SIAM J. MATRIX ANAL. APPL.
Vol. 42, No. 2, pp. 475–502

OPTIMAL COMPLEX RELAXATION PARAMETERS IN MULTIGRID FOR COMPLEX-SHIFTED LINEAR SYSTEMS*

L. ROBERT HOCKING† AND CHEN GREIF†

Abstract. We derive optimal complex relaxation parameters minimizing smoothing factors associated with multigrid using red-black successive overrelaxation or damped Jacobi smoothing applied to a class of linear systems arising from discretized linear partial differential equations with a complex shift. Our analysis yields analytical formulas for smoothing factors as a function of the complex relaxation parameter, which may then be efficiently numerically minimized. Our results are applicable to second-order discretizations in arbitrary dimensions, and generalize earlier work of Irad Yavneh on optimal relaxation parameters in the real case. Our analysis is based on deriving a novel connection between the performance of successive overrelaxation as a smoother and as a solver, and is validated by numerical experiments on problems in two and three spatial dimensions, using both vertex- and cell-centered multigrid, with both constant and variable coefficients. In the variable coefficient case we assign different relaxation parameters to different grids points, which our framework allows us to do efficiently.

Key words. multigrid, complex linear systems, local Fourier analysis, damped Jacobi, red-black SOR, relaxation parameter

AMS subject classifications. 65N55, 35J05, 65F10, 65N22

DOI. 10.1137/20M1342161

1. Introduction.

We are interested in the performance of multigrid for numerically solving d-dimensional complex-shifted linear partial differential equations (PDEs)

$$\mathcal{L}u = g, \qquad u, g : \Omega \subset \mathbb{R}^d \to \mathbb{C}, \tag{1.1}$$

where the domain Ω is a d-dimensional rectangle, and $\mathcal{L}$ is of the form

$$\mathcal{L} = L + s(\vec{x}), \tag{1.2}$$

where L is a real linear differential operator and $s : \Omega \to \mathbb{C}$ is a complex shift.

Multigrid [3, 29] is a highly efficient iterative solution technique based on separating the modes of the error into high frequencies and low frequencies, applying a smoother to reduce the amplitudes of the high-frequency modes, and applying a coarse grid correction to deal with the low frequencies. The smoothers that we are considering in this paper are damped Jacobi (ω-JAC) and red-black successive overrelaxation (RB-SOR), both of which involve determining relaxation parameters. This is the focus of our paper, and one of our main conclusions is that allowing these parameters to take on complex values yields significant gains in performance.

To discretize the PDEs that we consider in this paper (see section 2 for a few motivating examples), we use finite differences. Given the general form (1.1), we assume $\Omega = [0, 1]^d$ and discretize uniformly with N interior grid points along each

*Received by the editors June 1, 2020; accepted for publication (in revised form) by J. H. Adler December 16, 2020; published electronically April 7, 2021.
https://doi.org/10.1137/20M1342161
Funding: The work of the authors was partially supported by the Natural Sciences and Engineering Research Council of Canada.
†Department of Computer Science, The University of British Columbia, Vancouver, BC, V6T 1Z4 Canada (rhocking@cs.ubc.ca, greif@cs.ubc.ca).

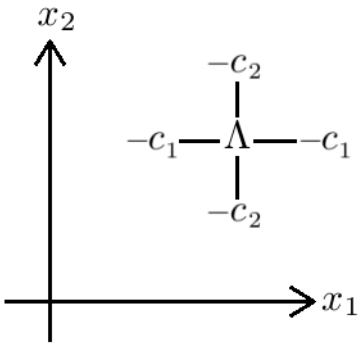


FIG. 1. *Visualization of the computational molecule associated with the discretization of* (1.1) *in two dimensions.*

dimension. We assume Dirichlet boundary conditions with mesh spacing h, and denote our discrete grid by G_h. We assume that second-order centered differences are applied, such that the computational molecule (stencil) of the corresponding discrete system

$$\mathcal{L}_h u_h = g_h, \quad \mathcal{L}_h \in \mathbb{C}^{N^d \times N^d}, \quad u_h, g_h \in \mathbb{C}^{N^d} \tag{1.3}$$

forms a $2d+1$ point star; see Figure 1 for an illustration.

In multiple dimensions it is convenient to label gridpoints by an index vector

$$\vec{I} \in \{1, 2, \ldots, N\}^d. \tag{1.4}$$

To compute the entries of the matrix $\mathcal{L}_h$, one must map these vectors onto single indices. Suppose this mapping is represented by

$$\mathrm{Ord} : \mathbb{Z}^d \to \mathbb{Z}.$$

Then for rows associated with interior gridpoints not immediately next to the boundary, the matrix entries are given by

$$[\mathcal{L}_h]_{\mathrm{Ord}(\vec{I}),\mathrm{Ord}(\vec{J})} = \begin{cases} \Lambda & \text{if } \vec{I} - \vec{J} = \vec{0}, \\ -c_j & \text{if } \vec{I} - \vec{J} = \pm e_j, \\ 0 & \text{otherwise,} \end{cases} \tag{1.5}$$

where $c_j > 0$, $\Lambda \in \mathbb{C}$, and e_j is the jth standard basis vector, $j \in \{1, 2, \ldots, d\}$.

Using the above notation, again for interior grid points the ω-JAC iteration is given by

$$\begin{aligned} \tilde{u}_h^{(n+1)}\left(\mathrm{Ord}(\vec{I})\right) &= \frac{1}{\Lambda}\left\{ g\left(\mathrm{Ord}(\vec{I})\right) \right. \\ &\qquad \left. + \sum_{j=1}^{d} c_j \left[u_h^{(n)}\left(\mathrm{Ord}(\vec{I}+e_j)\right) + u_h^{(n)}\left(\mathrm{Ord}(\vec{I}-e_j)\right)\right] \right\}, \\ u_h^{(n+1)}\left(\mathrm{Ord}(\vec{I})\right) &= \omega \tilde{u}_h^{(n+1)}\left(\mathrm{Ord}(\vec{I})\right) + (1-\omega) u_h^{(n)}\left(\mathrm{Ord}(\vec{I})\right), \end{aligned}$$

which is evaluated independently for each index vector $\vec{I}$ and hence can be computed in parallel. For gridpoints next to the boundary the above formulas are adjusted in a manner that is influenced by the choice of boundary conditions; this is straightforward and details are omitted.

RB-SOR relies on coloring the gridpoints as follows: *black points* are defined as those for which the index vector $\vec{I} = (I_1, \ldots, I_d)$ satisfies $I_1 + I_2 + \cdots + I_d$ odd, say, and *red points* are those for which $I_1 + I_2 + \cdots + I_d$ is of the opposite parity. Once the gridpoints have been colored, RB-SOR runs two sweeps of the above-defined ω-JAC iteration: it runs independently on each of the black points which are then updated, before running on the red points using the updated values at the black points, and updating again.

In this paper, we analyze multigrid with ω-JAC and RB-SOR smoothing applied to systems of the form (1.1) with stencil (1.5). Our analysis is inspired by and extends the beautiful work of Yavneh [34] on optimal relaxation parameters for RB-SOR in the real case; to the best of our knowledge, no such analysis has been performed in the complex case.

Naturally, working in the complex domain is more challenging, in the sense that the optimization problems involved are twice the dimension of the real case and the algebra is more difficult. Despite this, the analysis is still tractable and pays off. In several cases it leads to significantly improved convergence rates; there are even some examples for which there does not exist a real relaxation parameter leading to a convergent method, and yet, satisfactory convergence is obtained for a suitable complex relaxation parameter.

To achieve our goals, we use local Fourier analysis (LFA) to derive analytical smoothing factors for ω-JAC and RB-SOR, and then derive optimal complex relaxation parameters minimizing these smoothing factors. Our analysis is made easier by a connection we derive between the smoothing properties of RB-SOR and those of SOR as a standalone solver. For ω-JAC, we are able to derive analytical formulas for both the smoothing factor and optimal relaxation parameter. For RB-SOR, while we are able to derive an analytical formula for the smoothing factor, the minimization over viable complex relaxation parameters is done numerically. Fortunately, the optimization problem in this case is observed to have a unique local minimum, and hence minimization can be done efficiently at a negligible cost. Moreover, we show that the optimal relaxation parameters we derive depend on only three real parameters for any stencil of the form (1.5) in any dimension $d \geq 2$. Consequently, all possible optimal relaxation parameters may be precomputed in a three-dimensional (3D) lookup table, allowing for great computational efficiency.

We consider vertex-centered and cell-centered multigrid. In the former case, we assume that a direct discretization is applied on all grids, while in the latter case we assume that a Galerkin coarse grid operator is used, based on piecewise-constant prolongation and its transpose. Both setups have the property that the stencil on all coarse grids is of the form (1.5) if it is of that form on the finest grid, which means that the relaxation parameters we derive may be used to apply different relaxation parameters on different grids, as well as different points within a given grid in the case of variable coefficients. The viability of the approach is demonstrated in numerical experiments by way of significant gains in performance.

The remainder of this paper is organized as follows. In section 2 we present relevant examples and provide an overview of related work. In section 3 we apply LFA to our problem. Section 4 contains our analysis of ω-JAC smoothing, h-ellipticity, and their connection. Our analysis of RB-SOR smoothing appears in section 5. Our

results are validated in section 6 by numerical experiments. Finally, in section 7 we draw some conclusions and discuss potential directions for future work.

2. Examples and related work. A few examples that illustrate the breadth and importance of the class of problems (1.1)–(1.2) are in order. We will denote throughout the ubiquitous imaginary unit as

$$i = \sqrt{-1}.$$

An example of primary interest to us is the anisotropic Helmholtz equation [15]

$$-\nabla \cdot (D\nabla u) - k^2(\vec{x})(1 - i\alpha)u = g(\vec{x}), \qquad \vec{x} \in \Omega, \tag{2.1}$$

with appropriate boundary conditions, where $D \in \mathbb{R}^{d\times d}$ is a constant invertible diagonal matrix and $\alpha \geq 0$ represents the damping of the medium. Here $k(\vec{x}) : \Omega \to [0, \infty)$ is a nonnegative real function on Ω and $g(\vec{x})$ is a source term.

Another example is the time evolution of the Schrödinger wave equation [17]

$$i\hbar\Psi_t = -\frac{\hbar^2}{2m}\Delta\Psi + V\Psi.$$

After discretization in time with implicit time-stepping, this equation becomes a series of complex-shifted systems of the form (1.2).

The complex diffusion equation

$$u_t - \alpha\Delta u = f \quad \text{ with } \alpha \in \mathbb{C}$$

is also of interest and relevance; it has appeared, for example, in image processing applications [16].

Our framework is also applicable to the evaluation of functions of matrices via the matrix contour integral

$$f(L) = \int_{\mathcal{C}} f(z)(zI - L)^{-1}dz, \tag{2.2}$$

where L is a real linear differential operator and $\mathcal{C} \subset \mathbb{C}$ is a contour containing the spectrum of L, and on the interior of which f is analytic. Under some mild conditions, discretization leads to a series of complex-shifted linear systems of the form explored in this paper.

Multigrid techniques for complex linear systems of various kinds have been considered in the literature. In [26], the general complex case is analyzed within the context of algebraic multigrid (AMG), and complex intergrid transfer operators are developed. Complex symmetric systems are considered, for example, in [28]. While these AMG solvers consider ω-JAC and RB-SOR smoothing, only real relaxation parameters are considered despite the system being complex. Moreover, optimal values are not derived. While work on optimal complex relaxation parameters for SOR as a solver has appeared in the literature [21], to the best of our knowledge optimal complex relaxation parameters for RB-SOR as a smoother have not been derived.

The isotropic Helmholtz equation

$$-\Delta u - k^2(\vec{x})(1 - i\alpha)u = g(\vec{x}), \qquad \vec{x} \in \Omega, \tag{2.3}$$

is a special case of (2.1) with $D = I$, the identity matrix. It is a difficult problem that has received much attention in the literature, including the development of multigrid

methods. When $0 \leq \alpha \ll 1$, as is the case for some geophysical applications of interest, the problem is extremely challenging because in its discrete form it shifts a symmetric positive definite linear operator (the negative Laplacian) into the indefinite domain. Analysis in [7, 11] revealed the challenge of applying multigrid to this problem.

For larger values of α, multigrid methods are easier to design and they converge relatively rapidly. The papers [9, 10] consider preconditioning (2.3) using a version of the same equation with a larger α, approximately inverted using multigrid. This is referred to as the shifted-Laplacian preconditioning technique. Other fast solvers for the Helmholtz equations have been derived; see, for example, [8, 23, 24].

Recent years have seen significant research focusing on optimizing aspects of the shifted-Laplacian preconditioner, such as the size of the shift [14, 5], the choice of discretization [30], and the ingredients of the multigrid method. For example, in [30] operator-dependent prolongation [6] was compared with prolongation based on principles from algebraic multigrid, while also comparing smoothing based on ω-JAC with incomplete LU smoothing. In [20], the authors consider four-color SOR smoothing in the context of a GPU implementation. However, the above-mentioned papers only consider real relaxation parameters and do not derive optimal values.

3. Local Fourier analysis. There exist a number of theoretical tools for assessing the efficiency of multigrid: LFA, two-grid analysis, and three-grid analysis [29, 31, 32] are among the most popular. Recently, for parabolic or waveform multigrid, semi-algebraic mode analysis (SAMA) has been developed [12], which can handle convection-dominated convection-diffusion equations as well.

