

A fast method for the solution of the Helmholtz equation

Eldad Haber and Scott MacLachlan

September 15, 2010

Abstract

In this paper, we consider the numerical solution of the Helmholtz equation, arising from the study of the wave equation in the frequency domain. The approach proposed here differs from those recently considered in the literature, in that it is based on a decomposition that is exact when considered analytically, so the only degradation in computational performance is due to discretization and roundoff errors. In particular, we make use of a multiplicative decomposition of the solution of the Helmholtz equation into an analytical plane wave and a multiplier, which is the solution of a complex-valued advection-diffusion-reaction equation. The use of fast multigrid methods for the solution of this equation is investigated. Numerical results show that this is an efficient solution algorithm for a reasonable range of frequencies.

1 Introduction

In recent years, there has been a substantial interest in developing fast solvers for the Helmholtz equation [3, 4, 6–9, 18, 21]. This interest is motivated primarily by geophysical applications [13], although it also serves as an important model problem for many fields where fully indefinite partial differential equations arise. In this paper, we propose a new preconditioner for the Helmholtz equation.

The Helmholtz equation describes the propagation of a wave with frequency $\omega = 2\pi f$ within a non-homogeneous medium and can be written as

$$\begin{aligned} \Delta u + \kappa^2(\mathbf{x})\omega^2 u &= q & \mathbf{x} \in \Omega \\ \alpha \nabla u \cdot \mathbf{n} - Mu &= 0 & \mathbf{x} \in \partial\Omega, \end{aligned} \tag{1}$$

where u is the Fourier transform of the wavefield, $\kappa(\mathbf{x})$ is the “slowness” (or inverse of velocity) of the medium and can vary with position $\mathbf{x} = (x_1, x_2)^\top$, and q represents any sources. In this paper, we assume that $\Omega = [0, 1]^d$ where $d = 1, 2$, although the extension for 3D is straightforward. The equation is given with some consistent boundary conditions. Here, we consider both homogeneous Dirichlet and Sommerfeld boundary conditions,

$$\nabla u \cdot \mathbf{n} - i\omega\kappa u = 0,$$

that approximate boundary conditions at infinity. Other boundary conditions could be considered as well (see Section 2).

Upon discretization of the Helmholtz equation, a large, sparse, possibly complex-valued, and highly indefinite linear system is obtained. Here, we consider only the standard second-order finite-difference discretization, although other discretizations can be used as well. Typically, for boundary conditions that imitate appropriate behaviour at infinity, the system and solution are complex-valued. The solution of the Helmholtz equation is highly oscillatory and nonlocal, especially for problems with high frequencies. For these large frequencies, the wave length $\lambda = \frac{2\pi}{\omega\kappa} = \frac{1}{\kappa f}$ can be very small. This generates two related difficulties. First, a sufficiently fine grid has to be used in order to resolve all spatial frequencies. For the second-order finite-difference discretization considered here, the rule of thumb is that at least 10 points per wavelength are needed to accurately resolve the solution. For example, assuming $\kappa = 1$ and a frequency of $\omega = 32\pi$, the maximum mesh size is $1/160$. The problem is even worse if dispersion is considered and needs to be controlled. In this case, it is possible to show that the number of grid points per wavelength should increase when the frequency increases [2]. Such mesh sizes generate large linear systems even in 2D and especially in 3D.

A second, related problem is that the resulting linear systems are very difficult to solve. The performance of standard multigrid methods is hindered by the fact that standard smoothers, such as weighted-Jacobi or Gauss-Seidel, do not necessarily resolve high frequencies and may even be divergent. Thus, complementing these approaches with an appropriate coarse-grid correction process is very difficult. Due to the oscillatory and indefinite character of the equation, the solution of the Helmholtz equation is a challenging task, even though it is linear. Efficient solvers for this equation which yield mesh and frequency independence for the variable coefficient case are not well developed. Recent work includes the combination of multigrid with Krylov space relaxation [6], multigrid applied to the least-squares formulation [9], and the wave-ray multigrid approach [3,4]. Another approach that has attracted significant interest both recently and in the past is the family of “shifted-Laplacian” preconditioners (see, for example, [7,12,18,21] and the references cited therein). Despite many years of research, none of the above approaches is a “silver bullet” for this very difficult problem.

The goal of this work is to suggest a new family of preconditioners for the Helmholtz equation that enable quick convergence in linear time, largely independent of the frequency. Our approach is different from previous approaches, which are based on some variation of multigrid applied directly to a Helmholtz-type equation. Rather than working directly with the discretized Helmholtz equation, we reformulate the problem in continuous setting to obtain an equivalent system of equations made of a nonlinear part (an Eikonal equation) and a linear part (an advection-diffusion-reaction equation). We show that we are able to choose analytic solutions to the Eikonal equation that yield a linear system that is amenable to solution using multigrid principles. We then discretize and solve this linear system and use it as a preconditioner.

The rest of this paper is organized as follows. In Section 2, we present the basic idea for reformulation. In Section 3, we discuss the solution of the reformulated problem. Section 4

presents a multigrid method for the resulting complex-valued advection-diffusion-reaction equation. In Section 5, we suggest a new preconditioner for the Helmholtz equation based on our approximate solver, which is examined using Fourier analysis in Section 6. Finally, in Section 7, we conduct numerical experiments.

2 Reformulation

In this section, we present a reformulation of the Helmholtz equation that is more amenable to numerical solution. The basic idea is to decompose the solution using the so-called Rytov decomposition of the form

$$u(x) = a(\mathbf{x})e^{i\omega T(\mathbf{x})}. \quad (2)$$

This decomposition is a classical textbook approach (see [1,11]) and is typically used in order to *approximate* the *homogeneous* wave or Helmholtz equation. Here, in contrast, we use this decomposition to obtain an *exact* equivalent system for the *non-homogeneous* Helmholtz equation. Note that there is always such a decomposition and that the decomposition is not unique; for example, taking $T = 0$, we obtain the original Helmholtz system. The idea, then, is to choose an appropriate decomposition such that the resulting system is easier to solve.

Differentiating (2), we obtain

$$\begin{aligned} \nabla u &= e^{i\omega T} (\nabla a + i\omega a \nabla T), \\ \text{and } \Delta u &= e^{i\omega T} (\Delta a + 2i\omega \nabla a \cdot \nabla T + i\omega a \Delta T - \omega^2 a |\nabla T|^2). \end{aligned}$$

Substituting these in (1) and multiplying by $\exp(-i\omega T(x))$, we have

$$\Delta a + 2i\omega \nabla a \cdot \nabla T + i\omega a \Delta T - \omega^2 (|\nabla T|^2 - \kappa^2) a = q e^{-i\omega T(x)}. \quad (3)$$

Equation (3) suggests that the problem can be decomposed in the following way:

- Choose $T(\mathbf{x})$
- Set $\hat{q} = q \exp(-i\omega T(\mathbf{x}))$ and solve the resulting complex-valued advection-diffusion-reaction (ADR) equation for a :

$$\Delta a + 2i\omega \nabla T \cdot \nabla a + i\omega (\Delta T) a - \omega^2 (|\nabla T|^2 - \kappa^2) a = \hat{q}, \quad x \in \Omega. \quad (4)$$

The first question to be addressed is how should we choose T to obtain a better system for a ? One obvious choice is to choose T to satisfy the Eikonal Equation, $|\nabla T|^2 = \kappa^2$; however, this approach is difficult to realize when κ is highly variable, particularly given the need to compute ΔT in Equation (4). Instead, we consider using simple plane waves,

$$T = c\boldsymbol{\alpha} \cdot \mathbf{x}, \quad (5)$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)^\top$ with $\|\boldsymbol{\alpha}\| = 1$ and c is a constant scaling. Clearly, this T satisfies a simple Eikonal equation,

$$|\nabla T|^2 = c^2,$$

and the resulting ADR equation from (4) can be rewritten as

$$\Delta a + 2i\omega c \boldsymbol{\alpha} \cdot \nabla a - \omega^2(c^2 - \kappa^2)a = \hat{q} \quad x \in \Omega. \quad (6)$$

Now, if $c \gg \kappa$ then the system (6) has a strong negative-definite component. This may seem appealing at first; however, two difficulties arise when doing so. First, note that the right-hand side contains the factor $\exp(-i\omega c \boldsymbol{\alpha} \cdot \mathbf{x})$. This makes the right-hand side highly oscillatory and, thus, a very fine grid will be needed as c increases. Second, for large c , one obtains a boundary layer that requires special care for accurate solution of the problem. The goal, therefore, is to bias the eigenvalues of equation (6) by choosing the smallest c that keeps the value of $c^2 - \kappa^2$ non-negative. Thus, we propose to choose

$$c \geq \max_{\mathbf{x}}(\kappa(\mathbf{x})).$$

We further discuss the value of c and its effect on the solution later. For the moment, note that we obtain an ADR equation with a non-positive reaction term and this helps in biasing the eigenvalues towards the negative-definite side.

The ADR equation is, of course, not closed without appropriate boundary conditions. The boundary conditions for a are inherited from the boundary conditions for u . Homogeneous Dirichlet boundary conditions on u yield identical boundary conditions on a . More general Robin boundary conditions on u , such as Sommerfeld boundary conditions, yield

$$\mathbf{n} \cdot \nabla a + i\omega(c\mathbf{n} \cdot \boldsymbol{\alpha} - \kappa)a = 0 \quad x \in \partial\Omega.$$

Similarly, other more-involved boundary conditions for u produce corresponding boundary conditions for a .

