_1^4 - x_2^3 + 1$, the fine mesh $h = 2^{-10}$, and the coarse mesh $H = 2^{-5}$. Again, we take $m = 1, 2, \ldots, 7$ and present the numerical results in Figure 8. A nearly exponential rate of convergence is still observed for this challenging example. The difference between the Ritz–Galerkin method and the Petrov–Galerkin method is very mild.

It is worth noting that this example is constructed artificially, mixing different kinds of boundary conditions and rough coefficients without taking into account the analytical properties of this combination. Thus, the numerical results for this example demonstrate the effectiveness of our multiscale methods in a more general setting. Moreover, our right-hand side f is global, so most oversampling bubble parts would be nonzero.

5.6. Summary. We summarize what we have observed in these numerical examples. Both algorithms lead to a nearly exponential rate of convergence with respect to m, and we are able to use the offline-computed Galerkin basis to solve for multiple right-hand sides.

Moreover, it is observed that the difference between the Ritz–Galerkin and the Petrov–Galerkin approaches is very mild in most cases, but sometimes the Ritz–Galerkin method can have better performance. Therefore, we recommend using the Ritz–Galerkin approach in practice.

6. Proofs. This section presents the theoretical proofs in this paper. Some proofs are similar to those in the elliptic case. We will refer these proofs to the corresponding proofs in the elliptic case [13], while we will make relevant remarks on possible changes and modifications.

6.1. Proof of Proposition 2.1. In this subsection, we provide the proof of the qualitative version of the C^α estimate. It is a direct application of related results for elliptic equations.

Proof. We note that the Helmholtz PDE (1.1) is equivalent to

$$\begin{cases} -\nabla \cdot (A\nabla u) = f + k^2 V^2 u & \text{in } \Omega, \\ u = 0 \quad \text{on } \Gamma_D, \\ A\nabla u \cdot \nu = T_k u \quad \text{on } \Gamma_N \cup \Gamma_R. \end{cases} \tag{6.1}$$

Since $f \in L^2(\Omega)$, we know by the a priori estimate of the Helmholtz equation that $u \in H^1(\Omega)$. Therefore, we can regard (6.1) as an elliptic PDE with k^2V^2u known as a part of the right-hand side. This PDE has its right-hand side in $L^2(\Omega)$ and has u as its solution. We can invoke the result in Remark 6.5 of [27], which concludes that u lies in some Hölder space $C^\alpha(\Omega)$ such that

$$\|u\|_{C^\alpha(\Omega)} \le C(\|f\|_{L^2(\Omega)} + k^2\|u\|_{L^2(\Omega)})$$

for some Hölder exponent $\alpha \in (0,1)$ and C. ∎

6.2. Proof of Proposition 3.6. The proof relies on the fact that any function v on e belonging to $H^{1/2}(e) \bigcap C^\alpha(e)$ and vanishing at ∂e will be in the space $H^{1/2}_{00}(e)$; see Proposition 2.1 in [13] for detailed arguments of this fact. Then, $R_e\tilde{u}^{\mathsf{h}} \in H^{1/2}(e) \bigcap C^\alpha(e)$ and vanishes at ∂e, so it belongs to $H^{1/2}_{00}(e)$.

6.3. Proof of Theorem 3.8. We decompose the energy norm into the contribution from each element $T \in \mathcal{T}_H$:

$$\|u^{\mathsf{h}} - I_H u^{\mathsf{h}} - \sum_{e\in\mathcal{E}_H} v_e\|^2_{\mathcal{H}(\Omega)} = \sum_{T\in\mathcal{T}_H} \|u^{\mathsf{h}} - I_H u^{\mathsf{h}} - \sum_{e\sim T} v_e\|^2_{\mathcal{H}(T)},$$

where we have used the fact that $v_e = 0$ in T if e and T are not neighbors.

Let us fix an element T. For each $e \sim T$, the trace of the function $u^{\mathsf{h}} - I_H u^{\mathsf{h}} - \sum_{e\in T} v_e$ on e is $\tilde{u}^{\mathsf{h}} - I_H\tilde{u}^{\mathsf{h}} - \tilde{v}_e \in H^{1/2}_{00}(e)$. We can extend this trace to $\partial T\backslash e$ by 0 to get an $H^{1/2}(\partial T)$ boundary data. Then, this boundary data can be used to define a Helmholtz-harmonic function in T via the correspondence (3.8). Using the triangle inequality and the Cauchy–Schwarz inequality, we get

$$\|u^{\mathsf{h}} - I_H u^{\mathsf{h}} - \sum_{e\sim T} v_e\|^2_{\mathcal{H}(T)} \le C_{\mathrm{mesh}} \sum_{e\sim T} \|P_e(\tilde{u}^{\mathsf{h}} - I_H\tilde{u}^{\mathsf{h}}) - \tilde{v}_e\|^2_{\mathcal{H}^{1/2}_T(e)},$$

where the $\mathcal{H}^{1/2}_T(e)$ norm of a function $\tilde{\psi} \in H^{1/2}_{00}(e)$ is defined as

$$\|\tilde{\psi}\|^2_{\mathcal{H}^{1/2}_T(e)} := \int_T A|\nabla\psi|^2 + k^2|V\psi|^2. \tag{6.2}$$

The constant C_{mesh} depends on the mesh type only; for example, $C_{\mathrm{mesh}} = 4$ for the quadrilateral mesh and $C_{\mathrm{mesh}} = 3$ for the triangular mesh. Then, we sum the above inequality over all $T \in \mathcal{T}_H$, which yields

$$\begin{aligned}\|u^{\mathsf{h}} - I_H u^{\mathsf{h}} - \sum_{e\in\mathcal{E}_H} v_e\|^2_{\mathcal{H}(\Omega)} &\le C_{\mathrm{mesh}} \sum_{T\in\mathcal{T}_H}\sum_{e\sim T} \|P_e(\tilde{u}^{\mathsf{h}} - I_H\tilde{u}^{\mathsf{h}}) - \tilde{v}_e\|^2_{\mathcal{H}^{1/2}_T(e)} \\ &= C_{\mathrm{mesh}} \sum_{e\in\mathcal{E}_H} \|P_e(\tilde{u}^{\mathsf{h}} - I_H\tilde{u}^{\mathsf{h}}) - \tilde{v}_e\|^2_{\mathcal{H}^{1/2}(e)} \\ &\le C_{\mathrm{mesh}} \sum_{e\in\mathcal{E}_H} \epsilon_e^2. \end{aligned} \tag{6.3}$$

The proof is completed.