LFA and k-grid methods are based on the idea of neglecting boundary conditions by assuming an infinite grid, making certain simplifying assumptions regarding the multigrid iteration matrix, and then considering its effect on Fourier modes. Under these assumptions, for simple smoothers such as ω-JAC or SOR with lexicographic or red-black ordering, Fourier modes are either eigenfunctions of the multigrid iteration matrix or otherwise they form low-dimensional invariant eigenspaces, making the analysis relatively tractable.

To define LFA, we first partition Fourier modes into high-frequency and low-frequency modes, based on the following definition, which assumes a coarse-to-fine grid ratio of $H/h = 2$.

DEFINITION 3.1. *Given a Fourier mode* $\varphi_{h,\vec{\theta}}(\vec{x}) := e^{i\vec{\theta}\cdot\vec{x}/h}$ *with* $\vec{\theta} \in [-\pi, \pi]^d$, *we say that component* θ_j *of* $\vec{\theta}$ *is* high frequency *if* $\theta_j \in [-\pi, \pi]\backslash(-\pi/2, \pi/2)$ *and* low frequency *otherwise. We say that* $\varphi_{h,\vec{\theta}}(\vec{x})$ *is a low-frequency mode if every component of* $\vec{\theta}$ *is low frequency and is a high-frequency mode otherwise. We denote the set of all high-frequency modes by* $\Theta^d := [-\pi, \pi]^d\backslash(-\pi/2, \pi/2)^d$.

LFA is then based on the following two assumptions:

1. The grid G_h is assumed to be infinite, so that boundary conditions may be neglected.
2. The coarse grid correction perfectly eliminates low-frequency modes while leaving high-frequency modes unchanged, that is, denoting the ideal linear coarse grid operator by Q, we have

$$Q\varphi_{h,\vec{\theta}}(\vec{x}) = \begin{cases} \vec{0} & \text{if } \vec{\theta} \text{ is low frequency,} \\ \varphi_{h,\vec{\theta}}(\vec{x}) & \text{if } \vec{\theta} \text{ is high frequency.} \end{cases} \tag{3.1}$$

In contrast to the idealized coarse grid operator of LFA, two-grid analysis and three-

grid analysis are based on the assumption that the multigrid hierarchy contains only two and three grids, respectively, and that the matrix on the coarsest grid is inverted exactly. In practice, the coarse grid correction operator only approximately behaves like the LFA idealized coarse grid operator, and so two-grid and three-grid analysis are more accurate. In [18], it was observed that LFA diverged significantly from two-grid analysis, but the authors proposed a modified idealized coarse grid operator that brought them into close agreement.

The LFA framework may thus be less accurate than two-grid and three-grid analysis, but it serves us well for the problems we consider in this paper, providing a tractable analytical tool that proves to be effective at deriving analytical formulas for smoothing factors, which in turn allows us to efficiently obtain optimal LFA relaxation parameters. We show in our numerical experiments that the relaxation parameters we derive are reliable and significantly speed up multigrid convergence. In section 6.3 we experimentally compare multigrid convergence rates using optimal LFA relaxation parameters with optimal two-grid relaxation parameters.

The idealized multigrid iteration matrix M associated with LFA is given by

$$M = S^{\nu_2} Q S^{\nu_1}, \tag{3.2}$$

where Q is defined by (3.1), S denotes the action of a smoother—for example ω-JAC, Gauss–Seidel, or RB-SOR—and ν_1 and ν_2 denote the number of pre- and postsmoothing iterations, respectively. The smoothing factor is defined as

$$\mu = \rho(M)^{\frac{1}{\nu}}, \tag{3.3}$$

where $\nu = \nu_1 + \nu_2$ is total number of pre- and postsmoothing steps per multigrid iteration and $\rho(M)$ is the spectral radius of M. The smoothing factor gives us a measure of the average error reduction factor per smoothing step.

The eigenfunctions of ω-JAC are precisely the Fourier modes $\varphi_{h,\vec{\theta}}(\vec{x})$. Denoting by $s_\omega(\vec{\theta})$ the eigenvalue of ω-JAC corresponding to Fourier mode $\vec{\theta}$, we have

$$s_\omega(\vec{\theta}) = s_\omega(x(\vec{\theta})) = 1 - \omega\left(1 - \frac{x(\vec{\theta})}{\Lambda}\right), \tag{3.4}$$

where

$$x(\vec{\theta}) = 2\sum_{j=1}^{d} c_j \cos\theta_j. \tag{3.5}$$

We define for convenience

$$c = \sum_{j=1}^{d} c_j, \qquad c_m = \min(c_1, \ldots, c_d). \tag{3.6}$$

Noting that the set Θ^d defined in Definition 3.1 constrains at least one component θ_j of $\vec{\theta}$ to be high frequency, it follows that for this component we must have $\cos\theta_j \le 0$. It in turn follows that $x(\vec{\theta})$ given by (3.5) is maximized by choosing this θ_j to correspond to the c_j that is smallest in magnitude, namely c_m. From this one sees that $x(\vec{\theta})$ maps Θ^d onto the real interval $[-2c, 2(c - c_m)]$, hence

$$\mu(\omega) = \max_{\vec{\theta}\in\Theta^d} |s_\omega(x(\vec{\theta}))| = \max_{x\in[-2c,2(c-c_m)]} |s_\omega(x)|. \tag{3.7}$$

For RB-SOR, the situation is more delicate as the Fourier modes $\varphi_{h,\vec{\theta}}(\vec{x})$ are not eigenfunctions of the RB-SOR iteration matrix. In this case, we define $\tilde{\theta}$ by

$$\tilde{\theta}_j = \theta_j - \operatorname{sign}(\theta_j)\pi, \tag{3.8}$$

where we define $\operatorname{sign}(0) = -1$. It can then be shown (see [33], [29, sect. 4.5.1]) that for every $\vec{\theta} \in [-\pi, \pi]^d$, the two-dimensional (2D) subspace of modes spanned by $\varphi_{h,\vec{\theta}}(\vec{x})$ and $\varphi_{h,\tilde{\theta}}(\vec{x})$ is an invariant subspace of the RB-SOR iteration matrix. Then, defining $\widetilde{Q}(\vec{\theta})$ and $\widetilde{S}(\vec{\theta})$ to be 2×2 matrices corresponding to the action of the coarse grid correction and the RB-SOR iteration matrix, respectively, on this subspace, (3.3) reduces to

$$\mu(\omega) = \max_{\vec{\theta} \in [-\pi,\pi]^d} \rho\left(\widetilde{S}^{\nu_2}(\vec{\theta})\widetilde{Q}(\vec{\theta})\widetilde{S}^{\nu_1}(\vec{\theta})\right)^{\frac{1}{\nu}}. \tag{3.9}$$

The 2×2 matrices $\widetilde{Q}(\vec{\theta})$ and $\widetilde{S}(\vec{\theta})$ are sometimes called the *symbol matrices* associated with the matrices Q and S—see, e.g., [31, p. 116]—and we follow this convention here. The coarse grid correction symbol matrix $\widetilde{Q}(\vec{\theta})$ is given explicitly by

$$\widetilde{Q}(\vec{\theta}) = \begin{bmatrix} q(\vec{\theta}) & 0 \\ 0 & q(\tilde{\theta}) \end{bmatrix},$$

where

$$q(\vec{\theta}) = \begin{cases} 1 & \text{if } \vec{\theta} \text{ is high frequency,} \\ 0 & \text{if } \vec{\theta} \text{ is low frequency.} \end{cases}$$

It follows that

$$\mu(\omega) = \max_{\vec{\theta} \in [-\pi,\pi]^d} \rho\left(\widetilde{S}^{\nu_2}(\vec{\theta})\widetilde{Q}(\vec{\theta})\widetilde{S}^{\nu_1}(\vec{\theta})\right)^{\frac{1}{\nu}} = \max_{\vec{\theta} \in [-\pi,\pi]^d} \rho\left(\widetilde{Q}(\vec{\theta})\widetilde{S}^{\nu}(\vec{\theta})\right)^{\frac{1}{\nu}}. \tag{3.10}$$

Thus, within the framework of LFA the smoothing factor of RB-SOR depends on the total number of smoothing operations $\nu = \nu_1 + \nu_2$ but not on the number of pre- and postsmoothing operations ν_1 and ν_2 separately. The symbol matrix $\widetilde{S}(\vec{\theta})$ can be factored as

$$\widetilde{S}(\vec{\theta}) = \widetilde{S}_R(\vec{\theta})\widetilde{S}_B(\vec{\theta}), \tag{3.11}$$

where $\widetilde{S}_R$ and $\widetilde{S}_B$ denote the symbol matrices of the operators S_R and S_B corresponding to partial relaxation sweeps over the red and black points, respectively. For the class of operators (1.5), we have

$$\widetilde{S}_R = \frac{1}{2}\begin{bmatrix} s_\omega(x(\vec{\theta}))+1 & s_\omega(x(\tilde{\theta}))-1 \\ s_\omega(x(\vec{\theta}))-1 & s_\omega(x(\tilde{\theta}))+1 \end{bmatrix}, \quad \widetilde{S}_B = \frac{1}{2}\begin{bmatrix} s_\omega(x(\vec{\theta}))+1 & -s_\omega(x(\tilde{\theta}))+1 \\ -s_\omega(x(\vec{\theta}))+1 & s_\omega(x(\tilde{\theta}))+1 \end{bmatrix}, \tag{3.12}$$

where $s_\omega(x(\vec{\theta}))$ is given by (3.4) and $s_\omega(x(\tilde{\theta}))$ is given by

$$s_\omega(x(\tilde{\theta})) = 1 - \omega\left(1 + \frac{x(\vec{\theta})}{\Lambda}\right). \tag{3.13}$$

This follows from an argument identical to that found in [34].

4. Damped Jacobi smoothing and *h*-ellipticity. Although the main goal of our analysis in this paper is RB-SOR, we begin with the simpler case of ω-JAC. Our main objectives are to establish necessary conditions for $\mu(\omega) < 1$, derive the optimal complex value of ω, and gain some insight into the measure of *h-ellipticity* and its connection to ω-JAC.

4.1. *h*-ellipticity. Within the LFA framework, the eigenfunctions of $\mathcal{L}_h$ are the Fourier modes $\varphi_{h,\vec{\theta}}(\vec{x})$. Denoting by $\widetilde{\mathcal{L}}_h(\vec{\theta})$ the eigenvalue of $\mathcal{L}_h$ associated with $\varphi_{h,\vec{\theta}}(\vec{x})$, we may define the *h-ellipticity* of $\mathcal{L}_h$ [29, sect. 4.7] by

$$E_h(\mathcal{L}_h) = \left| \frac{\widetilde{\mathcal{L}}_h(\vec{\theta}^m)}{\widetilde{\mathcal{L}}_h(\vec{\theta}^M)} \right|, \tag{4.1}$$

where

$$\vec{\theta}^m = \operatorname{argmin}_{\vec{\theta}\in\Theta^d} |\widetilde{\mathcal{L}}_h(\vec{\theta})| \quad \text{and} \quad \vec{\theta}^M = \operatorname{argmax}_{\vec{\theta}\in[-\pi,\pi]^d} |\widetilde{\mathcal{L}}_h(\vec{\theta})|. \tag{4.2}$$

We have by construction $E_h(\mathcal{L}_h) \le 1$. If we additionally have

$$E_h(\mathcal{L}_h) \ge \text{const} > 0,$$

then—insofar as LFA is accurate—the existence of smoothers leading to h-independent convergence of multigrid is guaranteed. However, as $E_h(\mathcal{L}_h) \to 0$ the convergence of multigrid is expected to deteriorate, and in particular, $E_h(\mathcal{L}_h) = 0$ implies that no point-smoother (as opposed to line- or block-smoothers) can yield a convergent method [29, p. 126]. In our setting, we have

$$\widetilde{\mathcal{L}}_h(\vec{\theta}) = \Lambda - x(\vec{\theta}),$$

where $x(\vec{\theta})$ is given by (3.5). As we have already noted, $x(\vec{\theta})$ maps Θ^d onto the real interval $[-2c, 2(c - c_m)]$. From this, one readily computes

$$\widetilde{\mathcal{L}}_h(\vec{\theta}^m) = \begin{cases} \Lambda - 2(c - c_m) & \text{if } \operatorname{Re}(\Lambda) \ge 2(c - c_m), \\ \operatorname{Im}(\Lambda) & \text{if } -2c \le \operatorname{Re}(\Lambda) < 2(c - c_m), \\ \Lambda + 2c & \text{otherwise,} \end{cases} \tag{4.3}$$

while

$$\widetilde{\mathcal{L}}_h(\vec{\theta}^M) = \begin{cases} \Lambda + 2c & \text{if } \operatorname{Re}(\Lambda) > 0, \\ \Lambda - 2c & \text{otherwise.} \end{cases} \tag{4.4}$$

It then trivially follows that

$$E_h(\mathcal{L}_h) \ge \frac{|\operatorname{Im}(\Lambda)|}{|\Lambda| + 2c} > 0 \text{ if } \operatorname{Im}(\Lambda) \ne 0.$$

This simple observation tells us that for the class of problems under consideration adding an imaginary shift to the diagonal of $\mathcal{L}_h$ tends to improve h-ellipticity. Moreover, *any* nonzero shift guarantees $E_h(\mathcal{L}_h) > 0$, even for matrices that are not diagonally dominant. By contrast, if Λ is real and positive, we have $E_h(\mathcal{L}_h) = 0$ for $\Lambda \le 2(c - c_m)$. This could partially explain the success of the shifted-Laplacian family of preconditioners.

