Volume 00, Number 0, Pages 000–000
S (XX)0000-0

DOMAIN DECOMPOSITION PRECONDITIONING FOR HIGH-FREQUENCY HELMHOLTZ PROBLEMS WITH ABSORPTION

I.G. GRAHAM, E.A. SPENCE, AND E. VAINIKKO

Abstract. In this paper we give new results on domain decomposition preconditioners for GMRES when computing piecewise-linear finite-element approximations of the Helmholtz equation $-\Delta u - (k^2 + \mathrm{i}\varepsilon)u = f$, with absorption parameter $\varepsilon \in \mathbb{R}$. Multigrid approximations of this equation with $\varepsilon \neq 0$ are commonly used as preconditioners for the pure Helmholtz case ($\varepsilon = 0$). However a rigorous theory for such (so-called "shifted Laplace") preconditioners, either for the pure Helmholtz equation, or even the absorptive equation ($\varepsilon \neq 0$), is still missing. We present a new theory for the absorptive equation that provides rates of convergence for (left- or right-) preconditioned GMRES, via estimates of the norm and field of values of the preconditioned matrix. This theory uses a k- and ε-explicit coercivity result for the underlying sesquilinear form and shows, for example, that if $|\varepsilon| \sim k^2$, then classical overlapping additive Schwarz will perform optimally for the damped problem, provided the subdomain and coarse mesh diameters are carefully chosen. Extensive numerical experiments are given that support the theoretical results. While the theory applies to a certain weighted variant of GMRES, the experiments for both weighted and classical GMRES give comparable results. The theory for the absorptive case gives insight into how its domain decomposition approximations perform as preconditioners for the pure Helmholtz case $\varepsilon = 0$. At the end of the paper we propose a (scalable) multilevel preconditioner for the pure Helmholtz problem that has an empirical computation time complexity of about $\mathcal{O}(n^{4/3})$ for solving finite element systems of size $n = \mathcal{O}(k^3)$, where we have chosen the mesh diameter $h \sim k^{-3/2}$ to avoid the pollution effect. Experiments on problems with $h \sim k^{-1}$, i.e. a fixed number of grid points per wavelength, are also given.

1. Introduction

This paper is concerned with domain-decomposition preconditioning for finite-element discretisations of the boundary value problem

$$\begin{cases} -\Delta u - (k^2 + \mathrm{i}\varepsilon)u &= f \quad \text{in } \Omega, \\ \partial u/\partial n - \mathrm{i}\eta u &= g \quad \text{on } \Gamma, \end{cases} \tag{1.1}$$

with $k > 0$ and $\eta = \eta(k, \varepsilon)$, where either (i) Ω is a bounded domain in $\mathbb{R}^d$ with boundary Γ or (ii) Ω is the exterior of a bounded scatterer, Γ denotes an approximate far field boundary, and the problem is appended with a homogeneous Dirichlet condition on the boundary of the scatterer. Although the PDE in (1.1) is relevant in applications, our main motivation for studying this problem is its recent use in preconditioning the corresponding BVP for the Helmholtz equation:

$$\begin{cases} -\Delta u - k^2 u &= f \quad \text{in } \Omega, \\ \partial u/\partial n - \mathrm{i}k u &= g \quad \text{on } \Gamma, \end{cases} \tag{1.2}$$

Linear systems arising from finite element approximations of (1.1) with high wavenumber k are notoriously hard to solve. Because the system matrices are non-Hermitian and generally non-normal, general iterative methods like preconditioned (F)GMRES have to be employed. Analysing the convergence of these methods is hard, since an analysis of the spectrum of the system matrix alone is not sufficient for any rigorous convergence estimates.

The idea of preconditioning discretisations of (1.2) with approximate discretisations of (1.1) is often called "shifted Laplacian" preconditioning. From its origins in [17], this idea has had a large impact on the field of practical fast Helmholtz solvers. The main aim of the present paper is to provide theoretical underpinning for this idea, and to use this theoretical understanding to develop new preconditioners for (1.2).

Key words and phrases. Helmholtz equation, high frequency, absorption, iterative solvers, preconditioning, domain decomposition, GMRES.

We denote the system matrix arising from continuous piecewise linear ($P1$) Galerkin finite element approximations of (1.1) by A_ε (or simply A when $\varepsilon = 0$). For the solution of "pure Helmholtz" systems $A\mathbf{u} = \mathbf{f}$, the "shifted Laplacian" preconditioning strategy (written in left-preconditioning mode), involves iteratively solving the equivalent problem

$$B_\varepsilon^{-1} A\mathbf{u} = B_\varepsilon^{-1}\mathbf{f}, \tag{1.3}$$

where B_ε^{-1} is some readily computable approximation of A_ε^{-1} (for example a multigrid V-cycle). The rigorous analysis of the performance of this preconditioner is complicated, partly because it is based on a double approximation: $A^{-1} \approx A_\varepsilon^{-1} \approx B_\varepsilon^{-1}$, and partly because the convergence theory of GMRES for non-self-adjoint systems requires one to estimate either the field of values of the system matrix or the spectrum and its conditioning.

One natural approach is to write

$$I - B_\varepsilon^{-1}A \;=\; I - B_\varepsilon^{-1}A_\varepsilon \;+\; B_\varepsilon^{-1}A_\varepsilon(I - A_\varepsilon^{-1}A)\;, \tag{1.4}$$

and to recall that a sufficient (but by no means necessary) condition for GMRES to converge quickly is that the field of values of the system matrix should be bounded away from the origin and the norm of the system matrix should be bounded above. It is therefore clear from (1.4) that sufficient conditions for B_ε^{-1} to be a good preconditioner for A are:

(i) A_ε^{-1} is a good preconditioner for A

and

(ii) B_ε^{-1} is a good preconditioner for A_ε.

Achieving both (i) and (ii) simultaneously imposes contradictory requirements on ε. Indeed, it is natural to expect that (i) holds if $|\varepsilon|$ is sufficiently small, but that for (ii) to hold we need $|\varepsilon|$ sufficiently large. Most analyses of the performance of B_ε^{-1} as a preconditioner for A have focused on obtaining conditions under which property (i) holds and have concentrated on analysing spectra. While a detailed literature survey is given in [21, §1.1], an up-to-date summary of this is given at the end of this section.

In [21] we gave the first rigorous theory that identified conditions that ensure (i) above holds. There, under general conditions on the domain and mesh sequence, we showed that when $|\varepsilon|/k$ was bounded above by a sufficiently small constant then (i) holds.

The main theoretical purpose of the current paper is to obtain sufficient conditions for (ii) to hold in the case when B_ε^{-1} is chosen as a classical Additive Schwarz preconditioner for A_ε. We use the rigorous convergence theory of [12] (see also [4], [35, §1.3.2]), in which criteria for convergence of GMRES are given in terms of an upper bound on the norm of the system matrix and a lower bound on the distance of its field of values from the origin. In the Additive Schwarz construction, the domain is covered with overlapping subdomains with diameter denoted $H_{\rm sub}$ and also triangulated with a coarse mesh with diameter denoted H. (It is not necessary for $H_{\rm sub}$ and H to be related.) The overlap parameter is denoted δ, and $\delta \sim H$ corresponds to "generous overlap". Further technical requirements are given in §3.

We highlight at this stage that the conditions on $|\varepsilon|$ that we find for (ii) above to hold do not overlap with those described above for (i) to hold, and thus the combination of this paper with [21] does not provide a complete theory for preconditioning the Helmholtz equation with absorption. Nevertheless

(a) we believe that the present paper combined with [21] constitute the only rigorous results in the literature addressing when either of the properties (i) or (ii) above hold,

(b) the investigation into the property (ii) in the present paper, combined with the knowledge from [21] about the property (i), gives insight into how to design a good preconditioner for A (albeit one currently without a rigorous convergence theory); this is especially true when considering multilevel methods – see the discussion around Experiments 2 and 3 in §6.

1.1. **Summary of main theoretical results.** Throughout the paper we assume that $0 < |\varepsilon| \lesssim k^2$, so the ratio $|\varepsilon|/k^2$ is always bounded above, but may approach zero as $k \to \infty$. Our main theoretical results are Theorems 5.6, and 5.8 and their corollaries, which are proved in §5. Theorem 5.6 examines the left-preconditioned matrix: $B_\varepsilon^{-1}A_\varepsilon$, and obtains an upper bound on its norm and a lower bound on its field of values. The upper bound on the norm is $\mathcal{O}(k^2/|\varepsilon|)$, while the distance

of the field of values from the origin (in the case of generous overlap) has a lower bound of order $\mathcal{O}((|\varepsilon|/k^2)^2)$. These bounds are obtained subject to the subdomain and coarse mesh diameters satisfying bounds: $kH_{\text{sub}} \lesssim |\varepsilon|/k^2$ and $kH \lesssim (|\varepsilon|/k^2)^3$.

An important special case is that of maximum absorption $|\varepsilon| \sim k^2$. Then the results imply that the number of GMRES iterates will be bounded independently of k, provided H_{sub} and H both decrease with order k^{-1}. Thus, provided there is enough absorption in the system, the GMRES method will perform analogously to the Laplacian case, provided the coarse mesh decreases proportional to the wavelength (i.e. no further refinement for pollution is needed). Thus if the mesh diameter h of the fine grid decreases as $\mathcal{O}(k^{-3/2})$ (needed to remove pollution in the underlying discretization), then considerable coarsening can be carried out. We actually see in the numerical experiments in §6 that further coarsening beyond the k^{-1} theoretical limit may be possible, depending on the choice of ε.) Analogous results for right preconditioning (obtained by a duality argument) are given in Theorem 5.8. Then Corollaries 5.7 and 5.9 give the corresponding estimates for GMRES convergence for each of these preconditioners.

1.2. **How the theoretical results were obtained.** As in classical Schwarz theory, the proofs of Theorems 5.6 and 5.8 are obtained from a projection operator analysis (given in §4). However, in order to get good results for large k we do not use the classical approach of treating the Helmholtz operator as a perturbation of the Laplacian, as was done in [4] (see also [23], where this approach was used for the time-harmonic Maxwell equations). Rather, we exploit the coercivity of the problem with absorption (Lemma 2.4), leading to a projection analysis in the wavenumber-dependent inner product $(\cdot,\cdot)_{1,k}$. The norm of the projection operator corresponding to the two-level algorithm is estimated above in Theorem 4.3, while the distance of its field of values from the origin is estimated below in Theorem 4.17. The analysis depends on a technical estimate on the approximation power of the coarse space (Assumption 4.6). We prove this estimate for convex polygons (Theorem 4.7), and we also outline how to prove it for more general 2- and 3-d domains (Remark 4.9).

The estimates for the projection operators in §4 are converted to estimates for the norm and field of values of preconditioned Helmholtz matrices in §5. Because the analysis is performed in the "energy" inner product $\|\cdot\|_{1,k}$, the corresponding matrix estimates are obtained in the induced weighted Euclidean inner product. (A similar situation arises in the classical analysis [5].) We performed numerical experiments both for standard GMRES (with residual mininmization in the Euclidean norm) and for weighted GMRES (minimizing in the weighted norm), but in practice there was little difference in the results.

1.3. **Overview of numerical results.** A sequence of numerical experiments is given in §6 for solving systems with matrix A_ε with $h \sim k^{-3/2}$ ($n \sim k^3$, where n is the system dimension), yielding (empirically) pollution-free finite element solutions. In these experiments $H \sim H_{\text{sub}}$. First, we consider the performance of the preconditioner B_ε^{-1} (defined by the classical Additive Schwarz method), when applied to problems with coefficient matrix A_ε. As predicted by the theory, we see that B_ε^{-1} is an optimal preconditioner when $|\varepsilon| \sim k^2$ (i.e. the number of GMRES iterates is parameter independent), provided the coarse grid diameter H and subdomain diameter H_{sub} are sufficiently small. Experimentally, good results are also obtained even with larger H, H_{sub} when ε is large enough, and even with smaller ε when H, H_{sub} are small enough. We also test variants of the classical method, including Restricted Additive Schwarz (RAS) and the Hybrid variant of this (HRAS) (where coarse and local parts of the preconditioner are combined multiplicatively). Out of all the methods tested, HRAS performs the best.

Based on this empirical insight gained about preconditioning A_ε, we then investigate the performance of HRAS (with absorption ε) as a preconditioner for the pure Helmholtz problem with coefficient matrix A. We find that HRAS still works well, provided H and H_{sub} are small enough. There is surprisingly little variation in the performance with respect to the choice of ε. (In fact with $|\varepsilon| = k^\beta$, the performance is almost uniform in the range $\beta \in [0, 1.2]$ but there is some degradation as β approaches 2. This is surprising as the choice $\beta \sim 2$ is normally used in the multigrid context. We also test a variant of HRAS that uses impedance conditions on subdomain solves and this works well, especially for larger H, H_{sub}.

Finally, to solve problems with matrix A in the case of large k, we recommend an inner-outer preconditioner for use within FGMRES, where the outer solver is HRAS with $|\varepsilon| = k$ and $H \sim$

$H_{\rm sub} \sim k^{-1}$. The cost of the preconditioner is then dominated by the coarse grid problem, and for this we apply an inner iteration with preconditioner chosen as one-level HRAS with impedance boundary condition on local problems. With the best choice of ε appearing to be $|\varepsilon| \sim k$, we find this solver has a compute time of about $\mathcal{O}(k^4) \sim \mathcal{O}(n^{4/3})$ for the 2D problems tested, up to $k = 100$. This is a highly scalable preconditioner, whose action consists of inverting $\mathcal{O}(k^2)$ (parallel) finite-element systems of size $\mathcal{O}(k)$ and an additional $\mathcal{O}(k)$ finite-element systems of size $\mathcal{O}(k)$. Additional experiments, together with multilevel variations suitable for the case $h \sim k^{-1}$ are given in [26].

1.4. **Literature review.** We finish this section with a short literature survey on this topic, beginning with the literature on preconditioning with absorption, and then briefly discussing domain decomposition methods for wave problems.

The survey in [21, §1.1] focused on the spectral analyses in [17], [16], [43], [15, §5.1.2], [18], all of which concern the optimal choice of ε for A_ε to be a good preconditioner for A, i.e. for property (i) above to hold. Several authors have considered the question of when multigrid methods converge when applied to the problem with absorption (i.e. A_ε); this is related to (but not the same as) the question of when property (ii) above holds. Cools and Vanroose [9] computed the "minimal shift" (defined as the smallest value of ε for which every single eigenmode of the error is reduced through consecutive multigrid iterations) based on numerical evaluation of quantities arising from Fourier analysis, and found that (as a function of k) it is proportional to k^2. Cocquet and Gander [8] (following on from [18]) showed that, for a particular standard variant of multigrid applied to the 1-d Helmholtz equation with Dirichlet boundary conditions, one needs $|\varepsilon| \sim k^2$ to obtain convergence independent of k. They also analysed a less-standard variant of multigrid applied to general multi-dimensional Helmholtz problems with either Dirichlet or impedance boundary conditions, and showed that again one needs $|\varepsilon| \sim k^2$ for the method to be practical. Note that these analyses are concerned with the *convergence* of multigrid as a solver for A_ε, rather than using an approximation such as the V-cycle applied to the problem with absorption as a preconditioner for either the absorbing or the non-absorbing problem (A or A_ε respectively).

The study of non-overlapping domain-decomposition methods for wave problems has a long history, starting with the seminal paper of Benamou and Després [2]. Following that, optimized interface conditions were introduced [22], the success of which sparked substantial interest, for example [20], [11], and more recently the "source transfer" and related methods [7], [6], and [40]; these latter methods can be viewed as putting the "sweeping" method of [14] in a continuous (as opposed to discrete) setting. All these non-overlapping domain decomposition methods focus on the choice of good interface conditions but so far do not provide a systematic method of combining these with coarse grid operators or a convergence analysis explicit in subdomain or coarse grid size. There are also a few results on overlapping domain decomposition methods e.g. [41], [30], [31], with the latter explicitly using absorption; these demonstrated the potential of the methods analysed in this paper. Finally, we note that [44] introduces a new sweeping-style method for the Helmholtz equation, and also contains a good literature review of both domain-decomposition and sweeping-style methods.

2. Variational formulation

For ease of exposition, we restrict attention to the interior impedance problem (i.e. (1.2) is posed for Ω a bounded domain in $\mathbb{R}^d$ with boundary Γ). The results of the paper also hold for the truncated sound-soft scattering problem, and we outline in Remark 5.10 how to adapt them to this case.

Let Ω be a bounded, open, polygonal (Lipschitz polyhedral) domain in $\mathbb{R}^d$, $d = 2$ (or 3), with boundary Γ. We introduce the standard k-weighted inner product and norm on $H^1(\Omega)$:

$$(v,w)_{1,k} \;=\; (\nabla v, \nabla w)_{L^2(\Omega)} + k^2 (v,w)_{L^2(\Omega)} \quad \text{and} \quad \|v\|_{1,k} = (v,v)_{1,k}^{1/2} \,.$$

The standard variational formulation of (1.1) is: Given $f \in L^2(\Omega)$, $g \in L^2(\Gamma)$, $\varepsilon \in \mathbb{R}$ and $k > 0$ find $u \in H^1(\Omega)$ such that

$$a_\varepsilon(u,v) = F(v) \quad \text{for all } v \in H^1(\Omega), \tag{2.1}$$

where

$$a_\varepsilon(u,v) := \int_\Omega \nabla u \cdot \overline{\nabla v} - (k^2 + \mathrm{i}\varepsilon) \int_\Omega u\overline{v} - \mathrm{i}\eta \int_\Gamma u\overline{v}, \tag{2.2}$$

and

$$F(v) := \int_\Omega f\overline{v} + \int_\Gamma g\overline{v}. \tag{2.3}$$

In general η can be complex, with a natural choice being a square root of $k^2 + \mathrm{i}\varepsilon$. (more details are in Lemma 2.4). When $\varepsilon = 0$ and $\eta = k$ we are solving (1.2) and we simply write a instead of a_ε.

We consider the discretisation of problem (2.1) with $P1$ finite elements. Let $\mathcal{T}^h$ be a family of conforming meshes (triangles in 2D, tetrahedra in 3D), that are shape-regular as the mesh diameter $h \to 0$. A typical element of $\mathcal{T}^h$ is $\tau \in \mathcal{T}^h$ (a closed subset of $\overline{\Omega}$). Then our approximation space $\mathcal{V}^h$ is the space of all continuous functions on Ω that are piecewise affine with respect to $\mathcal{T}^h$. (The impedance boundary condition in (1.2) is implemented as a natural boundary condition.) The freedoms for $\mathcal{T}^h$ are the nodes, denoted $\mathcal{N}^h = \{x_j : j \in \mathcal{I}^h\}$, where $\mathcal{I}^h$ is a suitable index set. The standard basis for $\mathcal{V}^h$ is $\{\phi_j : j \in \mathcal{I}^h\}$ consisting of hat functions corresponding to the each of the nodes in $\mathcal{N}^h$.

