SIAM J. SCI. COMPUT.
Vol. 27, No. 4, pp. 1471–1492

A NOVEL MULTIGRID BASED PRECONDITIONER FOR HETEROGENEOUS HELMHOLTZ PROBLEMS*

Y. A. ERLANGGA†, C. W. OOSTERLEE†, AND C. VUIK†

Abstract. An iterative solution method, in the form of a preconditioner for a Krylov subspace method, is presented for the Helmholtz equation. The preconditioner is based on a Helmholtz-type differential operator with a complex term. A multigrid iteration is used for approximately inverting the preconditioner. The choice of multigrid components for the corresponding preconditioning matrix with a complex diagonal is validated with Fourier analysis. Multigrid analysis results are verified by numerical experiments. High wavenumber Helmholtz problems in heterogeneous media are solved indicating the performance of the preconditioner.

Key words. Helmholtz equation, nonconstant high wavenumber, complex multigrid preconditioner, Fourier analysis

AMS subject classifications. 65N55, 65F10, 65N22, 78A45, 76Q05

DOI. 10.1137/040615195

1. Introduction.

In this paper we present a novel preconditioner for high wavenumber Helmholtz problems in heterogeneous media. The preconditioner is based on the Helmholtz operator, where an imaginary term is added. This preconditioner can be handled by multigrid. This is somewhat surprising as multigrid, without enhancements, has convergence troubles for the original Helmholtz operator at high wavenumbers.

A part of this paper is therefore reserved for the analysis of the multigrid method for Helmholtz problems with a complex zeroth order term. This is done, for constant wavenumbers, by means of Fourier analysis. The preconditioned system leads to a favorably clustered spectrum for a Krylov subspace convergence acceleration. As the preconditioner is not based on a regular splitting of the original Helmholtz problem, it must be used in the setting of Krylov subspace methods. The particular example presented can be viewed as a generalization of the work by Bayliss, Goldstein, and Turkel [3] from the 1980s, where the Laplacian was used as a preconditioner for Helmholtz problems. This work has been generalized by Laird and Giles [17], proposing a Helmholtz preconditioner with a *positive* sign in front of the Helmholtz term. In [13] we have proposed a preconditioner with a purely imaginary shift added to the Laplacian. The method here is an improvement of that method.

In this paper we benefit from Fourier analysis in several ways. First of all, for idealized (homogeneous boundary conditions, constant coefficients) versions of the preconditioned system it is possible to visualize its spectrum for different values of the wavenumber, as Fourier analysis provides all eigenvalues. Second, for analyzing multigrid algorithms quantitatively, Fourier smoothing, two-, and three-grid analysis [6, 7, 23, 24, 30] are the tools of choice.

*Received by the editors September 16, 2004; accepted for publication (in revised form) May 19, 2005; published electronically January 27, 2006. The research is financially supported by Dutch Ministry of Economic Affairs project BTS01044.
http://www.siam.org/journals/sisc/27-4/61519.html
†Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Delft, The Netherlands (y.a.erlangga@math.tudelft.nl, c.w.oosterlee@math.tudelft.nl, c.vuik@math.tudelft.nl).

TABLE 1
Number of grid points employed, related to the wavenumber, so that $kh = 0.625$.

k	40	50	80	100	150	200	500	600
h	1/64	1/80	1/128	1/160	1/240	1/320	1/800	1/960

The outline of this paper is as follows. In section 2 the Helmholtz problem is introduced and the convergence difficulties of multigrid for this equation are detailed. The new preconditioner is introduced in section 3, where multigrid components for Helmholtz problems with a complex term (smoothing, operator-dependent prolongation) are presented. Fourier analysis to obtain quantitative performance estimates of components and methods is performed in section 4. Numerical experiments on two-dimensional high wavenumber heterogeneous Helmholtz problems are presented in section 5.

2. Helmholtz equation, standard multigrid. Consider the Helmholtz equation for a wave problem in a heterogeneous medium

$$\mathcal{A}\phi := -\partial_{xx}\phi - \partial_{yy}\phi - (1-\alpha i)k^2(x,y)\phi = g(x,y) \quad \text{in } \Omega \subset \mathbb{R}^2. \tag{1}$$

Here, $\phi = \phi(x,y)$ represents the solution, usually a pressure field, and g represents the source term. The medium is barely attenuative if $0 \le \alpha << 1$, with α indicating the fraction of damping in the medium ($i = \sqrt{-1}$, the imaginary unit). In geophysical applications, which are our main interest, this damping can be set up to 5% ($\alpha = 0.05$). The wavenumber $k = \omega_f/c$ is space-dependent because of a spatially dependent speed of sound $c(x,y)$ in a heterogeneous medium. With $\omega_f := 2\pi f$ the angular frequency (f is the frequency), wavelength ℓ is defined by $\ell = c/f$. The number of wavelengths in a domain of size L equals L/ℓ. n_w, the number of points per wavelength, is typically chosen to be 10–12 points. Wavenumber k can be large.

The dimensionless wavenumber $\overline{k}$ on a nondimensional $[0,1]^2$ domain is defined by $\overline{k} = 2\pi f L/c$. A dimensionless discretization step reads $h = \ell/(n_w L)$, and therefore for the angular frequency one finds $\omega_f = 2\pi/(n_w h) = 2\pi L/\ell$. With domain size $L = 1$, an accuracy requirement for second order discretizations is that $kh \le \pi/5 (\approx 0.63)$ for $n_w = 10$ points per wavelength, and $kh \le 0.53$ with $n_w = 12$ points per wavelength. In Table 1, the number of grid points used for several wavenumbers k is displayed. For each combination we have $kh = 0.625$. These mesh refinements assume a linear connection between k and h. However, in order to avoid a reduction of accuracy for a second order scheme due to the so-called *pollution effect* [26, 14] k^2h^3 should be set constant. As for an iterative solution method, keeping kh constant is more severe; we stay with kh as in Table 1. In this paper we emphasize the iterative solution rather than the accuracy of the discretization.

Typically, boundary conditions at the boundary $\Gamma = \partial\Omega$ are in the form of first- or second-order absorbing boundary conditions or of a perfectly matched layer (PML). We use approximate radiation (or nonreflecting) boundary conditions at an artificial boundary. The well-known second-order radiation boundary condition [12], to avoid unphysical reflections at boundaries, reads

$$\mathcal{A}_\Gamma\phi := \frac{\partial\phi}{\partial\nu} - ik\phi - \frac{i}{2k}\frac{\partial^2\phi}{\partial\tau^2} = 0 \quad \text{on } \Gamma, \tag{2}$$

with ν the outward normal direction to the boundary and τ pointing in the tangential direction. At the cornerpoints the suggestions in [2] to avoid corner reflections

have been adopted. Whereas these conditions are commonly applied for problems in homogeneous media, they are less obvious for inhomogeneous media with discontinuous wavenumbers at the boundaries. The reason is that these discontinuities may act as unphysical scatterers at the boundaries. This can be avoided by appropriately increasing damping, modeled by the imaginary part in (2). Another natural approach with an inhomogeneous medium at the boundaries may be the use of the PML.

If a discretization is applied to (1) and (2), a linear system of the form

$$A\boldsymbol{\phi} = \boldsymbol{g}, \qquad A \in \mathbb{C}^{N\times N}, \qquad \boldsymbol{\phi}, \boldsymbol{g} \in \mathbb{C}^{N}, \tag{3}$$

is obtained, where N is the number of unknowns in the computational domain Ω_h. Matrix A has complex components due to the discrete boundary operator (2) and the damping term in (1). A is in general symmetric with eigenvalues in the left and right half-plane, non-Hermitian, and, because of the accuracy requirements, also large for high wavenumbers. However, A is also sparse; its sparsity pattern depends on the discretization method used.

We consider here, in stencil notation, the well-known $O(h^2)$ 5-point discretization stencil:

$$A_h \triangleq \frac{1}{h^2}\begin{bmatrix} & -1 & \\ -1 & 4-(kh)^2(1-\alpha i) & -1 \\ & -1 & \end{bmatrix}. \tag{4}$$

We use matrix and stencil notation simultaneously: Matrix A (3) relates to the discretization of (1), (2), and discrete operator A_h (4) relates to the discretization of (1). The discrete solution is represented by $\boldsymbol{\phi}$ and ϕ_h, respectively. The eigenvalues (for constant k-problems with homogeneous Dirichlet boundary conditions)

$$\begin{aligned} \lambda_h^{\ell,m} &= \widetilde{\lambda}_h^{\ell,m} - k^2(1-\alpha i) \\ &\equiv \frac{2}{h^2}(2-\cos\ell\pi h - \cos m\pi h) - k^2(1-\alpha i) \quad (\ell, m = 1, 2, \ldots, \sqrt{N}-1) \end{aligned} \tag{5}$$

are not equal to zero as long as $k^2(1-\alpha i)$ is not equal to any of the eigenvalues of the corresponding discrete Laplacian $\widetilde{\lambda}_h^{\ell,m}$. Otherwise, the matrix is singular, and its null-space is spanned by the eigenfunctions

$$v_h^{\ell,m} = \sin\ell\pi x \sin m\pi y, \tag{6}$$

with ℓ, m for which $\lambda_h^{\ell,m} = 0$.

2.1. Multigrid convergence for the Helmholtz equation. Textbook multigrid methods are typically set up so that a smoothing method reduces high frequency components of an error between the numerical approximation and the exact discrete solution, and a coarse grid correction handles the low frequency error components. Whereas such methods are easily defined for elliptic Poisson-like equations, this is not the case for the Helmholtz equation without any damping in (1), $\alpha = 0$. Depending on the particular value of k^2, this equation gives rise to both smoothing and coarse grid correction difficulties. The matrix has eigenvalues in only the right half-plane as long as k^2 is less than the smallest eigenvalue of the Laplacian, $\widetilde{\lambda}_h^{1,1}$. For $k^2 > \widetilde{\lambda}_h^{1,1}$, the matrix does not have only positive eigenvalues. Pointwise Jacobi iteration with underrelaxation does not *converge* in that case, but since its smoothing properties are

satisfactory, the multigrid convergence will deteriorate only gradually for increasing k^2. By the time k^2 approaches the 6th eigenvalue $\widetilde{\lambda}_h^{\ell,m}$ ($k^2 \approx 150$), the standard multigrid method diverges. The Jacobi relaxation now diverges for smooth eigenfrequencies $v_h^{\ell,m}$ with $\widetilde{\lambda}_h^{\ell,m} < k^2$. Consequently, the multigrid method will still converge as long as the coarsest level used is fine enough to represent these smooth eigenfrequencies sufficiently. So, the coarsest level chosen limits the convergence. When k^2 gets larger more variables need to be represented on the coarsest level for standard multigrid convergence. Eventually, this does not result in an $\mathcal{O}(N)$ iterative method.

