Journal of Computational Physics 317 (2016) 47–65

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Multigrid-based 'shifted-Laplacian' preconditioning for the time-harmonic elastic wave equation

G. Rizzuti [a,c,*], W.A. Mulder [b,c]

[a] Delft Inversion B.V., Molengraafsingel 12, 2629 JD Delft, The Netherlands
[b] Shell Global Solutions International BV, P.O. Box 60, 2280 AB Rijswijk, The Netherlands
[c] Delft University of Technology, The Netherlands

ARTICLE INFO

Article history:
Received 8 December 2015
Received in revised form 18 March 2016
Accepted 25 April 2016
Available online 27 April 2016

Keywords:
Multigrid
Elastic
Wave equation
Frequency domain
Finite-difference

ABSTRACT

We investigate the numerical performance of an iterative solver for a frequency-domain finite-difference discretization of the isotropic elastic wave equation. The solver is based on the 'shifted-Laplacian' preconditioner, originally designed for the acoustic wave equation. This preconditioner represents a discretization of a heavily damped wave equation and can be efficiently inverted by a multigrid iteration. However, the application of multigrid to the elastic case is not straightforward because standard methods, such as point-Jacobi, fail to smooth the S-wave wavenumber components of the error when high P-to-S velocity ratios are present. We consider line smoothers as an alternative and apply local-mode analysis to evaluate the performance of the various components of the multigrid preconditioner. Numerical examples in 2-D demonstrate the efficacy of our method.

1. Introduction

The numerical solution of the wave equation is an important problem in several disciplines. The main motivation for the present work is given by geophysical applications concerning the imaging of the Earth's interior. Seismic waves, excited artificially at the surface or naturally in the subsurface, are recorded to infer its properties. Since the Earth is an elastic object, the elastic wave equation is of particular interest.

Imaging typically requires many wave simulations: for 3-D problems, where one could possibly deal with hundreds of millions of variables, efficient solvers are essential. Time- or frequency-domain methods can be considered. An attractive feature of the frequency domain is the fact that only a limited number of frequencies — well below what is prescribed by the Nyquist criterion — are actually required to obtain a satisfactory imaging result [1,2]. Moreover, a further level of parallelization is available since the computation of different sources and frequencies can be carried out independently. Due to the size of the problem, however, factorization methods such as LU decomposition with nested-dissection reordering [3], which are quite efficient in 2-D, become prohibitive in terms of required storage. For this reason, explicit time-marching methods are generally preferred, since it is not necessary to solve a large linear problem. For the frequency domain, iterative methods are an obvious alternative but are seriously hindered by the severe indefiniteness of the wave equation, especially for high frequencies. Therefore, a good preconditioner must be provided in order to obtain a reasonable convergence behaviour. In that case, iterative frequency-domain solvers can actually be competitive or even outperform the time-domain counterparts when a large number of compute cores is available, as advocated in recent work [4].

* Corresponding author.
E-mail addresses: rizzuti@delft-inversion.com (G. Rizzuti), Wim.Mulder@shell.com (W.A. Mulder).

http://dx.doi.org/10.1016/j.jcp.2016.04.049
0021-9991/

Several alternatives to iterative methods have been investigated. For example, approximate direct methods can be an efficient option, as shown by parallel multifrontal solver developments [5–7] for both the acoustic and elastic equation. One could also consider domain-decomposition methods as worked out for the acoustic case in [8] or [9]. An assessment of the relative merits of these methods is, however, beyond the scope of this paper.

Because of the difficulties mentioned above, the study of iterative solvers for the wave equation in the frequency domain is an active field of research. The present paper is based on the work of Erlangga et al. [10,11], who proposed preconditioning by a shifted Laplacian, solved by a multigrid iteration. This generalizes the work of [12] and [13]. The Helmholtz equation

$$\mathrm{H}\, p = f, \qquad \mathrm{H} = -k^2 - \Delta,$$

where k is the wavenumber and p the pressure wavefield, is replaced by a preconditioned system based on the complex shift of the Laplacian:

$$\mathrm{H}\,\mathrm{H}_\beta^{-1}\,\tilde{p} = f, \qquad \mathrm{H}_\beta = -k^2\,(1-\beta\,\mathrm{i}) - \Delta, \tag{1}$$

