How to Choose the Shift in the Shifted Laplace Preconditioner for the Helmholtz Equation Combined with Deflation

D. Lahaye and C. Vuik

Abstract In recent work we showed that the performance of the complex shifted Laplace preconditioner for the discretized Helmholtz equation can be significantly improved by combining it multiplicatively with a deflation procedure that employs multigrid vectors. In this chapter we argue that in this combination the preconditioner improves the convergence of the outer Krylov acceleration through a new mechanism. This mechanism allows for a much larger damping and facilitates the approximate solve with the preconditioner. The convergence of the outer Krylov acceleration is not significantly delayed and occasionally even accelerated. To provide a basis for these claims, we analyze for a one-dimensional problem a two-level variant of the method in which the preconditioner is applied after deflation and in which both the preconditioner and the coarse grid problem are inverted exactly. We show that in case that the mesh is sufficiently fine to resolve the wave length, the spectrum after deflation consists of a cluster surrounded by two tails that extend in both directions along the real axis. The action of the inverse of the preconditioner is to shrink the length of the tails while at the same time rotating them and shifting the center of the cluster towards the origin. A much larger damping parameter than in algorithms without deflation can be used.

1 Introduction

The Helmholtz equation is a classical model equation for the propagation of waves. Examples of its use in various branches of science and engineering are given in the references cited. Fast and scalable methods to solve the linear system that arise after discretization are urgently needed.

The advent of the complex shifted Laplacian in [1, 2] led to a breakthrough in solver capabilities. The basis of this work was laid in [3] and [4]. A work in which similar ideas are proposed albeit with a different perspective is [5]. The complex shifted Laplacian was reconsidered in [6–10] and led to a boost in tackling various

D. Lahaye (✉) • C. Vuik
DIAM, TU Delft, Mekelweg 4, 2611 CN Delft, The Netherlands
e-mail: d.j.p.lahaye@tudelft.nl; c.vuik@tudelft.nl

industrial applications as documented in [11–20]. For a survey we refer to [21]. Recent publications on various solution approaches include [22–31].

The convergence of the complex shifted Laplace preconditioners is analyzed in [10, 32, 33]. The preconditioner introduces damping and shifts small eigenvalues of the preconditioned system away from the origin such that the outer Krylov will be faster to converge. As the wavenumber increases while the number of grid points per wavelength is kept constant however, the number of small eigenvalues becomes too large for the preconditioner to handle effectively, and the required number of outer Krylov iterations increases linearly with the wavenumber. This motivated the development in [34] of a deflation approach aiming at removing small eigenvalues using a projection procedure. The paper [34] considers deflating the preconditioned operator using the columns of the coarse to fine grid interpolation operator as deflation vectors. The multilevel extension of the method requires a Krylov subspace acceleration at each level. The method is therefore called a multilevel Krylov method. Some form of approximation is required to avoid the explicit construction of the preconditioned operator and to render the method computationally feasible. By the approximation proposed in [34], the projection property of the deflation operator is lost. This renders the results of a model analysis of the method using Fourier modes more tedious to interpret.

The method we developed in [35] borrows ideas from [34]. However, instead of deflating the preconditioned operator, we instead deflate the original Helmholtz operator. We subsequently combined the deflation and complex shifted Laplacian multiplicatively. We thus avoid having to approximate a computationally expensive operator and preserve the projection property of the deflation operator. This construction allows to

- add a term to the deflation operator to shift a set of eigenvalues away from zero without significantly disturbing the non-zero eigenvalues. This in turn allows to extend the deflation method to multiple levels in a multigrid hierarchy. This multilevel extension can be interpreted as a multigrid method in which at least formally the complex shifted Laplacian acts as a smoother. As in [34], the method requires a Krylov acceleration at each level of the multigrid hierarchy;
- deduce the algebraic multiplicity of the zero eigenvalue of the deflated operator in a model problem analysis. This facilitates the computation of the non-zero eigenvalues;
- re-use implementations of the multigrid approximate inversion of the complex shifted Laplacian to code the operation with the deflation operator. In this re-use one has to construct the coarser grid operators by Galerkin coarsening, to provide a Krylov acceleration on the intermediate coarse levels and to provide a flexible Krylov method on the finest level. This can be done with for instance the PETSc software library [36].

In our model problem analysis we employ 10 or 20 grid points per wavelength on the finest level. We also assume that Dirichlet boundary conditions are employed. These conditions render the boundaries reflective for outward traveling waves and act as a worst case in terms of convergence for situations in which Sommerfeld or

other types of absorbing boundary conditions are imposed. The spectrum of the operator after applying deflation without any preconditioning is real-valued and consists of a tight cluster surrounded by two tails. These tails spread in opposite directions along the real axis as the wavenumber increases. Elements in each tail correspond to the elements in the near-kernel of the Helmholtz operator on either side of zero. The role of the preconditioner is to scale and rotate the eigenvalues of the deflated operator. The spectrum of the operator after applying both deflation and preconditioning is complex-valued and consists of a cluster surrounded by two tails. These tails spread along a line in opposite directions in the complex plane away from the cluster with increasing wavenumber. The abscissa and slope of this line as well as the spread of the eigenvalues along this line are functions of the damping parameter in the preconditioner. Our results convincingly show that the use of deflation allows to significantly increase the damping parameter. Results in [35] give evidence for the fact the use of deflation results in a reduction of outer Krylov iterations. Results in [37] illustrate how the reduction of iterations leads to a significant speed-up of the computations.

This paper is structured as follows: in Sect. 2 we present the one-dimensional problem we intend to solve. In Sect. 3 we discuss the eigenvalues distribution of the complex shifted Laplace preconditioned matrix for the case of Dirichlet and Sommerfeld boundary conditions. In Sect. 4 we combine the preconditioner multiplicatively with a deflation operator that employs multigrid vectors. In Sect. 5 we derive closed form expressions for the eigenvalues of the preconditioned deflated system matrix through a model problem analysis. In Sect. 6 we present numerical results. Finally, we draw conclusions in Sect. 7.

2 Problem Formulation

The Helmholtz equation for the unknown field $u(x)$ on the one-dimensional domain $\Omega = (0, 1)$ reads

$$- \Delta u - (1 - \alpha \iota) k^2 u = g \text{ on } \Omega \,, \tag{1}$$

where $\iota, \alpha \in \mathbb{R}^+$, $k(x, y)$ and $g(x, y)$ are the imaginary unit, the damping parameter, the wave number and the source function, respectively. Here we are primarily be interested in solving the hard case without damping, i.e., the case in which $\alpha = 0$. We use the case with damping to illustrate various arguments. The wave number k, the frequency f and the angular frequency $\omega = 2\pi f$, the speed of propagation $c(x, y)$ and the wavelength $\lambda = \frac{c(x,y)}{f}$ are related by

$$k = \frac{2\pi}{\lambda} = \frac{\omega}{c} \,. \tag{2}$$

On the boundary $\partial\Omega$ we impose either homogeneous Dirichlet or first order Sommerfeld radiation boundary conditions. The latter are given by

$$\frac{\partial u}{\partial n} - \iota k u = 0 \text{ on } \partial\Omega \,. \tag{3}$$

This condition renders the end points of the one-dimensional domain transparent for outgoing waves. The spectrum of the coefficient matrix is such that the problem with Dirichlet boundary conditions is harder to solve than the problem with Sommerfeld boundary conditions as observed in e.g. [32]. This statement remains valid even if deflation is deployed.

2.1 Finite Difference Discretization

The finite difference discretization of the above problems on a uniform mesh with mesh width h using the stencil

$$[A_h] = \frac{1}{h^2} \left[-1 \; 2 - (1 - \alpha\iota)\kappa^2 \; -1 \right] \text{ where } \kappa = k\,h \,, \tag{4}$$

results after elimination of the boundary conditions in the linear system

$$A_h x_h = b_h \,, \tag{5}$$

where

$$A_h = -\Delta_h - (1 - \alpha\iota)k^2 I_h \in \mathbb{C}^{(n-1)\times(n-1)} \,. \tag{6}$$

The discretization requires due care to avoid the pollution error [38, 39]. This can be done by either imposing a minimum number of grid points per wavelength or by imposing the more stringent condition that $k^2 h^3$ remains constant.

2.2 Eigenvalues of the Discrete Helmholtz Operator

The linear system matrix A_h is sparse and symmetric. In the case of no damping and a sufficiently high wave-number (and thus a sufficiently fine mesh), A_h is indefinite and has a non-trivial near-null space. In case that the constant wave number one-dimensional problem is supplied with Dirichlet boundary conditions and is discretized using a uniform mesh with mesh width $h = 1/n$, the eigenvalues of A_h are easy to compute. As other types of boundary conditions introduce some form of damping, the resulting spectrum is generally more favorable for the convergence

of the outer Krylov acceleration. This implies that the use of Dirichlet boundary conditions acts as a worst-case scenario in the analysis of the convergence of Krylov methods via the spectrum. With these assumptions, the eigenvalues of A_h are the negatively shifted eigenvalues of the discrete Poisson operator and are given by Sheikh et al. [35] and Ernst and Gander [40]

$$\lambda^\ell(A_h) = \frac{1}{h^2}(2 - 2c_\ell - \kappa^2),\tag{7}$$

for $1 \le \ell \le n - 1$, where

$$c_\ell = \cos(\ell \pi h).\tag{8}$$

The corresponding eigenvectors are the orthogonal set of discrete sine modes denoted by ϕ^ℓ where $1 \le \ell \le n-1$. Each ϕ^ℓ is thus a vector with $n-1$ components indexed by i and given by

$$\phi_i^\ell = \sin(i\,h\,\ell\,\pi) \text{ for } 1 \le i \le n - 1.\tag{9}$$

The mutual orthogonality of the ϕ^ℓ's implies that the matrix A_h is normal and that theory for the convergence of GMRES applies. In case that damping is included in the Helmholtz equation, a purely imaginary contribution is added to (7), shifting the eigenvalues away from the origin. This increase in distance away from the origin makes the damped version easier to solve. From (7), it follows that the eigenvalues of the h^2-scaled operator $h^2 A_h$ vary continuously between

$$\lambda^1(h^2 A_h) \approx -\kappa^2 \text{ and } \lambda^{n-1}(h^2 A_h) \approx 4 - \kappa^2\tag{10}$$

where $c_1 \approx 1$ and $c_{n-1} \approx -1$, respectively.

