A TWO-LEVEL SHIFTED LAPLACE PRECONDITIONER FOR
HELMHOLTZ PROBLEMS: FIELD-OF-VALUES ANALYSIS AND
WAVENUMBER-INDEPENDENT CONVERGENCE

LUIS GARCÍA RAMOS, REINHARD NABBEN

Abstract. One of the main tools for solving linear systems arising from the discretization
of the Helmholtz equation is the shifted Laplace preconditioner, which results from the discretization
of a perturbed Helmholtz problem $-\Delta u - (k^2 + i\varepsilon)u = f$ where $0 \neq \varepsilon \in \mathbb{R}$ is an absorption
parameter. In this work we revisit the idea of combining the shifted Laplace preconditioner with
two-level deflation and apply it to Helmholtz problems discretized with linear finite elements. We
use the convergence theory of GMRES based on the field of values to prove that GMRES applied
to the two-level preconditioned system with a shift parameter $\varepsilon \sim k^2$ converges in a number of
iterations independent of the wavenumber $k$, provided that the coarse mesh size $H$ satisfies a
condition of the form $Hk^2 \leq C$ for some constant $C$ depending on the domain but independent
of the wavenumber $k$. This behaviour is sharply different to the standalone shifted Laplacian,
for which wavenumber-independent GMRES convergence has been established only under the
condition that $\varepsilon \sim k$ by [M.J. Gander, I.G. Graham and E.A. Spence, Numer. Math., 131 (2015),
567-614]. Finally, we present numerical evidence that wavenumber-independent convergence of
GMRES also holds for pollution-free meshes, where the coarse mesh size satisfies $Hk^{3/2} \leq C$, and
inexact coarse grid solves.

1. Introduction. In this work we study the solution of linear systems of equations arising from the discretization of the Helmholtz equation. We concentrate here
on the interior Helmholtz problem with impedance boundary conditions, which for
a domain $\Omega \subset \mathbb{R}^d$ with boundary $\Gamma$, $k \in \mathbb{R}$, $f_1 \in L^2(\Omega)$ and $f_2 \in L^2(\Gamma)$ is defined as:

$$
\begin{cases}
-\Delta u - k^2 u &= f_1 \text{ in } \Omega, \\
\partial_n u - iku &= f_2 \text{ in } \Gamma.
\end{cases}
$$

The solution of these systems of equations is one of the main computational
bottlenecks for solving inverse problems in various applications, e.g., in exploration
geophysics and medical imaging. The standard variational formulation of (1.1) and
its corresponding Galerkin discretization with $P1$ finite elements on a simplicial
mesh $T_h$ of the domain $\Omega$ leads to a linear system

$$
Au = f,
$$

where $A \in \mathbb{C}^{N_h \times N_h}$, and $u, f \in \mathbb{C}^{N_h \times N_h}$. Due to the oscillatory character of the
solutions, in order to obtain an accurate approximation the number of gridpoints
in one dimension should be at least proportional to $k$, leading to linear systems
of size $N_h \sim k^d$ where $d$ is the spatial dimension. However, the Galerkin solutions
are affected by the pollution effect [1, 33] and this rule is not sufficient to maintain
accuracy when using discretizations with low-order finite elements for large
wavenumbers. In this case the number of points in one-dimension should be chosen
proportional to $k^{3/2}$ leading to very large linear systems of size $N_h \sim k^{3d/2}$. Moreover,
the matrix $A$ is non-Hermitian and indefinite, making the efficient solution of
(1.2) with standard iterative techniques a huge challenge; for a survey see [18].

The development of fast solvers for the Helmholtz equation has been an active
research area over the last decades. Notable works include the wave-ray method
[4,40], methods based on domain decomposition, and sweeping-type preconditioners
[10,11,53], see, e.g., the survey papers [12,18,22] for more references.

A very fruitful idea introduced in the landmark paper [16] is to precondition
(1.2) with the discretization of a Helmholtz problem with absorption (or shifted
Laplace problem) of the form

$$
\begin{cases}
-\Delta u - (k^2 + i\varepsilon)u &= f_1 \text{ in } \Omega, \\
\partial_n u - iku &= f_2 \text{ in } \Gamma,
\end{cases}
$$

arXiv:2006.08750v2 [math.NA] 17 Jun 2020
where \( \varepsilon \) is a real positive parameter. After discretization of (1.3) one obtains a linear system with coefficient matrix \( A_\varepsilon \). The complex shift \( i\varepsilon \) in (1.3) reduces the oscillations in the solutions and allows the preconditioner \( A_\varepsilon \) to be inverted with fast methods, e.g., multigrid or domain decomposition. This preconditioner is known as the complex shifted Laplacian (CSL) or shifted Laplace preconditioner, and is a building block for several state-of-the-art solvers for the Helmholtz equation, see, e.g., the book [37] for a recent overview of extensions and industrial applications of Helmholtz solvers based on the shifted Laplacian. The resulting preconditioned system

\[
AA_\varepsilon^{-1}g = f, \quad u = A_\varepsilon^{-1}g
\]

can be solved more efficiently than the original system with Krylov subspace methods for non-Hermitian systems. Here we only consider the solution of the linear system with the GMRES method. In practice, the preconditioner is inverted approximately with a fast method; denoting this approximation by \( D_\varepsilon^{-1} \approx A_\varepsilon^{-1} \), the linear system to be analyzed is

\[
AD_\varepsilon^{-1}g = f, \quad u = D_\varepsilon^{-1}g.
\]

Naturally there is a tradeoff in the choice of the shift \( \varepsilon \), since a small shift leads to faster convergence of the Krylov solver but only a large enough shift will allow the (approximate) inversion of \( A_\varepsilon \) with fast methods. To date, the most rigorous analysis on how to choose the shift has appeared in [21], and the series of papers [7, 25, 26]. The authors of [21] propose to separate the analysis of the shifted Laplacian in two questions:

(a) Assuming that \( A_\varepsilon \) is inverted exactly, determine conditions on \( \varepsilon \) for \( A_\varepsilon \) to be a good preconditioner for \( A \).

(b) Determine conditions on \( \varepsilon \) for \( D_\varepsilon^{-1} \) to be a good approximation for \( A_\varepsilon^{-1} \), in particular when using multigrid or domain decomposition methods.

More rigorously, one can use the identity

\[
I - AD_\varepsilon^{-1} = I - A_\varepsilon D_\varepsilon^{-1} + A_\varepsilon D_\varepsilon^{-1}(I - AA_\varepsilon^{-1}),
\]

to show that if both \( \|I - A_\varepsilon D_\varepsilon^{-1}\| \) and \( \|I - AA_\varepsilon^{-1}\| \) are small (i.e., the quantitative version of statements (a) and (b)) then GMRES applied to (1.4), is expected to converge fast.

Some of the early works on the shifted Laplacian [15, 16, 52] focused on answering question (a) using the GMRES residual bounds based on the spectrum of the matrix. In these papers the choice \( \varepsilon = \beta k^2 \) with \( \beta \in [0.5, 1] \) was recommended, based on an analysis of the spectrum of the preconditioned (continuous) Helmholtz operator with Dirichlet boundary conditions in 1D in [12] and the spectrum of more general preconditioned discrete Helmholtz operators in [52]. Likewise, a combination of questions (a) and (b) was investigated in [8] using Local Fourier Analysis, leading to an expression for the near-optimal value (for guaranteed multigrid V-cycle convergence and a near-optimal minimum number of GMRES iterations) depending on the wavenumber and gridsize. For more references on how to choose the shift in the CSL preconditioner see [21, Section 1.1].

The analysis of GMRES convergence based on the distribution of the eigenvalues of the preconditioned matrix rests on the assumption that the condition number of the matrix of eigenvectors of the preconditioned matrix is small. However, this factor is hard to estimate and for some problems it can be so large that the bound may not be informative at all. Moreover, it is known that the convergence of GMRES applied to a non-normal linear system cannot be predicted only by the spectrum of the matrix [27, 38]. The authors of [21] use instead the convergence theory of GMRES based on the field of values (which we summarize in section
3.) to study question (a) above. Their main result shows that under some natural assumptions on the geometry of the domain the condition $\varepsilon \lesssim k$ (i.e., $\varepsilon \leq Ck$ with a small enough constant $C$) is sufficient for the field of values of $\mathbf{A}\mathbf{A}_\varepsilon^{-1}$ to be bounded away from the origin as $k \to \infty$, and therefore under this condition the number of iterations of GMRES applied to the preconditioned system remains constant as $k$ is increased.

Question (b) has been studied in [7] for general 2D and 3D problems in the context of the multigrid method. There it is shown that choosing a shift $\varepsilon \sim k^2$ is a necessary and sufficient condition for the convergence of multigrid with a fixed number of (weighted Jacobi) smoothing steps applied to a linear system with coefficient matrix $\mathbf{A}_\varepsilon$. Regarding domain decomposition methods and question (b), the work [25] investigates the requirements for an additive Schwarz preconditioner $\mathbf{D}_\varepsilon^{-1} \approx \mathbf{A}_\varepsilon^{-1}$ (with Dirichlet transmission conditions between subdomains and coarse grid correction) to be effective as a preconditioner for a problem with coefficient matrix $\mathbf{A}_\varepsilon$, i.e., to obtain fast convergence of GMRES applied to a linear system with coefficient matrix $\mathbf{A}_\varepsilon \mathbf{D}_\varepsilon^{-1}$. There it is concluded that a sufficient condition is $\varepsilon \lesssim k^2$. The more recent paper [26] introduces an additive Schwarz method with local impedance boundary conditions as transmission conditions between subdomains and shows that it is possible to obtain wavenumber independent convergence also under the condition $\varepsilon \sim k^{1+\beta}$ for $\beta$ arbitrarily small. Together, these recent results imply that there exists a rigorously quantified gap between the condition for having a good preconditioner (which requires a small shift $\varepsilon \lesssim k$) and the requirement for being able to (approximately) invert $\mathbf{A}_\varepsilon$ (for which $\varepsilon \sim k^2$ or at least $\varepsilon \sim k^{1+\beta}$ with $\beta > 0$ arbitrarily small is necessary), thus motivating the need for more advanced preconditioning techniques.

In this paper we revisit the combination of the shifted Laplacian with two-level deflation, which was introduced in [47,48] based on previous work in [14], see also [24] for an analysis of the method proposed in [14]. We extend the simplified variant of two-level deflation from [47,48] to Helmholtz problems discretized with finite elements, and we contribute to the line of analysis proposed in [21] by studying the analogous of question (a) for the two-level shifted Laplacian combined with deflation. We use the framework for the analysis of two-level preconditioners from [29].

