NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. 2013; **20**:575–597
Published online 8 April 2013 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/nla.1881

Local Fourier analysis of the complex shifted Laplacian preconditioner for Helmholtz problems

Siegfried Cools and Wim Vanroose*,†

Department of Mathematics and Computer Science, University of Antwerp, Middelheimlaan 1, 2020 Antwerp, Belgium

SUMMARY

In this paper, we solve the Helmholtz equation with multigrid preconditioned Krylov subspace methods. The class of shifted Laplacian preconditioners is known to significantly speed up Krylov convergence. However, these preconditioners have a parameter $\beta \in \mathbb{R}$, a measure of the complex shift. Because of contradictory requirements for the multigrid and Krylov convergence, the choice of this shift parameter can be a bottleneck in applying the method. In this paper, we propose a wavenumber-dependent minimal complex shift parameter, which is predicted by a rigorous k-grid local Fourier analysis (LFA) of the multigrid scheme. We claim that, given any (regionally constant) wavenumber, this minimal complex shift parameter provides the reader with a parameter choice that leads to efficient Krylov convergence. Numerical experiments in one and two spatial dimensions validate the theoretical results. It appears that the proposed complex shift is both the minimal requirement for a multigrid V-cycle to converge and being near optimal in terms of Krylov iteration count.

Received 20 December 2011; Revised 21 January 2013; Accepted 18 February 2013

KEY WORDS: Helmholtz equation; indefinite systems; high wavenumber; Krylov method; shifted Laplacian preconditioner; local Fourier analysis

1. INTRODUCTION

The propagation of waves through a material can be mathematically modelled by the Helmholtz equation

$$\left(-\Delta - k(x)^2\right) u(x) = f(x),$$

on a d-dimensional subdomain $\Omega \subset \mathbb{R}^d$. Here $k(x)$ is the space-dependent wavenumber that models change in material properties as a function of $x \in \Omega$. For high wavenumbers k, the sparse system of linear equations that results from the discretization of this PDE is distinctly indefinite, causing most of the classic iterative methods to perform poorly. Incited by its efficiency on positive definite problems, the more advanced multigrid method has repeatedly received attention as possible solver for these systems. However, as stated in [1, 2], a direct application of the multigrid method to the discretized systems will inevitably result in divergence, as both the smoother and coarse-grid correction tend to introduce growing components in the error.

The aim of this paper was to understand the required modifications to the multigrid method such that it effectively solves the indefinite linear systems resulting from the discretization of the Helmholtz equation.

Over the past few years, many different methods have been proposed to solve this non-trivial problem, an overview of which can be found in [3]. Krylov subspace methods such as generalized

*Correspondence to: Wim Vanroose, Department of Mathematics and Computer Science, University of Antwerp, Middelheimlaan 1, 2020 Antwerp, Belgium.

†E-mail: wim.vanroose@ua.ac.be

minimal residual (GMRES) [4] or biconjugate gradient stabilized (BiCGStab) [5] are known for their ability to definitely converge to the solution. However, Krylov methods are generally not competitive without a good preconditioner. Consequently, in recent literature, a variety of specifically Helmholtz-tailored preconditioners have been proposed to speed up Krylov convergence. In this paper, we focus on the class of so-called shifted Laplacian preconditioners, which were introduced in [6] and further analysed in [7, 8], where they are shown to be very efficient Krylov method preconditioners, significantly reducing the number of Krylov iterations on both academic and realistic problems. Furthermore, unlike the original Helmholtz system, the shifted Laplacian preconditioning system can be solved effectively by a multigrid method.

Similar to many other Helmholtz solvers, for example, [9], however, the shifted Laplacian method introduces a parameter that is intrinsic to the method itself, namely the magnitude of the (real and/or imaginary) shift. Ideally no shift would be introduced, as preconditioning the original system with its exact inverse causes the Krylov method to immediately converge to the solution without iterating. Unfortunately, the use of a (complex) shift is vital to guarantee multigrid convergence, as stated in [3, 6, 7, 10]. In summary, one could say that the Krylov method prefers the shift to be as small as possible, whereas the multigrid solver only converges for a sufficiently large shift. This contradiction makes the choice of an optimal shift a non-trivial yet essential task in fine tuning the shifted Laplacian method.

Note that, whereas in this paper the shift is implemented directly into the discretization matrix, it can also be represented by using complex valued grid distances in the discretization [11]. Furthermore, a complex shift can be taken into account in algebraic multilevel method, as elaborated in [12]. The Krylov convergence rate has been explored as a function of a complex shift in [13].

The current paper aims at gaining more insight in the value of the complex shift. To date, this shift is determined primarily on a heuristical case-by-case basis, as very limited theoretical foundation is known. Furthermore, in recent literature, a discussion has risen on the question whether the choice of this complex shift can be independent of the wavenumber k. We will introduce the notion of a minimal complex shift, which we define as the smallest possible shift for multigrid to converge, and show that this shift must indeed depend directly on the wavenumber. Additionally, it will be verified that the proposed complex shift is optimal in terms of preconditioned Krylov iterations when exactly solving the preconditioning problem. When applying a limited amount of multigrid iterations to the preconditioning problem as is common practice, only approximately solving the preconditioning system will it appear that our definition provides a near-optimal value for the complex shift parameter w.r.t. Krylov convergence. Consequently, following the given definition, the reader is provided a near-optimal and generally safe choice for the complex shift.

The main tool used in this paper to analyse multigrid convergence and determine a realistic value for the complex shift is the local Fourier analysis (LFA), originally introduced by Brandt in 1977 [14]. Our analysis is primarily based on the LFA setting introduced in [15–17]. Moreover, the analysis in this text can be seen as an extension of the spectral analysis performed in [18] and is even more closely related to recent work such as [19, 20], the latter combining shifted Laplacian with the new technique of multigrid deflation for improved convergence.

The results in this paper are limited to problems with homogeneous Dirichlet boundary conditions. Realistic Helmholtz problems often have absorbing boundary conditions implemented with an absorbing layer such as perfectly matched layer [21]. Note that these absorbing boundary conditions often lead to a more favorable spectrum for Krylov convergence. The analysis performed here can thus be considered a worst-case convergence scenario for multigrid preconditioned Krylov methods.

This paper is organized as follows. In Section 2, we present a rigorous LFA analysis of the problem, ultimately allowing us to define and effectively calculate the minimal complex shift parameter. In Section 3, numerical results for actually solving the Helmholtz model problem with a variety of different wavenumbers and grid sizes are presented, which serve as a validation of the definition given in Section 2. These experiments verify that the theoretical complex shift is indeed minimal w.r.t. multigrid convergence, and we furthermore observe near optimality w.r.t. Krylov convergence. Additionally, we briefly discuss the similarity between our results and the observations made in [22]. We conclude the paper with a short review of the conclusions in Section 4.

DOI: 10.1002/nla

Note that throughout the text, we will use the symbol ι to denote the imaginary unity or $\sqrt{-1}$, to avoid confusion with the index designator i.

2. LOCAL FOURIER ANALYSIS

2.1. Problem statement and notation

The general aim of our research was to solve the d-dimensional indefinite Helmholtz equation on an open bounded domain $\Omega \subset \mathbb{R}^d$

$$-\Delta u + \sigma u = f \quad \text{on } \Omega, \tag{1}$$

where $\sigma \in \mathbb{R}_0^-$ is a distinctly negative constant. Observe that we denote the squared wavenumber as $-\sigma = k^2$.[‡] We do not attend to boundary conditions, as LFA does not take into account the domain boundaries; see Section 2.2. Using multigrid on a complex shifted Laplacian (CSL) preconditioner, we consider the problem of iteratively solving the related d-dimensional complex shifted Helmholtz equation

$$-\Delta u + \tilde{\sigma} u = g \quad \text{on } \Omega, \tag{2}$$

where $\tilde{\sigma} = \sigma(1 + \beta\iota) \in \mathbb{C}$, with complex shift parameter $\beta \in \mathbb{R}^+$. Note that the CSL preconditioner was originally introduced in [6] with a somewhat more general $\tilde{\sigma} = \sigma(\alpha + \beta\iota)$, but following the observations in [7], we will choose to permanently set $\alpha \equiv 1$. This is in a way a natural choice for α, as preserving the real shift keeps the preconditioner very close to the original problem. Equation (2) is typically discretized using a finite difference scheme, yielding a linear system $A\mathbf{u} = \mathbf{g}$. In our multigrid analysis, we assume a standard second-order finite difference discretization on an equidistant mesh with N grid points, where typically $N = 2^p$ for some $p \in \mathbb{N}_0$ and with mesh width $h = 1/N$. This renders for $d = 1$ a discretization matrix A^{1D} with stencil representation

$$A^{1D} = \frac{1}{h^2}\begin{pmatrix} -1 & 2 + \tilde{\sigma}h^2 & -1 \end{pmatrix}, \tag{3}$$

or, for $d = 2$, a discretization matrix A^{2D} given by

$$A^{2D} = \frac{1}{h^2}\begin{pmatrix} & -1 & \\ -1 & 4 + \tilde{\sigma}h^2 & -1 \\ & -1 & \end{pmatrix}, \tag{4}$$

which can easily be generalized for higher dimensions. Throughout the analysis, we consider the propagation of the initial fine-grid error $\mathbf{e}_1^{(0)}$ through a k-grid analysis error propagation matrix M_1^k as follows:

$$\mathbf{e}_1^{(m+1)} = M_1^k\, \mathbf{e}_1^{(m)}, \qquad m \geqslant 0, \tag{5}$$

where M_l^k designates the iteration matrix on grid level l corresponding to the multigrid cycle that employs k grid levels. Note that we designate the finest grid by the index $l = 1$, the one-level coarser grid by $l = 2$, and so on, implying a total of 2^{p-l+1} grid points on the lth grid. For the well-known two-grid cycle, M_1^2 can be expressed as [15]

$$M_1^2 = S_1^{\nu_2}\left(I_1 - I_2^1 A_2^{-1} I_1^2 A_1\right) S_1^{\nu_1}, \tag{6}$$

which is generalized by the following definition of the error propagation matrix M_l^k for a k-grid analysis [16, 17, 19]

$$\begin{aligned} M_l^k &= S_l^{\nu_2}\left(I_l - I_{l+1}^l\left(I_{l+1} - M_{l+1}^k\right) A_{l+1}^{-1} I_l^{l+1} A_l\right) S_l^{\nu_1}, \qquad l = 1, \ldots, k-1, \\ M_k^k &= 0, \end{aligned} \tag{7}$$

[‡]To avoid unnecessary terminological complications, in this text, we loosely refer to σ as 'the wavenumber'. However, the reader should keep in mind that we hereby intrinsically designate the negatively signed squared wavenumber $-k^2$.

DOI: 10.1002/nla

where S_l is the smoothing operator, A_l is the discretization matrix representation and I_l is the identity matrix on the lth coarsest grid G_l with mesh size $h_l = 2^{l-1}h$ $(l = 1, \ldots, k)$. For the one-dimensional (1D) problem, the lth coarsest grid G_l is given by

$$G_l = \{x \in \Omega \mid x = x_j = jh_l,\ j \in \mathbb{Z}\}, \tag{8}$$

whereas in two-dimensional (2D) it is defined as

$$G_l = \left\{\mathbf{x} \in \Omega \mid \mathbf{x} = \left(x_{j_1}, x_{j_2}\right) = (j_1 h_l, j_2 h_l),\ j_1, j_2 \in \mathbb{Z}\right\}. \tag{9}$$

Furthermore, $I_l^{l+1} : G_l \to G_{l+1}$ and $I_{l+1}^l : G_{l+1} \to G_l$ are the restriction and prolongation operators, respectively. In this paper, we will use full-weighting restriction and linear interpolation as the standard intergrid operators.