The Galerkin approximation of (2.1) in the space $\mathcal{V}^h$ is equivalent to the system

$$A_\varepsilon \mathbf{u} \;:=\; (S - (k^2 + \mathrm{i}\varepsilon)M - \mathrm{i}\eta N)\mathbf{u} \;=\; \mathbf{f}, \tag{2.4}$$

where

$$S_{\ell,m} = \int_\Omega \nabla\phi_\ell \cdot \nabla\phi_m, \quad M_{\ell,m} = \int_\Omega \phi_\ell\phi_m, \quad N_{\ell,m} = \int_\Gamma \phi_\ell\phi_m, \quad \ell, m \in \mathcal{I}^h \tag{2.5}$$

are, respectively, the stiffness matrix, the domain mass matrix, and the boundary mass matrix. Again we write the corresponding system matrix for (1.2) simply as A. Note that A and A_ε are symmetric but not Hermitian.

In this section we briefly provide the key properties of the sesquilinear form a_ε given in (2.2). This form depends on all three parameters ε, k and η, but only the first of these is reflected in the notation. Normally η will be chosen as a function of ε and k. We will assume throughout that

$$|\varepsilon| \lesssim k^2 \quad \text{and} \quad |\eta| \lesssim k. \tag{2.6}$$

(Here the notation $A \lesssim B$ (equivalently $B \gtrsim A$) means that A/B is bounded above by a constant independent of k, ε, and mesh diameters h, H_{sub}, H (the latter two introduced below). We write $A \sim B$ when $A \lesssim B$ an $B \lesssim A$.

The proof of the first result is a simple application of the Cauchy-Schwarz and multiplicative trace inequalities - see, e.g., [21, Lemma 3.1(i)].

Lemma 2.1 (Continuity)**.** *If $|\eta| \lesssim k$ then, given $k_0 > 0$, there exists a C_c independent of k and ε such that*

$$|a_\varepsilon(v,w)| \le C_c \|v\|_{1,k} \|w\|_{1,k} \tag{2.7}$$

for all $k \ge k_0$ and $v, w \in H^1(\Omega)$.

We now give a result about the coercivity of a_ε, which is a generalisation of [21, Lemma 3.1(ii)]. To state this we need to define $\sqrt{k^2 + \mathrm{i}\varepsilon}$, taking care to cater for both positive and negative ε. We need to consider both positive and negative ε since, whichever choice we make for the problem (1.1), the other forms the adjoint problem, and we need estimates on the solutions and sesquilinear forms for both problems (in particular, this is essential for analysing both left- and right-preconditioning).

Definition 2.2. *$z(k,\varepsilon) := \sqrt{k^2 + \mathrm{i}\varepsilon}$ is defined via the square root with the branch cut on the positive real axis. Note that this definition implies that, when $\varepsilon \neq 0$,*

$$\Im(z) > 0, \quad \operatorname{sign}(\varepsilon)\Re(z) > 0, \quad \text{and} \quad z(k,-\varepsilon) = -\overline{z(k,\varepsilon)}. \tag{2.8}$$

Proposition 2.3. *With $z(k,\varepsilon)$ defined above, for all $k > 0$,*

$$|z| \sim k \quad \text{and} \quad \frac{\Im(z)}{|z|} \sim \frac{|\varepsilon|}{k^2}. \tag{2.9}$$

Proof. Writing $z = p + \mathrm{i}q$, we see that the definition of z implies that

$$p = \sqrt{p^2} \text{ if } \varepsilon > 0, \quad p = -\sqrt{p^2} \text{ if } \varepsilon < 0, \quad \text{and} \quad q = \sqrt{q^2} \text{ for all } \varepsilon \neq 0,$$

where

$$p^2 = \frac{\sqrt{k^4+\varepsilon^2}+k^2}{2}, \qquad \text{and} \qquad q^2 = \frac{\sqrt{k^4+\varepsilon^2}-k^2}{2} \tag{2.10}$$

(and $\sqrt{\cdot}$ denotes the positive real square root). Using (2.6) we therefore see that $|p| \sim k$. Furthermore, the definition of z implies that $2pq = \varepsilon$, and thus $q = |q| \sim |\varepsilon|/|p| \sim |\varepsilon|/k$. Using (2.6) again, the estimates (2.9) follow. ☐

Lemma 2.4 (Coercivity)**.** *Let $z = z(k,\varepsilon)$ be as defined in Definition 2.2, and choose η in (2.2) to satisfy the inequality*

$$\Re(\overline{z}\eta) \geq 0. \tag{2.11}$$

Then there is a constant $\rho > 0$ independent of k and ε such that

$$|a_\varepsilon(v,v)| \;\geq\; \Im\left(\Theta a_\varepsilon(v,v)\right) \;\geq\; \rho\,\frac{|\varepsilon|}{k^2}\,\|v\|^2_{1,k} \tag{2.12}$$

for all $k > 0$ and $v \in H^1(\Omega)$, where $\Theta = -\overline{z}/|z|$.

Proof. Writing $z = p + iq$ and using the definition of a_ε, we have

$$a_\varepsilon(v,v) \;=\; \|\nabla v\|^2_{L^2(\Omega)} - (p+\mathrm{i}q)^2\|v\|^2_{L^2(\Omega)} - \mathrm{i}\eta\|v\|^2_{L^2(\Gamma)}\,.$$

Therefore

$$\Im\left[-(p-\mathrm{i}q)a_\varepsilon(v,v)\right] \;=\; q\|\nabla v\|^2_{L^2(\Omega)} + q(p^2+q^2)\|v\|^2_{L^2(\Omega)} + \Re\left[(p-\mathrm{i}q)\eta\right]\|v\|^2_{L^2(\Gamma)}\,.$$

Hence, dividing through by $|z| = \sqrt{p^2+q^2}$, and setting $\Theta = -\overline{z}/|z|$, we have

$$\Im\left[\Theta a_\varepsilon(v,v)\right] \;=\; \frac{\Im(z)}{|z|}\left[\|\nabla v\|^2_{L^2(\Omega)} + |z|^2\|v\|^2_{L^2(\Omega)}\right] + \frac{\Re\left(\overline{z}\eta\right)}{|z|}\|v\|^2_{L^2(\Gamma)}\,.$$

The result then follows from condition (2.11) and the second estimate in (2.9). ☐

Remark 2.5 (Choices of η satisfying (2.11))**.** *An obvious choice of η that satisfies the coercivity condition (2.11) is $\eta = z$, for then $\Re(\overline{z}\eta) = \Re(\overline{z}z) = |z|^2 > 0$, Another possible choice is $\eta = \mathrm{sign}(\varepsilon)k$, for then, by (2.8), we have $\Re(\overline{z}\eta) = \mathrm{sign}(\varepsilon)\Re(z)k > 0$. Note that both these choices satisfy the condition on $|\eta|$ in (2.6).*

The fact that the choice of η for coercivity to hold depends on the sign of ε is expected, since the sign of ε also dictates the properties of η required for the problem (1.1) to be well posed. Indeed, repeating the usual argument involving Green's identity (given for $\varepsilon = 0$ in, e.g., [39, Theorem 6.5]) we see that if $\varepsilon > 0$ we need $\Re(\eta) \geq 0$ for uniqueness and if $\varepsilon < 0$ we need $\Re(\eta) \leq 0$.

The condition for coercivity (2.11) is more restrictive that the conditions for uniqueness. Indeed, since $\Re(\overline{z}\eta) = \Re(z)\Re(\eta) + \Im(z)\Im(\eta)$, when $\varepsilon > 0$, a sufficient condition to ensure (2.11) is $\Re(\eta) > 0$, $\Im(\eta) \geq 0$. Similarly, when $\varepsilon < 0$ a sufficient condition for (2.11) is $\Re(\eta) < 0$, $\Im(\eta) \geq 0$.

This lemma immediately also gives us a result about the coercivity of the adjoint of a_ε, given by

$$a^*_\varepsilon(u,v) \;=\; \int_\Omega \nabla u\cdot\overline{\nabla v} - (k^2 - \mathrm{i}\varepsilon)\int_\Omega u\overline{v} + \mathrm{i}\overline{\eta}\int_\Gamma u\overline{v}.$$

Corollary 2.6. *Under assumption (2.11) we also have coercivity of the adjoint form:*

$$|a^*_\varepsilon(v,v)| \;\geq\; \Im\left(\Theta a_\varepsilon(v,v)\right) \;\geq\; \rho\,\frac{|\varepsilon|}{k^2}\,\|v\|^2_{1,k} \tag{2.13}$$

for all $k > 0$ and $v \in H^1(\Omega)$, where $\Theta = -\overline{z}/|z|$.

Proof. Note that the adjoint form is simply a copy of the original form a_ε, but with parameters ε and η replaced by $\widetilde{\varepsilon} = -\varepsilon$ and $\widetilde{\eta} = -\overline{\eta}$, and thus (by (2.8)) $\widetilde{z} = -\overline{z}$. The condition for coercivity of the adjoint form is then $\Re(\overline{\widetilde{z}}\widetilde{\eta}) \geq 0$, which is equivalent to condition (2.11). ☐

Remark 2.7. *Throughout the paper we will always assume that ε and η are chosen so that conditions (2.6) and (2.11) hold, and so the forms a_ε and a^*_ε always will satisfy the continuity and coercivity estimates (2.7), (2.12) and (2.13).*

3. Domain Decomposition

To define appropriate subspaces of $\mathcal{V}^h$, we start with a collection of open subsets $\{\widetilde{\Omega}_\ell : \ell = 1,\dots,N\}$ of $\mathbb{R}^d$ that form an overlapping cover of $\overline{\Omega}$, and we set $\Omega_\ell = \widetilde{\Omega}_\ell \cap \overline{\Omega}$. Each $\overline{\Omega}_\ell$ is assumed to be non-empty and $\overline{\Omega}_\ell$ is assumed to consist of a union of elements of the mesh $\mathcal{T}_h$. Then, for each $\ell = 1,\dots,N$, we set

$$\mathcal{V}_\ell = \{v_h \in \mathcal{V}^h : \mathrm{supp}(v_h) \subset \overline{\Omega}_\ell\}.$$

Note that, since functions in $\mathcal{V}^h$ are continuous, functions in $\mathcal{V}_\ell$ must vanish on the internal boundary $\partial\Omega_\ell\backslash\Gamma$, but are unconstrained on the external boundary $\partial\Omega_\ell \cap \Gamma$. The freedoms for $\mathcal{V}_\ell$ are denoted $\mathcal{N}^h(\Omega_\ell) = \{x_j : j \in \mathcal{I}^h(\Omega_\ell)\}$, where $\mathcal{I}^h(\Omega_\ell)$ is a suitable index set. The basis for $\mathcal{V}^h(\Omega_\ell)$ can then be written $\{\phi_j : j \in \mathcal{I}^h(\Omega_\ell)\}$.

Thus a solve of the Helmholtz problem (2.1) in the space $\mathcal{V}_\ell$ involves a Dirichlet boundary condition at internal boundaries and natural boundary condition at external boundaries (if any). The introduction of the absorption $\varepsilon \neq 0$ ensures such solves are always well-defined. Future work will consider the analysis of methods with other local boundary conditions (such as impedance or PML). Internal impedance conditions are considered in the experiments in §6.

For $j \in \mathcal{I}^h(\Omega_\ell)$ and $j' \in \mathcal{I}^h$, we define the restriction matrix $(R_\ell)_{j,j'} := \delta_{j,j'}$. The matrix $A_{\varepsilon,\ell} := R_\ell A_\varepsilon R_\ell^T$ is then just the minor of A_ε corresponding to rows and columns taken from $\mathcal{I}^h(\Omega_\ell)$. One-level domain decomposition methods are constructed from the inverses $A_{\varepsilon,\ell}^{-1}$. More precisely,

$$B_{\varepsilon,AS,local}^{-1} := \sum_\ell R_\ell^T A_{\varepsilon,\ell}^{-1} R_\ell \tag{3.1}$$

is the classical one-level preconditioner for A_ε with the subscript "*local*" indicating that the solves are on local subdomains Ω_ℓ.

For the theory, we need assumptions on the shape of the subdomains and the size of the overlap, and we require any point in $\overline{\Omega}$ to belong to a bounded number of overlapping subdomains. First, for simplicity we assume the subdomains are shape-regular Lipschitz polyhedra (polygons in 2D) of diameter $H_\ell = \mathrm{diam}(\Omega_\ell)$, with the volume of order $\sim H_\ell^d$ and surface area $\sim H_\ell^{d-1}$ respectively. The coarse mesh diameter $H_{\mathrm{sub}} := \max\{H_\ell : \ell = 1,\dots,N\}$ is then a parameter in our estimates. Each Ω_l is required to have a large enough interior boundary, i.e. we require that

$$|\partial\Omega_l \setminus \Gamma| \sim H_{\mathrm{sub}}^{d-1} \quad \text{for each } l. \tag{3.2}$$

Concerning the overlap, for each $\ell = 1,\dots,N$, let $\mathring{\Omega}_\ell$ denote the part of Ω_ℓ that is not overlapped by any other subdomains, and for $\mu > 0$ let $\Omega_{\ell,\mu}$ denote the set of points in Ω_ℓ that are a distance no more than μ from the boundary $\partial\Omega_\ell$. Then we assume that for some $\delta > 0$ and some $0 < c < 1$ fixed,

$$\Omega_{\ell,c\delta} \subset \Omega_\ell\backslash\mathring{\Omega}_\ell \subset \Omega_{\ell,\delta}. \tag{3.3}$$

Put more simply, the overlap is assumed to be uniformly of order δ; the case $\delta \sim H$ is called "generous overlap". Finally, we make the *finite overlap assumption*

$$\#\Lambda(\ell) \lesssim 1, \quad \text{where} \quad \Lambda(\ell) = \{\ell' : \Omega_\ell \cap \Omega_{\ell'} \neq \emptyset\}\ . \tag{3.4}$$

Two-level methods are obtained by adding a global coarse solve. Let $\{\mathcal{T}^H\}$ be a sequence of shape-regular, simplicial meshes on $\overline{\Omega}$, with mesh diameter H. We assume that each element of $\mathcal{T}^H$ consists of the union of a set of fine grid elements. The set of coarse mesh nodes is denoted by $\mathcal{I}^H$. The coarse space basis functions Φ_p are taken to be the continuous $P1$ hat functions on $\mathcal{T}^H$.

From these functions we define the coarse space $\mathcal{V}_0 := \mathrm{span}\{\Phi_p : p \in \mathcal{I}^H\}$, which is a subspace of $\mathcal{V}^h$. Now, if we introduce the restriction matrix

$$(R_0)_{pj} := \Phi_p(x_j^h)\ , \quad j \in \mathcal{I}^h, \quad p \in \mathcal{I}^H, \tag{3.5}$$

then the matrix

$$A_{\varepsilon,0} := R_0 A_\varepsilon R_0^T \tag{3.6}$$

is the stiffness matrix for problem (1.2) discretised in $\mathcal{V}_0$ using the basis $\{\Phi_p : p \in \mathcal{I}^H\}$. Note that, due to the coercivity result Lemma 2.4, both $A_{\varepsilon,0}$ and $A_{\varepsilon,\ell}$ are invertible for all mesh sizes h and

all choices of $\epsilon \neq 0$. This is easily seen, since, for example, if $A_{\varepsilon,0}\mathbf{v} = \mathbf{0}$, where $\mathbf{v}$ is a vector defined on the freedoms $\mathcal{I}^H$, then $0 \;=\; \mathbf{v}^* A_{\varepsilon,0}\mathbf{v} \;=\; a_\varepsilon(v_H, v_H)$, where $v_H = \sum_{p\in\mathcal{I}^H} v_p \Phi_p$ and so

$$0 \;=\; |a_\varepsilon(v_H, v_H)| \;\geq\; \rho\frac{|\varepsilon|}{k^2}\|v_H\|^2_{1,k},$$

which immediately implies $v_H = 0$, and thus $\mathbf{v} = \mathbf{0}$. Similar arguments apply to $A_{\varepsilon,\ell}$ and to the adjoints $A^*_{\varepsilon,\ell}$, $\ell = 0, \ldots, N$.

The classical Additive Schwarz method is

$$B^{-1}_{\varepsilon,AS} \;:=\; R_0^T A^{-1}_{\varepsilon,0} R_0 \;+\; B^{-1}_{\varepsilon,AS,local}, \tag{3.7}$$

(i.e. the sum of coarse solve and local solves) with $B^{-1}_{\varepsilon,AS,local}$ defined in (3.1).

4. Theory of Additive Schwarz Methods

The following theory establishes rigorously the powerful properties of the preconditioner (3.7) applied to A_ε if $|\varepsilon|$ is sufficiently large and $H_{\rm sub}$, H are sufficiently small.

This theory was inspired by reading again the results in [4] where non-self-adjoint problems that were "close to" self-adjoint coercive problems were considered. Although our problem here is not close to a self-adjoint coercive one, and our technical tools are very different, [4] provided a framework that we were able to adapt into the following results.

The first lemma is an extension of the familiar "stable splitting" property of domain decomposition spaces. This is well-known for the H^1 norm (see, e.g., [42]) but here we extend it to the case of the k-weighted energy norm.

Lemma 4.1. *For all $v_h \in \mathcal{V}^h$, there exist $v_\ell \in \mathcal{V}_\ell$ for each $\ell = 0, \cdots, N$ such that*

$$v_h = \sum_{\ell=0}^N v_\ell \quad \textit{and} \quad \sum_{\ell=0}^N \|v_\ell\|^2_{1,k} \;\lesssim\; \left(1 + \frac{H}{\delta}\right)\|v_h\|^2_{1,k}\ . \tag{4.1}$$

Proof. This is adapted from the proof of analogous results for Laplace problems; see, e.g., [42]. The proof starts by approximating v_h by the quasiinterpolant from the coarse space:

$$v_0 := \sum_{p\in\mathcal{I}^H} \widehat{v}_p \Phi^H_p$$

where

$$\widehat{v}_p = |\omega_p|^{-1}\int_{\omega_p} v_h \quad \text{and} \quad \omega_p = \text{supp}(\Phi^H_p)\ .$$

Then, using the shape regularity of $\mathcal{T}^H$ it is straightforward to show that

$$\|v_0\|_{L^2(\Omega)} \lesssim \|v_h\|_{L^2(\Omega)}. \tag{4.2}$$

Next, we take a partition of unity $\{\chi_\ell : \ell = 1, \ldots, N\}$ subordinate to the covering Ω_ℓ and set

$$v_\ell = I^h(\chi_\ell(v_h - v_0))\ , \tag{4.3}$$

where I^h denotes nodal interpolation onto $\mathcal{V}^h$. The first equality in (4.1) follows easily after summation. Moreover the estimate

$$\sum_{\ell=0}^N |v_\ell|^2_{H^1(\Omega)} \;\lesssim\; \left(1 + \frac{H}{\delta}\right)|v_h|^2_{H^1(\Omega)} \tag{4.4}$$

is familiar from results on self-adjoint coercive problems; see, e.g., [27, Theorem 3.8].