The following proposition, the proof of which is omitted, shows that for $|\Lambda|$ fixed, h-ellipticity is maximized when Λ is purely imaginary.

PROPOSITION 4.1. *Let $\Lambda = re^{i\phi}$ with $r > 0$ and $-\pi \le \phi \le \pi$. Then, for fixed r, $E_h(\mathcal{L}_h)$ is maximized at $\phi = \pm\frac{\pi}{2}$.*

4.2. ω-JAC smoothing. From (3.7) and (3.4) we have

$$\mu(\omega) = \max_{x\in[-2c,2(c-c_m)]} |s_\omega(x)| \geq |s_\omega(0)| = |1-\omega|.$$

Evidently, the smoothing factor $\mu(\omega)$ of multigrid with ω-JAC smoothing satisfies $\mu(\omega) < 1$ only if $|1-\omega| < 1$.

THEOREM 4.2. *Let $\mathcal{L}_h$ be a linear operator of the form* (1.5) *with $\Lambda \in \mathbb{C}$, and let $c, c_m \in \mathbb{R}$ be defined as in* (3.6). *Let $\mu(\omega)$ defined by* (3.7) *denote the smoothing factor of multigrid with ω-JAC smoothing applied to this system. Let*

$$\omega_{opt} = \operatorname{argmin}_{\omega\in\mathbb{C}} \mu(\omega) \quad \textit{and} \quad \mu_{opt} = \mu(\omega_{opt})$$

denote the optimal complex relaxation parameter minimizing $\mu(\omega)$ and corresponding optimal smoothing factor. Define

$$b_1 = \left(1 - 2\frac{c-c_m}{\Lambda}\right), \qquad b_2 = \left(1 + \frac{2c}{\Lambda}\right), \tag{4.5}$$

and assume $b_1, b_2 \neq 0$. Then there is a unique optimal complex relaxation parameter given by

$$\omega_{opt} = \frac{\frac{|b_1|}{b_1} + \frac{|b_2|}{b_2}}{|b_1| + |b_2|}. \tag{4.6}$$

The corresponding optimal complex smoothing factor is

$$\mu_{opt} = \frac{|b_1 - b_2|}{|b_1| + |b_2|}. \tag{4.7}$$

Proof. From (3.7) and (3.4) it follows that

$$\mu(\omega) = \max_{x\in[-2c,2(c-c_m)]} \left|1 - \omega\left(1 - \frac{x}{\Lambda}\right)\right|.$$

For fixed ω, $\{\omega\left(1-\frac{x}{\Lambda}\right)\}$ is a line segment in $\mathbb{C}$ with endpoints ωb_1 and ωb_2. The furthest point from $z = 1$ on this line segment must be one of the endpoints. Hence

$$\mu(\omega) = \max(\mu_1(\omega), \mu_2(\omega)) \quad \text{where } \mu_j(\omega) = |1 - \omega b_j|, \quad j = 1, 2. \tag{4.8}$$

Since $\mu(\omega)$ is the maximum of the moduli of two nonconstant holomorphic functions, the maximum principle [13, p. 88] implies that the optimal complex ω_{opt} must satisfy

$$\mu_1(\omega) = \mu_2(\omega). \tag{4.9}$$

For if the optimal value $\omega = \omega_{opt}$ were to obey $\mu_1(\omega_{opt}) > \mu_2(\omega_{opt})$ (say), the maximum principle and continuity imply that we could find another ω^*_{opt} in a neighborhood of ω_{opt} obeying $\mu_1(\omega_{opt}) > \mu_1(\omega^*_{opt}) > \mu_2(\omega^*_{opt})$, violating our assumption of the optimality of ω_{opt} (the same argument applies if the roles of μ_1 and μ_2 are reversed).

Next, we rewrite (4.9) as $f(\omega) = e^{i\theta}$, where $f(\omega)$ is the Möbius transform

$$f(\omega) = \frac{1 - b_1\omega}{1 - b_2\omega}$$

and $\theta \in [0, 2\pi)$. It follows from elementary properties of Möbius transforms [13, p. 63] that f^{-1} is given by

$$f^{-1}(\omega) = \frac{\omega - 1}{b_2\omega - b_1}.$$

Hence, any ω obeying (4.9) must be of the form

$$\omega(\theta) = f^{-1}(e^{i\theta}) = \frac{e^{i\theta} - 1}{b_2 e^{i\theta} - b_1} = \frac{1}{b_2}\left[1 + \frac{b_1 - b_2}{b_2 e^{i\theta} - b_1}\right].$$

Substitution into (4.8) gives

$$\mu(\omega(\theta)) = \mu_2(\omega(\theta)) = \left|\frac{b_1 - b_2}{b_2 e^{i\theta} - b_1}\right| \geq \frac{|b_1 - b_2|}{|b_1| + |b_2|}, \text{ with equality iff } e^{i\theta} = -\frac{b_1}{|b_1|}\frac{|b_2|}{b_2}.$$

It follows that μ_{opt} is given as stated in (4.7) in the statement of the theorem. It similarly follows that ω_{opt} is given by

$$\omega_{opt} = f^{-1}\left(-\frac{b_1}{|b_1|}\frac{|b_2|}{b_2}\right) = \frac{-\frac{b_1}{|b_1|}\frac{|b_2|}{b_2} - 1}{-b_2\frac{b_1}{|b_1|}\frac{|b_2|}{b_2} - b_1} = \frac{\frac{|b_2|}{b_2} + \frac{|b_1|}{b_1}}{|b_2| + |b_1|},$$

giving us (4.6) as stated in the theorem. ∎

It is also possible to analytically derive optimal *real* relaxation parameters and smoothing factors for our problem of interest. However, because the tools of complex analysis—notably the maximum principle—are no longer available, the proof is less clean and involves many cases. For the sake of brevity, it has been omitted.

In the cases $b_1 = 0$ or $b_2 = 0$ one may prove that $\mu(\omega) \geq 1$ for all $\omega \in \mathbb{C}$—optimal relaxation parameters are therefore not of interest.

Equipped with the results of our smoothing analysis, we now show how a known connection between damped Jacobi smoothing and h-ellipticity in the real symmetric case generalizes to our current setting.

For real symmetric matrices $\mathcal{L}_h$ with diagonal $D = a \cdot I$ and $\widetilde{\mathcal{L}}_h(\vec{\theta}^m)$, $\widetilde{\mathcal{L}}_h(\vec{\theta}^M) > 0$, it is known (see, e.g., [29, pp. 126–127]) that the smoothing factor of multigrid with damped Jacobi smoothing using relaxation parameter

$$\omega^* = \frac{2a}{\widetilde{\mathcal{L}}_h(\vec{\theta}^m) + \widetilde{\mathcal{L}}_h(\vec{\theta}^M)} \tag{4.10}$$

gives rise to the smoothing factor

$$\mu^* = \frac{1 - E_h(\mathcal{L}_h)}{1 + E_h(\mathcal{L}_h)} = \frac{\widetilde{\mathcal{L}}_h(\vec{\theta}^M) - \widetilde{\mathcal{L}}_h(\vec{\theta}^m)}{\widetilde{\mathcal{L}}_h(\vec{\theta}^M) + \widetilde{\mathcal{L}}_h(\vec{\theta}^m)}, \tag{4.11}$$

where $\widetilde{\mathcal{L}}_h(\vec{\theta}^m)$ and $\widetilde{\mathcal{L}}_h(\vec{\theta}^M)$ are given by (4.2). In our setting, it follows from (4.3), (4.4), and (4.5) that

$$\Lambda \cdot b_1 = \widetilde{\mathcal{L}}_h(\vec{\theta}^m) \quad \text{and} \quad \Lambda \cdot b_2 = \widetilde{\mathcal{L}}_h(\vec{\theta}^M) \qquad \text{if } \mathrm{Re}(\Lambda) > 2(c - c_m). \tag{4.12}$$

Hence, provided $\mathrm{Re}(\Lambda) > 2(c - c_m)$, the optimal smoothing factor (4.7) derived in Theorem 4.2 may be obtained from the rightmost equality of (4.11) by taking the absolute value of the numerator, the absolute value of each term in the denominator, substituting in (4.12), and cancelling common factors.

Also for $\mathrm{Re}(\Lambda) > 2(c - c_m)$, the optimal relaxation parameter (4.6) from Theorem 4.2 may be seen as a generalization of (4.10) by dividing the numerator and denominator of the latter by a, noting that the equality

$$\frac{|\widetilde{\mathcal{L}}_h(\vec{\theta}^m)|}{\widetilde{\mathcal{L}}_h(\vec{\theta}^m)} + \frac{|\widetilde{\mathcal{L}}_h(\vec{\theta}^M)|}{\widetilde{\mathcal{L}}_h(\vec{\theta}^M)} = 2$$

is valid for real symmetric matrices $\mathcal{L}_h$ with $\widetilde{\mathcal{L}}_h(\vec{\theta}^m)$, $\widetilde{\mathcal{L}}_h(\vec{\theta}^M) > 0$, substituting it into the numerator of (4.10), and then substituting (4.12) for $\mathcal{L}_h(\vec{\theta}^m)$ and $\mathcal{L}_h(\vec{\theta}^M)$.

5. Red-black SOR smoothing. We now move to consider the case of primary interest for us, namely RB-SOR smoothing. Expanding $\widetilde{S}(\vec{\theta}) = \widetilde{S}_R(\vec{\theta})\widetilde{S}_B(\vec{\theta})$ from (3.11) and (3.12) yields

$$\widetilde{S}(\vec{\theta}) = \begin{bmatrix} s_{11}(\vec{\theta}) & s_{12}(\vec{\theta}) \\ s_{21}(\vec{\theta}) & s_{22}(\vec{\theta}) \end{bmatrix},$$

where

$$\begin{aligned}
s_{11}(\vec{\theta}) &= \frac{1}{4}\left[(s_\omega(x(\vec{\theta})) + 1)^2 + (s_\omega(x(\tilde{\theta})) - 1)(1 - s_\omega(x(\vec{\theta})))\right], \\
s_{12}(\vec{\theta}) &= \frac{1}{4}\left[s_\omega^2(x(\tilde{\theta})) - 1 + (s_\omega(x(\vec{\theta})) + 1)(1 - s_\omega(x(\tilde{\theta})))\right], \\
s_{21}(\vec{\theta}) &= \frac{1}{4}\left[s_\omega^2(x(\vec{\theta})) - 1 + (s_\omega(x(\tilde{\theta})) + 1)(1 - s_\omega(x(\vec{\theta})))\right], \\
s_{22}(\vec{\theta}) &= \frac{1}{4}\left[(s_\omega(x(\tilde{\theta})) + 1)^2 + (s_\omega(x(\tilde{\theta})) - 1)(1 - s_\omega(x(\vec{\theta})))\right],
\end{aligned}$$

and $s_\omega(x(\vec{\theta}))$ and $s_\omega(x(\tilde{\theta}))$ are given by (3.4) and (3.13). The eigenvalues of $\widetilde{S}(\vec{\theta})$ satisfy the quadratic equation

$$\lambda^2 - \mathrm{tr}(\widetilde{S}(\vec{\theta}))\lambda + \det(\widetilde{S}(\vec{\theta})) = 0. \tag{5.1}$$

Thus, we compute the determinant and trace of $\widetilde{S}(\vec{\theta})$, which will be useful for us throughout our subsequent analysis. A short calculation yields

$$\det(\widetilde{S}(\vec{\theta})) = (1 - \omega)^2, \tag{5.2a}$$

$$\mathrm{tr}(\widetilde{S}(\vec{\theta})) = 2(1 - \omega) + \omega^2 \frac{x(\vec{\theta})^2}{\Lambda^2}. \tag{5.2b}$$

From this we immediately obtain the following restriction on the domain of ω, which is identical to the classical result [19] for SOR as a solver.

OBSERVATION 5.1. *$\mu(\omega) \geq 1$ unless $|\omega - 1| < 1$.*

Proof. If $\vec{\theta}$ and $\tilde{\theta}$ are both high frequency (as is the case, for example, with $\theta_j = \frac{\pi}{2}$ for all $j = 1, \ldots, d$), then $\widetilde{Q}(\vec{\theta}) = I$ and from (3.10) we have

$$\rho(\widetilde{Q}(\vec{\theta})\widetilde{S}(\vec{\theta})^\nu)^{\frac{1}{\nu}} = \rho(\widetilde{S}(\vec{\theta})),$$

where ν denotes the number of smoothing sweeps per multigrid iteration. Next, we have

$$|1 - \omega|^2 = |\det(\widetilde{S}(\vec{\theta}))| = |\lambda^+(\vec{\theta}) \cdot \lambda^-(\vec{\theta})| \leq \max(|\lambda^+(\vec{\theta})|, |\lambda^-(\vec{\theta})|)^2,$$

where $\lambda^+(\vec{\theta})$ and $\lambda^-(\vec{\theta})$ are the eigenvalues of $\widetilde{S}(\vec{\theta})$. Hence

$$|1 - \omega| \leq \max(|\lambda^+(\vec{\theta})|, |\lambda^-(\vec{\theta})|) = \rho(\widetilde{S}(\vec{\theta})) \leq \mu(\omega). \qquad \square$$

It is straightforward to show that at least one of $\vec{\theta}$, $\tilde{\theta}$ must be a high-frequency mode. Following [34], we follow the convention that $\tilde{\theta}$ is always high frequency, and we define the smoothing factors for $\vec{\theta}$ restricted to the high-frequency and low-frequency modes, respectively, by

$$\mu_H = \max_{\vec{\theta}\ \text{h.f.}} \rho(\widetilde{Q}(\vec{\theta})\widetilde{S}(\vec{\theta})^\nu)^{\frac{1}{\nu}}, \quad \mu_L = \max_{\vec{\theta}\ \text{l.f.}} \rho(\widetilde{Q}(\vec{\theta})\widetilde{S}(\vec{\theta})^\nu)^{\frac{1}{\nu}} \tag{5.3}$$

so that the overall smoothing factor is given by

$$\mu(\omega) = \max(\mu_H(\omega), \mu_L(\omega)). \tag{5.4}$$

Like [34], we go on to analyze the high-frequency and low-frequency cases separately.