The method that we use combines the shifted Laplacian with a projection-based preconditioner which removes the components of $\mathbf{A}\mathbf{A}_\varepsilon^{-1}$ that cause the slow convergence of a Krylov subspace method applied to the linear system. Algebraically this idea can be motivated as follows. Let $\mathbf{Q} \in \mathbb{C}^{n \times n}$ be a projection operator, i.e., a matrix such that $\mathbf{Q}^2 = \mathbf{Q}$. If $\mathcal{R}(\mathbf{Q})$ is the range of $\mathbf{Q}$ and $\mathcal{N}(\mathbf{Q})$ its nullspace, the direct sum decomposition $\mathbb{C}^n = \mathcal{R}(\mathbf{Q}) \oplus \mathcal{N}(\mathbf{Q})$ holds and $\mathbf{A}^{-1}$ can be split as

$$
\mathbf{A}^{-1} = \mathbf{A}^{-1}(\mathbf{I} - \mathbf{Q}) + \mathbf{A}^{-1}\mathbf{Q},
$$

note that $(\mathbf{I} - \mathbf{Q})$ is also a projection. Therefore, a sensible approximation for the inverse of $\mathbf{A}$ is

$$
\mathbf{B} = \mathbf{A}_\varepsilon^{-1}(\mathbf{I} - \mathbf{Q}) + \mathbf{A}^{-1}\mathbf{Q},
$$

(1.5)

The crucial point is that the projection can be chosen so that the term $\mathbf{A}^{-1}\mathbf{Q}$ on the right can be computed cheaply by solving a smaller linear system. If $m < n$ and $\mathcal{U}$ is a subspace of $\mathbb{C}^n$ spanned by the columns of a full rank matrix $\mathbf{U} \in \mathbb{C}^{n \times m}$ such that $\mathbf{U}^*\mathbf{A}\mathbf{U}$ is nonsingular (where the superscript * denotes the conjugate transpose), the projection $\mathbf{Q}$ with $\mathcal{R}(\mathbf{Q}) = \mathcal{A}\mathcal{U}$ and $\mathcal{N}(\mathbf{Q}) = \mathcal{U}^\perp$ (the orthogonal complement of $\mathcal{U}$ in the Euclidean inner product) has the form

$$
\mathbf{Q} = \mathbf{A}\mathbf{U}(\mathbf{U}^*\mathbf{A})^{-1}\mathbf{U}^*.
$$
which gives

\[ B = A^{-1}(I - AU(U^*AU)^{-1}U^*) + U(U^*AU)^{-1}U^* \]

and the term \( U(U^*AU)^{-1}U^* \) can be computed by inverting a smaller system with coefficient matrix \( U^*AU \). It follows easily using this projection representation that if \( B \) is used as a preconditioner for \( A \) then \( AB \) equals the identity when restricted to the subspace \( U \), so the spectrum of the preconditioned matrix \( AB \) contains one as an eigenvalue with multiplicity (at least) \( m = \dim(U) \). Moreover, if the subspace \( U \) contains the solution \( u \) we have \( f = Au \in AU \), and GMRES applied to the preconditioned linear system \( ABu = f \) with a zero initial guess will converge in one step. Projection-based preconditioners of the form (1.6) are related to classical deflation methods in which a projection operator is used to remove near-singular eigenspaces responsible for slowing down the convergence of a Krylov subspace iteration, the main difference being that in classical deflation the resulting deflated linear system is singular, i.e., the eigenvalues are shifted to zero, not to one. For more on the connection between two-level methods, projections and deflation see [23].

An important class of methods that lead to preconditioners of the form (1.6) are two-level (or two-grid) methods, which are the basis of the multigrid method [51]. In [47, 48], the authors consider a finite difference discretization of the problem on a grid \( G_h \) and a coarse grid \( G_H \subset G_h \), and choose the subspace \( U \) as the span of (the matrix representation of) the two-grid prolongation operator \( P \). Here we analyze an extension of this preconditioner to the finite element setting, where the mesh \( T_h \) is coarsened by choosing a mesh \( T_H \) such that the elements in \( T_H \) are unions of elements of \( T_h \). If \( V_h, V_H \) are spaces of \( P_1 \) finite elements to \( T_h \) and \( T_H \) respectively, we then have \( V_H \subset V_h \) and the prolongation operator is defined trivially by this inclusion. The resulting preconditioner is (see section 4. for more details)

\[ B_\varepsilon = A^{-1}_\varepsilon(I - APA^{-1}_H P^*) + PA^{-1}_H P^*, \]

leading to the preconditioned system

\[ AB_\varepsilon g = f. \]

The main result in our paper is Theorem 3.1, where we prove that it is possible to close the gap between the requirements for \( \varepsilon \), i.e., we show that wavenumber-independent GMRES convergence can be obtained with the two-level shifted Laplacian even in the case of a large shift \( \varepsilon \sim k^2 \), provided that the coarse grid size satisfies a condition of the form \( Hk^2 < C \) for some constant \( C > 0 \) depending only on the domain \( \Omega \) (but independent of the wavenumber \( k \)). Note that in this theorem we are assuming that the shifted Laplacian and the coarse grid system are inverted exactly.

This result also confirms what has been previously observed in the spectral analysis of a 1D model problem in [37, 48] where it has been shown (using Fourier analysis on a one-dimensional Helmholtz problem with Dirichlet boundary conditions) that when the shifted Laplacian is combined with two-level deflation with a complex shift \( \varepsilon = \beta k^2 \) it is possible to increase \( \beta \) without greatly affecting the spectrum of the preconditioned matrix.

We prove Theorem 3.1 for weighted GMRES in the inner product induced by the inverse of the domain mass matrix, and show that this norm is a natural norm to measure the residuals of the preconditioned system since it corresponds to the dual \( L^2 \) norm when \( C_N \) is identified with the space of coordinates of \( V'_h \) (Proposition 2.3, (c)). Fortunately, for a sequence of quasi-uniform meshes one can use a scaling argument and norm equivalences to show that the result also holds for GMRES in the Euclidean inner product (Corollary 3.10, part (a)).
This paper is organized as follows: In section 2, we review some basic results on the variational formulation of Helmholtz problems, focusing on the conditions for existence and uniqueness of solutions to these problems and their stability. In sections 3 and 4, we introduce the finite element formulation of the Helmholtz and shifted Laplace problems and the convergence theory of GMRES based on the field of values. In section 5, the two-grid preconditioner is introduced and we prove our main result. Finally, in section 6, we present some numerical experiments to illustrate our results.

2. Preliminaries: A recap of the variational formulation and finite element approximation of Helmholtz problems. In this section we review some basic results on the variational formulation of Helmholtz problems, focusing on the conditions for existence and uniqueness of solutions to these problems and their stability. In the second part we discuss the finite element approximation of Helmholtz problems with the Galerkin method. We refer the reader to [49] for a very good introduction to the variational formulation of Helmholtz problems. To simplify the notation, we will write $a \lesssim b$ to denote that there exists a constant $C$ such that $a \leq Cb$ independent of the parameters on which $a$ and $b$ may depend. Moreover, we write $a \sim b$ when $a \lesssim b$ and $b \lesssim a$.

Given a complex inner product space $V$ we denote its sesquilinear inner product by $(\cdot, \cdot)_V$. The antidual space (of continuous conjugate-linear functionals from $V$ to $\mathbb{C}$) is denoted by $V'$. The duality pairing $\langle \cdot, \cdot \rangle_{V' \times V} : V' \times V \to \mathbb{C}$ is defined for $f \in V'$, $v \in V$ as

$$\langle f, v \rangle_{V' \times V} = f(v).$$

The dual norm in the space $V'$ is defined by

$$\|f\|_{V'} = \sup_{0 \neq v \in V} \frac{|\langle f, v \rangle_{V' \times V}|}{\|v\|_V}.$$  

We will drop the subscripts $V', V$ when this introduces no ambiguities, and write only $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$. Let $\Omega \subset \mathbb{R}^d$ be a convex polyhedron in $\mathbb{R}^d$ (where $d = 1, 2, 3$) with boundary $\Gamma$. Recall that the Sobolev space $L^2(\Omega)$ of square integrable functions is equipped with the inner product

$$(u, v)_{L^2(\Omega)} = \left( \int_{\Omega} uv \right)^{1/2}.$$  

The higher order Sobolev spaces $H^m(\Omega)$ consist of functions $u \in L^2(\Omega)$ that have weak derivatives $\partial^\alpha u$ in $L^2(\Omega)$ for all multi-indices $\alpha$ with $|\alpha| \leq m$. The standard inner product in $H^m(\Omega)$ is given by

$$(u, v)_{m} = \sum_{|\alpha| \leq m} (\partial^\alpha u, \partial^\alpha v)_{L^2(\Omega)},$$

and the induced norm is denoted by $\|\cdot\|_{m, \Omega}$. In the space $H^m(\Omega)$ we also introduce the seminorm

$$|u|_m = \sum_{|\alpha| = m} \|\partial^\alpha u\|_{L^2(\Omega)}.$$  

The variational formulation of the Helmholtz problem requires the use of a special $k$-weighted inner product in the space $H^1(\Omega)$. Given $k \in \mathbb{R}$, we define the Helmholtz energy inner product on $H^1(\Omega)$ by

$$(u, v)_{k, 1, \Omega} = (\nabla u, \nabla v)_{L^2(\Omega)} + k^2 (u, v)_{L^2(\Omega)},$$

for any $u, v \in H^1(\Omega)$. The induced norm will be denoted by $\|\cdot\|_{k, 1, \Omega}$. After multiplying each side of (1.1) with a test function $v \in H^1(\Omega)$, integrating by parts and
substituting the boundary condition, the problem can be restated in the variational form

\[(2.1) \quad \text{Find } u \in H^1(\Omega) \text{ such that } a(u, v) = \langle f, v \rangle \text{ for all } v \in H^1(\Omega), \]

where the sesquilinear form \(a : H^1(\Omega) \times H^1(\Omega) \to \mathbb{C}\) and the antilinear functional \(f : L^2(\Omega) \to \mathbb{C}\) are given by

\[(2.2) \quad a(u, v) = \int_\Omega \nabla u \nabla v - k^2 \int_\Omega u v - ik \int_\Gamma u v,
\]

\[(2.3) \quad \langle f, v \rangle = \int_\Omega f_1 v + \int_\Gamma f_2 v.\]

The following lemma summarizes some properties of the sesquilinear form of the Helmholtz problem.