To obtain the second inequality in (4.1), we note that by definition of I^h, we have, for any $\tau \in \mathcal{T}^h$ with $\tau \subset \overline{\Omega_\ell}$, and any $x \in \tau$, we have

$$\begin{aligned} |v_\ell(x)| \;&=\; \left|\sum_{j\in\mathcal{I}^h(\tau)} (\chi_\ell(v_h - v_0))(x^h_j)\phi^h_j(x)\right| \;\leq\; \sum_{j\in\mathcal{I}^h(\tau)} |(v_h - v_0)(x^h_j)| \\ &\lesssim \left\{\sum_{j\in\mathcal{I}^h(\tau)} |(v_h - v_0)(x^h_j)|^2\right\}^{1/2} \;\sim\; |\tau|^{-1/2}\|v_h - v_0\|_{L^2(\tau)}, \end{aligned}$$

where $\{x_j : j \in \mathcal{I}^h(\tau)\}$ denotes the nodes on τ. Hence

$$\begin{aligned} \|v_\ell\|^2_{L^2(\Omega)} &= \sum_{\tau\subset\overline{\Omega_\ell}} \int_\tau |v_\ell|^2 \le \sum_{\tau\subset\overline{\Omega_\ell}} |\tau| \|v_\ell\|^2_{L^\infty(\tau)} \\ &\lesssim \sum_{\tau\subset\overline{\Omega_\ell}} |\tau||\tau|^{-1} \|v_h - v_0\|^2_{L^2(\tau)} = \|v_h - v_0\|^2_{L^2(\Omega_\ell)}. \end{aligned} \tag{4.5}$$

Thus, because of the finite overlap property (3.4), we have

$$\sum_{\ell=1}^N \|v_\ell\|^2_{L^2(\Omega)} \lesssim \|v_h - v_0\|^2_{L^2(\Omega)} \lesssim \|v_h\|^2_{L^2(\Omega)} + \|v_0\|^2_{L^2(\Omega)} . \tag{4.6}$$

Combination of this with (4.2) yields $\sum_{\ell=0}^N \|v_\ell\|^2_{L^2(\Omega)} \lesssim \|v_h\|^2_{L^2(\Omega)}$. Then multiplication by k^2 and combination with (4.4) gives the required result. □

The next lemma is a kind of converse to Lemma 4.1. Here the energy of a sum of components is estimated above by the sum of the energies.

Lemma 4.2. *For all choices of $v_\ell \in \mathcal{V}_\ell$, $\ell = 0, \cdots, N$, we have*

$$\left\| \sum_{\ell=0}^N v_\ell \right\|^2_{1,k} \lesssim \sum_{\ell=0}^N \|v_\ell\|^2_{1,k} . \tag{4.7}$$

Proof. Let $\sum_\ell$ denote the sum from $\ell = 1$ to N and and recall the notation $\Lambda(\ell)$ introduced in (3.4). Then, using several applications of the Cauchy-Schwarz inequality,

$$\begin{aligned} \left\| \sum_\ell v_\ell \right\|^2_{1,k} &= \left(\sum_\ell v_\ell, \sum_{\ell'} v_{\ell'} \right)_{1,k} = \sum_\ell \sum_{\ell'\in\Lambda(\ell)} (v_\ell, v_{\ell'})_{1,k} \\ &\le \sum_\ell \|v_\ell\|_{1,k} \left(\sum_{\ell'\in\Lambda(\ell)} \|v_{\ell'}\|_{1,k} \right) \\ &\le \left(\sum_\ell \|v_\ell\|^2_{1,k} \right)^{1/2} \left(\sum_\ell \left(\sum_{\ell'\in\Lambda(\ell)} \|v_{\ell'}\|_{1,k} \right)^2 \right)^{1/2} \\ &\le \left(\sum_\ell \|v_\ell\|^2_{1,k} \right)^{1/2} \left(\sum_\ell \#\Lambda(\ell) \sum_{\ell'\in\Lambda(\ell)} \|v_{\ell'}\|^2_{1,k} \right)^{1/2} \lesssim \sum_\ell \|v_\ell\|^2_{1,k} , \end{aligned} \tag{4.8}$$

where we used the finite overlap assumption (3.4). To obtain (4.7), we write

$$\begin{aligned} \left\| \sum_{\ell=0}^N v_\ell \right\|^2_{1,k} &= \left(\sum_{\ell=0}^N v_\ell, \sum_{\ell=0}^N v_\ell \right)_{1,k} \\ &= \|v_0\|^2_{1,k} + 2 \left(v_0, \sum_\ell v_\ell \right)_{1,k} + \left(\sum_\ell v_\ell, \sum_\ell v_\ell \right)_{1,k} . \end{aligned} \tag{4.9}$$

Using the Cauchy-Schwarz and the arithmetic-geometric mean inequalities on the middle term we can estimate (4.9) from above in the form

$$\lesssim \|v_0\|^2_{1,k} + \left\| \sum_\ell v_\ell \right\|^2_{1,k} ,$$

and the result follows from (4.8). □

Now for each $\ell = 0, \dots, N$, we define linear operators $Q_{\varepsilon,\ell} : H^1(\Omega) \to \mathcal{V}_\ell$ as follows. For each $v_h \in H^1(\Omega)$, $Q_{\varepsilon,\ell} v$ is defined to be the unique solution of the equation

$$a_\varepsilon(Q_{\varepsilon,\ell} v, w_{h,l}) \;=\; a_\varepsilon(v, w_{h,\ell}), \quad w_{h,\ell} \in \mathcal{V}_\ell. \tag{4.10}$$

and we then define

$$Q_\varepsilon = \sum_{\ell=0}^{N} Q_{\varepsilon,\ell} \ .$$

The matrix representation of Q_ε corresponds to the action of the preconditioner (3.7) on the matrix A_ε (this will be shown in Theorem 5.4 below). In Theorems 4.3 and 4.17 below we estimate the norm and field of values of Q_ε, and this yields corresponding estimates for the norm and field of values of the preconditioned matrix in Theorems 5.6. Such projection analysis is commonplace in domain decomposition; however, as far as we are aware, this is the first place where the projection operators are defined using the a_ε sesquilinear form and analysed in the wavenumber-dependent $\|\cdot\|_{1,k}$ energy norm.

Theorem 4.3. *(Upper bound on Q_ε)*

$$\|Q_\varepsilon v_h\|_{1,k} \;\lesssim\; \left(\frac{k^2}{|\varepsilon|}\right) \|v_h\|_{1,k} \quad \textit{for all} \quad v_h \in \mathcal{V}^h.$$

Proof. By the definition of Q_ε and Lemma 4.2, we have

$$\|Q_\varepsilon v_h\|_{1,k}^2 \;=\; \left\| \sum_{\ell=0}^{N} Q_{\varepsilon,\ell} v_h \right\|_{1,k}^2 \;\lesssim\; \sum_{\ell=0}^{N} \|Q_{\varepsilon,\ell} v_h\|_{1,k}^2 \ . \tag{4.11}$$

Furthermore, by applying Lemma 2.4 and the definition (4.10), we have

$$\begin{aligned}\sum_{\ell=0}^{N} \|Q_{\varepsilon,\ell} v_h\|_{1,k}^2 \;&\lesssim\; \left(\frac{k^2}{|\varepsilon|}\right) \sum_{\ell=0}^{N} \Im\left(\Theta a_\varepsilon(Q_{\varepsilon,\ell} v_h, Q_{\varepsilon,\ell} v_h)\right) \;=\; \left(\frac{k^2}{|\varepsilon|}\right) \Im\left(\Theta \sum_{\ell=0}^{N} a_\varepsilon(v_h, Q_{\varepsilon,\ell} v_h)\right) \\ &=\; \left(\frac{k^2}{|\varepsilon|}\right) \Im\left(\Theta a_\varepsilon\left(v_h, \sum_{\ell=0}^{N} Q_{\varepsilon,\ell} v_h\right)\right) \;\le\; \left(\frac{k^2}{|\varepsilon|}\right) \left| a_\varepsilon\left(v_h, \sum_{\ell=0}^{N} Q_{\varepsilon,\ell} v_h\right)\right|\end{aligned}$$

(recalling that $|\Theta| = 1$). Then, using Lemma 2.1, and then Lemma 4.2, we have

$$\begin{aligned}\sum_{\ell=0}^{N} \|Q_{\varepsilon,\ell} v_h\|_{1,k}^2 \;&\lesssim\; \left(\frac{k^2}{|\varepsilon|}\right) \|v_h\|_{1,k} \left\| \sum_{\ell=0}^{N} Q_{\varepsilon,\ell} v_h \right\|_{1,k} \\ &\lesssim\; \left(\frac{k^2}{|\varepsilon|}\right) \|v_h\|_{1,k} \left(\sum_{\ell=0}^{N} \|Q_{\varepsilon,\ell} v_h\|_{1,k}^2 \right)^{1/2} . \qquad (4.12)\end{aligned}$$

The result follows on combining (4.11) with (4.12). □

Remark 4.4. *The use of the estimate*

$$\Im\left(\Theta a_\varepsilon(v,v)\right) \;\gtrsim\; \frac{|\varepsilon|}{k^2} \|v\|_{1,k}^2,$$

which follows from (2.12), is crucial in the proof of Theorem 4.3. Indeed, the above proof uses the linearity of the imaginary part of $a(\cdot,\cdot)$ with respect to the second argument. The cruder estimate

$$|a_\varepsilon(v,v)| \;\gtrsim\; \frac{|\varepsilon|}{k^2} \|v\|_{1,k}^2,$$

which also follows from (2.12), could not be used to prove Theorem 4.3.

Lemma 4.5.

$$\left(1 + \frac{H}{\delta}\right)^{1/2} \left(\sum_{\ell=0}^{N} \|Q_{\varepsilon,\ell} v_h\|_{1,k}^2 \right)^{1/2} \;\gtrsim\; \frac{|\varepsilon|}{k^2} \|v_h\|_{1,k} \qquad \textit{for all } v_h \in \mathcal{V}^h.$$

Proof. We first recall the decomposition of v_h as given in Lemma 4.1. Then, using Lemma 2.4, the definition of $Q_{\varepsilon,\ell}$ and Lemma 2.1, we obtain:

$$\begin{aligned}\frac{|\varepsilon|}{k^2}\|v_h\|_{1,k}^2 \lesssim \Im\left[\Theta a_\varepsilon(v_h,v_h)\right] &= \sum_{l=0}^{N}\Im\left[\Theta a_\varepsilon(v_h,v_l)\right]\\ &= \sum_{l=0}^{N}\Im\left[\Theta a_\varepsilon(Q_{\varepsilon,\ell}v_h,v_l)\right] \lesssim \sum_{l=0}^{N}\|Q_{\varepsilon,\ell}v_h\|_{1,k}\|v_l\|_{1,k}\,.\end{aligned}$$

Then applying the Cauchy-Schwarz inequality and Lemma 4.1 yields

$$\frac{|\varepsilon|}{k^2}\|v_h\|_{1,k}^2 \lesssim \left(\sum_{\ell=0}^{N}\|Q_{\varepsilon,\ell}v_h\|_{1,k}^2\right)^{1/2}\left(\sum_{\ell=0}^{N}\|v_\ell\|_{1,k}^2\right)^{1/2} \lesssim \left(1+\frac{H}{\delta}\right)^{1/2}\left(\sum_{\ell=0}^{N}\|Q_{\varepsilon,\ell}v_h\|_{1,k}^2\right)^{1/2}\|v_h\|_{1,k}.$$

□

Our next key result (Lemma 4.10 below) is an estimate for the L^2-error in the coarse space projection operator $Q_{\varepsilon,0}$; this is crucially needed to get good estimates for the two-grid preconditioner represented by Q_ε. In order to prove this result we need to make an assumption about the approximability on the coarse grid of the solution of the adjoint problem.

Assumption 4.6 (Coarse-grid approximability of the adjoint problem)**.** *If ϕ is the solution of the adjoint problem*

$$-\Delta\phi-(k^2-\mathrm{i}\varepsilon)\phi=f \quad \text{on} \quad \Omega, \tag{4.13a}$$

$$\frac{\partial\phi}{\partial n}-\mathrm{i}\overline{\eta}\phi=0 \quad \text{on} \quad \Gamma\,, \tag{4.13b}$$

with $f\in L^2(\Omega)$, then

$$\inf_{\phi_0\in\mathcal{V}_0}\|\phi-\phi_0\|_{1,k} \lesssim kH\left(\frac{k}{|\varepsilon|}\right)\|f\|_{L^2(\Omega)}\,. \tag{4.13c}$$

Theorem 4.7. *Assumption 4.6 holds when Ω is a 2-d convex polygon, η satisfies (2.11), and the coarse grid is as described in §3 (with, in particular, H denoting the mesh diameter).*

Proof. If ϕ satisfies (4.13) and η satisfies (2.11), then the coercivity estimate (2.13) combined with the Lax–Milgram theorem implies that

$$\|\phi\|_{1,k} \lesssim \frac{k}{|\varepsilon|}\|f\|_{L^2(\Omega)}\,. \tag{4.14}$$

If Ω is a convex polygon, the regularity results in [28] can then be used to show that

$$\|\phi\|_{H^2(\Omega)} \lesssim \frac{k^2}{|\varepsilon|}\|f\|_{L^2(\Omega)}\,; \tag{4.15}$$

see [21, Lemma 2.12]. Now, with ϕ_0 denoting the Scott-Zhang quasi-interpolant on the coarse grid, we have

$$\inf_{\phi_0\in\mathcal{V}_0}\|\phi-\phi_0\|_{1,k} \lesssim H\|\phi\|_{H^2(\Omega)}+kH\|\phi\|_{H^1(\Omega)} \tag{4.16}$$

[37, Theorem 4.1], and the result (4.13c) follows from combining (4.14), (4.15), and (4.16). □

Remark 4.8 (Bounds on the adjoint problem)**.** *(i) In the proof of Theorem 4.7, we obtained the bound (4.14) from coercivity and the Lax–Milgram theorem. This bound can also be obtained from an argument involving Green's identity (with the latter giving better estimates in the case of an inhomogeneous boundary condition); see* [21, Remark 2.5] *(but note that the η in (2.3b) of that paper should be $\overline{\eta}$).*

(ii) The bounds (4.14) and (4.15) are the best currently-available bounds on the solution of (4.13) for $\varepsilon\gtrsim k$, but they are not optimal when $\varepsilon\ll k$ – see [21, Theorem 2.9].

Remark 4.9 (Establishing Assumption 4.6 for more general domains)**.** *(i) H^2-regularity of the Laplacian on convex polyhedra with homogeneous Dirichlet boundary conditions is proved in* [10, Corollary 18.18]. *The analogous result for inhomogeneous Neumann boundary conditions could then be used, following* [21, Lemma 2.12]*, to prove that* (4.15) *(and thus also Assumption 4.6) holds for the solution of* (4.13) *on convex polyhedra with quasi-uniform meshes.*

(ii) When Ω is a bounded, non-convex Lipschitz polyhedron in $\mathbb{R}^d$, $d = 2, 3$, it is natural to use a sequence of locally-refined meshes. In this case we expect Assumption 4.6 to hold where H is replaced by $(1/N)^{1/d}$, where N is the dimension of the subspace (so $(1/N)^{1/d}$ is the largest element diameter). The steps to prove this are outlined in [21, Assumption 3.7, Remark 3.8].

We now use Assumption 4.6 to prove the key lemma on the approximation power of $Q_{\varepsilon,0}$ measured in the L^2-norm on the domain.

Lemma 4.10. *(Estimate for $Q_{\varepsilon,0}$) For all $v \in H^1(\Omega)$,*

$$\|(I - Q_{\varepsilon,0})v\|_{L^2(\Omega)} \lesssim kH \left(\frac{k}{|\varepsilon|}\right) \|(I - Q_{\varepsilon,0})v\|_{1,k} \,. \tag{4.17}$$

Proof. In the proof, for simplicity, we write Q_0 instead of $Q_{\varepsilon,0}$. Recall that Q_0 is defined by the variational problem $a_\varepsilon(Q_0 v, w) = a_\varepsilon(v, w)$, for all $w \in \mathcal{V}_0$, and thus $e_0 := (I - Q_0)v$ satifies

$$a_\varepsilon(e_0, w) = 0 \quad \text{for all } w \in \mathcal{V}_0, \tag{4.18}$$

Let ϕ be the solution of the adjoint problem

$$-\Delta\phi - (k^2 - \mathrm{i}\varepsilon)\phi = e_0 \quad \text{on} \quad \Omega,$$
$$\frac{\partial\phi}{\partial n} + \mathrm{i}\overline{\eta}\phi = 0 \quad \text{on} \quad \Gamma.$$

Then, for all $w \in H^1(\Omega)$, we have $a_\varepsilon(w, \phi) = (w, e_0)_{L^2(\Omega)}$. Hence, using (4.18), we can write

$$\|e_0\|^2_{L^2(\Omega)} \;=\; |a_\varepsilon(e_0, \phi)| \;=\; |a_\varepsilon(e_0, \phi - \phi_0)| \tag{4.19}$$

for any $\phi_0 \in \mathcal{V}_0$. Now, by Assumption 4.6, there exists a $\phi_0 \in \mathcal{V}_0$ such that

$$\|\phi - \phi_0\|_{1,k} \lesssim kH \left(\frac{k}{|\varepsilon|}\right) \|f\|_{L^2(\Omega)} \,.$$

Therefore, using this last bound and continuity, we have

$$|a_\varepsilon(e_0, \phi - \phi_0)| \lesssim \|e_0\|_{1,k}\|\phi - \phi_0\|_{1,k} \lesssim \|e_0\|_{1,k}(kH) \left(\frac{k}{|\varepsilon|}\right) \|e_0\|_{L^2(\Omega)} \,, \tag{4.20}$$

and combining (4.20) and (4.19) we obtain (4.17). □

In what follows, we need both the Poincaré–Friedrichs inequality and the trace inequality on domains D of *characteristic length scale* L. By this we mean that D is assumed to have diameter $\sim L$, surface area $\sim L^{d-1}$ and volume $\sim L^d$. The estimates in the next two results are then explicit in L (with the hidden constants independent of L).

Theorem 4.11. *If D is a Lipschitz domain with characteristic length scale L, then the Poincaré-Friedrichs inequality is*

$$\|v\|_{L^2(D)} \;\lesssim\; L|v|_{H^1(D)}, \tag{4.21}$$

for all $v \in H^1(D)$ that vanish on a subset of ∂D with measure $\sim L^{d-1}$, and the multiplicative trace inequality is

$$\|v\|^2_{L^2(\partial D)} \;\lesssim\; \left(L^{-1}\|v\|_{L^2(D)} + |v|_{H^1(D)}\right)\|v\|_{L^2(D)} \,, \quad \textit{for all} \quad v \in H^1(D) \,. \tag{4.22}$$

Proof. For domains of size $\mathcal{O}(1)$, (4.21) is proved in, e.g., [34, Theorem 1.9], and (4.22) is proved in [28, Last equation on p. 41]. A scaling argument then yields (4.21) and (4.22). □

Combining (4.21) and (4.22) we obtain the following corollary.