2.2. *Basic principles of 1D local Fourier analysis*

We briefly sketch the key ideas behind LFA. LFA is based on the assumption that both relaxation and two-grid correction are local processes, in which each unknown is updated using only the information in neighbouring points. Furthermore, boundary conditions are neglected by extending all multigrid components to an infinite grid. It is presumed that the error $\mathbf{e}_l^{(m)}$ on the lth coarsest grid can be written as a formal linear combination of the Fourier modes $\varphi_l(\theta, x) = e^{\iota\theta x/h_l}$ with $x \in G_l$ and Fourier frequencies $\theta \in \mathbb{R}$. These frequencies may be restricted to the interval $\Theta = (-\pi, \pi] \subset \mathbb{R}$ as a consequence of the fact that for $x \in G_l$

$$\varphi_l(\theta + 2\pi, x) = \varphi_l(\theta, x). \tag{10}$$

The set of lth grid Fourier modes is typically denoted

$$\mathcal{E}_l = \operatorname{span}\left\{\varphi_l(\theta, x) = e^{\iota\theta x/h_l} \mid x \in G_l,\ \theta \in \Theta\right\}. \tag{11}$$

The Fourier modes are known [23] to be formal eigenfunctions of the operator A_l. More precisely, the general relation $A_l\varphi_l(\theta, x) = \tilde{A}_l(\theta)\varphi_l(\theta, x)$ holds, where the formal eigenvalue $\tilde{A}_l(\theta)$ is called the Fourier symbol of the operator A_l. Given a so-called low Fourier frequency $\theta^0 \in (-\pi/2, \pi/2]$, its complementary frequency θ^1 is defined as

$$\theta^1 = \theta^0 - \operatorname{sign}(\theta^0)\pi. \tag{12}$$

Using this notation, we can easily derive the following important property (for $1 \leqslant l < k$)

$$\varphi_l(\theta^0, x) = \varphi_{l+1}(2\theta^0, x) = \varphi_l(\theta^1, x), \qquad x \in G_{l+1}. \tag{13}$$

The Fourier modes $\varphi_l(\theta^0, x)$ and $\varphi_l(\theta^1, x)$ are called $(l+1)$th level harmonic modes. These Fourier modes coincide on the $(l+1)$th coarsest grid, where they are represented by a single coarse-grid mode with double frequency $\varphi_{l+1}(2\theta^0, x)$. In this way, each low-frequency mode $\varphi_l(\theta^0, x)$ is naturally associated with a high-frequency mode $\varphi_l(\theta^1, x)$ on the lth grid. It is convenient to denote the 2D subspace of $\mathcal{E}_l$ spanned by these $(l+1)$th level harmonics as

$$\mathcal{E}_l^{\theta^0} = \operatorname{span}\{\varphi_l(\theta^0, \cdot), \varphi_l(\theta^1, \cdot)\}, \qquad 1 \leqslant l < k, \tag{14}$$

with θ^1 as defined by (12). Note that every coarse-level subspace $\mathcal{E}_{l+1}^{\theta^0}$ can be decomposed into spaces spanned by finer-level harmonics, as (13) implies

$$\mathcal{E}_{l+1}^{\theta^0} = \mathcal{E}_l^{\theta^0/2} \cup \mathcal{E}_l^{\theta^1/2}, \qquad 1 \leqslant l < k-1. \tag{15}$$

The significance of these spaces $\mathcal{E}_l^{\theta}$ is that they are invariant under both smoothing operators and correction schemes under general assumptions. Because of the previous observations, throughout the Fourier symbol calculation, one can assume without loss of generality that each lth grid error

DOI: 10.1002/nla

$\mathbf{e}_l^{(m)}$ can be decomposed into components $e_{l,j}^{(m)} = \mathbf{e}_l^{(m)}(x_j)$ that consist of a single Fourier mode $\varphi_l(\theta, x_j)$ and thus can be represented as

$$e_{l,j}^{(m)} = A^{(m)}\,\varphi_l(\theta, x_j) = A^{(m)}\,e^{\imath j\theta}, \qquad \theta \in \Theta, \quad j \in \mathbb{Z}, \quad m \geqslant 0, \tag{16}$$

where the amplitude $A^{(m)}$ changes as a function of the iteration m. Moreover, using this expression together with the error relation in (5) and the stencil representation of the operators forming the error propagation matrix M_1^k, it can be derived that the error amplitudes in two subsequent iterations are related by an amplification factor function $\mathcal{G}_l(\theta, \sigma, \beta)$

$$A^{(m+1)} = \mathcal{G}_l(\theta, \sigma, \beta)A^{(m)}, \qquad m \geqslant 0, \tag{17}$$

which describes the evolution of the amplitude $A^{(m)}$ through consecutive iterations. It is our aim to elaborate an analytical expression for the amplification factor on the finest grid $\mathcal{G}_1(\theta, \sigma, \beta)$, which can be considered a continuation of the spectrum of M_1^k [10, 15]. For notational convenience, we tend to denote $\mathcal{G}_1(\theta, \sigma, \beta)$ as $\mathcal{G}(\theta, \sigma, \beta)$, dropping the fine-grid index. This amplification factor can be computed by calculating the Fourier symbols of each component of the G_1-grid error propagation matrix M_1^k. After the symbols are structured in so-called eigenmatrices for each component, combining harmonic frequencies, and the appropriate products of these matrices are taken, a $2^{k-1} \times 2^{k-1}$ eigenmatrix representation $\tilde{M}_1^k$ of the 1D error propagation matrix M_1^k is obtained. The amplification factor function $\mathcal{G}_1(\theta, \sigma, \beta)$ can then readily be computed as the spectral radius of $\tilde{M}_1^k$ [16, 23]

$$\mathcal{G}(\theta, \sigma, \beta) = \rho\left(\tilde{M}_1^k\right) = \max\left|\lambda\left(\tilde{M}_1^k\right)\right|. \tag{18}$$

A well-known condition for a general iterative method with given iteration matrix, say M, to have convergence, is that $\rho(M) < 1$. This convergence condition can be generalized to the following demand on the G_1-grid amplification factor

$$\max_{\theta \in \Theta} \mathcal{G}(\theta, \sigma, \beta) \leqslant 1, \tag{19}$$

clearly representing the analogue condition for a multigrid cycle. We denote the amplification factor as a function of the Fourier frequency θ, the wavenumber σ and the complex shift parameter β.

Definition of the minimal shift. With the help of the amplification factor, we define the *minimal complex shift parameter* $\beta_{\min}(\sigma)$ as a function of the wavenumber σ as

$$\beta_{\min} := \operatorname*{argmin}_{\beta \geqslant 0} \left\{ \max_{\theta \in \Theta} \mathcal{G}(\theta, \sigma, \beta) \leqslant 1 \right\}. \tag{20}$$

This definition can be interpreted through (19) as the smallest possible complex shift required for the multigrid method to converge, that is, it is the smallest value of β for which every single eigenmode of the error is reduced through consecutive multigrid iterations. Additionally, the numerical experiments presented in Section 3 will show that the complex shift parameter $\beta_{\min}$ as defined here is near optimal for any multigrid preconditioned Krylov method in terms of iteration count. This means that when the preconditioner is inverted exactly (up to discretization error) using a sufficiently large amount of multigrid steps, a minimal number of Krylov steps is required when choosing the value of the complex shift equal to $\beta_{\min}$.

2.3. The 1D Fourier symbols

In this section, we will effectively calculate the Fourier symbols of the different component operators of a 1D k-grid scheme.

DOI: 10.1002/nla

Discretization operator. The evolution of the error under the discretization operator A_l can be calculated from its stencil representation (3), yielding

$$e_{l,j}^{(m+1)} = \frac{1}{h_l^2}\left(-e_{l,j-1}^{(m)} + \left(2+\tilde{\sigma}h_l^2\right)\, e_{l,j}^{(m)} - e_{l,j+1}^{(m)}\right), \qquad j \in \mathbb{Z}.$$

Using expression (16), we found the amplitude relation to be

$$A^{(m+1)} = \left(-\frac{2}{h_l^2}\cos\theta + \frac{2}{h_l^2} + \tilde{\sigma}\right) A^{(m)},$$

and hence the discretization operator Fourier symbol $\tilde{A}_l(\theta)$ is

$$\tilde{A}_l(\theta) = \left(-\frac{2}{h_l^2}\cos\theta + \frac{2}{h_l^2} + \tilde{\sigma}\right). \tag{21}$$

Restriction operator. Using an analogous stencil argument, we can derive the error propagation under application of the full-weighting restriction operator I_l^{l+1} to be

$$e_{l+1,j}^{(m+1)} = \frac{1}{4}\, e_{l,2j-1}^{(m)} + \frac{1}{2}\, e_{l,2j}^{(m)} + \frac{1}{4}\, e_{l,2j+1}^{(m)}, \qquad j \in \mathbb{Z}.$$

Substituting the error components by (16), one obtains the Fourier symbol $\tilde{I}_l^{l+1}(\theta)$

$$\tilde{I}_l^{l+1}(\theta) = \frac{1}{2}(\cos\theta + 1). \tag{22}$$

Interpolation operator. The linear interpolation operator I_{l+1}^{l} propagates the coarse-grid error as

$$e_{l,2j}^{(m+1)} = e_{l+1,j}^{(m)} \quad \text{or} \quad e_{l,2j+1}^{(m+1)} = \frac{1}{2}e_{l+1,j}^{(m)} + \frac{1}{2}e_{l+1,j+1}^{(m)}, \qquad j \in \mathbb{Z},$$

where the first equation holds if $x \in G_l \cap G_{l+1}$ and the second if $x \in G_l \backslash G_{l+1}$. Once again substituting the error components by (16) and combining the aforementioned cases, the following expression for the Fourier symbol $\tilde{I}_{l+1}^{l}(\theta)$ is obtained:

$$\tilde{I}_{l+1}^{l}(\theta) = \frac{1}{2}(\cos\theta + 1). \tag{23}$$

Note that full-weighting restriction and linear interpolation are dual operators, yielding the exact same Fourier symbol $\tilde{I}_l^{l+1}(\theta) = \tilde{I}_{l+1}^{l}(\theta)$.

Smoothing operator. As a smoother, we will use standard weighted Jacobi relaxation throughout this text. It is easily derived (see e.g. [15]) that the ω-Jacobi relaxation matrix $S_l = I_l - \omega D_l^{-1} A_l$ relates the error in a grid point $x_j \in G_l$ through subsequent iterations by

$$e_{l,j}^{(m+1)} = (1-\omega)e_{l,j}^{(m)} + \frac{\omega}{2+\tilde{\sigma}h_l^2}\left(e_{l,j-1}^{(m)} + e_{l,j+1}^{(m)}\right), \qquad j \in \mathbb{Z}.$$

Presuming the error is of the form (16), one obtains the following amplitude relation and ω-Jacobi smoother Fourier symbol $\tilde{S}_l(\theta)$

$$\tilde{S}_l(\theta) = \left(1 - \omega + \frac{2\omega}{2+\tilde{\sigma}h_l^2}\cos\theta\right). \tag{24}$$

DOI: 10.1002/nla

For completeness and generality, we note that the following analysis can readily be conducted using a different smoother scheme like Gauss–Seidel relaxation, for which an analogous calculation shows the Fourier symbol $\tilde{S}_l(\theta)$ as given by

$$\tilde{S}_l(\theta) = \left(\frac{e^{i\theta}}{2 + \tilde{\sigma} h_l^2 - e^{-i\theta}} \right). \tag{25}$$

Once calculated, the 1D Fourier symbols of each component of M_1^k can now be structured in eigenmatrices, representing the action of an lth grid component on the subspace of $(l+1)$-grid harmonics, where $l = 1, \ldots, k-1$. This matrix building will be performed explicitly in the next sections for specific values of k to compute the G_1-grid amplification function $\mathcal{G}(\theta, \sigma, \beta)$ and, from this function, the minimal complex shift parameter $\beta_{\min}$.

2.4. A 1D local Fourier analysis of the weighted Jacobi smoother

To have some insight in the problem, we will initially perform a very basic LFA of the ω-Jacobi smoother. It is well known that the smoother, forming an essential part of any multigrid scheme, is unstable for the indefinite Helmholtz problem. This instability is caused by the smoothest eigenmodes, which have negative eigenvalues and consequently diverge under the action of the smoother, as described to some extend in [1, 2]. To make up for smoother instability, one could determine a complex shift purely on the basis of the divergence of the eigenmodes under application of the fine-grid smoother operator S_1. Hence, recalling that every 2D subspace of coarse-grid harmonics $\mathcal{E}_1^{\theta^0} \subset \mathcal{E}$ with $\theta^0 \in (-\pi/2, \pi/2]$ is left invariant under the actions of the smoother operator S_1, the fine-grid error propagation eigenmatrix can be written as

$$\begin{bmatrix} \tilde{S}_1(\theta^0) & 0 \\ 0 & \tilde{S}_1(\theta^1) \end{bmatrix}, \tag{26}$$

with $\tilde{S}_1(\theta)$ being the fine-grid smoother symbol derived in (24). As the matrix is diagonal, the calculation of the spectral radius in (18) is easy. This can then be substituted in definition (20) to obtain the minimal complex shift $\beta_{\min}$. Choosing this value for β, one ensures that all eigenmodes converge under the action of S_1. The result is shown in Figure 1, where the minimal complex shift parameter $\beta_{\min}$ is plotted as a function of the wavenumber σ for $N = 64$ grid points.