for some choice of the damping parameter $\beta > 0$. The key aspect of this preconditioning procedure relies on the inversion of the damped wave operator H_β by multigrid, whose attractive feature is its linear complexity. A standard multigrid method will generally fail when applied to the undamped case because of the large near-null space of the Helmholtz operator, which cannot be approximated very well on the coarser levels [14,15]. When strong damping is introduced, however, the computation of $\mathrm{H}_\beta^{-1}\,\tilde{p}$ can be accomplished by the multigrid method. Standard Krylov subspace methods such as GMRES [16], Bi-CGSTAB [17] or IDR(s) [18] can then be successfully applied to (1).

In this paper, we generalize the techniques presented by [11,10] to the elastic wave equation. Multigrid inversion of the damped elastic operator is, as one might expect, a harder problem than Helmholtz because of the different P- and S-wave modes. As P- and S-waves propagate with different velocities and, therefore, have different wavenumbers at a given frequency, smoothing should be tuned to perform well for both. On top of this, the near-null kernel comprises P- and S-waves and is more problematic than in the acoustic case. The value of β should be adjusted accordingly.

We organized this paper as follows: in Section 2 we discuss the numerical finite-difference discretization of the elastic wave equation. We will adopt two classical schemes, by Kelly [19] and by Virieux [20]. It will be shown that these schemes, besides having different numerical dispersion quality, exhibit distinct multigrid performance. Motivated by the success of the shifted-Laplacian preconditioning for the Helmholtz equation, we introduce an elastic version in Section 3 and study how it affects the spectral properties of the elastic wave equation. Similarly to the acoustic case, the preconditioned P- and S-wave eigenvalues are positioned along a circle in the positive real part of the complex plane, leading to a favourable setting for iterative methods. In Section 4, we will discuss the application of the multigrid technique to approximate the evaluation of this preconditioner. The performance of the various multigrid components will be analysed by local-mode analysis, briefly outlined for systems of equations. As it turns out, point-Jacobi smoother does not adequately smooth both the P- and S-wave modes of the error. We therefore introduce a suitable line smoother. This choice will be validated in combination with a multi-level multigrid analysis. Finally, in Section 5, we present numerical results for the homogeneous case as well as an elastic version of a highly heterogeneous example, the Marmousi model, showing good convergence properties.

2. Problem formulation and numerical discretization

We start this section with a brief discussion on the numerical discretization of the second-order formulation of the 2-D isotropic elastic wave equation. The unknowns will be represented by the particle displacement $\mathbf{u} = (u^x, u^z)$. In the frequency domain, the wave equation reads:

$$\begin{cases} -\rho\,\omega^2\,u^x - \partial_x\left[(\lambda+2\mu)\,\partial_x u^x + \lambda\,\partial_z u^z\right] - \partial_z\left(\mu\,\partial_z u^x + \mu\,\partial_x u^z\right) &= f^x, \\ -\rho\,\omega^2\,u^z - \partial_x\left(\mu\,\partial_z u^x + \mu\,\partial_x u^z\right) - \partial_z\left[\lambda\,\partial_x u^x + (\lambda+2\mu)\,\partial_z u^z\right] &= f^z, \end{cases} \tag{2}$$

where ω is the angular frequency, λ and μ are the Lamé parameters, ρ is the density and $\mathbf{f} = (f^x, f^z)$ a source term. This form will be used for the numerical experiments in Section 5. For the smoothing and multi-level multigrid analysis, we consider the homogeneous case. Then, the system (2) simplifies to:

$$\begin{cases} -k_S^2\,u^x - r^2\,\partial_{xx} u^x - \partial_{zz} u^x - (r^2-1)\,\partial_{xz} u^z &= g^x, \\ -k_S^2\,u^z - \partial_{xx} u^z - r^2\,\partial_{zz} u^z - (r^2-1)\,\partial_{xz} u^x &= g^z, \end{cases} \tag{3}$$

where $k_S = \omega/c_S$ is the wavenumber for the S-waves and $r = c_P/c_S > 1$ the P-to-S velocity ratio. The system (3) can be restated in matrix form as