In case that damping is added in the Helmholtz equation (1) by setting the damping parameter $\alpha > 0$, the imaginary component $\iota\alpha k^2$ is added to the eigenvalue expression (7). The eigenvectors remain unaltered. The eigenvalues are shifted upwards in the complex plane, and the problem becomes easier to solve iteratively.

In case that the Dirichlet boundary conditions are replaced by the Sommerfeld boundary conditions, both the eigenvalues and the eigenvectors change. An analytical computation of the spectrum in the limit of large k can be found in [41]. For the undamped case and for $k = 100$ and 10 grid points per wavelength, we computed the spectrum of the h^2-scaled matrix $h^2 A_h$ numerically. We plotted the sorted real and imaginary part of the eigenvalues as a function of the index ℓ in Fig. 1a and b, respectively. The sorting is such that different orderings are used in both figures. The real part closely resembles the expression for the Dirichlet case given by (7). The presence of a non-zero imaginary part in the eigenvalues render the use of Sommerfeld boundary conditions similar to the case of damping with Dirichlet boundary conditions with damping. The imaginary contribution shifts the eigenvalues away from the origin and renders the problem easier to solve

(a)

(b)

(c)

(d)
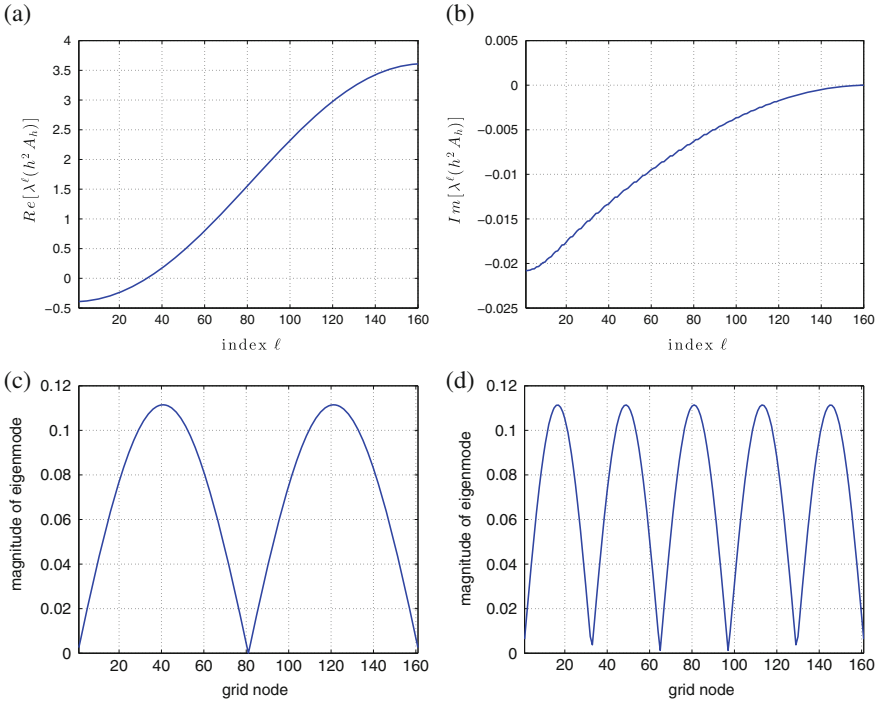
Fig. 1 Eigenvalues and magnitude of the second and fifth eigenmode of the discrete Helmholtz operator with Sommerfeld boundary conditions for $k = 100$ using 10 grid points per wavelength. (**a**) Sorted real part of eigenvalues. (**b**) Sorted imaginary part of eigenvalues. (**c**) Magnitude of second eigenmode. (**d**) Magnitude of fifth eigenmode

numerically. The magnitude of the second and fifth eigenvector is shown as a function of the grid index in Fig. 1c and d, respectively.

2.3 Multigrid Considerations

In the previous paragraph we assumed the mesh to be sufficiently fine to represent the wavelength. In this paper we will however consider an approach in which the Helmholtz equation without damping is discretized on a hierarchy of increasingly coarser meshes. This is the essential difference with CSLP precondition in previous work [1, 2] in which the original Helmholtz equation is discretized on the finest mesh only and in which the Helmholtz equation with damping only is coarsened.

The discretization of the undamped Helmholtz equation on the multigrid hierarchy motivates looking into properties of $h^2 A_h$ on the different levels of this hierarchy. We will derive bounds on the eigenvalues and a measure for the diagonal dominance of $h^2 A_h$ into account. For a fixed value of the wavenumber, each level

Table 1 Eigenvalue bounds and diagonal dominance measure of the h^2-scaled discretized Helmholtz operator $h^2 A_h$ for fixed wavenumber and for various values of the number of grid points per wavelength (gpw) on a multigrid hierarchy with five levels

κ	gpw	$\lambda^1(h^2 A_h) = -\kappa^2$	$\lambda^{n-1}(h^2 A_h) = 4 - \kappa^2$	$|2 - \kappa^2|$
0.3125	20	-0.0977	3.9023	1.9023
0.625	10	-0.3906	3.6094	1.6094
1.25	5	-1.5625	2.4375	0.4375
2.5	2.5	-6.25	-2.25	4.2500
5	1.25	-25	-21	23

of the hierarchy corresponds to a number of grid points per wavelength, and thus to a value of κ. Here we will consider the case in which the Helmholtz operator on the coarser levels is obtained via rediscretization and leave the case of Galerkin coarsening to later in the paper. To obtain bounds for the eigenvalues of the discrete Helmholtz operator on each level of the hierarchy in case of rediscretization, it suffices to substitute the corresponding value for κ into the bounds (10). As a measure for the diagonal dominance, we will adopt the absolute value of the diagonal element $|2-\kappa^2|$. For a multigrid hierarchy consisting of five levels obtained by standard $h \to 2h$ coarsening each level except for the coarsest, the eigenvalue bounds (10) and the value of $|2 - \kappa^2|$ are tabulated in Table 1. Motivating this measure for diagonal dominance is the fact that the weight of the off-diagonal elements does not change in traversing the hierarchy. The middle columns of Table 1 shows that in traversing the multigrid hierarchy from finest to coarsest level, the spectrum shifts in the negative direction and that on a sufficiently coarse level (here at 2.5 grid points per wavelength) even the largest eigenvalue becomes negative. From that level onward, the matrix ceases to be indefinite. The right-most column of Table 1 shows that the measure for the diagonal dominance initially decreases and increases again starting at a sufficiently coarse level (here again at 2.5 grid points per wavelength). At this coarsest level, the problem can be easily solved using the method of Jacobi for instance. For $k = 1000$ and for 10 grid points per wavelength for instance, the problem becomes definite and diagonally dominant starting from the third coarsest level onward. On these levels the use of complex solution algorithms such as the CSLP preconditioner is unnecessary. Similar ideas have already been discussed in [42]. We will return to this observation when discussing how to choose the damping parameter in the complex shifted Laplace preconditioner on the intermediate coarser levels.

3 Complex Shifted Laplace Preconditioner

In this section we introduce the complex shifted Laplace preconditioner [1, 2] and derive closed form expressions for the eigenvalues of the preconditioner and the preconditioned operator. We will in particular look into the effect of choosing a

very large damping parameter in the preconditioner. The information collected in
this section will serve as a reference for our model problem analysis in Sect. 5.

Denoting by β_2 the strictly positive damping parameter, the complex shifted
Laplace (CSLP) preconditioner can be written as

$$M_{h,\beta_2} = -\Delta_h - (1 - \iota\beta_2)\kappa^2 I_h \text{ where } \beta_2 \in \mathbb{R}^+ \setminus 0. \tag{11}$$

The value of β_2 needs to balance the quality of the preconditioner (favoring a
small value) with the ease of approximately inverting it (favoring a large value).
We assume that the boundary conditions implemented in A_h are imposed on M_{h,β_2}
as well. In both the case of Dirichlet and Sommerfeld boundary conditions, the
submatrices of M_{h,β_2} and A_h corresponding to the interior nodes differ by a scalar
multiple of the purely imaginary constant diagonal matrix $\iota\kappa^2 I_h$. In the absence and
presence of damping, the scalar involved is equal to β_2 and $\beta_2 - \alpha$, respectively.

3.1 Eigenvalues of the CSLP Preconditioner

Given that the matrices A_h and M_{h,β_2} differ by a purely imaginary constant diagonal
contribution, the eigenvalues of M_{h,β_2} are the eigenvalues of A_h shifted along the
imaginary axis. In both the case of Dirichlet and Sommerfeld boundary conditions,
the eigenvectors of M_{h,β_2} and A_h are the same. In the one-dimensional problem
with Dirichlet boundary conditions, we have that the eigenvalues of the h^2-scaled
preconditioner $h^2 M_{h,\beta_2}$ for $1 \leq \ell \leq n-1$ are given by

$$\lambda^\ell(h^2 M_{h,\beta_2}) = 2 - 2\,c_\ell - \kappa^2(1 - \iota\beta_2). \tag{12}$$

Let $\mu^\ell(h^2 M_{h,\beta_2})$ denote the inverse of this eigenvalue. Separating this inverse into a
real and imaginary part, we obtain

$$
\begin{aligned}
\mu^\ell(h^2 M_{h,\beta_2}) &= \frac{1}{2 - 2\,c_\ell - \kappa^2(1 - \iota\beta_2)} \\
&= \frac{2 - 2\,c_\ell - \kappa^2}{[2 - 2\,c_\ell - \kappa^2]^2 + \kappa^4\beta_2^2} - \iota\frac{\kappa^2\beta_2}{[2 - 2\,c_\ell - \kappa^2]^2 + \kappa^4\beta_2^2} \\
&= \frac{2 - 2\,c_\ell - \kappa^2}{|\lambda^\ell(h^2 M_{h,\beta_2})|} - \iota\frac{\kappa^2\beta_2}{|\lambda^\ell(h^2 M_{h,\beta_2})|} \\
&= \text{Re}[\mu^\ell(h^2 M_{h,\beta_2})] + \iota\text{Im}[\mu^\ell(h^2 M_{h,\beta_2})].
\end{aligned}
\tag{13}
$$

From these expressions we conclude that for $1 \leq \ell \leq n-1$

$$0 < \text{Re}[\mu^\ell(h^2 M_{h,\beta_2})] < 1 \quad \forall\beta_2 > 0, \tag{14}$$

$$-1 < \text{Im}[\mu^\ell(h^2 M_{h,\beta_2})] < 0 \quad \forall\beta_2 > 0, \tag{15}$$

and that in the limit for strong damping that

$$\mathrm{Re}[\mu^\ell(h^2 M_{h,\beta_2})] \to 0 \text{ as } \beta_2 \to +\infty, \tag{16}$$

$$\mathrm{Im}[\mu^\ell(h^2 M_{h,\beta_2})] \to 0 \text{ as } \beta_2 \to +\infty. \tag{17}$$

These results will be used to derive expressions for the eigenvalues of the preconditioned operator and the deflated preconditioned operator in the next paragraph and the next section, respectively.