**Lemma 2.1.** Let \(a\) be the sesquilinear form (2.2) of the Helmholtz problem. The following properties hold:

(a) [43, Lemma 8.1.6], [49, p.118] The form \(a\) is continuous with continuity constant \(C_c\) independent of \(k\), i.e., there exists \(C_c\) such that for all \(u, v \in H^1(\Omega), k \in \mathbb{R}\):

\[|a(u, v)| \leq C_c \|u\|_{1,k,\Omega} \|v\|_{1,k,\Omega} \]

(b) The form \(a\) satisfies the Gårding inequality (as an equality)

\[(2.4) \quad \|u\|^2_{1,k,\Omega} = \Re a(u, u) + 2k^2 \|u\|^2_{L^2(\Omega)}.\]

It can be shown [49, Lemma 6.17] that associated to the sesquilinear form \(a\) there exists a bounded operator \(A : H^1(\Omega) \to H^1(\Omega)'\) such that

\[(2.5) \quad a(u, v) = \langle Au, v \rangle,\]

using the operator \(A\), the variational problem can be rewritten as

\[(2.6) \quad \text{Find } u \in H^1(\Omega) \text{ such that } Au = f.\]

The next theorem gives a stability estimate for the Helmholtz problem. For the definition of the norm \(\cdot \|_{1/2,\Gamma}\) see [28, Section 6.2.4]

**Theorem 2.2.** Let \(\Omega \subset \mathbb{R}^d\) be a bounded convex domain with boundary \(\Gamma\), where \(d = 2, 3\). Given \(k_0 > 0\), there exists a constant \(C\) (depending only on \(\Omega\)) such that for any \(g \in L^2(\Omega), h \in L^2(\Gamma)\) and \(k > k_0\) the solution of the Helmholtz problem satisfies

\[(2.7) \quad \|u\|_{1,k,\Omega} \leq C(\|f_1\|_{L^2(\Omega)} + \|f_2\|_{L^2(\Gamma)}).\]

Moreover, if \(u \in H^2(\Omega):\)

\[(2.8) \quad |u|_{H^2(\Omega)} \leq C[(1 + k)(\|f_1\|_{L^2(\Omega)} + \|f_2\|_{L^2(\Gamma)}) + \|f_2\|_{1/2,\Gamma}].\]

**Proof.** See [43, Prop. 8.1.4] for the case \(d = 2\) and [31, Prop. 3.4, 3.6] for the case \(d = 3\). \(\square\)

**2.1. Finite element approximation of Helmholtz problems.** In this section we recall the Galerkin formulation of the Helmholtz problem. Let \(\{T_h\}_{h>0}\) be a family of conforming simplicial meshes of \(\Omega\), where \(h = \max_{K \in T_h} \text{diam}(K)\) denotes the mesh diameter. We assume that the family \(\{T_h\}_{h>0}\) is shape-regular, i.e.,

\[
\sup_{h>0} \max_{K \in T_h} \frac{\text{diam}(K)}{\rho(K)} \leq C < \infty.
\]
We let \( V_h \) be the space of P1 finite elements subordinate to \( T_h \) spanned by the standard nodal basis \( \Phi_h = (\phi_1, \ldots, \phi_{N_h}) \). To simplify the notation, in what follows we omit the subscripts and denote the \( L^2 \) inner product by \( \langle \cdot, \cdot \rangle \) and the corresponding norm by \( \| \cdot \| \). Every element \( u = \sum_{i=1}^{N_h} u_i \phi_i \in V_h \) can be represented by the vector of coordinates \( u = (u_1, \ldots, u_{N_h}) \in \mathbb{C}^{N_h} \), we write this correspondence as

\[
    u = \Phi_h u.
\]

Associated to \( \Phi_h \) there exists a canonical basis of antilinear functionals \( \Phi'_h = (\phi'_1, \ldots, \phi'_{N_h}) \) for the space \( V'_h \), that satisfies

\[
    \langle \phi'_i, \phi_j \rangle = \delta_{ij}, \text{ for } i, j = 1, \ldots, N_h,
\]

We write \( f = \Phi'_h f \) for the coordinate correspondence between \( \mathbb{C}^{N_h} \) and \( V'_h \). In this notation, we have

\[
    u = ((\phi'_1, u), \ldots, (\phi'_{N_h}, u))^*,
\]

here * denotes the conjugate transpose. Recall that if \( D \in \mathbb{C}^{N \times N} \) is a Hermitian positive definite (HPD) matrix, the inner product \( \langle \cdot, \cdot \rangle_D \) induced by \( D \) on \( \mathbb{C}^N \) is defined as

\[
    (x, y)_D = y^* D x,
\]

the corresponding inner product on \( \mathbb{C}^n \) will be denoted by \( \| \cdot \|_D \).

When \( \mathbb{C}^{N_h} \) is identified with the coordinate space of \( V_h \) via \( \Phi_h \), the domain mass matrix \( M \in \mathbb{C}^{N_h \times N_h} \) defined as

\[
    M_{ij} = \langle \phi_j, \phi_i \rangle_{L^2(\Omega)}, \quad 1 \leq i, j \leq N_h.
\]

induces a norm in \( \mathbb{C}^{N_h} \) corresponding to the \( L^2 \) norm in the space \( V_h \). This implies that for all \( u, v \in \mathbb{C}^{N_h} \) and \( u = \Phi_h u, v = \Phi_h v \in V_h \):

\[
    (u, v)_M = (u, v)_{L^2(\Omega)}.
\]

The Riesz representation theorem implies that for every \( f \in V'_h \) there exists a unique \( u_f \in V_h \) such that

\[
    \int_{\Omega} u_f \overline{v} = \langle f, v \rangle,
\]

for all \( v \in V_h \). The mapping \( \tau : V'_h \to V_h \) defined by \( \tau(f) = u_f \) is called the Riesz map (with respect to the \( L^2 \) inner product). The corresponding representation in \( \mathbb{C}^{N_h} \) of the duality pairing \( \langle \cdot, \cdot \rangle \), the \( L^2 \) Riesz map and the dual norm is explained in the next proposition, for completeness we include the proof from chapter 6 of [42].

**Proposition 2.3.** The following statements hold:

(a) The duality pairing \( \langle \cdot, \cdot \rangle : V'_h \times V_h \to \mathbb{C} \) is represented by the Euclidean product in \( \mathbb{C}^{N_h} \), that is, for \( f, u \in \mathbb{C}^{N_h} \), \( u = \Phi_h u \) and \( f = \Phi'_h f \):

\[
    \langle f, u \rangle = u^* f.
\]

(b) The matrix representation of the \( L^2 \) Riesz map \( \tau : V'_h \to V_h \) is the inverse of the mass matrix \( M \):

\[
    (2.10) \quad M^{-1} = \Phi_h \circ \tau \circ \Phi'_h \in \mathbb{C}^{N_h \times N_h}.
\]

(c) The norm in \( \mathbb{C}^{N_h} \) corresponding to the dual norm in \( V'_h \) is the norm induced by the inverse of the mass matrix \( M^{-1} \) from (2.10), that is, for \( f \in \mathbb{C}^{N_h} \) and \( f = \phi'_h f \) we have:

\[
    \| f \|_{M^{-1}} = \sup_{u \in \mathbb{C}^{N_h}} \frac{|u^* f|}{\| u \|_M} = \sup_{u \in V_h} \frac{|\langle f, u \rangle|}{\| u \|}.
\]
Proof. For part (a), let \( u = (u_1, \ldots, u_{N_h}) \) and \( f = (f_1, \ldots, f_{N_h}) \). We have
\[
\langle f, u \rangle = \left( \sum_{i=1}^{N_h} f_i \phi_i', \sum_{j=1}^{N_h} u_j \phi_j \right)
= \sum_{i,j=1}^{N_h} f_i \pi_j \langle \phi_i', \phi_j \rangle = \sum_{i=1}^{N_h} f_i \pi_i = u^* f.
\]

For part (b), suppose that \( M_{\tau} \in \mathbb{C}^{N_h \times N_h} \) is the matrix representation of the Riesz map \( \tau : V_h' \to V_h \). With part (a) we obtain
\[
u^* f = \langle f, u \rangle = (\tau f, u) = (M_{\tau} f, u)_{M} = u^* M_{\tau} f, \quad \text{for all } f, u \in \mathbb{C}^{N_h},
\]
which implies \( M_{\tau} \cdot M_{\tau} = I \), so \( M_{\tau} = M^{-1} \). For part (c), given \( f \in \mathbb{C}^{N_h} \) the Cauchy-Schwarz inequality in the inner product induced by \( M^{-1} \) implies that for every \( g \in \mathbb{C}^{N_h} \):
\[
|\langle f, g \rangle| \leq \|f\|_{M^{-1}} \|g\|_{M^{-1}},
\]
with equality when \( g \) is a scalar multiple of \( f \). Therefore,
\[
\|f\|_{M^{-1}} = \frac{|\langle f, f \rangle|}{\|f\|_{M^{-1}}} = \sup_{g \in \mathbb{C}^{N_h}} \frac{|\langle f, g \rangle|}{\|g\|_{M^{-1}}}
= \sup_{g \in \mathbb{C}^{N_h}} \frac{|(M^{-1} g)^* f|}{\|g\|_{M^{-1}}}
= \sup_{u \in \mathbb{C}^{N_h}} \frac{|(M^{-1} Mu)^* f|}{\|Mu\|_{M^{-1}}}
= \sup_{u \in \mathbb{C}^{N_h}} \frac{u^* f}{\|u\|_M} = \sup_{u \in V_h} \|u\| \frac{|\langle f, u \rangle|}{\|u\|}.
\]

The Galerkin problem in \( V_h \) takes the form
\[
\text{(2.11) Find } u \in V_h \text{ such that } a(u, v) = f(v) \text{ for all } v \in V_h.
\]

Using the operator \( A_h : V_h \to V_h' \) defined as in (2.5) we see that the Galerkin problem is equivalent to finding a solution \( u \in V_h \) to the functional equation
\[
A_h u = f|_{V_h},
\]
where the right hand side is the restriction of \( f \) to \( V_h \). If \( u = \Phi u \) and \( f = \Phi f \), we obtain the linear-algebraic formulation of the Galerkin problem
\[
\text{(2.12) } Au = f,
\]
where the matrix \( A \in \mathbb{C}^{N_h \times N_h} \) and the vector \( f \in \mathbb{C}^{N_h} \) are given by
\[
A_{ij} = a(\phi_j, \phi_i), \quad 1 \leq i, j \leq N_h.
\]
\[
f = ((f, \phi_1), \ldots, (f, \phi_{N_h}))^T.
\]

In the case of Helmholtz problems with the sesquilinear form \( a \) given by (2.2), the matrix \( A \) from the linear system has the form
\[
A = S - k^2 M - ik N,
\]
where \( M \) is the mass matrix (2.9), and \( S, N \) are defined by
\[
S_{ij} = (\nabla \phi_j, \nabla \phi_i), \quad N_{ij} = (\phi_j, \phi_i)_{L^2(\partial D)} , \quad 1 \leq i, j \leq N_h.
\]
To finish this section, we discuss the approximability properties of the space $\mathcal{V}_h$. We assume that $\mathcal{V}_h$ is a space of piecewise linear Lagrange finite elements on a simplicial mesh (triangular or tetrahedral, in 2D or 3D respectively). Under this assumption, the Scott-Zhang interpolation operator $\Pi_{SZ}: H^1(\Omega) \to \mathcal{V}_h$ is well defined (see [17, Section 1.6.2]). Using the norm equivalence
\[
\|u\|_{1,k,\Omega} \sim k \|u\|_{L^2(\Omega)} + \|\nabla u\|_{L^2(\Omega)}
\]
and standard interpolation estimates (see [17, Lemma 1.130]) it can be shown that the Scott-Zhang interpolation operator $\Pi_{SZ}$ has the property that for all $w \in H^2(\Omega)$:
\[
(2.13) \quad \|w - \Pi_{SZ} w\|_{1,k,\Omega} \lesssim h \|w\|_{H^2(\Omega)} + hk \|w\|_{H^1(\Omega)}
\]
and for $w \in H^1(\Omega)$
\[
(2.14) \quad \|w - \Pi_{SZ} w\|_{1,k,\Omega} \lesssim (1 + hk) \|w\|_{H^1(\Omega)}.
\]