6.4. Proof of Theorem 3.10. This is the key theorem underlying the exponential convergence for approximating u^{h}. To prove it, we need to analyze the spectrum of the operator R_e for each edge e. The treatments for interior edges and edges connected to the boundary are slightly different due to the different boundary conditions involved. We will explain the proof for interior edges in detail and comment on the changes needed to be made for edges connected to the boundary.

Since this theorem is stated for all edges, we start by discussing some geometric relations that hold uniformly for all interior edges.

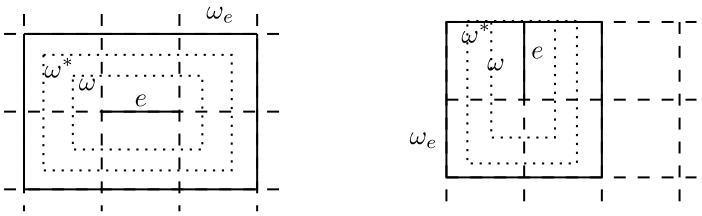

FIG. 9. *Geometric relation* $e \subset \omega \subset \omega^* \subset \omega_e$.

6.4.1. Geometric relation: Interior edges. Suppose that e is an interior edge so that e lies strictly in the interior domain of ω_e; see Figure 2. We describe some geometric relation[6] between e and ω_e that will be needed in our analysis. Figure 9 illustrates our ideas for a uniform quadrilateral mesh. For each interior edge e, there exist two concentric rectangles $\omega \subset \omega^*$ with the center being the midpoint m_e of e such that $e \subset \omega \subset \omega^* \subset \omega_e$; the center m_e is the center of gravity of ω and ω^*. We require $\omega^* \cap \partial\Omega = \emptyset$. Moreover, one side of ω and ω^* should be parallel to e. We introduce three parameters l_1, l_2, l_3 to specify and describe the geometry:

1. With respect to the center m_e, the two rectangles ω and ω^* are scaling equivalent such that there exists $l_1 > 1$, $\omega^* - m_e = l_1 \cdot (\omega - m_e)$. Here, we use the notation that $t \cdot X := \{tx : x \in X\}$ for a set X and a scalar t. For our choice of ω_e, the parameter l_1 can be selected to only depend on c_0 and c_1 in subsection 3.1.1.
2. The ratio of ω's larger side length over the smaller side length is bounded by a uniform constant $l_2 > 1$ that depends on c_0 and c_1 only.
3. There is a constant $l_3 > 1$ depending on c_0 and c_1 only such that $l_3 \cdot e \subset \omega$.

We note that l_1, l_2, l_3 are universal constants for all interior edges. All three parameters depend on c_0, c_1 only. We introduce these parameters in order to get a uniform treatment for every interior edge. Indeed, several constants in our estimates depend on l_1, l_2, l_3, but not on k and H, uniformly for all interior edges.

6.4.2. Main idea of the proof. In the following, we explain the main ideas of our proof. Recall the target is to show the left singular values of R_e decay nearly exponentially fast. Similar to the rationale behind (3.17), it suffices to show that there exists an $m-1$ dimensional space $W_{m,e} \subset H_{00}^{1/2}(e)$ such that

$$\min_{\tilde{v}_e \in W_{m,e}} \|R_e v - \tilde{v}_e\|_{\mathcal{H}^{1/2}(e)} \leq C_\epsilon \exp\left(-m^{(\frac{1}{d+1}-\epsilon)}\right) \|v\|_{\mathcal{H}(\omega_e)} \tag{6.4}$$

for any $v \in U(\omega_e)$. We also use $U(\omega')$ to denote the function space in ω' defined via (3.15), with ω_e replaced by any ω'. Our proof contains two main steps, summarized in the following two lemmas.

LEMMA 6.1. *For $d > 0$ and any $v \in U(\omega^*)$, there exists an $m-1$ dimensional space $\Phi_{m,e} \subset U(\omega)$ such that*

$$\min_{\chi \in \Phi_{m,e}} \|v - \chi\|_{\mathcal{H}(\omega)} \leq C_\epsilon \exp\left(-m^{(\frac{1}{d+1}-\epsilon)}\right) \|v\|_{\mathcal{H}(\omega^*)} \tag{6.5}$$

for some C_ϵ independent of k and H.

[6] It is similar to that in subsection 3.3.1 of [13].

LEMMA 6.2. *For $d=2$ and any $v \in H^1(\omega)$ and $\nabla \cdot (A\nabla v) \in L^2(\omega)$, it holds that*

$$\|R_e v\|_{\mathcal{H}^{1/2}(e)} \leq C \left(\|v\|_{\mathcal{H}(\omega)} + H \|\nabla \cdot (A\nabla v) + k^2 V^2 v\|_{L^2(\omega)} \right) \tag{6.6}$$

for some C independent of k and H.

Remark 6.3. Here in Lemma 6.1 and 6.2, the constants are independent of k because in the local domain, the operator behaves similarly to an elliptic operator. Moreover, for edges connected to the boundary that we will discuss in subsection 6.4.7, the boundary condition is of order 1 after rescaling due to the assumption on our mesh size $Hk \lesssim 1$. Thus, eventually no k-dependence will be involved for our estimates in the local domain. This is different from the global C^α regularity estimate in the proof of Proposition 2.1. See subsection 6.4.6 for details, where we only use the C^α estimate of an elliptic equation.

We will defer the proofs of the two lemmas to subsections 6.4.3 and 6.4.6 and describe how to prove Theorem 3.10 using them here.