Corollary 4.12. *If D is a Lipschitz domain with characteristic length scale L and v vanishes on a subset of ∂D of measure $\sim L^{d-1}$, then*

$$\|v\|_{L^2(\partial D)} \;\lesssim\; L^{1/2}|u|_{H^1(D)} \,. \tag{4.23}$$

At various places in this paper we make use of the simple "Cauchy inequality":

$$2ab \le \delta a^2 + \frac{b^2}{\delta}, \quad a, b, \delta > 0. \tag{4.24}$$

In particular, using this (with $\delta = 1$) and the multiplicative trace inequality (4.22), we obtain another corollary to Theorem 4.11.

Corollary 4.13. *If D is a Lipschitz domain (with characteristic length scale $\mathcal{O}(1)$) then*

$$k^{1/2}\|v\|_{L^2(\partial D)} \lesssim \|v\|_{1,k}\ , \quad \text{for all} \quad v \in H^1(D) \quad \text{and} \quad k \ge 1. \tag{4.25}$$

Our goal for the rest of the section is to bound the field of values $(v_h, Q_\varepsilon v_h)_{1,k}/\|v_h\|_{1,k}^2$ away from the origin in the complex plane. (Note that the field of values is computed with respect to the $(\cdot,\cdot)_{1,k}$ inner product.) We do this by estimating $|(v_h, Q_\varepsilon v_h)_{1,k}|$ below by $\sum_{l=0}^N \|Q_{\varepsilon,l} v_h\|_{1,k}^2$ plus "remainder" terms (which turn out to be higher order, i.e. bounded by a positive power of H or $H_{\rm sub}$), and then use Lemma 4.5 to bound the sum below by $\|v_h\|_{1,k}^2$. Lemma 4.14 sets up the "remainder" terms, $R_{\varepsilon,\ell}(v_h)$, Lemmas 4.15 and 4.16 estimate these, and the final result is then given in Theorem 4.17.

Lemma 4.14. *For $\ell = 0, \dots, N$, set*

$$R_{\varepsilon,\ell}(v_h) \ := \ ((I - Q_{\varepsilon,\ell})v_h, Q_{\varepsilon,\ell} v_h)_{1,k}\,. \tag{4.26}$$

Then

$$(v_h, Q_\varepsilon v_h)_{1,k} = \sum_{\ell=0}^N \left\{ \|Q_{\varepsilon,\ell} v_h\|_{1,k}^2 + R_{\varepsilon,\ell}(v_h) \right\}. \tag{4.27}$$

Furthermore, $R_{\varepsilon,\ell}$ satisfies

$$|R_{\varepsilon,\ell}(v_h)| \ \lesssim \ D_{\varepsilon,\ell}(v_h) + B_{\varepsilon,\ell}(v_h), \tag{4.28}$$

where the "domain" and "boundary" contributions to the bound are given by

$$D_{\varepsilon,\ell}(v_h) = k^2 \|(I - Q_{\varepsilon,\ell})v_h\|_{L^2(\Omega_\ell)} \|Q_{\varepsilon,\ell} v_h\|_{L^2(\Omega_\ell)}, \tag{4.29}$$

$$B_{\varepsilon,\ell}(v_h) = k \|(I - Q_{\varepsilon,\ell})v_h\|_{L^2(\Gamma_\ell)} \|Q_{\varepsilon,\ell} v_h\|_{L^2(\Gamma_\ell)}. \tag{4.30}$$

and $\Omega_0 = \Omega$, $\Gamma_0 = \Gamma$, and $\Gamma_\ell = \Gamma \cap \partial\Omega_\ell$, for $\ell = 1, \dots, N$.

Proof. By the definition of Q_ε,

$$(v_h, Q_\varepsilon v_h)_{1,k} = \sum_{\ell=0}^N (v_h, Q_{\varepsilon,\ell} v_h)_{1,k} = \sum_{\ell=0}^N \left\{ \|Q_{\varepsilon,\ell} v_h\|_{1,k}^2 + ((I - Q_{\varepsilon,\ell})v_h, Q_{\varepsilon,\ell} v_h)_{1,k} \right\}\ ,$$

yielding (4.27). To obtain (4.28), we recall that.

$$(u, v)_{1,k} = a_\varepsilon(u, v) + (2k^2 + \mathrm{i}\varepsilon)(u, v)_{L^2(\Omega)} + \mathrm{i}\eta(u, v)_{L^2(\Gamma)}.$$

Then, since

$$a_\varepsilon((I - Q_{\varepsilon,\ell})v_h, Q_{\varepsilon,\ell} v_h) = 0$$

(from the definition of $Q_{\varepsilon,\ell}$ (4.10)), we have

$$R_{\varepsilon,\ell}(v_h) = (2k^2 + \mathrm{i}\varepsilon)((I - Q_{\varepsilon,\ell})v_h, Q_{\varepsilon,\ell} v_h)_{L^2(\Omega_\ell)} + \mathrm{i}\eta((I - Q_{\varepsilon,\ell})v_h, Q_{\varepsilon,\ell} v_h)_{L^2(\Gamma_\ell)},$$

where we have also used the fact that $Q_{\varepsilon,\ell} v_h$ has support only on Ω_l. The desired "domain" and "boundary" estimates (4.29) and (4.30) then follow after using (2.6). ☐

We now bound $D_{\varepsilon,\ell}(v_h)$ and $B_{\varepsilon,\ell}(v_h)$, using the following strategy. First Lemma 4.15 bounds $D_{\varepsilon,0}(v_h)$ in terms of a positive power of H, which is obtained by using Lemma 4.10 to estimate the $\|(I - Q_{\varepsilon,0})v_h\|_{L^2(\Omega)}$ component of $D_{\varepsilon,0}(v_h)$. Then, in Lemma 4.16 we bound $\sum_{\ell=1}^N D_{\varepsilon,\ell}(v_h)$ in terms of a positive power of $H_{\rm sub}$, by applying the Poincaré–Friedrichs inequality (4.21) to each of the $\|Q_{\varepsilon,\ell} v_h\|_{L^2(\Omega_\ell)}$ terms in this sum. These two lemmas also provide bounds on $B_{\varepsilon,0}(v_h)$ and $\sum_{\ell=1}^N B_{\varepsilon,\ell}(v_h)$, respectively, where similar ideas are used, except this time in conjunction with trace inequalities. Recalling that H is the coarse mesh diameter and $H_{\rm sub}$ the subdomain diameter, we are then able to control the error terms by making H and $H_{\rm sub}$ sufficiently small (Theorem 4.17); it turns out that the required condition on H is more stringent than that on $H_{\rm sub}$.

Lemma 4.15 (Bounds on $D_{\varepsilon,0}$ and $B_{\varepsilon,0}$)**.** *For any $\alpha, \alpha' \geq 0$ and any $v_h \in \mathcal{V}^h$,*

$$D_{\varepsilon,0}(v_h) \lesssim kH \left(\frac{k^2}{|\varepsilon|}\right) \left[\left(\frac{k^2}{|\varepsilon|}\right)^{\alpha} \|Q_{\varepsilon,0} v_h\|_{1,k}^2 + \left(\frac{k^2}{|\varepsilon|}\right)^{-\alpha} \|v_h\|_{1,k}^2\right], \tag{4.31}$$

$$B_{\varepsilon,0}(v_h) \lesssim (kH)^{1/2} \left(\frac{k^2}{|\varepsilon|}\right)^{1/2} \left[\left(\frac{k^2}{|\varepsilon|}\right)^{\alpha'} \|Q_{\varepsilon,0} v_h\|_{1,k}^2 + \left(\frac{k^2}{|\varepsilon|}\right)^{-\alpha'} \|v_h\|_{1,k}^2\right], \tag{4.32}$$

and thus (taking $\alpha' = \alpha$)

$$\begin{aligned} &D_{\varepsilon,0}(v_h) + B_{\varepsilon,0}(v_h) \\ &\lesssim (kH)^{1/2} \left(\frac{k^2}{|\varepsilon|}\right)^{1/2} \left[1 + (kH)^{1/2} \left(\frac{k^2}{|\varepsilon|}\right)^{1/2}\right] \left[\left(\frac{k^2}{|\varepsilon|}\right)^{\alpha} \|Q_{\varepsilon,0} v_h\|_{1,k}^2 + \left(\frac{k^2}{|\varepsilon|}\right)^{-\alpha} \|v_h\|_{1,k}^2\right]. \end{aligned} \tag{4.33}$$

Proof. For the bound on $D_{\varepsilon,0}(v_h)$, we use (4.17), the triangle inequality, and the Cauchy inequality (4.24) to obtain

$$\begin{aligned} D_{\varepsilon,0}(v_h) &\lesssim kH \left(\frac{k^2}{|\varepsilon|}\right) \|(I - Q_{\varepsilon,0}) v_h\|_{1,k} \|Q_{\varepsilon,0} v_h\|_{1,k}, \\ &\lesssim kH \left(\frac{k^2}{|\varepsilon|}\right) \left[\|Q_{\varepsilon,0} v_h\|_{1,k}^2 + \|Q_{\varepsilon,0} v_h\|_{1,k} \|v_h\|_{1,k}\right], \qquad (4.34) \\ &\lesssim kH \left(\frac{k^2}{|\varepsilon|}\right) \left[\|Q_{\varepsilon,0} v_h\|_{1,k}^2 + \left(\frac{k^2}{|\varepsilon|}\right)^{\alpha} \|Q_{\varepsilon,0} v_h\|_{1,k}^2 + \left(\frac{k^2}{|\varepsilon|}\right)^{-\alpha} \|v_h\|_{1,k}^2\right], \qquad (4.35) \end{aligned}$$

for any $\alpha \geq 0$. Since $|\varepsilon| \lesssim k^2$, (4.31) follows.

For the bound on $B_{\varepsilon,0}(v_h)$, we apply the multiplicative trace inequality (4.22) on Ω (so $L \sim 1$), to obtain

$$B_{\varepsilon,0}(v_h) \lesssim k \|(I - Q_{\varepsilon,0}) v_h\|_{L^2(\Omega)}^{1/2} \|(I - Q_{\varepsilon,0}) v_h\|_{H^1(\Omega)}^{1/2} \|Q_{\varepsilon,0} v_h\|_{L^2(\Gamma)}.$$

Using (4.17), we then have

$$B_{\varepsilon,0}(v_h) \lesssim k (kH)^{1/2} \left(\frac{k}{|\varepsilon|}\right)^{1/2} \|(I - Q_{\varepsilon,0}) v_h\|_{1,k}^{1/2} \|(I - Q_{\varepsilon,0}) v_h\|_{H^1(\Omega)}^{1/2} \|Q_{\varepsilon,0} v_h\|_{L^2(\Gamma)}$$

and then using (4.25) and the triangle inequality we obtain

$$\begin{aligned} B_{\varepsilon,0}(v_h) &\lesssim (kH)^{1/2} \left(\frac{k^2}{|\varepsilon|}\right)^{1/2} \|(I - Q_{\varepsilon,0}) v_h\|_{1,k} \|Q_{\varepsilon,0} v_h\|_{1,k} \\ &\lesssim (kH)^{1/2} \left(\frac{k^2}{|\varepsilon|}\right)^{1/2} \left[\|Q_{\varepsilon,0} v_h\|_{1,k}^2 + \|v_h\|_{1,k} \|Q_{\varepsilon,0} v_h\|_{1,k}\right]. \end{aligned}$$

This last inequality is the analogue of (4.34), and proceeding as before we obtain (4.32). □

Lemma 4.16 (Bounds on $\sum D_{\varepsilon,\ell}$, $\sum B_{\varepsilon,\ell}$)**.** *For any $\alpha, \alpha' \geq 0$ and any $v_h \in \mathcal{V}^h$,*

$$\sum_{\ell=1}^{N} D_{\varepsilon,\ell}(v_h) \lesssim kH_{sub} \left[\left(\frac{k^2}{|\varepsilon|}\right)^{\alpha} \sum_{\ell=1}^{N} \|Q_{\varepsilon,l} v_h\|_{1,k}^2 + \left(\frac{k^2}{|\varepsilon|}\right)^{-\alpha} \|v_h\|_{1,k}^2\right] \tag{4.36}$$

and

$$\sum_{\ell=1}^{N} B_{\varepsilon,\ell}(v_h) \lesssim kH_{sub} \left[\left(\frac{k^2}{|\varepsilon|}\right)^{\alpha'} \sum_{\ell=1}^{N} \|Q_{\varepsilon,l} v_h\|_{1,k}^2 + \left(\frac{k^2}{|\varepsilon|}\right)^{-\alpha'} \|v_h\|_{1,k}^2\right]. \tag{4.37}$$

Therefore (letting $\alpha' = \alpha$),

$$\sum_{\ell=1}^{N} \left[D_{\varepsilon,\ell}(v_h) + B_{\varepsilon,\ell}(v_h)\right] \lesssim kH_{sub} \left[\left(\frac{k^2}{|\varepsilon|}\right)^{\alpha} \sum_{\ell=1}^{N} \|Q_{\varepsilon,l} v_h\|_{1,k}^2 + \left(\frac{k^2}{|\varepsilon|}\right)^{-\alpha} \|v_h\|_{1,k}^2\right]. \tag{4.38}$$

Proof. Let $\ell = 1, \ldots, N$. Recalling both that $Q_{\varepsilon,\ell} v_h$ vanishes on $\partial\Omega_\ell \backslash \Gamma$ and the assumption (3.2), we can use the Poincaré inequality (4.21) on Ω_ℓ, and then use the triangle inequality to obtain

$$\begin{aligned} D_{\varepsilon,\ell}(v_h) &\lesssim k^2 H_{\text{sub}} \left\| (I - Q_{\varepsilon,\ell}) v_h \right\|_{L^2(\Omega_\ell)} |Q_{\varepsilon,\ell} v_h|_{H^1(\Omega_l)}, \\ &\lesssim k H_{\text{sub}} \left[\|Q_{\varepsilon,\ell} v_h\|_{1,k}^2 + k \|v_h\|_{L^2(\Omega_\ell)} \|Q_{\varepsilon,\ell} v_h\|_{1,k} \right] \end{aligned}$$

(where the $1,k$-norm is over the support of $Q_{\varepsilon,\ell} v_h$, which is Ω_ℓ). Using (4.24) we obtain

$$D_{\varepsilon,\ell}(v_h) \lesssim k H_{\text{sub}} \left[\left(\frac{k^2}{|\varepsilon|} \right)^{\alpha} \|Q_{\varepsilon,\ell} v_h\|_{1,k}^2 + k^2 \left(\frac{k^2}{|\varepsilon|} \right)^{-\alpha} \|v_h\|_{L^2(\Omega_\ell)}^2 \right],$$

with $\alpha \geq 0$. Summing from $\ell = 1$ to N, and using the finite-overlap property (3.4), gives (4.36).

From (4.30) we have

$$B_{\varepsilon,\ell}(v_h) \lesssim k \left[\|Q_{\varepsilon,\ell} v_h\|_{L^2(\Gamma_\ell)}^2 + \|v_h\|_{L^2(\Gamma_\ell)} \|Q_{\varepsilon,\ell} v_h\|_{L^2(\Gamma_\ell)} \right]$$

and then using (4.23) we have

$$B_{\varepsilon,\ell}(v_h) \lesssim k \left[H_{\text{sub}} |Q_{\varepsilon,\ell} v_h|_{H^1(\Omega_\ell)}^2 + H_{\text{sub}}^{1/2} \|v\|_{L^2(\Gamma_\ell)} |Q_{\varepsilon,\ell} v_h|_{H^1(\Omega_\ell)} \right].$$

Summing from $\ell = 1$ to N we then obtain

$$\sum_{\ell=1}^N B_{\varepsilon,\ell}(v_h) \lesssim k H_{\text{sub}} \sum_{\ell=1}^N |Q_{\varepsilon,\ell} v_h|_{H^1(\Omega_\ell)}^2 + k H_{\text{sub}}^{1/2} \sum_{\ell=1}^N \|v_h\|_{L^2(\Gamma_\ell)} |Q_{\varepsilon,\ell} v_h|_{H^1(\Omega_\ell)}. \tag{4.39}$$

(Note that the sums in the last inequality could be restricted to those ℓ with $\Gamma \cap \partial\Omega_\ell \neq \emptyset$, but this is not used in the following.) Using the Cauchy-Schwarz inequality, then (4.23) and finally (4.24), we have

$$\begin{aligned} k H_{\text{sub}}^{1/2} \sum_{\ell=1}^N \|v_h\|_{L^2(\Gamma_l)} |Q_{\varepsilon,\ell} v_h|_{H^1(\Omega_\ell)} &\lesssim k H_{\text{sub}}^{1/2} \left(\sum_{\ell=1}^N \|v_h\|_{L^2(\Gamma_\ell)}^2 \right)^{1/2} \left(\sum_{\ell=1}^N |Q_{\varepsilon,\ell} v_h|_{H^1(\Omega_\ell)}^2 \right)^{1/2}, \\ &\lesssim k H_{\text{sub}} \|v_h\|_{1,k} \left(\sum_{\ell=1}^N \|Q_{\varepsilon,\ell} v_h\|_{1,k}^2 \right)^{1/2}, \\ &\lesssim k H_{\text{sub}} \left[\left(\frac{k^2}{|\varepsilon|} \right)^{\alpha'} \sum_{\ell=1}^N \|Q_{\varepsilon,\ell} v_h\|_{1,k}^2 + \left(\frac{k^2}{|\varepsilon|} \right)^{-\alpha'} \|v_h\|_{1,k}^2 \right]. \end{aligned} \tag{4.40}$$

Inserting (4.40) into (4.39), we obtain the result (4.37). ∎

Our main result in the rest of this section is the following estimate from below on the field of values of Q_ε.

Theorem 4.17 (Bound below on the field of values)**.** *There exists a constant $\mathcal{C}_1 > 0$ such that*

$$|(v_h, Q_\varepsilon v_h)_{1,k}| \gtrsim \left(1 + \frac{H}{\delta} \right)^{-1} \left(\frac{|\varepsilon|}{k^2} \right)^2 \|v_h\|_{1,k}^2, \quad \textit{for all} \quad v_h \in \mathcal{V}^h, \tag{4.41}$$

when

$$\max \left\{ k H_{sub},\ k H \left(1 + \frac{H}{\delta} \right) \left(\frac{k^2}{|\varepsilon|} \right)^2 \right\} \leq \mathcal{C}_1 \left(1 + \frac{H}{\delta} \right)^{-1} \left(\frac{|\varepsilon|}{k^2} \right). \tag{4.42}$$

Note that the condition on the coarse mesh diameter H_{sub} is more stringent than the condition on the subdomain diameter H; one finds similar criteria in domain-decomposition theory for coercive elliptic PDEs; see, e.g., [27].