2.2. The shifted Laplace preconditioner and GMRES. Given $\varepsilon > 0$, we consider the Helmholtz problem with absorption (or “shifted Laplace” problem)
\[
(2.15) \quad \begin{cases}
-\Delta u - (k^2 + i\varepsilon) u &= f_1 \quad \text{in } \Omega, \\
\partial_n u -iku &= f_2 \quad \text{in } \Gamma.
\end{cases}
\]
with corresponding variational formulation
\[
(2.16) \quad \text{Find } u \in H^1(\Omega) \text{ such that } a_\varepsilon(u, v) = \langle f, v \rangle \text{ for all } v \in H^1(\Omega),
\]
with the sesquilinear form $a_\varepsilon$ and the antilinear functional $f$ given by
\[
(2.17) \quad a_\varepsilon(u, v) = \int_\Omega \nabla u \nabla \overline{v} - (k^2 + i\varepsilon) \int_\Omega u \overline{v} - ik \int_\Gamma u \overline{v},
\]
\[
(2.18) \quad \langle f, v \rangle = \int_\Omega f_1 \overline{v} + \int_\Gamma f_2 \overline{v}.
\]
The next theorem summarizes the properties of the sesquilinear form $a_\varepsilon$.

**Lemma 2.4.** [21, Lemma 3.1] Let $a_\varepsilon$ be the sesquilinear form of the shifted Laplace problem. The following properties hold:

1. The form $a_\varepsilon$ is continuous, that is, if $0 < \varepsilon \lesssim k^2$ then given $k_0 > 0$ there exists a constant $C_\varepsilon$ independent of $k, \varepsilon$ such that for all $k > k_0$, and $u, v \in H^1(\Omega)$
\[
|a_\varepsilon(u, v)| \leq C_\varepsilon \|u\|_{1,k,\Omega} \|v\|_{1,k,\Omega}.
\]
2. The form $a_\varepsilon$ is coercive, that is, if $0 < \varepsilon \lesssim k^2$ there exists a constant $\alpha > 0$ independent of $k, \varepsilon$ such that for all $k > 0$ and $u \in H^1(\Omega)$
\[
|a_\varepsilon(u, u)| \geq \alpha \frac{\varepsilon}{k^2} \|u\|_{1,k,\Omega}^2.
\]

The matrix $A_\varepsilon$ corresponding to the discrete shifted Laplace problem has the form
\[
A_\varepsilon = S - (k^2 + i\varepsilon) M - ik N = A - i\varepsilon M.
\]
We will use in our analysis a bound for the GMRES residuals based on the field of values. Recall that given a Hermitian positive definite (HPD) matrix $D \in \mathbb{C}^{N \times N}$ and an arbitrary $C \in \mathbb{C}^{N \times N}$, the field of values of $C$ in the inner product induced by $D$ is the set
\[
\mathcal{F}_D(C) = \left\{ \frac{(Cx, x)_D}{(x, x)_D} : x \in \mathbb{C}^N, x \neq 0 \right\}.
\]
Note that the spectrum of $C$ is contained in $\mathcal{F}_D(C)$ for any HPD matrix $D$. The Toeplitz-Hausdorff theorem states that the field of values is a convex, compact set \cite{32}, hence the following quantity is well defined:

$$\nu_D(C) = \min_{z \in \mathcal{F}_D(C)} |z|.$$ 

The next theorem by Elman \cite{9} shows that quantities related to the field of values can be used to bound the residuals of a minimum residual method in an arbitrary inner product.

**Theorem 2.5.** Let $r_0$ be the initial residual of the GMRES method applied to the linear system $Cu = f$

in the inner product induced by $D$. The $n$-th residual $r_n$ satisfies:

$$\frac{\|r_n\|_D}{\|r_0\|_D} \leq \left(1 - \frac{\nu_D(C)^2}{\|C\|_D^2}\right)^{n/2},$$

The proof of the GMRES bound based on the field of values relies on the fact that the residual minimization problem solved by a GMRES iteration in step $j$ can be restricted from the $j$-dimensional Krylov subspace to a one-dimensional subspace, so in general one cannot expect the bound (2.19) to be sharp in the intermediate steps of an iteration. Nevertheless, for (preconditioned) linear systems that result from finite element discretizations of PDEs, one can estimate the quantities $\nu_D(C)$ and $\|C\|$ using properties of the continuous problem and the finite element discretization, and in this way obtain rigorous proofs of parameter-independent GMRES convergence, see, e.g., \cite{2, 21, 29, 30, 41, 50}. Other convergence bounds for GMRES based on the field of values are surveyed in \cite{39}.

We close this section by recalling the main result in \cite{21} restricted to the kind of problems and domains that we are considering here (i.e., the interior impedance problem on convex polyhedral domains discretized with P1 finite elements). We remark that the analysis in \cite{21} includes also the exterior scattering problem and more general domains (star-shaped domains).

**Theorem 2.6** (Theorem 1.5 in \cite{21}). Let $\Omega$ be a convex polyhedron and suppose that the matrices $A$ and $A_\varepsilon$ result from the discretization of the Helmholtz and shifted Laplace problems 1.1 and 2.15 with P1 finite elements on a quasi-uniform sequence of meshes $\{T_h\}_{h \geq 0}$. Let $\varepsilon \lesssim k^2$, and $k_0, C > 0$. Then, there exist constants $C_1, C_2$ (independent of $h,k,\varepsilon$ but depending on $k_0, C$) such that, for $k > k_0$ with $hk^2 \geq C$,

$$\|I - AA_\varepsilon^{-1}\| \leq C_1 \frac{\varepsilon}{k},$$

$$\|I - A_\varepsilon^{-1}A\| \leq C_2 \frac{\varepsilon}{k},$$

and the GMRES method applied to the linear systems

$$A_\varepsilon^{-1}Au = A_\varepsilon^{-1}f, \; AA_\varepsilon^{-1}f = g,$$

converges in a number of iterations independent of $k$.

**3. A two-level preconditioner for the Helmholtz equation based on the shifted Laplacian.** In order to introduce the two-level preconditioner for the Helmholtz problem, we first review the basics of the multigrid method for finite element problems, following the presentation in \cite{3}. Let $V_H \subset V_h$ be a subspace of finite element functions of dimension $N_H$, corresponding to a coarse grid $T_H \subset T_h$. 

We denote by $\Phi_H$ and $\Phi'_H$ the coordinate mappings for $V_H$ and $V'_H$. Since $V_H \subset V_h'$ and $V'_h \subset V'_H$, the prolongation and restriction operators $P: V_H \rightarrow V_h'$ and $R: V'_h \rightarrow V'_H$ can be defined trivially, that is, $Pv = v$ for $v \in V_H$ and $Rf = f|_{V'_H}$ for $v \in V'_h$. Moreover, the coordinate mappings satisfy

$$\mathbb{C}^{N_h} \xrightarrow{\Phi_H} V_H \subset V_h \xrightarrow{\Phi^{-1}_h} \mathbb{C}^{N_h}, \text{ and } \mathbb{C}^{N_h} \xrightarrow{\Phi'_H} V'_h \subset V'_H \xrightarrow{\Phi'^{-1}_H} \mathbb{C}^{N_h},$$

and this gives the matrix form of the prolongation and restriction operators:

$$P = \Phi^{-1}_h \circ \Phi_H \in \mathbb{C}^{N_h \times N_h} \text{ and } R = (\Phi'_H)^{-1} \circ \Phi'_h \in \mathbb{C}^{N_h \times N_h}.$$ 

For $v_H \in V_H$ and $f \in V'_h$ we have

$$\langle f, Pv_H \rangle = \langle f, v_H \rangle = \langle Rf, v_H \rangle,$$

combining this relation and part (a) of Proposition 2.3 we conclude that the matrix form of the restriction operator is the Hermitian transpose of the prolongation: $R = P^*$. The Galerkin coarse grid matrix is defined as

$$A_H = RAP = P^*AP \in \mathbb{C}^{N_h \times N_h}.$$ 

Using the definition of the prolongation and restriction operators, it can be shown that $A_H$ corresponds to the Galerkin operator in the coarse space $V_H$ (Lemma 9.1 in [7]). The two-grid preconditioner $B_\varepsilon$ that we will study has the matrix form

$$B_\varepsilon = A_\varepsilon^{-1}(I - APA^{-1}_H) + PA^{-1}_H,$$

where $A_\varepsilon$ is the discrete shifted Laplacian in the space $V_h$. Note that it is not immediate that the preconditioner is non-singular, but this is in fact the case since it has been shown in [23] that a two-level preconditioner of the form is non-singular if and only if the matrix $P^*A_\varepsilon P$ is non-singular, and this holds because of the coercivity of $a_\varepsilon(\cdot,\cdot)$. We can now state our main result.

**Theorem 3.1.** There exists a constant $C > 0$ depending only on the domain $\Omega$ such that if the coarse grid size $H$ satisfies $HK^2 < C$ the GMRES method in the inner product induced by the inverse mass matrix $M^{-1}$ (2.10) applied to the preconditioned system

$$AB_\varepsilon g = f,$$

converges in a number of iterations independent of the wavenumber $k$.

**Outline of the proof of Theorem 3.1:** To prove Theorem 3.1 we will show that for a sufficiently small $H$ the field of values $\mathcal{F}_{M^{-1}}(AB_\varepsilon)$ is contained in a circle centered at 1 with radius independent of the wavenumber $k$. The norm induced by $M^{-1}$ is a natural norm to measure the residuals of the preconditioned system since it corresponds to the dual $L^2$ norm when $C^{N_h}$ is identified with the space of coordinates of $V_h'$ (see Proposition 2.3, (c)). Recalling that $M^{-1}$ is the matrix representation of the $L^2$ Riesz map (Proposition 2.3, (b)), we let $\hat{u} = M^{-1}g$, $\hat{A}_\varepsilon = AB_\varepsilon M$, and write the preconditioned system

$$AB_\varepsilon g = f,$$

as the equivalent system

$$\hat{A}_\varepsilon \hat{u} = f.$$
The latter linear system encodes a functional equation
\[ \hat{A}_\varepsilon \hat{u} = f, \]
where \( \hat{u} \in \mathcal{V}_h \) and \( f \in \mathcal{V}_h' \). This form allows us to formulate the preconditioned problem as a variational problem in \( \mathcal{V}_h \). The proof of Theorem 3.1 is divided in several parts:

1. First, we show in proposition 3.2 that the linear system of equations (3.3) corresponds to the formulation of a variational problem in the space \( \mathcal{V}_h \) for a certain sesquilinear form \( \hat{a}_\varepsilon(\cdot,\cdot) \).