$$\mathbf{H}\,\mathbf{u} = \mathbf{g}, \qquad \mathbf{H} = -k_S^2\,\mathbf{I} - \boldsymbol{\mathcal{D}}, \qquad \boldsymbol{\mathcal{D}} = \begin{pmatrix} r^2\,\partial_{xx} + \partial_{zz} & (r^2-1)\,\partial_{xz} \\ (r^2-1)\,\partial_{xz} & \partial_{xx} + r^2\,\partial_{zz} \end{pmatrix}, \tag{4}$$

to represent the elastic Helmholtz operator. Since in the 2-D elastic case we deal with a system of equations, $\mathbf{H}$ is a 2×2-block operator and we specify its scalar blocks by

$$\mathbf{H} = \begin{pmatrix} \mathrm{H}_{xx} & \mathrm{H}_{xz} \\ \mathrm{H}_{zx} & \mathrm{H}_{zz} \end{pmatrix}.$$

For the smoothing and multi-level multigrid analysis treated in the following sections, it is convenient to consider an explicit formulation, analogous to (4), for the acoustic case with $\mu = 0$. Equation (4) should be adjusted by introducing the wavenumber for the P-waves $k_S = r\,k_P$, where $k_P = \omega/c_P$, and taking the limit $r \to +\infty$ in (4). The resulting elastic operator reads:

$$\mathbf{H} = -k_P^2\,\mathbf{I} - \boldsymbol{\mathcal{D}}, \qquad \boldsymbol{\mathcal{D}} = \begin{pmatrix} \partial_{xx} & \partial_{xz} \\ \partial_{xz} & \partial_{zz} \end{pmatrix}. \tag{5}$$

Both the forms (4) and (5) have to be considered when dealing with elastic media where acoustic regions with $\mu = 0$ are present.

Several finite-difference schemes are available to discretize the system (2). We will represent the unknown functions by the values assumed on the collocation grid $\mathrm{G}^h = h\,\mathbb{Z}^2$ ($\mathbb{Z}$ being the set of integers), where h is the discretization step-length or spacing, and denote a discretized quantity by $u_{k,l} \approx u(kh, lh)$. For simplicity, we will only consider isotropic grids with equal horizontal and vertical spacings. For the treatment of these numerical methods, the stencil notation is particularly convenient. For example, we will indicate a 9-point compact scheme by

$$s \triangleq \begin{bmatrix} s(-1,-1) & s(0,-1) & s(1,-1) \\ s(-1,\ 0) & s(0,\ 0) & s(1,\ 0) \\ s(-1,\ 1) & s(0,\ 1) & s(1,\ 1) \end{bmatrix},$$

meaning that the action of the operator s on a grid function is defined by the discrete cross-correlation:

$$(s \star u)_{i,j} = \sum_{m,n} s(m,n)\,u_{i+m,j+n}. \tag{6}$$

Note that the stencil is expressed with z downwards, following the geophysical rather than the numerical convention. The discretization $\mathbf{H}^h$ of the operator $\mathbf{H}$, in the homogeneous case, may be represented by the four stencils:

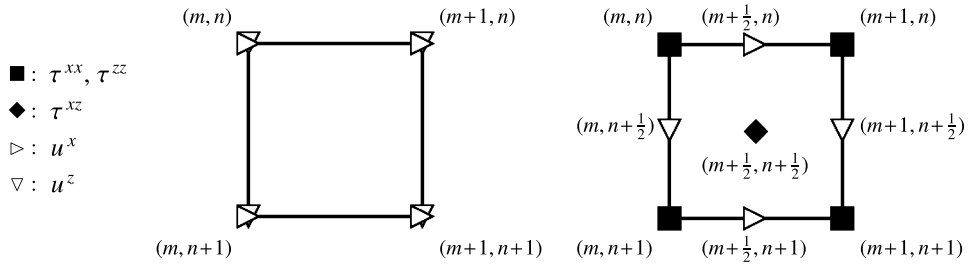
$$\begin{aligned} \mathrm{H}^h_{xx} &\triangleq -k_S^2 - r^2\,\partial^h_{xx} - \partial^h_{zz}, & \mathrm{H}^h_{xz} &\triangleq -(r^2-1)\,\partial^h_{xz;z}, \\ \mathrm{H}^h_{zx} &\triangleq -(r^2-1)\,\partial^h_{xz;x}, & \mathrm{H}^h_{zz} &\triangleq -k_S^2 - \partial^h_{xx} - r^2\,\partial^h_{zz}, \end{aligned} \tag{7}$$