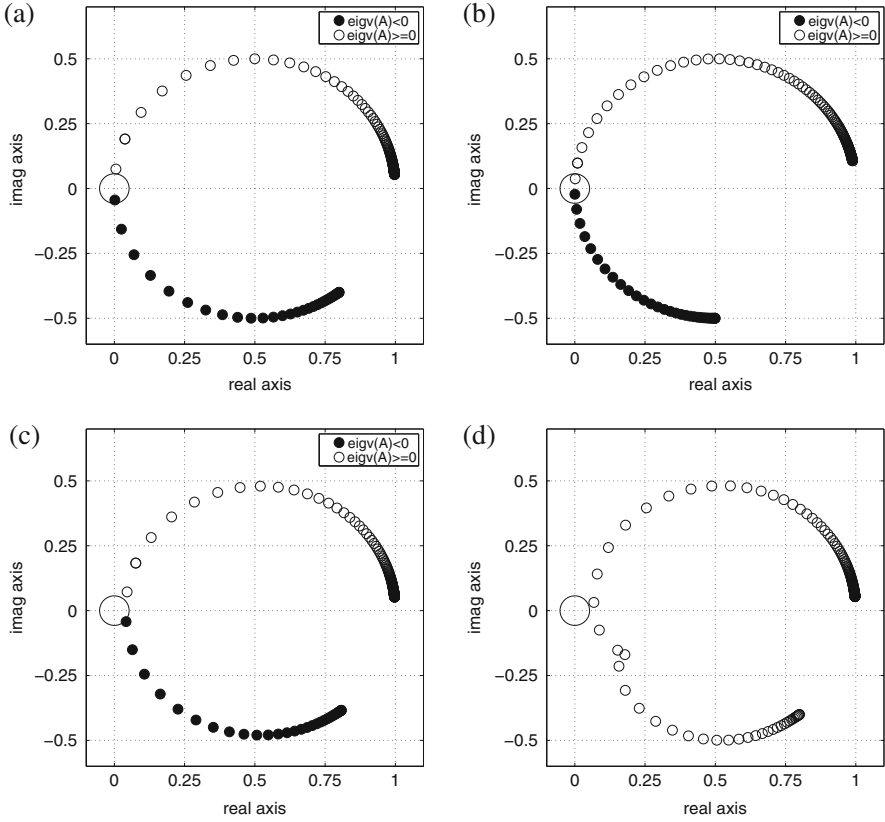## 3.2 Eigenvalues of the CSLP Preconditioned Operator

In deriving the eigenvalues of the preconditioned operator, we will assume the preconditioner to be inverted exactly. We will consider both the case of Dirichlet and Sommerfeld boundary conditions. In the former case, M_{h,β_2}^{-1} and A_h share the set of discrete sine modes given by (9). In the absence of damping, the eigenvalues of the preconditioned operator $M_{h,\beta_2}^{-1} A_h$ are the scaled and rotated eigenvalues of A_h given by

$$\lambda^\ell(M_{h,\beta_2}^{-1} A_h) = \mu^\ell(M_{h,\beta_2}) \, \lambda^\ell(A_h) \tag{18}$$

$$= \mathrm{Re}[\mu^\ell(h^2 M_{h,\beta_2})] \, \lambda^\ell(A_h) + \iota \, \mathrm{Im}[\mu^\ell(h^2 M_{h,\beta_2})] \, \lambda^\ell(A_h)$$

$$= \frac{\lambda^\ell(A_h)(2 - 2 c_\ell - \kappa^2)}{[2 - 2 c_\ell - \kappa^2]^2 + \kappa^4 \beta_2^2} - \iota \frac{\lambda^\ell(A_h)\kappa^2 \beta_2}{[2 - 2 c_\ell - \kappa^2]^2 + \kappa^4 \beta_2^2}.$$

This computation can be generalized to include non-zero damping (i.e., $\alpha = 0$) in the Helmholtz equation. In the case of Sommerfeld boundary conditions, we will resort to the numerical computations of the eigenvalues.

In Fig. 2 we plotted the eigenvalues $\lambda^\ell(M_{h,\beta_2}^{-1} A_h)$ for $1 \leq \ell \leq n-1$ in the complex plane for $k = 1000$ and 10 grid points per wavelength for four cases. In all four cases we highlighted a small region around the origin with a circle. In Fig. 2a we consider the case of Dirichlet boundary conditions without damping using $\beta_2 = 0.5$. We used shaded and non-shaded symbols to distinguish the eigenvalues that correspond to the index ℓ for which $\lambda^\ell(A_h)$ is negative and positive, respectively. Clearly both the real and imaginary part of $M_{h,\beta_2}^{-1} A_h$ are small for those values of ℓ for which $\lambda^\ell(A_h)$ shows a change of sign, i.e., for the values of ℓ that correspond to the near-kernel of A_h. These small eigenvalues hamper the convergence of the outer Krylov acceleration.

In Fig. 2b we consider again the case of Dirichlet boundary conditions without damping, this time using the larger value $\beta_2 = 1$. Comparing this figure with Fig. 2a confirms that for larger values of β_2 the eigenvalues $\mu^\ell(M_{h,\beta_2})$ and therefore the eigenvalues $\lambda^\ell(M_{h,\beta_2}^{-1} A_h)$ shift towards the origin. This causes the quality of

Fig. 2 Eigenvalues of the CSLP preconditioned operator for various preconditioning strategies for $k = 1000$ using 10 grid points per wavelength. (**a**) No damping and $\beta_2 = 0.5$. (**b**) No damping and $\beta_2 = 1$. (**c**) $\alpha = 0.02$ and $\beta_2 = 0.5$. (**d**) Sommerfeld and $\beta_2 = 0.5$

the preconditioner to degrade. The analysis in [32] shows that a similar shift of eigenvalues towards the origin occurs as k increases while β_2 and κ is kept constant.

In Fig. 2c we consider once more the case of Dirichlet boundary conditions, this time using a damping coefficient $\alpha = 0.02$. Comparing this figure with Fig. 2a shows that by introducing damping in the Helmholtz equation, the eigenvalues close to the origin shift towards the right in the complex plane. The increase of the magnitude of the eigenvalues that are small in size renders the preconditioned systems easier to solve.

In Fig. 2d finally we consider the case of Sommerfeld boundary conditions. This figure closely resembles to Fig. 2c. The introduction of the Sommerfeld boundary conditions is seen to introduce damping that causes a shift of small eigenvalues away from the origin. The preconditioned system again becomes easier to solve.

4 Combining Deflation and Precondition Multiplicatively

In this section we describe how we combine the complex shifted Laplacian preconditioner (CSLP) with a deflation technique. This approach is motivated by the fact that the convergence of the CSLP preconditioned Krylov acceleration is hampered by a few eigenvalues that are small in size. This is especially a problem in cases without damping. The objective of deflation is to remove these undesirable eigenvalues by a projection procedure. We describe the deflation technique on two levels, its extension to multiple levels, and the multiplicative combination of the preconditioner and the deflation technique.

4.1 Deflation by Two-Grid Vectors

Assuming p to be a non-zero natural number, we discretize the computational domain $\Omega = (0, 1)$ by a uniform mesh with $n = 2^p$ elements and mesh width $h = 1/n$ resulting in a fine mesh Ω^h. The discretization of the Helmholtz equation results after elimination of the boundary nodes in a discrete operator $A_h \in \mathbb{C}^{(n-1)\times(n-1)}$. Standard $h \to H = 2h$ coarsening of the fine mesh Ω^h results in a coarse mesh Ω^H with $n/2 - 1$ internal nodes. We denote by $Z_{h,H} \in \mathbb{R}^{(n-1)\times(n/2-1)}$ the coarse-to-fine grid interpolation operator. We employ a linear interpolation operator that, for fine grid points not belonging to the coarse grid, has the stencil

$$[Z_{h,H}] = \frac{1}{4} \left[1\ 2\ 1 \right]_H^h .$$ (19)

The columns of $Z_{h,H}$ are referred to as the deflation vectors. A deflation technique that uses these vectors is referred as two-grid deflation. The restriction operator is set equal to the full-weighting restriction operator. With this choice the restriction is equal to the transpose of the interpolation. This construction fits the theoretical framework of deflation methods.

The coarse grid operator E_H is constructed by Galerkin coarsening

$$E_H = Z_{h,H}^T A_h Z_{h,H} \in \mathbb{C}^{(n/2-1)\times(n/2-1)} .$$ (20)

The spectral properties of E_H will be discussed in the next section. We then define the coarse grid solve operator $Q_{h,H}$ as

$$Q_{h,H} = Z_{h,H} E_H^{-1} Z_{h,H}^T \in \mathbb{C}^{(n-1)\times(n-1)} ,$$ (21)

and the deflation operator $P_{h,H}$ as

$$P_{h,H} = I_h - A_h Q_{h,H} \in \mathbb{C}^{(n-1)\times(n-1)} .$$ (22)

The construction of E_H by Galerkin coarsening is such that $P_{h,H}$ satisfies the relation $P_{h,H}^2 = P_{h,H}$. $P_{h,H}$ is thus a projection and has eigenvalues 0 and 1. The matrix $P_{h,H}$ corresponds to the residual propagation matrix in a basic iterative solution method based on the splitting $A_h = Q_{h,H} - (Q_{h,H} - A_h)$ for the linear system (5).