As we show next, solving the Eikonal and the ADR equations as a replacement to the Helmholtz equation is appealing. This is because fast and efficient solvers enable an $\mathcal{O}(n)$ solution to the ADR equation [20, 22].

3 The advection-diffusion-reaction equation

We first consider the solution of the advection-diffusion-reaction equation (6), using boundary conditions that are inherited from the Helmholtz equation. We consider two different discretizations for ∇a , the second-order centered difference

$$a_{x_1} \approx \frac{1}{2h}(a_{i+1,j} - a_{i-1,j}), \quad a_{x_2} \approx \frac{1}{2h}(a_{i,j+1} - a_{i,j-1}),$$

and, to have a stable discretization for ∇a , the second-order upwind discretization; assuming $\alpha_1, \alpha_2 > 0$, we compute the derivatives as

$$a_{x_1} \approx \frac{1}{2h}(3a_{ij} - 4a_{i-1,j} + a_{i-2,j}), \quad a_{x_2} \approx \frac{1}{2h}(3a_{ij} - 4a_{i,j-1} + a_{i,j-2}).$$

Boundary conditions are implemented by using ghost points as is typically done in finite-difference formulations [20]. We then use a standard five-point stencil to discretize the Laplacian, obtaining a system of the form

$$Ma = (A_1 + 2ic\omega A_2 - \omega^2 A_3)a = \widehat{b}, \quad (7)$$

where A_1 is a discretization of the Laplacian and A_2 is the discretization of $\boldsymbol{\alpha} \cdot \nabla$ and A_3 is a diagonal matrix which contains the discretization of $0 \leq \beta = c^2 - \kappa^2$ on its diagonal.

To better understand the discretization and its properties, we use local Fourier analysis (LFA). We assume periodic boundary conditions and constant coefficients. We then have that A_1 , A_2 , and A_3 are represented by the stencils

$$A_1 = \frac{1}{h^2} \begin{pmatrix} & 1 & \\ 1 & -4 & 1 \\ & & 1 \end{pmatrix} \quad (8)$$

$$A_2 = \frac{1}{2h} \begin{pmatrix} & \alpha_2 & \\ -\alpha_1 & & \alpha_1 \\ & \alpha_2 & \end{pmatrix} \quad \text{for central differences} \quad (9)$$

$$A_2 = \frac{1}{2h} \begin{pmatrix} & & 0 & \\ \alpha_1 & -4\alpha_1 & 3(\alpha_1 + \alpha_2) & 0 \\ & & -4\alpha_2 & \\ & & & \alpha_2 \end{pmatrix} \quad \text{for upwind} \quad (10)$$

$$A_3 = \begin{pmatrix} \beta \\ \end{pmatrix} \quad (11)$$

To see that this equation can be efficiently solved using multigrid methods, we first consider LFA to examine the h -ellipticity of the system (see, for example, [20] for a detailed discussion of this analysis). Consider the usual Fourier component, $\exp(i\frac{\boldsymbol{\theta}^\top \mathbf{x}}{h})$, with $\boldsymbol{\theta} = (\theta_1, \theta_2)^\top$. The symbols of A_1 , A_2 and A_3 are

$$\begin{aligned} \widehat{A}_1 &= \frac{1}{h^2}(2 \cos(\theta_1) + 2 \cos(\theta_2) - 4) \\ \widehat{A}_2 &= \frac{i}{h}(\alpha_1 \sin(\theta_1) + \alpha_2 \sin(\theta_2)) \quad \text{for central differences} \\ \widehat{A}_2 &= \frac{1}{2h}(\alpha_1(3 - 4e^{-i\theta_1} + e^{-2i\theta_1}) + \alpha_2(3 - 4e^{-i\theta_2} + e^{-2i\theta_2})) \quad \text{for upwind} \\ \widehat{A}_3 &= \beta \end{aligned}$$

and, thus, the symbol of the system is

$$\widehat{B}_h = \widehat{A}_1 + 2i\omega c \widehat{A}_2 - \omega^2 \beta. \quad (12)$$

In order to have an amenable discretization for multigrid methods, one requires that the discretization is h -elliptic. Recall that the h -ellipticity measure is defined as

$$E_h = \frac{\min\{|\widehat{B}_h(\boldsymbol{\theta})|; \boldsymbol{\theta} \in T^{\text{high}}\}}{\max\{|\widehat{B}_h(\boldsymbol{\theta})|; \boldsymbol{\theta} \in [-\pi, \pi)^2\}},$$

where $T^{\text{high}} = [-\pi, \pi)^2 \setminus [-\frac{1}{2}\pi, \frac{1}{2}\pi)^2$ and that the discrete system given by B_h is said to be h -elliptic if E_h is bounded uniformly away from 0, independently of h .

Consider first the case where central differences are used for the advection term, and assume for simplicity that $\kappa = 1$ and $c = \kappa$, which implies that $\beta = 0$. Note that for $\beta > 0$ the system has better properties (as it is more negative definite) and that we can always choose c such that $\beta \geq 0$. Thus, the case where $\beta = 0$ is the worst-case scenario. The symbol of the operator is, then,

$$\widehat{B}_h = \frac{1}{h^2} ((2 \cos(\theta_1) + 2 \cos(\theta_2) - 4) - 2\omega h(\alpha_1 \sin(\theta_1) + \alpha_2 \sin(\theta_2))).$$

To further simplify the discussion, we take $\alpha_1 = 1$ and $\alpha_2 = 0$. The minimum of the numerator of the h -ellipticity measure clearly occurs when $\theta_2 = \pi/2$, yielding

$$\widehat{B}_h(\theta_1, \theta_2 = \pi/2) = \frac{1}{h^2} (2 \cos(\theta_1) - 4 - 2\omega h \sin(\theta_1)).$$

Clearly, if the product of $\xi = \omega h$ is not bounded, then the advection term will dominate and, since $\sin(\theta_1)$ can be both positive and negative, the minimum of the function over the high frequencies will go to zero and h -ellipticity measure will vanish. On the other hand, if we make sure that ξ is bounded from above, then the symbol will stay h -elliptic. For the particular case we analyze here, it can be observed that as long as $\xi \approx 1.12$ or smaller, the h -ellipticity of the system is still larger than 0.25. This observation implies that, for this discretization, standard multigrid should work as long as the coarse grid is not too coarse and one should bound h from above. We thus conclude that the central difference approximation to the advection term is not optimal, but it can still yield a convergent multigrid scheme if done carefully.

Next, we consider second-order upwinding for the advection term, again taking $c = \kappa = 1$, $\alpha_1 = 1$, and $\alpha_2 = 0$. Using a similar analysis to the central difference case, we obtain that

$$\widehat{B}_h(\theta_1, \theta_2 = \pi/2) = \frac{1}{h^2} (2 \cos(\theta_1) - 4 + i\omega h(4e^{i\theta_1} - 3 - e^{2i\theta_1})).$$

For this discretization, it is more difficult to analyze the h -ellipticity measure and we turn to numerical evaluation. A plot of the h -ellipticity measure, E_h , as a function of ξ is presented in Figure 1. As can be observed, this discretization maintains its h -ellipticity for all values of ξ . We have performed similar numerical studies for other values of α and c and over both values of θ . The results are similar to the simple case presented above. This is not surprising, as it is well known that upwind schemes are h -elliptic [20].

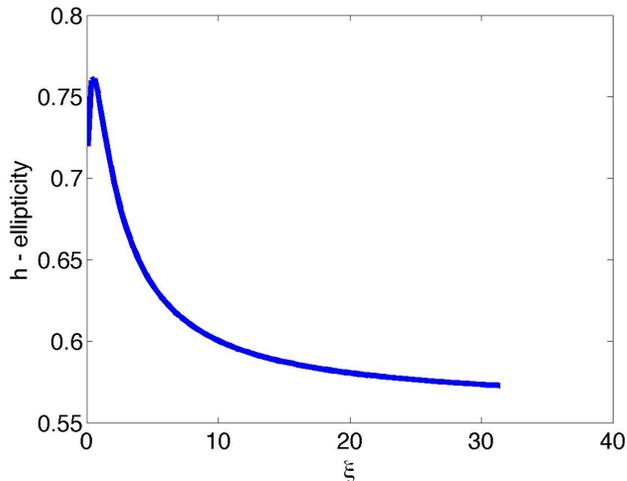


Figure 1: h -ellipticity as a function of ξ for the upwind discretization

The boundedness of the h -ellipticity measure away from 0 for these discretizations implies that the eigenvalues that correspond to high frequencies are bounded away from 0, and that it is possible to generate fast (multigrid) algorithms for the solution of the problem (see [20] and the following section for details). Realizing these algorithms, however, still requires care. Next, we consider smoothing and two-level local Fourier analysis of a simple multigrid algorithm for the solution of these equations.

4 A multigrid method for complex-valued advection-diffusion-reaction equations

The h -ellipticity bounds above, particularly for the upwind discretization, imply the existence of an optimally scaling multigrid solution algorithm for the ADR problem. Indeed, significant literature exists on efficient solution algorithms for ADR equations with constant real-valued coefficients [5, 14–17, 20, 22]. To our knowledge, however, none of the existing literature deals with the case of complex-valued coefficients. Here, we consider the most straight-forward approach, directly applying a standard algorithm from the real-valued coefficient case to the complex-valued coefficient problem of interest here. We stress, however, that this is the most naïve approach, which is considered only as a proof of concept. Significant effort is needed to develop optimal and robust solution strategies for the full range of parameters to be considered; this will be the focus of future research.