and is determined by the particular choices for the second-order differential operators ∂^h_{xx}, $\partial^h_{xz;x}$, $\partial^h_{xz;z}$, ∂^h_{zz}. A classical 9-point finite-difference scheme by Kelly et al. [19] is summarized by

$$\partial^h_{xx} \triangleq \frac{1}{h^2}\begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \partial^h_{xz;x} = \partial^h_{xz;z} \triangleq \frac{1}{4h^2}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}, \quad \partial^h_{zz} \triangleq \frac{1}{h^2}\begin{bmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{bmatrix}. \tag{8}$$

The discretized system reads

$$\begin{cases} -k_S^2 h^2\,u^x_{m,n} - r^2\left(u^x_{m+1,n} - 2\,u^x_{m,n} + u^x_{m-1,n}\right) - \left(u^x_{m,n+1} - 2\,u^x_{m,n} + u^x_{m,n-1}\right) \\ \qquad\qquad - \dfrac{1}{4}(r^2-1)\left(u^z_{m+1,n+1} - u^z_{m+1,n-1} - u^z_{m-1,n+1} + u^z_{m-1,n-1}\right) = h^2\,g^x_{m,n} \\ -k_S^2 h^2\,u^z_{m,n} - \left(u^z_{m+1,n} - 2\,u^z_{m,n} + u^z_{m-1,n}\right) - r^2\left(u^z_{m,n+1} - 2\,u^z_{m,n} + u^z_{m,n-1}\right) \\ \qquad\qquad - \dfrac{1}{4}(r^2-1)\left(u^x_{m+1,n+1} - u^x_{m+1,n-1} - u^x_{m-1,n+1} + u^x_{m-1,n-1}\right) = h^2\,g^z_{m,n}. \end{cases} \tag{9}$$



(a) Legend (b) Regular grid (c) Staggered grid

Fig. 1. Discretization grid according to (b) Kelly's and (c) Virieux's formulation. For the staggered-grid scheme, the elastic wavefields are defined on different grids. The grid points where these wavefield are defined, are denoted by different symbols according to (a). Note that z points downwards.

Another popular discretization scheme is based on a staggered-grid formulation, originally developed for the first-order wave equation by Virieux [20]. According to this scheme, the particle displacement components and the stress fields $\tau^{ab} = \sum_{cd} C^{ab,cd}\partial_c u^d$, $C^{ab,cd}$ being the stiffness tensor, are discretized on different grids that are shifted with respect to G^h. More specifically, the vertical component u^z is defined on the grid $G^h + (0, h/2)$, whereas the horizontal component u^x is positioned on $G^h + (h/2, 0)$. The stress components τ^{xx} and τ^{zz} are defined on the regular grid G^h, while τ^{xz} is on $G^h + (h/2, h/2)$. This configuration, pictured in Fig. 1, leads to the finite-difference scheme:

$$\begin{cases} -k_S^2 h^2 u^x_{m+\frac{1}{2},n} - r^2\left(u^x_{m+\frac{3}{2},n} - 2u^x_{m+\frac{1}{2},n} + u^x_{m-\frac{1}{2},n}\right) - \left(u^x_{m+\frac{1}{2},n+1} - 2u^x_{m+\frac{1}{2},n} + u^x_{m+\frac{1}{2},n-1}\right) \\ \qquad\qquad - (r^2-1)\left(u^z_{m+1,n+\frac{1}{2}} - u^z_{m+1,n-\frac{1}{2}} - u^z_{m,n+\frac{1}{2}} + u^z_{m,n-\frac{1}{2}}\right) = h^2 g^x_{m+\frac{1}{2},n} \\ -k_S^2 h^2 u^z_{m,n+\frac{1}{2}} - \left(u^z_{m+1,n+\frac{1}{2}} - 2u^z_{m,n+\frac{1}{2}} + u^z_{m-1,n+\frac{1}{2}}\right) - r^2\left(u^z_{m,n+\frac{3}{2}} - 2u^z_{m,n+\frac{1}{2}} + u^z_{m,n-\frac{1}{2}}\right) \\ \qquad\qquad - (r^2-1)\left(u^x_{m+\frac{1}{2},n+1} - u^x_{m-\frac{1}{2},n+1} - u^x_{m+\frac{1}{2},n} + u^x_{m-\frac{1}{2},n}\right) = h^2 g^z_{m,n+\frac{1}{2}}. \end{cases} \tag{10}$$