By applying deflation, the linear system (5) is transformed into

$$P_{h,H} A_h x_h = P_{h,H} b_h \,. \tag{23}$$

The columns of $Z_{h,H}$ lie in the kernel of the deflated operator [43], i.e.,

$$P_{h,H} A_h Z_{h,H} = 0_{(n-1)\times(n/2-1)} \,. \tag{24}$$

The matrix $P_{h,H} A_h$ is thus singular and has a zero eigenvalue with multiplicity $n/2 - 1$. The computation of the remaining $n/2$ eigenvalues will be shown in the next section. The solution of the linear system (23) is defined up to a component in the kernel of $P_{h,H} A_h$. Such a solution can be found by a Krylov subspace method on the condition that in the application of $P_{h,H}$ the coarse linear system with E_H is solved to full precision at each iteration. What this condition implies and how it can be alleviated will be discussed in the next paragraph.

4.2 Multilevel Extension

For large problems in two or three dimensions, the exact inversion of the coarser grid matrix E_H is impractical and one has to resort to approximate coarser grid solves. Without proper care, this will however lead to the zero eigenvalue of $P_{h,H} A_h$ to be replaced by a cluster of near-zero eigenvalues. Such a cluster impedes the fast convergence of the outer Krylov acceleration. This can be avoided introducing a shift over a distance γ with $Q_{h,H}$ in the deflation operator $P_{h,H}$ and to define $P_{h,H,\gamma}$ as

$$P_{h,H,\gamma} = P_{h,H} + \gamma Q_{h,H} = I_h - A_h Q_{h,H} + \gamma Q_{h,H} \,. \tag{25}$$

With this definition, the equivalent of (24) for $P_{h,H,\gamma}$ is

$$P_{h,H,\gamma} A_h Z_{h,H} = \gamma Z_{h,H} \,, \tag{26}$$

i.e., γ is an eigenvalue with multiplicity $n/2 - 1$ of deflated matrix $P_{h,H,\gamma} A_h$. The value of γ is chosen once a choice for the preconditioner is made. We will give details in the next paragraph. The shift away from zero of the eigenvalues of the deflated matrix allows to solve the coarse grid system with coefficient matrix E_H approximately for instance by a recursive application of the two-level algorithm described. The use of a Krylov subspace solver on the coarser level requires to resort

to a flexible Krylov subspace solver on the fine level. The depth of the multigrid cycle can be limited in accordance to the discussion given in Sect. 2.

4.3 Multiplicative Combining of Preconditioning and Deflation

The CSLP preconditioner M_{h,β_2} and the deflation operator $P_{h,H,\gamma}$ including the shift with γ can be combined multiplicatively to construct a composite preconditioner. If the precondition is applied after the deflation, the linear system to be solved can be written as

$$B_{h,H,\beta_2,\gamma}\, x = (M_{h,\beta_2}^{-1}\, P_{h,H} + \gamma\, Q_{h,H})\, b\,, \tag{27}$$

where $B_{h,H,\beta_2,\gamma}$ is the preconditioned deflated operator

$$B_{h,H,\beta_2,\gamma} = (M_{h,\beta_2}^{-1}\, P_{h,H} + \gamma\, Q_{h,H})\, A_h\,. \tag{28}$$

In case that $\gamma = 1$, the matrix $I - B_{h,H,\beta_2,\gamma}$ is the error propagation matrix of a two-grid $V(0, 1)$ cycle applied to the linear system (5) with Galerkin coarsening and with M_{h,β_2} assuming at least formally the role of the smoother. In case that $\gamma \neq 1$, the composite preconditioner can be implemented as the additive combination of previously described $V(0, 1)$ cycle and a shift with $\gamma = 1$. Closed form expressions for the eigenvalues of $B_{h,H,\beta_2,\gamma}$ defined by (28) for $\gamma = 0$ and $\gamma \neq 0$ will be derived in the next section.

4.4 Comments on a Practical Implementation

An implementation of a multigrid approximate inversion CSLP as preconditioner can be easily extended to its combined use with the above described deflation technique. The multigrid components already in place can be recycled. A flexible Krylov acceleration on each level is required.

5 Model Problem Analysis

In this section we first derive closed form expressions for the eigenvalues of the Galerkin coarse grid operator E_H and the deflation operator $P_{h,H}$ defined by (20) and (22), respectively. Next we extend this analysis of the eigenvalues of the deflated operator $P_{h,H} A_h$ and the preconditioned deflated operator $M_{h,\beta_2}^{-1} P_{h,H} A_h$ given in the left-hand side of (23) and (27) with $\gamma = 0$, respectively. We consider the one-dimensional problem with Dirichlet boundary conditions with and without damping.

Based on the arguments on the resemblance of the eigenvalues in the problem with damping and with Sommerfeld boundary conditions in Sect. 2, we assume here that the problem with damping in the Helmholtz equation offers a good representation of the problem with Sommerfeld boundary conditions. We will derive expression for the eigenvalues by computing the action of these operators on the set of discrete sine modes defined by (9). This analysis is referred to a Rigorous Fourier analysis to distinguish it from a Local Fourier analysis in which the influence of the boundary conditions is not taken into account. Assuming Dirichlet boundary conditions, the set of sine modes ϕ_h^ℓ given by (9) forms a basis in which both the discrete operator A_h and the preconditioner M_{h,β_2} are diagonal. The analysis of the coarse operator E_H and the deflation operator $P_{h,H}$ requires care in handling the grid aliasing effect in the intergrid transfer operators $Z_{h,H}$ and $Z_{h,H}^T$. The eigenvalue expressions resulting from our analysis are fractions in which the eigenvalues of the coarse grid operator E_H appear in the denominator. These expressions form the basis for a subsequent analysis. The scattering of the eigenvalues along both sides of the real axis in case of 10 grid points per wavelength for instance can then be related to the near-kernel eigenvalues of the coarse grid operator E_H.

We assume the one-dimensional problem on $\Omega = (0, 1)$ with Dirichlet boundary conditions to be discretized by a uniform mesh with mesh width h. The coarse mesh obtained by standard coarsening then has a mesh width $H = 2h$. The use of Dirichlet boundary conditions was motivated in Sect. 2. We will perform a two-level analysis and assume that the Galerkin coarse grid operator E_H defined by (20) is inverted exactly. By reordering the eigenvectors of A_h defined by (9) in a standard way in $(\ell, n - \ell)$ pairs [44], we obtain the basis

$$V_h = \{(\phi_h^\ell, \phi_h^{n-\ell}) \mid \ell = 1, \ldots, n/2 - 1\} \cup \{\phi_h^{n/2}\}. \tag{29}$$

The modes ϕ_h^ℓ and $\phi_h^{n-\ell}$ form a pair by coarse grid aliasing. In the basis (29) first the deflation operator $P_{h,H}$, subsequently the deflated operator $P_{h,H} A_h$ and finally the preconditioned deflated operator $M_{h,\beta_2}^{-1} P_{h,H} A_h$ can be written in a block diagonal form. For a generic $(n - 1) \times (n - 1)$ matrix B, we will denote this block decomposition as

$$B = \left[(B)^\ell \right]_{1 \le \ell \le n/2}, \tag{30}$$

where for $1 \le \ell \le n/2 - 1$ the block $(B)^\ell$ is of size 2×2 and where for $\ell = n/2$ the block B^ℓ is a number. From this block diagonal form the eigenvalues of B can be computed with relative ease. For the restriction operator $Z_{h,H}^T$ and the coarse grid operator E_H that have size $(n/2 - 1) \times (n - 1)$ and $(n/2 - 1) \times (n/2 - 1)$ the size of the diagonal blocks reduces to 1×2 and 1×1, respectively.

In the following we will subsequently compute the eigenvalues of the Galerkin coarse grid operator E_H, the deflation operator $P_{h,H}$, the deflated operator $P_{h,H} A_h$ and finally the preconditioned deflated operator without shift $M_{h,\beta_2}^{-1} P_{h,H} A_h$ and with shift $(M_{h,\beta_2}^{-1} P_{h,H} + \gamma Q_{h,H}) A_h$. As before, we will especially look into large values of the damping parameter β_2.

5.1   Eigenvalues of the Coarse Grid Operator E_H

The block diagonal representation of the interpolation operator $Z_{h,H}^T$ in the basis (29) can be obtained by a standard computation [44]. Using the fact that $c_{n-\ell} = -c_\ell$, one obtains for $1 \leq \ell \leq n/2 - 1$ the 1×2 blocks

$$(Z_{h,H}^T)^\ell = 1/2 \left[(1 + c_\ell) \ -(1 - c_\ell) \right].$$ (31)

Given the $n/2$-th sine mode $\phi^{n/2}$ is equal to zero in all the coarse grid nodes and given the stencil (19), we have that

$$(Z_{h,H}^T)^{n/2} = 1/2.$$ (32)

The diagonal block of the discrete operator A_h in the basis (29) are for $1 \leq \ell \leq n/2 - 1$ given by

$$(A_h)^\ell = \begin{pmatrix} \lambda^\ell(A_h) & 0 \\ 0 & \lambda^{n-\ell}(A_h) \end{pmatrix} = \frac{1}{h^2} \begin{pmatrix} 2 - 2c_\ell - \kappa^2 & 0 \\ 0 & 2 + 2c_\ell - \kappa^2 \end{pmatrix},$$ (33)

and for $\ell = n/2$ by

$$(A_h)^{n/2} = (\kappa^2 - 2)/h^2.$$ (34)

The 1×1 diagonal blocks $(E_H)^\ell$ of the Galerkin coarse grid operator are obtained by the Galerkin coarsening of the individual blocks and results for all coarse grid values of ℓ including $\ell = n/2$ in

$$(E_H)^\ell = (Z_{h,H}^T)^\ell (A_h)^\ell (Z_{H,h})^\ell = \frac{1}{2h^2} [2(1 - c_\ell^2) - \kappa^2(1 + c_\ell^2)].$$ (35)

Given that in the basis (29) the operator E_H is diagonal, we have that the ℓ-th eigenvalue $\lambda^\ell(E_H)$ is equal to the ℓ-th diagonal block $(E_H)^\ell$. The eigenvalues of the H^2-scaled operator E_H are then for $1 \leq \ell \leq n/2$ given by

$$\lambda^\ell(H^2 E_H) = -(2 + 2c_\ell^2)\kappa^2 + 4 - 4c_\ell^2.$$ (36)