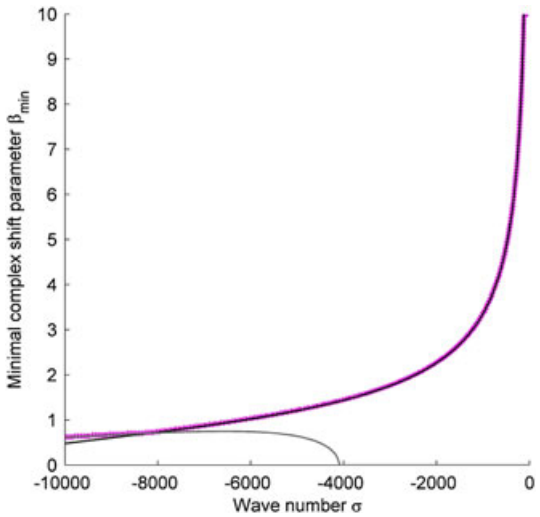

Figure 1. The minimal complex shift parameter $\beta_{\min}$ from the 1D smoother analysis with $N = 64$ and $\omega = 2/3$ as a function of the wavenumber σ. Solid black line: analytical lower limit (27) corresponding to Fourier frequency $\theta = 0$. Solid grey line: analytical lower limit (28) corresponding to the frequency $\theta = \pi$.

DOI: 10.1002/nla

One observes from Figure 1 that in order to compensate for smoother instability, values of $\beta_{\min}$ are generally very large with $\lim_{\sigma\to 0}\beta_{\min} = +\infty$, suggesting extremely large complex shifts are needed to eliminate smoother instability. Because of the simple form of the eigenmatrix (26), the lower limit $\beta_{\min}$ can easily be calculated analytically. The frequency θ maximising $\left|\tilde{S}_1(\theta)\right|$ over $(-\pi,\pi]$ can be found by setting the first derivative of $\left|\tilde{S}_1(\theta)\right|$ equal to zero, rendering

$$-8\omega^2\cos\theta\sin\theta - 4\omega(1-\omega)(2+\sigma h^2)\sin\theta = 0,$$

which reveals $\theta = 0$, $\theta = \pi$ and $\theta = \pm\arccos\{-(1-\omega)(2+\sigma h^2)/2\omega\}$ as local extrema. A second derivative check confirms $\theta = 0$ and $\theta = \pi$ to maximize the smoother symbol $\left|\tilde{S}_1(\theta)\right|$. The value of $\beta_{\min}$ can be derived by substituting both maxima in inequality (19). Presuming $\theta = 0$, the expression for $\mathcal{G}$ reduces to

$$\mathcal{G}(0,\sigma,\beta) = \left|1-\omega+\frac{2\omega}{2+\tilde{\sigma}h^2}\right| \leqslant 1,$$

which can be elaborated using the definition $\tilde{\sigma} = \sigma(1+\beta\imath)$ to yield a lower limit on the complex shift

$$\sqrt{\frac{4}{(\omega-2)\sigma h^2} - 1} \leqslant \beta. \tag{27}$$

A similar lower limit can be derived for the maximum in $\theta = \pi$, which eventually comes down to

$$\sqrt{-\frac{\omega(4+\sigma h^2)^2 - 2(2+\sigma h^2)(4+\sigma h^2)}{(\omega-2)(\sigma h^2)^2}} \leqslant \beta. \tag{28}$$

The minimal complex shift $\beta_{\min}$ is now defined as the minimal value of β satisfying both inequalities. Note that from the left-hand side of inequality (27), it readily follows that $\lim_{\sigma\to 0}\beta_{\min} = +\infty$, as we observed from Figure 1.

Although this short analysis of the smoother is useful to obtain some intuition, the value of the minimal complex shift $\beta_{\min}$ is clearly overestimated because we only consider the smoother operator, while in a multigrid setting the smoothest error components are removed by the coarse-grid correction. Indeed, to guarantee multigrid convergence, a much smaller value of β may be chosen compared with the shift suggested by Figure 1. In the next sections, the second component of the multigrid method, the correction scheme, will be taken into account to attain a realistic curve for $\beta_{\min}$ as a function of the wavenumber.

2.5. A 1D two-grid local Fourier analysis

A more realistic curve for the minimal complex shift of a multigrid cycle is obtained by considering a basic two-grid analysis with only one presmoothing step, for which the fine-grid error propagation matrix is given by

$$M_1^2 = \left(I_1 - I_2^1 A_2^{-1} I_1^2 A_1\right) S_1. \tag{29}$$

As suggested earlier, every 2D subspace of coarse-grid harmonics $\mathcal{E}_1^{\theta^0} \subset \mathcal{E}$ with $\theta^0 \in (-\pi/2, \pi/2]$ is left invariant under the actions of both the smoother operator S_1 and the pure two-grid correction operator $I_2 - I_2^1 A_2^{-1} I_1^2 A_1$. The action of the total two-grid error propagation matrix M_1^2 on $\mathcal{E}_1^{\theta^0}$ is given by its 2×2 eigenmatrix

$$\tilde{M}_1^2 = \left[I_1 - \begin{bmatrix}\tilde{I}_2^1(\theta^0)\\ \tilde{I}_2^1(\theta^1)\end{bmatrix}\tilde{A}_2(2\theta^0)^{-1}\begin{bmatrix}\tilde{I}_1^2(\theta^0)\\ \tilde{I}_1^2(\theta^1)\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\tilde{A}_1(\theta^0)\\ \tilde{A}_1(\theta^1)\end{bmatrix}^D\right]\begin{bmatrix}\tilde{S}_1(\theta^0)\\ \tilde{S}_1(\theta^1)\end{bmatrix}^D. \tag{30}$$

We use the superscript D notation to designate the operation that transforms a vector into a diagonal matrix by placing the entries of the superscribed vector along the main diagonal. The spectral

DOI: 10.1002/nla

radius of this expression, and thus the function $\mathcal{G}(\theta, \sigma, \beta)$, can easily be calculated analytically or numerically.

In Figure 2, we show the amplification factor for two choices of σ. In Figure 2(a, b), where $\sigma = -500$, we notice the appearance of a resonance that is caused by the coarse-grid correction. Only this one resonant mode tends to diverge, whereas for the majority of the frequencies θ, the amplification factor is significantly smaller than 1. The appearance of such a resonance was already discussed in [3] and originates from the inversion of the coarse-grid discretization symbol $\tilde{A}_2(2\theta^0)$ in (30). Using expression (21), one can approximate the frequency θ corresponding to the semi-asymptote by

$$\theta \approx \pm \arcsin \sqrt{-\sigma h_2^2/4} = \pm \arcsin \sqrt{-\sigma h^2}, \tag{31}$$

forcing the real part of the denominator of $\mathcal{G}$, that is, the real part of the symbol $\tilde{A}_2(2\theta^0)$, to zero and thus maximising the function $\mathcal{G}(\theta, \sigma, \beta)$ for fixed values of σ and β. Note that approximation (31) only makes sense when $|\sigma| < 1/h^2$. This discussion suggests that for small values of $|\sigma|$, the divergence of the coarse-grid correction scheme is to be the determining factor in the choice of the complex shift $\beta_{\min}$.

Alternatively, for high values of $|\sigma|$, the maximum of $G(\theta, \sigma, \beta)$ appears to be spread broadly around $\theta = 0$, as one perceives from Figure 2(c, d), which implies that a large range of smooth eigenmodes tend to diverge. The maximum here is not due to the coarse-grid correction, but originates primarily from the divergence of the smoothing operator. To substantiate this statement, we consider the smoothing operator Fourier symbol $\tilde{S}_1(\theta)$ calculated in (24). Presuming $-4/h^2 < \sigma < -1/h^2$, approximation (31) no longer holds (no semi-asymptotes are generated), and it can easily be established using derivative arguments that without a complex shift (i.e. with $\tilde{\sigma} = \sigma$), expression (24) reaches a maximum in either $\theta = 0$ (if $\sigma > -2/h^2$) or its complementary frequency $\theta = \pi$ (if

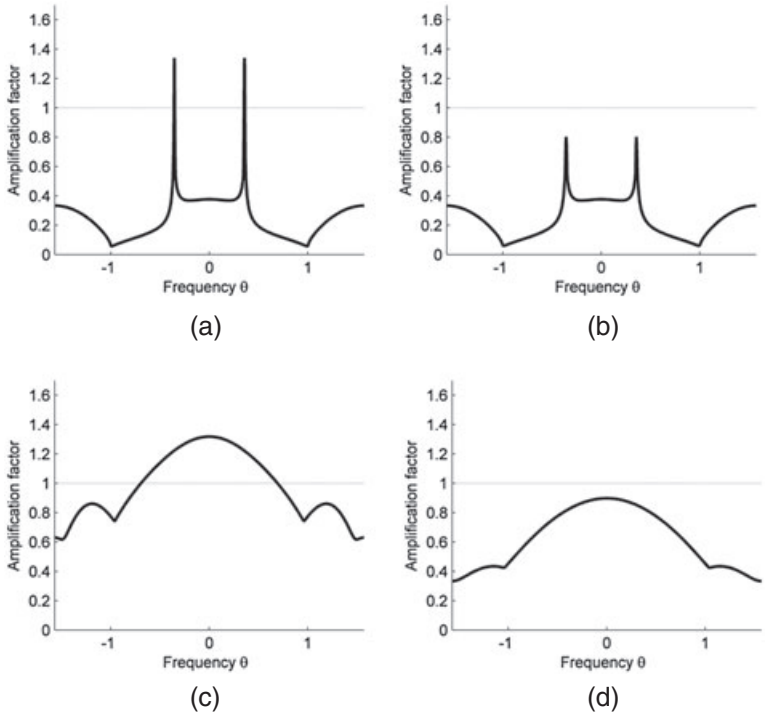

Figure 2. Amplification factor function $G(\theta, \sigma, \beta)$ for the 1D two-grid analysis with $N = 64$ as a function of $\theta \in (-\pi/2, \pi/2]$ for fixed wavenumber and complex shift. Figures (a) and (b): wavenumber $\sigma = -500$ and complex shift parameter $\beta = 0.02$ (a) and $\beta = 0.04$ (b). Figures (c) and (d): $\sigma = -5000$ and $\beta = 0.2$ (c) and $\beta = 0.8$ (d).

DOI: 10.1002/nla

$\sigma < -2/h^2$). Note that the maximization of $\tilde{S}_1(\theta)$ was already discussed in the previous section. Additionally, one can verify that the absolute value of this maximum is always larger than 1, implying divergence of the smoothest mode when applying the smoothing operator S_1. It is clear that for large values of $|\sigma|$, the main factor determining the value of $\beta_{\min}$ is the divergence caused by the smoother operator, rather than the coarse-grid correction.

From analytical expression (30) for the amplification factor $\mathcal{G}(\theta, \sigma, \beta)$, one can now numerically calculate the value of $\beta_{\min}$ given a fixed wavenumber σ. The maximization over $\theta \in \Theta$ is performed using an equidistant discretization of the frequency domain with a sufficiently small frequency step. Subsequently, the minimal complex shift argument $\beta_{\min}$ is computed using 10 steps of the elementary bisection method, providing an accuracy of at least three decimals to $\beta_{\min}$. Figure 3 represents the minimal complex shift parameter $\beta_{\min}$ as a function of the wavenumber σ for $N = 64$ grid points. Notice the significant increase of the curve around $\sigma = -1/h^2$, corresponding to the difference in the underlying component generating $\beta_{\min}$ as discussed earlier: for wavenumbers $\sigma < -1/h^2$, the divergence of the smoothing operator is the determining factor in selecting $\beta_{\min}$,[§] whereas for $\sigma > -1/h^2$, the application of the coarse-grid correction operator is decisive.

Furthermore, we added to the figure some results for $\beta_{\min}$ on the basis of the experimentally measured convergence rate. The grey points in overplot represent these experimentally measured values of $\beta_{\min}$, determined by numerical calculation of the asymptotic convergence factors for the two-grid scheme on a random initial fine-grid error $\mathbf{e}_1^{(0)}$, while subjecting those factors to criterion (19). The experimental values are perfectly matched by the theoretical curve. The black points represent some similarly computed experimental values for a full V-cycle.