In addition to this feature, the Helmholtz equation also brings a multigrid coarse grid correction difficulty. Eigenvalues close to the origin may undergo a sign change after discretization on a coarser grid. If a sign change occurs, the coarse grid solution does not give a convergence acceleration to the finer grid problem but gives a severe convergence degradation (or even divergence) instead. In [11] this phenomenon is analyzed and a remedy for the coarse grid correction related to these problematic eigenvalues is proposed. The efficient treatment in [11] is that the multigrid method is combined with Krylov subspace iteration methods. GMRES is proposed as a smoother and as a cure for the problematic coarse grid correction.

Standard multigrid will also fail for k^2-values very close to eigenvalues. In that case subspace correction techniques should be employed [9].

An advanced multigrid based solution method for the Helmholtz equation is the wave-ray multigrid method [8]. The method has been adapted for a first-order system least-squares version of the Helmholtz equation in [18]. Wave-ray multigrid has been developed for Helmholtz problems with constant or smoothly varying wavenumbers. A thorough overview for the numerical solution of the Helmholtz equation is presented in [25].

3. Shifted Laplacian preconditioner. To solve (3), iterative methods based on the Krylov subspace are of interest. In particular, we choose preconditioned Bi-CGSTAB. In [13], Bi-CGSTAB is preferred over other Krylov subspace methods as the convergence for Helmholtz problems is reported typically faster than that of GMRES. We have also tested advanced versions such as Bi-CGSTAB(2) [22] and GMRESR [28], but Bi-CGSTAB remains the method of choice, especially for the Helmholtz equation without damping ($\alpha = 0$). A preconditioner $M \in \mathbb{C}^{N\times N}$ for A is developed such that the preconditioned system

$$AM^{-1}\boldsymbol{\psi} = \boldsymbol{g}, \qquad \boldsymbol{\psi} = M\boldsymbol{\phi}, \tag{7}$$

has better spectral properties than the original system. The preconditioner M proposed here is based on the following operator:

$$\mathcal{M} \equiv -\partial_{xx} - \partial_{yy} - (\beta_1 - \beta_2 i)k^2(x,y), \qquad \beta_1, \beta_2 \in \mathbb{R}, \tag{8}$$

with (β_1, β_2) parameters that can be chosen freely and with i the imaginary unit. Boundary conditions are set identical to those for the original Helmholtz problem (2).

A large imaginary value for the "wavenumber" physically corresponds to adding some form of damping for preconditioning. In the time domain the heat equation is sometimes used as the preconditioner for the Laplacian. Also the Jacobi iteration for the Laplacian can be interpreted as a time stepping procedure for the heat equation. For the wave equation such an iteration is a less common approach. A large imaginary Helmholtz term can be seen as a time-dependent term that is transformed to Fourier space.

If one is interested in only the interior problem without any damping term and real-valued boundary conditions, the original problem would be real-valued. Introducing a complex shift then introduces complex arithmetic into a real problem. Besides the extra work, it necessitates possibly different iterative solvers. Here, however, we are interested in geophysics applications with outgoing waves.

The basic choice in this paper is $(\beta_1, \beta_2) = (1, 1)$. Tuning of multigrid components is especially necessary for $\beta_2 < 1$, for example, for $\beta_2 = 0.5$, to be presented below. In [13] we have proposed a positive *purely imaginary* shift $(\beta_1, \beta_2) = (0, 1)$ to the Laplacian for a satisfactory convergence. Preconditioner (8) is an improvement of this preconditioner with $\beta_1 = 1$.

We perform Fourier analysis to visualize the effect of the choice of (β_1, β_2) in the preconditioner on the clustering of the eigenvalues of the preconditioned system. For this we consider operator (7) with homogeneous Dirichlet boundary conditions, wavenumber k constant, and a discrete version of Helmholtz operator (1), A_h, and of preconditioner (8), M_h. This particular choice of the boundary conditions especially simplifies the analysis. For radiation boundary conditions the Helmholtz operator is nonnormal. Hence eigenvalue analysis alone would not be sufficient for analyzing preconditioned Krylov subspace methods. Here, we perform the analysis as a first indication of what we can expect from the solver to be developed and concentrate on the eigenvalues.

For both A_h and M_h we choose the 5-point stencil, as in (4). The components (6) are eigenfunctions of these discrete operators with constant coefficients. With these eigenfunctions $A_h M_h^{-1}$ is diagonalizable and the eigenvalues are easily determined. In the first tests we do not include damping in A_h, $\alpha = 0$ in (1), (4).

Figure 1 presents spectra of $A_h M_h^{-1}$ for $(\beta_1, \beta_2) = (0, 0)$ (Laplacian preconditioner), $(\beta_1, \beta_2) = (-1, 0)$ (Laird preconditioner [17]), $(\beta_1, \beta_2) = (0, 1)$ (preconditioner from [13]), $(\beta_1, \beta_2) = (1, 1)$ (basic parameter choice), $(\beta_1, \beta_2) = (1, 0.5)$, and $(\beta_1, \beta_2) = (1, 0.3)$ (more advanced parameters). The results are for $k = 40$ ($k^2 = 1600$) and $h = 1/64$. Similar eigenvalue distributions are observed for finer grids.

From the spectra presented with the new preconditioner, the lower pictures of Figure 1 are favorable as their real parts vary between 0 and 1. The Laplacian preconditioner in Figure 1(a) exhibits large isolated eigenvalues; for the Laird preconditioner the eigenvalues in Figure 1(b) are distributed between -1 and 1 on the real axis. The preconditioners with complex Helmholtz terms give rise to a curved spectrum. Whereas the real part of the spectrum in Figure 1(c) still includes a part of the negative real axis, this is not the case for the (β_1, β_2)-preconditioners with $\beta_1 = 1$. The difference between Figures 1(d), 1(e), and 1(f) is that, with a smaller value of β_2, fewer outliers close to the origin are observed. This is favorable for the convergence of the preconditioned Krylov method. The approximate inversion of the preconditioner itself by multigrid, however, will be shown to be harder for smaller values of β_2. In Figure 2 the spectra for $k = 100 (k^2 = 10^4)$ are presented on a grid with $h = 1/160$ for $\beta_1 = 1$ and β_2 varying between 1 and 0.3. The spectra are very similar to those in Figure 1. More eigenvalues lie, however, in the vicinity of the origin due to the higher wavenumber and the correspondingly finer grid. Figure 3 presents the distribution of eigenvalues for the case that 5% damping ($\alpha = 0.05$) is set in $\mathcal{A}$. Parameters in the preconditioner are $(\beta_1, \beta_2) = (1, 0.5)$. Again the 5-point stencil as in (4) is used for discretization. Figure 3(a) presents the spectrum for $k = 40, h = 1/64$, and Figure 3(b) presents the spectrum for $k = 100, h = 1/160$. An interesting observation is that now the eigenvalues move away from the origin into the right half-plane. This

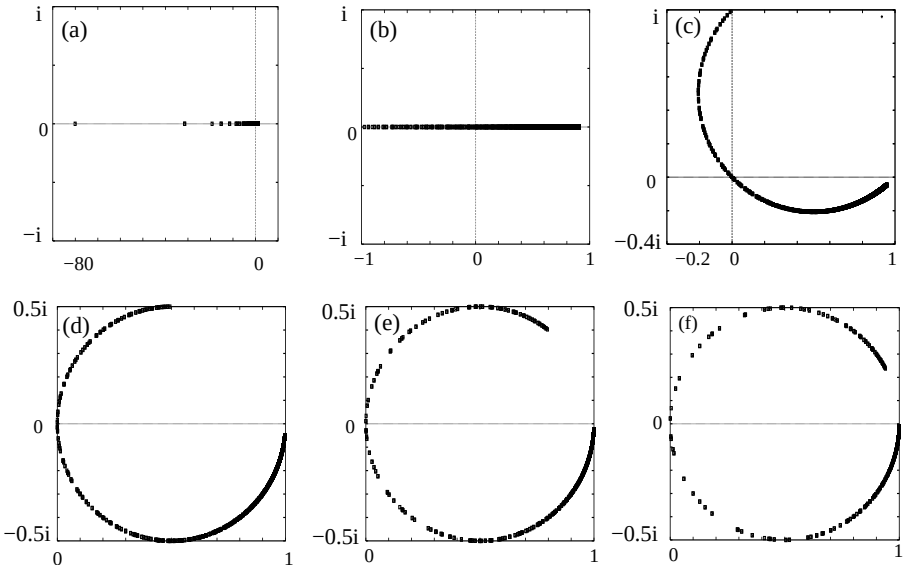


FIG. 1. *Spectral pictures of* $A_h M_h^{-1}$ *with* $\alpha = 0$ *and different values of* (β_1, β_2) *in* (8). (a) $(\beta_1, \beta_2) = (0, 0)$, (b) $(-1, 0)$, (c) $(0, 1)$, (d) $(1, 1)$, (e) $(1, 0.5)$, *and* (f) $(1, 0.3)$.

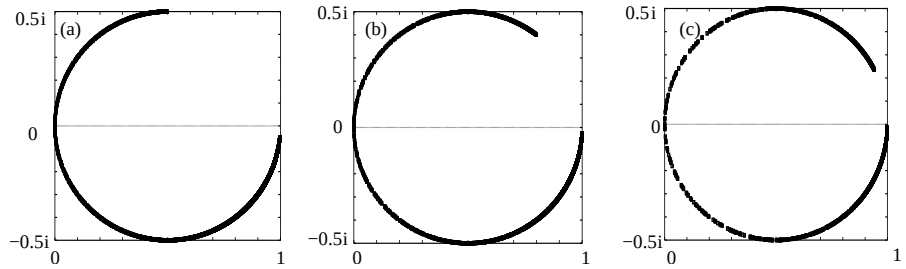


FIG. 2. *Spectral pictures of* $A_h M_h^{-1}$ *for* $k = 100$, $h = 1/160$, *and* $\alpha = 0$; (a) $(\beta_1, \beta_2) = (1, 1)$, (b) $(\beta_1, \beta_2) = (1, 0.5)$, *and* (c) $(\beta_1, \beta_2) = (1, 0.3)$.

is beneficial for iterative solution methods. From the spectra in Figure 3 it is expected that the Bi-CGSTAB (and GMRES) convergence in the case of damping will be considerably faster than for the undamped case.