2. Next, we prove in Lemma 3.3 that the form \( \hat{a}_\varepsilon \) and the field of values \( \mathcal{F}_{M^{-1}}(AB\varepsilon) \) are related by
\[
\mathcal{F}_{M^{-1}}(AB\varepsilon) = \left\{ \frac{\hat{a}_\varepsilon(u,u)}{(u,u)} : u \in \mathcal{V}_h \setminus \{0\} \right\}.
\]

3. In Lemma 3.5 we establish some properties of the sesquilinear form \( \hat{a}_\varepsilon \). In particular, we show that there exists an operator \( J : \mathcal{V}_h \to \mathcal{V}_h \) such that
\[
\hat{a}_\varepsilon(u,u) = 1 + i\varepsilon \frac{(Ju,u)}{(u,u)},
\]
for all \( u \in \mathcal{V}_h \), thus reducing the analysis of the field of values (3.4) to that of the field of values of the operator \( J \) in the \( L^2 \) inner product.

4. Lemmas 3.6 and 3.7 show that the form \( a(\cdot,\cdot) \) is coercive when restricted to a subspace of \( \mathcal{V}_h \). Here we use a variation of a duality argument due to Schatz [46], who used it to prove existence and quasioptimality of the Galerkin solution of a variational problem with a sesquilinear form that satisfies a Gårding inequality. This argument has appeared in several forms in the literature of finite element analysis of Helmholtz problems (see, e.g., [45], [19] and the references therein), and has been used before also in the analysis of two-grid methods for indefinite problems (see [5,30]).

5. Finally, we use the coercivity result for \( a(\cdot,\cdot) \) in Lemma 3.8 and the coercivity of \( a_\varepsilon(\cdot,\cdot) \) in Lemma 3.9 to give estimates that can be combined to bound the norm of \( \varepsilon J \) independently of the wavenumber, under the restriction that \( Hk^2 \leq C \) for a constant \( C \) independent of \( k \).

To proceed with the first step we formulate the preconditioned problem as a variational problem in \( \mathcal{V}_h \) using the strategy from [30], outlined in the next two propositions.

**Proposition 3.2.** Let \( a \) be the sesquilinear form from the Helmholtz problem defined in (2.2), let \( \tau : \mathcal{V}_h' \to \mathcal{V}_h \) the \( L^2 \)-Riesz map, and \( P : \mathcal{V}_H \to \mathcal{V}_h, \; R : \mathcal{V}_h' \to \mathcal{V}_H' \) be the prolongation and restriction operators respectively. The following statements hold:

(a) Let \( Q : \mathcal{V}_h \to \mathcal{V}_H \) be the solution operator to the problem: For \( u \in \mathcal{V}_h \) find \( Qu \in \mathcal{V}_H \) such that
\[
a(Qu,v) = \langle R\tau^{-1}u,v \rangle \quad \text{for all} \; v \in \mathcal{V}_H,
\]
then, the matrix form of the operator \( Q : \mathcal{V}_h \to \mathcal{V}_h \) is
\[
Q = PA_H^{-1}P^*M \in \mathbb{C}^{N_h \times N_h}.
\]

(b) Let \( N : \mathcal{V}_h \to \mathcal{V}_H \) be the solution operator to the adjoint-type problem: For \( u \in \mathcal{V}_h \) find \( Nu \in \mathcal{V}_H \) such that
\[
a(v,Nu) = a(v,u) \quad \text{for all} \; v \in \mathcal{V}_H,
\]
then, the matrix form of the operator \( N : \mathcal{V}_h \to \mathcal{V}_h \) is
\[
N = PA_H^{-1}P^*A^*.
\]
(c) Let \( J : V_h \to V_h \) be the solution operator to the problem: For \( u \in V_h \) find \( Ju \in V_h \) such that for all \( v \in V_h \)

\[
(3.8) \quad a_{\varepsilon}(Ju, v) = \langle \tau^{-1}u, (I - N)v \rangle,
\]

then, the matrix form of the operator \( J \) is

\[
J = A_{\varepsilon}^{-1}(I - N^*)M = A_{\varepsilon}^{-1}(I - APA_{H}^{-1}P^*)M.
\]

Proof. To prove (a), let \( u \in V_h, \ u = \Phi_h^{-1}u \in \mathbb{C}^N_h \) and \( w = \Phi_h^{-1}Qu \in \mathbb{C}^N_h \), recalling that \( M \) is the matrix representation of the inverse Riesz map \( \tau^{-1} \) we have that the condition (3.5) is equivalent to

\[
v^*A_Hw = v^*P^*Mu \quad \text{for all} \quad v \in \mathbb{C}^N_h,
\]

this gives \( w = A_H^{-1}PMu \) and \( Qu = PA_H^{-1}PMu \).

To show (b), let \( u \in V_h, \ u = \Phi_h^{-1}u \in \mathbb{C}^N_h \) and \( w = \Phi_h^{-1}Nu \in \mathbb{C}^N_h \). Then condition (3.7) is equivalent to

\[
A_H^*w = A_H^*P^*u \quad \text{and} \quad Nu = Pw = PA_H^{-1}P^*u.
\]

To prove (c), let \( u \in V_h, \ u = \Phi_h^{-1}u \in \mathbb{C}^N_h \) and \( Ju = \Phi_h^{-1}Ju \in \mathbb{C}^N_h \). Using the results of parts (a) and (b), we see that (3.8) is equivalent to

\[
v^*A_{\varepsilon}Ju = ((I - N^*)v)^*M u = v^*(I - N^*)Mu \quad \text{for all} \quad v \in \mathbb{C}^N_h,
\]

hence \( Ju = A_{\varepsilon}^{-1}(I - N^*)Mu = A_{\varepsilon}^{-1}(I - APA_{H}^{-1}P^*)Mu. \)

Lemma 3.3. Let \( A, B, M \) be the discrete Helmholtz operator, the two-grid preconditioner and the mass matrix in \( V_h \) respectively, and define \( \hat{A}_{\varepsilon} = AB_{\varepsilon}M \). The following properties hold:

(a) If \( \hat{a}_{\varepsilon} : V_h \times V_h \to \mathbb{C} \) is defined as

\[
(3.9) \quad \hat{a}_{\varepsilon}(u, v) = a((J + Q)u, v),
\]

then, for \( u, v \in \mathbb{C}^N_h \) and \( u = \Phi_h u, v = \Phi_h v \in V_h \):

\[
\hat{a}_{\varepsilon}(u, v) = \hat{a}_{\varepsilon}(u, v).
\]

(b) Given \( f \in V_h' \), consider the problem:

\[
(3.10) \quad \text{Find } u \in V_h \text{ such that } \hat{a}(u, v) = \langle f, v \rangle \text{ for all } v \in V_h.
\]

Then, the preconditioned system \( \hat{A}_{\varepsilon}u = f \) is the linear algebraic formulation of (3.10).

Proof. Part (a) follows from the definition of \( \hat{A}_{\varepsilon} \) and the matrix representations of \( J \) and \( Q \) given in the previous proposition. Part (b) is a straightforward consequence of (a). \( \square \)

Lemma 3.4. Let \( A \in \mathbb{C}^N_h \) be the discrete Helmholtz operator of the Galerkin problem in \( V_h \), \( M \in \mathbb{C}^N_h \) the mass matrix for the finite element space \( V_h \) and \( A_{\varepsilon} \in \mathbb{C}^N_h \) the discrete shifted Laplacian. Let \( B_{\varepsilon} \) be the two-grid preconditioner defined by

\[
B_{\varepsilon} = A_{\varepsilon}^{-1}(I - APA_{H}^{-1}P^*) + PA_H^{-1}P^*,
\]

and \( \hat{a}_{\varepsilon} \) the sesquilinear form defined in Proposition 3.4 (a). Then, the field of values of \( AB_{\varepsilon} \) in the inner product induced by \( M^{-1} \) is

\[
F_{M^{-1}}(AB_{\varepsilon}) = \left\{ \frac{\hat{a}_{\varepsilon}(u, u)}{(u, u)} : 0 \neq u \in V_h \right\}.
\]
Proof. Let \( z = \frac{(AB_ε,g)}{\|g\|_{M^{-1}}} \in \mathcal{F}_{M^{-1}}(AB_ε) \), \( u = M^{-1}g \) and \( u = Φ_h u \). Then,

\[
z = \frac{(AB_ε,g)}{\|g\|_{M^{-1}}} = \frac{g^* M^{-1} AB_ε \ g}{\|g\|_{M^{-1}}} = \frac{(M^{-1}g)^* AB_ε \ g}{\|g\|_{M^{-1}}} = \frac{u^* AB_ε \ Mu}{\|Mu\|_{M^{-1}}} = u^* \hat{A}_ε u \frac{\|Mu\|_{M^{-1}}}{(u,u)},
\]

where in the last step we used part (a) of Lemma 3.3. Since the correspondence

\[
g \mapsto u = M^{-1}g \mapsto u = Φ_h u
\]

is a bijection between \( \mathbb{C}^{N_h} \setminus \{0\} \) and \( V_h \setminus \{0\} \), the equality of sets (3.11) follows.

**Proposition 3.5 (Properties of \( \hat{a}_ε, Q, N \)).** Let \( \hat{a}_ε : V_h \times V_h \to \mathbb{C} \) be the sesquilinear form defined in Lemma 3.4, \( \tau : V'_h \to V_h \) the \( L^2 \) Riesz map and \( Q, N, J \) the operators introduced in Proposition 3.2.