The following corollary restricts attention to a commonly encountered situation.

Corollary 4.18. *Suppose $\delta \sim H_{sub} \sim H$. There exists a constant $\mathcal{C}_1 > 0$ such that*

$$|(v_h, Q_\varepsilon v_h)_{1,k}| \gtrsim \left(\frac{|\varepsilon|}{k^2} \right)^2 \|v_h\|_{1,k}^2, \quad \textit{for all} \quad v_h \in \mathcal{V}^h, \tag{4.43}$$

when

$$kH \;\leq\; \mathcal{C}_1 \left(\frac{|\varepsilon|}{k^2}\right)^3.$$

Proof of Theorem 4.17. By Lemma 4.14,

$$|(v_h, Q_\varepsilon v_h)_{1,k}| \;\gtrsim\; \sum_{\ell=0}^{N} \|Q_{\varepsilon,\ell} v_h\|_{1,k}^2 - \sum_{\ell=0}^{N} \left(D_{\varepsilon,\ell}(v_h) + B_{\varepsilon,\ell}(v_h)\right).$$

Then, using the bounds (4.33) and (4.38) we have

$$\begin{aligned}
|(v_h, Q_\varepsilon v_h)_{1,k}| \;\gtrsim\; & \sum_{\ell=0}^{N} \|Q_{\varepsilon,\ell} v_h\|_{1,k}^2 \\
& - (kH)^{1/2} \left(\frac{k^2}{|\varepsilon|}\right)^{1/2} \left(1 + (kH)^{1/2} \left(\frac{k^2}{|\varepsilon|}\right)^{1/2}\right) \left[\left(\frac{k^2}{|\varepsilon|}\right)^{\alpha} \|Q_{\varepsilon,0} v_h\|_{1,k}^2 + \left(\frac{k^2}{|\varepsilon|}\right)^{-\alpha} \|v_h\|_{1,k}^2\right] \\
& - (kH_{\rm sub}) \left[\left(\frac{k^2}{|\varepsilon|}\right)^{\alpha'} \sum_{\ell=1}^{N} \|Q_{\varepsilon,\ell} v_h\|_{1,k}^2 + \left(\frac{k^2}{|\varepsilon|}\right)^{-\alpha'} \|v_h\|_{1,k}^2\right]
\end{aligned}$$

for $\alpha, \alpha' \geq 0$. Therefore, there exist $C_1, C_2 > 0$ (sufficiently small) such that

$$(kH)^{1/2} \left(\frac{k^2}{|\varepsilon|}\right)^{(1+2\alpha)/2} \left(1 + (kH)^{1/2} \left(\frac{k^2}{|\varepsilon|}\right)^{1/2}\right) \;\leq\; C_1, \tag{4.44}$$

and

$$(kH_{\rm sub}) \left(\frac{k^2}{|\varepsilon|}\right)^{\alpha'} \;\leq\; C_2 \tag{4.45}$$

ensure that

$$\begin{aligned}
|(v_h, Q_\varepsilon v_h)_{1,k}| \;\gtrsim\; & \sum_{\ell=0}^{N} \|Q_{\varepsilon,\ell} v_h\|_{1,k}^2 - (kH_{\rm sub}) \left(\frac{k^2}{|\varepsilon|}\right)^{-\alpha'} \|v_h\|_{1,k}^2 \\
& - (kH)^{1/2} \left(\frac{k^2}{|\varepsilon|}\right)^{1/2} \left(1 + (kH)^{1/2} \left(\frac{k^2}{|\varepsilon|}\right)^{1/2}\right) \left(\frac{k^2}{|\varepsilon|}\right)^{-\alpha} \|v_h\|_{1,k}^2.
\end{aligned} \tag{4.46}$$

Since $\alpha \geq 0$, there exists a $\widetilde{C}_1 > 0$ such that

$$(kH)^{1/2} \left(\frac{k^2}{|\varepsilon|}\right)^{(1+2\alpha)/2} \;\leq\; \widetilde{C}_1 \tag{4.47}$$

ensures that (4.44) holds; i.e. (4.46) holds under (4.47) and (4.45).

Using in (4.46) the bound in Lemma 4.5, we obtain

$$\begin{aligned}
|(v_h, Q_\varepsilon v_h)_{1,k}| \;\gtrsim\; & \left(1 + \frac{H}{\delta}\right)^{-1} \left(\frac{|\varepsilon|}{k^2}\right)^2 \|v_h\|_{1,k}^2 - (kH_{\rm sub}) \left(\frac{k^2}{|\varepsilon|}\right)^{-\alpha'} \|v_h\|_{1,k}^2 \\
& - (kH)^{1/2} \left(\frac{k^2}{|\varepsilon|}\right)^{1/2} \left(1 + (kH)^{1/2} \left(\frac{k^2}{|\varepsilon|}\right)^{1/2}\right) \left(\frac{k^2}{|\varepsilon|}\right)^{-\alpha} \|v_h\|_{1,k}^2.
\end{aligned}$$

Therefore, there exist $C_3, C_4 > 0$ (sufficiently small) so that the conditions

$$(kH)^{1/2} \left(\frac{k^2}{|\varepsilon|}\right)^{(1-2\alpha)/2} \left(1 + (kH)^{1/2} \left(\frac{k^2}{|\varepsilon|}\right)^{1/2}\right) \;\leq\; C_3 \left(1 + \frac{H}{\delta}\right)^{-1} \left(\frac{|\varepsilon|}{k^2}\right)^2 \tag{4.48}$$

and

$$(kH_{\rm sub}) \left(\frac{k^2}{|\varepsilon|}\right)^{-\alpha'} \;\leq\; C_4 \left(1 + \frac{H}{\delta}\right)^{-1} \left(\frac{|\varepsilon|}{k^2}\right)^2, \tag{4.49}$$

together with (4.46) and (4.47), ensure that the result (4.41) holds.

Now, condition (4.48) can be rewritten as

$$(kH)^{1/2}\left(\frac{k^2}{|\varepsilon|}\right)^{1/2}\left(1+(kH)^{1/2}\left(\frac{k^2}{|\varepsilon|}\right)^{1/2}\right) \;\le\; C_3\left(1+\frac{H}{\delta}\right)^{-1}\left(\frac{|\varepsilon|}{k^2}\right)^{2-\alpha}.$$

Now, since $\varepsilon \lesssim k^2$,

$$\left(1+\frac{H}{\delta}\right)^{-1}\left(\frac{|\varepsilon|}{k^2}\right)^{\beta} \lesssim 1$$

for any $\beta \ge 0$. Therefore, if $\alpha \le 2$, then there exists a $\widetilde{C_3} > 0$ such that the condition (4.48) is ensured by the condition

$$(kH)^{1/2}\left(\frac{k^2}{|\varepsilon|}\right)^{1/2} \le \widetilde{C}_3\left(1+\frac{H}{\delta}\right)^{-1}\left(\frac{|\varepsilon|}{k^2}\right)^{2-\alpha}.$$

i.e.

$$(kH)^{1/2}\left(\frac{k^2}{|\varepsilon|}\right)^{(5-2\alpha)/2} \le \widetilde{C_3}\left(1+\frac{H}{\delta}\right)^{-1}. \tag{4.50}$$

In summary (from (4.47), (4.45), (4.50), and (4.49)) we have shown that there exist $\widetilde{C}_1, C_2, \widetilde{C}_3, C_4 > 0$ such that the required result (4.41), holds if the following four conditions hold:

$$(kH)^{1/2}\left(\frac{k^2}{|\varepsilon|}\right)^{(1+2\alpha)/2} \;\le\; \widetilde{C}_1, \tag{4.51}$$

$$(kH_{\rm sub})\left(\frac{k^2}{|\varepsilon|}\right)^{\alpha'} \;\le\; C_2, \tag{4.52}$$

$$(kH)^{1/2}\left(\frac{k^2}{|\varepsilon|}\right)^{(5-2\alpha)/2}\left(1+\frac{H}{\delta}\right) \le \widetilde{C}_3, \tag{4.53}$$

and

$$(kH_{\rm sub})\left(\frac{k^2}{|\varepsilon|}\right)^{2-\alpha'}\left(1+\frac{H}{\delta}\right) \le C_4, \tag{4.54}$$

where $0 \le \alpha \le 2$ and $\alpha' \ge 0$.

The optimal choice of α to balance the exponents in (4.51) and (4.53) (ignoring the factor $(1+H/\delta)$) is $\alpha = 1$, and the optimal choice of α' to balance the exponents in (4.52) and (4.54) (again ignoring $(1+H/\delta)$) is $\alpha' = 1$. With these values of α and α', the four conditions above are ensured by the condition (4.42). □

Remark 4.19 (One-level methods)**.** *Inspecting the proof of Theorem 4.17, we see that the bound from below on the field of values relies on the bound in Lemma 4.5, which in turn relies on the second bound in Lemma 4.1. In the case of the one-level method (i.e. A_ε is preconditioned with (3.1)), the constant on the right-hand side of the analogue of the second bound in (4.1) does not ~ 1 when $\delta \sim H$; instead it blows up as $H \to 0$. This is why we do not currently have a result analogous to Theorem 4.17 for the one-level method.*

5. Matrices and convergence of GMRES

In this section we interpret the results of Theorems 4.3 and 4.17 in terms of matrices and explain their implications for the convergence of GMRES for the Helmholtz equation. Let us begin by recalling the convergence theory for GMRES due originally to Elman [13] and Eisenstadt, Elman and Schultz [12], and used in the context of domain decomposition methods in [4]. The most convenient statement for our purposes is [1]. We consider any abstract linear system

$$C\mathbf{x} = \mathbf{d} \tag{5.1}$$

in $\mathbb{C}^n$, where C is an $n \times n$ nonsingular complex matrix. Choose an initial guess $\mathbf{x}^0$, introduce the residual $\mathbf{r}^0 = \mathbf{d} - C\mathbf{x}^0$ and the usual Krylov spaces:

$$\mathcal{K}^m(C, \mathbf{r}^0) := \operatorname{span}\{C^j\mathbf{r}^0 : j = 0, \ldots, m-1\}\ .$$

Let $\langle\cdot,\cdot\rangle_D$ denote the inner product on $\mathbb{C}^n$ induced by some Hermitian positive definite matrix D, i.e.

$$\langle \mathbf{V}, \mathbf{W}\rangle_D := \mathbf{W}^* D\mathbf{V} \tag{5.2}$$

with induced norm $\|\cdot\|_D$, where * denotes Hermitian transpose. For $m \geq 1$, define $\mathbf{x}^m$ to be the unique element of $\mathcal{K}^m$ satisfying the minimal residual property:

$$\|\mathbf{r}^m\|_D := \|\mathbf{d} - C\mathbf{x}^m\|_D \;=\; \min_{\mathbf{x}\in\mathcal{K}^m(C,\mathbf{r}^0)} \|\mathbf{d} - C\mathbf{x}\|_D,$$

When $D = I$ this is just the usual GMRES algorithm, and we write $\|\cdot\| = \|\cdot\|_I$, but for more general D it is the weighted GMRES method [19] in which case its implementation requires the application of the weighted Arnoldi process [29]. In §6 we give results for standard GMRES which corresponds to the case $D = I$ and also for a weighted variant with respect to a certain matrix D defined by (5.4) below. The following theorem is a simple generalisation of the classical convergence result stated in [1].

Theorem 5.1. *Suppose* $0 \notin W_D(C)$. *Then*

$$\frac{\|\mathbf{r}^m\|_D}{\|\mathbf{r}^0\|_D} \;\leq\; \sin^m(\beta)\;, \quad \textit{where} \quad \cos(\beta) := \frac{\operatorname{dist}(0, W_D(C))}{\|C\|_D}\;, \tag{5.3}$$

where $W_D(C)$ *denotes the* field of values *(also called the* numerical range *of* C*) with respect to the inner product induced by* D*, i.e.*

$$W_D(C) \;=\; \{\langle \mathbf{x}, C\mathbf{x}\rangle_D : \mathbf{x} \in \mathbb{C}^n, \|\mathbf{x}\|_D = 1\}.$$

Proof. For the "standard" case $D = I$ the result is stated in [1]. For general D, write $\widetilde{C} = D^{1/2}CD^{-1/2}$, $\widetilde{\mathbf{d}} = D^{1/2}\mathbf{d}$, $\widetilde{\mathbf{x}} = D^{1/2}\mathbf{x}$, $\widetilde{\mathbf{x}}^m = D^{1/2}\mathbf{x}^m$, and $\widetilde{\mathbf{r}^0} = D^{1/2}\mathbf{r}_0$. Then it is easy to see that $\widetilde{x}^m \in \mathcal{K}(\widetilde{C}, \widetilde{r}^0)$ and it satisfies the "standard" GMRES criterion for the transformed system but in the Euclidean norm:

$$\|\widetilde{\mathbf{r}}^m\| := \|\widetilde{\mathbf{d}} - \widetilde{C}\widetilde{\mathbf{x}}^m\| \;=\; \min_{\widetilde{\mathbf{x}}\in\mathcal{K}^m(\widetilde{C},\widetilde{\mathbf{r}^0})} \|\widetilde{\mathbf{d}} - \widetilde{C}\widetilde{\mathbf{x}}\|.$$

Then we know that the result (5.3) holds with $D = I$, $C = \widetilde{C}$ and $\mathbf{r}^m = \widetilde{\mathbf{r}}^m$. It is then simple to transform this back to obtain (5.3) in the case of general D. □

Remark 5.2. *Note that for all* $\mathbf{x} \in \mathbb{C}^n$ *with* $\|\mathbf{x}\|_D = 1$*, we have*

$$0 \;\leq\; \operatorname{dist}(0, W_D(C)) \;\leq\; |\langle \mathbf{x}, C\mathbf{x}\rangle_D| \leq \|C\|_D$$

and so the second formula in (5.3) *necessarily defines an angle* β *in the range* $[0, \pi/2]$*. Thus, for good GMRES convergence we aim to ensure that* $\operatorname{dist}(0, W_D(C))$ *is bounded well away from zero and that* $\|C\|_D$ *is as small as possible. Theorem 5.1 could therefore be viewed as a generalisation to the case of GMRES of the familiar condition number criterion for the convergence of the conjugate gradient method for positive definite systems. The result of Theorem 5.1 is stated without proof in* [4]*, with a reference to* [13]*; however* [13] *is concerned only with standard GMRES in the Euclidean inner product.*

Remark 5.3. *As we see in Theorems 5.6 and 5.8 below, the analysis of §4 provides us with estimates for the norm and field of values of the preconditioned matrix in the weighted norm induced by the real symmetric positive matrix* D_k *defined in* (5.4) *below. Other analyses of domain decomposition methods for non self-adjoint or non-positive definite PDEs (e.g.* [5], [36]*) have arrived at analogous estimates in weighted norms, although the weights appearing in these previous analyses are different, being associated with either the standard* H^1 *norm or semi-norm, and not the* k*- weighted energy norm, appropriate for Helmholtz problems, used here.*

We now use the theory in §4 to obtain results about the iterative solution of the linear systems arising from the Helmholtz equation. We start by interpreting the operators $Q_{\varepsilon,\ell}$ defined in (4.10) in terms of matrices.

Theorem 5.4. *Let* $v_h = \sum_{j\in\mathcal{I}^h} V_j\phi_j \;\in\; \mathcal{V}^h$*. Then*

$$(i)\quad Q_{\varepsilon,\ell}v_h \;=\; \sum_{j\in\mathcal{I}^h(\Omega_\ell)} \left(R_\ell^T A_{\varepsilon,\ell}^{-1} R_\ell A_\varepsilon \mathbf{V}\right)_j \phi_j\;, \quad \ell = 1,\ldots,N\;,$$

$$(ii) \quad Q_{\varepsilon,0} v_h \;=\; \sum_{p \in \mathcal{I}^H} \left(R_0^T A_{\varepsilon,0}^{-1} R_0 A_\varepsilon \mathbf{V} \right)_p \Phi_p \;,$$

with $A_{\varepsilon,\ell}$, $\ell = 0, \ldots, N$ *defined in* (3.1) *and* (3.6).

Proof. These results are similar to those for symmetric elliptic problems found for example in [42], so we will be brief. For (i), let $\ell \in \{1, \ldots, N\}$, let $w_{h,\ell}$ and $y_{h,\ell}$ be arbitrary elements of $\mathcal{V}_\ell$, and denote their coefficient vectors $\mathbf{W}$ and $\mathbf{Y}$ (with nodal values on all of $\mathcal{I}^h$). Then $\mathbf{W} = R_\ell^T \mathbf{w}$ and $\mathbf{Y} = R_\ell^T \mathbf{y}$, where $\mathbf{w}, \mathbf{y}$ have nodal values on $\mathcal{I}^h(\Omega_\ell)$. The definitions of A_ε and $A_{\varepsilon,\ell}$, (2.4) and (3.1), then imply that $a_\varepsilon(y_{h,\ell}, w_{h,\ell}) = \mathbf{W}^* A_\varepsilon \mathbf{Y} = \mathbf{w}^* A_{\varepsilon,\ell} \mathbf{y}$. So if $\mathbf{y} := A_{\varepsilon,\ell}^{-1} R_\ell A_\varepsilon \mathbf{V}$ for some $\mathbf{V} \in \mathbb{C}^n$, we have

$$a_\varepsilon(y_{h,\ell}, w_{h,\ell}) \;=\; \mathbf{w}^* R_\ell A_\varepsilon \mathbf{V} \;=\; (R_\ell^T \mathbf{w})^* A_\varepsilon \mathbf{V} = \; \mathbf{W}^* A_\varepsilon \mathbf{V} \;=\; a_\varepsilon(v_h, w_{h,\ell}),$$

where $y_{h,\ell}$ is the finite element function with nodal values $\mathbf{y}$. Thus, by definition of $Q_{\varepsilon,\ell}$, we have $y_{h,\ell} = Q_{\varepsilon,\ell} v_h$, which implies the result (i). The proof of (ii) is similar. □

The main results of the previous section - Theorems 4.3 and 4.17 - give estimates for the norm and the field of values of the operator Q_ε on the space $\mathcal{V}^h$, with respect to the inner product $(\cdot,\cdot)_{1,k}$ and its associated norm. In the following we translate these results into norm and field of values estimates for the preconditioned matrix $B_{\varepsilon,AS}^{-1} A_\varepsilon$ in the weighted inner product $\langle \cdot, \cdot \rangle_{D_k}$, where the weight matrix is :

$$D_k := S + k^2 M \;, \tag{5.4}$$

and S and M are defined in (2.5). In fact D_k is the matrix representing the $(\cdot,\cdot)_{1,k}$ inner product on the finite element space $\mathcal{V}^h$ in the sense that if $v_h, w_h \in \mathcal{V}^h$ with coefficient vectors $\mathbf{V}, \mathbf{W}$ then

$$(v_h, w_h)_{1,k} \;=\; \langle \mathbf{V}, \mathbf{W} \rangle_{D_k}. \tag{5.5}$$

Theorem 5.5. *Let* $v_h = \sum_{j \in \mathcal{I}^h} V_j \phi_h \;\in\; \mathcal{V}^h$. *Then*

$$\begin{aligned} (i) \quad (v_h, Q_\varepsilon v_h)_{1,k} \;&=\; \langle \mathbf{V}, B_{\varepsilon,AS}^{-1} A_\varepsilon \mathbf{V} \rangle_{D_k} \\ (ii) \quad \|Q_\varepsilon v_h\|_{1,k} \;&=\; \|B_{\varepsilon,AS}^{-1} A_\varepsilon \mathbf{V}\|_{D_k} \end{aligned}$$

Proof. For arbitrary $w_h, v_h \in \mathcal{V}^h$, with coefficient vectors $\mathbf{W}$ and $\mathbf{V}$, using Theorem 5.4, we have

$$(w_h, Q_{\varepsilon,\ell} v_h)_{1,k} \;=\; \langle \mathbf{W}, R_\ell^T A_{\varepsilon,\ell}^{-1} R_\ell A_\varepsilon \mathbf{V} \rangle_{D_k}, \qquad \ell = 0, \ldots, N.$$

Summing these over $\ell = 0, \ldots, N$ and using (3.1), (3.6), and (3.7), we obtain

$$(w_h, Q_\varepsilon v_h)_{1,k} \;=\; \langle \mathbf{W}, B_{\varepsilon,AS}^{-1} A_\varepsilon \mathbf{V} \rangle_{D_k},$$

from which (i) and (ii) follow immediately. □

The following main result now follows from Theorems 4.3, 4.17, and 5.5.