Proof of Theorem 3.10. From the above discussion, it remains to show (6.4). For $v \in \mathcal{H}(\omega_e)$, we have $v \in \mathcal{H}(\omega^*)$ and $\|v\|_{\mathcal{H}(\omega^*)} \leq \|v\|_{\mathcal{H}(\omega_e)}$. By Lemma 6.1, we get

$$\min_{\chi \in \Phi_{m,e}} \|v - \chi\|_{\mathcal{H}(\omega)} \leq C_\epsilon \exp\left(-m^{(\frac{1}{d+1} - \epsilon)} \right) \|v\|_{\mathcal{H}(\omega_e)}.$$

Now, since $v - \chi$ satisfies the condition in Lemma 6.2 and v and χ both vanish under the operator $v \to \nabla \cdot (A\nabla v) + k^2 V^2 v$, we obtain

$$\min_{\chi \in \Phi_{m,e}} \|R_e v - R_e \chi\|_{\mathcal{H}^{1/2}(e)} \leq C C_\epsilon \exp\left(-m^{(\frac{1}{d+1} - \epsilon)} \right) \|v\|_{\mathcal{H}(\omega_e)}.$$

Thus, taking $W_{m,e} = R_e \Phi_{m,e}$ completes the proof. ∎

6.4.3. Proof of Lemma 6.1. The proof of this lemma is inspired by Theorem 3.3 in [2], which states a similar result but for elliptic equations only. We generalize it here for the Helmholtz equation.

First, by our geometric construction, $\omega^* - m_e = l_1 \cdot (\omega - m_e)$. We denote a sequence of domains $\omega = \omega_0 \subset \omega_1 \subset \cdots \subset \omega_{N-1} \subset \omega_N = \omega^*$ such that they are concentric and that $\omega_j - m_e = (1+t)(\omega_{j-1} - m_e)$ for $j = 1, 2, \ldots, N$. Here, $t = l_1^{1/N} - 1$. Then, there are two important lemmas, whose proofs are presented in subsections 6.4.4 and 6.4.5.

LEMMA 6.4. *For each $0 \leq j \leq N$ and any $n \in \mathbb{N}$, there is an n-dimensional space $W_n(\omega_j) \subset U(\omega_j)$ such that for all $v \in U(\omega_j)$, it holds that*

$$\inf_{w \in W_n(\omega_j)} \|v - w\|_{L^2(\omega_j)} \leq C H n^{-1/d} \|v\|_{\mathcal{H}(\omega_j)}, \tag{6.7}$$

where C is a generic constant independent of k, H, t, and n.

LEMMA 6.5. *For each $1 \leq j \leq N$ and every $v \in U(\omega_j)$, it holds that*

$$\|v\|_{\mathcal{H}(\omega_{j-1})} \leq C/(tH) \|v\|_{L^2(\omega_j)}, \tag{6.8}$$

where C is a generic constant independent of k, H, and t.

With the two lemmas, we are ready to prove Lemma 6.1.

Proof of Lemma 6.1. Choose $n = \lfloor m/N \rfloor$. The proof relies on an iteration argument. We start from $j = N$. By (6.7) and (6.8), we get an n-dimensional space

$W_n(\omega_N) \subset U(\omega_N)$ and a function $w_N \in W_n(\omega_N)$ such that

$$\|v - w_N\|_{\mathcal{H}(\omega_{N-1})} \leq C/(tH)\|v - w_N\|_{L^2(\omega_N)} \leq Ct^{-1}n^{-1/d}\|v\|_{\mathcal{H}(\omega_N)},$$

where we have used the fact that the infimum in (6.7) is attained since it is a finite dimensional optimization problem. Here, by abuse of notation, the value of the constant C varies in different places. It is a generic constant independent of k, H, t, and n.

Now, we iterate the above process. The function $v - w_N \in U(\omega_{N-1})$, so again by (6.7) and (6.8), we get an n-dimensional space $W_n(\omega_{N-1}) \subset U(\omega_{N-1})$ and a function $w_{N-1} \in W_n(\omega_{N-1})$ such that

$$\|v - w_N - w_{N-1}\|_{\mathcal{H}(\omega_{N-2})} \leq Ct^{-1}n^{-1/d}\|v - w_N\|_{\mathcal{H}(\omega_{N-1})} \leq (Ct^{-1}n^{-1/d})^2\|v\|_{\mathcal{H}(\omega_N)}.$$

Repeating the above procedure, we get

$$\left\| v - \sum_{j=1}^{N} w_j \right\|_{\mathcal{H}(\omega)} \leq (Ct^{-1}n^{-1/d})^N \|v\|_{\mathcal{H}(\omega^*)},$$

where each $w_j \in U(\omega_j) \subset U(\omega_0) = U(\omega)$. Therefore, there exists an $m \geq nN$ dimensional space $\Phi_{m,e} \subset U(\omega)$ such that

$$\inf_{w \in \Phi_{m,e}} \|v - w\|_{\mathcal{H}(\omega)} \leq (Ct^{-1}n^{-1/d})^N \|v\|_{\mathcal{H}(\omega^*)}.$$

For a parameter q to be determined later, choose $N = \left\lfloor m^{\frac{q}{q+1}} \right\rfloor$; then we obtain

$$(Ct^{-1}n^{-1/d})^N \leq \left(Ct^{-1}\left(\frac{m}{N}\right)^{-1/d} \right)^N = \exp\left(N\left(\frac{1}{d}\log\left(\frac{N}{m}\right) + \log C - \log t\right)\right). \tag{6.9}$$

Using $N \leq m^{\frac{q}{q+1}}$ and $t = l_1^{1/N} - 1 = \exp(\frac{1}{N}\log l_1) - 1 \geq \frac{1}{N}\log l_1 \geq m^{-\frac{q}{q+1}}\log l_1$, we can bound the right-hand side of (6.9) as

$$\begin{aligned}(Ct^{-1}n^{-1/d})^N &\leq \exp\left(-m^{\frac{q}{q+1}}\left(\left(\frac{1}{d} - q\right)\frac{1}{q+1}\log m - \log C + \log\log l_1\right)\right) \\ &\leq C_q \exp\left(-m^{\frac{q}{q+1}}\right)\end{aligned} \tag{6.10}$$

for some constant C_q that depends on q, d, C, l_1 if $q < 1/d$. Here, in the last inequality, we used the fact that when $q < 1/d$, there exists an M_q such that if $m \geq M_q$, then

$$\left(\frac{1}{d} - q\right)\frac{1}{q+1}\log m - \log C + \log\log l_1 \geq 1,$$

and thus, $(Ct^{-1}n^{-1/d})^N \leq \exp\left(-m^{\frac{q}{q+1}}\right)$ for $m \geq M_q$. By choosing

$$C_q = \max_{1 \leq m < M_q} \exp\left(-m^{\frac{q}{q+1}}\left(\left(\frac{1}{d} - q\right)\frac{1}{q+1}\log m - \log C + \log\log l_1\right)\right)\exp(m^{\frac{q}{q+1}}) + 1,$$

we can prove that (6.10) is valid.