If we define the shifted particle displacement by $\bar{u}^x_{m,n} = u^x_{m+1/2,n}$ and $\bar{u}^z_{m,n} = u^z_{m,n+1/2}$, the stencil representation (7) acting on $\bar{u}^x$ and $\bar{u}^z$ can then be summarized by

$$\begin{aligned} \partial^h_{xx} &\triangleq \frac{1}{h^2}\begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix}, & \partial^h_{xz;x} &\triangleq \frac{1}{h^2}\begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ -1 & 1 & 0 \end{bmatrix}, \\ \partial^h_{xz;z} &\triangleq \frac{1}{h^2}\begin{bmatrix} 0 & 1 & -1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, & \partial^h_{zz} &\triangleq \frac{1}{h^2}\begin{bmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{bmatrix}. \end{aligned} \tag{11}$$

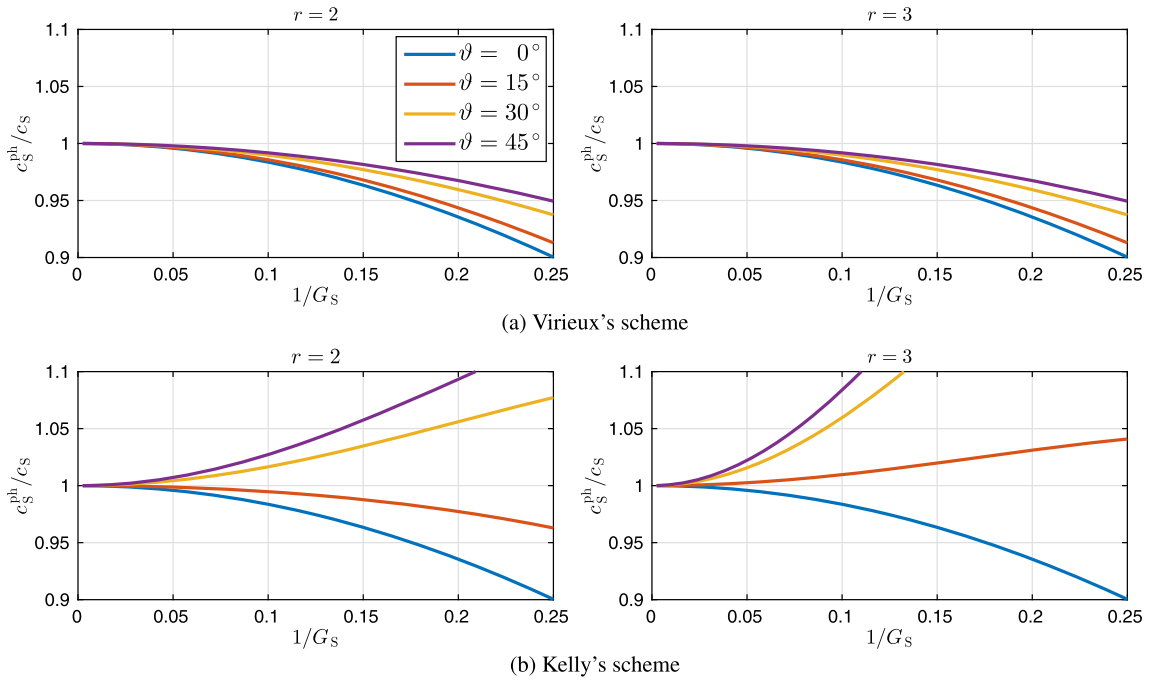
The difference between the mixed-derivative stencils $\partial^h_{xz;x}$ and $\partial^h_{xz;z}$ arises from the definition of the particle components on different grids. In the following sections, we will drop the notation $\bar{\mathbf{u}}$ and use $\mathbf{u}$ instead, to indicate the shifted discretized wavefield when dealing with the staggered-grid scheme.