In the end points of the range from $\ell = 1$ to $\ell = n/2$, these expressions reduce to

$$\lambda^1(H^2 E_H) \approx -4\kappa^2 < 0 \text{ and } \lambda^{n/2}(H^2 E_H) \approx 4 - 2\kappa^2 > 0,$$ (37)

where $c_{\ell=1} \approx 1$ and $c_{\ell=n/2} = 0$, respectively. In the range of ℓ considered, close to zero eigenvalues of $\lambda^\ell(H^2 E_H)$ thus exist. The expressions (36) are the coarse grid equivalents of (10) and can be generalized to the case with damping by introducing a shift with $\iota \, \alpha \, \kappa^2$.
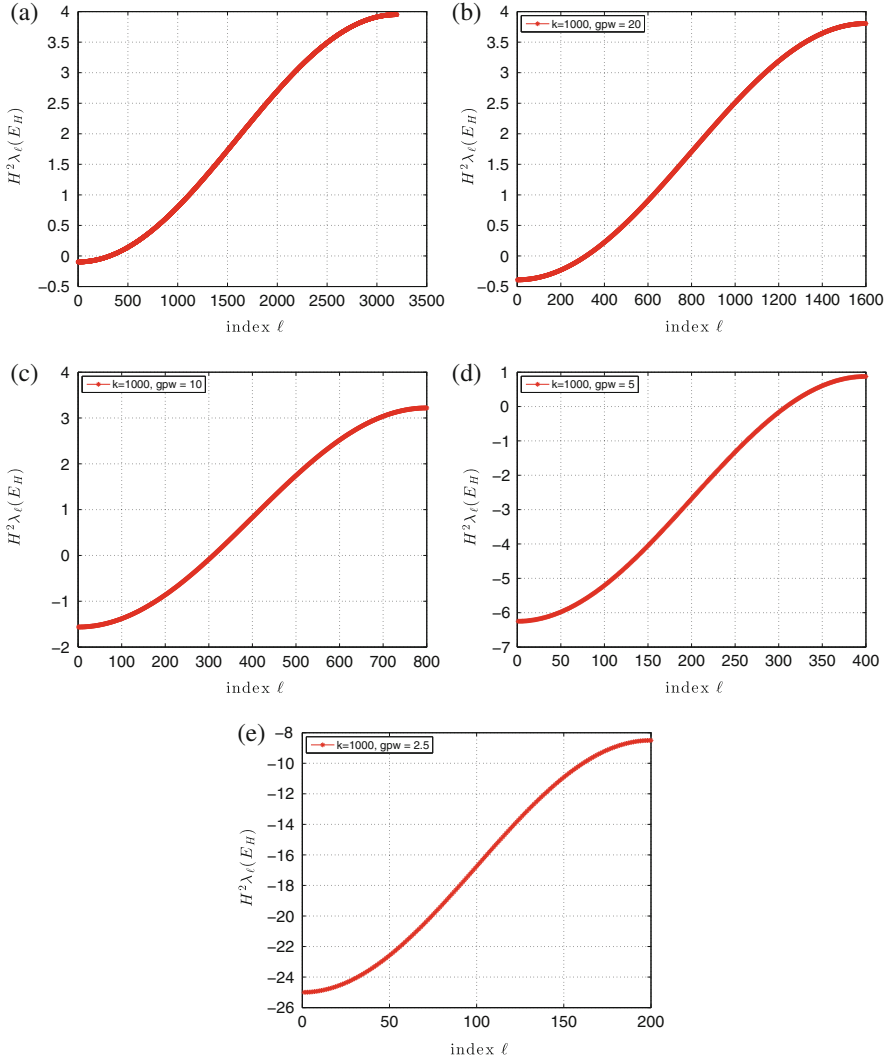
Fig. 3 Eigenvalues of the Helmholtz Galerkin coarse grid operator E_H for $k = 1000$ as a function of the index ℓ on a multigrid hierarchy consisting of five levels. On the finest level 40 grid points per wave length are employed. On each coarser level the number of grid points per wave length is halved. (**a**) 40 gpw; (**b**) 20 gpw; (**c**) 10 gpw; (**d**) 5 gpw; (**e**) 2.5 gpw

In Fig. 3 we plotted $\lambda^\ell(H^2 E_H)$ given by (37) as a function of ℓ for $k = 1000$ using various grid points per wavelength ranging from 40 (corresponding to $\kappa = 0.15625$) in the top left of the figure to 2.5 (corresponding to $\kappa = 2.5$) in the bottom of the figure. This figure clearly shows that in traversing the multigrid hierarchy from finest to coarser level (and thus increasing κ) the eigenvalues of the coarse grid operator

E_H shift towards the left on the real plane until all eigenvalues become negative and bounded away from zero on sufficiently coarse grids (sufficiently large values of κ). This is in accordance with the bounds (37).

The fact that in problems without damping the matrix E_H on fine and intermediate levels has several close to zero eigenvalues will play a central role in the subsequent analysis. By introducing damping, the issue of these small eigenvalues will be alleviated to some extent.

5.2   Eigenvalues of the Deflation Operator $P_{h,H}$

The diagonal blocks $(P_{h,H})^\ell$ of the deflation operator $P_{h,H}$ are for $1 \le \ell \le n/2 - 1$ given by

$$(P_{h,H})^\ell = I - (Z_{h,H})^\ell \, [(E_H)^\ell]^{-1} \, (Z_{h,H}^T)^\ell \, (A_h)^\ell \,, \tag{38}$$

and for $\ell = n/2$ by

$$(P_{h,H})^{n/2} = 1 \,. \tag{39}$$

As $P_{h,H}$ is a deflation operator, the individual 2×2 blocks $(P_{h,H})^\ell$ are projections as well and therefore have 0 and 1 as eigenvalue. Less immediate results will follow next.

5.3   Eigenvalues of the Deflated Operator $P_{h,H} A_h$

The diagonal blocks $(P_{h,H} A_h)^\ell$ of the deflated operator $P_{h,H} A_h$ are for $1 \le \ell \le n/2 - 1$ given by

$$\begin{aligned}
(P_{h,H} A_h)^\ell &= (P_{h,H})^\ell \, (A_h)^\ell \\
&= (A_h)^\ell - (Z_{h,H})^\ell \, [(E_H)^\ell]^{-1} \, (Z_{h,H}^T)^\ell \, [(A_h)^\ell]^2 \,,
\end{aligned} \tag{40}$$

and for $\ell = n/2$ by

$$(P_{h,H} A_h)^{n/2} = (2 - \kappa^2)/h^2 \,. \tag{41}$$

Property (24) translates on the 2×2 block level to $(P_{h,H} A_h)^\ell (Z_{h,H})^\ell = 0_{2 \times 1}$. Block $(P_{h,H} A_h)^\ell$ thus has a zero eigenvalue with multiplicity one. The remaining non-zero eigenvalue is then equal to the trace $\text{Tr}[(P_{h,H} A_h)^\ell]$. Computations show that for

$1 \leq \ell \leq n/2 - 1$ the elements of the h^2-scaled block $h^2 (P_{h,H} A_h)^\ell$ are given by

$$h^2 (P_{h,H} A_h)^\ell = \frac{1}{\lambda^\ell (H^2 E_H)} \begin{pmatrix} pa^\ell_{11,h,H} & pa^\ell_{12,h,H} \\ pa^\ell_{21,h,H} & pa^\ell_{22,h,H} \end{pmatrix}, \tag{42}$$

where the four matrix elements $pa^\ell_{ij,h,H}$ for $1 \leq i,j \leq 2$ are polynomials of second degree in κ^2. The diagonal elements $pa^\ell_{11,h,H}$ and $pa^\ell_{22,h,H}$ are more precisely given by

$$pa^\ell_{11,h,H} = (c_\ell - 1)^2 (2c_\ell - 2 + \kappa^2)(2c_\ell + 2 - \kappa^2) \tag{43}$$

$$pa^\ell_{22,h,H} = (c_\ell + 1)^2 (2c_\ell - 2 + \kappa^2)(2c_\ell + 2 - \kappa^2). \tag{44}$$

Observe that these expressions only differ by the sign in the first factor. As the off-diagonal elements $pa^\ell_{12,h,H}$ and $pa^\ell_{21,h,H}$ are not required to compute the trace, their detailed expression is omitted here. The non-zero eigenvalue of the ℓ-th block $h^2 (P_{h,H} A_h)^\ell$ is then given by

$$\begin{aligned} \lambda^\ell \left(h^2 P_{h,H} A_h\right) &= \mathrm{Tr}[h^2 (P_{h,H} A_h)^\ell] \\ &= \frac{1}{\lambda^\ell (H^2 E_H)} [pa^\ell_{11,h,H} + pa^\ell_{22,h,H}] \\ &= \frac{2}{\lambda^\ell (H^2 E_H)} (c_\ell^2 + 1)(2c_\ell - 2 + \kappa^2)(2c_\ell + 2 - \kappa^2). \end{aligned} \tag{45}$$

Give that the deflated operator involves a coarse grid solve, it is natural that the eigenvalue $\lambda^\ell (H^2 E_H)$ of the coarse grid operator appears in the denominator. In the range from $\ell = 1$ to $\ell = n/2$, the eigenvalues (45) decrease from

$$\lambda^1 \left(h^2 P_{h,H} A_h\right) \approx 4 - \kappa^2 \text{ to } \lambda^{n/2} \left(h^2 P_{h,H} A_h\right) = 2 - \kappa^2 \tag{46}$$

where $c_{\ell=1} \approx 1$ and $c_{\ell=n/2} = 0$, respectively. This decrease is however not monotone. Indeed, for those values of ℓ that corresponds to the near-kernel of $H^2 E_H$, the numerator of (45) is finite and the denominator very small. The eigenvalues $\lambda^\ell \left(h^2 P_{h,H} A_h\right)$ thus become very large for those values of ℓ. Stated differently, the closest-to-zero eigenvalue of $H^2 E_H$ causes of a vertical asymptote to appear in the plot of $\lambda^\ell \left(h^2 P_{h,H} A_h\right)$ versus ℓ.