Now, we choose $q < 1/d$ and denote $\frac{q}{q+1} = \frac{1}{d+1} - \epsilon$ for some $\epsilon > 0$. There is a one-to-one correspondence between q and small positive ϵ, so we can also write the error estimate in (6.10) in terms of ϵ as

$$(Ct^{-1}n^{-1/d})^N \leq C_\epsilon \exp\left(-m^{\frac{1}{d+1}-\epsilon}\right).$$

This completes the proof. ∎

6.4.4. Proof of Lemma 6.4. First, using the spectrum of the Laplacian operator with Neumann's boundary condition, there exists an n-dimensional space $S_n \subset H^1(\omega_j)$ such that for any $v \in H^1(\omega_j)$,

$$\inf_{w\in S_n} \|v-w\|_{L^2(\omega_j)} \le CHn^{-1/d}\|v\|_{H^1(\omega_j)} \le CHn^{-1/d}\|v\|_{\mathcal{H}(\omega_j)}\,, \tag{6.11}$$

where C is a generic constant independent of k, H, t, and n. Equivalently, this implies the identity embedding operator $Q : (\mathcal{H}(\omega_j), \|\cdot\|_{\mathcal{H}(\omega_j)}) \to (L^2(\omega_j), \|\cdot\|_{L^2(\omega_j)})$ such that $Qv = v$ is compact and the nth largest left singular value $\mu_n \le CHn^{-1/d}$.

Now, since $U(\omega_j)$ is a closed subspace of $(\mathcal{H}(\omega_j), \|\cdot\|_{\mathcal{H}(\omega_j)})$, we can view Q as an operator from $(U(\omega_j), \|\cdot\|_{\mathcal{H}(\omega_j)})$ to $(L^2(\omega_j), \|\cdot\|_{L^2(\omega_j)})$. Denote its singular values in a nonincreasing order by $\{\mu_n{}'\}$. Using the max-min theorem for singular values, we obtain

$$\begin{aligned}\mu_n' &= \max_{S_n\subset U(\omega_j),\dim(S_n)=n}\ \min_{v\in S_n,\|v\|_{\mathcal{H}(\omega_j)}=1} \|Qv\|_{L^2(\omega_j)}\\ &\le \max_{S_n\subset \mathcal{H}(\omega_j),\dim(S_n)=n}\ \min_{v\in S_n,\|v\|_{\mathcal{H}(\omega_j)}=1} \|Qv\|_{L^2(\omega_j)} = \mu_n\,.\end{aligned}$$

Thus, $\mu_n' \le CHn^{-1/d}$. Therefore, there is an n-dimensional space $W_n(\omega_j) \subset U(\omega_j)$ such that for all $v \in U(\omega_j)$, it holds that

$$\inf_{w\in S_n} \|v-w\|_{L^2(\omega_j)} \le CHn^{-1/d}\|v\|_{H^1(\omega_j)} \le CHn^{-1/d}\|v\|_{\mathcal{H}(\omega_j)}\,.$$

The proof is completed.

6.4.5. Proof of Lemma 6.5. We introduce a cutoff function $\eta \in C^1(\omega_j)$ such that $0 \le \eta \le 1$, and $\eta = 1$ in ω_{j-1}, as well as $|\nabla\eta(x)| \le C/(tH)$ for some constant C independent of k, H, and t.

For any $v \in U(\omega_j)$, we use the test function $\eta^2 v$ and the weak form to get

$$(A\nabla v, \nabla(\eta^2 v))_{\omega_j} - k^2(Vv, V\eta^2 v)_{\omega_j} = 0\,, \tag{6.12}$$

where we have used the definition of $U(\omega_j)$ (see the beginning of subsection 6.4.2) and the property of our construction that $\partial\omega_j \cap (\Gamma_N \cup \Gamma_R) = \emptyset$.

Using the relation $\|A^{1/2}\eta\nabla v\|^2_{L^2(\omega_j)} = (A\nabla v, \eta^2\nabla v)_{\omega_j}$ and the above formula, we obtain

$$\begin{aligned}\|A^{1/2}\eta\nabla v\|^2_{L^2(\omega_j)} &= -2(A^{1/2}\eta\nabla v, A^{1/2}v\nabla\eta)_{\omega_j} + k^2(Vv, V\eta^2 v)_{\omega_j}\\ &\le \frac{1}{2}\|A^{1/2}\eta\nabla v\|^2_{L^2(\omega_j)} + 2\|A^{1/2}v\nabla\eta\|^2_{L^2(\omega_j)} + k^2V_{\max}^2\|v\|^2_{L^2(\omega_j)}\,,\end{aligned} \tag{6.13}$$

which leads to $\|A^{1/2}\eta\nabla v\|^2_{L^2(\omega_j)} \le 4\|A^{1/2}v\nabla\eta\|^2_{L^2(\omega_j)} + 2k^2V_{\max}^2\|v\|^2_{L^2(\omega_j)}$. Therefore, using the fact that $\eta = 1$ in ω_{j-1}, we have

$$\begin{aligned}\|v\|^2_{\mathcal{H}(\omega_{j-1})} &\le \|A^{1/2}\eta\nabla v\|^2_{L^2(\omega_j)} + k^2V_{\max}^2\|v\|^2_{L^2(\omega_j)}\\ &\le 4\|A^{1/2}v\nabla\eta\|^2_{L^2(\omega_j)} + 3k^2V_{\max}^2\|v\|^2_{L^2(\omega_j)}\\ &\le \left(\frac{4C^2}{(tH)^2} + 3k^2V_{\max}^2\right)\|v\|^2_{L^2(\omega_j)}\\ &\le \frac{C'^2}{(tH)^2}\|v\|^2_{L^2(\omega_j)}\end{aligned} \tag{6.14}$$

for some C' independent of k, H, and t, where we have used Assumption 1 such that $kV_{\max}H \le C''$ for $C'' = A_{\min}^{1/2}/(\sqrt{2}C_P)$. This completes the proof.