5.1. The case of $\vec{\theta}$ high frequency. Our analysis of the high-frequency case is based on deriving a connection between the performance of RB-SOR applied to the matrix $\mathcal{L}_h$ defined by (1.3) when it is used as a *smoother* in conjunction with multigrid, with its performance when applied as a direct *solver* to a associated matrix $\widehat{\mathcal{L}}_h$. Because the case of a complex relaxation parameter in the context of a solver is already well understood [21], we can leverage this connection and apply known results about the solver case to the smoother case.

Given a linear operator $\mathcal{L}_h$ of the form (1.5) acting on the grid G_h with ordering Ord, let us define the associated operator $\widehat{\mathcal{L}}_h$ acting on the same grid G_h by

$$[\widehat{\mathcal{L}}_h]_{\text{Ord}(\vec{I}),\text{Ord}(\vec{J})} = \begin{cases} \Lambda & \text{if } \vec{I} - \vec{J} = \vec{0}, \\ -\left(c_j - \frac{c_m}{d}\right) & \text{if } \vec{I} - \vec{J} = \pm e_j, \\ 0 & \text{otherwise,} \end{cases} \tag{5.5}$$

where c_m is given by (3.6). The boundary conditions of $\widehat{\mathcal{L}}_h$ are taken to be the same as those of $\mathcal{L}_h$, and the index vectors $\vec{I}, \vec{J}$ are restricted to the same set (1.4). We can think of $\widehat{\mathcal{L}}_h$ as a version of $\mathcal{L}_h$ in which the diagonal dominance has been improved by decreasing the magnitude of all off-diagonal elements.

Let us denote by $\widehat{\mathcal{M}}_h(\omega)$ the iteration matrix of SOR with relaxation parameter ω applied to (5.5). Let us denote by $\widehat{\rho}(\omega)$ the spectral radius of $\widehat{\mathcal{M}}_h(\omega)$ in the limit as $h \to 0$, that is

$$\widehat{\rho}(\omega) = \lim_{h \to 0} \rho(\widehat{\mathcal{M}}_h(\omega)). \tag{5.6}$$

The following proposition establishes the promised link between $\mu_H(\omega)$ and $\widehat{\rho}(\omega)$.

PROPOSITION 5.2. *Let $\mathcal{L}_h$ be any matrix of the form* (1.5)*, and let $\widehat{\mathcal{L}}_h$ be the associated matrix* (5.5)*. Let $\mu_H(\omega)$ and $\widehat{\rho}(\omega)$ be defined by* (5.3) *and* (5.6)*, respectively. Then we have the equality*

$$\mu_H(\omega) = \widehat{\rho}(\omega).$$

Proof. Our basic approach is to show that both $\mu_H(\omega)$ and $\widehat{\rho}(\omega)$ can be expressed in the same way in terms of the roots of the same quadratic polynomial. We begin with $\mu_H(\omega)$. Following [34], we note that as $\widetilde{Q}(\vec{\theta}) = I$ from (5.3) we have

$$\mu_H = \max_{\vec{\theta}\ \text{h.f.}} \rho(\widetilde{Q}(\vec{\theta})\widetilde{S}(\vec{\theta})^\nu)^{\frac{1}{\nu}} = \max_{\vec{\theta}\ \text{h.f.}} \rho(\widetilde{S}(\vec{\theta})^\nu)^{\frac{1}{\nu}} = \max_{\vec{\theta}\ \text{h.f.}} \rho(\widetilde{S}(\vec{\theta}))$$

independent of the number of smoothing sweeps per multigrid iteration, ν. Next, note that although we have assumed that $\vec{\theta}$ is high frequency (and hence must have

at least one high-frequency component), it also must have at least one low-frequency component. Otherwise, $\tilde{\theta}$ would be low frequency, violating our convention. Thus we must have $\cos\theta_j \geq 0$ for at least one component θ_j of $\vec{\theta}$, and similarly we must have $\cos\theta_{j'} \leq 0$ for some other (possibly distinct) component $\theta_{j'}$ of $\vec{\theta}$. Setting $\cos\theta_j = 0$ for the c_j corresponding to c_m satisfies both conditions at once—the maximum and minimum values of $x(\vec{\theta})$ are obtained by setting all remaining cosines equal to $+1$ and -1, respectively, yielding $-2(c-c_m) \leq x \leq 2(c-c_m)$ (with c, c_m defined as in (3.6)). Hence

$$\mu_H(\omega) = \max_{x\in[-2(c-c_m),2(c-c_m)]} \left\{\max(|\lambda^+(x,\omega)|, |\lambda^-(x,\omega)|)\right\}$$

where $\lambda^{\pm}(x,\omega)$ are the two roots of the quadratic equation (5.1). Substituting the expressions (5.2a) and (5.2b) for $\mathrm{tr}(S)$ and $\det(S)$ gives

$$\lambda^2 - \left(2(1-\omega) + \omega^2\frac{x^2}{\Lambda^2}\right)\lambda + (1-\omega)^2 = 0.$$

Next, we consider $\widehat{\rho}(\omega)$. It will be useful for us to consider the diagonal of the matrix $\widehat{\mathcal{L}}_h$, which we denote by $\widehat{\mathcal{D}}_h$. Applying a red-black ordering to $\widehat{\mathcal{L}}_h$ puts it in the block 2×2 form

$$\widehat{\mathcal{L}}_h = \begin{bmatrix} D_{rr} & L_{rb} \\ L_{br} & D_{bb} \end{bmatrix},$$

where D_{rr} and D_{bb} denote the restriction of $\widehat{\mathcal{D}}_h$ to red and black points, respectively. Since D_{rr} and D_{bb} are diagonal matrices, this tells us that $\widehat{\mathcal{L}}_h$ is a two-cyclic matrix that satisfies Young's property A as defined in [36, p. 93]. It follows [36, Theorem 2.3] that the eigenvalues $\{\lambda_h\}$ of $\widehat{\mathcal{M}}_h(\omega)$ are related to the eigenvalues $\{\delta_h\}$ of the Jacobi iteration matrix $\widehat{\mathcal{J}}_h = I - (\widehat{\mathcal{D}}_h)^{-1}\widehat{\mathcal{L}}_h$ by Young's relation

$$(\lambda_h + \omega - 1)^2 = \omega^2\delta_h^2\lambda_h. \tag{5.7}$$

At the same time, the eigenvalues $\{\delta_h\}$ of the Jacobi iteration matrix $\widehat{\mathcal{J}}_h$ are given by

$$\{\delta_h\} = \frac{\mathcal{I}_h}{\Lambda}, \quad \text{where } \mathcal{I}_h = \left\{ 2\sum_{j=1}^{d}\left(c_j - \frac{c_m}{d}\right)\cos\theta_j : \theta_j \in \left\{\frac{2\pi\ell}{N+1}\right\}_{\ell=1}^{N}\right\},$$

and in particular, the largest magnitude eigenvalue of the Jacobi iteration matrix is given by $\pm\delta_h^*$, where

$$\delta_h^* = 2\frac{c-c_m}{\Lambda}\cos(2\pi h), \tag{5.8}$$

a fact which will be useful in our subsequent analysis.

To see this, note that the Jacobi matrix can be factored as

$$\widehat{\mathcal{J}}_h = \frac{1}{\Lambda}\bigoplus_{j=1}^{d}\left(c_j - \frac{c_m}{d}\right)T_N[1,0,1],$$

where $T_N[1,0,1]$ is the tridiagonal Toeplitz matrix of size $N\times N$ with 1's on its 1st and -1st diagonals and zeros elsewhere. The eigenvalues of $T_N[1,0,1]$ are known analytically; see, for example, [22, Theorem 2.2]. The spectrum of $\widehat{\mathcal{J}}_h$ may then be found by applying well-known properties of Kronecker sums.

Since the roots of polynomials depend continuously on their coefficients [25], and since $\mathcal{I}_h$ is dense in $[-2(c-c_m), 2(c-c_m)]$ as $h \to 0$, we have

$$\begin{aligned}\widehat{\rho}(\omega) &= \lim_{h\to 0}\left\{\max_{x\in\mathcal{I}_h}\left\{\max\left(|\lambda^+(x,\omega)|, |\lambda^-(x,\omega)|\right)\right\}\right\} \\ &= \max_{x\in[-2(c-c_m),2(c-c_m)]}\left\{\max(|\lambda^+(x,\omega)|, |\lambda^-(x,\omega)|)\right\},\end{aligned}$$

where $\lambda^-(x,\omega)$ and $\lambda^+(x,\omega)$ are the roots of the quadratic in λ,

$$(\lambda+\omega-1)^2 = \omega^2\frac{x^2}{\Lambda^2}\lambda.$$

Expanding this quadratic shows it is the same as the one we obtained for $\mu_H(\omega)$. ∎

Having established the above connection, we can now import known results about the behavior of SOR with complex ω as a solver in order to analyze the present situation with smoothing.

THEOREM 5.3. *Let $\mu_H(\omega)$ be defined by* (5.3)*. Then we have*

$$\mu_H(\omega) = \max\left|\frac{\omega(c-c_m)}{\Lambda} \pm \sqrt{(1-\omega)+\frac{\omega^2(c-c_m)^2}{\Lambda^2}}\right|^2, \tag{5.9}$$

where the maximum is taken over the positive and negative values of the ± sign. Moreover, the choice

$$\omega_{ub} = \frac{2}{1+\sqrt{1-4\frac{(c-c_m)^2}{\Lambda^2}}}$$

minimizes $\mu_H(\omega)$, and we have

$$\mu_H(\omega_{ub}) = |1-\omega_{ub}| = \left|\frac{1-\sqrt{1-4\frac{(c-c_m)^2}{\Lambda^2}}}{1+\sqrt{1-4\frac{(c-c_m)^2}{\Lambda^2}}}\right|.$$

Proof. By Proposition 5.2, it suffices to prove the same statements for $\widehat{\rho}(\omega)$. However, this analysis has already been done in [21]. To prove (5.9), the key result is [21, Lemma 4.1], which requires that $\widehat{\mathcal{L}}_h$ obey property A and the spectrum of the Jacobi iteration matrix $\widehat{\mathcal{J}}_h$ obey

$$\sigma(\widehat{\mathcal{J}}_h) \subseteq [-\zeta_h, \zeta_h] \subset \mathbb{C} \quad \text{and} \quad \pm\zeta_h \in \sigma(\widehat{\mathcal{J}}_h) \text{ for some } \zeta_h \in \mathbb{C}, \tag{5.10}$$

where $[-\zeta_h, \zeta_h]$ denotes the line segment joining $\pm\zeta_h \in \mathbb{C}$. If this requirement is satisfied, the largest-magnitude eigenvalue of the SOR iteration matrix may be found, independently of ω, by substituting the value $\delta_h = \pm\zeta_h$ into Young's relation (5.7) and solving for λ_h. In our situation $\widehat{\mathcal{L}}_h$ indeed satisfies property A, and (5.10) is obeyed with $\zeta_h = \delta_h^*$ satisfying (5.8). Substituting δ_h^* into (5.7) and square-rooting both sides yields

$$\lambda_h^* \pm 2\omega\frac{(c-c_m)}{\Lambda}\cos(2\pi h)(\lambda_h^*)^{\frac{1}{2}} + (\omega-1) = 0,$$

where λ_h^* denotes the largest-magnitude eigenvalue of the SOR iteration matrix. Using continuity to take h to 0 yields

$$\mu_H(\omega) = \lim_{h\to 0} \max \left| \frac{\omega(c-c_m)\cos(2\pi h)}{\Lambda} \pm \sqrt{(1-\omega) + \frac{\omega^2(c-c_m)^2\cos^2(2\pi h)}{\Lambda^2}} \right|^2,$$

which is equivalent to the claimed expression.

The expressions for ω_{ub} and $\mu_H(\omega_{ub})$ then follow from [21, Theorem 4.1], once again using continuity to take h to 0. ∎

The relaxation parameter ω_{ub} reduces, in the case of $\Lambda = 2c \in \mathbb{R}$, where c is given by (3.6), to the parameter of the same name found by Yavneh in [34, Theorem 2.1] when analyzing RB-SOR in the real case. In that case, ω_{ub} is an upper bound on the optimal real relaxation parameter $\omega_{opt} \in (1, \omega_{ub})$. In our setting, since $\mathbb{C}$ has no natural ordering, it no longer makes sense to talk about upper bounds; however, we have stuck with this notation for the sake of consistency with [34].

The expression for $\mu_H(\omega_{ub})$ given above also reduces, under the same assumptions, to the lower bound for the optimal smoothing factor noted in [34, Corollary 2.2]—in our case it remains a lower bound. Yavneh also noted in [34] similarities between the expressions he derived and optimal relaxation parameters and convergence rates of SOR as a solver, but stopped short of deriving an explicit connection, as we have done in Proposition 5.2.

However, while in the real case [34] it was found that ω_{ub} gives a good approximation of the true optimal relaxation parameter ω_{opt}, we will see that in the complex case, this is no longer true in general.