Theorem 5.6 (Main result for left preconditioning)**.**

$$(i) \qquad \|B_{\varepsilon,AS}^{-1} A_\varepsilon\|_{D_k} \;\lesssim\; \left(\frac{k^2}{|\varepsilon|} \right) \quad \textit{for all} \quad H, H_{sub}.$$

Furthermore, there exists a constant $\mathcal{C}_1$ *such that*

$$(ii) \qquad |\langle \mathbf{V}, B_{\varepsilon,AS}^{-1} A_\varepsilon \mathbf{V} \rangle_{D_k}| \;\gtrsim\; \left(1 + \frac{H}{\delta} \right)^{-1} \left(\frac{|\varepsilon|}{k^2} \right)^2 \|\mathbf{V}\|_{D_k}^2, \qquad \textit{for all } \mathbf{V} \in \mathbb{C}^n,$$

when

$$\max \left\{ kH_{sub},\; kH \left(1 + \frac{H}{\delta} \right) \left(\frac{k^2}{|\varepsilon|} \right)^2 \right\} \;\leq\; \mathcal{C}_1 \left(1 + \frac{H}{\delta} \right)^{-1} \left(\frac{|\varepsilon|}{k^2} \right). \tag{5.6}$$

Combining Theorem 5.1 and Theorem 5.6 we obtain:

Corollary 5.7 (GMRES convergence for left preconditioning)**.** *Consider the weighted GMRES method where the residual is minimised in the norm induced by D_k (see, e.g., [29]). Let $\mathbf{r}^m$ denote the mth iterate of GMRES applied to the system A_ε, left preconditioned with $B_{\varepsilon,AS}^{-1}$. Then*

$$\frac{\|\mathbf{r}^m\|_{D_k}}{\|\mathbf{r}^0\|_{D_k}} \lesssim \left(1-\left(1+\frac{H}{\delta}\right)^{-2}\left(\frac{|\varepsilon|}{k^2}\right)^6\right)^{m/2}, \tag{5.7}$$

provided condition (5.6) holds.

As a particular example of Corollary 5.7 we see that, provided $|\varepsilon| \sim k^2$, $H, H_{\rm sub} \sim k^{-1}$ and $\delta \sim H$, then GMRES will converge with the number of iterations independent of all parameters. This property is illustrated in the numerical experiments in the next section, all of which concern the case $\delta \sim H_{\rm sub} \sim H$. These experiments also explore the sharpnesss of the result (5.7) in the two cases: (i) $|\varepsilon|$ decreases below k^2 for fixed H and (ii) H increases above k^{-1} for fixed ε. Our experiments show that there may be room to improve the theoretical results.

While Corollary 5.7 provides rigorous estimates only for weighted GMRES, we see in §6 that there is, in fact, very little difference between the results with weighted GMRES and standard GMRES, so most of our experiments are for standard GMRES. The difficulty of proving results about standard GMRES in the context of domain decomposition for non self-adjoint problems was previously investigated by other researchers; see, e.g., [5].

In the next section we also explore the use of $B_{\varepsilon,AS}^{-1}$ as a preconditioner for A. A particularly effective preconditioner is obtained with $H \sim k^{-1}$ and $\varepsilon \approx k$. However this preconditioner has a complexity dominated by the cost of inverting the coarse mesh problem. A multilevel variant where the coarse problem is approximated by an inner GMRES iteration (within the FGMRES format) is also proposed and is demonstrated to be very efficient for solving finite element approximations of the Helmholtz equation with $h \sim k^{-3/2}$.

Some of our experiments below use right preconditioning rather than left preconditioning. Nevertheless, using the coerciveness for the adjoint form in Corollary 2.6, we can obtain the following result about right preconditioning, however in the inner product induced by D_k^{-1}. From this, the analogue of Corollary 5.7, with D_k replaced by D_k^{-1}, follows.

Theorem 5.8 (Main result for right preconditioning)**.** *With the same notation as in Theorem 4.17, we have*

$$(i) \qquad \|A_\varepsilon B_{\varepsilon,AS}^{-1}\|_{D_k^{-1}} \lesssim \left(\frac{k^2}{|\varepsilon|}\right) \quad \text{for all} \quad H, H_{sub}.$$

Furthermore, provided condition (5.6) holds,

$$(ii) \qquad |\langle \mathbf{V}, A_\varepsilon B_{\varepsilon,AS}^{-1}\mathbf{V}\rangle_{D_k^{-1}}| \gtrsim \left(1+\frac{H}{\delta}\right)^{-1}\left(\frac{|\varepsilon|}{k^2}\right)^2 \|\mathbf{V}\|_{D_k^{-1}}^2, \quad \text{for all } \mathbf{V} \in \mathbb{C}^n.$$

Proof. To simplify the notation, we write B_ε^{-1} instead of $B_{\varepsilon,AS}^{-1}$. An easy calculation shows that for all $\mathbf{V} \in \mathbf{C}^n$ and with $\mathbf{W} = D_k^{-1}\mathbf{V}$, we have

$$\frac{|\langle \mathbf{V}, A_\varepsilon B_\varepsilon^{-1}\mathbf{V}\rangle_{D_k^{-1}}|}{\langle \mathbf{V},\mathbf{V}\rangle_{D_k^{-1}}} = \frac{|\langle (B_\varepsilon^*)^{-1}A_\varepsilon^*\mathbf{W},\mathbf{W}\rangle_{D_k}|}{\langle \mathbf{W},\mathbf{W}\rangle_{D_k}} = \frac{|\langle \mathbf{W}, (B_\varepsilon^*)^{-1}A_\varepsilon^*\mathbf{W}\rangle_{D_k}|}{\langle \mathbf{W},\mathbf{W}\rangle_{D_k}},$$

where A^*, $(B_\varepsilon^*)^{-1}$ are the Hermitian transposes of A, B_ε^{-1} respectively. The coercivity of the adjoint form proved in Corollary 2.6 then ensures that the estimate in Theorem 5.6 (ii) also holds for the adjoint matrix and the result (ii) then follows. The result (i) is obtained analogously from taking the adjoint and using Theorem 5.6 (i). ☐

Corollary 5.9 (GMRES convergence for right preconditioning)**.** *Under the same assumptions, the result of Corollary 5.7 still holds when left preconditioning is replaced by right preconditioning.*

Remark 5.10 (The truncated sound-soft scattering problem)**.** *We now outline how the results can be adapted to hold for the truncated sound-soft scattering problem. By this, we mean the exterior, homogeneous Dirichlet problem, with the radiation condition imposed as an impedance boundary condition on a far-field boundary. That is,*

$$-\Delta u - (k^2 + \mathrm{i}\varepsilon)u = f \quad \text{in } \Omega, \tag{5.8a}$$

$$\frac{\partial u}{\partial n} - \mathrm{i}\eta u = g \quad \text{on } \partial\Omega_R, \tag{5.8b}$$

$$u = 0 \quad \text{on } \partial\Omega_D, \tag{5.8c}$$

where Ω_D is the scatterer and Ω_R is a bounded Lipschitz domain with $\Omega_D \subset \Omega_R$. With $f = 0$ and an appropriate choice of g, the solution of the above problem is a well-known approximation to the sound-soft scattering problem (see, e.g., [21, Problem 2.4] *for more details).*

The variational formulation of this problem is almost identical to that of the interior impedance problem in §2, except now the Hilbert space is $\{v \in H^1(\Omega) : v = 0 \text{ on } \partial\Omega_D\}$ and the integrals over Γ in (2.2) and (2.3) are over $\partial\Omega_R$. The essential Dirichlet boundary condition means that the nodes on $\partial\Omega_D$ are no longer freedoms. Thus, the domain decomposition technique for this problem are almost the same as those described in detail above, except that the subdomains will have Dirichlet conditions not only at interior boundaries, but also on any part of their boundary that intersects with $\partial\Omega_D$. The rest of the results in §2-§4 go through as before, and therefore analogues of Theorems 5.6 and 5.8 and Corollaries 5.7 and 5.9 hold for the truncated problem.

6. Numerical Experiments

Our numerical experiments discuss the solution of (2.4) on the unit square, with $\eta = k$, discretised by the continuous linear finite element method on a uniform triangular mesh of diameter h. The problem to be solved is thus specified by the choice of h and ε which we denote by

$$h_{\text{prob}} \quad \text{and} \quad \varepsilon_{\text{prob}} \,. \tag{6.1}$$

We will discuss the case $\varepsilon_{\text{prob}} > 0$ in Experiment 1 (with results in Tables 1, 2, 3); the empirical observations from these results are then used to motivate a preconditioner for the pure Helmholtz problem ($\varepsilon_{\text{prob}} = 0$) in Experiments 2 and 3 (with results in Tables 4, 5, 6, 7).

We will be focused on solving systems with $h_{\text{prob}} \sim k^{-3/2}$ (the discretisation level generally believed to remove the pollution effect; see, e.g., the literature reviews in [21, Remark 4.2] and [24, §1.2.2]), however the case $h_{\text{prob}} \sim k$ (a fixed number of grid points per wavelength) appears as a relevant subproblem when we construct multilevel methods in Experiment 3 below.

In the general theory given in §3, coarse grid size H and subdomain size H_{sub} are unrelated. but in our experiments here we construct local subdomains by taking each of the elements of the coarse grid and extending them to obtain an overlapping cover with overlap parameter δ. This is chosen as large as possible, but with the restriction no two extended subdomains can touch unless they came from touching elements of the original coarse grid. In this scenario $\delta \sim H$ (generous overlap), $H_{\text{sub}} \sim H$ and our preconditioners are thus determined by choices of H and ε, which we denote by

$$H_{\text{prec}} \quad \text{and} \quad \varepsilon_{\text{prec}} \,. \tag{6.2}$$

In our preconditioners the coarse grid problem is of size $\sim H_{\text{prec}}^{-2}$ and there are $\sim H_{\text{prec}}^{-2}$ local problems of size $(H_{\text{prec}}/h_{\text{prob}})^2$. If there were no overlap, the method would be "perfectly load balanced" (i.e. local problems of the same size as the coarse problem) when $H_{\text{prec}} = h_{\text{prob}}^{1/2}$. Thus, for load balancing,

$$H_{\text{prec}} \sim k^{-3/4} \quad \text{when} \quad h_{\text{prob}} \sim k^{-3/2}. \tag{6.3}$$

(Because of overlap, the local problems are larger than estimated in (6.3), and the method is in fact loadbalanced for somewhat finer coarse meshes than those predicted in (6.3). We will investigate both cases when $\varepsilon_{\text{prob}}$ and $\varepsilon_{\text{prec}}$ are equal and cases when $\varepsilon_{\text{prec}} > \varepsilon_{\text{prob}} = 0$. The question of greatest practical interest is: if $\varepsilon_{\text{prob}} = 0$, how to choose $\varepsilon_{\text{prec}}$ and H_{prec} in order to maximise the efficiency of the preconditioner? This question is addressed towards the end of these experiments. but first we illustrate the theoretical results in §5 which are about the case $\varepsilon_{\text{prec}} = \varepsilon_{\text{prob}} \neq 0$.

The first preconditioner considered is the *Classical Additive Schwarz* (`AS`) preconditioner defined in (3.7). We will also be interested in variants of this that replace the local component (3.1) with something else. The first variant involves averaging in the overlap of the subdomains. For each fine grid node x_j $(j \in \mathcal{I}^h)$, let L_j denote the number of subdomains which contain x_j. Then the

local operator is:

$$(B^{-1}_{\varepsilon,AVE,local}\mathbf{v})_j \;=\; \frac{1}{L_j}\sum_{\ell:\, x_j\in\Omega_\ell}\left(R_\ell^T A^{-1}_{\varepsilon,\ell}R_\ell\mathbf{v}\right)_j \;, \quad \text{for each} \quad j\in\mathcal{I}^h \;,$$

and the corresponding *Averaged Additive Schwarz* (`AVE`) preconditioner is:

$$B^{-1}_{\varepsilon,AVE} \;=\; R_0^T A^{-1}_{\varepsilon,0}R_0 \;+\; B^{-1}_{\varepsilon,AVE,local}. \tag{6.4}$$

The second variant is the *Restrictive Additive Schwarz* (`RAS`) preconditioner, which is well-known in the literature [3], [31]. Here to define the local operator, for each $j\in\mathcal{I}^h$, choose a single $\ell=\ell(j)$ with the property that $x_j\in\Omega^{\ell(j)}$. Then the action of the local contribution, for each vector of fine grid freedoms $\mathbf{v}$, is:

$$(B^{-1}_{\varepsilon,RAS,local}\mathbf{v})_j \;=\; \left(R^T_{\ell(j)}A^{-1}_{\varepsilon,\ell(j)}R_{\ell(j)}\mathbf{v}\right)_j \;, \quad \text{for each} \quad j\in\mathcal{I}^h \;, \tag{6.5}$$

and the RAS preconditioner is

$$B^{-1}_{\varepsilon,RAS} \;=\; R_0^T A^{-1}_{\varepsilon,0}R_0 \;+\; B^{-1}_{\varepsilon,RAS,local}\;. \tag{6.6}$$

All three of these variants of Additive Schwarz can be used in a hybrid way. This means that instead of doing all the local and coarse grid problems independently (and thus potentially in parallel), we first do a coarse solve and then perform the local solves on the residual of the coarse solve. This was first introduced in [32]. As described in [25], this is closely related to the deflation method [33], which has been used recently to good effect in the context of shifted Laplacian combined with multigrid [38]. We will show results for the Hybrid RAS (`HRAS`) preconditioner which takes the form

$$B^{-1}_{\varepsilon,HRAS} := R_0^T A^{-1}_{\varepsilon,0}R_0 + P_0^T\left(B^{-1}_{\varepsilon,RAS,local}\right)P_0\;, \tag{6.7}$$

where

$$P_0 \;=\; I - A_\varepsilon R_0^T A_0^{-1}R_0\;.$$

All our results are obtained with GMRES without restarts, with the implementation done in python. The results in Experiment 1 are obtained both with left preconditioning and with right preconditioning (flexible GMRES); the relevant python codes are `scipy.sparse.linalg.gmres` and `pyamg.krylov.fgmres` respectively. The other two experiments are done with right preconditioning. In all experiments we use standard GMRES, which minimises the residual in the standard Euclidean inner product. However, motivated by Theorem 5.6 and Corollary 5.7, Experiment 1 also discusses the results of applying left preconditioned GMRES in the inner product induced by the matrix D_k. For this we use the algorithm described in [29], editing an existing GMRES code to work with the inner product induced by D_k. In all experiments the starting guess is zero and the residual reduction tolerance is set at 10^{-6}.

6.1. **Experiment 1.** Here we solve (2.4) with $\mathbf{f}=\mathbf{1}$, $h_{\rm prob}=k^{-3/2}$,with various choices of $\varepsilon_{\rm prob}=\varepsilon_{\rm prec}$ and $H_{\rm prec}$. We use triangular coarse grids. The results in Section 5 tell us that, provided

$$\varepsilon_{\rm prob}=\varepsilon_{\rm prob}\sim k^2 \quad \text{and} \quad kH_{\rm prec}\sim 1,$$

then the number of GMRES iterations with the preconditioner `AS` will remain bounded as $k\to\infty$. Our first set of results are for the regular (Euclidean inner product) GMRES algorithm. Tables 1, 2 and 3 give results for a range of $\varepsilon_{\rm prob}=\varepsilon_{\rm prec}$ and $H_{\rm prec}$, assuming that $H_{\rm prec}=k^{-\alpha}$ for different choices of α. The number of iterations of the Classical Additive Schwarz method is denoted by $\#_{AS}$, while the number of iterations for the variants using averaging, RAS and Hybrid RAS are denoted $\#_{AVE}$, $\#_{RAS}$ and $\#_{HRAS}$.