4. Multigrid for the preconditioner.

4.1. Multigrid components. Geometric multigrid converges satisfactorily for the Helmholtz operator (8) for certain choices of β_1 and β_2 (assumed in [15], see also [16]). In this section, we detail the multigrid components that can be specified for approximately inverting a discrete version of $\mathcal{M}$ in (8). We consider a 5-point discretization and denote the equation for the preconditioner by $M_h \phi_h = \psi_h$. Standard multigrid coarsening, i.e., doubling the mesh size h in every direction, is chosen.

For smoothing the pointwise Jacobi relaxation with underrelaxation (ω-JAC) is chosen. This smoother is well parallelizable, which is an important aspect for our research (w.r.t. a generalization to three dimensions). In principle, one can choose the underrelaxation parameter $\omega \in \mathbb{C}$, but the Fourier analysis indicates that there is no real benefit for the problems considered. So, we choose $\omega \in \mathbb{R}$.

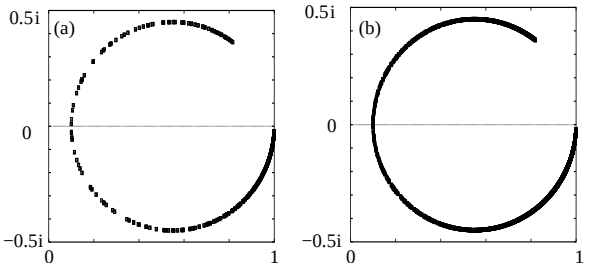


FIG. 3. *Spectral pictures of AM^{-1} with 5 % damping in A and $(\beta_1, \beta_2) = (1, 0.5)$; (a) $k = 40, h = 1/64$ and (b) $k = 100, h = 1/160$.*

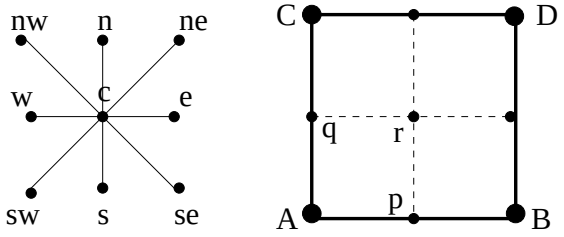


FIG. 4. *Left: Nine point stencil with numbering. Right: Coarse grid cell and four fine cells, (Coarse grid indices designated by capital letters and fine grid indices designated by lower case letters).*

The coarse grid correction components are also based on established operators. For the discrete coarse grid operators $M_{2h}, M_{4h}, \ldots$, the Galerkin coarse grid operator is used:

$$M_{2h} := R_h^{2h} M_h P_{2h}^h, \quad M_{4h} := R_{2h}^{4h} M_{2h} P_{4h}^{2h}, \quad \text{etc.}$$

In the Fourier analysis to follow, this discretization will be compared to a direct coarse grid discretization of (1). The Galerkin coarse grid discretization is a natural choice for heterogeneous problems. Also with boundary conditions containing first and second derivatives, it is convenient to choose the Galerkin coarse grid discretization as it defines the appropriate coarse grid boundary stencils automatically. The transfer operators used in building the coarse grid operators are the same as those used for transferring coarse and fine grid quantities to fine and coarse grids, respectively. The prolongation operator considered is an operator-dependent interpolation based on de Zeeuw's transfer operators [31]. Originally, this prolongation was set up for general (possibly unsymmetric) real-valued matrices with a splitting of matrix M into a symmetric and an antisymmetric part, $M_s = \frac{1}{2}(M + M^T)$, $M_t = M - M_s$ in [31]. However, since the discretization here leads to a complex symmetric matrix, the prolongation is adapted and briefly explained for such matrices with nine diagonals. The numbering in a stencil for the explanation of the prolongation is as shown in Figure 4 (left side). The right side of Figure 4 shows one coarse and four fine grid cells with indices for the explanation of the interpolation weights. Capital letters denote coarse grid points and lower case letters denote fine grid points. Operator element m_p^w, for example, denotes the west element of operator M_h at point p on the fine grid. The corrections from the coarse to the fine grid are obtained by interpolation among nearest coarse grid neighbors. The operator-dependent interpolation weights,

w, to determine the fine grid correction quantities e_h are derived with the following formulas.

- For fine grid points p in Figure 4, $e_{h,p} = w_A e_{H,A} + w_B e_{H,B}$.
  $w_A = \min(1, \max(0, w_w))$; $w_B = \min(1, \max(0, w_e))$,

where

$$d_w = \max(|m_p^{sw} + m_p^w + m_p^{nw}|, |m_p^{sw}|, |m_p^{nw}|), \tag{9}$$

$$d_e = \max(|m_p^{se} + m_p^s + m_p^{ne}|, |m_p^{se}|, |m_p^{ne}|), \tag{10}$$

$$w_w = \frac{d_w}{d_w + d_e}, \qquad w_e = \frac{d_e}{d_w + d_e}. \tag{11}$$

- For fine grid points q in Figure 4, $e_{h,q} = w_A e_{H,A} + w_C e_{H,C}$.
  $w_A = \min(1, \max(0, w_s))$; $w_C = \min(1, \max(0, w_n))$,

with

$$d_n = \max(|m_q^{nw} + m_q^n + m_q^{ne}|, |m_q^{nw}|, |m_q^{ne}|), \tag{12}$$

$$d_s = \max(|m_q^{sw} + m_q^s + m_q^{se}|, |m_q^{sw}|, |m_q^{se}|), \tag{13}$$

$$w_s = \frac{d_s}{d_s + d_n}, \qquad w_n = \frac{d_n}{d_s + d_n}. \tag{14}$$

On the remaining points the prolongation is defined as follows:

$$\text{On fine grid points that are also coarse points, } e_h(A) = e_{2h}(A). \tag{15}$$

$$\text{On points } r,\ e_h(r) \text{ is determined so that } \quad M_h P_{2h}^h e_{2h} = 0 \text{ at } r. \tag{16}$$

The interpolation weights are the same as in [31] but are specially tailored to the symmetric complex Helmholtz equation, i.e., the unsymmetric components in [31] have been removed. $|.|$ denotes the modulus, in this case, leading to real-valued interpolation weights. As for symmetric problems with jumping coefficients, the prolongation operator by de Zeeuw [31] is very similar to the original operator-dependent prolongation in [1]. In [1], for d_w, for example, the lumped sum of three elements, $m_p^{sw} + m_p^w + m_p^{nw}$, is chosen. For satisfactory convergence it is, however, important to consider the *modulus* of the operator elements, as in (9), (10), (12), and (13), in the definition of the interpolation weights. This prolongation is also valid at boundaries.

The full weighting operator is employed as the restriction operator. So, we do not choose the adjoint of the prolongation operator, which is commonly used but is not absolutely necessary, as already stated in [1] (an example where the restriction is not the adjoint of the prolongation operator has been given in [10]). We choose the combination of a full weighting restriction and the operator-dependent interpolation, as it brings a robust convergence for a variety of Helmholtz problems with constant and nonconstant coefficients. For constant coefficients and mildly varying wavenumbers, bilinear interpolation also gives very satisfactory convergence results, but for strongly varying coefficients, as in the Marmousi problem discussed in section 5.3, a robust and efficient convergence on different grid sizes and for many frequencies is observed for the combination of the transfer operators chosen.

4.2. Fourier analysis. Fourier smoothing and two-grid analysis, two classical multigrid analysis tools, have been used for quantitative estimates of the smoothing properties and of the other multigrid components in a two-grid method [5, 6, 7, 23, 24]. Consider a discretization of (7), (8), $M_h \phi_h = \psi_h$, where ϕ_h represents the exact discrete solution. The error $w_h^l = \phi_h^l - \phi_h$ after the lth iteration is transformed by a

two-grid cycle as

$$w_h^{l+1} = T_h^{2h}\, w_h^l, \qquad T_h^{2h} = S_h^{\nu_2} K_h^{2h} S_h^{\nu_1}, \qquad K_h^{2h} = I_h - P_{2h}^h (M_{2h})^{-1} R_h^{2h} M_h. \tag{17}$$

M_h, M_{2h} correspond to discretizations of (8) on the h-, $2h$-grid, S_h is the smoothing operator on the fine grid, and I_h is the identity operator. ν_l $(l = 1, 2)$ represents the number of pre- and postsmoothing steps, and R_h^{2h} and P_{2h}^h denote the restriction and prolongation operator, respectively. In the analysis we assume an equidistant grid with $\sqrt{N}$ points in each direction. The $O(h^2)$-discrete complex Helmholtz operator from (8) with constant wavenumber and Dirichlet boundary conditions belongs to the class of symmetric stencils. For these stencils it is possible to apply Fourier analysis on the basis of discrete sine-eigenfunctions $v_h^{\ell,m}$, $\ell, m = 1, \ldots, \sqrt{N} - 1$ (6), instead of the local Fourier analysis with exponential functions. For problems with symmetric stencils and homogeneous Dirichlet boundary conditions, this analysis can predict h-dependent convergence factors. From the discussion of multigrid methods for the original Helmholtz equation, it seems necessary to gain insight into the h-dependency of the multigrid methods developed also for the complex Helmholtz operator. (The definition of the operator-dependent prolongation and the Galerkin coarse grid stencils in section 4.1 also leads to symmetric operators that can be analyzed within this framework.)

For the pointwise Jacobi smoother, the $v_h^{\ell,m}$ (6) are also eigenfunctions of the smoothing operator. This is not true for the two-grid iteration operator T_h^{2h}. However, 4-dimensional linearly independent spaces, the *harmonics*,

$$E_h^{\ell,m} = \left[v_h^{\ell,m},\ v_h^{\sqrt{N}-\ell,\sqrt{N}-m},\ -v_h^{\sqrt{N}-\ell,m},\ -v_h^{\ell,\sqrt{N}-m} \right] \quad \text{for} \quad \ell, m = 1, \ldots, \tfrac{\sqrt{N}}{2} \tag{18}$$

are invariant under these operators. One can show [23, 24] that

$$\begin{aligned} &M_h : \operatorname{span}[v_h^{\ell,m}] \to \operatorname{span}[v_h^{\ell,m}], \quad (M_{2h})^{-1} : \operatorname{span}[v_{2h}^{\ell,m}] \to \operatorname{span}[v_{2h}^{\ell,m}], \\ &S_h : \operatorname{span}[v_h^{\ell,m}] \to \operatorname{span}[v_h^{\ell,m}], \\ &R_h^{2h} : E_h^{\ell,m} \to \operatorname{span}[v_{2h}^{\ell,m}], \quad P_{2h}^h : \operatorname{span}[v_{2h}^{\ell,m}] \to E_h^{\ell,m}, \end{aligned}$$

and $T_h^{2h} : E_h^{\ell,m} \to E_h^{\ell,m}$ $(\ell, m = 1, \ldots, \frac{\sqrt{N}}{2})$. Therefore, the representation of T_h^{2h} with respect to $E_h^{\ell,m}$ leads to a block-diagonal matrix, $\widetilde{T}_h^{2h}$,

$$T_h^{2h} \mathrel{\hat{=}} \left[\widehat{T}_h^{2h}(\ell, m) \right]_{\ell,m=1,\ldots,\frac{\sqrt{N}}{2}} =: \widetilde{T}_h^{2h} \ . \tag{19}$$

Here the blocks $\widehat{T}_h^{2h}(\ell, m)$ are 4×4 matrices if $\ell, m < \frac{\sqrt{N}}{2}$, and are 2×2 (1×1) matrices if either $\ell = \frac{\sqrt{N}}{2}$ or $m = \frac{\sqrt{N}}{2}$ $(\ell = \frac{\sqrt{N}}{2}$ and $m = \frac{\sqrt{N}}{2})$. The two-grid convergence factor is defined as

$$\rho_{2g} := \max_{1 \le \ell, m \le \frac{\sqrt{N}}{2}} \rho\Big(\widehat{T}_h^{2h}(\ell, m) \Big) \ . \tag{20}$$

Thus, the spectral radii of at most 4×4 matrices $\widehat{T}_h^{2h}(\ell, m)$ have to be determined, and their maximum with respect to ℓ and m has to be found.