5.2. The case of $\vec{\theta}$ low frequency. In this section we analyze $\mu_L(\omega)$ given by (5.3). If $\vec{\theta}$ is low frequency, then we have

$$\widetilde{Q}(\vec{\theta}) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

It follows that

$$\mu_L(\omega) = \max_{\vec{\theta}\ \text{l.f.}} \rho(\widetilde{Q}(\vec{\theta})\widetilde{S}(\vec{\theta})^\nu)^{\frac{1}{\nu}} = \max_{\vec{\theta}\ \text{l.f.}} |s_{22}^{(\nu)}(\vec{\theta})|,$$

where $s_{22}^{(\nu)}(\vec{\theta})$ denotes the (2,2) component of the matrix $\widetilde{S}^\nu(\vec{\theta})$. We restrict our attention to $\nu = 1, 2$, but our approach generalizes to higher values of ν.

First, consider $\nu = 1$. In this case we obtain after some algebra

$$\mu_L(\omega) = \max_{\vec{\theta}\ \text{l.f.}} \left| 1 + \omega \left(1 + \frac{x(\vec{\theta})}{\Lambda}\right)\left(\omega\frac{x(\vec{\theta})}{2\Lambda} - 1\right)\right|,$$

where $\vec{\theta}$ being low frequency means that $\vec{\theta} \in [-\frac{\pi}{2}, \frac{\pi}{2}]^d$. Since every cosine in $x(\vec{\theta})$ given by (3.5) must now be nonnegative, we have $x([-\frac{\pi}{2}, \frac{\pi}{2}]^d) = [0, 2c]$, so that

$$\mu_L(\omega) = \max_{x\in[0,2c]} \left|a_0 + a_1 x + a_2 x^2\right|, \tag{5.11}$$

where c is defined in (3.6) and

$$a_0 = 1 - \omega, \qquad a_1 = \frac{\omega}{\Lambda}\left(\frac{\omega}{2} - 1\right), \qquad a_2 = \frac{\omega^2}{2\Lambda^2}. \tag{5.12}$$

Now, for $\nu = 2$ we obtain

$$\mu_L(\omega) = \max_{x\in[0,2c]} \sqrt{|a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4|}, \tag{5.13}$$

where

$$\begin{aligned} a_0 &= \omega^2 - 2\omega + 1, \qquad a_1 = -\frac{\omega}{\Lambda}\left(\omega^2 - 3\omega + 2\right), \\ a_2 &= \frac{2\omega^2}{\Lambda^2}\left(1-\omega\right), \quad a_3 = \frac{\omega^3}{2\Lambda^3}\left(\omega - 2\right), \quad a_4 = \frac{\omega^4}{2\Lambda^4}. \end{aligned} \tag{5.14}$$

We can now derive an expression for $\mu_L(\omega)$. The proof of the theorem that follows is straightforward and relies on the simple observation that the condition for critical points of the magnitude of a complex-valued function f of a real variable x is $\mathrm{Re}\left(\mathrm{f}\frac{\mathrm{d}\bar{\mathrm{f}}}{\mathrm{dx}}\right) = 0$, which trivially follows from the identity $|f|^2 = f\bar{f}$.

THEOREM 5.4. *Let $\mu_L(\omega)$ be defined by* (5.3)*, and assume $\nu \in \{1, 2\}$. Let $p_2(x)$ and $p_4(x)$ denote the quadratic and quartic functions from* (5.11) *and* (5.13)*, with coefficients given by* (5.12) *and* (5.14)*. Then we have*

$$\mu_L(\omega) = \begin{cases} \max_{x\in X_1} |p_2(x)| & \text{if } \nu = 1, \\ \max_{x\in X_2} \sqrt{|p_4(x)|} & \text{if } \nu = 2, \end{cases} \tag{5.15}$$

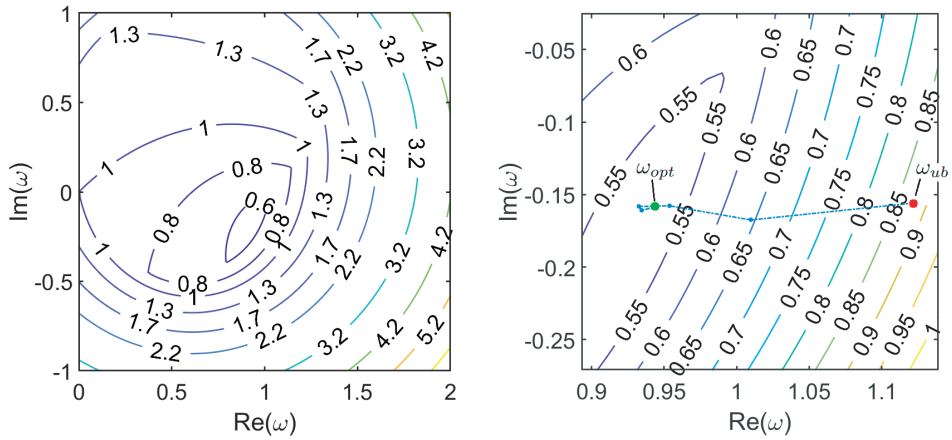
where X_1 denotes the real roots of the cubic

$$0 = \mathrm{Re}(a_0\bar{a}_1) + (2\,\mathrm{Re}(a_0\bar{a}_2) + |a_1|^2)x + 3\,\mathrm{Re}(a_1\bar{a}_2)x^2 + 2|a_2|^2x^3$$

falling in the interval $[0, 2c]$ together with the endpoints of said interval, and X_2 denotes the same for the degree-seven polynomial

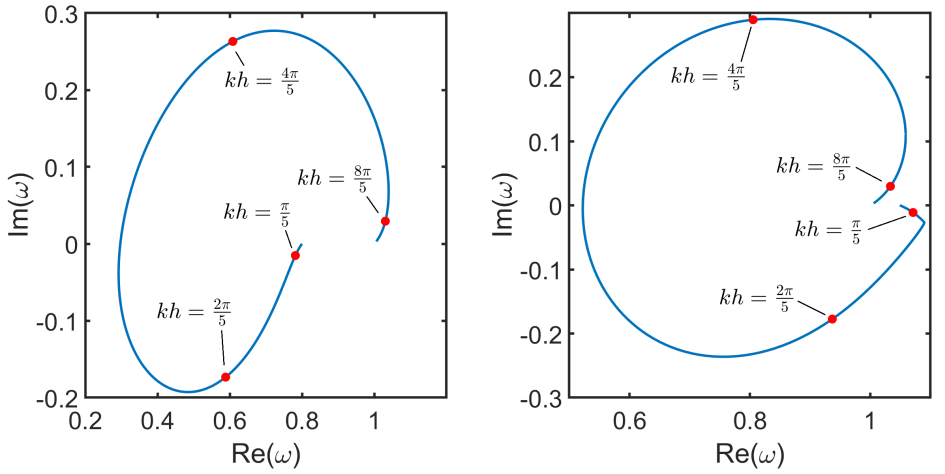
$$\begin{aligned} 0 = {} & \mathrm{Re}(a_0\bar{a}_1) + \left(2\,\mathrm{Re}(a_0\bar{a}_2) + |a_1|^2\right)x \\ & + 3\,\mathrm{Re}\left(a_0\bar{a}_3 + a_1\bar{a}_2\right)x^2 + \left(4\,\mathrm{Re}(a_0\bar{a}_4 + a_1\bar{a}_3) + 2|a_2|^2\right)x^3 \\ & + 5\,\mathrm{Re}\left(a_1\bar{a}_4 + a_2\bar{a}_3\right)x^4 + \left(6\,\mathrm{Re}(a_2\bar{a}_4) + 3|a_3|^2\right)x^5 + 7\,\mathrm{Re}(a_3\bar{a}_4)x^6 + 4|a_4|^2x^7. \end{aligned}$$

5.3. Computation of smoothing factors and optimal relaxation parameters. Given Theorems 5.3 and 5.4, the smoothing factor $\mu(\omega)$ may be computed from (5.4). Figure 2 shows a plot of $\mu(\omega)$ so computed when $\mathcal{L}_h$ is the differential operator associated with the isotopic Helmholtz equation (2.3) with $\alpha = 0.5$, and $k(\vec{x}) \equiv k$ a constant and $\nu = 1$. The case $kh = \frac{4\pi}{5}$ is shown in Figure 2(a). Notice that $\mu(\omega)$ seems to have a single local minimum which is also the global minimum—this is also the case for a considerable number of other values of kh that we tried, and for varying degrees of anisotropy in the anisotropic case (2.1). Although minimizing (5.4) analytically appears intractable, this suggests that the problem is well suited to efficient numerical minimization. We found that computing the optimal ω_{opt} numerically using the MATLAB `fminsearch` function with initial guess $\omega_0 = \omega_{ub}$ given by Theorem 5.3 is highly effective. This is illustrated in Figure 2(b), where we have zoomed in on a region of Figure 2(a) and superimposed ω_{ub} and ω_{opt}, as well as intermediate steps in the optimization (in this case, only five iterations were required for convergence). In some cases, such as $kh = \frac{\pi}{5}$, ω_{opt} and ω_{ub} are so close that optimization is barely necessary. That said, given that iterations are extremely cheap, we allow a relatively large maximum number of iterations.



(a) Isocontours of $\mu(\omega)$. (b) Minimization iterates.

FIG. 2. *The surface $\mu(\omega)$ for RB-SOR applied to the isotropic Helmholtz equation* (2.3) *with $\alpha = 0.5$, $kh = 4\pi/5$, and $\nu = 1$. The surface is depicted in* (a). *The convergence of the numerical minimization procedure is illustrated in* (b), *where we use ω_{ub} as an initial guess to compute ω_{opt}. Convergence is reached within five iterations; see the nearly horizontal line cutting through the figure, with the value of the smoothing factor quickly decreasing from approximately* 0.87 *to less than* 0.55.



(a) ω-JAC. (b) RB-SOR.

FIG. 3. *Path in $\mathbb{C}$ traced out by the optimal complex relaxation parameter of multigrid applied to the isotropic Helmholtz problem* (2.3) *with $\alpha = 0.5$, and $k(\vec{x}) \equiv k$ constant for $d = 2$, for $kh \geq 0$. Points of interest corresponding to the optimal relaxation parameters on the top four finest grids are marked (assuming $kh = \frac{\pi}{5}$ on the finest grid).*

With a little algebra, it follows from Theorems 4.2, 5.3, and 5.4 that as $|\Lambda| \to \infty$, we have $\omega_{opt} \to 1$ for both ω-JAC and RB-SOR. In Figure 3 we visualize this by plotting the path in $\mathbb{C}$ traced out by the optimal relaxation parameters of ω-JAC and RB-SOR for this problem for $kh \geq 0$. As $kh \to \infty$ we have $\omega_{opt} \to 1$ for both smoothers, as in this case $|\Lambda| \to \infty$ as $kh \to \infty$.

We have observed that it is not enough to do smoothing analysis and derive optimal relaxation parameters on the top grid in the multigrid hierarchy; rather we should do local Fourier analysis separately on all grids and use different relaxation parameters on different grids. In the case of vertex-centered multigrid with a direct disretization on coarse grids, the computational molecule on coarse grids will also be a $2d+1$ point star of the form (1.5). The only difference is that any occurrences of h in the stencil on the fine grid will on the ℓth coarse grid have to be replaced with $2^\ell h$ (the case $\ell = 0$ denoting the finest grid).

We also consider cell-centered discretizations using d-dimensional piecewise constant interpolation P (defined for $d = 2, 3$ in [27] and which may be straightforwardly generalized), restriction given by $R = \frac{1}{2^d} P^T$, and using a Galerkin coarse grid operator based on R and P. Note that the Galerkin coarse grid operator incorporates a rescaling [2], [29, p. 524]. As has been observed in [27], in this case for $2d+1$-point stencils of the form (1.5), the sparsity pattern is preserved by the Galerkin coarse grid operator. Moreover, one may readily show that if on the fine grid the imaginary part of the operator is confined to the diagonal, the same is true on all coarse grids. Thus, our optimal relaxation parameters may be applied separately on all grids, as in the case of vertex centered multigrid with a direct discretization.

5.4. Variable coefficients and universal lookup table. For constant coefficient PDEs discretized with stencils of the form (1.5), the procedure described in section 5.3 is sufficient for computing ω_{opt} in negligible time. For variable coefficient problems, however, we in general want to use different relaxation parameters at different grid points, and this approach is no longer adequate. In order to handle this case, we observe that the smoothing factors in Theorems 5.3 and 5.4 are a function of only three quantities: $c > 0$ and $c_m > 0$ given by (3.6) and the stencil diagonal $\Lambda \in \mathbb{C}$. Moreover, we can always rescale (1.3) so that the row-sum of the off-diagonal matrix elements is equal to 1 in magnitude. This fixes $c = 0.5$ and leaves us with μ_{opt} and ω_{opt} as a function of $c_m > 0$ and $\Lambda \in \mathbb{C}$ only. Writing $\Lambda = re^{i\phi}$, we conclude that ω_{opt} and μ_{opt} can be written in the form

$$(\omega_{opt}, \mu_{opt}) = F(r, c_m, \phi)$$

for some function $F : \mathbb{R}^3 \to \mathbb{C} \times \mathbb{R}$ which may be computed using the procedure in section 5.3. To cover all cases of interest, it is sufficient to sample F on a finite box in $\mathbb{R}^3$, because c_m and ϕ are bounded while $r \gg 1$ represents a highly diagonally dominant and hence trivial scenario. We can thus precompute a universal lookup table from which the pair $(\omega_{opt}, \mu_{opt})$ may be efficiently computed for *any* PDE with stencil of the form (1.5) regardless of the dimension of the PDE or the specific values of Λ and $c_1, c_2, \ldots, c_d$. This is a key feature in our ability to efficiently solve problems with variable coefficients; see section 6.