For real-valued ADR equations, the primary concern in discretization is the stability of the resulting discrete problem, which depends on the size of the mesh-Péclet (or grid-Reynolds) number, Pe , the ratio of the diffusive and advective coefficients weighted by the grid size [20, §7.1.2]. When the mesh-Péclet number is large ($Pe > 2$), upwind discretizations are stable

and, thus, preferable, while upwind or central difference discretizations are acceptable for small mesh-Péclet numbers. For Equation (6), we easily get the bound $Pe \leq 2c(\omega h)$, and see that the stability depends both on the choice of c and the value of $\xi = \omega h$. However, there are two important considerations for the setting considered here. First, it is not clear whether the stability concerns that apply for the case of a real-valued coefficient are even relevant in the complex-valued case considered here. Secondly, it is not clear which direction is “upwind” in the setting of a complex-valued coefficient of the advection term.

Here, we ignore these issues of stability, and examine the multigrid solution of the complex-valued ADR equations with both the upwind and central-difference discretizations of the advection terms. We choose the simplest possible definition of “upwind” in this case, ignoring the factor of i and using the signs of the entries in $c\boldsymbol{\alpha}$. This approach is primarily justified by the fact that we will consider the solution of the ADR equation only as a preconditioner for the direct discretization of the Helmholtz equation, as discussed in Section 5. The strong dependency of the h -ellipticity bound on ωh in the central-difference case suggests that we should expect difficulty in defining an optimal multigrid approach for this discretization. In contrast, the strong lower bound on E_h for the upwind discretization, as seen in Figure 1, suggests that we should be more successful in this setting.

To examine the potential performance of multigrid methods for Equation (6), we first consider local Fourier analysis [20]. Here, we restrict attention to the case where $\kappa = 1$, $\boldsymbol{\alpha} = \frac{1}{\sqrt{2}}(1, 1)^\top$ and $c > 0$; other cases can be treated similarly so long as attention is paid to the identification of an “upwind” direction. From the standard case, along with the h -ellipticity analysis above, we expect that “downwind”-ordered Gauss-Seidel relaxation (which, in this case, is the same as lexicographical Gauss-Seidel) should be a good relaxation scheme for these matrices when $c\omega h$ is small.

Under the assumption of a uniform infinite grid, the usual Fourier functions, $\exp\left(i\frac{\boldsymbol{\theta}^\top \mathbf{x}}{h}\right)$, with $\boldsymbol{\theta} = (\theta_1, \theta_2)^\top$, $-\frac{\pi}{2} \leq \theta_1, \theta_2 < \frac{3\pi}{2}$, are the eigenvectors of the lexicographical Gauss-Seidel sweep [20]. To find the corresponding eigenvalues, we split the stencils in Equations (8)-(11) into their diagonal, lower-triangular, and upper-triangular parts, and consider a typical eigenvector equation for the relaxation sweep. For the upwind discretization, we have

$$\begin{aligned} \lambda \left[\frac{1}{h^2} (v_{i-1,j} + v_{i,j-1}) - \left(\frac{4}{h^2} + \omega^2(c^2 - \kappa^2) \right) v_{i,j} \right. \\ \left. + \frac{i\omega c}{h} (\alpha_1(3v_{i,j} - 4v_{i-1,j} + v_{i-2,j}) + \alpha_2(3v_{i,j} - 4v_{i,j-1} + v_{i,j-2})) \right] \\ = -\frac{1}{h^2} (v_{i+1,j} + v_{i,j+1}). \end{aligned}$$

Now, making the Fourier ansatz, we can find the value of λ , the amplification factor of the relaxation for the Fourier mode of frequency $\boldsymbol{\theta}$, as

$$\lambda = \frac{-e^{i\theta_1} - e^{i\theta_2}}{e^{-i\theta_1} + e^{-i\theta_2} - (4 + \omega^2 h^2 (c^2 - \kappa^2)) + i\omega h c (\alpha_1(3 - 4e^{-i\theta_1} + e^{-2i\theta_1}) + \alpha_2(3 - 4e^{-i\theta_2} + e^{-2i\theta_2}))}.$$

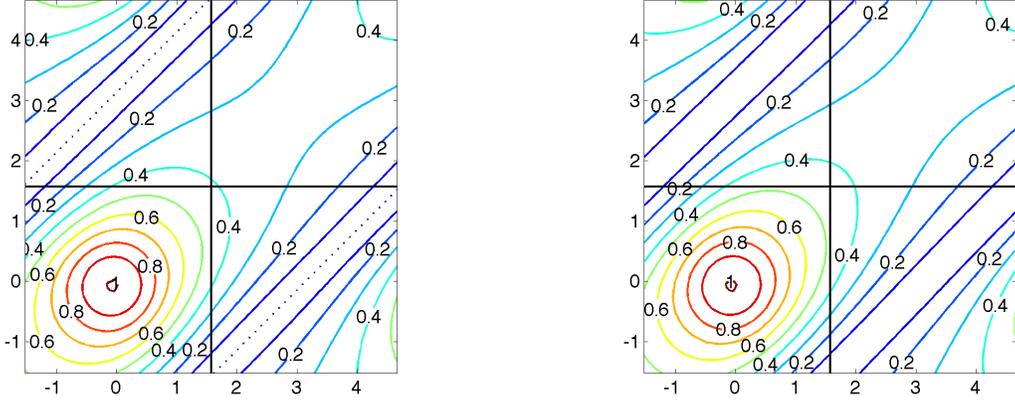


Figure 2: The LFA amplification factors for “downwind” (lexicographic) Gauss-Seidel relaxation applied to the discretization of Equation (6) when $c = 1$ and $\omega h = \frac{\pi}{32}$, as a function of the Fourier mode, $\boldsymbol{\theta}$. The upwind discretization is shown at left, while the central-difference discretization is shown at right.

Similarly, for the central-difference discretization, we have

$$\lambda = \frac{-e^{i\theta_1} - e^{i\theta_2} - i\omega h c (\alpha_1 e^{i\theta_1} + \alpha_2 e^{i\theta_2})}{e^{-i\theta_1} + e^{-i\theta_2} - (4 + \omega^2 h^2 (c^2 - \kappa^2)) - i\omega h c (\alpha_1 e^{-i\theta_1} + \alpha_2 e^{-i\theta_2})}.$$

Figure 2 shows contour plots of these error-amplification factors predicted by the LFA for downwind Gauss-Seidel when $c = 1$ and $\omega h = \frac{\pi}{32}$ for both the upwind and central-difference discretizations. Note that in both pictures, there is a small range of eigenvalues around $\boldsymbol{\theta} = \mathbf{0}$ where the amplification factor for relaxation is greater than one. The LFA smoothing factor, measured as the largest magnitude eigenvalue over frequencies $\boldsymbol{\theta} \in [-\frac{\pi}{2}, \frac{3\pi}{2})^2 \setminus [-\frac{\pi}{2}, \frac{\pi}{2})^2$, is 0.506 for the central-difference discretization and 0.526 for the upwind discretization.

This smoothing analysis can be combined with analysis of the coarse-grid correction stage by considering the coupling of Fourier modes in the coarse-grid correction process. We consider only standard geometric coarsening by a factor of two in each direction, with constant-coefficient interpolation and restriction operators. Thus, the coarse-grid correction phase couples groups of four Fourier modes (the so-called harmonics); given a low-frequency mode, $\boldsymbol{\theta} = (\theta_1, \theta_2)^\top \in [-\frac{\pi}{2}, \frac{\pi}{2})^2$, the high-frequency modes with frequencies $(\theta_1 + \pi, \theta_2)^\top$, $(\theta_1, \theta_2 + \pi)^\top$, and $(\theta_1 + \pi, \theta_2 + \pi)^\top$ are coupled in the coarse-grid correction process. Accounting for this coupling, we can compute the LFA eigenvalues of the two-grid cycle combining the downwind Gauss-Seidel relaxation with a standard Galerkin coarse-grid correction operator, using bilinear interpolation and full-weighting restriction. Figure 3 shows these eigenvalues for a cycle using one pre-relaxation and one post-relaxation step. Note that for these values of c and ωh , the diverging modes of the relaxation sweep are effectively attenuated by the coarse-grid correction process. The largest-magnitude LFA eigenvalues are 0.177 for the upwind discretization and 0.156 for the central-difference discretization.

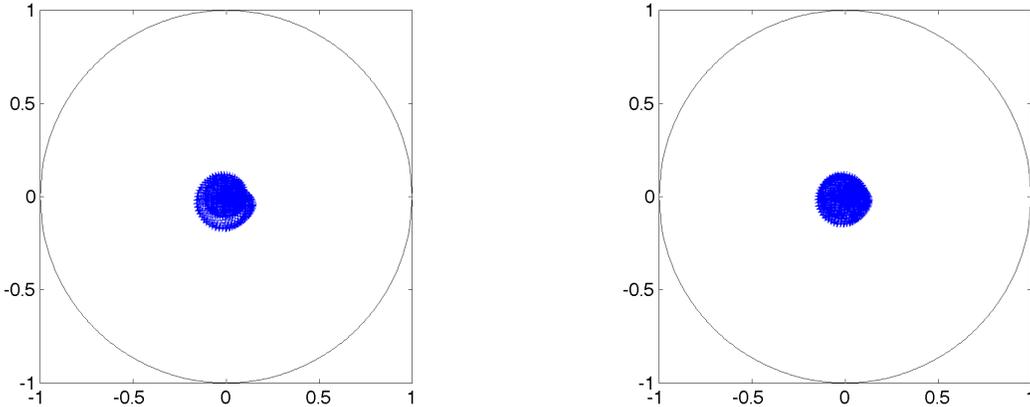


Figure 3: The LFA spectrum of the two-grid error-propagation operator using downwind Gauss-Seidel relaxation with bilinear interpolation and full-weighting restriction for $c = 1$ and $\omega h = \frac{\pi}{32}$. The upwind discretization is shown at left, while the central-difference discretization is shown at right.