It is clear from Figure 3 that the aforementioned analysis yields accurate results for the two-grid scheme. However, it is also clear that the addition of multiple coarser grids completely alters the choice of $\beta_{\min}$ for small values of $|\sigma|$. To obtain a sufficiently accurate simulation of the full V-cycle, a higher-order k-grid scheme should be applied. In particular, we would like to guarantee a near-to-exact theoretical prediction of $\beta_{\min}$ around $\sigma = -(0.625/h)^2$ for the V-cycle, respecting the $kh \leqslant 0.625$ criterion from [24] for a minimum of 10 grid points per wavelength on the standard $[0, 1]$ domain. From this perspective, a four-grid LFA analysis will appear to be satisfactory (see next section). We emphasize however that this paper primarily focusses on the iterative solution,

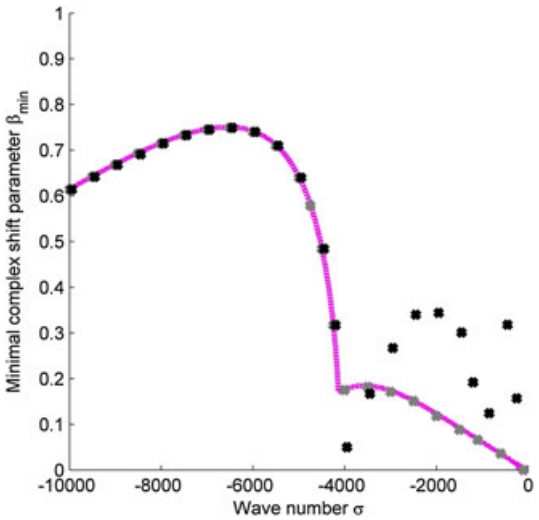

Figure 3. The minimal complex shift parameter $\beta_{\min}$ from the 1D two-grid analysis with $N = 64$ and $\omega = 2/3$ as a function of the wavenumber σ. Black and grey points represent experimentally measured values of $\beta_{\min}$ (see text).

[§]Note the striking resemblance between the high wavenumber regime of Figure 3 and the analytical lower limit (28) corresponding to $\theta = \pi$ displayed in Figure 1 (solid grey line).

DOI: 10.1002/nla

rather than the accuracy of the discretization. Indeed, the high wavenumber regime $\sigma < -4000$ of the $\beta_{\min}$ curve shown in Figure 3 is primarily useful from a theoretical point of view. However, it must be noted that when descending to coarser levels in the multigrid hierarchy, wavenumbers look larger compared with the finer levels.

2.6. *One-dimensional three-grid and four-grid local Fourier analyses*

In this section, we extend the two-grid LFA analysis to a more general k-grid analysis, motivated by the inaccurate results of a two-grid analysis for low values of $|\sigma|$. For notational purposes, we restrict ourselves to a rigorous explanation of the $k = 3$ case; however, the four-grid analysis is completely analogous. Results are shown in Figure 4 for both $k = 3$ and $k = 4$.

The three-grid fine-grid error propagation matrix is given by (7) to be

$$M_1^3 = \left(I_1 - I_2^1 \left(I_2 - \left(I_2 - I_3^2 A_3^{-1} I_2^3 A_2\right) S_2\right) A_2^{-1} I_1^2 A_1\right) S_1. \tag{32}$$

For any $\theta^0 \in (-\pi/2, \pi/2]$, the three-grid operator leaves the four-dimensional (4D) subspace $\mathcal{E}_2^{\theta^0} \subset \mathcal{E}$ of G_3 harmonics invariant, implying that its eigenmatrix $\tilde{M}_1^3$ is a 4×4 square matrix

$$\tilde{M}_1^3 = \left[I_1 - \begin{bmatrix} \tilde{I}_2^1(\theta^{0,0}) & 0 \\ \tilde{I}_2^1(\theta^{0,1}) & 0 \\ 0 & \tilde{I}_2^1(\theta^{1,0}) \\ 0 & \tilde{I}_2^1(\theta^{1,1}) \end{bmatrix} \left[I_2 - \tilde{M}_2^3\right] \begin{bmatrix} \tilde{A}_2(\theta^0) \\ \tilde{A}_2(\theta^1) \end{bmatrix}^{D-1} \begin{bmatrix} \tilde{I}_2^1(\theta^{0,0}) & 0 \\ \tilde{I}_2^1(\theta^{0,1}) & 0 \\ 0 & \tilde{I}_2^1(\theta^{1,0}) \\ 0 & \tilde{I}_2^1(\theta^{1,1}) \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \tilde{A}_1(\theta^{0,0}) \\ \tilde{A}_1(\theta^{0,1}) \\ \tilde{A}_1(\theta^{1,0}) \\ \tilde{A}_1(\theta^{1,1}) \end{bmatrix}^D \right] \begin{bmatrix} \tilde{S}_1(\theta^{0,0}) \\ \tilde{S}_1(\theta^{0,1}) \\ \tilde{S}_1(\theta^{1,0}) \\ \tilde{S}_1(\theta^{1,1}) \end{bmatrix}^D,$$

$$\tilde{M}_2^3 = \left[I_2 - \begin{bmatrix} \tilde{I}_3^2(\theta^0) \\ \tilde{I}_3^2(\theta^1) \end{bmatrix} \tilde{A}_3(2\theta^0)^{-1} \begin{bmatrix} \tilde{I}_2^3(\theta^0) \\ \tilde{I}_2^3(\theta^1) \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \tilde{A}_2(\theta^0) \\ \tilde{A}_2(\theta^1) \end{bmatrix}^D \right] \begin{bmatrix} \tilde{S}_2(\theta^0) \\ \tilde{S}_2(\theta^1) \end{bmatrix}^D, \tag{33}$$

where we introduce as a short notation $\theta^{0,0} = \theta^0/2$ and $\theta^{1,0} = \theta^1/2$, and additionally $\theta^{0,1} = \theta^0/2 + \pi$ and $\theta^{1,1} = \theta^1/2 + \pi$. The amplification factor function $\mathcal{G}(\theta, \sigma, \beta)$ is per definition the spectral radius of the eigenmatrix M_1^3.

We remark that the resonances in the function $\mathcal{G}$ discussed in the previous section indeed reappear in the three-level approximation. However, approximating the resonant frequency analogue to (31) becomes a significantly non-trivial matter for $k > 2$ because of the inversion of multiple Fourier symbols.

Given a wavenumber σ, the value of the corresponding minimal complex shift $\beta_{\min}$ can then be calculated numerically as described in the previous section. The theoretical results from the three-grid and four-grid analyses are shown in Figure 4 for $N = 64$. The black dots again represent

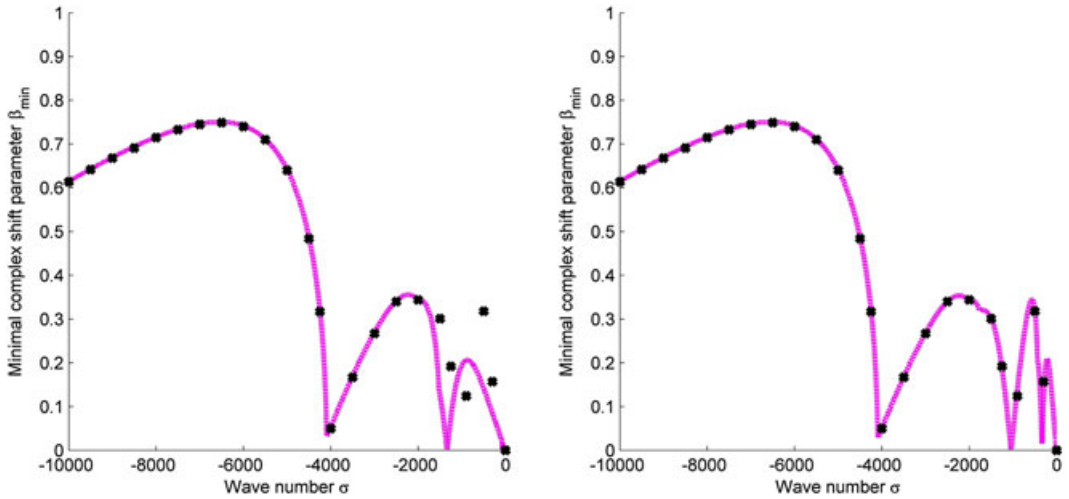

Figure 4. The minimal complex shift parameter $\beta_{\min}$ from the 1D three-grid (left) and four-grid (right) analyses with $N = 64$ and $\omega = 2/3$ as a function of the wavenumber σ.

DOI: 10.1002/nla

experimentally measured values of $\beta_{\min}$ for a full V-cycle. Note that the theoretical curve from the four-grid analysis precisely matches the experimental values around $\sigma = -(0.625/h)^2 = -1500$.

From Figure 4, one observes that adding an additional level in the LFA reveals a new 'kink' coupled by a local minimum in the $\beta_{\min}$ curve. Note how every additional local minimum revealed by the k-grid operator M_1^k is situated at the exact value of σ for which the coarsest grid operator A_k turns completely negative definite. Indeed, the two-grid operator displays a local minimum around $\sigma = -4096$, the three-grid analysis reveals an additional local minimum around $\sigma = -1024$, and the four-grid analysis adds a minimum at $\sigma = -256$. Taking into account more levels in the k-grid analysis would indeed result in further refinement of the $\beta_{\min}$ curve in the low-wavenumber regime. However, we have restricted our analysis to the four-grid scheme, as is customary in the literature.

2.7. *Basic principles of 2D local Fourier analysis*

The LFA performed in the previous sections can easily be extended to 2D problems. In this case, the error $\mathbf{e}_l^{(m)}$ can be written as a formal linear combination of the 2D Fourier modes $\varphi_l(\boldsymbol{\theta}, \mathbf{x}) = e^{\iota\left(\theta_1 x_{j_1} + \theta_2 x_{j_2}\right)/h_l}$ with $\mathbf{x} = \left(x_{j_1}, x_{j_2}\right) \in G_l$, which is now defined by (9), and a couple of frequencies $\boldsymbol{\theta} = (\theta_1, \theta_2) \in \Theta^2 = (-\pi, \pi]^2 \subset \mathbb{R}^2$. In analogy to (11), the subspace of lth grid Fourier modes is denoted:

$$\mathcal{E}_l = \text{span}\left\{\varphi_l(\mathbf{x}, \boldsymbol{\theta}) = e^{\iota\left(\theta_1 x_{j_1} + \theta_2 x_{j_2}\right)/h_l} \,|\, \mathbf{x} = \left(x_{j_1}, x_{j_2}\right) \in G_l,\ \boldsymbol{\theta} = (\theta_1, \theta_2) \in \Theta^2\right\}. \tag{34}$$

Considering a low frequency $\boldsymbol{\theta}^{00} = \left(\theta_1^{00}, \theta_2^{00}\right) \in (-\pi/2, \pi/2]^2$ and defining its 2D complementary frequencies as

$$\begin{aligned} \boldsymbol{\theta}^{10} &= \boldsymbol{\theta}^{00} - \left(\text{sign}\left(\theta_1^{00}\right)\pi, 0\right), \\ \boldsymbol{\theta}^{01} &= \boldsymbol{\theta}^{00} - \left(0, \text{sign}\left(\theta_2^{00}\right)\pi\right), \\ \boldsymbol{\theta}^{11} &= \boldsymbol{\theta}^{00} - \left(\text{sign}\left(\theta_1^{00}\right)\pi, \text{sign}\left(\theta_2^{00}\pi\right)\right), \end{aligned} \tag{35}$$

one can derive the harmonics property for $1 \leqslant l < k$

$$\varphi_l(\boldsymbol{\theta}^{00}, \mathbf{x}) = \varphi_l(\boldsymbol{\theta}^{10}, \mathbf{x}) = \varphi_{l+1}(2\boldsymbol{\theta}^{00}, \mathbf{x}) = \varphi_l(\boldsymbol{\theta}^{01}, \mathbf{x}) = \varphi_l(\boldsymbol{\theta}^{11}, \mathbf{x}), \qquad \mathbf{x} \in G_{l+1}. \tag{36}$$

The low-frequency Fourier mode $\varphi_l(\boldsymbol{\theta}^{00}, \mathbf{x})$ and three high-frequency modes $\varphi_l(\boldsymbol{\theta}^{10}, \mathbf{x})$, $\varphi_l(\boldsymbol{\theta}^{01}, \mathbf{x})$ and $\varphi_l(\boldsymbol{\theta}^{11}, \mathbf{x})$ coincide on the $(l+1)$th coarsest grid and thus are called $(l+1)$th level harmonics. Analogous to (14), we denote the 4D subspace of $\mathcal{E}_l$ spanned by these $(l+1)$th level harmonics as

$$\mathcal{E}_l^{\boldsymbol{\theta}^{00}} = \text{span}\{\varphi_l(\boldsymbol{\theta}^{00}, \cdot), \varphi_l(\boldsymbol{\theta}^{10}, \cdot), \varphi_l(\boldsymbol{\theta}^{01}, \cdot), \varphi_l(\boldsymbol{\theta}^{11}, \cdot)\}, \qquad 1 \leqslant l < k. \tag{37}$$

Again, these spaces are invariant under general smoothing operators and correction schemes. Given these notations, the 2D LFA itself is completely similar to the 1D case, with every lth grid error component $e_{l,(j_1,j_2)}^{(m)}$ being represented as a single Fourier mode

$$e_{l,(j_1,j_2)}^{(m)} = A^{(m)}\, \varphi_l\left(\boldsymbol{\theta}, \left(x_{j_1}, x_{j_2}\right)\right), \qquad \boldsymbol{\theta} \in \Theta^2, \quad j_1, j_2 \in \mathbb{Z}, \quad m \geqslant 0, \tag{38}$$

from which a relation between the amplitudes in subsequent iterations can be derived

$$A^{(m+1)} = \mathcal{G}_l(\boldsymbol{\theta}, \sigma, \beta) A^{(m)}, \qquad m \geqslant 0. \tag{39}$$

The finest-grid amplification factor $\mathcal{G}_1(\boldsymbol{\theta}, \sigma, \beta)$ of the 2D operator M_1^k can be calculated according to (18), where $\tilde{M}_1^k$ now represents the $4^{k-1} \times 4^{k-1}$ eigenmatrix representation of M_1^k. The Fourier symbols defining this eigenmatrix are derived in the next section.