(a) For all \( u \in V_h \)

\[
a(Qu,u) = \langle τ^{-1}u, Nu \rangle.
\]

(b) For all \( u \in V_h, v \in V_H \)

\[
a(v,(I-N)u) = 0.
\]

(c) For all \( u \in V_h \)

\[
(3.12) \quad \hat{a}_ε(u,u) = (u,u) + iε(Ju,u).
\]

**Proof.** We begin with part (a). From the definition of \( Q \) and \( N \) we have, for all \( u \in V_h \), and \( w, v \in V_H \):

\[
(3.13) \quad a(v,u) = a(v,Nu),
\]

\[
(3.14) \quad a(Qu,w) = \langle Rτ^{-1}u, w \rangle.
\]

Substituting \( v = Qu \) in (3.13) we obtain for \( u \in V_h \)

\[
a(Qu,u) = a(Qu,Nu),
\]

and setting \( w = Nu \) in (3.14) gives for \( u \in V_h \)

\[
a(Qu,Nu) = \langle Rτ^{-1}u, Nu \rangle = \langle τ^{-1}u, Nu \rangle.
\]

therefore \( a(Qu,u) = \langle τ^{-1}u, Nu \rangle \) holds for all \( u \in V_h \). The statement (b) is a consequence of the definition of the operator \( N \). To prove (c), recall that \( a(u,v) = a_ε(u,v) + iε(u,v) \), then using the definition of \( Q, J \) and part (a) we get

\[
\hat{a}_ε(u,u) = a((J + Q),u) = a(Ju,u) + a(Qu,u)
\]

\[
= a_ε(Ju,u) + iε(Ju,u) + a(Qu,u)
\]

\[
= \langle τ^{-1}u, (I-N)u \rangle + iε(Ju,u) + \langle τ^{-1}u, Nu \rangle
\]

\[
= \langle τ^{-1}u, u \rangle + iε(Ju,u) = \langle u,u \rangle + iε(Ju,u).
\]

The following two lemmas show that the sesquilinear form \( a \) is coercive when restricted to the range of the operator \( N \).
Lemma 3.6 (Bound on $L^2$ norm of $I - N$). For every $u \in \mathcal{V}_h$

(3.15) \[\| (I - N)u \|_{L^2(\Omega)} \lesssim Hk \| (I - N)u \|_{1,k,\Omega} .\]

Proof. We use a duality argument. Let $v = (I - N)u$ and $\phi$ the solution to the problem: Find $\phi \in H^1(\Omega)$ such that

\[ a(\phi, w) = (v, w) \text{ for all } w \in H^1(\Omega). \]

If $\Pi_{SZ} : H^1(\Omega) \to \mathcal{V}_H$ is the Scott-Zhang interpolation operator, for $w = v$ we have

\[ \| (I - N)u \|_{L^2(\Omega)}^2 = |(v, w)| = |a(\phi, (I - N)u)| = |a(\phi - \Pi_{SZ}\phi, (I - N)u)| \]

(Lemma 3.5, part (b)) \[ \lesssim \| \phi - \Pi_{SZ}\phi \|_{1,k,\Omega} \| (I - N)u \|_{1,k,\Omega} \]

(continuity of $a$).

Combining these estimates we obtain

\[ \| (I - N)u \|_{L^2(\Omega)}^2 \lesssim Hk \| (I - N)u \|_{L^2(\Omega)} \| (I - N)u \|_{1,k,\Omega}, \]

and dividing both sides of the inequality by $\| (I - N)u \|_{L^2(\Omega)}$ yields (3.15).

In the next lemma we show that if the coarse grid is sufficiently fine, the bilinear form $a$ is coercive when restricted to the range of the operator $I - N$.

Lemma 3.7. There exist constants $C, \alpha$ independent of $h, H, k$ such that if $Hk^2 < C$

\[ \alpha \| (I - N)u \|_{1,k,\Omega}^2 \leq \Re a((I - N)u, (I - N)u). \]

Proof. The sesquilinear form $a$ satisfies the Gårding inequality

\[ \Re a((I - N)u, (I - N)u) = \| (I - N)u \|_{1,k,\Omega}^2 - 2k^2 \| (I - N)u \|_{L^2(\Omega)}^2. \]

Combining this with Lemma 3.6 gives

\[ \| (I - N)u \|_{1,k,\Omega}^2 - 2k^2 \| (I - N)u \|_{L^2(\Omega)}^2 \geq \| (I - N)u \|_{1,k,\Omega}^2 - |\tilde{C}H^2k^4| \| (I - N)u \|_{1,k,\Omega}^2 \]

\[ = (1 - \tilde{C}H^2k^4) \| (I - N)u \|_{1,k,\Omega}^2 \]

where the constant $\tilde{C}$ comes from the estimate in Lemma 3.6. Let $\alpha \in (0, 1)$ and define $C = [(1 - \alpha)\tilde{C}^{-1}]^{1/2} > 0$. It is easy to see that if $Hk^2 < C$ we have $(1 - \tilde{C}H^2K^4) > \alpha$, therefore we have

\[ \alpha \| (I - N)u \|_{1,k,\Omega}^2 \leq \Re a((I - N)u, (I - N)u). \]

Lemma 3.8 (Using coercivity to bound $\| (I - N)Ju \|_{1,k,\Omega}$). Suppose that $H$ satisfies the requirements of the previous lemma. Then, for all $u \in \mathcal{V}_h$

(3.16) \[ \| (I - N)u \|_{1,k,\Omega} \lesssim (1 + Hk) \| u \|_{1,k,\Omega} \]
Proof. Let $H$ be such that $Hk^2$ is sufficiently small, i.e. $Hk^2 < C$ where $C$ is the constant from the previous lemma. For $u \in \mathcal{V}_h$, let $\Pi_{SZ} u$ be the Scott-Zhang interpolant of $u$ in $\mathcal{V}_H$. Combining Lemma 3.7, the orthogonality condition in part (b) of Lemma 3.5 and the continuity of $a$ we have

$$
\|(I - N)u\|_{1,k}^2 \lesssim \Re(a(I - N)u, (I - N)u) \quad \text{(Lemma 3.7)}
$$

$$
= \Re(a(u - \Pi_C u, (I - N)u) \quad \text{(Lemma 3.5, (b))}
$$

$$
\lesssim \|u - \Pi_C u\|_{1,k} \|(I - N)u\|_{1,k} \quad \text{(by continuity of $a$)}.
$$

With the estimate (2.14) for the Scott-Zhang interpolation operator, we obtain

$$
\|u - \Pi_{SZ} u\|_{1,k,\Omega} \lesssim (1 + Hk)\|u\|_{H^1(\Omega)}
$$

$$
\leq (1 + Hk)\|u\|_{1,k,\Omega}.
$$

Therefore,

$$
\|(I - N)u\|_{2,k,\Omega}^2 \lesssim \|u - \Pi_C u\|_{1,k,\Omega} \|(I - N)u\|_{1,k,\Omega}
$$

$$
\lesssim (1 + Hk)\|u\|_{1,k,\Omega} \|(I - N)u\|_{1,k,\Omega}.
$$

Dividing both sides by $\|(I - N)u\|_{1,k,\Omega}$ gives (3.16). \qed

Proof. For all $u \in L^2$, we have

$$
\|Ju\|_{L^2(\Omega)} \lesssim \frac{Hk^2}{\varepsilon} \|u\|_{L^2(\Omega)}
$$

(3.17)

Proof. We estimate as follows:

$$
\|Ju\|_{1,k,\Omega}^2 \leq \alpha^{-1} \frac{k^2}{\varepsilon} |a_\varepsilon(Ju, Ju)| \quad \text{(by coercivity of $a_\varepsilon$)}
$$

$$
= \alpha^{-1} \frac{k^2}{\varepsilon} |\langle \tau^{-1} u, (I - N)Ju \rangle| \quad \text{(by definition of $J$)}
$$

$$
= \alpha^{-1} \frac{k^2}{\varepsilon} |(u, (I - N)Ju)_{L^2(\Omega)}| \quad \text{(by definition of $\tau$)}
$$

$$
\leq \alpha^{-1} \frac{k^2}{\varepsilon} \|u\|_{L^2(\Omega)} \|(I - N)Ju\|_{L^2(\Omega)} \quad \text{(by the Cauchy-Schwarz ineq.)}
$$

$$
\lesssim \frac{Hk^3}{\varepsilon} \|u\|_{L^2(\Omega)} \|(I - N)Ju\|_{1,k,\Omega} \quad \text{(by Lemma 3.6)}
$$

$$
\lesssim \frac{Hk^3}{\varepsilon} (1 + Hk) \|u\|_{L^2(\Omega)} \|Ju\|_{1,k} \quad \text{(by Lemma 3.8)},
$$

dividing each side of the inequality by $\|Ju\|_{1,k}$ we obtain

$$
\|Ju\|_{1,k,\Omega} \lesssim \frac{Hk^3}{\varepsilon} (1 + Hk)\|u\|_{L^2(\Omega)}
$$

and, since $Hk \lesssim 1$,

$$
\|Ju\|_{L^2(\Omega)} \leq k^{-1} \|Ju\|_{1,k,\Omega} \lesssim \frac{Hk^2}{\varepsilon} \|u\|_{L^2(\Omega)}.
$$

We can now prove our main result.

Proof of Theorem 3.1. Combining Lemma 3.4 and part (c) of Proposition 3.5, we have that the field of values of the preconditioned matrix $A B_{\varepsilon}$ in the inner product induced by $M^{-1}$ is the set

$$
\mathcal{F}_{M^{-1}}(A B_{\varepsilon}) = \left\{ 1 + i\varepsilon \frac{(Ju, u)}{(u, u)} : u \in \mathcal{V}_h \setminus \{0\} \right\}.
$$

(3.18)
By Lemma 3.9, if $\varepsilon \lesssim k^2$ and $Hk^2$ is sufficiently small we have

$$\varepsilon \|Ju\| \lesssim Hk^2 \|u\|,$$

therefore, under this restriction on $H$, we have

$$\frac{\varepsilon |(Ju, u)|}{(u, u)} \leq \varepsilon \frac{\|Ju\| \|u\|}{\|u\|^2} \lesssim Hk^2,$$

so choosing $Hk^2$ sufficiently small the field of values $F_{M^{-1}}(AB_\varepsilon)$ lies inside a circle centered at 1 that does not contain the origin, with radius independent of the wavenumber $k$. Therefore, under this restriction on the coarse grid size, the distance of the field of values to the origin $\nu_{M^{-1}}(AB_\varepsilon)$ is independent of $k$. Moreover, the inequality (see Chapter 1 of [32])

$$\|AB_\varepsilon\|_{M^{-1}} \leq 2 \max_{z \in F_{M^{-1}}(AB_\varepsilon)} |z|,$$

implies that the norm $\|AB_\varepsilon\|_{M^{-1}}$ is bounded independently of $k$ as well. Therefore, the quantity

$$\frac{\nu_{M^{-1}}(AB_\varepsilon)}{\|AB_\varepsilon\|_{M^{-1}}}$$

is bounded away from zero independently of $k$. Using the residual bound (2.19) we conclude that, for $Hk^2$ sufficiently small, if the GMRES method in the inner product induced by $M^{-1}$ is applied to the linear system

$$AB_\varepsilon g = f,$$

the number of iterations required to obtain a reduction of the relative residual by a fixed tolerance is bounded by a constant, independent of the wavenumber $k$. \hfill \Box

In the next corollary we show that Theorem 3.1 also holds for the Euclidean inner product in the case of quasi-uniform meshes.

**Corollary 3.10.** Suppose, in addition to the hypothesis of Theorem 3.1, that the sequence of meshes $\{T_h\}_{h>0}$ is quasi-uniform. Then, there exists a constant $C > 0$ depending only on the domain $\Omega$ such that if the coarse grid size $H$ satisfies $Hk^2 \leq C$ the GMRES method in the Euclidean inner product applied to the preconditioned system

$$AB_\varepsilon g = f,$$

converges in a number of iterations independent of the wavenumber $k$.

**Proof.** Recall that for a sequence of quasi-uniform meshes the following norm equivalence holds:

$$\|v\|_M \sim h^{d/2} \|v\|_I,$$

for all $v \in C^N_h$ and $h > 0$, with the hidden constants independent of $h$ [3].