The definition of the smoothing factor μ is closely related. The smoothing factor measures the reduction of high frequency error components by an iterative method.

It is based on a coarse grid correction operator that annihilates the low frequency error components completely and keeps the high frequency components unchanged. K_h^{2h} is replaced by a projection operator Q_h^{2h} mapping onto the space of high frequencies, i.e., a block-diagonal matrix with $\widehat{Q}_h^{2h}$ at most 4×4-diagonal blocks defined by $\mathrm{diag}(0,1,1,1)$. So, μ is computed as ρ_{2g} (20) with $\widehat{K}_h^{2h}$ in $\widehat{T}_h^{2h}$ replaced by $\widehat{Q}_h^{2h}$.

Recently, three-grid Fourier analysis was proposed in [30]. An issue that can be analyzed in some more detail with a third grid is the coarse grid correction. If a large difference occurs between the two-grid and the three-grid convergence factors, ρ_{2g} and ρ_{3g}, this is an indication for a problematic coarse grid correction. For the complex Helmholtz preconditioner it is important to analyze the coarse grid correction carefully. The error transformation by a three-grid cycle is given by

$$\begin{aligned} w_h^{l+1} &= T_h^{4h}\, w_h^l \quad \text{with} \\ T_h^{4h} = S_h^{\nu_2} K_h^{4h} S_h^{\nu_1} \quad &\text{and} \quad K_h^{4h} = I_h - P_{2h}^h (I_{2h} - (T_{2h}^{4h})^\gamma)(M_{2h})^{-1} R_h^{2h} M_h. \end{aligned} \tag{21}$$

Here T_{2h}^{4h}, defined by (17), reads $T_{2h}^{4h} = S_{2h}^{\nu_2}(I_{2h} - P_{4h}^{2h}(M_{4h})^{-1}R_{4h}^{2h})S_{2h}^{\nu_1}$. M_{4h} corresponds to $4h$-grid discretization of (8); S_{2h} is the smoothing operator, and I_{2h} is the identity on the $2h$-grid; and R_{2h}^{4h} and P_{4h}^{2h} are transfer operators between the different grids. The $2h$-equation is solved approximately in a three-grid cycle (21) by performing γ two-grid iterations T_{2h}^{4h} with zero initial approximation; see also [23, 30].

The three-grid analysis is a recursive application of the two-grid analysis. Four frequencies are coupled not only in the transition from the h- to the $2h$-grid but also in the transition from the $2h$- to the $4h$-grid. Thus the three-grid error transformation operator couples 16 Fourier frequencies. As a consequence, T_h^{4h} is unitarily equivalent to a block-diagonal matrix $\widetilde{T}_h^{4h}$ with at most 16×16 blocks, $\widehat{T}_h^{4h}(\ell, m)$. The block matrices are composed of the Fourier symbols from the two-grid analysis, which is due to the recursive application of the two-grid analysis. One may compute the three-grid factor ρ_{3g} as the supremum of the spectral radii from the 16×16 block matrices, $\widehat{T}_h^{4h}(\ell, m)$.

For more details about the three-grid analysis, we refer to [30]. Three-grid Fourier analysis software, based on the exponential functions, is freely available; see http://www.mgnet.org/mgnet-codes-wienands.html.

4.3. Fourier analysis and multigrid results. We first compare the numerical multigrid convergence with asymptotic convergence factors $\mu, \rho_{2g}, \rho_{3g}$ from Fourier analysis. For this, we consider here solely the preconditioner $\mathcal{M}$ (8). (The behavior of the complete solution method will be considered in the next section.) Wavenumber k is taken as a constant here and a square domain with an equidistant grid is used. The second-order boundary conditions (2) are set in the numerical experiments to mimic reality.

An interesting aspect is that almost identical convergence factors are obtained, both from the analysis and from the actual experiments, for constant values of kh. They are set as in Table 1. The results are validated from $k = 40$ up to $k = 600$, the highest wavenumber tested is ($k^2 = 3.6 \times 10^5$). During testing the following abbreviations are used: "ω-JAC" is the Jacobi smoother with underrelaxation, "Galerkin" is the Galerkin coarse grid discretization, and "direct" is a direct coarse grid discretization of the PDE. "Direct" has not been implemented in the numerical code, but it can be used in the analysis framework.

Multigrid coarsening is continued until fewer than 10×10 points are processed on the coarsest grid. The number of levels is h- and therefore also k-dependent, as

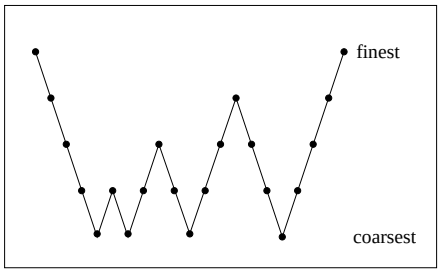


FIG. 5. *An F-cycle for five grids.*

TABLE 2

Comparison of asymptotic convergence from Fourier analysis with numerical multigrid convergence, $(\beta_1, \beta_2) = (0, 1)$. μ *is the smoothing factor;* ρ_{2g}, ρ_{3g} *are the two- and three-grid convergence factors from Fourier analysis; and* ρ_h *is the numerical multigrid convergence factor. The smoother is* ω*-JAC with* $\omega = 0.8$.

(ν_1, ν_2)	μ	ρ_{2g}	ρ_{3g}, $\gamma = 2$	ρ_h, F-cycle
(1,0)	0.60	0.60	0.60	0.58
(1,1)	0.36	0.36	0.36	0.34

kh is kept constant on the finest grid, and varies between 5 and 9 grids.

The F-cycle (see Figure 5) is always used in the numerical tests; the V-cycle's performance was generally too poor; and the W-cycle is considered too expensive on the very fine grids processed at high wavenumbers. The F-cycle often shows the robustness of the W-cycle at the efficiency of the V-cycle. In the three-grid analysis, $\gamma = 2$, the W-cycle analysis is used.

Remark. The Fourier analysis applied directly to the Helmholtz equation (1) with $\alpha = 0$ and the specified mesh sizes gives a satisfactory smoothing factor, but the two- and three-grid analysis convergence factors and also the actual multigrid results show a strong divergence, as expected.

The case $(\beta_1, \beta_2) = (0, 1)$. We start with $(\beta_1, \beta_2) = (0, 1)$, as in [13]. This case is not of the highest interest as a preconditioner, as the Bi-CGSTAB convergence for the corresponding preconditioned system is worse than with $\beta_1 = 1$ (shown in the next section). This case $(\beta_1, \beta_2) = (0, 1)$ serves as a reference for the comparison between Fourier analysis and numerical convergence. The underrelaxation parameter ω is set to $\omega = 0.8$, as this is the optimal choice for the Laplacian [24]. The agreement between the smoothing two- and three-grid Fourier analysis results with one and two smoothing iterations and the numerical convergence is excellent, presented in Table 2. The results obtained are very similar to the convergence factors for the Laplacian with ω-JAC.

Remark. For the case $(\beta_1, \beta_2) = (0, 1)$, one can adopt the well-known multigrid components: direct PDE coarse grid discretization and red-black Gauss–Seidel relaxation. This gives $\rho_{3g} = 0.16$ for $\gamma = 1$ and $\rho_{3g} = 0.08$ for $\gamma = 2$ with two smoothing iterations, very similar to the Laplacian situation. Red-black Gauss–Seidel relaxation is, however, not as robust as the ω-JAC relaxation for the $\beta_1 = 1$ cases. Furthermore, the cost in CPU time on a Linux PC of one red-black Gauss–Seidel iteration is about twice that of a Jacobi iteration.

The case $(\beta_1, \beta_2) = (1, 1)$. The second test is for $(\beta_1, \beta_2) = (1, 1)$. In this test we employ ω-JAC smoothing with $\omega = 0.7$ in an F(1,1)-cycle ($\nu_1 = \nu_2 = 1$). It is necessary to adapt the relaxation parameter ω for satisfactory numerical convergence. The

TABLE 3

Comparison of convergence $(\beta_1, \beta_2) = (1, 1)$ with Fourier analysis convergence ($\gamma = 1$), ω-JAC, $\omega = 0.7$, and $F(1, 1)$-cycle. Coarse grid discretizations are compared. (The direct discretization has not been implemented).

Coarse discr.	μ	ρ_{2g}	ρ_{3g}, $\gamma = 2$	ρ_h, F(1,1)
Galerkin	0.47	0.47	0.47	0.45
Direct	0.47	0.47	0.47	-

TABLE 4

Fourier analysis convergence factors compared to multigrid convergence $(\beta_1, \beta_2) = (1, 0.5)$. The smoother is ω-JAC with $\omega = 0.5$. (The direct discretization has not been implemented).

Coarse discr.	μ	ρ_{2g}	ρ_{3g}, $\gamma = 2$	ρ_h, F(1,1)
Galerkin	0.60	0.60	0.60	0.61
Direct	0.60	0.60	0.60	-

performance of the ω-JAC smoother is not sensitive with respect to choosing somewhat smaller values of ω. We compare the Galerkin discretization with the direct coarse grid PDE discretization. Analysis results with two smoothing iterations are shown in Table 3, and they are compared to the numerical F(1,1) multigrid convergence.

Convergence factors well below 0.5 are obtained with the F(1,1)-cycle and ω-JAC relaxation with $\omega = 0.7$. The Fourier analysis results with the Galerkin coarse grid discretization are very similar to those obtained with a direct coarse grid PDE discretization.