Definition of the minimal shift. Once calculated, the amplification factor function allows us to compute the minimal complex shift parameter, defined in 2D as

$$\beta_{\min} := \underset{\beta \geqslant 0}{\text{argmin}} \left\{ \max_{\boldsymbol{\theta} \in \Theta^2} \mathcal{G}(\boldsymbol{\theta}, \sigma, \beta) \leqslant 1 \right\}, \tag{40}$$

which is the analogue of definition (20) for the 2D case.

DOI: 10.1002/nla

2.8. The 2D Fourier symbols

Later, we briefly sum the Fourier symbols of the different component operators of the 2D k-grid operator M_1^k. The elaboration of these symbols is completely similar to the calculations in Section 2.3, and is omitted here in favour of readability.

Discretization operator. Using stencil representation (4) and expression (38), one can readily find the discretization operator Fourier symbol $\tilde{A}_l(\boldsymbol{\theta})$ to be

$$\tilde{A}_l(\boldsymbol{\theta}) = -\frac{2}{h_l^2}\cos\theta_1 - \frac{2}{h_l^2}\cos\theta_2 + \frac{4}{h_l^2} + \tilde{\sigma}. \tag{41}$$

Restriction operator. Using an analogous stencil argument, one can derive the Fourier symbol of the 2D full-weighting restriction operator I_l^{l+1} as

$$\tilde{I}_l^{l+1}(\boldsymbol{\theta}) = \frac{1}{4}(\cos\theta_1\cos\theta_2 + \cos\theta_1 + \cos\theta_2 + 1). \tag{42}$$

Interpolation operator. As is the case in the 1D analysis, the Fourier symbol of the 2D linear interpolation operator I_{l+1}^l can be derived to yield exactly the same expression as the full-weighting restriction, that is,

$$\tilde{I}_{l+1}^l(\boldsymbol{\theta}) = \frac{1}{4}(\cos\theta_1\cos\theta_2 + \cos\theta_1 + \cos\theta_2 + 1). \tag{43}$$

Smoothing operator. Presuming the error is of the form (38), one obtains the following ω-Jacobi smoother Fourier symbol for the 2D case

$$\tilde{S}_l(\boldsymbol{\theta}) = 1 - \omega + \frac{2\omega}{4 + \tilde{\sigma}h_l^2}(\cos\theta_1 + \cos\theta_2). \tag{44}$$

2.9. Two-dimensional three-grid and four-grid local Fourier analyses

Once calculated, the 2D Fourier symbols of each component of M_1^k can again be structured in eigenmatrices, representing the action of the lth grid components on the subspace of $(l+1)$th level harmonics, where $l = 1, \ldots, k-1$. The product of these matrices yields the eigenmatrix $\tilde{M}_1^k$, from which the amplification function $\mathcal{G}(\boldsymbol{\theta}, \sigma, \beta)$ can be calculated. For any given σ, the value of the minimal complex shift parameter $\beta_{\min}$ can then be computed numerically as described in Section 2.5. We have effectively performed the 2D computation for $k = 2, 3$ and 4. The results from the three-grid and four-grid analyses for $N = 64$ are shown in Figure 5. Clearly, the observations made in the 1D case are again visible here.

We notice a significant increase of the $\beta_{\min}$ value for wavenumbers $\sigma = -2/h^2$ and smaller. For these wavenumbers, the maximum of the 2D amplification factor is determined solely by the divergence of the smoother operator. For values of $\sigma > -2/h^2$, the resonances caused by the correction scheme that appear in the 1D problem near a single frequency given by Equation (31) now appears for a range of couples (θ_1, θ_2) for which the real part of the (combination of) coarse-grid symbol(s) is approximately zero. As in the 1D case, a four-grid analysis accurately simulates $\beta_{\min}$ values extracted from V-cycle experiments up to wavenumbers as small as $\sigma = -1500$.

2.10. Extensions and general remarks

Note that all previous results were constructed under the assumption that only one pre-smoothing or post-smoothing step is applied, that is, $\nu = \nu_1 + \nu_2 = 1$. However, often multiple smoothing steps are used to obtain a more accurate or faster converging iterative solution to the given problem. The minimal complex shift obviously depends on the number of smoothing steps ν (cf. (7)). This observation is depicted in Figure 6, which shows the $\beta_{\min}$ curve for $\nu = 1, \ldots, 4$. The instability of the ω-Jacobi smoother operator, caused by divergence of the smoothest modes as described in

DOI: 10.1002/nla

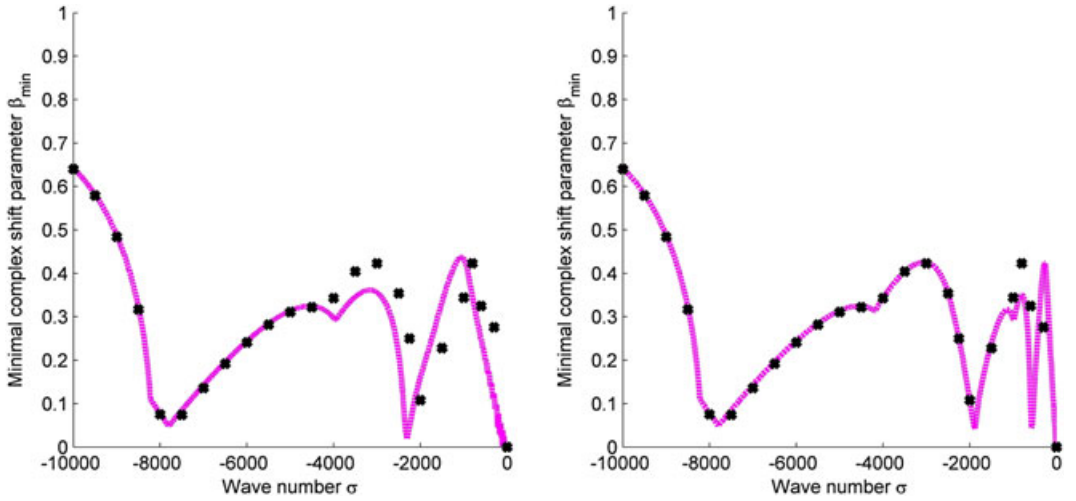

Figure 5. The minimal complex shift parameter $\beta_{\min}$ from the 2D three-grid (left) and four-grid (right) analyses with $N = 64$ and $\omega = 2/3$ as a function of the wavenumber σ.

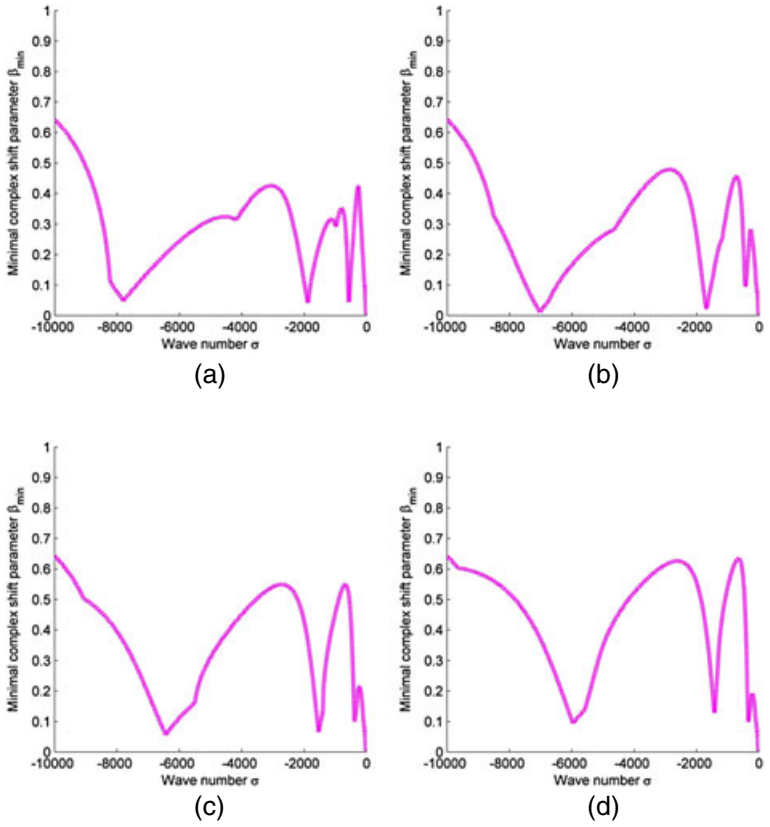

Figure 6. The minimal complex shift parameter $\beta_{\min}$ from the 2D four-grid analysis with $N = 64$ and $\omega = 2/3$ as a function of the wavenumber σ, where $\nu = 1$ (a), $\nu = 2$ (b), $\nu = 3$ (c) and $\nu = 4$ (d).

[1, 2], requires the complex shift to rise significantly for some wavenumbers σ when applying multiple smoothing steps. Altering the number of smoothing steps clearly has a significant impact on the $\beta_{\min}$ curve. As a general tendency, one could state that the value of $\beta_{\min}$ rises (at most) linearly as a function of the number of smoothing steps applied.

DOI: 10.1002/nla

Another parameter of the analysis is the weight $\omega \in [0, 1]$ of the Jacobi smoother. When applying the ω-Jacobi smoother, it is convenient to use the standard $\omega = 2/3$ weight, which is known to be optimal for a 1D Poisson problem [15]. However, for the more general Helmholtz equation with $\sigma \neq 0$, it can be shown [1] that the optimal Jacobi weight for the 1D problem is given by

$$\omega_{\text{opt}} = \frac{2 + \sigma h^2}{3 + \sigma h^2}, \tag{45}$$

implying that the smoother weight value should be smaller when considering larger values of $|\sigma|$. As shown in Figure 7, altering the smoother weight does not have more than a marginal effect on the $\beta_{\min}$ curve, causing it to rise only slightly as ω increases.[¶]

Furthermore, a comment should be made on the discretization dependency of the theoretical curves. Note that the value of $\beta_{\min}$ is indeed dependent on the finest-mesh stepsize $h = h_1$. However, it is clear from the Fourier symbol calculations that this value never appears separately from the wavenumber σ, that is, $\beta_{\min}$ intrinsically depends on the product σh^2. Hence, the value of $\beta_{\min}$ remains unchanged as long as σh^2 is constant. In practice, this implies that the $\beta_{\min}$ curve for N evaluated in σ has the exact same value as the curve for $2N$ evaluated in 4σ. Doubling the number of finest-grid points therefore implies a stretching of the $\beta_{\min}$ curve over the wavenumber domain by a factor of 4. Vice versa, the value of $\beta_{\min}$ in $\sigma/4$ for a $N/2$ discretization is identical to the value of the $\beta_{\min}$ curve in σ for N grid points, yielding an easy theoretical lower limit for the value of β when considering a multigrid V-cycle on a coarser level.