For larger values of ξ , the divergence of Gauss-Seidel relaxation becomes problematic. Figure 4 shows the LFA amplification factor for downstream Gauss-Seidel relaxation when $c = 1$ and $\omega h = \frac{\pi}{4}$. Note that, for both discretizations, there is now a substantial region where the amplification factor is greater than 1. Furthermore, while the amplification factors look qualitatively similar (and have similar LFA smoothing factors of 0.765 and 0.620 for central differences and upwind, respectively), the quantitative behaviour of the LFA two-grid convergence factors for these problems is quite different, with a predicted factor of 1.56 for the upwind discretization, but a predicted factor of 39.29 for the central-difference discretization. This underscores the added stability of the upwind discretization. Looking more closely at the LFA spectra, shown in Figure 5, we see that the eigenvalues for the upwind discretization remain bounded away from 1 for the multigrid error-propagation operator, corresponding to a preconditioned system for which the eigenvalues are bounded away from zero. In contrast, not only are the eigenvalues outside the unit circle for the central-difference discretization more spread out, but they also cluster near 1; this suggests that a preconditioned Krylov iteration for this problem is likely to have convergence issues.

In the above calculations, we have considered only the case where $c = \kappa = 1$. From the point-of-view of the mesh-Péclet number, choosing as small a value of c as possible seems reasonable. However, the reaction term in (6) also plays an important role in the stability and h -ellipticity of the resulting matrices. In particular, for either discretization, for large-enough c , the problem becomes negative-definite. For $\omega h = \frac{\pi}{4}$, taking $c = 2.5$ but keeping $\kappa = 1$ gives the LFA amplification factors and eigenvalues shown in Figure 6. Note, in particular, that for this value of c , the downwind Gauss-Seidel relaxation is convergent for both of these problems, but the LFA spectrum of the two-grid operator is more spread out. This

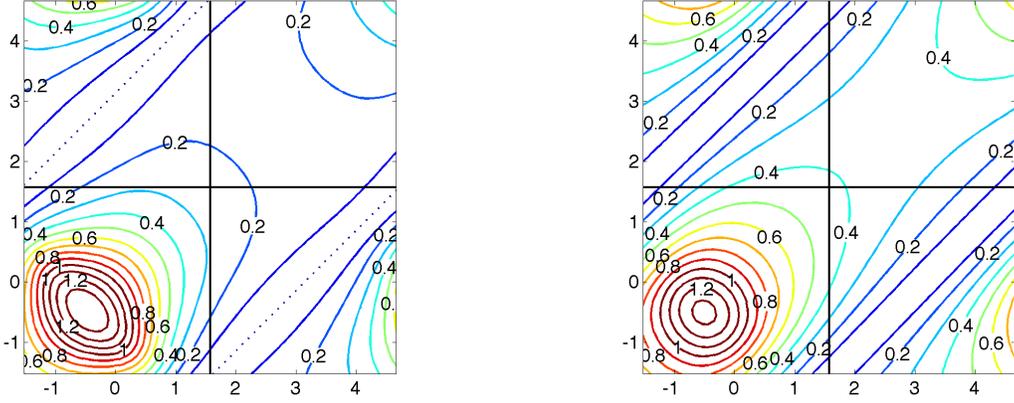


Figure 4: LFA amplification factors for “downwind” (lexicographic) Gauss-Seidel relaxation applied to the discretization of Equation (6) when $c = 1$ and $\omega h = \frac{\pi}{4}$, as a function of the Fourier mode, θ . The upwind discretization is shown at left, while the central-difference discretization is shown at right.

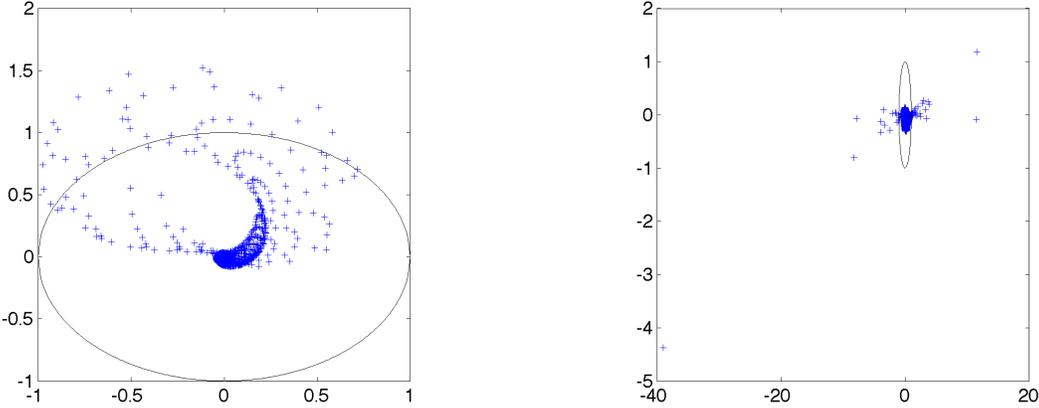


Figure 5: The LFA spectrum of the two-grid error-propagation operator using downwind Gauss-Seidel relaxation with bilinear interpolation and full-weighting restriction for $c = 1$ and $\omega h = \frac{\pi}{4}$. The upwind discretization is shown at left, while the central-difference discretization is shown at right. Note that the axis limits here are different from one-another and from those shown in other figures; the unit circle is given as a reference.

	$c = 1$ $\omega h = \frac{\pi}{32}$	$c = 1$ $\omega h = \frac{\pi}{4}$	$c = 2.5$ $\omega h = \frac{\pi}{4}$
central differences	0.105	28.72	0.588
upwind	0.125	1.354	0.446

Table 1: Two-level computed asymptotic convergence factors, averaged over 100 iterations, for multigrid V(1,1) cycles using downwind (lexicographical) Gauss-Seidel relaxation and Galerkin coarsening with $h = \frac{1}{32}$.

	$c = 1$ $\omega h = \frac{\pi}{32}$	$c = 1$ $\omega h = \frac{\pi}{4}$	$c = 2.5$ $\omega h = \frac{\pi}{4}$
central differences	0.296	4.98×10^3	0.626
upwind	0.312	1.89×10^4	0.670

Table 2: Multigrid computed asymptotic convergence factors, averaged over 100 iterations, for V(1,1) cycles using downwind (lexicographical) Gauss-Seidel relaxation and Galerkin coarsening with $h = \frac{1}{32}$. Due to strong divergence, results for $c = 1$ and $\omega h = \frac{\pi}{4}$ are averaged over only 50 iterations.

is because the slowest-to-converge modes of relaxation are no longer those associated with the smoothest Fourier modes but are biased towards $\boldsymbol{\theta} = \left(-\frac{\pi}{4}, -\frac{\pi}{4}\right)^\top$. The LFA smoothing factors are 0.750 and 0.366 for the central-difference and upwind discretizations, respectively, while the LFA two-grid convergence factors are 0.642 for central differences and 0.779 for upwind.

Numerical experiments confirm the expectations given by the Fourier analysis, that the multigrid V-cycle converges quickly for small $\xi = \omega h$, or for larger ξ if c is appropriately large. When ξ is large, but $c = \kappa = 1$, we see divergence. Table 1 shows the two-level convergence factors, averaged over 100 iterations for matrices with $h = \frac{1}{32}$, given a zero right-hand side and a random initial guess. Note that these are relatively consistent with the predictions made by LFA. Multigrid V-cycle results (coarsening down to a 3×3 matrix) are given in Table 2. While there is some degradation in the multigrid convergence factors, the same qualitative behaviour is seen.