The case $(\beta_1, \beta_2) = (1, 0.5)$. The preconditioner of choice in this paper is based on the parameters $(\beta_1, \beta_2) = (1, 0.5)$. For this parameter set it is possible to define a converging multigrid iteration by means of an F(1,1)-cycle, ω-JAC relaxation with $\omega = 0.5$, and a Galerkin coarse grid discretization. The underrelaxation parameter needs to be adapted for a robust convergence for a variety of heterogeneous Helmholtz problems. For values $\beta_2 < 0.5$ it is very difficult to define a satisfactory converging multigrid F(1,1)-cycle with the components at hand. They are therefore not considered.

Table 4 compares the Galerkin with the direct PDE coarse grid discretization. Also here, the operator-dependent interpolation and full weighting restriction are chosen, and two smoothing iterations are applied. The smoothing factors and two- and three-grid factors are very similar, which is an indication for the proper choice of coarse grid correction components for the problems under investigation. The numerical convergence with the F(1,1)-cycle is again very similar to the Fourier results.

In the following three remarks, we explain the satisfactory convergence of a standard multigrid method for the complex Helmholtz equation and $\beta_1 = 1$ with some heuristic arguments. The remarks show that, when β_2 is chosen small but such that the multigrid method still converges, the coarse grid stencils do not represent the fine grid problem well. In particular, the main diagonal operator elements in (22) on different coarse grids are smaller than the off-diagonal elements. This will be the reason why the V-cycle does not perform as well as the F-cycle and why damped Jacobi relaxation is a more robust smoother than red-black Gauss–Seidel relaxation for the case $\beta_2 = 0.5$.

Remark: Smoothing. The Fourier symbol of ω-JAC for the complex Helmholtz equation reads

TABLE 5
Smoothing factors μ_h for ω-JAC on different coarse grids and various (β_1, β_2)-values.

(β_1, β_2)	ω in ω-JAC	h			
		1/64	1/32	1/16	1/8
$(1, 0)$	0.7	0.47	0.75	2.31	0.18
$(0, 1)$	0.8	0.36	0.32	0.13	0.05
$(1, 1)$	0.7	0.47	0.56	0.35	0.13
$(1, 0.5)$	0.5	0.60	0.77	0.81	0.32

$$S_h = 1 - \frac{\omega}{4 - (\beta_1 - \beta_2 i)(hk)^2} \left(4 - (\beta_1 - \beta_2 i)(hk)^2 - 2\cos \ell\pi h - 2\cos m\pi h\right),$$
$$\ell, m = 1, \ldots, \sqrt{N} - 1.$$

We consider the case $k = 40, h = 1/64$ and take ω as in the previous experiments. Table 5 presents smoothing factors on four consecutive grids for $(\beta_1, \beta_2) = (1, 0)$ (the original Helmholtz equation) and for $(\beta_1, \beta_2) = (0, 1), (1, 1)$, and $(1, 0.5)$. For simplicity, a direct PDE discretization on the coarse grids has been used. From Table 5, one confirms that for $h = 1/16, \omega$-JAC diverges for the original Helmholtz operator (also found with other relaxation parameters). This is in accordance with the remarks in [9, 11] that smoothing problems do not occur on the very fine or the very coarse grids but do occur on the intermediate grids. Furthermore, it can be observed that the $(\beta_1, \beta_2) = (0, 1)$-preconditioner resembles a Laplacian-type situation, with excellent smoothing factors on all grids. The preconditioners with $\beta_1 = 1, \beta_2 \neq 0$ give smoothing factors less than one on every grid. The (1,1)-preconditioner exhibits better smoothing factors than the set $(\beta_1, \beta_2) = (1, 0.5)$, which represents a limit case for which smoothing factors are still below one.

Remark: Simplified coarse grid analysis. Some insight into the coarse grid correction can be gained from the so-called "simplified coarse grid analysis" or first-differential-approximation analysis [7, 9, 24]. As in [11] we apply this analysis for a one-dimensional (1D) Helmholtz operator. Assuming that transfer operators do not have any effect on the lowest frequencies, the quantity $1 - \lambda_h^\ell / \lambda_{2h}^{2\ell}$ (ℓ small) gives some insight into the relation between the discrete fine and coarse grid operators. This quantity should be close to zero and is an indication of the suitability of a coarse grid operator in a multigrid method. For the original 1D Helmholtz equation and $\alpha = 0$ (no damping), this quantity reads [11]

$$1 - \lambda_h^\ell / \lambda_{2h}^{2\ell} = \frac{\sin^4(\ell h\pi/2)}{\sin^2(\ell h\pi/2)\cos^2(\ell h\pi/2) - (kh/2)^2}, \qquad \ell = 1, \ldots, N.$$

It may give rise to a problematic coarse grid correction in the range where

$$\sin^2(\ell h\pi/2)\cos^2(\ell h\pi/2) \approx (kh/2)^2$$

and ℓ is associated with a smooth mode. For a 1D version of the *complex* Helmholtz operator, this quantity reads

$$1 - \lambda_h^\ell / \lambda_{2h}^{2\ell} = \frac{\sin^4(\ell h\pi/2)}{\sin^2(\ell h\pi/2)\cos^2(\ell h\pi/2) - (kh/2)^2(\beta_1 - \beta_2 i)}$$
$$= \frac{\sin^4(\ell h\pi/2)\left(\sin^2(\ell h\pi/2)\cos^2(\ell h\pi/2) - (kh/2)^2(\beta_1 + \beta_2 i)\right)}{\left(\sin^2(\ell h\pi/2)\cos^2(\ell h\pi/2) - (kh/2)^2\beta_1\right)^2 + (kh/2)^2\beta_2^2},$$
$$\ell = 1, \ldots, N.$$

This expression for the complex Helmholtz operator is close to zero for the (β_1, β_2)-sets under consideration: the denominator does not reach zero, and the numerator contains the term $\sin^4 \ell h\pi/2$ which is very small for smooth eigenmodes.

Remark: h-ellipticity. When a Galerkin coarse grid discretization is used, it is difficult to gain insight into the coarse grid correction, as the coarse grid stencil elements are constructed with nontrivial formulas. Therefore, we discuss here for the case $(\beta_1, \beta_2) = (1, 0.5)$ two coarse grid discretizations. With $h = 1/64, k = 40$, $\alpha = 0$ in (4), we obtain by direct PDE discretization similar coarse grid stencils as the fine grid stencil with grid sizes $2h$ or $4h$, respectively. In that case, only the central stencil element contains an imaginary contribution. When the Galerkin coarse grid operator is employed, the imaginary part is distributed over all entries. With operator-dependent interpolation and full weighting restriction we find

$$A_{2h} \triangleq \begin{bmatrix} -282.9 + 15.3i & -665.8 + 80.6i & -282.9 + 15.3i \\ -665.8 + 80.6i & 2164.5 + 461.2i & -665.8 + 80.6i \\ -282.9 + 15.3i & -665.8 + 80.6i & -282.9 + 15.3i \end{bmatrix},$$

$$A_{4h} \triangleq \begin{bmatrix} -129.5 + 43.0i & -290.1 + 135.0i & -129.5 + 43.0i \\ -290.1 + 135.0i & -101.4 + 483.2i & -290.1 + 135.0i \\ -129.5 + 43.0i & -290.1 + 135.0i & -129.5 + 43.0i \end{bmatrix}. \tag{22}$$

The h-ellipticity measures are 0.28 and 0.18, indicating the suitability of the stencils for pointwise smoothing [7, 24]. For the direct PDE discretization, the h-ellipticity measures are 0.13 and 0.45 for the $2h$- and $4h$-discretizations, respectively. The fact that these qualitative measures are not close to zero means that pointwise smoothers can be constructed for these stencils. From these complicated coarse grid stencils it is, however, difficult to judge between the different smoothers, relaxation parameters, etc., but the three-grid Fourier analysis helps to some extent. We obtain very satisfactory multigrid convergence with simple multigrid components, although the coarse grid discretization (22) seems awkward. At least it does not spoil the h-independent multigrid convergence. One merely needs to choose the underrelaxation parameter in the smoother with some care.

4.4. Multigrid for the preconditioner. One multigrid iteration is taken for approximating the inverse of the operator in (8). After some experimentation it was found that it is sufficient to employ a multigrid iteration with a convergence factor $\rho_h \approx 0.6$ for the preconditioner. To some extend this can also be observed qualitatively from spectral pictures obtained by Fourier analysis (again, constant k, Dirichlet boundary conditions). Starting with a regular splitting of M_h,

$$C_h \phi_h^{l+1} = (C_h - M_h)\phi_h^l + \psi_h, \quad \text{or} \quad \phi_h^{l+1} = (I_h - C_h^{-1} M_h)\phi_h^l + C_h^{-1}\psi_h. \tag{23}$$

This splitting is considered to represent a multigrid iteration, with iteration matrix $(I_h - C_h^{-1}M_h)$ and C_h^{-1} an *approximation* of M_h^{-1}. T_h^{2h} in (17) represents the two-grid version of a multigrid iteration matrix. Therefore, we equate $T_h^{2h} = I_h - C_h^{-1}M_h$. Matrix $\widetilde{T}_h^{2h}$ in (19) is a block matrix related to T_h^{2h}: $\widetilde{T}_h^{2h} = U_h T_h^{2h} U_h^{-1}$, where U_h is a unitary matrix with four consecutive rows defined by the orthogonal eigenvectors related to (6). U_h transforms the two-grid iteration matrix into the block-diagonal matrix $\widetilde{T}_h^{2h}$. Clearly,

$$\widetilde{T}_h^{2h} = I_h - U_h C_h^{-1} M_h U_h^{-1}, \text{ and}$$
$$U_h C_h^{-1} M_h U_h^{-1} = U_h C_h^{-1} U_h^{-1} U_h M_h U_h^{-1} =: \widetilde{C}_h^{-1} \widetilde{M}_h$$