Although the analysis in this paper is restricted to the constant-k model problem given by (1), the results displayed here can be easily extended to cover more realistic space-dependent k settings. Indeed, for problems with a moderately space-dependent wavenumber, for example, the regionally space-dependent *wedge model* [7], the previous analysis may be conducted on each distinct stencil, resulting in each domain region being associated with a different minimal complex shift $\beta_{\min}$. Combining these lower limits, an indisputably safe choice for the complex shift parameter β would be the largest possible regional $\beta_{\min}$ appearing in the problem, which leads to a stable multigrid scheme in every region of the domain.

To conclude this section, we present a short discussion on the choice of $\beta_{\min}$ that is currently being used in the literature. A widely accepted choice for the complex shift parameter is $\beta = 0.5$, which was first introduced in [7] and has been used ever since by many researchers in the field. In this paper by Erlangga *et al.*, one reads that

> The preconditioner of choice in this paper is based on the parameters $(\beta_1, \beta_2) = (1, 0.5)$. (...) For values $\beta_2 < 0.5$ it is very difficult to define a satisfactory converging multigrid F(1,1)-cycle with the

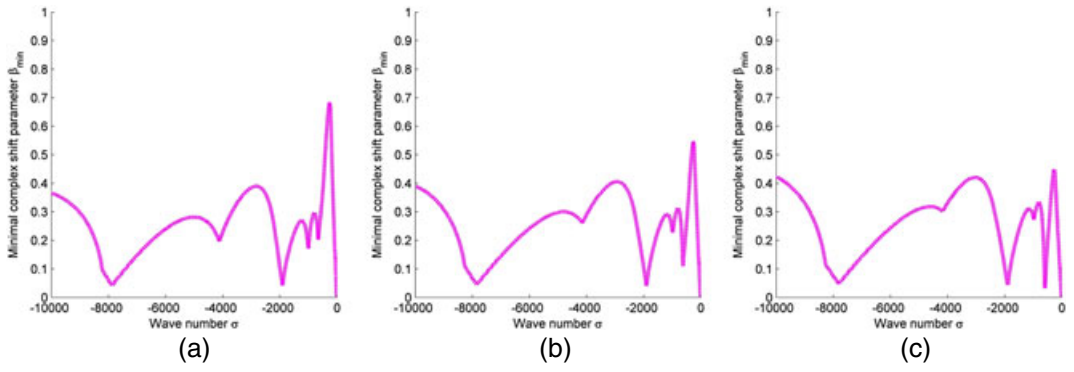

Figure 7. The minimal complex shift parameter $\beta_{\min}$ from the 2D four-grid analysis with $N = 64$ and $\nu = 1$ as a function of the wavenumber σ, where $\omega = 0.2$ (a), $\omega = 0.4$ (b) and $\omega = 0.6$ (c).

[¶]We remark that function values $\beta_{\min}$ for wavenumbers in the region $|\sigma| < 1000$ should not be taken into account, as the four-grid scheme is not guaranteed to correctly predict these values—see preceding sections.

DOI: 10.1002/nla

> components at hand. They are therefore not considered. (...) From the results in Table 6 we conclude that the preferred methods among the choices are the preconditioners with $\beta_1 = 1$. (...) Fastest [multigrid preconditioned Krylov] convergence is obtained for $(\beta_1, \beta_2) = (1, 0.5)$.

where β_1 and β_2 refer to the real and complex shift parameters, designated in this text by α and β, respectively. Conclusions were drawn from a variety of numerical experiments with a fixed wavenumber σ and mesh width satisfying $kh = 0.625$. The results can be compared with the theoretical value for $\beta_{\min}$ suggested by Figure 6(b) at $\sigma = -1600$. Note that we apply a V-cycle and use the smoother weight $\omega = 2/3$, contrary to the F-cycle and $\omega = 0.5$ used in [7]; however, these differences have only a minor effect on the $\beta_{\min}$ curve. It is clear that for all wavenumbers $\sigma \geqslant -1600$ (indicating the use of at least 10 grid points per wavelength), the corresponding minimal complex shift is indeed smaller than 0.5. In the case where $\sigma = -1600$, $\beta_{\min}$ equals 0.2, suggesting an even smaller complex shift may be used. Hence, respecting the $kh \leqslant 0.625$ criterion, the choice for $\beta = 0.5$ always guarantees V(1,1)-cycle convergence. The supposed near optimality of $\beta_{\min}$ with respect to the number of Krylov iterations is discussed in the next section.

3. NUMERICAL RESULTS

In this section, we present some experiments that will be used to confirm and validate the theoretical results obtained in Section 2, as well as provide the reader with some valuable insight on the definition choice of the minimal complex shift as stated earlier. The value of $\beta_{\min}$ predicted by the LFA curves will indeed prove a valuable lower limit for β in practical applications. Note that some of the experimental figures obtained in this section are mainly intended as a validation of the theoretical results, and their further use in practical applications is rather limited.

3.1. Minimality of the complex shift w.r.t. multigrid convergence

To validate the results from the previous section, we use a multigrid algorithm on a CSL preconditioner for some well-known Krylov methods [7], which we apply to the discretized model problem (1) with a standard all-one right-hand side $f = 1$. This leads to an outer Krylov iteration, where in each preconditioning phase a number of μ two-grid or V-cycle inner steps are applied to solve the preconditioning system. Note that all experiments presented in this chapter are performed using left preconditioning. Two standard Krylov methods for solving this model problem considered here are GMRES [4] and BiCGStab [5]. The number of multigrid steps per preconditioning phase is denoted μ, where, for example, GMRES with $\mu = 5$ indicates that five inner multigrid iterations are used within every outer GMRES step. As initial guess for the solution, we use a standard all-zero vector $\mathbf{v_0}$.

A first result is displayed in Figure 8, where the number of Krylov iterations is plotted for different choices of μ, the number of two-grid iterations per Krylov step. A relative reduction tolerance

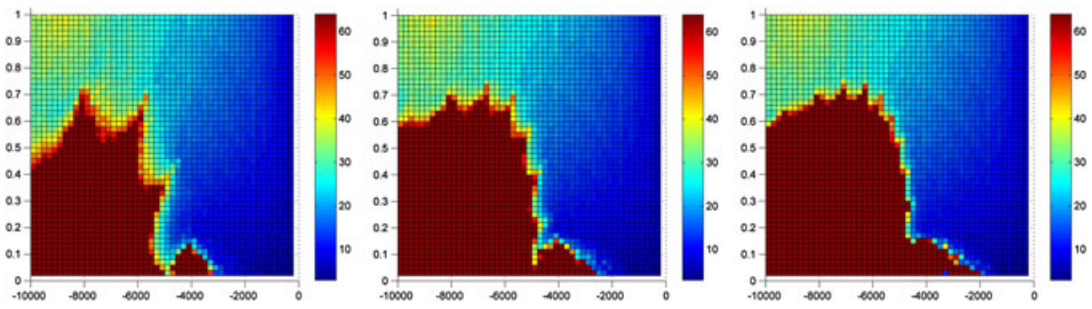

Figure 8. Two-grid preconditioned Krylov method iteration count (colour) as a function of the wavenumber σ and the complex shift β for the 1D model problem with $N = 64$. The applied method is BiCGStab with $\mu = 10$ (a), $\mu = 20$ (b) and $\mu = 30$ (c).

DOI: 10.1002/nla

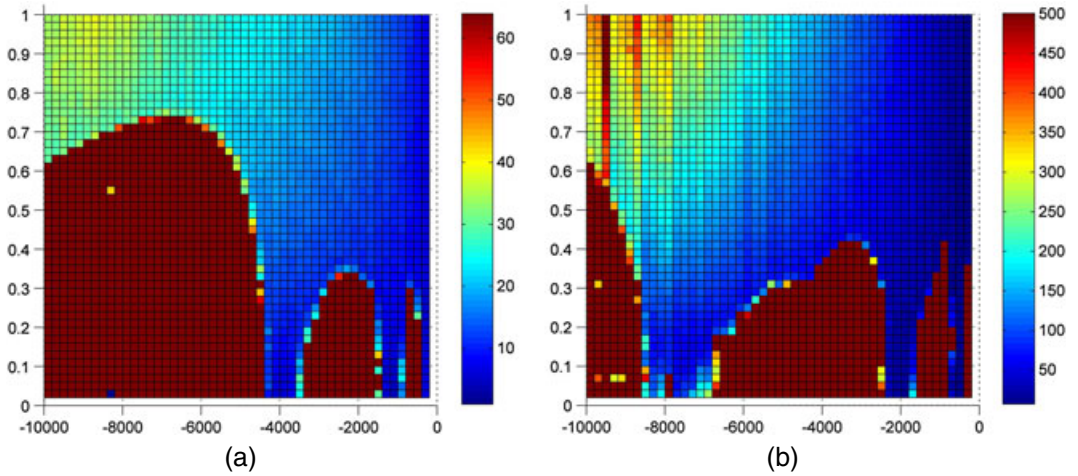

Figure 9. V(1,0)-cycle-preconditioned Krylov method iteration count (colour) as a function of the wavenumber σ and the complex shift β for the 1D (a) and 2D (b) model problems, both with $N = 64$. The applied method is BiCGStab with $\mu = 100$ (a) and $\mu = 60$ (b).

of 10^{-6} is used on the initial residual, and the maximum number of Krylov iterations is capped at 64, being the number of unknowns in the system. For most values of $\beta < \beta_{\min}$, the number of Krylov iterations required to reach the solution is excessively large. Note that for some $\beta < \beta_{\min}$, the destructive effect of the divergent modes on the global convergence appears only when sufficiently many two-grid iterations are applied. This is because divergence is a limit concept, implying that the error may not increase during the initial m_0 multigrid iterations. However, for all shifts $\beta < \beta_{\min}$, the number of Krylov iterations increases dramatically as μ grows larger, in which case we approach the theoretical curve. These experimental results confirm that choosing the minimal complex shift at least as large as $\beta_{\min}$ always ensures a safe choice for β, independently of the number of multigrid cycles performed. The experiment can easily be extended to a full V(1,0)-cycle in 1D and 2D as displayed by Figure 9, supporting the theoretical results shown in Figures 4 and 5 for the higher-order k-grid schemes.

We can state that $\beta_{\min}$ is indeed *minimal* with respect to multigrid convergence, meaning it can be considered a lower limit to *guarantee* convergence. Consequently, for a large (infinite) number of multigrid applications, $\beta_{\min}$ is the smallest possible complex shift for any multigrid preconditioned Krylov method to converge.

3.2. *Near optimality of the complex shift w.r.t. Krylov convergence*

Another important observation can be made, relating the minimal complex shift parameter $\beta_{\min}$ to Krylov convergence behaviour. For practical purposes, we consider the slightly smaller $N = 32$ model problem, with theoretical four-grid $\beta_{\min}$ curve as shown in Figure 10(a).

Temporarily focussing on a specific wavenumber σ, Figure 11 shows the number of Krylov iterations as a function of the complex shift β for four different numbers of multigrid applications. Note that for $\beta \ll \beta_{\min}$, the number of Krylov iterations is typically large as discussed in the previous section. Additionally however, one clearly observes the existence of a minimum number of Krylov iterations corresponding to a certain complex shift. It can be derived from the figure that for $\sigma = -1000$, the minimum is reached at $\beta = 0.30$ (a), $\beta = 0.36$ (b), $\beta = 0.40$ (c) and $\beta = 0.42$ (d) for $\mu = 1$ (a) , $\mu = 3$ (b), $\mu = 5$ (c) and $\mu = 10$ (d) two-grid iterations, respectively. These shifts are *optimal* for the Krylov convergence, in the sense that they reduce the global number of iterations to a minimum. Note that the theoretical minimal complex shift for $\sigma = -1000$ equals $\beta_{\min} \approx 0.42$. Hence, it appears that for a large number of multigrid iterations, the 'iteration-minimum-β' approximates the theoretical $\beta_{\min}$.