6. Numerical experiments. In our numerical experiments we choose to primarily focus on the anisotropic Helmholtz equation (2.1) in two and three dimensions. We consider

$$-\sum_{j=1}^{d} \epsilon_j u_{x_j x_j} - k^2(\vec{x})(1 - i\alpha)u = g(\vec{x}), \tag{6.1}$$

where each ϵ_j is positive, $\sum_{j=1}^{d} \epsilon_j = d$, and $d \in \{2, 3\}$. Equation (6.1) is a special case of (2.1); the nonnegative constant α is the dissipation of the medium. As before, $k(\vec{x}) : \Omega \to [0, \infty)$ is a positive real function on Ω and $g(\vec{x})$ is a source term.

6.1. Experimental setup. We discretize (6.1) using second-order centered differences on a uniform grid G_h with mesh spacing h and Dirichlet boundary conditions. Unless specified otherwise, our domain is $[0,1]^d$, with N interior gridpoints along each dimension. We will consider vertex-centered and cell-centered multigrid; unless specified otherwise vertex-centered multigrid is used.

Our multigrid code was written in MATLAB. All intergrid transfer operators are standard full-weighting and bilinear or trilinear interpolation (for vertex-centered multigrid), and piecewise-constant interpolation and its scaled transpose (for cell-centered multigrid). All experiments use multigrid F-cycles. For constant coefficient experiments, our coarsest grid consists of a single point. For variable coefficients, our coarsest grid has $N = 15$ points along each axis (vertex-centered multigrid) or $N = 16$ points (cell-centered multigrid). We measure the convergence rate of multigrid as the average reduction factor in the norm of the residual over the final ten iterations. As a stopping criterion, we run until the initial residual norm is decreased by 10^{-10} or until a maximum number of iterations has been reached. As a right-hand side, for every experiment we have considered both a point source and a random right-hand side generated by applying the matrix from (1.3) to a random solution vector. We report the results of the random right-hand side. The results are averaged over ten runs of multigrid.

In order to avoid the pollution effect [1], constraints must be imposed on the size of kh. In the special case of isotropic Helmholtz, namely $\epsilon_1 = \epsilon_2 = \cdots = \epsilon_d = 1$, this is rather straightforward. The local wavenumber $k(\vec{x})$ is given in terms of the spatially varying speed of sound $c(\vec{x})$ and wave frequency f by $k(\vec{x}) = \frac{2\pi f}{c(\vec{x})}$. This implies that the wavelength λ is given locally by

$$\lambda(\vec{x}) = \frac{c(\vec{x})}{f} = \frac{2\pi}{k(\vec{x})}. \tag{6.2}$$

In order to maintain an accurate discrete solution, one requires that the number of points per wavelength n_w obeys $n_w \geq m$, where m is an integer number that may depend on the order of the discretization. For the isotropic case this, combined with (6.2), translates into the requirement

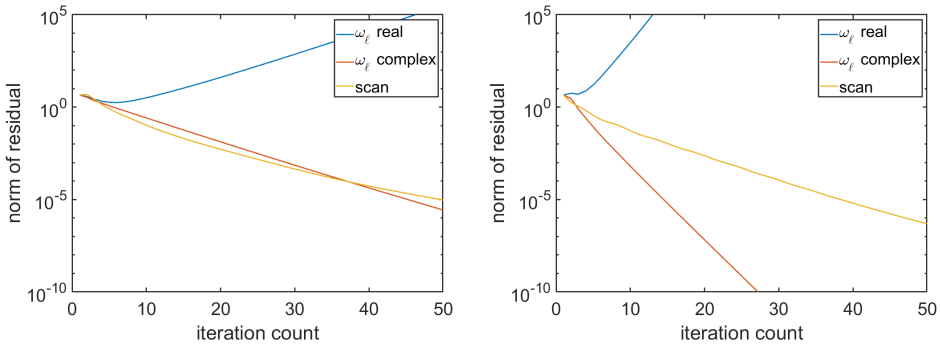
$$k_M h \leq \frac{2\pi}{m}, \qquad \text{where } k_M = \max_{\vec{x}\in G_h} k(\vec{x}). \tag{6.3}$$

Going back to the anisotropic case, the wavelength is now direction dependent. The transformation of variables $x_j = \sqrt{\epsilon_j} x'_j$ for $j = 1, \ldots, d$ turns (6.1) back into an isotropic Helmholtz equation in the transformed coordinate system; in the original coordinate system, along each x_j-axis the wavelength is thus stretched by a factor of $\sqrt{\epsilon_j}$. It thus follows from this and (6.2) that in this case the requirement (6.3) becomes

$$k_M h \leq \sqrt{\epsilon_m}\frac{2\pi}{m}, \quad \text{where} \quad k_M = \max_{\vec{x}\in G_h} k(\vec{x}) \quad \text{and} \quad \epsilon_m = \min(\epsilon_1, \ldots, \epsilon_d). \tag{6.4}$$

In [10, p. 1472], a second-order discretization is considered and m is taken to be within the range of 10 to 14. In this section, we use $m = 10$.

6.2. Experimental comparison of relaxation parameter selection strategies. As a first experiment we consider the anisotropic Helmholtz equation (6.1) in two dimensions with $\alpha = 0.5$, $k(\vec{x}) \equiv k$, $\epsilon_1 \equiv \epsilon$, and $\epsilon_2 \equiv 2-\epsilon$ constant. Various values



(a) ω-JAC. (b) RB-SOR.

FIG. 4. *Convergence curves of multigrid $F(1,1)$ cycles with ω-JAC and RB-SOR smoothing for the anisotropic Helmholtz problem* (6.1) *in two dimensions with $\epsilon_1 = 1/3$, $\epsilon_2 = 5/3$, and $\alpha = 0.5$, using three different strategies for the selection of the relaxation parameter(s). In the legend, ω_ℓ real and complex are based, respectively, on selecting optimal real and complex relaxation parameters on all grids, and "scan" means we perform an exhaustive scanning strategy.*

of $\epsilon \leq 1$ are considered. We use an $N \times N$ grid with $N = 2^8 - 1 = 255$, and select k so that $kh = \sqrt{\epsilon}\frac{\pi}{5}$ on the finest grid (that is, so that (6.4) is satisfied). A random right-hand side is used. The strategy that we advocate is performing LFA on all grids and selecting the optimal complex relaxation parameter for each grid separately. For each level $\ell = 0, \ldots, n$, choose $\omega_\ell \in \mathbb{C}$ minimizing the ℓth grid smoothing factor μ_ℓ. We have confirmed experimentally that this adaptive approach significantly improves upon the cheaper approach of performing LFA on the top grid only (that is, choose a single $\omega \in \mathbb{C}$ so that μ_0, the smoothing factor on the finest grid, is minimized), and the additional (modest) computational overhead pays off.

To demonstrate the effectiveness of our approach, we compare it with the traditional strategy of using a real ω. For fairness of comparison, we use a similar adaptive strategy here: we perform LFA on all grids with a real relaxation parameter. For each level $\ell = 0, \ldots, n$, we choose $\omega_\ell \in \mathbb{R}$ minimizing μ_ℓ.

To further demonstrate the effectiveness of our adaptive approach, we include results based on performing the impractical strategy of *exhaustive scanning*, which attempts to find the best single relaxation parameter ω—to be used on all grids—by brute force. Specifically, we run multigrid for 20 iterations each ω on a dense 100×100 grid subdividing $[0, 2] \times [-1, 1] \subset \mathbb{C}$, feed the best ω into the minimization routine, and keep the result if it is better than the best ω sampled in the grid.

For RB-SOR, we compare with two additional strategies. The first is a base case where we replace RB-SOR with red-black Gauss–Seidel (that is, set $\omega = 1$ on all levels). The second is to use ω_{ub} from Theorem 5.3 as an approximation for $\omega_{opt} \in \mathbb{C}$ minimizing μ_ℓ on each level ℓ. This was shown to be highly effective in the real case [34], so it is natural to ask whether it remains so in our setting. For the case of ω-JAC, we also compare with the relaxation parameter $\omega = 0.5$ as in [9].

Figure 4 illustrates multigrid convergence curves for a fixed choice of ϵ and different strategies, while Table 1 gives the multigrid convergence rate as well as the smoothing factor on the top four grids for various values of ϵ and different numbers of relaxation sweeps ν and for ω-JAC and RB-SOR smoothing with the various strategies described above. A pattern that we see emerge is that in most cases all strategies

do well on the finest grid, but have more difficulty on coarser grids; we speculate that this is due to a loss of diagonal dominance.

Our proposed strategy stands out in its ability to keep the smoothing factor small on the coarser grids. For RB-SOR, using optimal *real* relaxation parameters on all grids leads to serious increase in the smoothing factors on lower grids, resulting at best in significantly reduced performance and at worst in divergence. In fact, this strategy appears to do no better than red-black Gauss–Seidel.

For ω-JAC in the isotropic case $\epsilon = 1$, our strategy is still better but the effect is less pronounced, suggesting that ω-JAC is less sensitive to the choice of relaxation parameter in this case; this has also been observed in [9, p. 1476].

For the anisotropic case, we observe that ω-JAC appears to have much more difficulty than RB-SOR. For the latter, in this case using ω_{ub} as an approximation for ω_{opt} is insufficient, but using it as an initial guess for a minimization procedure on all grids as proposed in section 5.3 yields excellent convergence results.

TABLE 1

RB-SOR and ω-JAC performance for model problem (6.1) *in two dimensions with $\alpha = 0.5$, $\epsilon_1 = 2-\epsilon_2 = \epsilon$, and varying values of ϵ and ν, and a random right-hand side. The experimental multigrid convergence rate ρ and smoothing factors μ_ℓ for $\ell = 0, 1, 2, 3$ are given for a few different relaxation parameter selection strategies. The convergence rate ρ is averaged over ten runs of multigrid with different random right-hand sides. Our grid is 255×255, so that the number of unknowns is $255^2 = 65025$.*

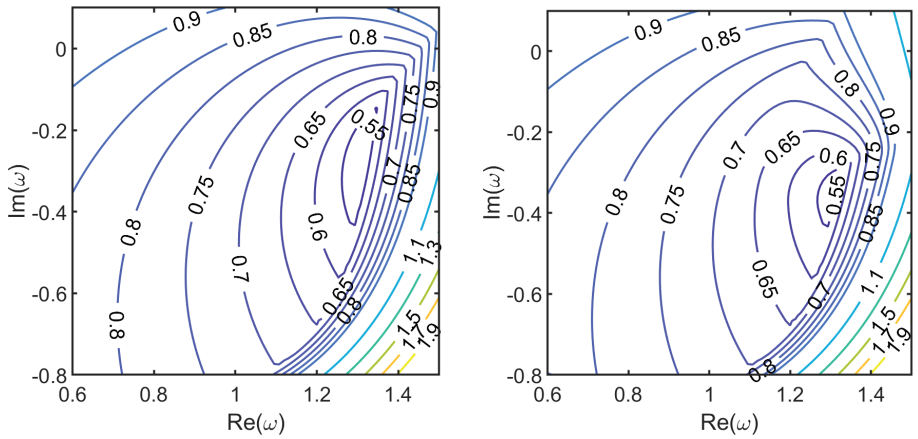
RB-SOR

Strategy	ϵ	ν	μ_0^ν	μ_1^ν	μ_2^ν	μ_3^ν	ρ
RB-GS ($\omega = 1$)	1	1	0.307	0.617	0.920	0.092	0.550
	1/3	2	0.551	0.838	5.595	0.217	> 1
	1/10	2	0.921	0.978	1.262	3.834	> 1
$\omega_\ell = \text{argmin}_{\omega\in\mathbb{R}}\, \mu_\ell(\omega)$	1	1	0.181	0.606	0.792	0.060	0.609
	1/3	2	0.172	0.758	1.000	0.209	> 1
	1/10	2	0.404	0.884	1.000	1.000	> 1
Exhaustive scanning	1	1	0.372	0.667	1.186	0.226	0.369
	1/3	2	0.757	0.854	1.191	0.642	0.828
	1/10	2	0.937	0.959	1.053	0.826	> 1
$\omega_\ell = (\omega_{ub})_\ell$	1	1	0.194	0.878	0.820	0.091	0.497
	1/3	2	0.165	0.813	2.138	0.115	> 1
	1/10	2	0.347	0.551	1.415	1.289	> 1
Our proposed strategy	1	1	0.180	0.512	0.409	0.048	0.347
	1/3	2	0.160	0.505	0.426	0.033	0.415
$\omega_\ell = \text{argmin}_{\omega\in\mathbb{C}}\, \mu_\ell(\omega)$	1/10	2	0.326	0.540	0.701	0.380	0.657

ω-JAC

Strategy	ϵ	ν	μ_0^ν	μ_1^ν	μ_2^ν	μ_3^ν	ρ
$\omega = 0.5$ (proposed in [9])	1	2	0.604	0.778	0.813	0.326	0.595
	1/3	2	0.866	0.956	1.514	0.595	> 1
	1/10	2	0.960	0.989	1.125	1.952	> 1
$\omega_\ell = \text{argmin}_{\omega\in\mathbb{R}}\, \mu_\ell(\omega)$	1	2	0.423	0.703	0.812	0.017	0.525
	1/3	2	0.761	0.922	1.000	0.425	> 1
	1/10	2	0.923	0.978	1.000	1.000	> 1
Exhaustive scanning	1	2	0.441	1.078	1.083	0.163	0.440
	1/3	2	0.819	0.892	2.502	0.909	0.821
	1/10	2	0.963	0.979	1.049	1.822	> 1
Our proposed strategy	1	2	0.423	0.664	0.452	0.016	0.421
	1/3	2	0.761	0.889	0.891	0.315	0.765
$\omega_\ell = \text{argmin}_{\omega\in\mathbb{C}}\, \mu_\ell(\omega)$	1/10	2	0.922	0.967	0.972	0.891	0.955

6.3. LFA versus two-grid analysis. Given that more accurate predictive tools such as two-grid or three-grid analysis exist, it is natural to ask if we would do



(a) Squared LFA smoothing factor $\mu(\omega)^2$. (b) Two-grid spectral radius $\rho^{2G}(\omega)$.