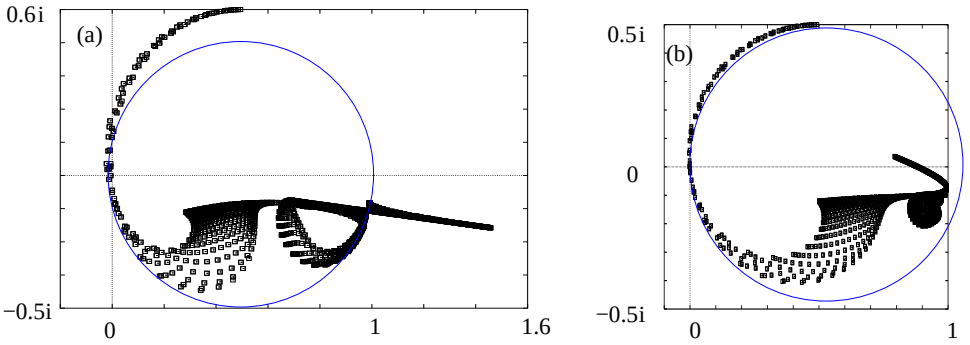


FIG. 6. *Spectral pictures of the preconditioned system with one two-grid iteration used for preconditioning* $(\beta_1, \beta_2) = (1, 1)$, $k = 40, h = 1/64$*:* (a) *one* ω*-JAC relaxation and* (b) *two* ω*-JAC relaxations,* $\omega = 0.7$*. (The eigenvalues with the exact inversion lie at the circles.)*

is in block-diagonal form. We have $\widetilde{C}_h^{-1}\widetilde{M}_h\widetilde{M}_h^{-1} = (I_h - \widetilde{T}_h^{2h})\widetilde{M}_h^{-1}$. So, the expression for the block-diagonal form $\widetilde{A}_h\widetilde{C}_h^{-1}$ ($\widetilde{C}_h^{-1}$ is the approximation of $\widetilde{M}_h^{-1}$) from (7) reads

$$\widetilde{A}_h\widetilde{C}_h^{-1} = \widetilde{A}_h(I_h - \widetilde{T}_h^{2h})\widetilde{M}_h^{-1}. \tag{24}$$

As all the symbols of the operators in the right-hand side of (24) can be formed easily with Fourier two-grid analysis, the corresponding eigenvalues can be visualized for various multigrid cycles. These spectra can be compared to those in Figure 1, where operator M_h from (8) is inverted exactly. Figure 6, for example, presents the spectrum of the $(\beta_1, \beta_2) = (1, 1)$-preconditioned system where a two-grid iteration is used for preconditioning for wavenumber $k = 40$ ($h = 1/64$). Figure 6(a) shows the spectrum for one ω-JAC ($\omega = 0.7$) smoothing iteration for which $\rho_{2g} \approx 0.7$, whereas Figure 6(b) shows the two-grid spectral picture with two ω-JAC smoothing iterations, $\nu_1 + \nu_2 = 2$, operator-dependent interpolation, full weighting restriction, and Galerkin coarse grid discretization ($\rho_{2g} = 0.45$). Figure 6(b) shows a spectrum that coincides well with the spectrum related to the exact inversion in Figure 1(d), whereas in Figure 6(a) eigenvalues are also outside the circle obtained with the exact inversion.

Figure 7 presents the spectra with a two-grid iteration for the $(\beta_1, \beta_2) = (1, 0.5)$-preconditioner and Galerkin coarsening, with ω-JAC relaxation ($\omega = 0.5$). Figure 7(a) is for $\nu = 1$; Figure 7(b) is for $\nu = 2$. Also for this approximate inversion of the preconditioner the spectrum obtained in Figure 7(b) compares well with the exact inversion in Figure 1(e), indicating that one multigrid iteration with two ω-JAC smoothing steps may be sufficient for approximating M_h^{-1}.

Indeed, a numerical comparison between inverting the preconditioner by several multigrid iterations versus by only one multigrid iteration did not lead to substantially different numbers of Krylov subspace iterations for solving the Helmholtz problem.

5. Applications. In this section the overall solution method, preconditioned Bi-CGSTAB for the indefinite heterogeneous Helmholtz problems (1) with the complex Helmholtz (β_1, β_2)-preconditioner, is evaluated. One multigrid F(1,1)-cycle is used for approximately inverting the preconditioner equation with the complex Helmholtz operator. Three problems of increasing difficulty are discussed.

5.1. Constant wavenumber. For constant wavenumbers k the Bi-CGSTAB convergence for the Helmholtz equation with the three preconditioners is presented.

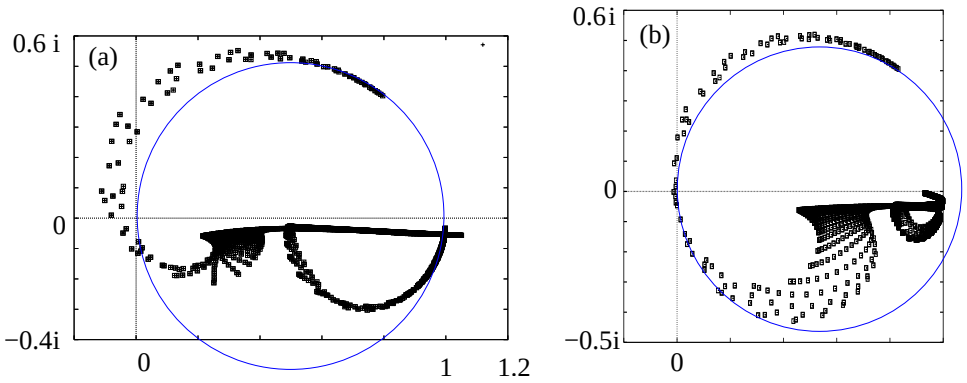


FIG. 7. *Spectral pictures of preconditioned system with one two-grid iteration used for preconditioning* $(\beta_1, \beta_2) = (1, 0.5)$, $k = 40, h = 1/64$*:* (a) *one* ω*-JAC relaxation and* (b) *two* ω*-JAC relaxations,* $\omega = 0.5$. *(The eigenvalues with exact inversion lie at the circles.)*

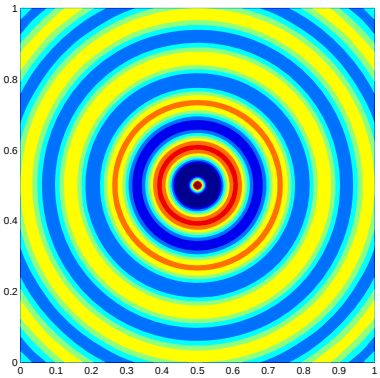


FIG. 8. *Numerical solution at* $k = 50$ *for the model problem with* k *constant.*

We consider a square domain $\Omega = (0, 1)^2$. A point source is located at the center of the domain. The solution satisfies the second-order conditions (2). In these experiments the finest grid size for each wavenumber is as shown in Table 1. The numerical solution corresponding to $k = 50$ is presented in Figure 8. Unphysical reflections at the boundaries are not present due to the boundary treatment.

A zero initial guess has been used during the computations. The Bi-CGSTAB iteration is terminated as soon as the initial residual is reduced by 7 orders of magnitude. Note that each Bi-CGSTAB iteration involves two preconditioning steps.

For all three preconditioners, $(\beta_1, \beta_2) = (0, 1), (1, 1)$, and $(1, 0.5)$, the method chosen to approximately invert the preconditioner consists of one multigrid F(1,1)-cycle with ω-JAC, operator-dependent interpolation plus full weighting as the transfer operators, and a Galerkin coarse grid discretization. The only difference is the value of the underrelaxation parameter in ω-JAC, which is $\omega = 0.8$ for $(\beta_1, \beta_2) = (0, 1)$, $\omega = 0.7$ for $(\beta_1, \beta_2) = (1, 1)$, and $\omega = 0.5$ for $(\beta_1, \beta_2) = (1, 0.5)$. The results for different values of k and $(\beta_1, \beta_2) = (0, 1)$ are presented in the upper part of Table 6. In the middle part of Table 6, the Bi-CGSTAB convergence with the $(\beta_1, \beta_2) = (1, 1)$-preconditioner is presented. In the lower lines of Table 6 the $(\beta_1, \beta_2) = (1, 0.5)$-preconditioner is employed. Next to the results for the Helmholtz equation without

TABLE 6

Number of preconditioned Bi-CGSTAB iterations and CPU time in seconds (in parentheses) to reduce the initial residual by 7 orders. Damping parameter α is varied in the Helmholtz problem.

		k				
(β_1, β_2)	α from (1)	40	50	80	100	150
$(0, 1)$	$\alpha = 0$	57 (0.44)	73 (0.92)	112 (4.3)	126 (7.7)	188 (28.5)
	2.5% damping	48 (0.38)	61 (0.77)	84 (3.3)	93 (5.6)	121 (18.5)
	5% damping	45 (0.35)	55 (0.70)	69 (2.7)	75 (4.7)	97 (14.9)
(1,1)	$\alpha = 0$	36 (0.30)	39 (0.51)	54 (2.2)	74 (4.5)	90 (13.9)
	2.5% damping	33 (0.27)	37 (0.48)	44 (1.8)	51 (3.2)	61 (9.6)
	5% damping	28 (0.24)	30 (0.39)	36 (1.5)	41 (2.6)	49 (7.5)
(1,0.5)	$\alpha = 0$	26 (0.21)	31 (0.40)	44 (1.8)	52 (3.3)	73 (10.8)
	2.5% damping	24 (0.20)	26 (0.35)	33 (1.4)	39 (2.5)	47 (7.3)
	5% damping	21 (0.18)	23 (0.32)	28 (1.2)	32 (2.1)	37 (5.8)

TABLE 7

High wavenumbers, number of Bi-CGSTAB iterations, and CPU time in seconds (in parentheses) needed to reduce the initial residual by 7 orders with and without damping in the Helmholtz problem.

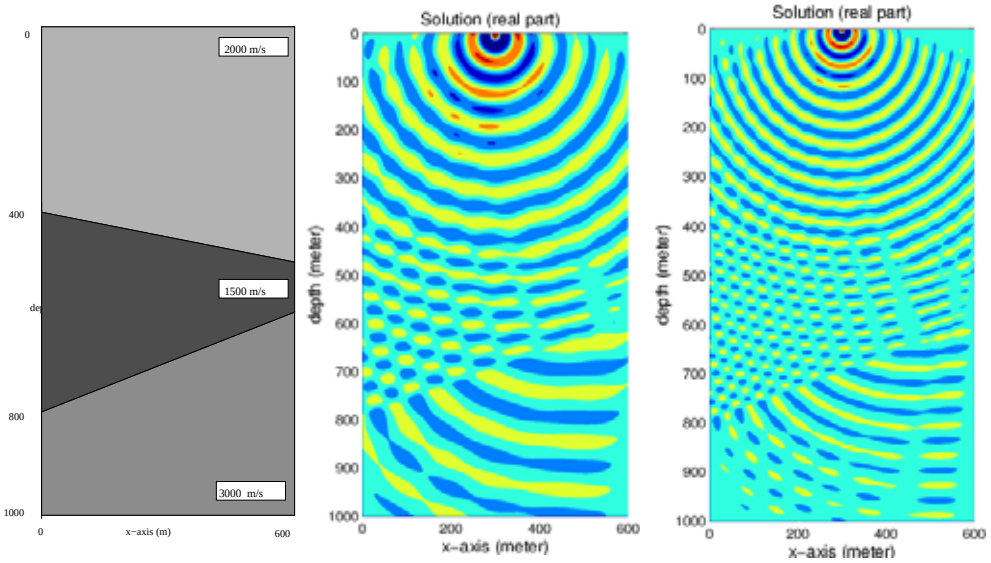
		k		
(β_1, β_2)	α in (1)	200	500	600
(1,1)	$\alpha = 0$	114 (30.8)	291 (515)	352 (890)
	2.5 % damping	74 (20.2)	125 (227)	145 (372)
	5 % damping	56 (15.5)	95 (174)	80 (205)
(1,0.5)	$\alpha = 0$	92 (25.4)	250 (425)	298 (726)
	2.5 % damping	57 (15.2)	91 (164)	102 (252)
	5 % damping	44 (11.9)	64 (115)	66 (165)

any damping ($\alpha = 0$), we also show the convergence with 2.5% ($\alpha = 0.025$) and 5% ($\alpha = 0.05$) damping. The number of Bi-CGSTAB iterations are presented as well as the CPU time on a Pentium 4 PC with 2.4 Ghz and 2 Gb RAM. From the results in Table 6 we conclude that the preferred methods among the choices are the preconditioners with $\beta_1 = 1$. This was already expected from the spectra in Figure 1. Fastest convergence is obtained for $(\beta_1, \beta_2) = (1, 0.5)$. The components of the multigrid iteration for this preconditioner have been validated with the help of the Fourier analysis.