6.4.6. Proof of Lemma 6.2. We use Lemma 3.10 of [13], which implies that

$$\|R_e v\|_{\mathcal{H}^{1/2}(e)} \leq C \left(\|A^{1/2} \nabla v\|_{L^2(\omega)} + H \|\nabla \cdot (A \nabla v)\|_{L^2(\omega)} \right) \tag{6.15}$$

for some C independent of k, H. Indeed, Lemma 3.10 of [13] implies that C can depend on the C^α estimate constant of v in ω. The discussion in Remark 6.3 implies that this constant is independent of k.

By a triangular inequality, we have

$$\begin{aligned} H\|\nabla \cdot (A\nabla v)\|_{L^2(\omega)} &\leq H\|k^2 V^2 v\|_{L^2(\omega)} + H\|\nabla \cdot (A\nabla v) + k^2 V^2 v\|_{L^2(\omega)} \\ &\leq C' \|kVv\|_{L^2(\omega)} + H\|\nabla \cdot (A\nabla v) + k^2 V^2 v\|_{L^2(\omega)}, \end{aligned}$$

where we have used Assumption 1 such that $kV_{\max} H \leq C'$ for $C' = A_{\min}^{1/2}/(\sqrt{2} C_P)$. Now, using the definition of the $\mathcal{H}(\omega)$ norm, we have

$$\|A^{1/2} \nabla v\|_{L^2(\omega)} + C' \|kVv\|_{L^2(\omega)} \leq C'' \|v\|_{\mathcal{H}(\omega)}$$

for some generic constant C'' that does not depend on anything else. Combining the above inequalities concludes the proof.

6.4.7. For edges connected to the boundary. The above proofs are for interior edges. For edges connected to the boundary, we need a different geometric relation, as depicted in the right of Figure 9. The quantitative characterization of this geometric relation is the same as that in subsection 3.3.2 of [13], which introduces three other parameters l_4, l_5, l_6 to describe the geometry associated with edges, similar to l_1, l_2, l_3 for interior edges.

The main idea of the proof for this case is the same as that for the interior edges. We need to prove Lemmas 6.1 and 6.2 for edges connected to the boundary. The proof of Lemma 6.2 remains nearly the same. A technical part is that the constant in the inequality depends on the local C^α estimate. According to the discussion in Remark 6.3, the local C^α constant is independent of k for edges connected to the boundary. To prove Lemma 6.1, we again use the same strategy in subsection 6.4.3, by establishing Lemmas 6.4 and 6.5 and then using an iteration argument. The iteration argument and the proof for Lemma 6.4 remain unchanged. For Lemma 6.5, the only slight change is (6.12), which becomes

$$(A\nabla v, \nabla(\eta^2 v))_{\omega_j} - k^2 (Vv, V\eta^2 v)_{\omega_j} = (T_k v, \eta^2 v)_{\partial\omega_j \cap (\Gamma_N \cup \Gamma_R)} \tag{6.16}$$

due to the boundary conditions involved. However, since $\mathrm{Re}\,(T_k v, \eta^2 v)_{\partial\omega_j \cap (\Gamma_N \cup \Gamma_R)} \leq 0$, the conclusion of Lemma 6.5 still holds.

Therefore, the result also holds for edges connected to the boundary.

Remark 6.6. We have assumed that Ω is a polygonal domain, so the shape of the local domains around the boundary is well behaved. In particular, a uniform Poincaré inequality will hold for these domains (in general, the constant in the Poincaré inequality depends on the shape of the domain). This guarantees that we can obtain a uniform constant in Theorem 3.10 for both interior edges and edges connected to the boundary.

6.5. Proof of Proposition 3.12. First, we have the bound on the oversampling bubble part in (3.22):

$$\|u_{\omega_e}^{\mathsf{b}}\|_{\mathcal{H}(\omega_e)} \leq \frac{3C_P'}{A_{\min}^{1/2}} H \|f\|_{L^2(\omega_e)}. \tag{6.17}$$

Applying Lemma 6.2 and the definition of $u^{\mathsf{b}}_{\omega_e}$ leads to

$$\begin{aligned}\|R_e u^{\mathsf{b}}_{\omega_e}\|_{\mathcal{H}^{1/2}(e)} &\le C\left(\|u^{\mathsf{b}}_{\omega_e}\|_{\mathcal{H}(\omega)} + H\|\nabla\cdot(A\nabla u^{\mathsf{b}}_{\omega_e}) + k^2V^2u^{\mathsf{b}}_{\omega_e}\|_{L^2(\omega)}\right)\\ &\le C'H\|f\|_{L^2(\omega_e)}\,,\end{aligned} \tag{6.18}$$

where C' is a constant independent of k and H.

6.6. Proof of Theorem 4.3.

Proof. Define $e_S = u^{\mathsf{h}} - u^{\mathsf{s}} - u_S \in V^{\mathsf{h}}$. Take $\psi = N_k^\star(e_S)$. It holds that

$$\|e_S\|^2_{L^2(\Omega)} = a(e_S,\psi) = a(e_S,\psi - v)$$

for any $v \in S$ due to the property of the Galerkin solution. Thus, using the boundedness of $a(\cdot,\cdot)$, we obtain that

$$\|e_S\|^2_{L^2(\Omega)} \le C_c\|e_S\|_{\mathcal{H}(\Omega)}\|\psi - v\|_{\mathcal{H}(\Omega)} = C_c\|e_S\|_{\mathcal{H}(\Omega)}\|\overline{\psi} - \overline{v}\|_{\mathcal{H}(\Omega)}\,. \tag{6.19}$$

As $\overline{\psi} = N_k\overline{e_S}$ according to the definition of the adjoint problem in subsection 2.2, we can take infimum of v over S, using the fact that $S = \overline{S}$, the definition (4.3), and the inequality (6.19), to get

$$\|e_S\|^2_{L^2(\Omega)} \le C_c\|e_S\|_{\mathcal{H}(\Omega)}\cdot\eta(S)\|\overline{e_S}\|_{L^2(\Omega)}\,,$$

which leads to the desired $L^2(\Omega)$ error estimate: $\|e_S\|_{L^2(\Omega)} \le C_c\eta(S)\|e_S\|_{\mathcal{H}(\Omega)}$.