FIG. 5. *Comparison of the two-grid spectral radius as a function of relaxation parameter ω with the squared smoothing factor as a function of ω, for the anisotropic Helmholtz problem* (6.1) *with $\epsilon_1 = \epsilon$, $\epsilon_2 = 2 - \epsilon$ for $\epsilon = \frac{1}{3}$, with $kh = \frac{2}{5}\pi$, $\alpha = 0.5$, and $\nu = 2$.*

TABLE 2

Comparison of the use of optimal LFA relaxation parameters on each grid with optimal two-grid relaxation parameters on each grid for the anisotropic Helmholtz problem (6.1) *in two dimensions for various values of ϵ and ν, assuming RB-SOR smoothing, using a random right-hand side. In the case of LFA, the νth power μ_ℓ^ν of the smoothing factor is provided for $\ell = 0, 1, 2$ (that is, the top three grids). For the two-grid case, the spectral radius ρ_ℓ^{2G} is provided for the same values of ℓ. In both cases, the (average) multigrid convergence rate ρ and number of iterations (Iter) to converge are provided.*

		LFA					Two-grid				
ϵ	ν	μ_0^ν	μ_1^ν	μ_2^ν	Iter	ρ	ρ_0^{2G}	ρ_1^{2G}	ρ_2^{2G}	Iter	ρ
1	1	0.18	0.51	0.41	20	0.35	0.26	0.53	0.39	20	0.35
1/3	2	0.16	0.51	0.43	24	0.42	0.25	0.53	0.42	18	0.32
1/10	2	0.35	0.54	0.70	41.3	0.66	0.44	0.61	0.70	36.4	0.62

better deriving relaxation parameters minimizing these two- or three-grid spectral radii rather than minimizing LFA smoothing factors, and if so how much better? While a full answer is beyond the scope of the current work, we attempt to provide some insight in Figure 5 and Table 2. In Figure 5 we have taken the anisotropic Helmholtz problem (6.1) with $\epsilon_1 = \epsilon$, $\epsilon_2 = 2 - \epsilon$ for $\epsilon = \frac{1}{3}$, with $kh = \frac{2}{5}\pi$, $\alpha = 0.5$, $\nu = 2$ and plotted, as a function of relaxation parameter ω, both the square of the LFA smoothing factor[1] and two-grid spectral radius (computed numerically using a 300×300 grid of Fourier modes). We observe remarkable similarity in this case, which is representative of the situation for other examples we tried as well. The optimal relaxation parameter for LFA is given by $\omega_{opt}^{LFA} = 1.312 - 0.262i$, while for two-grid it is $\omega_{opt}^{2G} = 1.304 - 0.350i$. The square of the optimal LFA smoothing factor is given by $\mu(\omega_{opt}^{LFA})^2 = 0.506$. The optimal two-grid spectral radius is $\rho^{2G}(\omega_{opt}^{2G}) = 0.524$ whereas the two-grid spectral radius using the optimal LFA relaxation parameter

[1]Because the convention for LFA smoothing factors is to incorporate a power of $\frac{1}{\nu}$ in order to measure the error reduction per smoothing sweep—whereas two-grid spectral radii do not do this—the smoothing factor must be raised to the power of ν in order to compare the two objects.

is $\rho^{2G}(\omega_{opt}^{LFA}) = 0.586$. Thus, the optimal relaxation parameters predicted by LFA and two-grid analysis are in this case remarkably similar, but LFA appears slightly optimistic. Using the optimal LFA relaxation parameter as an approximation for the optimal two-level relaxation parameter results in a modest increase in the two-level spectral radii.

In Table 2, we take the same setup as the previous experiment (section 6.2) but compare the use of optimal LFA relaxation parameters on each grid with optimal two-grid relaxation parameters on each grid. We note that the greatest differences appear to be for highly anisotropic problems with small complex shifts. The magnitude of the shift appears to be particularly important. In particular, on the third grid where the shift is relatively large—the shift on the ℓth grid is proportional to $2^{2\ell}h^2$—we have very close agreement. While the optimal two-grid relaxation parameters exhibit some modest gains in the anisotropic case, their computation is prohibitive and hence impractical—numerical computation of the two-grid spectral radius based on maximizing over a d-dimensional grid of Fourier modes, with M grid points along each dimension, involves computing the spectral radius of $O(M^d)$ matrices of size $2^d \times 2^d$. By contrast, LFA relaxation parameters come at a negligible cost.

6.4. Experimental validation for a 3D model problem with constant coefficients. In this experiment we consider (6.1) in three dimensions with $\epsilon_1 = \epsilon$, $\epsilon_2 = \epsilon_3 = \frac{1}{2}(3 - \epsilon)$, and $\alpha = 0.5$.

We fix our relaxation parameter selection strategy to be optimal LFA complex relaxation parameters computed separately on each grid, the utility of this strategy having already been demonstrated in sections 6.2 and 6.3. Our goal in this example is to demonstrate that multigrid obtains h-independent convergence rates and to compare the efficiency of ω-JAC to RB-SOR smoothing. To that end, we compute multigrid convergence rates as well as the number of iterations and runtime required for convergence, for increasing values of N. The right-hand side is random. The results are shown in Table 3. In almost every case (the exception is $\epsilon = 1/5$ with ω-JAC smoothing), the convergence rate as well as iterations required for convergence are effectively constant, while the running time goes up by roughly a factor of eight each time N doubles—this is to be expected in three dimensions.

6.5. Experimental validation for a 2D problem with variable coefficients. In this experiment, we consider a generalized version of the problem (6.1) in two dimensions, with ϵ_1 and ϵ_2 now functions of $\vec{x} \in [0,1]^2$ and the constraint $\epsilon_1 + \epsilon_2 = 2$ dropped. Specifically, we take a random right-hand side, fix $\alpha = 0.5$ and $\nu = 2$, and consider

$$\epsilon_1 = 1 + \epsilon \sin(2\pi x), \quad \epsilon_2 = 1 + \epsilon \sin(2\pi y). \tag{6.5}$$

Our focus in this example is on demonstrating convergence, and as such, we study the convergence behavior for the relatively large value $kh = \frac{\pi}{5}$. We note that in reality, in order to avoid the pollution effect [1], we would need to restrict kh to adhere to a constraint similar to (6.3), which would make convergence faster.

Both ω-JAC and RB-SOR smoothing are considered for various grid sizes N. However, we now use a spatially varying relaxation parameter based on freezing coefficients. For ω-JAC, the explicit formula (4.6) from Theorem 4.2 allows us to efficiently precompute a spatially varying $\omega_{opt}(\vec{x})$ on each grid level. For RB-SOR, the same is accomplished by applying trilinear interpolation to the universal lookup table described in section 5.4. Results are given in Table 4. We observe that while multigrid

TABLE 3

RB-SOR and ω-JAC performance for the anisotropic Helmholtz problem (6.1) *in three dimensions with $\alpha = 0.5$, $\epsilon_1 = \epsilon$, and $\epsilon_2 = \epsilon_3 = \frac{1}{2}(3 - \epsilon)$ and a random right-hand side. Different mesh sizes and different values of ϵ are considered. The experimentally observed multigrid convergence rate ρ, number of iterations to converge (Iter), and time (t) in seconds to converge are given. Results are averaged over ten runs of multigrid with different random right-hand sides. For RB-SOR, we use $\nu = 1$ if $\epsilon = 1$ and $\nu = 2$ otherwise. For ω-JAC, $\nu = 2$ is used irrespective of ϵ. We use an $N \times N \times N$ grid where $N = 2^\ell - 1$ and consider $6 \le \ell \le 8$. The finest grid considered is $255 \times 255 \times 255$, so that the number of unknowns is approximately 1.65 million.*

RB-SOR

	$\epsilon = 1$			$\epsilon = 1/3$			$\epsilon = 1/5$		
ℓ	ρ	Iter	t (s)	ρ	Iter	t (s)	ρ	Iter	t (s)
6	0.42	23	3.5	0.40	20	5.7	0.49	26.4	6.8
7	0.42	23	31.0	0.40	21	54.5	0.52	28	71.2
8	0.42	23	281.4	0.41	21	495.9	0.52	28	651.2

ω-JAC

	$\epsilon = 1$			$\epsilon = 1/3$			$\epsilon = 1/5$		
ℓ	ρ	Iter	t (s)	ρ	Iter	t (s)	ρ	Iter	t (s)
6	0.59	37	6.3	0.85	109	18.6	0.90	176	30.9
7	0.59	37	54.5	0.86	111	166.8	0.93	214	320.1
8	0.58	36	457.9	0.86	110.8	1447.0	0.93	228	2890.3

convergence rates are not as stable as in the constant coefficient case, they do appear to indicate a satisfactory level of scalability. RB-SOR significantly outperforms ω-JAC, yielding reasonable convergence rates even in the highly anisotropic case $\epsilon = 0.9$ (for which ω-JAC diverges).

TABLE 4

RB-SOR and ω-JAC performance for model problem (6.1) *in two dimensions with variable coefficients* (6.5). *The experimentally observed multigrid convergence rate ρ, number of iterations to converge (Iter), and time (t) in seconds to converge are given. Results are averaged over ten runs of multigrid with different random right-hand sides. For both smoothers, we fix $\nu = 2$ and consider different values of ϵ. We use an $N \times N$ grid where $N = 2^\ell - 1$ and consider $7 \le \ell \le 12$. Divergence is denoted by DIV.*

RB-SOR

	$\epsilon = 0.5$			$\epsilon = 0.7$			$\epsilon = 0.9$		
ℓ	Iter	ρ	t (s)	Iter	ρ	t (s)	Iter	ρ	t (s)
7	22	0.388	0.3	32.8	0.559	0.4	51.3	0.706	0.7
8	25	0.431	1.0	39	0.598	1.7	55.1	0.728	2.4
9	27	0.455	5.2	45	0.641	9.1	56.4	0.737	11.4
10	28	0.468	24.6	48	0.671	42.1	56.8	0.739	50.8
11	29	0.485	107.0	49	0.644	182.4	57	0.739	210.7
12	30	0.473	459.7	54	0.658	832.3	57	0.732	907.4

ω-JAC

	$\epsilon = 0.5$			$\epsilon = 0.7$			$\epsilon = 0.9$		
ℓ	Iter	ρ	t (s)	Iter	ρ	t (s)	Iter	ρ	t (s)
7	62.6	0.768	0.9	116.9	0.850	1.7	DIV	> 1	DIV
8	69.5	0.798	3.7	130	0.875	7.0	DIV	> 1	DIV
9	71.7	0.809	16.6	139.3	0.889	32.8	DIV	> 1	DIV
10	74.2	0.815	73.1	143.2	0.895	141.2	DIV	> 1	DIV
11	74.3	0.817	303.9	143	0.895	572.3	DIV	> 1	DIV
12	74.4	0.818	1267.4	144	0.897	2474.0	DIV	> 1	DIV

6.6. The Marmousi problem. In our experiments we have thus far considered vertex-centered multigrid applied to problems with constant or smoothly varying coefficients. As a final experiment, we show how our framework may be applied—using

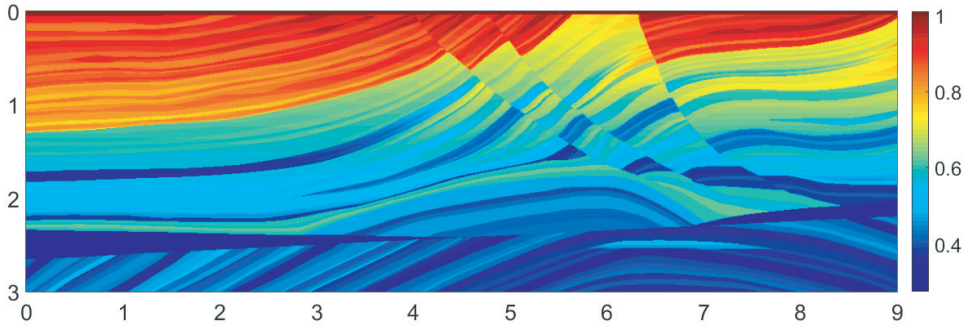


FIG. 6. *The wavenumber $k(\vec{x})$ for the Marmousi problem* [4]*, rescaled to take on a maximum value of* 1 *(units in the x and y directions are in kilometers).*

cell-centered multigrid—to a well-known benchmark problem with jumping coefficients. Specifically, we consider the 2D isotropic Helmholtz problem (2.3) with $k(\vec{x})$ given by the Marmousi dataset [4], illustrated in Figure 6. The Marmousi dataset is a 3km deep by 9km wide geophysical dataset typically used for seismic imaging modelling—we take our units as kilometers and discretize the rectangle $[0,9] \times [0,3]$ using a uniform $3N \times N$ grid for increasing values of N (we have $h = 3/N$ in this case), and each time rescale $k(\vec{x})$ so that its maximum value k_M obeys $k_M h = \frac{\pi}{5}$ (that is, the worst case value still meeting the accuracy constraint (6.3)). As in our other experiments, we use a random right-hand side.