In Table 1, the results for $\alpha=1$ confirm the result of Corollaries 5.7 and 5.9. The other parts of this table show that in fact when $\varepsilon_{\rm prob}=\varepsilon_{\rm prec}=k^2$ then the iteration counts remain bounded as k increases for a range of $H_{\rm prec}$ chosen to decrease more slowly with k than the theoretical requirement of $\mathcal{O}(k^{-1})$. Thus, if there is enough absorption, the preconditioner still works well for solving the shifted system, even for much coarser coarse meshes than those predicted by Corollaries 5.7 and 5.9. The case $H_{\rm prec}=k^{-0.8}$ is close to being load balanced (see (6.3) and the remarks following). To give an idea of the sizes of the systems involved, when $H_{\rm prec}=k^{-0.8}$ and $k=120$

	Left preconditioning				Right preconditioning			
	$\alpha = 1$				$\alpha = 1$			
k	$\#_{AS}$	$\#_{AVE}$	$\#_{RAS}$	$\#_{HRAS}$	$\#_{AS}$	$\#_{AVE}$	$\#_{RAS}$	$\#_{HRAS}$
10	21	15	15	8	21	15	15	8
20	20	15	15	8	19	15	15	8
40	21	16	16	9	19	16	15	8
60	21	16	16	9	19	16	15	8
80	26	18	16	9	23	17	15	8
100	21	17	16	9	19	16	15	8
$\alpha = 0.9$					$\alpha = 0.9$			
k	$\#_{AS}$	$\#_{AVE}$	$\#_{RAS}$	$\#_{HRAS}$	$\#_{AS}$	$\#_{AVE}$	$\#_{RAS}$	$\#_{HRAS}$
10	19	15	15	8	19	15	15	8
20	23	18	18	9	21	18	17	8
40	27	21	19	10	24	19	17	9
60	25	20	20	10	21	20	18	9
80	25	21	20	10	21	20	18	9
100	25	21	20	10	21	20	18	9
$\alpha = 0.8$					$\alpha = 0.8$			
k	$\#_{AS}$	$\#_{AVE}$	$\#_{RAS}$	$\#_{HRAS}$	$\#_{AS}$	$\#_{AVE}$	$\#_{RAS}$	$\#_{HRAS}$
10	19	15	14	8	18	15	14	8
20	21	18	17	9	20	18	17	9
40	23	22	19	10	20	20	17	10
60	21	20	19	11	18	19	17	10
80	21	20	19	11	18	19	17	10
100	22	23	19	11	18	20	17	10

TABLE 1. Number of iterations for various preconditioners with $h_{\text{prob}} = k^{-3/2}$, $\varepsilon_{\text{prob}} = \varepsilon_{\text{prec}} = k^2$, $H_{\text{prec}} = k^{-\alpha}$

the size of the fine grid problem is $n = 1782225$ while the size of the coarse grid problem is 2116 and there are 2025 local problems of maximal size 3364 to be solved. We also note the overall improvement as we compare different preconditioners in the sequence `AS`, `AVE`, `RAS`, `HRAS`.

Then Tables 2 and 3 repeat the same experiments for the cases $\varepsilon_{\text{prob}} = \varepsilon_{\text{prec}} = k$ and $\varepsilon_{\text{prob}} = \varepsilon_{\text{prec}} = 1$. We observe here that when $H_{\text{prec}} = k^{-1}$, both methods continue to work quite well (although the number of iterations does grow mildly with k), however for coarser coarse meshes $H_{\text{prec}} = k^{-\alpha}$ with $\alpha < -1$, the method quickly becomes unusable. The general superiority of `HRAS` over the other methods is striking. A * in the tables indicates that the number of iterations was above 200.

To finish Experiment 1, we repeated the experiments above with left preconditioning but where the GMRES algorithm minimises the residual in the norm induced by D_k. The resulting iteration counts were almost identical to those given in Tables 1, 2 and 3, so we do not give them here.

Note that the results in Table 2, especially the columns corresponding to $\alpha = 1$ and HRAS, show that B_k^{-1} is a good (although admittedly not perfect) preconditioner for A_k. Based on [21] this strongly suggests B_k^{-1} will be a good preconditioner for A (recall properties (i) and (ii) in the introduction); we see that this is indeed the case in the next experiment which is about preconditioners for A.

While Experiment 1 illustrates very well our theoretical results about preconditioning the problem with absorption (i.e. the matrix A_ε), the ultimate goal of our work is to determine the best preconditioner for the problem without absorption (i.e. the matrix A). Therefore, the rest of our experiments focus on investigating this question, and so from now on we take $\varepsilon_{\text{prob}} = 0$. Also, in the experiments above, `HRAS` outperformed all the other preconditioners and so in the rest of our experiments we restrict attention to `HRAS`.

Moreover, remembering that the local solves in $B^{-1}_{\varepsilon,RAS,local}$ are solutions of local problems with a Dirichlet condition on interior boundaries of subdomains, and noting that these are not expected to perform well for genuine wave propagation (i.e. ε small), we also consider the use of impedance

	Left preconditioning				Right preconditioning			
	$\alpha = 1$				$\alpha = 1$			
k	$\#_{\texttt{AS}}$	$\#_{\texttt{AVE}}$	$\#_{\texttt{RAS}}$	$\#_{\texttt{HRAS}}$	$\#_{\texttt{AS}}$	$\#_{\texttt{AVE}}$	$\#_{\texttt{RAS}}$	$\#_{\texttt{HRAS}}$
10	25	16	16	10	25	17	16	9
20	25	20	20	11	24	20	20	11
40	35	30	31	16	32	28	28	14
60	46	41	42	22	40	38	37	19
80	67	56	55	30	56	50	48	24
100	75	69	70	38	64	63	61	31
	$\alpha = 0.9$				$\alpha = 0.9$			
k	$\#_{\texttt{AS}}$	$\#_{\texttt{AVE}}$	$\#_{\texttt{RAS}}$	$\#_{\texttt{HRAS}}$	$\#_{\texttt{AS}}$	$\#_{\texttt{AVE}}$	$\#_{\texttt{RAS}}$	$\#_{\texttt{HRAS}}$
10	22	18	18	10	22	18	18	10
20	37	27	28	14	35	27	27	13
40	118	45	85	24	107	42	83	21
60	*	171	192	40	*	175	187	35
80	*	*	*	61	*	*	*	54
100	*	*	*	97	*	*	*	86
	$\alpha = 0.8$				$\alpha = 0.8$			
k	$\#_{\texttt{AS}}$	$\#_{\texttt{AVE}}$	$\#_{\texttt{RAS}}$	$\#_{\texttt{HRAS}}$	$\#_{\texttt{AS}}$	$\#_{\texttt{AVE}}$	$\#_{\texttt{RAS}}$	$\#_{\texttt{HRAS}}$
10	23	18	19	12	23	18	18	12
20	40	33	35	18	37	33	34	17
40	173	140	190	122	153	130	177	116
60	*	*	*	*	*	*	*	*
80	*	*	*	*	*	*	*	*
100	*	*	*	*	*	*	*	*

TABLE 2. Number of iterations for various preconditioners with $h_{\rm prob} = k^{-3/2}$, $\varepsilon_{\rm prob} = \varepsilon_{\rm prec} = k$, $H_{\rm prec} = k^{-\alpha}$

boundary conditions on the local solves. We therefore introduce the sesquilinear form local to the subdomain Ω_ℓ, defined as the following local equivalent of (2.2):

$$a_{\varepsilon,Imp,\ell}(v,w) \;=\; \int_{\Omega_\ell} \nabla v.\nabla\overline{w} - (k^2 + i\varepsilon)\int_{\Omega_\ell} v\overline{w} - \mathrm{i}k\int_{\partial\Omega_\ell} v\overline{w},$$

(remember that we are choosing $\eta = k$ in all the experiments). We let $A_{\varepsilon,Imp,\ell}$ be the stiffness matrix arising from this form, i.e.

$$(A_{\varepsilon,Imp,\ell})_{j,j'} \;=\; a_{\varepsilon,Imp,\ell}(\phi_{j'},\phi_j)\;, \quad j,j' \in \mathcal{I}(\overline{\Omega_\ell}).$$

This can be used as a local operator in any of the preconditioners introduced above. For example if it is inserted into the HRAS operator (6.7), then the one-level variant is

$$(B^{-1}_{\varepsilon,Imp,RAS,local}\mathbf{v})_j \;=\; \left(\widetilde{R}^T_{\ell(j)}A^{-1}_{\varepsilon,Imp,\ell(j)}\widetilde{R}_{\ell(j)}\mathbf{v}\right)_j\;, \quad \text{for each} \quad j \in \mathcal{I}^h\;. \tag{6.8}$$

Here (noting the distinction with (3.1)), $\widetilde{R}_\ell$ denotes the restriction operator $(\widetilde{R}_\ell)_{j,j'} = \delta_{j,j'}$, (as before) j' ranges over all $\mathcal{I}^h$, but now j runs over all indices such that $x_j \in \overline{\Omega}_\ell\backslash\Gamma$. The hybrid two-level variant is

$$B^{-1}_{\varepsilon,Imp,HRAS} := R_0^T A^{-1}_{\varepsilon,0}R_0 + P_0^T\left(B^{-1}_{\varepsilon,Imp,RAS,local}\right)P_0\;. \tag{6.9}$$

We refer to these as the one- and two-level ImpHRAS preconditioners.

6.2. **Experiment 2.** In Tables 4 and 5 below, we illustrate the performance of these preconditioners with various choices of $\varepsilon_{\rm prec}$ when solving the problem (2.4) with $\varepsilon_{\rm prob} = 0$ and $h_{\rm prob} = k^{-3/2}$. Here we use rectangular coarse grids and subproblems and employ right (FGMRES) preconditioning (although the performance with triangular grids and left preconditioning is similar). In order to ensure our problem (2.4) has physical significance, we choose the data f, g in (2.3) so that the

	Left preconditioning				Right preconditioning			
	$\alpha = 1$				$\alpha = 1$			
k	$\#_{\texttt{AS}}$	$\#_{\texttt{AVE}}$	$\#_{\texttt{RAS}}$	$\#_{\texttt{HRAS}}$	$\#_{\texttt{AS}}$	$\#_{\texttt{AVE}}$	$\#_{\texttt{RAS}}$	$\#_{\texttt{HRAS}}$
10	26	17	17	10	26	17	17	10
20	25	21	21	12	24	21	20	11
40	37	32	32	17	33	30	30	15
60	49	44	45	24	43	41	40	20
80	74	63	61	33	62	56	53	27
100	83	78	79	43	71	70	68	34
$\alpha = 0.9$					$\alpha = 0.9$			
k	$\#_{\texttt{AS}}$	$\#_{\texttt{AVE}}$	$\#_{\texttt{RAS}}$	$\#_{\texttt{HRAS}}$	$\#_{\texttt{AS}}$	$\#_{\texttt{AVE}}$	$\#_{\texttt{RAS}}$	$\#_{\texttt{HRAS}}$
10	23	18	18	10	22	18	18	10
20	38	28	30	14	37	28	28	13
40	134	49	96	26	121	45	93	23
60	*	*	*	44	*	*	*	39
80	*	*	*	71	*	*	*	62
100	*	*	*	115	*	*	*	101
$\alpha = 0.8$					$\alpha = 0.8$			
k	$\#_{\texttt{AS}}$	$\#_{\texttt{AVE}}$	$\#_{\texttt{RAS}}$	$\#_{\texttt{HRAS}}$	$\#_{\texttt{AS}}$	$\#_{\texttt{AVE}}$	$\#_{\texttt{RAS}}$	$\#_{\texttt{HRAS}}$
10	23	20	19	13	23	20	19	13
20	42	35	39	19	40	35	36	17
40	*	194	*	182	187	186	*	181
60	*	*	*	*	*	*	*	*
80	*	*	*	*	*	*	*	*
100	*	*	*	*	*	*	*	*

TABLE 3. Number of iterations for various preconditioners with $h_{\text{prob}} = k^{-3/2}$, $\varepsilon_{\text{prob}} = \varepsilon_{\text{prec}} = 1$, $H_{\text{prec}} = k^{-\alpha}$

exact solution of problem (2.1) is a plane wave $u(x) = \exp(\mathrm{i}kx.\hat{d})$ where $\hat{d} = (1/\sqrt{2}, 1/\sqrt{2})^T$. In Tables 4 and 5 respectively, iteration counts for the two level versions of `HRAS` and `ImpHRAS` are given, with the counts for the corresponding one level method given as subscripts. From these tables we make the following observations:

(1) When $H_{\text{prec}} = k^{-1}$, the two-level versions of both `HRAS` and `ImpHRAS` perform quite well, although the number of iterations does grow mildly with k. The corresponding one-level versions perform poorly, showing that the coarse grid operator is doing a good job in this scenario. The choice of $\varepsilon_{\text{prec}}$ has minimal effect (except that it appears that $\varepsilon_{\text{prec}}$ should not be chosen much bigger than $k^{1.5}$.

(2) When $H_{\text{prec}} = k^{-\alpha}$ and $\alpha < 1$, `HRAS` becomes unusable. `ImpHRAS` also degrades as α decreases but then starts to improve again, and at $\alpha = 0.6$ provides a reasonably efficient solver with very slow growth of iterations with k. Here the two-grid variant is not much better than the one-grid variant, due to the fact that the coarse grid problem has become very coarse (when $k = 80$, $n = 511225$, the size of the coarse grid problem is 196, and the size of each of the largest local problem is 10404.

At the end of Experiment 1, we argued that the fact that B_k^{-1} was a good preconditioner for A_k, along with the results of [21], suggested that B_k^{-1} would be a good preconditioner for A. Having performed Experiment 2, we can now compare preconditioning A_k with B_k^{-1} to preconditioning A with B_k^{-1}. Relevant results are in the column in Table 2 for right-preconditioning with HRAS and $\alpha = 1$ and the column in Table 4 with $\alpha = 1$ and $\beta = 1$. Although the iteration counts are slightly different, a linear least-squares fit shows that the rate of growth with k is very similar in each case ($k^{0.60}$ versus $k^{0.53}$) which is in line with the intuition above.

6.3. **Experiment 3.** From observations above, we can identify a possible multilevel strategy for preconditioning the problem with $h_{\text{prob}} = k^{-3/2}$ and $\varepsilon_{\text{prob}} = 0$. We do this only in 2D using the experiments above, but it is possible to carry out a similar analysis in 3D.

$\alpha = 1$							
$k\backslash\beta$	0	0.4	0.8	1	1.2	1.6	2.0
10	10_{34}	10_{34}	11_{34}	11_{34}	12_{34}	15_{33}	19_{34}
20	12_{92}	12_{92}	12_{92}	12_{92}	13_{92}	19_{92}	37_{93}
40	18_*	18_*	18_*	18_*	18_*	25_*	63_*
60	25_*	25_*	25_*	25_*	25_*	32_*	86_*
80	34_*	34_*	33_*	33_*	32_*	39_*	110_*
100	45_*	45_*	44_*	43_*	42_*	47_*	136_*
$\alpha = 0.8$							
$k\backslash\beta$	0	0.4	0.8	1	1.2	1.6	2.0
10	16_{25}	16_{24}	16_{24}	16_{24}	17_{24}	18_{24}	19_{26}
20	23_{59}	23_{59}	23_{59}	23_{59}	24_{59}	30_{58}	39_{61}
40	$*_*$	$*_*$	175_*	139_*	96_*	65_{160}	76_{132}
60	$*_*$	$*_*$	$*_*$	$*_*$	169_*	114_*	119_{197}
80	$*_*$	$*_*$	$*_*$	$*_*$	$*_*$	166_*	165_*
100	$*_*$	$*_*$	$*_*$	$*_*$	$*_*$	$*_*$	$*_*$
$\alpha = 0.6$							
$k\backslash\beta$	0	0.4	0.8	1	1.2	1.6	2.0
10	16_{18}	16_{18}	16_{18}	16_{18}	16_{18}	16_{19}	19_{22}
20	55_{72}	55_{71}	53_{67}	51_{63}	48_{58}	39_{46}	39_{43}
40	140_{124}	138_{129}	131_{135}	125_{133}	114_{125}	86_{94}	81_{76}
60	$*_*$	$*_*$	$*_*$	$*_*$	$*_*$	147_{141}	113_{102}
80	$*_*$	$*_*$	$*_*$	$*_*$	$*_*$	178_{160}	135_{121}
100	$*_*$	$*_*$	$*_*$	$*_*$			

TABLE 4. Number of iterations for HRAS with $\varepsilon_{\rm prob} = 0$, $\varepsilon_{\rm prec} = k^\beta$, $H_{\rm prec} = k^{-\alpha}$, and right preconditioning.

The first step is to use a two level HRAS with, say, $\varepsilon_{\rm prec} = k$ and $H_{\rm prec} = \mathcal{O}(k^{-1})$. Denote the number of iterations of this method by $I_1(k)$. (A least squares linear fit of the data for the two-level method in Column 5 of the first pane of Table 4, indicates that $I_1(k) \approx \mathcal{O}(k^{0.4})$.) Each application of the preconditioner requires the solution of one system of size $H_{\rm prec}^{-2} = \mathcal{O}(k^2)$ and $\mathcal{O}(k^2)$ systems of size $(H_{\rm prec}/h_{\rm prob})^2 \approx \mathcal{O}(k)$. Each matrix-vector multiplication with the system matrix needs $\mathcal{O}(k^3)$ operations. The total cost is then approximately:

$$I_1(k)\left[\underbrace{C(k^2)}_{\text{Coarse solve}} + \underbrace{\mathcal{O}(k^2)C(k)}_{\text{local solves}} + \underbrace{\mathcal{O}(k^3)}_{\text{matrix}-\text{vector multiplications}}\right], \tag{6.10}$$

where $C(m)$ denotes an estimate of the cost of backsolving with a factorized $m \times m$ finite element system in 2D (the theoretical upper bound is $C(m) \sim m^{3/2}$). (Only backsolves need be counted since these appear in every iteration while the factorization needs only be done once.)

The local solves in (6.10) can in principle be done in parallel, so that the main bottleneck is likely to be the (relatively large) coarse solve. For this reason we consider replacing the direct coarse solve with an inner iteration within an FGMRES set-up. Thus we need an efficient iterative method for the coarse problem, which itself is a Helmholtz finite element system with $h_{\rm prob} \sim k^{-1}$, $\varepsilon_{\rm prob} = k$. Table 6 gives some experiments with preconditioned iterative methods for problems of this form in the two cases $h_{\rm prob} = \pi/10k$ and $h_{\rm prob} = \pi/5k$ (20 and 10 grid points per wavelength respectively), using ImpHRAS with $\varepsilon_{\rm prec} = k$ and $H_{\rm prec} = k^{-1/2}$. Iteration counts are given for the two-level variant (and the one-level variant as subscripts). These show that the one-level method works just as well as (sometimes even better than) the two level method. With the number of iterations denoted by $I_2(k)$, a least-squares linear fit to the one-level data for $h_{\rm prob} = \pi/5k$ yields the estimate $I_2(k) \sim k^{0.3}$.