Table 6 shows that the Bi-CGSTAB convergence with some damping in the Helmholtz problem is considerably faster than for $\alpha = 0$. This was already expected from the spectra in Figure 3. Furthermore, the number of iterations in the case of damping grows only slowly for increasing wavenumbers, especially for the $(\beta_1, \beta_2) = (1, 0.5)$-preconditioner.

The difference between the two preconditioners with $\beta_1 = 1$ is more pronounced if we compute higher wavenumbers. The Bi-CGSTAB convergence and CPU time for the higher wavenumbers, without and with damping in the Helmholtz problem, are presented in Table 7. Also for the higher wavenumbers damping in the Helmholtz problem by means of $\alpha \neq 0$ improves the convergence significantly. Very satisfactory convergence is found for high wavenumbers on fine grids.

5.2. The wedge model. A problem of intermediate difficulty is the wedge model. It is used to evaluate the preconditioner's behavior for a simple heterogeneous medium. The problem is adopted from [19]. The domain is defined to be a rectangle of dimension $600 \times 1000\ m^2$. The second-order boundary conditions (2) are set, and a point source is located at the center of the upper surface (which is assigned



(a) (b) (c)

FIG. 9. *Wedge problem:* (a) *Problem geometry with velocity profile indicated,* (b) *real part of numerical solution at* 30 *Hz, and* (c) *real part of numerical solution at* 50 *Hz.*

to be $y = 0$) with frequency, $f = \overline{k}c/(2\pi)L$, varying from 10 to 60 Hz (where c is the speed of sound). The corresponding values of the local dimensionless wavenumbers $\overline{k}$ vary between 20 (smallest for 10 Hz) and 240 (biggest for 60 Hz). For the problem at 10 Hz approximately 18 points per wavelength are used. Figure 9(a) presents the domain, the wedge, and the variation of c in the medium. The variation of c is due to the different local properties of the medium. The real part of the numerical solution for the wedge problem at 30 Hz and 50 Hz is plotted in Figures 9(b) and 9(c).

In the preconditioner (8) wavenumber $k(x, y)$ is chosen as in the original problem. Also the boundary conditions in the preconditioner are as for the original problem. The number of Bi-CGSTAB iterations with one multigrid iteration for the preconditioner with $(\beta_1, \beta_2) = (0, 1), (1, 1)$, and $(1, 0.5)$ are displayed in Table 8 for frequencies ranging from 10 to 60 Hz on corresponding grid sizes. Results with and without damping in the Helmholtz problem are presented. The only difference in the multigrid methods for the preconditioner is the value of the relaxation parameter: for $(\beta_1, \beta_2) = (0, 1)$ $\omega = 0.8$, for $(\beta_1, \beta_2) = (1, 1)$ $\omega = 0.7$, and for $(\beta_1, \beta_2) = (1, 0.5)$ $\omega = 0.5$. A zero initial guess has been used as a starting approximation. The convergence results for $(\beta_1, \beta_2) = (1, 0.5)$ are best, also without any damping in the original problem. The convergence with the (1,0.5)-preconditioner is about 1.5 times faster than with the (1,1)-preconditioner and about 3 times faster than with the (0,1)-preconditioner. The Bi-CGSTAB convergence for the wedge problem for $\alpha = 0$ and different frequencies are also visualized for $(\beta_1, \beta_2) = (1, 0.5)$ in Figure 10.

5.3. The Marmousi problem. This example is a part of the full Marmousi problem which mimics subsurface geology [4]; see also [19]. The domain is rectangular with a dimension of $6000 \times 1600\ m^2$. A point source is placed at the center of the upper surface. The values for the speed of sound c are irregularly structured throughout the

TABLE 8

Bi-CGSTAB convergence for the wedge problem with and without damping and the three multigrid based (β_1, β_2)-preconditioners compared. The number of Bi-CGSTAB iterations and the CPU time in seconds (in parentheses) are shown.

f (Hz)	Grid	Damping	(β_1, β_2) (0,1)	(1,1)	(1,0.5)
10	75 × 125	0.0%	52 (1.2)	30 (0.67)	19 (0.42)
		2.5%	48 (1.1)	27 (0.62)	17 (0.39)
		5.0%	42 (0.91)	25 (0.57)	16 (0.38)
20	149 × 249	0.0%	91 (8.8)	45 (4.5)	27 (2.8)
		2.5%	75 (7.2)	39 (4.0)	23 (2.4)
		5.0%	65 (6.3)	35 (3.5)	20 (2.1)
30	232 × 386	0.0%	128 (30.6)	64 (15.8)	37 (9.4)
		2.5%	94 (22.8)	49 (12.3)	29 (7.5)
		5.0%	86 (21.0)	42 (10.7)	25 (6.6)
40	301 × 501	0.0%	161 (66.1)	80 (33.5)	49 (20.8)
		2.5%	116 (48.0)	60 (25.4)	35 (15.2)
		5.0%	91 (37.9)	46 (19.8)	28 (12.4)
50	376 × 626	0.0%	205 (134.5)	98 (65.5)	58 (38.7)
		2.5%	135 (89.0)	67 (45.5)	37 (24.8)
		5.0%	99 (66.5)	54 (37.1)	32 (22.0)
60	481 × 801	0.0%	232 (247.3)	118 (127.6)	66 (71.9)
		2.5%	147 (159.1)	74 (81.1)	42 (47.1)
		5.0%	110 (119.6)	58 (64.5)	32 (36.7)

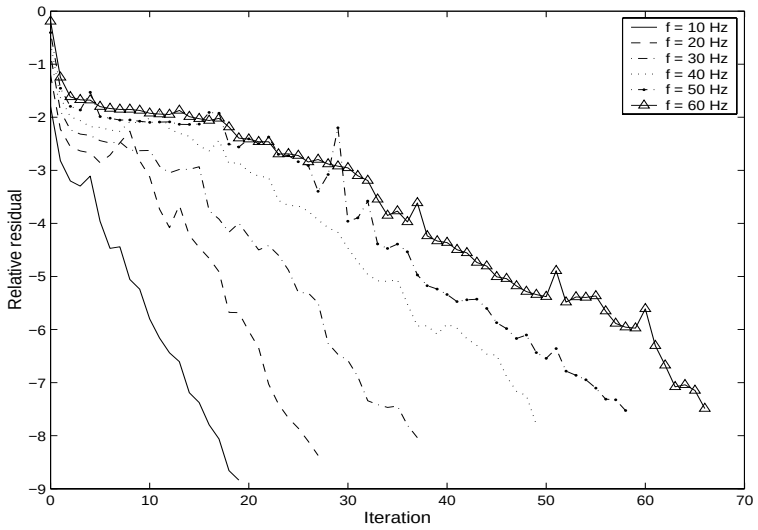


FIG. 10. *Bi-CGSTAB convergence plot for $(\beta_1, \beta_2) = (1, 0.5)$ for the wedge problem at different frequencies, $\alpha = 0$.*

domain; see Figure 11(a). The minimum number of points per wavelength equals 17. The frequency is varied between 1 and 30 Hz.

Preconditioning consists of one multigrid iteration for the complex Helmholtz equation with the multigrid components prescribed. The underrelaxation parameter in ω-JAC is varied as usual depending on (β_1, β_2). In the preconditioner again the wavenumbers $k(x, y)$ are as in the original problem. Also the boundary conditions are as in the original problem. Table 9 presents the number of Bi-CGSTAB iterations to solve the indefinite Helmholtz–Marmousi problem with the CPU times required shown

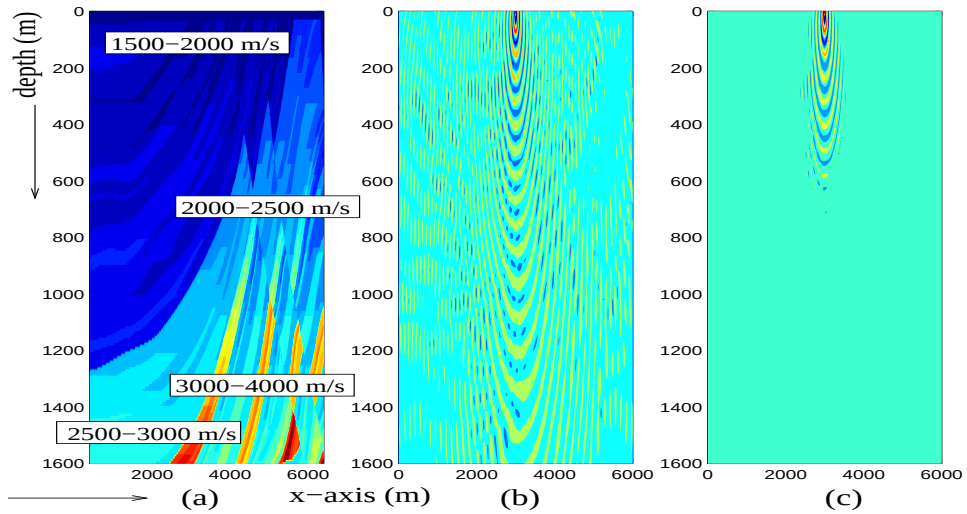


FIG. 11. *Marmousi problem (not to scale).* (a) *Velocity distribution in meter/s,* (b) *real part of the solution for* $f = 20$ *Hz, no damping, and* (c) *real part of the solution for* $f = 20$ *Hz,* 2.5% *damping.*