In Fig. 4 we plotted $\lambda^\ell \left(h^2 P_{h,H} A_h\right)$ given by (45) as a function of ℓ for $k = 1000$. As in Fig. 3, we consider a sequence of five grids in which the number of grid points per wavelength ranges from 40 on the finest to 2.5 on the coarsest. On each level we consider a two-level construction of the deflation operator. For illustration purposes, we superimposed in each plot of $\lambda^\ell \left(h^2 P_{h,H} A_h\right)$ a plot of $\lambda^\ell \left(H^2 E_H\right)$. On the y-axis we labeled the extreme values $2 - \kappa^2$ and $4 - \kappa^2$. In the various subfigures of Fig. 4, the eigenvalues are seen to be bounded by $2 - \kappa^2$ and $4 - \kappa^2$, except for values close
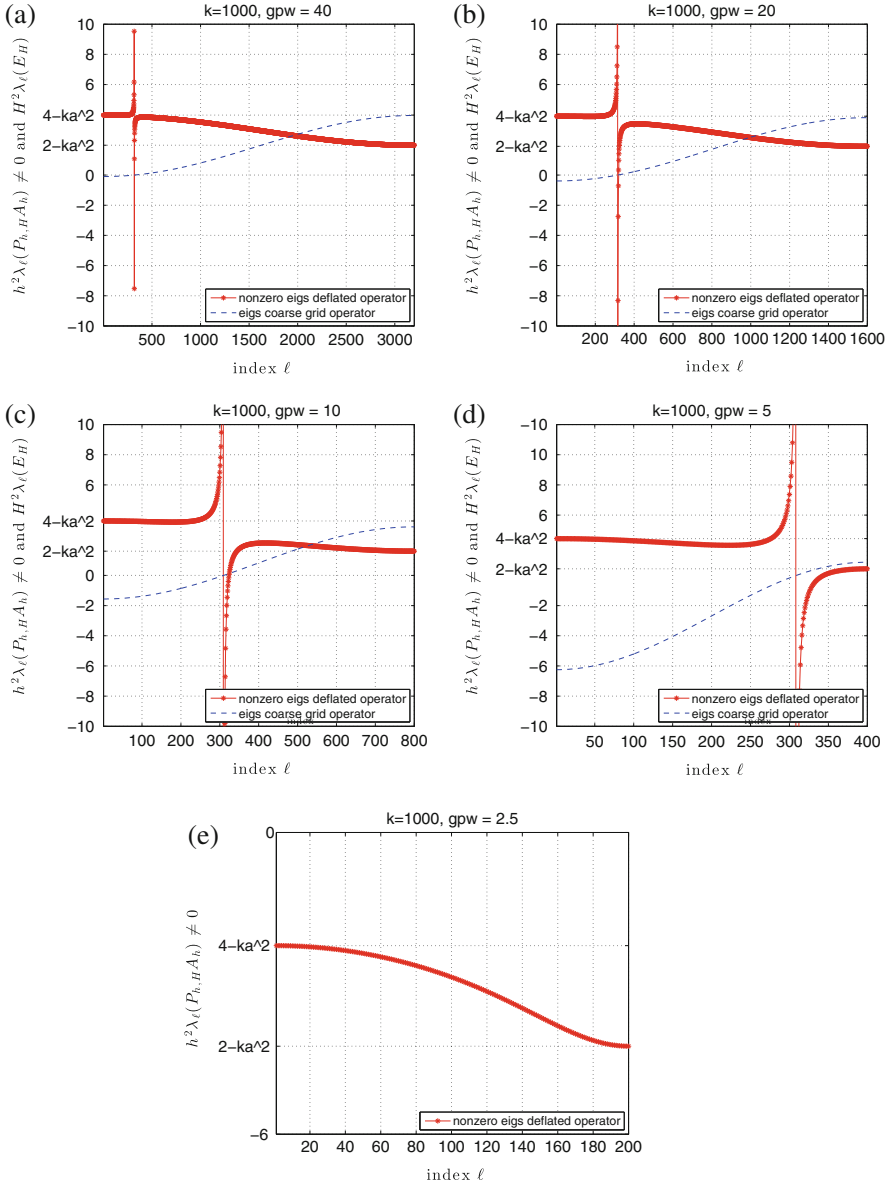
Fig. 4 Eigenvalues of the Helmholtz Galerkin coarse grid operator E_H (*dashed line*) and the two-grid deflated Helmholtz operator $P_{h,H}A_h$ (*solid line*) for $k = 1000$ as a function of the index ℓ on a multigrid hierarchy consisting of five levels. On the finest level 40 grid points per wave length are employed. On each coarser level the number of grid points per wave length is halved. (**a**) 40 gpw; (**b**) 20 gpw; (**c**) 10 gpw; (**d**) 5 gpw; (**e**) 2.5 gpw

to a vertical asymptote. The value of ℓ for which this asymptote occurs, is seen to coincide with the value of ℓ for which $\lambda^\ell \left(H^2 E_H \right) \approx 0$. This value of ℓ shifts towards the right on coarser meshes until disappearing completely. This agrees with our discussion of $\lambda^\ell \left(H^2 E_H \right)$ in the previous paragraph. The number of eigenvalues large in size of $h^2 P_{h,H} A_h$ is proportional to the number of close-to-zero eigenvalues of $H^2 E_H$. This number is small on the finest mesh considered in Fig. 4, increases on intermediate coarser meshes and is zero on the coarsest mesh.

The previous discussion implies that in a plot $\lambda^\ell \left(h^2 P_{h,H} A_h \right)$ on the real axis (instead of versus ℓ as before), the spectrum appears clustered between $2 - \kappa^2$ and $4 - \kappa^2$, except for two tails that spread along both sides of the real axis. The spread of these tails is inversely proportional to the size of the smallest eigenvalues of $H^2 E_H$. The number of elements in these tails is proportional to the number close-to-zero eigenvalues of E_H. For a fixed value of the wavenumber, the spread and number of elements in the tail vary with the number of grid points per wavelength employed.

5.4 Eigenvalues of the Preconditioned Deflated Operator $M_{h,\beta_2}^{-1} P_{h,H} A_h$

The diagonal blocks of the preconditioned deflated operator $(M_{h,\beta_2}^{-1} P_{h,H} A_h)^\ell$ are for $1 \leq \ell \leq n/2 - 1$ given by

$$(M_{h,\beta_2}^{-1} P_{h,H} A_h)^\ell = (M_{h,\beta_2}^{-1})^\ell \, (P_{h,H} A_h)^\ell . \tag{47}$$

and for $\ell = n/2$ by

$$(M_{h,\beta_2}^{-1} P_{h,H} A_h)^{n/2} = \frac{2 - \kappa^2}{2 - \kappa^2(1 - \iota\beta_2)} . \tag{48}$$

From the singularity of the block $(P_{h,H} A_h)^\ell$ and (47) follows that the ℓ-th diagonal block $(M_{h,\beta_2}^{-1} P_{h,H} A_h)^\ell$ is singular as well. Its non-zero eigenvalue can thus be computed by merely computing its trace. The diagonal blocks of the h^2-scaled preconditioner $h^2 M_{h,\beta_2}$ in the basis (29) are for $1 \leq \ell \leq n/2 - 1$ given by

$$(h^2 M_{h,\beta_2})^\ell = \begin{pmatrix} \lambda^\ell(h^2 M_{h,\beta_2}) & 0 \\ 0 & \lambda^{n-\ell}(h^2 M_{h,\beta_2}) \end{pmatrix} \tag{49}$$

Assuming the preconditioner to be inverted exactly, the diagonal blocks of the inverse of the preconditioner are given by

$$(h^{-2} M_{h,\beta_2}^{-1})^\ell = \begin{pmatrix} \mu^\ell(h^2 M_{h,\beta_2}) & 0 \\ 0 & \mu^{n-\ell}(h^2 M_{h,\beta_2}) \end{pmatrix} . \tag{50}$$

The non-zero eigenvalue of the ℓ-th diagonal block $(M_{h,\beta_2}^{-1} P_{h,H} A_h)^\ell$ is then given by

$$\lambda^\ell(M_{h,\beta_2}^{-1} P_{h,H} A_h) = \mathrm{Tr}[(M_{h,\beta_2}^{-1} P_{h,H} A_h)^\ell] \tag{51}$$

$$= \frac{1}{\lambda^\ell(H^2 E_H)} \big[\mu^\ell(h^2 M_{h,\beta_2}) pa_{11,h,H}^\ell +$$

$$\mu^{n-\ell}(h^2 M_{h,\beta_2}) pa_{22,h,H}^\ell \big].$$

Observe that the eigenvalues of the Galerkin coarse grid operator E_H appear in the denominator. The coefficients $pa_{11,h,H}^\ell$ and $pa_{22,h,H}^\ell$ are real-valued. It is thus easy to split the non-zero eigenvalue $\lambda^\ell(M_{h,\beta_2}^{-1} P_{h,H} A_h)$ is a real and imaginary part and obtain for $1 \le \ell \le n/2 - 1$

$$\mathrm{Re}\left[\lambda^\ell(M_{h,\beta_2}^{-1} P_{h,H} A_h)\right] = \frac{1}{\lambda^\ell(H^2 E_H)} \bigg[\mathrm{Re}\left[\mu^\ell(h^2 M_{h,\beta_2})\right] pa_{11,h,H}^\ell +$$

$$\mathrm{Re}\left[\mu^{n-\ell}(h^2 M_{h,\beta_2})\right] pa_{22,h,H}^\ell \bigg], \tag{52}$$

and

$$\mathrm{Im}\left[\lambda^\ell(M_{h,\beta_2}^{-1} P_{h,H} A_h)\right] = \frac{1}{\lambda^\ell(H^2 E_H)} \bigg[\mathrm{Im}\left[\mu^\ell(h^2 M_{h,\beta_2})\right] pa_{11,h,H}^\ell +$$

$$\mathrm{Im}\left[\mu^{n-\ell}(h^2 M_{h,\beta_2})\right] pa_{22,h,H}^\ell \bigg]. \tag{53}$$

Next we will use the results derived in Sect. 3 to find upper bounds for this real and imaginary part. These bounds will allow us to argue how the preconditioner transforms the eigenvalues of the deflated operator and how in particular the value of the damping parameter β_2 affects this transformation.