As discussed above, the two-grid LFA spectra in Figure 5 suggest that, at least for the upwind discretization, the multigrid method may make an effective preconditioner for a Krylov method. In Table 3, we present iteration counts for multigrid-preconditioned GMRES to reduce the norm of the residual by a relative factor of 10^{-7} when $c = 1$, for various values of ωh . We note that the standalone multigrid method is divergent for this range of parameters, but that it can be effective when used as a preconditioner. For small values of ωh , the performance is similar for both discretizations, but for larger values of ωh , the performance for the central-difference discretization degrades quickly with increased ωh . Also, note that these results are computed with right-hand sides given by applying the system matrix to a

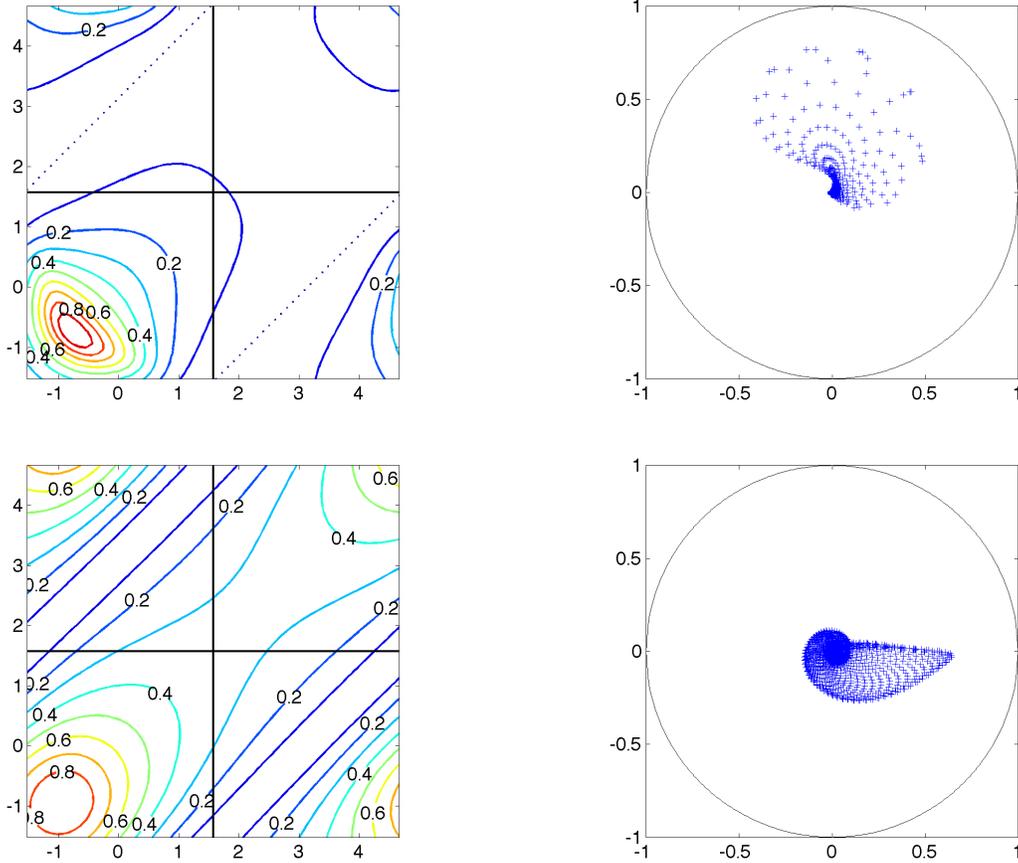


Figure 6: At left, the LFA amplification factor for “downwind” (lexicographic) Gauss-Seidel relaxation applied to the discretization of Equation (6) when $c = 2.5$ and $\omega h = \frac{\pi}{4}$, as a function of the Fourier mode, θ . At right, the LFA spectrum of the two-grid error-propagation operator using this smoother with bilinear interpolation and full-weighting restriction. The top row of figures corresponds to the upwind discretization of the advection term, while the bottom row corresponds to the central-difference discretization.

		$\omega h = \frac{\pi}{16}$	$\omega h = \frac{\pi}{12}$	$\omega h = \frac{\pi}{8}$	$\omega h = \frac{\pi}{6}$	$\omega h = \frac{\pi}{5}$	$\omega h = \frac{\pi}{4}$
upwind differences	$h = \frac{1}{32}$	8	14	24	27	28	30
	$h = \frac{1}{64}$	18	34	51	58	64	78
	$h = \frac{1}{96}$	30	47	75	93	110*	> 200
central differences	$h = \frac{1}{32}$	7	11	20	33	43	62
	$h = \frac{1}{64}$	17	28	58	98	121	186
	$h = \frac{1}{96}$	30	48	97	167	> 200	> 200

Table 3: Number of iterations of multigrid-preconditioned GMRES needed to achieve a residual reduction of 10^{-7} for $c = 1$ with various values of ωh . Results marked by a * denote a lack of convergence in GMRES, due to a detected rank deficiency.

	$h = \frac{1}{32}$	$h = \frac{1}{64}$	$h = \frac{1}{128}$	$h = \frac{1}{256}$	$h = \frac{1}{512}$	$h = \frac{1}{1024}$
upwind differences	30	51	46	31	23	18
central differences	62	58	46	31	23	18

Table 4: Number of iterations of multigrid-preconditioned GMRES needed to achieve a residual reduction of 10^{-7} for $c = 1$ with various values of h , with fixed $\omega = 8\pi$.

randomly generated solution vector; as such, the iteration counts are somewhat subject to the random-number generation. However, our tests show only slight variation (by an iteration or two) over different realizations. For the upwind discretization, these results are qualitatively similar to those seen for the shifted-Laplace preconditioners considered in [7, 18, 21], where results degrade as ωh increases and, for fixed ωh , scale like $1/h$. However, the h -ellipticity bounds discussed above suggest that better performance is possible, with iteration counts that are independent of h for fixed ωh , and only slight dependence on ωh , at least in the upwind-discretization case.

In contrast, if we fix ω and vary h , the results for both discretizations show performance similar to what is expected for standard multigrid applied to an elliptic problem. In Table 4, we show iteration counts for $\omega = 8\pi$ as the grid is refined. Note, in particular, that for the coarsest meshes considered, these results reflect those shown in Table 3; however, for the finest meshes, the performance is nearly identical between the two discretizations, and is clearly bounded (and improving) as $h \rightarrow 0$.

5 Preconditioning

The Rytov decomposition could be used directly to obtain approximate solutions for the continuous Helmholtz equation; however, it also provides a unique approach for preconditioning the standard, second-order finite-difference discretization of the Helmholtz equation, as discussed here.

Assume that we have the discrete Helmholtz equation,

$$Au = b, \tag{13}$$

where A is a discretization of the Helmholtz operator with the appropriate boundary conditions and b is the discretization of the right-hand side. Assume also that we chose a solution to the Eikonal equation of the form $T = c \boldsymbol{\alpha} \cdot \mathbf{x}$. Finally, let

$$Ba = (A_1 + 2ic\omega A_2 - \omega^2 A_3)a = M^{-1}b \tag{14}$$

be a discretization of the ADR equation discussed above, where the diagonal matrix M is defined by

$$M = \text{diag}(e^{i\omega c \boldsymbol{\alpha} \cdot \mathbf{x}}),$$

where \mathbf{x}_j is the location of the node corresponding to the j^{th} degree of freedom, $u_j \approx u(\mathbf{x}_j)$.

The decomposition $u = ae^{i\omega c \boldsymbol{\alpha} \cdot \mathbf{x}}$ implies that

$$u = Ma; \tag{15}$$

substituting (15) into the discrete Helmholtz equation (13) and left-multiplying with M^{-1} gives

$$M^{-1}AMa = M^{-1}b. \tag{16}$$

Noting that, in exact arithmetic (no approximation error), Equations (14) and (16) have the same solution and the same right-hand sides, we can now use the matrix $B = A_1 + 2i\omega A_2 - \omega^2 A_3$ as a preconditioner for the scaled Helmholtz equation (16). In the absence of discretization and round-off errors, this preconditioner is *exact!* This is the fundamental difference between our approach and other approaches known to us, particularly the family of shifted-Laplace preconditioners [7, 12, 18, 21]. Of course, in the presence of discretization errors, we expect this preconditioner to be only approximate. In particular, if the discretization errors are significant, the performance may deteriorate as h increases. However, as a preconditioner, we hope that there are few enough mistreated eigenvalues that the outer Krylov solver remains efficient.

6 Analysis of the preconditioner

To further understand this preconditioner, we analyze it in 2D. First, we consider the problem with simple Dirichlet boundary conditions, constant coefficients $\kappa = c = 1$, and $\boldsymbol{\alpha}_1 = 2^{-\frac{1}{2}}(1, 1)^\top$. We take $\omega = 2n\pi$ and discretize the equations on a fixed grid with meshsize $h = 1/64$. We then generate the preconditioned matrix for the central and upwind discretizations and numerically evaluate the eigenvalues of the preconditioned system. The results are plotted in Figure 7.

These figures show that when the product $\xi = \omega h$ is sufficiently small, the eigenvalues tend to concentrate around the point $z = 1$ in the complex plane. For larger values of