For the $\mathcal{H}(\Omega)$ error, the property of Galerkin's solution implies that for any $v \in S$, we have

$$\begin{aligned}\|e_S\|^2_{\mathcal{H}(\Omega)} &= \operatorname{Re} a(e_S,e_S) + \{\|e_S\|^2_{\mathcal{H}(\Omega)} - \operatorname{Re} a(e_S,e_S)\}\\ &= \operatorname{Re} a(e_S,u^{\mathsf{h}} - u^{\mathsf{s}} - v) + 2\|kV(x)e_S\|^2_{L^2(\Omega)} + \operatorname{Re}(T_k e_S,e_S)_{\Gamma_N\cup\Gamma_R}\\ &\le C_c\|e_S\|_{\mathcal{H}(\Omega)}\|u^{\mathsf{h}} - u^{\mathsf{s}} - v\|_{\mathcal{H}(\Omega)} + 2(kV_{\max}C_c\eta(S))^2\|e_S\|^2_{\mathcal{H}(\Omega)}\,,\end{aligned} \tag{6.20}$$

where we have used the fact that $\operatorname{Re}(T_k e_S,e_S)_{\Gamma_N\cup\Gamma_R} \le 0$ and the $L^2(\Omega)$ error estimate that we established earlier.

By the assumption $k\eta^{\mathsf{h}}(S) \le 1/(2C_cV_{\max})$, the last term in (6.20) is bounded by $\frac{1}{2}\|e_S\|^2_{\mathcal{H}(\Omega)}$. Thus, due to the arbitrariness of v, we arrive at

$$\|e_S\|_{\mathcal{H}(\Omega)} \le 2C_c\inf_{v\in S}\|u^{\mathsf{h}} - v\|_{\mathcal{H}(\Omega)}\,.$$

This completes the proof for the first part. Next, we move to the proof for the discrete inf-sup stability. For any $v \in S$, set $z = 2N_k^\star(k^2V^2v) \in \mathcal{H}(\Omega)$ so that $a(v,z) = 2k^2(V^2v,v)_\Omega$. Plugging v and $v+z$ into the sesquilinear form yields

$$\begin{aligned}a(v,v+z) &= a(v,v) + a(v,z)\\ &= (A\nabla v,\nabla v)_\Omega - k^2(V^2v,v)_\Omega - (T_kv,v)_{\Gamma_N\cup\Gamma_R} + 2k^2(V^2v,v)_\Omega\\ &= \|v\|^2_{\mathcal{H}(\Omega)} - (T_kv,v)_{\Gamma_N\cup\Gamma_R}\,.\end{aligned}$$

By the definition of T_k, $\operatorname{Re}(T_kv,v)_{\Gamma_N\cup\Gamma_R} \le 0$, so it holds that

$$\operatorname{Re} a(v,v+z) \ge \|v\|^2_{\mathcal{H}(\Omega)}\,.$$

Now, by the definition of the adjoint problem, we have $\overline{z} = 2N_k(k^2V^2\overline{v})$. Let $z_S \in S$ achieve the best approximation in (4.3) for $f = 2k^2V^2\overline{v}$, so that

$$\|\overline{z}^{\mathsf{h}} - \overline{z}^{\mathsf{s}} - z_S\|_{\mathcal{H}(\Omega)} \le \eta(S)\|2k^2V^2\overline{v}\|_{L^2(\Omega)} \le 2kV_{\max}\eta(S)\|v\|_{\mathcal{H}(\Omega)}\,. \tag{6.21}$$

We can choose $v' = v + \overline{z_S} \in S$ to compute

$$\operatorname{Re} a(v,v+\overline{z_S}) = \operatorname{Re} a(v,v+z) - \operatorname{Re} a(v,z-\overline{z_S}) \ge \|v\|^2_{\mathcal{H}(\Omega)} - C_c\|v\|_{\mathcal{H}(\Omega)}\|\overline{z} - z_S\|_{\mathcal{H}(\Omega)}\,.$$

We use the bound in (6.21) and the triangle inequality to get

$$|a(v, v+\overline{z_S})| \geq \|v\|^2_{\mathcal{H}(\Omega)}(1-2C_c k V_{\max}\eta(S)) - C_c\|v\|_{\mathcal{H}(\Omega)}(\|z^{\mathsf{s}}\|_{\mathcal{H}(\Omega)} + \|z^{\mathsf{b}}\|_{\mathcal{H}(\Omega)})\,.$$

Meanwhile, by a triangle inequality, we get

$$\|v+\overline{z_S}\|_{\mathcal{H}(\Omega)} \leq \|v\|_{\mathcal{H}(\Omega)} + \|z^{\mathsf{h}} - z^{\mathsf{s}} - \overline{z_S}\|_{\mathcal{H}(\Omega)} + \|z\|_{\mathcal{H}(\Omega)} + \|z^{\mathsf{s}}\|_{\mathcal{H}(\Omega)} + \|z^{\mathsf{b}}\|_{\mathcal{H}(\Omega)}\,.$$

Finally, we are left to estimate the energy norm of z and its fine scale parts. By the stability estimate in (2.3), we have

$$\|z\|_{\mathcal{H}(\Omega)} \leq C_{\mathrm{stab}}(k)\|2k^2V^2v\|_{L^2(\Omega)} \leq 2C_{\mathrm{stab}}(k)kV_{\max}\|v\|_{\mathcal{H}(\Omega)}\,,$$

and by the bound on the fine part as given by (3.19), it holds that

$$\|z^{\mathsf{s}}\|_{\mathcal{H}(\Omega)} + \|z^{\mathsf{b}}\|_{\mathcal{H}(\Omega)} \leq C_s H\|2k^2V^2\overline{v}\|_{L^2(\Omega)} \leq 2C_s HkV_{\max}\|v\|_{\mathcal{H}(\Omega)}\,.$$

Therefore, we obtain

$$\begin{aligned}\sup_{v'\in S\setminus\{0\}} \frac{|a(v,v')|}{\|v\|_{\mathcal{H}(\Omega)}\|v'\|_{\mathcal{H}(\Omega)}} &\geq \frac{|a(v,v+\overline{z_S})|}{\|v\|_{\mathcal{H}(\Omega)}\|v+\overline{z_S}\|_{\mathcal{H}(\Omega)}}\\ &\geq \frac{(1-2\eta(S)C_ckV_{\max}-2C_cC_sHkV_{\max})\|v\|^2_{\mathcal{H}(\Omega)}}{(1+2\eta(S)kV_{\max}+2C_{\mathrm{stab}}(k)kV_{\max}+2C_sHkV_{\max})\|v\|^2_{\mathcal{H}(\Omega)}}\,.\end{aligned}$$

Using the assumptions that $\eta(S)kV_{\max} \leq 1/(4C_c)$ and $C_sHkV_{\max} \leq 1/(8C_c)$, we obtain the desired estimate. ∎

7. Concluding remarks. In this paper, we have developed a multiscale framework for solving the Helmholtz equation in heterogeneous media and high frequency regimes. The coarse-fine scale decomposition of the solution space is motivated by the MsFEM. In our algorithm, the coarse scale Helmholtz-harmonic part and the fine scale bubble part are computed separately. Their own structures are carefully explored, such as the low complexity of the coarse part and the locality of the fine part. A nearly exponential rate of convergence is proved rigorously and is confirmed numerically for a wide range of the Helmholtz equations with rough coefficients, high contrast, and mixed boundary conditions.