Using this fact and the characterization of the norm $\|\cdot\|_{M^{-1}}$ as the dual norm of $\|\cdot\|_M$ (Theorem 2.3, part (c)), it can be shown that

$$\|f\|_{M^{-1}} \sim h^{-d/2} \|f\|_I,$$

for all $f \in C^N_h$ and $h > 0$, with the hidden constants independent of $h$. A straightforward computation shows that

$$AB_\varepsilon = I + i\varepsilon \hat{J},$$
where $J = MA^{-1}(I - APA_H^{-1}P^*)$, therefore, the field of values of $AB_\varepsilon$ in the Euclidean inner product equals

$$\mathcal{F}_I(AB_\varepsilon) = \left\{ 1 + i\varepsilon \frac{(\hat{J}f,f)_I}{(f,f)_I} : 0 \neq f \in \mathbb{C}^{N_h} \right\}.$$  

(3.19)

Using the norm equivalence between $\| \cdot \|_I$ and $\| \cdot \|_{M^{-1}}$ from above we have

$$\|J\|_I \sim \|J\|_{M^{-1}},$$  

(3.20)

and combining part (c) of Proposition 3.2 with part (b) of Proposition 2.3, we have that the matrix $\hat{J}$ is the representation of the operator $\hat{J} = J^{-1}J\tau$ and it follows from Theorem 5.9 that

$$\|\hat{J}\|_{M^{-1}} = \sup_{f \in \mathbb{C}^{N_h}} \|\hat{J}f\|_{M^{-1}} = \sup_{u \in \mathbb{C}^{N_h}} \|Ju\|_M = \sup_{u \in \mathbb{V}_h} \|Ju\|_{L^2(\Omega)} \leq \frac{Hk^2}{\varepsilon}. $$

(3.21)

Therefore, for all $0 \neq f \in \mathbb{C}^{N_h}$ we have

$$\varepsilon \frac{|(\hat{J}f,f)_I|}{(f,f)_I} \leq \varepsilon \frac{\|\hat{J}f\|_I \|f\|_I}{\|f\|^2} \leq \varepsilon \frac{\|\hat{J}f\|_I}{\|f\|_I} \leq Hk^2,$$

where we have used the Cauchy-Schwarz inequality in the first step and (3.20) together with (3.21) in the last step. Using (3.19), we conclude that choosing $Hk^2$ sufficiently small the field of values $\mathcal{F}_I(AB_\varepsilon)$ lies inside a circle centered at 1 that does not contain the origin, with radius independent of the wavenumber $k$. The rest of the proof is similar to the last part of the proof of Theorem 5.9. \qed

4. Numerical Experiments. In this section we present the results of some numerical experiments that illustrate our theoretical results. The experiments were performed using MATLAB 2017b on a Macbook Pro with a 2,4 GHz Intel Core i5 processor. For the discretization with finite elements we have used the package FEM [6].

Experiment 1. In our first experiment we study the Helmholtz problem (1.1) on the domain $\Omega = (0,1)$. Although this problem leads to linear systems that are small and do not need to be solved with iterative methods, we use them here for illustrative purposes since for higher dimensional problems and large values of $k$ the computation of the field of values is very expensive. According to our theory, we choose for the discretization the number of interior gridpoints for the coarse mesh equal to $\lceil \frac{k^2}{2} \rceil$ which leads to a coarse problem of size $N_H = \lceil \frac{k^2}{2} \rceil + 2$ and a fine problem of size $N_h = 2N_H - 1$. We plot the field of values of the matrices $AA_\varepsilon^{-1}$ and $AB_\varepsilon$ using the method of Johnson [35], for increasing wavenumbers $k$ and various choices of $\varepsilon$. The main point of this experiment is to show that for increasing wavenumbers $k$ and a shift $\varepsilon \sim k^2$ the set $\mathcal{F}_I(AB_\varepsilon)$ remains bounded away from the origin, as predicted by the theory, in contrast to $\mathcal{F}_I(AB_\varepsilon)$, which moves closer to the origin as $k$ is increased. The results of this experiment are shown in Figures 1 and 2. Note that in this case the field of values $\mathcal{F}_I(AB_\varepsilon)$ is practically equal to a single point.

We repeat this computation choosing a number of interior gridpoints for the coarse mesh equal to $\lceil \frac{k^{3/2}}{2} \rceil$, which leads to a coarse problem of size $N_H = \lceil \frac{k^{3/2}}{2} \rceil + 2$ and a fine problem of size $N_h = 2N_H - 1$. The results of this experiment are shown in Figures 3 and 4. We see that that under a less restrictive condition on the meshsize the field of values of $AB_\varepsilon$ remains bounded away from zero as $k$ is increased. This is not predicted by our theory, but can be explained from the fact that it has been shown in [34] that the condition $N_h \sim k^{3/2}$ is sufficient to obtain a 'pollution-free' solution to the Helmholtz problem with the Galerkin method in 1-D. However, this has not been proved in higher dimensions.
Figure 1: Field of values of $\mathbf{A} \mathbf{A}^{-1}$ (left) and $\mathbf{A} \mathbf{B}_\epsilon$ (right) for a 1D Helmholtz problem and various values of $k$, with $\epsilon = k^2$ and $N_H \sim k^2$.

Figure 2: Field of values of $\mathbf{A} \mathbf{A}^{-1}$ (left) and $\mathbf{A} \mathbf{B}_\epsilon$ (right) for a 1D Helmholtz problem and various values of $k$, with $\epsilon = 5k^2$ and $N_H \sim k^2$.

Experiment 2. For our next experiment, we use GMRES to solve the interior impedance problem (1.1) on the unit square $(0,1)^2$, discretized with a uniform triangular grid. The right hand side is the constant vector of ones, and the initial guess the zero vector. We compare the complex shifted Laplace (CSL) preconditioner and the two-level preconditioner (TL) for various values of the shift $\epsilon$ and distinct coarsening levels. The CSL preconditioner is inverted with one multigrid F(1,1) cycle with $\omega$-Jacobi smoothing on all levels, where $\omega = 0.6$. The grid is chosen as follows: starting with a coarse grid with 3 points in each direction, we refine the grid uniformly until we obtain a number of points (in one dimension) larger than $\lceil \alpha k^{3/2} \rceil$ where $\alpha = 0.6$. The two-level preconditioner is tested using three different coarse grids. If $h$ is the fine meshsize in one dimension, the coarse meshsizes for the three different methods TL-1, TL-2 and TL-3 are $H = 2h, 4h, 8h$ respectively. Since the coarse grid matrices are still large, we use an incomplete LU factorization with drop tolerance of $10^{-6}$ to simulate the exact solve of the coarse grid systems. The results are shown in table 1. Although the meshsize scales with $k^{-3/2}$ (not with $k^{-2}$, as required by the theory), the number of iterations remains constant when the coarse meshes of meshsize $2h$ and $4h$ are used. For the coarse meshsize $8h$ the number of iterations increases linearly, although at a much slower rate than the number.
of iterations of GMRES preconditioned by the standalone shifted Laplacian. This experiment shows that the theoretical results are not sharp and that wave-number independent convergence can be obtained also with pollution-free meshes where the mesh size scales with \( k^{-3/2} \). Note that increasing the shift \( \varepsilon \) leads to an increase in the number of iterations with the standalone CSL preconditioner, and for the wavenumbers \( k = 80 \) and \( k = 100 \) with the shift \( \varepsilon = 5k^2 \) the GMRES method fails to reach the stopping criterion after 200 iterations. In contrast, the number of iterations remains bounded for the two-level preconditioner even for larger \( \varepsilon \).

Experiment 3. In our next experiment, we solve again the interior impedance problem (1.1) on the unit square \((0,1)^2\) with the same discretization, initial guess and right hand side as in our previous experiment, but this time we use a multilevel extension of the preconditioner, i.e., a multilevel Krylov method [13, 47]. This setting is not included in our theory but is more practically relevant since in realistic applications a two-grid preconditioner can be too expensive to apply. The multilevel preconditioner is implemented within a flexible GMRES (FGMRES) iteration [44], and every inexact coarse-grid solve (with a fixed small number of iterations) is performed by another FGMRES iteration, continuing (recursively) through all the
\[ \varepsilon = k^2 \quad \varepsilon = 2k^2 \quad \varepsilon = 5k^2 \]

| k | CSL | TL-1 | TL-2 | TL-3 | CSL | TL-1 | TL-2 | TL-3 | CSL | TL-1 | TL-2 | TL-3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 | 14 | 7 | 8 | 10 | 19 | 7 | 9 | 12 | 29 | 7 | 11 | 17 |
| 20 | 26 | 7 | 10 | 16 | 38 | 7 | 11 | 20 | 65 | 7 | 13 | 29 |
| 40 | 45 | 6 | 9 | 17 | 79 | 6 | 9 | 19 | 146 | 6 | 9 | 22 |
| 80 | 78 | 7 | 12 | 36 | 149 | 7 | 12 | 42 | - | 7 | 12 | 46 |
| 100 | 99 | 6 | 9 | 21 | 184 | 6 | 9 | 22 | - | 6 | 9 | 23 |

Table 1: Experiment 2. Number of GMRES iterations for the Helmholtz linear system preconditioned by the CSL and the two-level method (TL) for various values of the shift \( \varepsilon \) and different levels of coarsening.

| \( \varepsilon = k^2 \) | \( \varepsilon = 2k^2 \) |
|---|---|
| k | CSL | MK(8,4,2) | MK(6,4,2) | CSL | MK(8,4,2) | MK(6,4,2) |
|---|---|---|---|---|---|---|
| 20 | 26 | 0.42 | 7 | 0.77 | 7 | 0.54 | 7 | 0.70 | 7 | 0.48 |
| 40 | 45 | 3.18 | 6 | 3.54 | 6 | 2.39 | 79 | 6.79 | 7 | 2.83 | 6 | 2.14 |
| 80 | 78 | 31.78 | 7 | 8.85 | 7 | 6.75 | 149 | 93.73 | 7 | 9.19 | 7 | 7.09 |
| 120 | 130 | 377.22 | 7 | 32.81 | 7 | 30.14 | - | - | 7 | 33.28 | 8 | 30.02 |
| 160 | 142 | 8121.02 | 6 | 124.87 | 8 | 129.24 | - | - | 7 | 150.37 | 8 | 127.34 |

Table 2: Experiment 3. Number of GMRES iterations and computation time (in seconds) for the Helmholtz linear system preconditioned by the CSL and the multilevel Krylov method (MK).

Grids until the coarsest grid is reached. For more details on the implementation of multilevel Krylov methods we refer the reader to [13, 36].

To set up a multilevel Krylov method requires fixing a number of iterations for the intermediate levels. We denote by MK\((l,m,n)\) a method with \(l,m,n\) iterations in the second, third and fourth level grid respectively, and one iteration in the remaining coarser grid levels. In this experiment we compare the CSL with the multilevel Krylov methods MK\((8,4,2)\) and MK\((6,4,2)\). The results are shown in Table 2. Similarly as in the two-level case, the multilevel preconditioner outperforms the CSL and requires a constant number of iterations to reach the desired tolerance, even though the intermediate coarse solves are done only inexactly with a small number of iterations. Note also that decreasing the number of iterations in the second level only increases the number of outer iterations by one or two, and the computation times of the two methods MK\((8,4,2)\) and MK\((6,4,2)\) are very similar.