Vertex-centered multigrid using ω-JAC smoothing in conjunction with a Galerkin coarse grid operator based on operator-dependent prolongation [6] has been observed in the literature to have difficulties in obtaining a convergent second-order scheme for $\alpha < 0.5$; see, e.g., [9, sect. 4.3]. In [30] it was shown that ILU smoothing with a fourth-order scheme can handle the case $\alpha = 0.4$.

Here we use cell-centered multigrid using piecewise-constant interpolation and a Galerkin coarse grid operator, as discussed in section 5.3. This approach has been observed to be well suited to problems with jumping coefficients [27]. An advantage in our case is that this leads to stencils of the form (1.5) on all grids, so that our framework applies on every level, as noted in section 5.3. We consider $\alpha = 0.5$ and $\alpha = 0.4$.

TABLE 5

Cell-centered multigrid with RB-SOR smoothing for the isotropic Helmholtz problem (2.3) *in two dimensions, with $k(\vec{x})$ given by the Marmousi dataset defined on the rectangle $[0,9] \times [0,3]$ and illustrated in Figure* 6*, discretized on a uniform $3N \times N$ grid where $N = 2^\ell$ with $7 \le \ell \le 11$. The experimentally observed multigrid convergence rate ρ, number of iterations to converge (Iter), and time (t) in seconds to converge are given. Results are averaged over ten runs of multigrid with different random right-hand sides. We fix $\nu = 2$ and consider $\alpha = 0.5$ and $\alpha = 0.4$.*

	$\alpha = 0.5$			$\alpha = 0.4$		
ℓ	Iter	ρ	t (s)	Iter	ρ	t (s)
7	33	0.553	1.2	58.4	0.722	2.1
8	31.6	0.538	5.1	49	0.683	7.7
9	30	0.521	20.2	42.7	0.652	28.5
10	30	0.528	82.9	40.7	0.634	112.7
11	30	0.534	338.9	41	0.642	466.6

As evident from Table 5, in our setting we are able to obtain fast scalable con-

vergence for a second-order scheme using RB-SOR smoothing, while avoiding the computational overhead of operator-dependent prolongation and maintaining a consistent sparsity pattern on all grids.

7. Conclusions and future work. We have considered optimal complex relaxation parameters minimizing smoothing factors of multigrid with damped Jacobi and red-black SOR smoothing in arbitrary dimensions. Most of our efforts are focused on the challenging case of RB-SOR. We have generalized work done in [34] for the real case, and have shown that the complex case reveals significant new analytical challenges. Our analysis is based on deriving a connection between the performance of RB-SOR as a smoother and a solver; a similar connection for the real case was noted in [34], but was neither proved nor used as an analytical tool.

In Theorem 5.3 we have derived an approximate optimal relaxation parameter ω_{ub} generalizing the approximate optimal relaxation parameter in [34]. However, its relationship with the true optimal parameter ω_{opt} is subtle; while in the real case it is sufficient to use ω_{ub} as a proxy for ω_{opt}, in the complex case we use ω_{ub} as an initial guess for a minimization routine run to find ω_{opt}.

We have demonstrated the utility of complex relaxation parameters for a number of model problems, using vertex-centered and cell-centered multigrid, including problems in three dimensions and problems with variable coefficients. For the latter, a universal lookup table allows us to efficiently precompute spatially varying relaxation parameters.

One of our central conclusions is that *complex* relaxation parameters should be used for problems with complex coefficients. Numerical experiments with anisotropic Helmholtz equations appear to confirm this view—we show that for RB-SOR, using the optimal LFA complex relaxation parameter on each grid level yields significantly faster convergence compared to using the optimal real relaxation parameter.

Our framework seems well suited for cell-centered multigrid with a Galerkin coarse grid operator based on piecewise-constant interpolation and its transpose, where the sparsity pattern of the stencil is preserved on coarse grids and the complex shift remains confined to the diagonal. We have demonstrated fast convergence in this case for the well-known Marmousi benchmark problem with jumping coefficients.

A shortcoming of our work is that it cannot be applied to the case where a Galerkin coarse grid operator is used in conjunction with vertex-centered discretization. In this case, the stencil on coarse grids will no longer be a $2d + 1$-point star—rather it will become a 3^d-point box. Some work for 9-point symmetric stencils in the 2D real case has been done in [35].

We have limited ourselves to the consideration of multigrid as a solver. A direction for future research is the extension of our analysis to the case where multigrid is used as a preconditioner for Krylov subspace methods.

A webpage containing MATLAB code with an implementation of the algorithms described in this paper is available at http://www.cs.ubc.ca/~greif/Publications/hg2020.html.

Acknowledgment. We are deeply grateful to two referees whose excellent comments and suggestions have greatly improved the quality of this paper.

REFERENCES

[1] A. Bayliss, C. I. Goldstein, and E. Turkel, *On accuracy conditions of the numerical computation of waves*, J. Comput. Phys., 59 (1985), pp. 396–404, https://doi.org/10.1016/

0021-9991(85)90119-6.
[2] D. Braess, *Towards algebraic multigrid for elliptic problems of second order*, Computing, 55 (1995), pp. 379–393, https://doi.org/10.1007/BF02238488.
[3] A. Brandt, *Multi-level adaptive solutions to boundary-value problems*, Math. Comp., 31 (1977), pp. 333–390, https://doi.org/10.2307/2006422.
[4] A. Brougois, M. Bourget, P. Lailly, M. Poulet, P. Ricarte, and R. Versteeg, *Marmousi, model and data*, in Conference Proceedings, EAEG Workshop – Practical Aspects of Seismic Data Inversion, 1990, cp-108-00002, https://doi.org/10.3997/2214-4609.201411190.
[5] S. Cools and W. Vanroose, *Local Fourier analysis of the complex shifted Laplacian preconditioner for Helmholtz problems*, Numer. Linear Algebra Appl., 20 (2013), pp. 575–597, https://doi.org/10.1002/nla.1881.
[6] P. M. de Zeeuw, *Matrix-dependent prolongations and restrictions in a blackbox multigrid solver*, J. Comput. Appl. Math., 33 (1990), pp. 1–27, https://doi.org/10.1016/0377-0427(90)90252-U.
[7] H. C. Elman, O. G. Ernst, and D. P. O'Leary, *A multigrid method enhanced by Krylov subspace iteration for discrete Helmhotz equations*, SIAM J. Sci. Comput., 23 (2001), pp. 1291–1315, https://doi.org/10.1137/S1064827501357190.
[8] B. Engquist and L. Ying, *Sweeping preconditioner for the Helmholtz equation: Moving perfectly matched layers*, Multiscale Model. Simul., 9 (2011), pp. 686–710, https://doi.org/10.1137/100804644.
[9] Y. A. Erlangga, C. W. Oosterlee, and C. Vuik, *A novel multigrid based preconditioner for heterogeneous Helmholtz problems*, SIAM J. Sci. Comput., 27 (2006), pp. 1471–1492, https://doi.org/10.1137/040615195.
[10] Y. A. Erlangga, C. Vuik, and C. W. Oosterlee, *On a class of preconditioners for solving the Helmholtz equation*, Appl. Numer. Math., 50 (2004), pp. 409–425.
[11] O. G. Ernst and M. J. Gander, *Why it is difficult to solve Helmholtz problems with classical iterative methods*, in Numerical Analysis of Multiscale Problems, Lect. Notes Comput. Sci. Eng. 83, Springer, Heidelberg, 2012, pp. 325–363, https://doi.org/10.1007/978-3-642-22061-6_10.
[12] S. Friedhoff and S. P. MacLachlan, *A generalized predictive analysis tool for multigrid methods*, Numer. Linear Algebra Appl., 22 (2015), pp. 618–647, https://doi.org/10.1002/nla.1977.
[13] T. W. Gamelin, *Complex Analysis*, Springer, 2001.
[14] M. J. Gander, I. G. Graham, and E. A. Spence, *Applying GMRES to the Helmholtz equation with shifted Laplacian preconditioning: What is the largest shift for which wavenumber-independent convergence is guaranteed?*, Numer. Math., 131 (2015), pp. 567–614, https://doi.org/10.1007/s00211-015-0700-2.
[15] M. J. Gander and H. Zhang, *A class of iterative solvers for the Helmholtz equation: Factorizations, sweeping preconditioners, source transfer, single layer potentials, polarized traces, and optimized Schwarz methods*, SIAM Rev., 61 (2019), pp. 3–76, https://doi.org/10.1137/16M109781X.
[16] G. Gilboa, N. Sochen, and Y. Y. Zeevi, *Image enhancement and denoising by complex diffusion processes*, IEEE Trans. Pattern Anal. Mach. Intell., 26 (2004), pp. 1020–1036, https://doi.org/10.1109/TPAMI.2004.47.
[17] D. J. Griffiths, *Introduction to Quantum Mechanics*, 2nd ed., Pearson Prentice Hall, 2004, http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN/0131118927.
[18] Y. He and S. Maclachlan, *Two-level Fourier analysis of multigrid for higher-order finite-element discretizations of the Laplacian*, Numer. Linear Algebra Appl., 27 (2020), e2285, https://doi.org/10.1002/nla.2285.
[19] W. M. Kahan, *Gauss-Seidel Methods of Solving Large Systems of Linear Equations*, Ph.D thesis, University of Toronto, Toronto, ON, 1958, http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&res_dat=xri:pqdiss&rft_dat=xri:pqdiss:NK01046.
[20] H. Knibbe, C. W. Oosterlee, and C. Vuik, *GPU implementation of a Helmholtz Krylov solver preconditioned by a shifted Laplace multigrid method*, J. Comput. Appl. Math., 236 (2011), pp. 281–293, https://doi.org/10.1016/j.cam.2011.07.021.
[21] B. Kredell, *On complex successive overrelaxation*, BIT Numer. Math., 2 (1962), pp. 143–152.
[22] D. Kulkarni, D. Schmidt, and S.-K. Tsui, *Eigenvalues of tridiagonal pseudo-Toeplitz matrices*, Linear Algebra Appl., 297 (1999), pp. 63–80, https://doi.org/10.1016/S0024-3795(99)00114-7.
[23] F. Liu and L. Ying, *Additive sweeping preconditioner for the Helmholtz equation*, Multiscale

Model. Simul., 14 (2016), pp. 799–822, https://doi.org/10.1137/15M1017144.
[24] F. Liu and L. Ying, *Recursive sweeping preconditioner for the three-dimensional Helmholtz equation*, SIAM J. Sci. Comput., 38 (2016), pp. A814–A832, https://doi.org/10.1137/15M1010154.
[25] G. Lozada-Cruz, *A simple application of the implicit function theorem*, Bol. Asoc. Mat. Venez., 19 (2012), pp. 71–76.
[26] S. P. MacLachlan and C. W. Oosterlee, *Algebraic multigrid solvers for complex-valued matrices*, SIAM J. Sci. Comput., 30 (2008), pp. 1548–1571, https://doi.org/10.1137/070687232.
[27] M. Mohr and R. Wienands, *Cell-centred multigrid revisited*, Comput. Vis. Sci., 7 (2004), pp. 129–140, https://doi.org/10.1007/s00791-004-0137-0.
[28] S. Reitzinger, U. Schreiber, and U. van Rienen, *Algebraic multigrid for complex symmetric matrices and applications*, J. Comput. Appl. Math., 155 (2003), pp. 405–421, https://doi.org/10.1016/S0377-0427(02)00877-4.
[29] U. Trottenberg, C. W. Oosterlee, and A. Schüller, *Multigrid*, Academic Press, San Diego, 2001.
[30] N. Umetani, S. P. MacLachlan, and C. W. Oosterlee, *A multigrid-based shifted Laplacian preconditioner for a fourth-order Helmholtz discretization*, Numer. Linear Algebra Appl., 16 (2009), pp. 603–626, https://doi.org/10.1002/nla.634.
[31] R. Wienands and W. Joppich, *Practical Fourier Analysis for Multigrid Methods*, with 1 CD-ROM (Windows and UNIX), Numer. Insights 4, Chapman & Hall/CRC, Boca Raton, FL, 2005.
[32] R. Wienands and C. W. Oosterlee, *On three-grid Fourier analysis for multigrid*, SIAM J. Sci. Comput., 23 (2001), pp. 651–671, https://doi.org/10.1137/S106482750037367X.
[33] I. Yavneh, *Multigrid smoothing factors for red-black Gauss–Seidel relaxation applied to a class of elliptic operators*, SIAM J. Numer. Anal., 32 (1995), pp. 1126–1138, https://doi.org/10.1137/0732051.
[34] I. Yavneh, *On red-black SOR smoothing in multigrid*, SIAM J. Sci. Comput., 17 (1996), pp. 180–192, https://doi.org/10.1137/0917013.
[35] I. Yavneh and E. Olvovsky, *Multigrid smoothing for symmetric nine-point stencils*, Appl. Math. Comput., 92 (1998), pp. 229–246.
[36] D. Young, *Iterative methods for solving partial difference equations of elliptic type*, Trans. Am. Math. Soc., 76 (1954), pp. 92–111, http://www.jstor.org/stable/1990745.