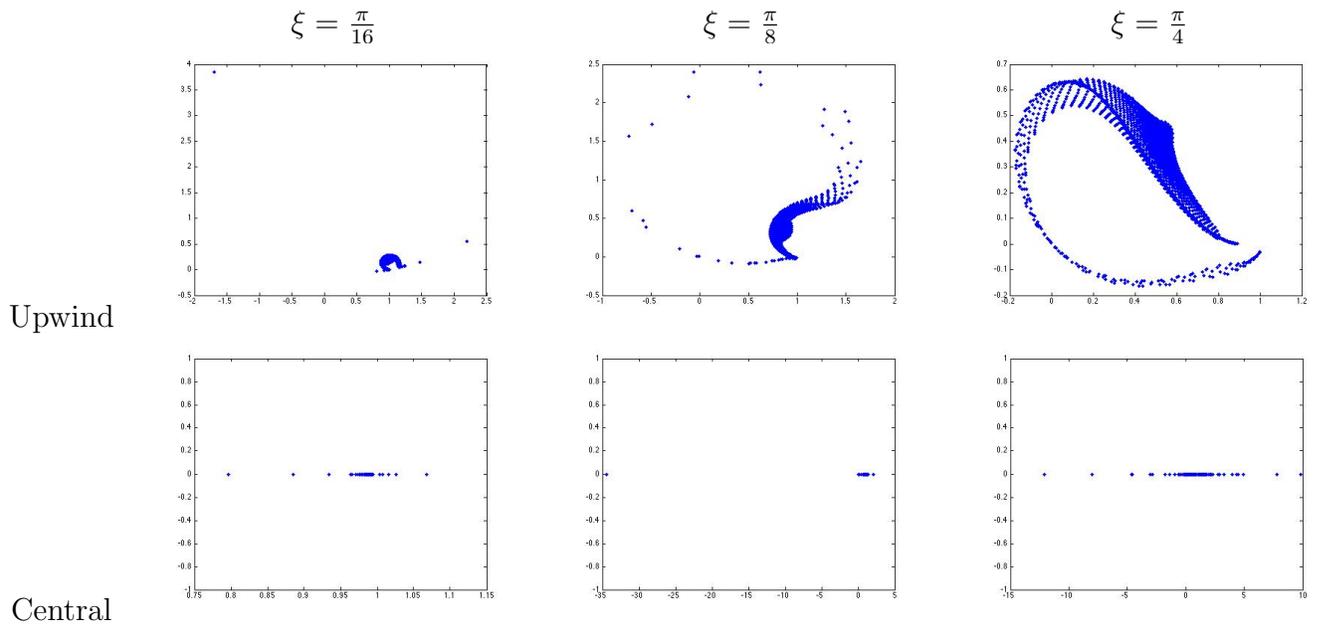


Figure 7: Eigenvalues of the preconditioned Helmholtz system with $h = 64^{-1}$. The top row corresponds to the upwind discretization of the advection term, while the bottom row corresponds to central differences. Note that for the central difference discretization the eigenvalues are real.

ξ , there is deterioration, and the eigenvalues are spreading. Nevertheless, for the upwind discretization, the eigenvalues remain away from 0, which indicates that our approach could lead to a very efficient preconditioner. For the central difference discretization, however, while the eigenvalues are real (as both the system matrix, A , and the preconditioning matrix, B , are Hermitian), they are much more spread out than those for the upwind discretization. Overall, it appears that the upwind preconditioner is better suited for this problem.

To better understand these observations, we use local Fourier analysis of the preconditioned system. It seems at first that this approach is not viable, because the multiplication of the Helmholtz equation by the matrix $M = \text{diag}(e^{i\omega c \boldsymbol{\alpha} \cdot \mathbf{x}})$ yields non-constant coefficients in the discretized operator. However, we will show by direct calculation that the similarity transformation with the infinite-grid extension of M in the preconditioning step does not change the diagonalization property of the Fourier basis (although it does change which eigenvector is associated with each eigenvalue). The goal of this analysis is to see the effect of the preconditioner on eigenvalues which are not 0. For simplicity, we again choose $\kappa = c = 1$.

We can view the preconditioned operator in two steps. First, we "scale" the discrete Helmholtz matrix

$$A_s = \text{diag}(e^{-i\omega \boldsymbol{\alpha} \cdot \mathbf{x}})(\Delta^h + \omega^2) \text{diag}(e^{i\omega \boldsymbol{\alpha} \cdot \mathbf{x}}).$$

To analyze the effect of the scaling, we look at its effect on a Fourier component of the form

$$v = e^{i\frac{\boldsymbol{\theta} \cdot \mathbf{x}}{h}} \quad \boldsymbol{\theta} \in (-\pi, \pi]^2.$$

It is straightforward to verify that the vector v is an eigenvector, since

$$\begin{aligned} h^2 A_s v &= h^2 \text{diag}(e^{-i\omega \boldsymbol{\alpha} \cdot \mathbf{x}})(\Delta^h + \omega^2) \text{diag}(e^{i\omega \boldsymbol{\alpha} \cdot \mathbf{x}})v = h^2 e^{-i\omega \boldsymbol{\alpha} \cdot \mathbf{x}} (\Delta^h + \omega^2) e^{i(\frac{\omega h \boldsymbol{\alpha} + \boldsymbol{\theta}}{h}) \cdot \mathbf{x}} \\ &= e^{-i\omega \boldsymbol{\alpha} \cdot \mathbf{x}} (2 \cos(\omega h \alpha_1 + \theta_1) + 2 \cos(\omega h \alpha_2 + \theta_2) - 4 + h^2 \omega^2) e^{i\frac{\omega h \boldsymbol{\alpha} + \boldsymbol{\theta}}{h} \cdot \mathbf{x}} \\ &= (2 \cos(\omega h \alpha_1 + \theta_1) + 2 \cos(\omega h \alpha_2 + \theta_2) - 4 + h^2 \omega^2) e^{i\frac{\boldsymbol{\theta} \cdot \mathbf{x}}{h}}. \end{aligned}$$

Next, consider the action of the convection-diffusion discretized using the upwind method.

$$\begin{aligned} h^2 B v &= h^2 (\Delta^h + 2i\omega \boldsymbol{\alpha} \cdot \nabla_h) v \\ &= (2 \cos(\theta_1) + 2 \cos(\theta_2) - 4 + i\omega h (\alpha_1 (3 - 4e^{-i\theta_1} + e^{-2i\theta_1}) + \alpha_2 (3 - 4e^{-i\theta_2} + e^{-2i\theta_2}))) v. \end{aligned}$$

Thus, we have as usual that $v = e^{i\frac{\boldsymbol{\theta} \cdot \mathbf{x}}{h}}$ is an eigenvector of the preconditioned system $B^{-1} A_s$, with eigenvalues

$$\lambda^{\text{upwind}}(\boldsymbol{\theta}) = \frac{2 \cos(\omega h \alpha_1 + \theta_1) + 2 \cos(\omega h \alpha_2 + \theta_2) - 4 + h^2 \omega^2}{2 \cos(\theta_1) + 2 \cos(\theta_2) - 4 + i\omega h (\alpha_1 (3 - 4e^{-i\theta_1} + e^{-2i\theta_1}) + \alpha_2 (3 - 4e^{-i\theta_2} + e^{-2i\theta_2}))}.$$

Similar calculations for the central discretization of the gradient yield

$$\lambda^{\text{central}}(\boldsymbol{\theta}) = \frac{2 \cos(\omega h \alpha_1 + \theta_1) + 2 \cos(\omega h \alpha_2 + \theta_2) - 4 + h^2 \omega^2}{2 \cos(\theta_1) + 2 \cos(\theta_2) - 4 - 2\omega h (\alpha_1 \sin(\theta_1) + \alpha_2 \sin(\theta_2))}.$$

To see the effect of this preconditioner, we need to look at its behavior for different values of $\xi = \omega h$, and for different frequencies, $\boldsymbol{\theta}$. Fixing $\boldsymbol{\alpha} = 2^{-\frac{1}{2}}(1, 1)^\top$, in Figure 8, we plot the

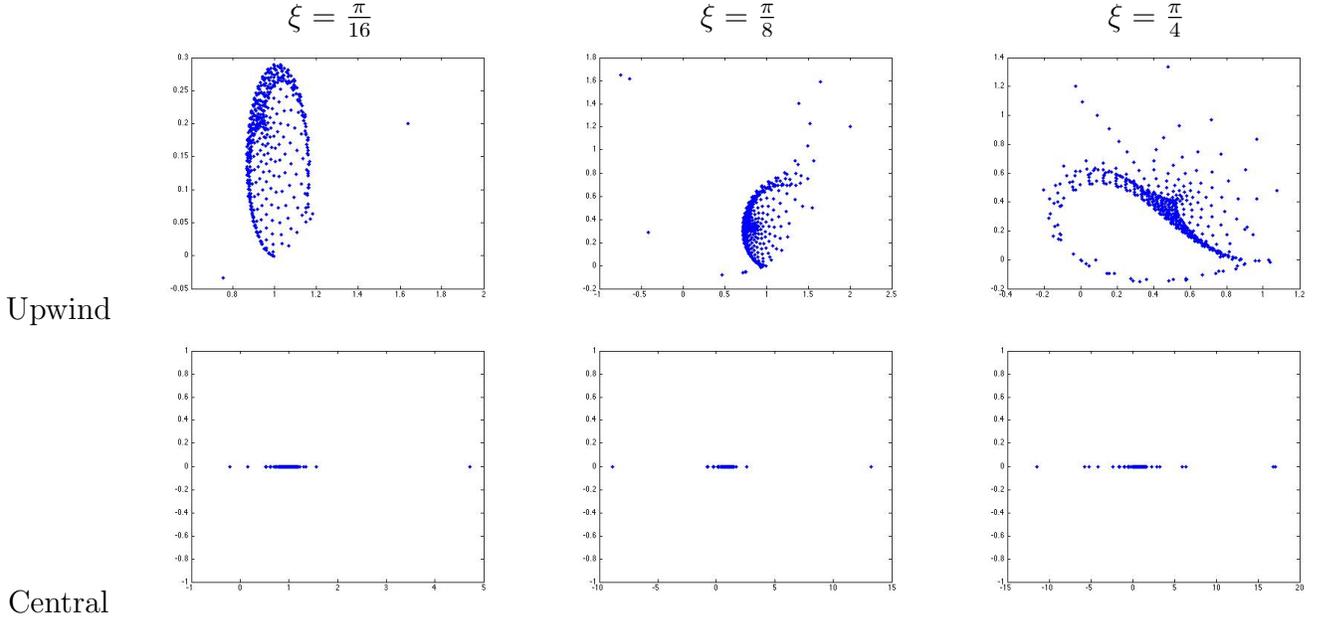


Figure 8: Eigenvalues of the preconditioned Helmholtz system, computed using Fourier analysis, for different value of $\xi = \omega h$. The top row corresponds to the upwind discretization of the advection term, while the bottom row corresponds to central differences. Note that for the central difference discretization the eigenvalues are real.

eigenvalues of the system for different ξ 's. Notice that while there is some difference in the eigenvalues shown here and those computed in Figure 7, particularly in the extreme parts of the spectrum, the general shape and clustering of the eigenvalues is quite similar in these two calculations.