If we use this method (in its one level form) to approximate the coarse solve in (6.10), then each application of the preconditioner requires the solution of $\mathcal{O}(k)$ systems of size $\mathcal{O}(k)$ and the total

$\alpha = 1$							
$k\backslash\beta$	0	0.4	0.8	1	1.2	1.6	2.0
10	15_{45}	15_{45}	15_{45}	15_{46}	15_{46}	17_{47}	21_{49}
20	17_{104}	17_{104}	17_{105}	17_{105}	18_{105}	20_{107}	34_{113}
40	21_{*}	21_{*}	21_{*}	21_{*}	21_{*}	26_{*}	56_{*}
60	27_{*}	27_{*}	27_{*}	27_{*}	27_{*}	33_{*}	78_{*}
80	36_{*}	36_{*}	35_{*}	35_{*}	34_{*}	40_{*}	101_{*}
100	47_{*}	47_{*}	46_{*}	45_{*}	43_{*}	48_{*}	123_{*}
$\alpha = 0.8$							
$k\backslash\beta$	0	0.4	0.8	1	1.2	1.6	2.0
10	15_{27}	15_{27}	15_{27}	15_{27}	15_{27}	16_{28}	19_{29}
20	23_{56}	23_{56}	23_{56}	23_{56}	23_{56}	25_{57}	34_{62}
40	51_{106}	51_{106}	51_{106}	51_{106}	49_{106}	48_{108}	63_{116}
60	107_{150}	106_{150}	105_{150}	104_{150}	99_{150}	85_{151}	99_{164}
80	187_{194}	185_{193}	183_{193}	178_{193}	168_{193}	132_{194}	138_{*}
100	$*_{*}$	$*_{*}$	$*_{*}$	$*_{*}$	$*_{*}$	185_{*}	179_{*}
$\alpha = 0.6$							
$k\backslash\beta$	0	0.4	0.8	1	1.2	1.6	2.0
10	14_{18}	14_{18}	14_{18}	14_{18}	14_{19}	15_{20}	17_{23}
20	27_{31}	27_{31}	26_{31}	26_{31}	26_{32}	28_{33}	36_{42}
40	51_{51}	51_{51}	50_{51}	50_{51}	48_{51}	50_{51}	73_{66}
60	72_{71}	71_{71}	70_{71}	69_{71}	69_{70}	70_{67}	104_{91}
80	74_{85}	74_{85}	74_{85}	74_{84}	74_{83}	77_{77}	126_{111}
100	84_{98}	84_{98}	84_{98}	84_{97}	84_{95}	86_{87}	148_{131}

TABLE 5. Number of iterations for `ImpHRAS` with $\varepsilon_{\text{prob}} = 0$, $\varepsilon_{\text{prec}} = k^{\beta}$, $H_{\text{prec}} = k^{-\alpha}$, and right preconditioning.

k	$h_{\text{prob}} = \pi/5k$	$h_{\text{prob}} = \pi/10k$
10	9_{10}	9_{9}
20	14_{15}	14_{15}
40	21_{24}	22_{24}
60	30_{32}	31_{32}
80	35_{35}	37_{35}
100	39_{38}	41_{39}
120	42_{40}	45_{43}
140	46_{43}	49_{46}

TABLE 6. Number of iterations for `ImpHRAS` with $\varepsilon_{\text{prob}} = k = \varepsilon_{\text{prec}}$, $H_{\text{prec}} = k^{-0.5}$, and right preconditioning.

cost of the solution is about $I_2(k)\left(kC(k) + \mathcal{O}(k^2)\right)$. Using this to replace the first term on the right-hand side of (6.10), the cost of the resulting inner-outer algorithm would be approximately

$$I_1(k)\left[I_2(k)\left(kC(k) + k^2\right) + \mathcal{O}(k^2)C(k) + \mathcal{O}(k^3)\right] \ . \tag{6.11}$$

Now it is well-known that in 2D fast direct solvers for finite element systems of size k perform with $\mathcal{O}(k)$ complexity for $k \leq 10^5$. So for practically relevant wavenumbers we expect a complexity for the inner-outer algorithm of the form $I_1(k)\left[k^2 I_2(k) + \mathcal{O}(k^3)\right]$.

In Table 7, we illustrate the performance of this inner-outer method (which we denote `IO-ImpHRAS`) implemented within the FGMRES framework. The outer tolerance is 10^{-6} (as before, and the inner tolerance τ is as indicated). Below each iteration count we present also the total running time of our reference NumPy implementation; this includes the setup time of all (sub)matrices. We also give an average time for each outer iteration.

$\tau = 0.01$							
$k\backslash\beta$	**0**	**0.4**	**0.8**	**1**	**1.2**	**1.6**	**2.0**
10	**15(5)**	**15(5)**	**15(5)**	**15(5)**	**15(5)**	**17(4)**	**21(3)**
	0.69 [0.02]	0.75 [0.02]	0.73 [0.03]	0.77 [0.03]	0.74 [0.03]	0.75 [0.02]	0.82 [0.02]
20	**17(7)**	**17(7)**	**17(7)**	**17 (6)**	**18(6)**	**20(4)**	**34 (2)**
	4.49 [0.12]	4.38 [0.12]	4.34 [0.11]	4.35 [0.11]	4.42 [0.11]	4.38 [0.10]	5.18 [0.08]
40	**21(11)**	**21(11)**	**21(11)**	**21(10)**	**21(9)**	**26(6)**	**56 (2)**
	62.9 [0.96]	62.8 [0.96]	63.1 [0.96]	62.4 [0.93]	62.1 [0.91]	63.3 [0.81]	82.8 [0.73]
60	**27(15)**	**27(15)**	**27(15)**	**27(14)**	**27(12)**	**33(6)**	**78(2)**
	420.9 [4.13]	422 [4.13]	423 [4.10]	421 [4.01]	416 [3.87]	426 [3.46]	560 [3.21]
80	**36(17)**	**36(17)**	**35(16)**	**35(15)**	**34(12)**	**40(6)**	**101(2)**
	1536 [11.1]	1564 [11.1]	1608 [10.89]	1555 [10.5]	1540 [10.0]	1542 [8.91]	2052 [8.54]
100	**47(21)**	**47(21)**	**46(20)**	**45(18)**	**43(15)**	**48(7)**	**123(2)**
	4061 [20.9]	4281 [21.1]	4073 [19.6]	3992 [20.5]	4078 [18.4]	3880 [17.1]	5130 [17.5]
$\tau = 0.1$							
$k\backslash\beta$	**0**	**0.4**	**0.8**	**1**	**1.2**	**1.6**	**2.0**
10	**15(3)**	**15(3)**	**15(3)**	**15(3)**	**15(3)**	**17(2)**	**19(2)**
	0.60 [0.02]	0.60 [0.02]	0.61 [0.02]	0.60 [0.02]	0.59 [0.02]	0.60 [0.02]	0.73 [0.02]
20	**17(4)**	**17(4)**	**17(4)**	**17(4)**	**18(4)**	**20(3)**	**34(1)**
	3.88 [0.10]	3.84 [0.09]	3.78 [0.09]	3.91 [0.09]	3.93 [0.09]	3.98 [0.09]	4.86 [0.08]
40	**21(7)**	**21(7)**	**21(7)**	**21(6)**	**21(6)**	**26(4)**	**56(1)**
	56.9 [0.83]	56.9 [0.83]	56.8 [0.82]	56.4 [0.81]	56.8 [0.83]	60.3 [0.76]	80.6 [0.70]
60	**27(9)**	**27(9)**	**27(8)**	**27(8)**	**27(7)**	**34(4)**	**82(1)**
	386 [3.65]	384 [3.63]	384 [3.58]	382 [3.55]	379 [3.46]	396 [3.24]	543 [3.13]
80	**36(10)**	**35(10)**	**35(10)**	**35(9)**	**34(8)**	**40(4)**	**104(1)**
	1361 [9.55]	1389 [9.51]	1398 [9.40]	1344 [9.34]	1368 [9.10]	1332 [8.49]	1926 [8.35]
100	**47(13)**	**46(13)**	**46(12)**	**45(11)**	**43(10)**	**48(4)**	**126(1)**
	3699 [18.6]	3723 [19.3]	3700 [18.1]	3535 [17.1]	3687 [16.7]	3530 [16.1]	4935 [16.9]
$\tau = 0.5$							
$k\backslash\beta$	**0**	**0.4**	**0.8**	**1**	**1.2**	**1.6**	**2.0**
10	**17(2)**	**17(1)**	**18(1)**	**18(1)**	**18(1)**	**19(1)**	**23(1)**
	0.61 [0.02]	0.59 [0.02]	0.66 [0.01]	0.66 [0.02]	0.65 [0.02]	0.66 [0.01]	0.65 [0.01]
20	**19(2)**	**19(2)**	**19(2)**	**19(2)**	**19(2)**	**25(1)**	**36(1)**
	3.86 [0.08]	3.72 [0.08]	3.72 [0.08]	3.68 [0.08]	3.66 [0.08]	4.00 [0.07]	4.96 [0.07]
40	**22(4)**	**22(4)**	**22(4)**	**22(3)**	**22(3)**	**28(2)**	**61(1)**
	54.8 [0.73]	54.9 [0.73]	54.8 [0.72]	54.7 [0.71]	54.8 [0.71]	58.0 [0.69]	80.4 [0.68]
60	**28(5)**	**28(5)**	**28(5)**	**28(5)**	**28(4)**	**35(2)**	**82(1)**
	370 [3.20]	371 [3.20]	372 [3.19]	370 [3.16]	369 [3.11]	383 [3.00]	539 [3.10]
80	**36(6)**	**36(6)**	**36(6)**	**36(5)**	**35(5)**	**42(2)**	**104(1)**
	1288 [8.62]	1375 [8.69]	1300 [8.59]	1316 [8.51]	1273 [8.38]	1323 [8.08]	1909 [8.19]
100	**46(8)**	**46(8)**	**46(7)**	**45(7)**	**44(6)**	**49(2)**	**126(1)**
	3533 [16.5]	3678 [16.01]	3586 [16.4]	3471 [15.9]	3483 [16.2]	3503 [15.5]	4832 [16.4]

TABLE 7. $\texttt{IO} - \texttt{ImpHRAS}$ with $\varepsilon_{\text{prob}} = 0$ and $\varepsilon_{\text{prec}} = k^\beta$. Bold font: number of outer (inner) iterations). Non-bold font: total time in seconds [with an average time for each outer iteration in square brackets]

We see from the table that the commonly-used choice $\varepsilon = k^2$ performs much worse that the choice of $\varepsilon = k^\beta$ for the values of $\beta < 2$ considered here; of these latter choices, $\varepsilon = k$ seems to be the best choice for this composite algorithm. The best times were obtained for a fairly large inner tolerance τ; the results for $\varepsilon = k, \tau = 0.5$ show a growth of the total compute time with k that is close to $\mathcal{O}(k^4) = \mathcal{O}(n^{4/3})$. Note that the runs were performed on a single CPU core; the method is highly parallelisable and parallel implementation results will be presented in a future paper.

Acknowledgements. We thank Melina Freitag, Stefan Güttel, Jennifer Pestana and Andy Wathen for valuable advice about various aspects of weighted GMRES. We also thank Clemens Pechstein for valuable comments.

REFERENCES

[1] B. Beckermann, S. A. Goreinov, and E. E. Tyrtyshnikov. Some remarks on the Elman estimate for GMRES. *SIAM journal on Matrix Analysis and Applications*, 27(3):772–778, 2006.

[2] J-D. Benamou and B. Després. A domain decomposition method for the Helmholtz equation and related optimal control problems. *Journal of Computational Physics*, 136(1):68–82, 1997.

[3] X-C. Cai and M. Sarkis. A restricted additive Schwarz preconditioner for general sparse linear systems. *SIAM Journal on Scientific Computing*, 21(2):792–797, 1999.

[4] X-C. Cai and O. B. Widlund. Domain decomposition algorithms for indefinite elliptic problems. *SIAM Journal on Scientific and Statistical Computing*, 13(1):243–258, 1992.

[5] X-C. Cai and J. Zou. Some observations on the l^2 convergence of the additive Schwarz preconditioned GMRES method. *Numerical linear algebra with applications*, 9(5):379–397, 2002.

[6] Z. Chen and X. Xiang. A source transfer domain decomposition method for Helmholtz equations in unbounded domain Part II: Extensions. *Numerical Mathematics: Theory, Methods and Applications*, 6(03):538–555, 2013.

[7] Z. Chen and X. Xiang. A source transfer domain decomposition method for Helmholtz equations in unbounded domain. *SIAM Journal on Numerical Analysis*, 51(4):2331–2356, 2013.

[8] P-H. Cocquet and M. Gander. Analysis of multigrid performance for finite element discretizations of the shifted Helmholtz equation. *preprint*, 2014.

[9] S. Cools and W. Vanroose. Local Fourier Analysis of the complex shifted Laplacian preconditioner for Helmholtz problems. *Numerical Linear Algebra with Applications*, 20:575–597, 2013.
[10] M. Dauge. *Elliptic boundary value problems on corner domains*. Number 1341 in Lecture Notes in Mathematics. Springer-Verlag, 1988.
[11] V. Dolean, M. J. Gander, and L. Gerardo-Giorda. Optimized Schwarz Methods for Maxwell's Equations. *SIAM Journal on Scientific Computing*, 31(3):2193–2213, 2009.
[12] S. C. Eisenstat, H. C. Elman, and M. H. Schultz. Variational iterative methods for nonsymmetric systems of linear equations. *SIAM Journal on Numerical Analysis*, pages 345–357, 1983.
[13] H. C. Elman. *Iterative Methods for Sparse Nonsymmetric Systems of Linear Equations*. PhD thesis, Yale University, 1982.
[14] B. Engquist and L. Ying. Sweeping preconditioner for the Helmholtz equation: hierarchical matrix representation. *Comm. Pure Appl. Math.*, 64:697–735, 2011.
[15] Y. A. Erlangga. Advances in iterative methods and preconditioners for the Helmholtz equation. *Archives of Computational Methods in Engineering*, 15(1):37–66, 2008.
[16] Y. A. Erlangga, C. W. Oosterlee, and C. Vuik. A novel multigrid based precconditioner for heterogeneous Helmholtz problems. *SIAM J. Sci. Comp.*, 27:1471–1492, 2006.
[17] Y. A. Erlangga, C. Vuik, and C. W. Oosterlee. On a class of preconditioners for solving the Helmholtz equation. *Applied Numerical Mathematics*, 50(3):409–425, 2004.
[18] O. G. Ernst and M. J. Gander. Why it is difficult to solve Helmholtz problems with classical iterative methods. In I. G. Graham, T. Y. Hou, O. Lakkis, and R. Scheichl, editors, *Numerical Analysis of Multiscale Problems*, volume 83 of *Lecture Notes in Computational Science and Engineering*, pages 325–363. Springer, 2012.
[19] A. Essai. Weighted FOM and GMRES for solving nonsymmetric linear systems. *Numerical Algorithms*, 18(3-4):277–292, 1998.
[20] M. J. Gander, L. Halpern, and F. Magoules. An optimized Schwarz method with two-sided Robin transmission conditions for the Helmholtz equation. *International journal for numerical methods in fluids*, 55(2):163–175, 2007.
[21] M.J. Gander, I.G. Graham, and E.A. Spence. Applying GMRES to the Helmholtz equation with shifted Laplacian preconditioning: What is the largest shift for which wavenumber-independent convergence is guaranteed? *Numerische Mathematik*, 131(3):567–614, 2015.
[22] M.J. Gander, F. Magoulès, and F. Nataf. Optimized Schwarz methods without overlap for the Helmholtz equation. *SIAM Journal on Scientific Computing*, 24(1):38–60, 2002.
[23] J. Gopalakrishnan and J. Pasciak. Overlapping schwarz preconditioners for indefinite time harmonic maxwell equations. *Mathematics of Computation*, 72(241):1–15, 2003.
[24] I. G. Graham, M. Löhndorf, J. M. Melenk, and E. A. Spence. When is the error in the h-BEM for solving the Helmholtz equation bounded independently of k? *BIT Numer. Math.*, 55(1):171–214, 2015.
[25] I. G. Graham and R. Scheichl. Robust domain decomposition algorithms for multiscale PDEs. *Numerical Methods for Partial Differential Equations*, 23(4):859–878, 2007.
[26] I. G. Graham, E. A. Spence, and E. Vainikko. Recent results on domain decomposition preconditioning for the high-frequency Helmholtz equation using absorption. In Domenico Lahaye, Jok Tang, and Kees Vuik, editors, *Modern solvers for Helmholtz problems*. Birkhauser series in Geosystems Mathematics, 2016.
[27] I.G. Graham, P.O. Lechner, and R. Scheichl. Domain decomposition for multiscale PDEs. *Numerische Mathematik*, 106(4):589–626, 2007.
[28] P. Grisvard. *Elliptic problems in nonsmooth domains*. Pitman, Boston, 1985.
[29] S. Güttel and J. Pestana. Some observations on weighted GMRES. *Numerical Algorithms*, 67(4):733–752, 2014.
[30] J-H. Kimn and M. Sarkis. Restricted overlapping balancing domain decomposition methods and restricted coarse problems for the Helmholtz problem. *Computer Methods in Applied Mechanics and Engineering*, 196(8):1507–1514, 2007.
[31] J-H. Kimn and M. Sarkis. Shifted Laplacian RAS solvers for the Helmholtz equation. In *Domain Decomposition Methods in Science and Engineering XX*, pages 151–158. Springer, 2013.
[32] J. Mandel and M. Brezina. Balancing domain decomposition for problems with large jumps in coefficients. *Mathematics of Computation*, 65(216):1387–1401, 1996.
[33] R. Nabben and C. Vuik. A comparison of deflation and coarse grid correction applied to porous media flow. *SIAM Journal on Numerical Analysis*, 42(4):1631–1647, 2004.
[34] J. Nečas. *Les méthodes directes en théorie des équations elliptiques*. Masson, 1967.
[35] M. A. Olshanskii and E. E. Tyrtyshnikov. *Iterative methods for linear systems: theory and applications*. SIAM, 2014.
[36] M. Sarkis and D. B. Szyld. Optimal left and right additive Schwarz preconditioning for minimal residual methods with Euclidean and energy norms. *Computer Methods in Applied Mechanics and Engineering*, 196(8):1612–1621, 2007.
[37] L. R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Mathematics of Computation*, 54(190):483–493, 1990.
[38] A. H. Sheikh, D. Lahaye, and C. Vuik. On the convergence of shifted Laplace preconditioner combined with multilevel deflation. *Numerical Linear Algebra with Applications*, 20:645–662, 2013.
[39] E. A. Spence. "When all else fails, integrate by parts" – an overview of new and old variational formulations for linear elliptic PDEs. In A. S. Fokas and B. Pelloni, editors, *Unified transform method for boundary value problems: applications and advances*. SIAM, 2015.

[40] C. C. Stolk. A rapidly converging domain decomposition method for the Helmholtz equation. *Journal of Computational Physics*, 241:240–252, 2013.
[41] A. Toselli. Some results on overlapping Schwarz methods for the Helmholtz equation employing perfectly matched layers. In *Domain Decomposition Methods in Sciences and Engineering: Eleventh International Conference London, UK*, pages 539–545. Citeseer, 1998.
[42] A. Toselli and O. Widlund. *Domain decomposition methods: algorithms and theory*. Springer, 2005.
[43] M. B. Van Gijzen, Y. A. Erlangga, and C. Vuik. Spectral analysis of the discrete Helmholtz operator preconditioned with a shifted Laplacian. *SIAM Journal on Scientific Computing*, 29(5):1942–1958, 2007.
[44] L. Zepeda-Núnez and L. Demanet. The method of polarized traces for the 2D Helmholtz equation. *arXiv preprint arXiv:1410.5910*, 2014.

Dept of Mathematical Sciences, University of Bath, Bath BA2 7AY, UK.
E-mail address: I.G.Graham@bath.ac.uk

Dept of Mathematical Sciences, University of Bath, Bath BA2 7AY, UK.
E-mail address: E.A.Spence@bath.ac.uk

Institute of Computer Science, University of Tartu, 50409, Estonia
E-mail address: eero.vainikko@ut.ee