DOI: 10.1002/nla

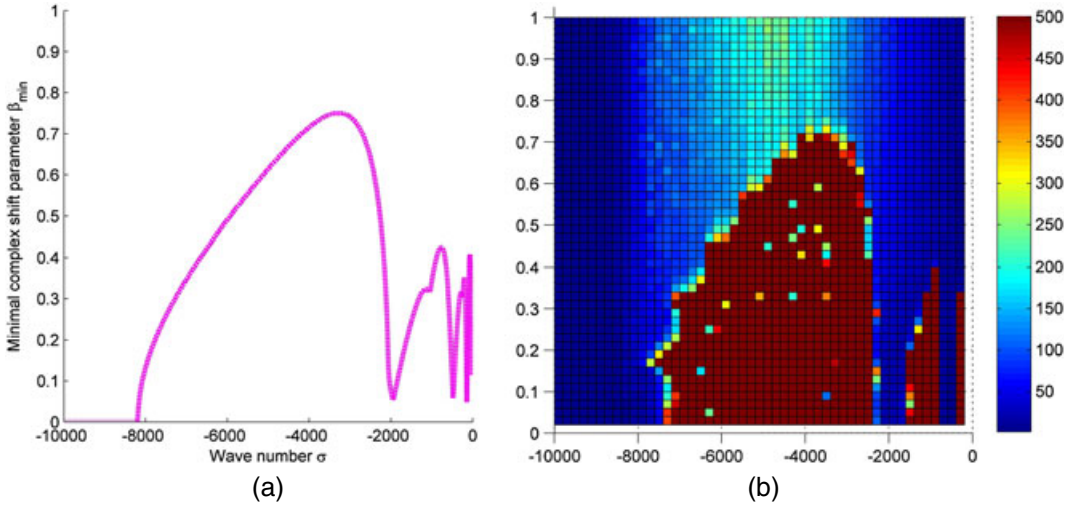

Figure 10. V(1,0)-cycle-preconditioned GMRES iteration count with $\mu = 30$ (colour) as a function of the wavenumber σ and the complex shift β for the 2D model problem with $N = 32$ (b), in comparison with the theoretical four-grid LFA complex shift parameter $\beta_{\min}$ (a).

A similar tendency applies to other values of the wavenumber, as shown in Figure 12, where the complex shift corresponding to the minimum amount of Krylov iterations is plotted as a function of the wavenumber σ. Comparing this 'iteration-minimum-β' with the theoretical $\beta_{\min}$ curve, one observes that the latter is being approximated by the iteration-minimum curves for increasing numbers of multigrid applications. Consequently, we call $\beta_{\min}$ *near optimal* with respect to the Krylov convergence, implying that for sufficiently large numbers of multigrid iterations, the global number of Krylov iterations corresponding to $\beta_{\min}$ will be minimal. However, when approximately solving the preconditioning problem by using only a small number of multigrid cycles, as is common practice, the iteration-minimum-β may be slightly smaller than $\beta_{\min}$, hence the term 'near optimality'. Nonetheless, choosing $\beta = \beta_{\min}$ as the complex shift provides a 100% multigrid stable and low-Krylov iteration solution method.

As an extra validation of these results, Figure 13 shows the number of Krylov iterations as a function of β for a more realistic 256×256-grid 2D model problem preconditioned by different amounts of V(1,0)-cycles. Again, the near optimality of $\beta_{\min}$ is clearly visible, as the shifts minimising the number of Krylov iterations $\beta = 0.26$ (a), $\beta = 0.34$ (b) and $\beta = 0.34$ (c) approximate the theoretical minimal complex shift parameter $\beta \approx 0.34$.

3.3. On the notion of near optimality

As described in the previous section, the LFA from Section 2 does not always yield the most optimal shift parameter β for a given multigrid preconditioned Krylov solver. Indeed, for some problems, choosing $\beta < \beta_{\min}$ leads to faster Krylov convergence, despite the slightly diverging multigrid preconditioner. The analysis from Section 2 can however be extended to the full preconditioned Krylov iteration matrix as presented in [25]. For a preconditioning k-grid scheme, the eigenmatrix containing the Fourier symbols corresponding to the full iteration matrix K_1^k can be written as

$$\tilde{K}_1^k = \tilde{A}_1(\sigma)\left(I_1 - \tilde{M}_1^k\right)\tilde{A}_1^{-1}(\tilde{\sigma}), \tag{46}$$

where $\tilde{A}_1(\sigma)$ is the (diagonal) eigenmatrix corresponding to the fine-grid discretization matrix $A_1(\sigma)$ of the original Helmholtz problem, $\tilde{A}_1(\tilde{\sigma})$ is the Fourier representation of the perturbed preconditioning matrix $A_1(\tilde{\sigma})$ and $\tilde{M}_1^k$ is the k-grid eigenmatrix as presented in Section 2. For completeness, we note that the preceding equation corresponds to an MG-Krylov scheme where right preconditioning is used.

DOI: 10.1002/nla

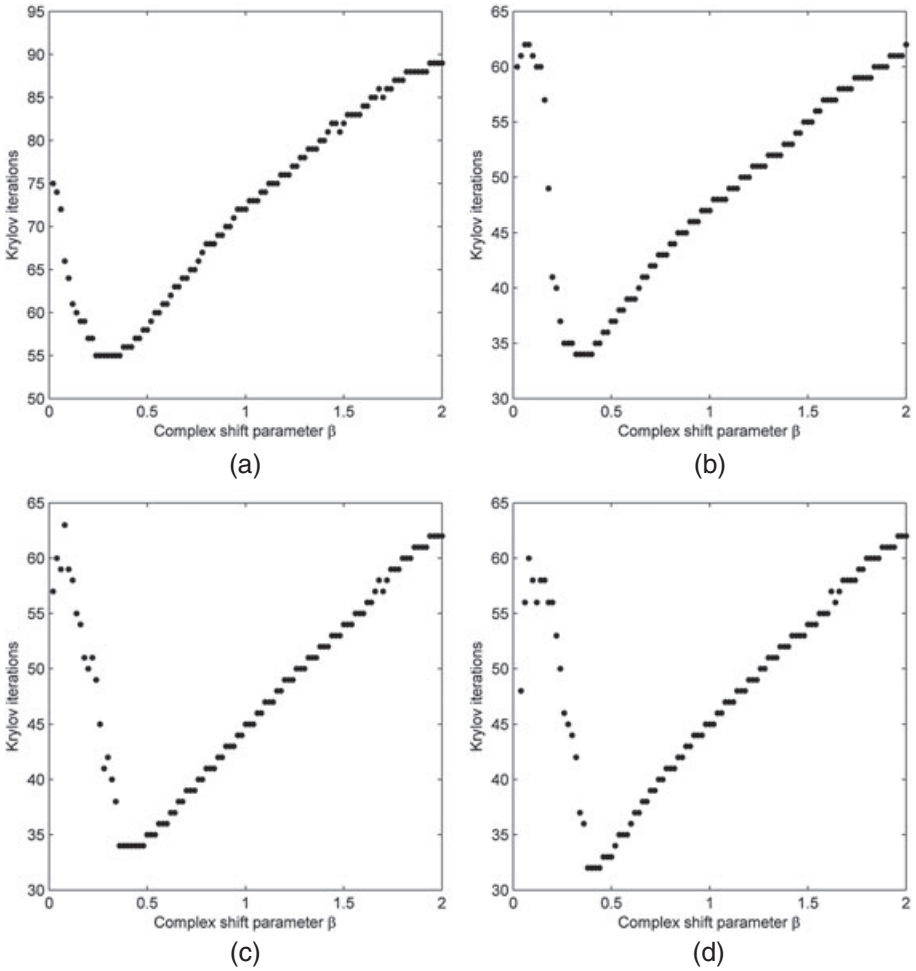

Figure 11. Two-grid preconditioned Krylov method iteration count for $\sigma = -1000$ in function of the complex shift β for the 2D model problem with $N = 32$. The applied method is GMRES with $\mu = 1$ (a), $\mu = 3$ (b), $\mu = 5$ (c) and $\mu = 10$ (d). Corresponding minima can be found at $\beta = 0.30$ (a), $\beta = 0.36$ (b), $\beta = 0.40$ (c) and $\beta = 0.42$ (d).

It is shown in [25] that, using (46), a sharp theoretical upper bound on the total convergence factor ρ_i can be determined when applying MG-GMRES(m) to the Poisson problem. In the setting of the Helmholtz problem, the determination of a sharp upper bound for ρ_i will appear to be rather non-trivial (see further). However, a sufficient condition for a given multigrid preconditioned Krylov method to converge is the so-called half-plane condition (HPC), which states that the field of values of the iteration matrix $\left\{\bar{x}^{\mathrm{T}} K_1^k x / \bar{x}^{\mathrm{T}} x : x \in \mathbb{C}^{(N-1)^d}\right\}$ should be contained in an open half-plane $\{z : Re(e^{-i\varphi} z) > 0\}$ for some $\varphi \in [0, 2\pi]$. This condition can be directly transferred into the LFA setting by demanding that the eigenvalues of $\tilde{K}_1^k$ lie within an open half-plane.

Figure 14 now shows a new minimal β curve based upon the HPC convergence criterion, effectively displaying the smallest possible value of β for which the HPC is satisfied and yielding an absolute lower limit on the shift for two-grid preconditioned Krylov convergence in function of the wavenumber σ. Note that the lower limit for β displayed in Figure 14 is distinctly smaller than the corresponding iteration-minimum-β derived from Figure 12. Consequently, the shift parameter minimising the general number of Krylov iterations can in practice be found in between the lower bound given by the HPC condition (implying MG-Krylov convergence) and the upper bound given

DOI: 10.1002/nla

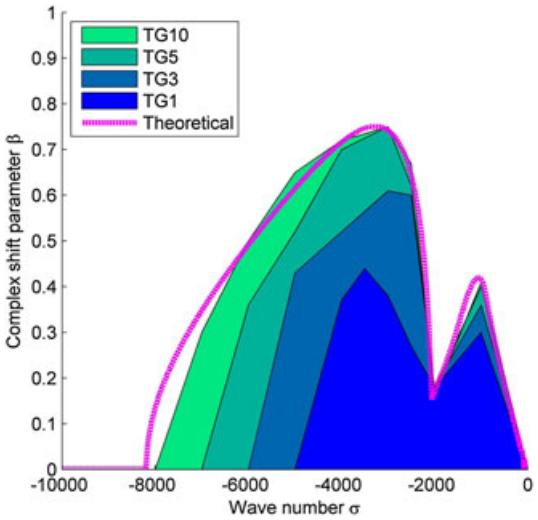

Figure 12. Two-grid preconditioned Krylov method iteration-minimum β as a function of the wavenumber σ for the 2D model problem with $N = 32$. The applied method is GMRES with $\mu = 1$, $\mu = 3$, $\mu = 5$ and $\mu = 10$.

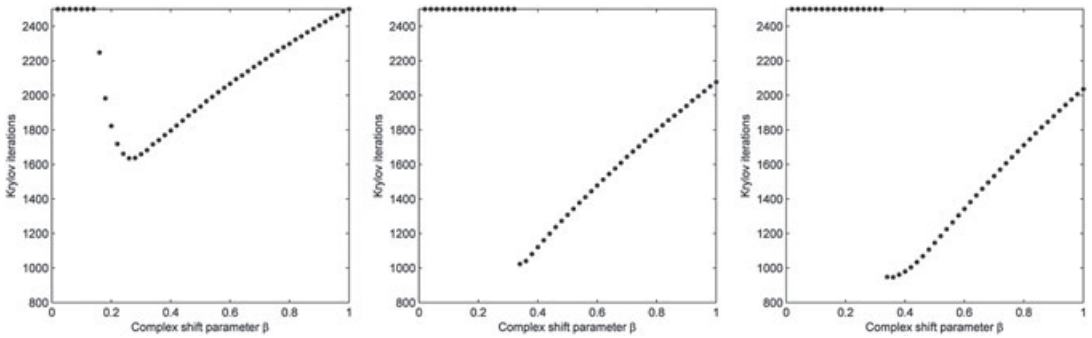

Figure 13. V(1,0)-preconditioned Krylov method iteration count for $\sigma = -64\,000$ as a function of the complex shift β for the 2D model problem with $N = 256$. The applied method is GMRES with $\mu = 3$ (a), $\mu = 5$ (b) and $\mu = 10$ (c). Minima can be found at $\beta = 0.26$ (a), $\beta = 0.34$ (b) and $\beta = 0.34$ (c).

by the LFA analysis in Section 2 (ensuring multigrid convergence), tending more towards the latter one as μ grows.