We see that, as we observed in our numerical experiments, for small values of ξ the preconditioner is indeed very efficient. On the other hand, as ξ becomes larger, the eigenvalues spread and some get closer to 0, which implies worse conditioning. By the time $\xi = \frac{\pi}{4}$, this approach seems to lose some of its appeal as a preconditioner. These results are not surprising as, for large values of ξ , the discrete Helmholtz equation does not resolve the high frequencies and, thus, it is not faithful to the continuous PDE. Since our preconditioner is based on an equivalence between two continuous forms, we require both PDEs to be faithfully represented on the computational grid.

7 Numerical experiments

In this section, we conduct numerical experiments that demonstrate the strengths and weaknesses of our proposed method. Following the papers [6, 7], we conduct experiments in two dimensions, where all problems are solved on the interval $[0, 1]^2$, and we experiment with frequencies $\omega = 2\pi f$ with $f \in \{\frac{1}{2}, 1, 2, 4, 8, 16\}$. For each of our test problems, we conduct tests

ω n	32	64	128	256	512
Dirichlet					
π	3 6	2 5	2 4	2 4	2 3
2π	4 8	3 6	3 5	3 4	2 4
4π	8 22	6 10	4 7	3 5	3 4
8π	29 94	15 55	8 19	6 9	4 6
16π	<i>nc nc</i>	92 <i>nc</i>	32 56	14 48	8 19
32π	<i>nc nc</i>	<i>nc nc</i>	<i>nc nc</i>	<i>nc nc</i>	38 79
Sommerfeld					
π	9	8	8	7	7
2π	11	9	9	9	9
4π	24	13	10	9	9
8π	60	43	19	12	10
16π	<i>nc</i>	<i>nc</i>	66	32	16
32π	<i>nc</i>	<i>nc</i>	<i>nc</i>	<i>nc</i>	63

Table 5: Number of exact-preconditioned GMRES iterations for a uniform media as a function of the mesh size, n , and the frequency, ω , for the central-difference approximation of the advection term (left number in top half) and an upwind approximation (right number in top half and data in lower half). In these results, *nc* indicates no convergence after 100 iterations.

using various meshsizes and frequencies. We use both homogeneous Dirichlet boundary conditions and a combination of Neumann with Sommerfeld radiation boundary conditions. In these tests, a homogeneous Neumann boundary condition is applied on the boundary $x_2 = 0$ and, for all other sides of the box, the Sommerfeld boundary condition is implemented. This is a common set of boundary conditions for geophysical problems. For Dirichlet boundary conditions, we use the central approximation to the advection term as well as the upwind scheme. We use only the upwind discretization for the Sommerfeld boundary conditions. For the solution of the linear system, we then use full GMRES [19] as a driving iterative technique. We terminate the iteration when the relative residual is smaller than 10^{-7} or when the number of iterations exceeds either 100, in the case of exact preconditioning, or 200, in the case of multigrid preconditioning.

7.1 Constant coefficient experiment

Our first experiment deals with constant coefficients, and we set $\kappa = c = 1$. We choose a point source for the right-hand side. Results for the exact preconditioner (i.e., with direct inversion of the preconditioning matrix, B) with both Dirichlet and Sommerfeld boundary conditions are recorded in Table 5. Results for the multigrid preconditioner (i.e., using a single multigrid V-cycle per iteration to approximate the inversion of the preconditioning matrix, B) with both Dirichlet and Sommerfeld boundary conditions are recorded in Table 6.

ω	n	32	64	128	256	512
Dirichlet						
π		8 8	7 8	7 8	7 8	7 8
2π		8 10	8 9	8 8	8 8	8 8
4π		19 19	18 17	16 17	15 15	14 14
8π		87 116	92 99	82 85	76 77	68 73
16π		<i>nc nc</i>	<i>nc nc</i>	<i>nc nc</i>	101 116	89 95
32π		<i>nc nc</i>	<i>nc nc</i>	<i>nc nc</i>	<i>nc nc</i>	93 106
Sommerfeld						
π		8	7	8	8	8
2π		9	9	8	8	8
4π		20	19	17	16	14
8π		119	101	91	79	76
16π		<i>nc</i>	<i>nc</i>	<i>nc</i>	126	99
32π		<i>nc</i>	<i>nc</i>	<i>nc</i>	<i>nc</i>	115

Table 6: Number of multigrid-preconditioned GMRES iterations for a uniform media as a function of the mesh size, n , and the frequency, ω , for the central-difference approximation of the advection term (left number in top half) and an upwind approximation (right number in top half and data in lower half). In these results, *nc* indicates no convergence after 200 iterations.

The results in Table 5 clearly indicate that as long as $\omega h = \xi \leq 0.2 \approx \frac{\pi}{15}$ we obtain an efficient preconditioner. For $\xi \leq 0.1 \approx \frac{\pi}{30}$, we obtain a truly remarkable preconditioner that converges in very few iterations. The experiment above is conducted with a point source for the right-hand side; we have also run the experiment with random right-hand side and obtained similar results.

7.2 The wedge example

One common test problem is the so-called wedge problem. In this case, we assume a non-uniform slowness model, κ , with binary values of $\frac{1}{2}$ and 1. The model and the solution for a point source with Sommerfeld boundary conditions is plotted in Figure 9.

We conduct the same experiments as in the previous subsection, varying the frequencies and mesh. The results are recorded in Table 7. Results for the multigrid preconditioner (i.e., using a single multigrid V-cycle per iteration to approximate the inversion of the preconditioning matrix, B) with both Dirichlet and Sommerfeld boundary conditions are recorded in Table 8.

As can be observed, the performance of our preconditioner in this case is almost identical to the performance obtained for the constant coefficient problem.

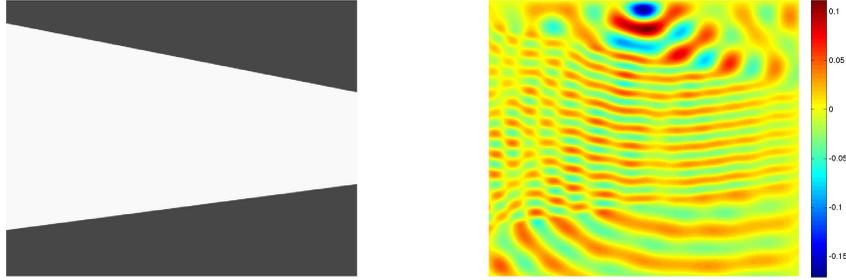


Figure 9: The wedge model (left) and its response to a point source function with frequency of 16π (right).

ω n	32	64	128	256	512
Dirichlet					
π	3 6	2 5	2 4	2 3	2 3
2π	4 8	3 6	3 5	2 4	2 3
4π	10 24	6 11	4 7	4 5	3 4
8π	37 99	17 58	10 18	6 9	4 6
16π	<i>nc nc</i>	<i>nc nc</i>	42 96	19 28	10 18
32π	<i>nc nc</i>	<i>nc nc</i>	<i>nc nc</i>	<i>nc nc</i>	46 89
Sommerfeld					
π	11	8	8	8	7
2π	25	10	9	9	8
4π	24	14	12	10	10
8π	96	53	21	14	12
16π	<i>nc</i>	<i>nc</i>	96	40	20
32π	<i>nc</i>	<i>nc</i>	<i>nc</i>	<i>nc</i>	93

Table 7: Number of exact-preconditioned GMRES iterations for the wedge model as a function of the mesh size, n , and the frequency, ω , for the central-difference approximation of the advection term (left number in top half) and an upwind approximation (right number in top half and data in lower half). In these results, *nc* indicates no convergence after 100 iterations.

ω n	32	64	128	256	512
Dirichlet					
π	8 9	7 8	7 8	7 8	7 8
2π	9 11	9 10	9 10	8 10	8 10
4π	15 28	15 27	15 26	15 25	14 25
8π	64 97	63 77	54 71	54 69	52 62
16π	<i>nc nc</i>	<i>nc nc</i>	<i>nc nc</i>	<i>nc nc</i>	99 107
32π	<i>nc nc</i>	<i>nc nc</i>	<i>nc nc</i>	<i>nc nc</i>	178 165
Sommerfeld					
π	9	8	8	8	8
2π	10	9	8	8	8
4π	19	17	16	15	15
8π	116	99	90	77	75
16π	<i>nc</i>	<i>nc</i>	<i>nc</i>	122	91
32π	<i>nc</i>	<i>nc</i>	<i>nc</i>	<i>nc</i>	103

Table 8: Number of multigrid-preconditioned GMRES iterations for the wedge example as a function of the mesh size, n , and the frequency, ω , for the central-difference approximation of the advection term (left number in top half) and an upwind approximation (right number in top half and data in lower half). In these results, *nc* indicates no convergence after 200 iterations.