Perhaps surprisingly, our framework implies that designing an accurate multiscale method for the Helmholtz equation is not much more different from that for the elliptic equation. Many techniques in the elliptic case can be successfully adapted once the mesh size satisfies $H = O(1/k)$, a condition that does not suffer from the pollution effect. This work also demonstrates the broad applicability of our exponentially convergent multiscale framework proposed originally in [13].

Most discussions in this paper are concerned with dimension $d = 2$. In our future work, we will generalize the methodology to dimension $d = 3$, where we can use nodal, edge, and face bases to approximate the local solution in the nonoverlapped domain decomposition.

It is also of future interest to extend this methodology systematically to other equations such as the Schrodinger equation, where the problem is time-dependent and the potential function could introduce indefiniteness into the system. On the other hand, developing a better theoretical understanding of the behavior of the multiscale framework with respect to high contrast in the media is also an exciting direction for further exploration.

REFERENCES

[1] A. K. Aziz, R. B. Kellogg, and A. B. Stephens, *A two point boundary value problem with a rapidly oscillating solution*, Numer. Math., 53 (1988), pp. 107–121.

[2] 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.

[3] I. Babuška, R. Lipton, P. Sinz, and M. Stuebner, *Multiscale-spectral GFEM and optimal oversampling*, Comput. Methods Appl. Mech. Engrg., 364 (2020), 112960.

[4] I. Babuška and J. Osborn, *Can a finite element method perform arbitrarily badly?*, Math. Comput., 69 (2000), pp. 443–462.

[5] I. M. Babuska and S. A. Sauter, *Is the pollution effect of the FEM avoidable for the Helmholtz equation considering high wave numbers?*, SIAM J. Numer. Anal., 34 (1997), pp. 2392–2423.

[6] M. Bernkopf, T. Chaumont-Frelet, and J. M. Melenk, *Wavenumber-Explicit Stability and Convergence Analysis of hp Finite Element Discretizations of Helmholtz Problems in Piecewise Smooth Media*, preprint, arXiv:2209.03601, 2022.

[7] T. Betcke, S. N. Chandler-Wilde, I. G. Graham, S. Langdon, and M. Lindner, *Condition number estimates for combined potential integral operators in acoustics and their boundary element discretisation*, Numer. Methods Partial Differential Equations, 27 (2011), pp. 31–69.

[8] D. L. Brown, D. Gallistl, and D. Peterseim, *Multiscale Petrov-Galerkin method for high-frequency heterogeneous Helmholtz equations,* in Meshfree Methods for Partial Differential Equations VIII, Springer, Cham, 2017, pp. 85–115.

[9] A. Buhr and K. Smetana, *Randomized local model order reduction*, SIAM J. Sci. Comput., 40 (2018), pp. A2120–A2151.

[10] K. Chen, Q. Li, J. Lu, and S. J. Wright, *Randomized sampling for basis function construction in generalized finite element methods*, Multiscale Model. Simul., 18 (2020), pp. 1153–1177.

[11] Y. Chen and T. Y. Hou, *Function approximation via the subsampled Poincaré inequality*, Discrete Contin. Dyn. Syst. Ser. A, 41 (2021), pp. 169–199.

[12] Y. Chen and T. Y. Hou, *Multiscale elliptic pde upscaling and function approximation via subsampled data*, Multiscale Model. Simul., 20 (2022), pp. 188–219.

[13] Y. Chen, T. Y. Hou, and Y. Wang, *Exponential convergence for multiscale linear elliptic PDEs via adaptive edge basis functions*, Multiscale Model. Simul., 19 (2021), pp. 980–1010.

[14] E. T. Chung, Y. Efendiev, and W. T. Leung, *Constraint energy minimizing generalized multiscale finite element method*, Comput. Methods Appl. Mech. Engrg., 339 (2018), pp. 298–319.

[15] Y. Efendiev, J. Galvis, and T. Y. Hou, *Generalized multiscale finite element methods (GMsFEM)*, J. Comput. Phys., 251 (2013), pp. 116–135.

[16] B. Engquist and L. Ying, *Sweeping preconditioner for the Helmholtz equation: Hierarchical matrix representation*, Commun. Pure Appl. Math., 64 (2011), pp. 697–735.

[17] B. Engquist and L. Ying, *Sweeping preconditioner for the Helmholtz equation: Moving perfectly matched layers*, Multiscale Model. Simul., 9 (2011), pp. 686–710.

[18] B. Engquist and H. Zhao, *Approximate separability of the Green's function of the Helmholtz equation in the high frequency limit*, Commun. Pure Appl. Math., 71 (2018), pp. 2220–2274.

[19] S. Esterhazy and J. M. Melenk, *On stability of discretizations of the Helmholtz equation,* in Numerical Analysis of Multiscale Problems, Springer, Berlin, 2012, pp. 285–324.

[20] P. Freese, M. Hauck, and D. Peterseim, *Super-Localized Orthogonal Decomposition for High-Frequency Helmholtz Problems*, preprint, arXiv:2112.11368, 2021.

[21] S. Fu, E. Chung, and G. Li, *Edge multiscale methods for elliptic problems with heterogeneous coefficients*, J. Comput. Phys., 396 (2019), pp. 228–242.

[22] S. Fu and K. Gao, *A fast solver for the Helmholtz equation based on the generalized multiscale finite-element method*, Geophys. J. Int., 211 (2017), pp. 797–813.