We start by considering the real part (52). The inequality (14) states that both $\mathrm{Re}\left[\mu^\ell(h^2 M_{h,\beta_2})\right]$ and $\mathrm{Re}\left[\mu^{n-\ell}(h^2 M_{h,\beta_2})\right]$ are bounded above by 1. We thus have that

$$\mathrm{Re}\left[\lambda^\ell(M_{h,\beta_2}^{-1} P_{h,H} A_h)\right] \le \frac{1}{\lambda^\ell(H^2 E_H)} \bigg[pa_{11,h,H}^\ell + pa_{22,h,H}^\ell \bigg] = \lambda^\ell(h^2 P_{h,H} A_h), \tag{54}$$

where we have used expression (45) for $\lambda^\ell(h^2 P_{h,H} A_h)$. The distance between $\mathrm{Re}\left[\lambda^\ell(M_{h,\beta_2}^{-1} P_{h,H} A_h)\right]$ and $\lambda^\ell(h^2 P_{h,H} A_h)$ can be increased by taking large values of β_2. This is particularly interesting for those values of ℓ for which $\lambda^\ell(h^2 P_{h,H} A_h)$ is large in size, i.e., for those values of ℓ corresponding to the near-null space of E_H. By taking large values of β_2, these large values of $\lambda^\ell(h^2 P_{h,H} A_h)$ can be reduced, i.e., brought back to the center of the cluster of the eigenvalues by the action of

the preconditioner. Eigenvalues $\lambda^\ell(h^2 P_{h,H} A_h)$ that lie in the interval from $2 - \kappa^2$ to $4 - \kappa^2$ are mapped to eigenvalues with a real part in a bounded interval. The length of this interval shrinks and its midpoint shift to zero of β_2 increases. Despite of this shift to zero, a larger damping than in the case without deflation can be chosen.

Next we consider the imaginary part (53). We use the expression in (18) to rewrite the imaginary parts $\text{Im}\left[\mu^\ell(h^2 M_{h,\beta_2})\right]$ and $\text{Im}\left[\mu^{n-\ell}(h^2 M_{h,\beta_2})\right]$ to obtain that

$$\text{Im}\left[\lambda^\ell(M_{h,\beta_2}^{-1} P_{h,H} A_h)\right] = \frac{-\beta_2 \kappa^2}{\lambda^\ell(H^2 E_H)}\left[\frac{pa_{11,h,H}^\ell}{|\lambda^\ell(h^2 M_{h,\beta_2})|} + \frac{pa_{22,h,H}^\ell}{|\lambda^{n-\ell}(h^2 M_{h,\beta_2})|}\right]. \quad (55)$$

On meshes with a sufficient number of grid points per wavelength, κ^2 is a small number. Expression (55) thus yields a small imaginary part except for those values of ℓ for which $\lambda^\ell(H^2 E_H) \approx 0$ and thus also $\lambda^\ell(h^2 M_{h,\beta_2}) \approx 0$. Eigenvalues $\lambda^\ell(h^2 P_{h,H} A_h)$ inside and outside the interval from $2 - \kappa^2$ to $4 - \kappa^2$ are mapped to eigenvalues with an imaginary part that is small and that increases proportionally to β_2, respectively.

In Fig. 5 we plotted the non-zero eigenvalues of $M_{h,\beta_2}^{-1} P_{h,H} A_h$ in the complex plane for $k = 1000$ and $\beta_2 = 1$. In traversing the hierarchy from the finest to the coarsest level, the range in the real part of the eigenvalues is seen to first increases to subsequently decrease starting from the third coarsest level with five grid point per wavelength. This in accordance with our previous discussion.

We can summarize the discussion by stating the action of the preconditioner is to contract and rotate the eigenvalues of the deflated operator. This is illustrated in Fig. 6 in which the eigenvalues $\lambda^\ell(M_{h,\beta_2}^{-1} P_{h,H} A_h)$ are plotted in the complex plane for $k = 1000$ and ten grid point per wavelength.

6 Numerical Results

In this section we present numerical results for the one-dimensional problem on the unit interval and the two-dimensional problem on the unit square. For both problem we consider the problem without damping supplied with homogeneous Dirichlet boundary conditions discretized using either 10 or 20 grid points per wavelength. We adopt a two-level variant of the deflation operator and assume the both preconditioner on the finest level and the coarse grid operator to be inverted exactly. As outer Krylov we run full GMRES with a zero initial guess. We declare convergence at the k-th iteration if the relative residual norm $\|A_h x_h^k - b_h\|_2 / \|b_h\|_2$ drops below 10^{-6}. We compare the following five algorithmic variants. The first variant merely employs A-DEF1 (without CSLP) as a preconditioner. The second, third and fourth variant combine A-DEF1 and CSLP multiplicative with β_2 equal to 0.5, 1 and 10, respectively. The fifth variant employs $\beta_2 = 10$ and approximates the CSLP preconditioner by its diagonal. The required numbered GMRES iterations for the one and two-dimensional problem are given in Tables 2 and 3, respectively. For
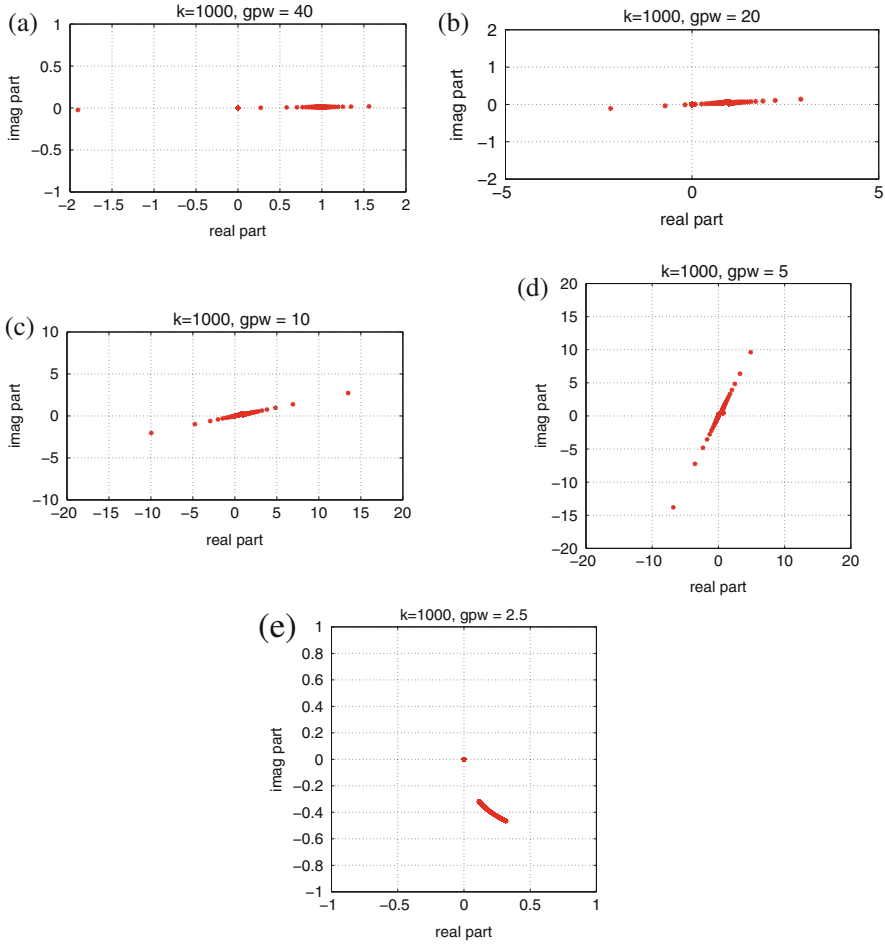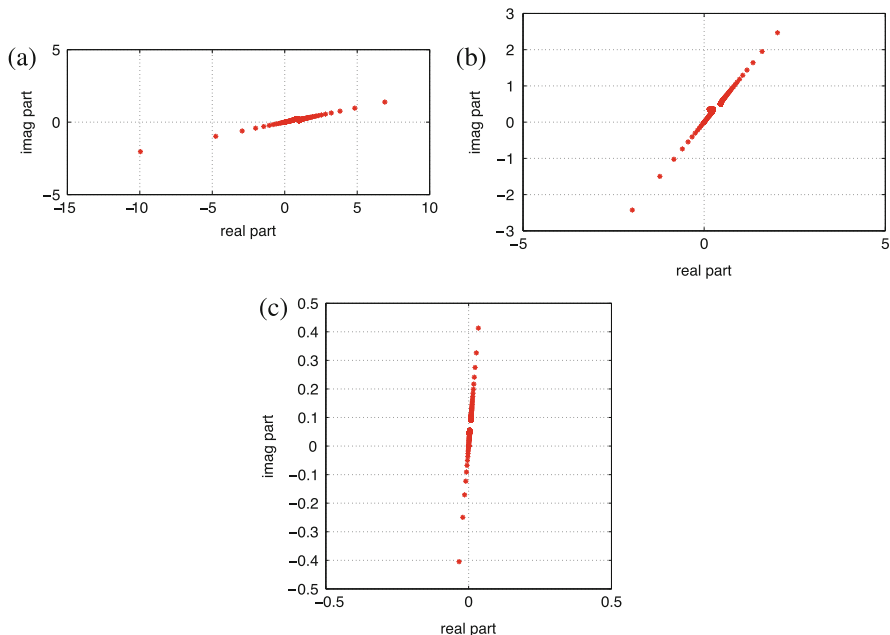
Fig. 5 Eigenvalues of the preconditioned deflated operator $M_{h,\beta_2}^{-1} P_{h,H} A_h$ in the complex plane for $k = 1000$ and $\beta_2 = 1$ on a multigrid hierarchy consisting of five levels. On the finest level 40 grid points per wave length are employed. On each coarser level the number of grid points per wave length is halved. (**a**) 40 gpw; (**b**) 20 gpw; (**c**) 10 gpw; (**d**) 5 gpw; (**e**) 2.5 gpw

the second and third variant we compare the multiplicative combination of A-DEF1 and CSLP with merely using CSLP as a preconditioner.

From Tables 2 and 3 we conclude that for the one and two-dimensional problem the combined use of A-DEF1 and CSLP

- results in a lower iteration count than either A-DEF1 or CSLP used separately. This reduction grows with the wave number;
- allows to use a large damping parameter β_2 without significantly increasing the iteration count;

Fig. 6   Eigenvalues of the preconditioned deflated operator $M_{h,\beta_2}^{-1} P_{h,H} A_h$ on a fixed mesh for various values of the damping parameter β_2 using 10 grid points per wavelength for $k = 1000$. (**a**) $\beta_2 = 1$; (**b**) $\beta_2 = 10$; (**c**) $\beta_2 = 100$

- allows to set $\beta_2 = 10$ and to approximate the CSLP preconditioner by its diagonal without significantly increasing the iteration count.

7   Conclusions

In this paper we considered a solution method for the Helmholtz equation that combines the complex shifted Laplace preconditioner with a deflation technique that employs multigrid vectors. We derived closed form expressions for the eigenvalues of the deflated preconditioned operator through a model problem analysis. From this analysis we conclude that a much larger damping parameter can be used without adversely affecting the convergence of outer Krylov acceleration. Further research is required to tune the algorithmic to large scale applications.