Aiming to rigourously define the iteration-minimum-β for a general multigrid preconditioned Krylov method, the subsequent analysis from [25] can be transferred to a Helmholtz setting to obtain an upper bound for the convergence factor ρ_i. Indeed, by containing the spectrum of K_1^k within an ellipse $E(c,d,a)$ (excluding the origin) with centre c, focal distance d and major-semi axis a, it is shown in [25] that $\rho_{i\gg 1}$ can be heuristically estimated by

$$\rho_i \approx \frac{a + \sqrt{a^2 - d^2}}{c + \sqrt{c^2 - d^2}}. \tag{47}$$

This estimate is generally sharp for the Poisson case as the interior of $E(c,d,a)$ is well covered by the spectrum of K_1^k. For most Helmholtz problems however, the ellipse constructed is rather pathological as the spectrum of K_1^k only covers a fraction of $E(c,d,a)$ because of its highly irregular shape. Consequently, estimate (47) given earlier is generally *not* sharp for Helmholtz problems, as illustrated by Table I. One observes that the iteration-minimum-β value is clearly reflected in the

DOI: 10.1002/nla

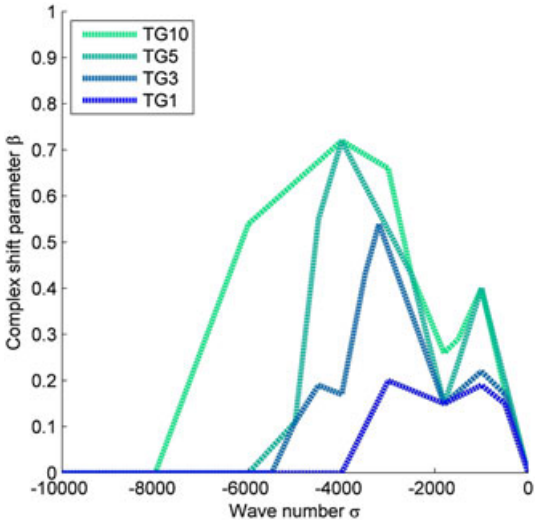

Figure 14. Two-grid preconditioned Krylov method half-plane condition minimum β as a function of the wavenumber σ for the 2D model problem with $N = 32$. Results given for $\mu = 1$, $\mu = 3$, $\mu = 5$ and $\mu = 10$.

Table I. Experimental and predicted two-grid preconditioned Krylov convergence factors for the 2D model problem with $N = 32$, $\mu = 1$ and wavenumber $\sigma = -1000$.

β	0.1	0.2	0.3	0.4	0.5
ρ_{TH}	1.00	1.00	0.99	0.99	0.99
ρ_{EX}	0.78	0.75	0.74	0.75	0.76

Experimental factors calculated based upon Figure 11; theoretical factors based upon estimate (47).

experimental convergence factor estimates but is not captured by the theoretical estimates. In conclusion, the ellipse-fitting methodology described in [25] unfortunately does not yield appropriate results for general Helmholtz problems. Future work may include the derivation of a sharp theoretical convergence factor estimate for general Helmholtz problems, leading to a rigorous prediction of the iteration-minimum-β.

3.4. General remarks on the number of Krylov iterations of a stable MG-Krylov solver

This concluding section provides the reader with some general intuition regarding the number of Krylov iterations for a sufficiently shifted (thus stable) multigrid preconditioned Krylov problem and the dependency of this number of iterations on the wavenumber σ. As demonstrated by experimental Figures 8, 9, 10(b) and 11, the general number of Krylov iterations required to accurately solve the 2D model problem with a fixed and sufficiently large complex shift β highly depends on the wavenumber σ. For sufficiently large and fixed complex shifts, one distinctly perceives the number of iterations to gradually rise to a maximum around $\sigma = -4/h^2$. Upon reaching this maximum, the number of iterations decreases slowly, exhibiting a steep descend for values of σ around $-8/h^2$. The same observation has been reported in [22], where an eigenvalue analysis was performed to rigourously anticipate the Krylov convergence behaviour. Note that we have used standard Dirichlet boundary conditions, as opposed to the absorbing boundary conditions used in [22]; however, the conclusions are identical. The large difference in the number of Krylov iterations required is due to the indefinite nature of the problem, reaching a maximum around $\sigma = -4/h^2$ (where the problem is heavily indefinite and thus hard to solve) and causing the number of iterations to suddenly

DOI: 10.1002/nla

drop for values of σ around $-8/h^2$, that is, where the 2D problem turns negative definite (and thus again easy to solve), as Figure 10(b) clearly illustrates (see colours). A rigorous explanation for this behaviour is beyond the scope of this text. For additional information and a thorough analysis on the subject, we cordially refer the reader to [22, 26].

4. CONCLUSIONS AND DISCUSSION

In this paper, we have analysed the convergence of a CSL preconditioned Krylov solver for the Helmholtz problem. A multigrid method is used to solve the shifted Laplacian preconditioning system, which is a Helmholtz problem where the wavenumber k^2 is scaled by $(1 + \beta\iota)$. This results in an operator $-\Delta + \tilde{\sigma}$, where $\tilde{\sigma} = -k^2(1 + \beta\iota)$ is complex valued.

The asymptotic multigrid convergence rate of two-grid, three-grid and four-grid schemes were analysed theoretically with the aid of LFA. It is found that the convergence rate is mainly determined by the amplification factor maximum, which appears at a single resonance frequency. The resulting convergence rate depends on the grid distance h, the complex shift parameter β and the wavenumber k. By increasing the complex shift β, the maximum at the resonance decreases and the convergence rate improves. In general, the larger β, the better the multigrid convergence rate.

However, the larger we choose the value of β, the further the complex shifted problem deviates from the original Helmholtz problem and the worse it will perform as a preconditioner. Indeed, a balance needs to be found between fast Krylov and multigrid convergence.

From the expression of the convergence rate, it is possible to define $\beta_{\min}$ as the smallest value of β for which the multilevel solution method of the shifted problem is stable. If β is taken smaller than $\beta_{\min}$, unstable modes are bound to destroy the multigrid solver.

When solving the preconditioner problem exactly using multigrid (up to discretization error order), the complex shift should always be taken larger than $\beta_{\min}$. This ensures multigrid will converge when applied to the shifted problem. Experiments show that in this case the choice of $\beta = \beta_{\min}$ is optimal, leading to a minimal number of Krylov iterations. However, when solving the preconditioner problem approximately using only a limited number of V-cycles, experimental results show that a minimum number of Krylov iterations are reached when choosing the shift parameter β slightly smaller than $\beta_{\min}$. We conclude that $\beta_{\min}$ is a safe and near-optimal choice for the complex shift parameter, ensuring multigrid stability and solving the problem by using a (nearly) minimal number of Krylov iterations.

Additionally, we have shown that $\beta_{\min}$ depends in an irregular way on the wavenumber k and the grid distance h, and that it is furthermore dependent on the number of pre-smoothing and post-smoothing steps. Choosing β around 0.5, as is common practice in the literature, ensures that the standard V(1,1)-cycle convergences for all $kh \leqslant 0.625$, as this β is distinctly larger than all corresponding $\beta_{\min}$. This indeed legitimizes the choice of $\beta = 0.5$ for practical purposes.

For problems with regional space-dependent wavenumbers $k(x)$, each region can be associated with a certain minimal complex shift $\beta_{\min}$. In that case, an indisputably safe choice for the complex shift parameter β would be the largest possible regional $\beta_{\min}$ appearing in the problem.

Note that we have used homogeneous Dirichlet boundary conditions throughout the paper, whereas many applications use absorbing boundary conditions. This shifts the eigenvalues of the original Helmholtz problem away from the real axis and generally leads to better Krylov convergence. Consequently, the analysis in this paper is a discussion of the worst-case convergence scenario, and in practice much better Krylov convergence will be found than shown by the presented experiments. As LFA makes no restrictions on the boundaries however, the theoretical results presented within this text are generally valid.

ACKNOWLEDGEMENTS

The authors acknowledge the financial support from FWO-Flanders through project G.0.120.08. W. V. is also supported by FWO *krediet aan navorser* project number 1.5.145.10. Additionally, we would like to thank Bram Reps for fruitful discussions on the subject.

DOI: 10.1002/nla

REFERENCES

1. Elman H, Ernst O, O'Leary D. A multigrid method enhanced by Krylov subspace iteration for discrete Helmholtz equations. *SIAM Journal on Scientific Computing* 2002; **23**(4):1291–1315.
2. Brandt A, Ta'asan S. Multigrid method for nearly singular and slightly indefinite problems. *Multigrid Methods II, Lecture Notes in Mathematics* 1986; **1228**:99–121.
3. Ernst O, Gander M. Why it is difficult to solve Helmholtz problems with classical iterative methods. *Lecture Notes in Computational Science and Engineering* 2012; **83**:325–363.
4. Saad Y, Schultz M. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing* 1986; **7**(3):856–869.
5. Van der Vorst H. BiCGSTAB: a fast and smoothly converging variant of BiCG for the solution of nonsymmetric linear systems. *SIAM Journal on Scientific Computing* 1992; **13**(2):631–644.
6. Erlangga Y, Oosterlee C, Vuik C. On a class of preconditioners for solving the Helmholtz equation. *Applied Numerical Mathematics* 2004; **50**(3–4):409–425.
7. Erlangga Y, Oosterlee C, Vuik C. A novel multigrid based preconditioner for heterogeneous Helmholtz problems. *SIAM Journal on Scientific Computing* 2006; **27**(4):1471–1492.
8. Erlangga Y, Vuik C, Oosterlee C. Comparison of multigrid and incomplete LU shifted-Laplace preconditioners for the inhomogeneous Helmholtz equation. *Applied Numerical Mathematics* 2006; **56**(5):648–666.
9. Haber E, MacLachlan S. A fast method for the solution of the Helmholtz equation. *Journal of Computational Physics* 2011; **230**(12):4403–4418.
10. Cools S. Complex shifted multigrid methods for the indefinite Helmholtz equation. *Master Thesis (unpublished)*, University of Antwerp, Belgium, 2011.
11. Reps B, Vanroose W, Zubair H. On the indefinite Helmholtz equation: complex stretched absorbing boundary layers, iterative analysis, and preconditioning. *Journal of Computational Physics* 2010; **229**(22):8384–8405.
12. Bollhöfer M, Grote M, Schenk O. Algebraic multilevel preconditioner for the Helmholtz equation in heterogeneous media. *SIAM Journal on Scientific Computing* 2009; **31**(5):3781–3805.
13. Osei-Kuffuor D, Saad Y. Preconditioning Helmholtz linear systems. *Applied Numerical Mathematics* 2010; **60**(4):420–431.
14. Brandt A. Multi-level adaptive solutions to boundary-value problems. *Mathematics of Computation* 1977; **31**(138):333–390.
15. Briggs W, Henson V, McCormick S. *A Multigrid Tutorial*. Society for Industrial Mathematics: Philadelphia, 2000.
16. Trottenberg U, Oosterlee C, Schüller A. *Multigrid*. Academic Press: New York, 2001.
17. Wienands R, Oosterlee C. On three-grid Fourier analysis for multigrid. *SIAM Journal on Scientific Computing* 2002; **23**(2):651–671.
18. van Gijzen H, Erlangga Y, Vuik C. Spectral analysis of the discrete Helmholtz operator preconditioned with a shifted Laplacian. *SIAM Journal on Scientific Computing* 2007; **29**(5):1942–1985.
19. Rodrigo C, Gaspar F, Oosterlee C, Yavneh I. Accuracy measures and Fourier analysis for the full multigrid algorithm. *SIAM Journal on Scientific Computing* 2010; **32**(5):3108–3129.
20. Sheikh A, Lahaye D, Vuik C. A scalable Helmholtz solver combining the shifted Laplace preconditioner with multigrid deflation. *Reports of the Department of Applied Mathematical Analysis (unpublished)*, Delft University of Technology, 2011.
21. Berenger J. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics* 1994; **114**(2):185–200.
22. Reps B, Vanroose W. Analyzing the wave number dependency of the convergence rate of a multigrid preconditioned Krylov method for the Helmholtz equation with an absorbing layer. *Numerical Linear Algebra with Applications* 2012; **19**(2):232–252.
23. Stüben K, Trottenberg U. Multigrid methods: fundamental algorithms, model problem analysis and applications. *Multigrid Methods, Lecture Notes in Mathematics* 1982; **960**:1–176.
24. Bayliss A, Goldstein C, Turkel E. On accuracy conditions for the numerical computation of waves. *Journal of Computational Physics* 1985; **59**(3):396–404.
25. Wienands R, Oosterlee C, Washio T. Fourier analysis of GMRES(m) preconditioned by multigrid. *SIAM Journal on Scientific Computing* 2000; **22**(4):582–603.
26. Simoncini V, Szyld D. Recent computational developments in Krylov subspace methods for linear systems. *Numerical Linear Algebra with Applications* 2007; **14**(1):1–59.

DOI: 10.1002/nla