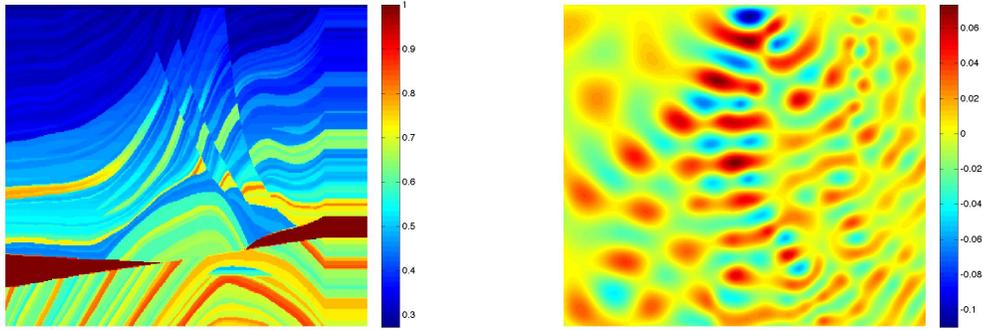


Figure 10: The Marmousi model (left) and its response to a point source function with frequency of 16π (right).

7.3 The Marmousi model

In our third experiment, we attempt to solve a standard test problem in geophysical prospecting, namely the Marmousi model [10]. The velocity model here is rather complicated and has a contrast ratio between the smallest and largest velocities of 1 to 5.5. The model and the result of a point source wave spreading within the media are plotted in Figure 10.

The results for the Marmousi model are summarized in Table 9. Results for the multigrid preconditioner (i.e., using a single multigrid V-cycle per iteration to approximate the inversion of the preconditioning matrix, B) with both Dirichlet and Sommerfeld boundary conditions are recorded in Table 10.

As can be observed in the table, the number of iterations is generally not sensitive to the overall complicated structure of the Marmousi model, and we obtain a good preconditioner as long as $\xi \leq 0.2$.

8 Conclusions

In this paper, we have developed a preconditioner for Helmholtz equation. While our development focuses on the 2D case, the extension to 3D is straightforward. The preconditioner is based on the Rytov decomposition of the solution, which yields an Eikonal equation and a complex-valued advection-diffusion-reaction equation. We use an analytic solution for the Eikonal equation and a multigrid method for the advection-diffusion-reaction equation. We show that the linear system obtained is h-elliptic for the upwind discretization but loses its h-ellipticity for coarse grids when central differences are used. The preconditioner seems to yield reasonable results and converges in very few iterations as long as the discretization of the finest problem is sufficiently resolved, which implies that the value of $\xi = \omega h$ is sufficiently small. Although this restriction may seem problematic at first, one should note that reliable error estimates requires that $\omega^3 h^2$ is bounded [2], which dictates a small value for

ω	n	32	64	128	256	512
π		3 6	2 5	2 4	2 3	2 3
2π		4 8	3 6	3 5	3 4	3 3
4π		8 19	6 9	5 6	3 5	3 4
8π		31 91	16 43	9 16	7 9	4 6
16π		<i>nc nc</i>	98 <i>nc</i>	34 86	15 35	9 15
32π		<i>nc nc</i>	<i>nc nc</i>	<i>nc nc</i>	<i>nc nc</i>	43 89
Sommerfeld						
π		9	9	8	8	8
2π		12	10	10	9	9
4π		22	15	13	12	11
8π		74	41	20	15	14
16π		<i>nc</i>	<i>nc</i>	93	33	19
32π		<i>nc</i>	<i>nc</i>	<i>nc</i>	<i>nc</i>	67

Table 9: Number of exact-preconditioned GMRES iterations for the Marmousi model as a function of the mesh size, n , and the frequency, ω , for the central-difference approximation of the advection term (left number in top half) and an upwind approximation (right number in top half and data in lower half). In these results, *nc* indicates no convergence after 100 iterations.

ωh if a reliable solution is needed. Our multigrid method is not ξ -independent; although theory suggest that it is possible to obtain a convergent multigrid method for all ξ 's (due to h-ellipticity), we have observed that a simple Gauss-Seidel smoother does not lead to such method. Therefore, we intend to explore other multigrid strategies that will allow for better performance for even higher wave numbers.

References

- [1] K. Aki and P. Richards. *Quantitative Seismology*. Academic Press, 1980.
- [2] A. Bayliss, C.I. Goldstein, and E. Turkel. On accuracy conditions for the numerical computation of waves. *Journal of Computational Physics*, 59:396–404, 1985.
- [3] A. Brandt and I. Livshits. Wave ray multigrid method for standing wave equations. *ETNA*, 6:162–181, 1997.
- [4] A. Brandt and S. Ta'asan. Multigrid methods for nearly singular and slightly indefinite problems. In W. Hackbusch and U. Trottenberg, editors, *Multigrid Methods II*, pages 99–121, Berlin, 1986. Springer–Verlag.

ω n	32	64	128	256	512
Dirichlet					
π	6 7	6 7	5 6	4 6	4 5
2π	9 10	6 8	6 7	5 6	5 6
4π	21 23	9 11	8 9	7 9	7 9
8π	92 96	26 24	19 21	20 22	19 22
16π	<i>nc nc</i>	176 188	110 114	106 113	101 108
32π	<i>nc nc</i>	<i>nc nc</i>	<i>nc nc</i>	<i>nc nc</i>	166 175
Sommerfeld					
π	7	7	7	7	6
2π	11	9	9	9	9
4π	21	19	17	16	16
8π	119	103	92	81	77
16π	<i>nc</i>	<i>nc</i>	<i>nc</i>	129	99
32π	<i>nc</i>	<i>nc</i>	<i>nc</i>	<i>nc</i>	118

Table 10: Number of multigrid-preconditioned GMRES iterations for the Marmousi model as a function of the mesh size, n , and the frequency, ω , for the central-difference approximation of the advection term (left number in top half) and an upwind approximation (right number in top half and data in lower half). In these results, *nc* indicates no convergence after 200 iterations.

- [5] M. Brezina, T. Manteuffel, S. McCormick, J. Ruge, and G. Sanders. Towards adaptive smoothed aggregation (α SA) for nonsymmetric problems. *SIAM J. Sci. Comput.*, 32:14–39, 2010.
- [6] Howard C. Elman, Oliver G. Ernst, and Dianne P. O’Leary. A multigrid method enhanced by subspace iteration for discrete Helmholtz equations. *SIAM Journal on Scientific Computing*, 23:1291–1315, 2001.
- [7] Y. A. Erlangga, C. W. Oosterlee, and C. Vuik. A novel multigrid based preconditioner for heterogeneous Helmholtz problems. *SIAM Journal on Scientific Computing*, 27:1471–1492, 2006.
- [8] Seongjai Kim and Soohyun Kim. Multigrid simulation for high-frequency solutions of the Helmholtz problem in heterogeneous media. *SIAM J. Sci. Comput.*, 24(2):684–701, 2002.
- [9] B. Lee, T. A. Manteuffel, S. F. McCormick, and J. Ruge. First order system least-squares for Helmholtz equation. *SIAM Journal on Scientific Computing*, 21:1927–1949, 2000.
- [10] G. S. Martin, R. Wiley, and K. J. Marfurt. Marmousi2: An elastic upgrade for marmousi. *The Leading Edge*, 25:156–158, 2006.
- [11] A. Ben Menahem and S. Jit. *Seismic waves and sources*. Springer-Verlag, 1981.
- [12] M. Magolu monga Made, R. Beauwens, and G. Warzee. Preconditioning of discrete helmholtz operators perturbed by a diagonal complex matrix. *Communications in Numerical Methods in Engineering*, 11:801–817, 2000.
- [13] WA Mulder and RE Plessix. A comparison between one-way and two-way wave-equation migration. *Geophysics*, 69(6):1491–1504, 2004.
- [14] Yvan Notay. A robust algebraic multilevel preconditioner for non-symmetric M -matrices. *Numer. Linear Algebra Appl.*, 7(5):243–267, 2000.
- [15] Maxim A. Olshanskii and Arnold Reusken. Convergence analysis of a multigrid method for a convection-dominated model problem. *SIAM J. Numer. Anal.*, 42(3):1261–1291, 2004.
- [16] Arnold Reusken. Fourier analysis of a robust multigrid method for convection-diffusion equations. *Numer. Math.*, 71(3):365–397, 1995.
- [17] Arnold Reusken. Convergence analysis of a multigrid method for convection-diffusion equations. *Numer. Math.*, 91(2):323–349, 2002.
- [18] C. D Riyanti, Y. A Erlangga, R. E Plessix, W. A Mulder, C Vuik, and C Oosterlee. A new iterative solver for the time-harmonic wave equation. *Geophysics*, 71(5):E57–E63, 2006.

- [19] Y. Saad. *Iterative Methods for Sparse Linear Systems*. PWS Publishing Company, 1996.
- [20] U. Trottenberg, C. Oosterlee, and A. Schuller. *Multigrid*. Academic Press, 2001.
- [21] N. Umetani, S.P. MacLachlan, and C.W. Oosterlee. A multigrid-based shifted-Laplacian preconditioner for a fourth-order Helmholtz discretization. *Numer. Linear Alg. Appl.*, 16(8):603–626, 2009.
- [22] I. Yavneh. *Multigrid Techniques for Incompressible Flows*. PhD thesis, Weizmann Institute of Science, Rehovot, Israel, 1991.