Table 2 Iteration count for various methods for the 1D problem without damping

1D without damping					
		$\beta_2 = 0.5$	$\beta_2 = 1$	$\beta_2 = 10$	$\beta_2 = 10$
k	A-DEF1	CSLP/CSLP+A-DEF1	CSLP/CSLP+A-DEF1	CSLP+A-DEF1	JACOBI+A-DEF1
10 gpw					
10	5	7/3	8/4	5	5
20	9	10/5	12/6	7	7
40	15	16/8	20/8	9	9
80	15	23/8	33/9	9	9
160	20	36/13	55/14	14	12
320	30	61/19	97/20	19	19
640	45	108/33	179/33	34	32
20 gpw					
10	9	7/3	8/4	5	6
20	13	10/4	12/4	5	6
40	14	15/5	19/5	6	6
80	15	22/6	33/6	6	7
160	19	37/8	56/8	8	8
320	18	59/9	95/9	9	9
640	28	104/14	174/14	14	15
1280	36	190/23	328/23	23	23

Table 3 Iteration count for various methods for the 2D problem without damping

2D without damping					
		$\beta_2 = 0.5$	$\beta_2 = 1$	$\beta_2 = 10$	$\beta_2 = 10$
k	A-DEF1	CSLP/CSLP+A-DEF1	CSLP/CSLP+A-DEF1	CSLP+A-DEF1	JACOBI+A-DEF1
10 gpw					
10	18	9/5	11/5	9	11
20	24	17/7	22/8	10	11
40	36	45/16	64/16	19	21
80	68	130/43	210/41	45	46
20 gpw					
5	18	5/3	6/3	5	9
10	17	9/3	11/3	3	5
15	21	12/4	16/4	5	9
20	24	16/6	22/6	7	10
30	20	29/5	40/5	5	12

References

1. Y. A. Erlangga, C. Vuik, and C. W. Oosterlee. On a class of preconditioners for solving the Helmholtz equation. *Appl. Numer. Math.*, 50(3–4):409–425, 2004.

2. Y. A. Erlangga, C. W. Oosterlee, and C. Vuik. A novel multigrid based preconditioner for heterogeneous Helmholt problems. *SIAM J. Sci. Comput*, 27:1471–1492, 2006.

3. C. I. Goldstein A. Bayliss and E. Turkel. An iterative method for the Helmholtz equation. *Journal of Computational Physics*, 49:443 – 457, 1983.

4. L. A. Laird and M. B. Giles. Preconditioned iterative solution of the 2D Helmholtz equation. Technical report, Comp. Lab. Oxford University UK, 2002. NA-02/12.

5. M. M. M. Made. Incomplete factorization-based preconditionings for solving the Helmholtz equation. *International Journal for Numerical Methods in Engineering*, 50:1077–1101, 2001.

6. Y. A. Erlangga and R. Nabben. On a multilevel Krylov method for the Helmholtz equation preconditioned by shifted Laplacian. *Electronic Transactions on Numerical Analysis*, 31:403–424, 2008.

7. B. Reps, W. Vanroose, and H. Bin Zubair. On the indefinite Helmholtz equation: Complex stretched absorbing boundary layers, iterative analysis, and preconditioning. *J. Comput. Phys.*, 229:8384–8405, November 2010.

8. M. Bollhöfer, M. J. Grote, and O. Schenk. Algebraic multilevel preconditioner for the Helmholtz equation in heterogeneous media. *SIAM Journal on Scientific Computing*, 31:3781–3805, 2009.

9. T. Airaksinen, E. Heikkola, A. Pennanen, and J. Toivanen. An algebraic multigrid based shifted-Laplacian preconditioner for the Helmholtz equation. *Journal of Computational Physics*, 226:1196 – 1210, 2007.

10. 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*, pages 1–48, 2015.

11. J. Zhu, X. W. Ping, R. S. Chen, Z. H. Fan, and D. Z. Ding. An incomplete factorization preconditioner based on shifted Laplace operators for FEM analysis of microwave structures. *Microwave and Optical Technology Letters*, 52:1036–1042, 2010.

12. T. Airaksinen, A. Pennanen, and J. Toivanen. A damping preconditioner for time-harmonic wave equations in fluid and elastic material. *J. Comput. Phys.*, 228:1466–1479, March 2009.

13. C. D. Riyanti, A. Kononov, Y. A. Erlangga, C. Vuik, C. Oosterlee, R.E. Plessix, and W.A. Mulder. A parallel multigrid-based preconditioner for the 3D heterogeneous high-frequency Helmholtz equation. *Journal of Computational Physics*, 224:431–448, 2007.

14. R. E. Plessix. A Helmholtz iterative solver for 3D seismic-imaging problems. *Geophysics*, 72:SM185–SM194, 2007.

15. R. E. Plessix. Three-dimensional frequency-domain full-waveform inversion with an iterative solver. *Geophysics*, 74(6):149–157, 2009.

16. N. Umetani, S. P. MacLachlan, and C. W. Oosterlee. A multigrid-based shifted Laplacian preconditioner for a fourth-order Helmholtz discretization. *Numerical Linear Algebra with Applications*, 16:603–626, 2009.

17. T. Airaksinen and S. Mönkölä. Comparison between the shifted-Laplacian preconditioning and the controllability methods for computational acoustics. *J. Comput. Appl. Math.*, 234:1796–1802, July 2010.

18. L. Zepeda-Núñez and L. Demanet. The method of polarized traces for the 2D Helmholtz equation. *Journal of Computational Physics*, 308:347–388, 2016.

19. D. Osei-Kuffuor and Y. Saad. Preconditioning Helmholtz linear systems. *Appl. Numer. Math.*, 60:420–431, April 2010.

20. H. Calandra, S. Gratton, R. Lago, X. Pinel, and X. Vasseur. Two-level preconditioned Krylov subspace methods for the solution of three-dimensional heterogeneous Helmholtz problems in seismics. *Numerical Analysis and Applications*, 5:175–181, 2012.

21. Y.A. Erlangga. Advances in iterative methods and preconditioners for the Helmholtz equation. *Archives of Computational Methods in Engineering*, 15:37–66, 2008.
22. Huangxin Chen, Haijun Wu, and Xuejun Xu. Multilevel preconditioner with stable coarse grid corrections for the helmholtz equation. *SIAM Journal on Scientific Computing*, 37(1):A221–A244, 2015.
23. L. Conen, V. Dolean, Rolf R. Krause, and F. Nataf. A coarse space for heterogeneous helmholtz problems based on the Dirichlet-to-Neumann operator. *Journal of Computational and Applied Mathematics*, 271:83–99, 2014.
24. M. Ganesh and C. Morgenstern. An efficient multigrid algorithm for heterogeneous acoustic media sign-indefinite high-order FEM models. *Numerical Linear Algebra with Applications*, 2016.
25. I. Livshits. Multiple galerkin adaptive algebraic multigrid algorithm for the helmholtz equations. *SIAM Journal on Scientific Computing*, 37(5):S195–S215, 2015.
26. L. N. Olson and J. B. Schroder. Smoothed aggregation for helmholtz problems. *Numerical Linear Algebra with Applications*, 17(2-3):361–386, 2010.
27. C. C. Stolk. A dispersion minimizing scheme for the 3-d Helmholtz equation based on ray theory. *Journal of Computational Physics*, 314:618–646, 2016.
28. J. Poulson, B. Engquist, S. Li, and L. Ying. A parallel sweeping preconditioner for heterogeneous 3d Helmholtz equations. *SIAM Journal on Scientific Computing*, 35(3):C194–C212, 2013.
29. P. Tsuji, B. Engquist, and L. Ying. A sweeping preconditioner for time-harmonic Maxwell equations with finite elements. *Journal of Computational Physics*, 231(9):3770–3783, 2012.
30. A. Vion and C. Geuzaine. Double sweep preconditioner for optimized schwarz methods applied to the helmholtz problem. *Journal of Computational Physics*, 266:171–190, 2014.
31. AH Sheikh, D Lahaye, L Garcia Ramos, R Nabben, and C Vuik. Accelerating the shifted laplace preconditioner for the helmholtz equation by multilevel deflation. *Journal of Computational Physics*, 322:473–490, 2016.
32. 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*, 29:1942–1958, 2007.
33. S. Cools and W. Vanroose. Local Fourier analysis of the complex shifted Laplacian preconditioner for helmholtz problems. *Numerical Linear Algebra with Applications*, 20(4):575–597, 2013.
34. Y. A. Erlangga and R. Nabben. Deflation and balancing preconditioners for Krylov subspace methods applied to nonsymmetric matrices. *SIAM J. Matrix Anal. Appl.*, 30(2):684–699, 2008.
35. 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(4):645–662, 2013.
36. S. Balay, J. Brown, K. Buschelman, V. Eijkhout, W. D. Gropp, D. Kaushik, M. G. Knepley, L. Curfman McInnes, B. F. Smith, and H. Zhang. PETSc users manual. Technical Report ANL-95/11 - Revision 3.4, Argonne National Laboratory, 2013.
37. A. H. Sheikh. *Development of the Helmholtz Solver Based On A Shifted Laplace Preconditioner and A Multilevel Deflation Technique*. PhD thesis, DIAM, TU Delft, 2014.
38. A. Bayliss, C.I. Goldstein, and E. Turkel. On accuracy conditions for the numerical computation of waves. *Journal of Computational Physics*, 59(3):396 – 404, 1985.
39. I. M. Babuska and S. A Sauter. Is the pollution effect of the FEM avoidable for the Helmholtz equation considering high wave numbers? *SIAM review*, 42(3):451–484, 2000.
40. 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 Berlin Heidelberg, 2012.

41. M. Gander. Fourier analysis of Helmholtz problems with Robin boundary conditions. Private Communications 2016.
42. I. Livshits. Use of shifted laplacian operators for solving indefinite helmholtz equations. *Numerical Mathematics: Theory, Methods and Applications*, 8(01):136–148, 2015.
43. J. M. Tang. *Two Level Preconditioned conjugate gradient methods with applications to bubbly flow problems.* PhD thesis, DIAM, TU Delft, 2008.
44. U. Trottenberg, C. W. Oosterlee, and A. Schüller. *Multigrid.* Academic Press, London, 2000.