**Experiment 4.** In our final experiment we solve the impedance problem on the square \( \Omega = (0,1)^2 \) with a space-dependent wavenumber. This problem is adapted from [15]. The space-dependent wavenumber is given by

\[
k(x, y) = \begin{cases} 
(4/3)k_{ref} & \text{if } 0 \leq y < 0.2x + 0.2 \\
2k_{ref} & \text{if } 0.2x + 0.2 \leq y < -0.2x + 0.8 \\
k_{ref} & \text{if } -0.2x + 0.8 \leq y < 1 
\end{cases}
\]

where \( k_{ref} > 0 \) is a reference wavenumber. This function is depicted in Figure 5. As the reference wavenumber \( k_{ref} \) is varied, we choose the number of points for a uniform triangular mesh similarly as in the previous problems, with the number of points in one direction proportional to \( \lceil \alpha k_{ref}^{3/2} \rceil \) with \( \alpha = 1.1 \). The larger value of \( \alpha \) is chosen to take into account the fact that the maximum wavenumber over the domain is \( 2k_{ref} \). Similarly to the previous experiments, the CSL preconditioner is compared here with the multilevel Krylov methods MK\((8,4,2)\) and MK\((6,4,2)\). The results are shown in Table 3. Similarly to the previous experiments, the number of iterations of the CSL preconditioner grows linearly with the wavenumber, and the number of iterations with either of the multilevel Krylov methods remains constant.
5. Conclusions. In this paper we have presented a two-level shifted Laplace preconditioner for Helmholtz problems discretised with finite elements. We used the convergence theory of GMRES based on the field of values to rigorously establish that GMRES will converge in a number of iterations independent of the wavenumber if a condition of the form $HK^2 \leq C$ holds, for a constant $C$ independent of the wavenumber $k$ but possibly dependent on the size of the complex shift $\varepsilon$. We have also shown in numerical experiments that wavenumber independent convergence can also be obtained under the weaker condition $HK^{3/2} \leq C$, using a multilevel extension of the preconditioner (a multilevel Krylov method with inexact coarse grid solves), and for a test problem with heterogeneous wavenumber.

REFERENCES

[1] I. M. Babuška and S. A. Sauter, Is the Pollution Effect of the FEM Avoidable for the Helmholtz Equation Considering High Wave Numbers?, SIAM Rev., 42 (2000), pp. 451–484.
[2] M. Benzi and M. A. Olshanskii, Field-of-Values Convergence Analysis of Augmented Lagrangian Preconditioners for the Linearized Navier–Stokes Problem, SIAM J. Numer. Anal., 49 (2011), pp. 770–788.
[3] D. Braess, Finite Elements, Theory, Fast Solvers, and Applications in Solid Mechanics, Cambridge University Press, Apr. 2007.
[4] A. Brandt and I. Livshits, Wave-ray multigrid method for standing wave equations, ETNA, 6 (1997), pp. 162–181 (electronic).
[5] X. Cai and O. B. Widlund, Domain Decomposition Algorithms for Indefinite Elliptic Problems, SISC, 13 (1992), pp. 243–258.
[6] L. Chen, iFEM: an innovative finite element methods package in MATLAB, tech. report, University of California at Irvine, 2009.
[7] P.-H. Cocquet and M. J. Gander, How Large a Shift is Needed in the Shifted Helmholtz Preconditioner for its Effective Inversion by Multigrid?, SIAM J. Sci. Comput., 39 (2017), pp. A438–A478.
[8] S. Cools and W. Vanroose, Local Fourier analysis of the complex shifted Laplacian preconditioner for Helmholtz problems, Numer. Linear Algebra Appl. (NLAA), 20 (2013), pp. 575–597.
[9] H. H. Elman, *Iterative Methods for Large, Sparse, Nonsymmetric Systems of Linear Equations*, PhD thesis, Yale University, 1982.

[10] B. Engquist and L. Ying, *Sweeping preconditioner for the Helmholtz equation: Hierarchical matrix representation*, Communications on Pure and Applied Mathematics, 64 (2011), pp. 697–735.

[11] ———, *Sweeping Preconditioner for the Helmholtz Equation: Moving Perfectly Matched Layers*, 9 (2011), pp. 686–710.

[12] Y. A. Erlangga, *Advances in iterative methods and preconditioners for the Helmholtz equation*, Arch Comput Method Eng, 15 (2007), pp. 37–66.

[13] Y. A. Erlangga and R. Nabben, *Multilevel Projection-Based Nested Krylov Iteration for Boundary Value Problems*, SIAM J. Sci. Comput., 30 (2008), pp. 1572–1595.

[14] ———, *On a multilevel Krylov method for the Helmholtz equation preconditioned by shifted Laplacian*, ETNA, (2008).

[15] Y. A. Erlangga, C. Vuik, and C. W. Oosterlee, *On a class of preconditioners for solving the Helmholtz equation (extended version)*, arXiv, (2011), pp. arXiv:1105.2112–324.

[16] L. C. Evans, *Partial Differential Equations*, American Mathematical Soc., 2010.

[17] 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?*, Numer. Math., 131 (2015), pp. 567–614.

[18] M. J. Gander and H. Zhang, *A Class of Iterative Solvers for the Helmholtz Equation: Factorizations, Sweeping Preconditioners, Source Transfer, Single Layer Potentials, Polarized Traces, and Optimized Schwarz Methods*, SIAM Rev., 61 (2019), pp. 3–76.

[19] L. García Ramos, R. Keil, and R. Nabben, *Projections, Deflation, and Multigrid for Nonsymmetric Matrices*, SIAM J. Matrix Anal. & Appl., 41 (2020), pp. 83–105.

[20] L. García Ramos and N. Nabben, *On the Spectrum of Deflated Matrices with Applications to the Deflated Shifted Laplace Preconditioner for the Helmholtz Equation*, SIAM J. Matrix Anal. & Appl., 39 (2018), pp. 262–286.

[21] I. G. Graham, E. A. Spence, and E. Vainikko, *Domain decomposition preconditioning for high-frequency Helmholtz problems with absorption*, Math. Comp., 86 (2017), pp. 2089–2127.

[22] I. G. Graham, E. A. Spence, and J. Zou, *Domain Decomposition with local impedance conditions for the Helmholtz equation*, arXiv, (2018).

[23] A. Greenbaum, V. Pták, and Z. Strakoš, *Any Nonincreasing Convergence Curve is Possible for GMRES*, SIAM J. Matrix Anal. & Appl., 17 (1996), pp. 465–469.

[24] W. Hackbusch, *Elliptic Differential Equations*, vol. 18 of Springer Series in Computational Mathematics, Springer Berlin Heidelberg, Berlin, Heidelberg, 1992.

[25] A. Hannukainen, *Field of values analysis of preconditioners for the Helmholtz equation in lossy media*, arXiv, (2011).

[26] ———, *Field of values analysis of a two-level preconditioner for the Helmholtz equation*, arXiv, (2012).

[27] U. Hetmaniuk, *Stability estimates for a class of Helmholtz problems*, Commun. Math. Sci., 5 (2007), pp. 665–678.

[28] R. A. Horn and C. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, June 1994.

[29] F. Ihlenburg, *Finite Element Analysis of Acoustic Scattering*, vol. 132 of Applied Mathematical Sciences, Springer Science & Business Media, New York, 1998.

[30] F. Ihlenburg and I. M. Babuška, *Finite element solution of the Helmholtz equation with high wave number Part I: The h-version of the FEM*, Computers & Mathematics with Applications, 30 (1995), pp. 9–37.

[31] C. Johnson, *Numerical Determination of the Field of Values of a General Complex Matrix*, SIAM J. Numer. Anal., 15 (1978), pp. 595–602.

[32] R. Keil, R. Nabben, and D. B. Szyld, *Adaptive Multilevel Krylov Methods*, ETNA, 51 (2019), pp. 512–528.

[33] D. Lahaye, J. M. Tang, and C. Vuik, *Modern Solvers for Helmholtz Problems*, Birkhäuser, Mar. 2017.

[34] J. Liesen and Z. Strakoš, *Krylov Subspace Methods, Principles and Analysis*, Oxford Uni-
versity Press, Oct. 2012.

[39] J. Liesen and P. Tichý, The field of values bounds on ideal GMRES, arXiv, (2018).

[40] I. Livshits and A. Brandt, Accuracy properties of the wave-ray multigrid algorithm for Helmholtz equations, SIAM J. Sci. Comput., 28 (2006), pp. 1228–1251.

[41] D. Loghin and A. J. Wathen, Analysis of Preconditioners for Saddle-Point Problems, SIAM J. Sci. Comput., 25 (2004), pp. 2029–2049.

[42] J. Málek and Z. Strakoš, Preconditioning and the Conjugate Gradient Method in the Context of Solving PDEs, SIAM, Philadelphia, PA, Dec. 2014.

[43] J. M. Melenk, On Generalized Finite Element Methods, PhD thesis, University of Maryland, 1995.

[44] Y. Saad, A flexible inner-outer preconditioned GMRES algorithm, SIAM J. Sci. Comput., 14 (1993), pp. 461–469.

[45] S. A. Sauter, A refined finite element convergence theory for highly indefinite Helmholtz problems, Computing, (2006).

[46] A. H. Schatz, An observation concerning Ritz-Galerkin methods with indefinite bilinear forms, Math. Comp., 28 (1974), pp. 959–962.

[47] A. H. Sheikh, D. Lahaye, L. García Ramos, R. Nabben, and C. Vuik, Accelerating the shifted Laplace preconditoner for the Helmholtz equation by multilevel deflation, Journal of Computational Physics, 322 (2016), pp. 473–490.

[48] A. H. Sheikh, D. Lahaye, and C. Vuik, On the convergence of shifted Laplace preconditioner combined with multilevel deflation, Numer. Linear Algebra Appl. (NLAA), 20 (2013), pp. 645–662.

[49] E. A. Spence, Overview of Variational Formulations for Linear Elliptic PDEs, in Unified Transform for Boundary Value Problems, Society for Industrial and Applied Mathematics, Philadelphia, PA, Jan. 2015, pp. 93–159.

[50] G. Starke, Field-of-values analysis of preconditioned iterative methods for nonsymmetric elliptic problems, Numer. Math., 78 (1997), pp. 103–117.

[51] U. Trottenberg, C. W. Oosterlee, and A. Schüller, Multigrid, Academic Press, 2001.

[52] M. B. van Gijzen, Y. A. Erlangga, and C. Vuik, Spectral analysis of the discrete Helmholtz operator preconditioned with a shifted Laplacian, SIAM J. Sci. Comput., 29 (2007), pp. 1942–1958.

[53] L. A. Zepeda-Núñez and L. Demanet, The method of polarized traces for the 2D Helmholtz equation, (2015), pp. 1–54.