TABLE 9

Bi-CGSTAB convergence for the Marmousi problem with and without damping and the three multigrid based (β_1, β_2)*-preconditioners are presented. The number of Bi-CGSTAB iterations and the CPU time in seconds (in parentheses) are shown.*

f (Hz)	Grid	Damping	(β_1, β_2) (0,1)	(1,1)	(1,0.5)
1	751 × 201	0.0%	74 (37.5)	54 (27.6)	38 (19.7)
		2.5%	67 (34.1)	55 (28.2)	32 (17.0)
		5.0%	64 (32.5)	53 (27.3)	31 (16.5)
10	751 × 201	0.0%	180 (89.2)	84 (42.4)	47 (24.2)
		2.5%	119 (59.4)	59 (30.1)	33 (17.5)
		5.0%	96 (48.3)	48 (24.8)	28 (15.0)
20	1501 × 401	0.0%	414 (832.3)	168 (308.7)	104 (212.1)
		2.5%	203 (410.8)	88 (179.9)	55 (115.3)
		5.0%	145 (294.7)	64 (133.4)	37 (79.3)
30	2001 × 534	0.0%	458 (1724.8)	211 (799.4)	136 (519.4)
		2.5%	197 (745.6)	94 (361.1)	58 (226.8)
		5.0%	119 (455.3)	61 (238.4)	38 (151.9)

in parentheses. Results are presented for $\alpha = 0, 0.025$, and 0.05. A zero initial guess has been used. The $(\beta_1, \beta_2) = (1, 0.5)$-preconditioner shows a satisfactory and robust convergence for this problem with irregularly varying wavenumbers. For $\alpha = 0.05$ the number of iterations increases only very slowly for increasing frequencies. With the $(\beta_1, \beta_2) = (1, 0.5)$-preconditioner the CPU time is reduced by a factor of 3 compared to the performance of the $(\beta_1, \beta_2) = (0, 1)$-preconditioner for the challenging problems. The difference with the $(\beta_1, \beta_2) = (1, 1)$ is less pronounced but still significant.

The real parts of the solutions at 20 Hz for $\alpha = 0$ and $\alpha = 0.025$ are presented in Figures 11(b) and 11(c). The effect of damping of the solution is significant, as can be deduced from these global pictures. However, in the actual geophysics applications, some damping is present. Figure 11(c) may therefore be a more realistic solution for the real application. An adaptation of the solution method presented to a variant

of the Helmholtz problem in which the damping parameter α is varying *locally* is an easy generalization in our preconditioner.

6. Conclusions. In this paper a complex Helmholtz preconditioner has been proposed for handling indefinite Helmholtz problems in heterogeneous media. In the preconditioner we advocate the use of a Helmholtz operator with a negative real term, as in the original Helmholtz problem, plus a positive imaginary Helmholtz part. Multigrid is proposed for approximately inverting the complex Helmholtz operator in the preconditioner. As the original Helmholtz problems are defined on square domains with structured grids (a common choice for many geophysical applications), a geometric multigrid method based on cartesian grids can be defined for the complex Helmholtz preconditioner. Extension of the geometric multigrid to the complex Helmholtz operator is straightforward. We can employ very similar multigrid components as those for the Laplacian, with some variation of the underrelaxation parameter in the point-wise Jacobi smoother, and an operator-dependent prolongation operator to deal with highly varying wavenumbers. These components are validated by Fourier analysis tools. The smallest size of the β_2-parameter in front of the imaginary Helmholtz term in the preconditioner, for which the multigrid method can be successfully employed, has been determined. After a parameter study all parameters appearing in the method are fixed (i.e., $\beta_1 = 1, \beta_2 = 0.5, \omega = 0.5$ in ω-JAC).

Bi-CGSTAB, preconditioned with a multigrid iteration for the complex Helmholtz operator, proves to be an efficient and robust iterative solution method to solve heterogeneous high wavenumber Helmholtz problems. The applications ranged from constant wavenumber to irregular heterogeneity structures in a medium. The multigrid components have been chosen such that the solution method is well parallelizable. The method proposed and the corresponding analysis are easily generalized to three dimensions. This is currently being done.

Acknowledgments. The use of Roman Wienands freely available Fourier analysis software *LFA00_scalar* and the particular version *RFA00_scalar* has been extremely helpful and is gratefully acknowledged.

The authors thank Rene-Edouard Plessix and Wim A. Mulder at Shell International Exploration and Production, Rijswijk, The Netherlands, for fruitful discussions and for providing the Marmousi data.

REFERENCES

[1] R. E. Alcouffe, A. Brandt, J. E. Dendy Jr., and J. W. Painter, *The multi-grid method for the diffusion equation with strongly discontinuous coefficients*, SIAM J. Sci. Comput., 2 (1981), pp. 430–454.

[2] A. Bamberger, P. Joly, and J. E. Roberts, *Second-order absorbing boundary conditions for the wave equation: A solution for the corner problem*, SIAM J. Numer. Anal., 27 (1990), pp. 323–352.

[3] A. Bayliss, C. I. Goldstein, and E. Turkel, *An iterative method for Helmholtz equation*, J. Comput. Phys., 49 (1983), pp. 443–457.

[4] A. Bourgeois, M. Bourget, P. Lailly, M. Poulet, P. Ricarte, and R. Versteeg, *Marmousi, model and data*, in The Marmousi Experience: EAEG, Workshop on Practical Aspects of Seismic Data Inversion, 1991, pp. 5–16.

[5] K. Brackenridge, *Multigrid and cyclic reduction applied to the Helmholtz equation*, in Proceedings of the 6th Annual Copper Mountain Conference on Multigrid Methods (CP3224), N. D. Melson, T. A. Manteuffel, and S. F. Mc Cormick, eds., NASA, Hampton, VA, 1993, pp. 31–41.

[6] A. Brandt, *Multi-level adaptive solutions to boundary-value problems*, Math. Comp., 31 (1977), pp. 333–390.

[7] A. Brandt, *Multigrid Techniques: 1984 Guide with Applications to Fluid Dynamics*, GMD-Studie 85, Gesellech. Math. Datenverarbeitung, Sankt Augustin, Germany, 1984.
[8] A. Brandt and I. Livshits, *Wave-ray multigrid methods for standing wave equations*, Electron. Trans. Numer. Analysis, 6 (1997), pp. 162–181.
[9] A. Brandt and S. Ta'asan, *Multigrid method for nearly singular and slightly indefinite problems*, in Proceedings EMG'85 Cologne, Multigrid Methods II, W. Hackbusch, and U. Trottenberg, eds., Springer, Berlin, 1986, pp. 99–121.
[10] J. E. Dendy Jr., *Blackbox multigrid for nonsymmetric problems*, Appl. Math. Comput., 13 (1983), pp. 261–283.
[11] H. R. Elman, O. G. Ernst, and D. P. O'Leary, *A multigrid method enhanced by Krylov subspace iteration for discrete Helmholtz equations*, SIAM J. Sci. Comput., 23 (2001), pp. 1291–1315.
[12] B. Engquist and A. Majda, *Absorbing boundary conditions for the numerical simulation of waves*, Math. Comp., 31 (1997), pp. 629–651.
[13] Y. A. Erlangga, C. Vuik, and C. W. Oosterlee, *On a class of preconditioners for the Helmholtz equation*, Appl. Numer. Math., 50 (2004), pp. 409–425.
[14] F. Ihlenburg and I. Babuska, *Finite element solution to the Helmholtz equation with high wave numbers*, Comput. Math. Appl., 30 (1995), pp. 9–37.
[15] S. Kim and S. Kim, *Multigrid simulation for high-frequency solutions of the Helmholtz problem in heterogeneous media*, SIAM J. Sci. Comput., 24 (2002), pp. 684–701.
[16] D. Lahaye, H. De Gersem, S. Vandewalle, and K. Hameyer, *Algebraic multigrid for complex symmetric systems*, IEEE Trans. Magn., 36 (2000), pp. 1535–1538.
[17] A. L. Laird and M. B. Giles, *Preconditioned Iterative Solution of the 2D Helmholtz Equation*, report 02/12, Oxford Computer Laboratory, Oxford, UK, 2002.
[18] B. Lee, T. A. Manteuffel, S. F. McCormick, and J. Ruge, *First-order system least-squares for the Helmholtz equation*, SIAM J. Sci. Comput., 21 (2000), pp. 1927–1949.
[19] R. E. Plessix and W. A. Mulder, *Separation-of-variables as a preconditioner for an iterative Helmholtz solver*, Appl. Numer. Math., 44 (2003), pp. 385–400.
[20] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear system*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.
[21] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, 2003.
[22] G. L. G. Sleijpen and D. R. Fokkema, *BiCGstab(ℓ) for Linear Equations involving Unsymmetric Matrices with Complex Spectrum*, Electron. Trans. Numer. Anal., 1 (1993), pp. 11–32.
[23] K. Stüben and U. Trottenberg, *Multigrid methods: fundamental algorithms, model problem analysis and applications*, in Multigrid Methods, Lecture Notes in Math. 960, W. Hackbusch and U. Trottenberg, eds., Springer, Berlin, Germany, 1982, pp. 1–176.
[24] U. Trottenberg, C. W. Oosterlee, and A. Schüller, *Multigrid*, Academic Press, London, 2001.
[25] E. Turkel, *Numerical difficulties solving time harmonic equations*, in Multiscale Computational Methods in Chemistry and Physics, A. Brandt, et. al., eds., IOS Press, Tokyo, 2001, pp. 319–337.
[26] A. Bayliss, C. I. Goldstein, and E. Turkel, *On accuracy conditions for the numerical computation of waves*, J. Comput. Phys., 59 (1985), pp. 396–404.
[27] H. A. van der Vorst, *Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems*, SIAM J. Sci. Comput., 13 (1992), pp. 631–644.
[28] H. A. Van der Vorst and C. Vuik, *GMRESR: A family of nested GMRES methods*, Numer. Linear Algebra Appl., 1 (1994), pp. 369–386.
[29] R. Wienands, C. W. Oosterlee, and T. Washio, *Fourier analysis of GMRES(m) preconditioned by multigrid*, SIAM J. Sci. Comput., 22 (2000), pp. 582–603.
[30] R. Wienands and C. W. Oosterlee, *On three-grid Fourier analysis for multigrid*, SIAM J. Sci. Comput., 23 (2001), pp. 651–671.
[31] P. M. de Zeeuw, *Matrix-dependent prolongations and restrictions in a blackbox multigrid solver*, J. Comput. Appl. Math., 33 (1990), pp. 1–27.