[23] S. Fu, G. Li, R. Craster, and S. Guenneau, *Wavelet-Based Edge Multiscale Finite Element Method for Helmholtz Problems in Perforated Domains*, preprint, arXiv:1906.08453, 2019.

[24] D. Gallistl and D. Peterseim, *Stable multiscale Petrov-Galerkin finite element method for high frequency acoustic scattering*, Comput. Methods Appl. Mech. Engrg., 295 (2015), pp. 1–17.

[25] I. Graham and S. Sauter, *Stability and finite element error analysis for the Helmholtz equation with variable coefficients*, Math. Comput., 89 (2020), pp. 105–138.
[26] I. G. Graham, O. R. Pembery, and E. A. Spence, *The Helmholtz equation in heterogeneous media: A priori bounds, well-posedness, and resonances*, J. Differential Equations, 266 (2019), pp. 2869–2923.
[27] J. A. Griepentrog and L. Recke, *Linear elliptic boundary value problems with non-smooth data: Normal solvability on Sobolev-Campanato spaces*, Math. Nachr., 225 (2001), pp. 39–74.
[28] M. Hauck and D. Peterseim, *Multi-Resolution Localized Orthogonal Decomposition for Helmholtz Problems*, preprint, arXiv:2104.11190, 2021.
[29] P. Henning and D. Peterseim, *Oversampling for the multiscale finite element method*, Multiscale Model. Simul., 11 (2013), pp. 1149–1175.
[30] U. Hetmaniuk and A. Klawonn, *Error estimates for a two-dimensional special finite element method based on component mode synthesis*, Electron. Trans. Numer. Anal., 41 (2014), pp. 109–132.
[31] U. L. Hetmaniuk and R. B. Lehoucq, *A special finite element method based on component mode synthesis*, ESAIM Math. Model. Numer. Anal., 44 (2010), pp. 401–420.
[32] T. Y. Hou and P. Liu, *Optimal local multi-scale basis functions for linear elliptic equations with rough coefficient*, Discrete Contin. Dyn. Syst., 36 (2016), pp. 4451–4476.
[33] 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.
[34] D. Lafontaine, E. A. Spence, and J. Wunsch, *For Most Frequencies, Strong Trapping Has a Weak Effect in Frequency-Domain Scattering*, preprint, arXiv:1903.12172, 2019.
[35] D. Lafontaine, E. A. Spence, and J. Wunsch, *Wavenumber-explicit convergence of the hp-FEM for the full-space heterogeneous Helmholtz equation with smooth coefficients*, Comput. Math. Appl., 113 (2022), pp. 59–69.
[36] C. Ma, C. Alber, and R. Scheichl, *Wavenumber Explicit Convergence of a Multiscale GFEM for Heterogeneous Helmholtz Problems*, preprint, arXiv:2112.10544, 2021.
[37] C. Ma and R. Scheichl, *Error Estimates for Fully Discrete Generalized FEMs with Locally Optimal Spectral Approximations*, preprint, arXiv:2107.09988, 2021.
[38] C. Ma, R. Scheichl, and T. Dodwell, *Novel Design and Analysis of Generalized FE Methods Based on Locally Optimal Spectral Approximations*, preprint, arXiv:2103.09545, 2021.
[39] A. Målqvist and D. Peterseim, *Localization of elliptic multiscale problems*, Math. Comput., 83 (2014), pp. 2583–2603.
[40] J. Melenk and S. Sauter, *Convergence analysis for finite element discretizations of the Helmholtz equation with Dirichlet-to-Neumann boundary conditions*, Math. Comput., 79 (2010), pp. 1871–1914.
[41] J. M. Melenk and S. Sauter, *Wavenumber explicit convergence analysis for Galerkin discretizations of the Helmholtz equation*, SIAM J. Numer. Anal., 49 (2011), pp. 1210–1243.
[42] A. Moiola and E. A. Spence, *Acoustic transmission problems: Wavenumber-explicit bounds and resonance-free regions*, Math. Models Methods Appl. Sci., 29 (2019), pp. 317–354.
[43] A. A. Oberai and P. M. Pinsky, *A multiscale finite element method for the Helmholtz equation*, Comput. Methods Appl. Mech. Engrg., 154 (1998), pp. 281–297.
[44] M. Ohlberger and B. Verfurth, *A new heterogeneous multiscale method for the Helmholtz equation with high contrast*, Multiscale Model. Simul., 16 (2018), pp. 385–411.
[45] H. Owhadi, *Multigrid with rough coefficients and multiresolution operator decomposition from hierarchical information games*, SIAM Rev., 59 (2017), pp. 99–149.
[46] H. Owhadi and C. Scovel, *Operator-Adapted Wavelets, Fast Solvers, and Numerical Homogenization: From a Game Theoretic Approach to Numerical Approximation and Algorithm Design*, Cambridge University Press, Cambridge, 2019.
[47] H. Owhadi, L. Zhang, and L. Berlyand, *Polyharmonic homogenization, rough polyharmonic splines and sparse super-localization*, ESAIM Math. Model. Numer. Anal., 48 (2014), pp. 517–552.
[48] D. Peterseim, *Eliminating the pollution effect in Helmholtz problems by local subscale correction*, Math. Comput., 86 (2017), pp. 1005–1036.
[49] D. Peterseim and B. Verfürth, *Computational high frequency scattering from high-contrast heterogeneous media*, Math. Comput., 89 (2020), pp. 2649–2674.
[50] J. Poulson, B. Engquist, S. Li, and L. Ying, *A parallel sweeping preconditioner for heterogeneous 3D Helmholtz equations*, SIAM J. Sci. Comput., 35 (2013), pp. C194–C212.
[51] S. Sauter and C. Torres, *Stability estimate for the Helmholtz equation with rapidly jumping coefficients*, Z. Angew. Math. Phys., 69 (2018), 139.

[52] J. Schleuss and K. Smetana, *Optimal Local Approximation Spaces for Parabolic Problems*, preprint, arXiv:2012.02759, 2020.
[53] K. Smetana and A. T. Patera, *Optimal local approximation spaces for component-based static condensation procedures*, SIAM J. Sci. Comput., 38 (2016), pp. A3318–A3356.
[54] L. Tartar, *An Introduction to Sobolev Spaces and Interpolation Spaces*, Springer, Berlin, 2007.