The numerical dispersion properties of the staggered-grid scheme are significantly better than Kelly's and the S-wave phase and group velocities are independent of the velocity ratio. Indeed, the accuracy of Kelly's scheme gets worse for increasing values of r (Fig. 2). These numerical properties are determined by how the scheme realizes, in its discrete version, the decoupling of P- and S-wave modes in a homogeneous medium. The staggered-grid scheme does this 'consistently', meaning that the second-order stencils (11) can be derived from forward and backward first-order difference operators, e.g.,

$$\partial^h_{x+} \triangleq \frac{1}{h}\begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \partial^h_{x-} \triangleq \frac{1}{h}\begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{12}$$

and

$$\partial^h_{xx} = \partial^h_{x+} * \partial^h_{x-}, \quad \partial^h_{xz;x} = \partial^h_{x-} * \partial^h_{z+}, \quad \partial^h_{xz;z} = \partial^h_{x+} * \partial^h_{z-}, \quad \partial^h_{zz} = \partial^h_{z+} * \partial^h_{z-}, \tag{13}$$



(a) Virieux's scheme

(b) Kelly's scheme

Fig. 2. Numerical S-wave phase velocity ratios c_S^{ph}/c_S for (a) Virieux's and (b) Kelly's scheme as a function of the inverse of G_S (number of points per S-wave wavelength) and propagation angles ϑ. The numerical dispersion properties of Kelly's scheme become worse for increasing velocity ratios r, while for the staggered-grid scheme, they remain independent of r.

where '$*$' is the convolution operation. As a consequence, the P- and S-wave potentials satisfy a discretized version of the Helmholtz equation, defined by the same stencil in (11), and the S-wave phase velocity is therefore independent of r. Kelly's scheme does not have this property.

In principle, better options than the one presented here are available in terms of numerical accuracy. In [21], 9-point compact stencils are specifically designed to minimize numerical dispersion by averaging the differential operator with its rotation, along with lumped mass matrix techniques. A similar rationale has been applied by [22–24] to obtain very accurate 25-point finite-difference and finite-element schemes. The compactness of these stencils is intended to reduce the fill-in of factorization methods such as LU decomposition. It should be noted that this is not necessary when employing iterative methods and one can easily adopt higher-order versions of Kelly's and Virieux's schemes with wider stencils. This will, however, affect the behaviour of the multigrid solver.

In this paper, we will discuss multigrid with the staggered-grid scheme (10) in mind. Even though the 'unphysical' coupling of Kelly's scheme is more advantageous for multigrid applications, as we will see later, its use is questionable due to the inaccuracy of the scheme. If there are no acoustic regions ($\mu = 0$) in the medium, multigrid performs comparably for the two schemes. Generally speaking, the theory and methods developed in the following sections can be straightforwardly applied to any finite-difference or finite-element discretization of the wave equation on structured meshes.

3. A 'shifted-Laplacian' preconditioner for the elastic wave equation

Given a certain discretization of the wave equation system (4),

$$\mathbf{H}^h \mathbf{u}^h = \mathbf{g}^h, \quad \mathbf{H}^h = -k_S^2 \mathbf{I} - \boldsymbol{\mathcal{D}}^h, \tag{14}$$

we introduce the following preconditioning

$$\mathbf{H}^h \left(\mathbf{H}_\beta^h\right)^{-1} \tilde{\mathbf{u}}^h = \mathbf{g}^h, \quad \mathbf{H}_\beta^h = -k_S^2 (1 - \beta\,\mathrm{i})\,\mathbf{I} - \boldsymbol{\mathcal{D}}^h, \tag{15}$$

based on the very same ideas presented by [10,11]. From a physical point of view, the linear system $\mathbf{H}_\beta^h$ corresponds to a damped version of the elastic wave equation. We begin a preliminary analysis of this preconditioning by testing the spectral properties of the system (15), which, ideally, requires the inverse of the damped system. A few iterations of multigrid will achieve only an approximation of this inverse, thus, the following analysis assumes idealized circumstances. We make use of local-mode analysis [25–27], which will be discussed in more detail in Section 4.2.1. Essentially, it consists in a diagonal

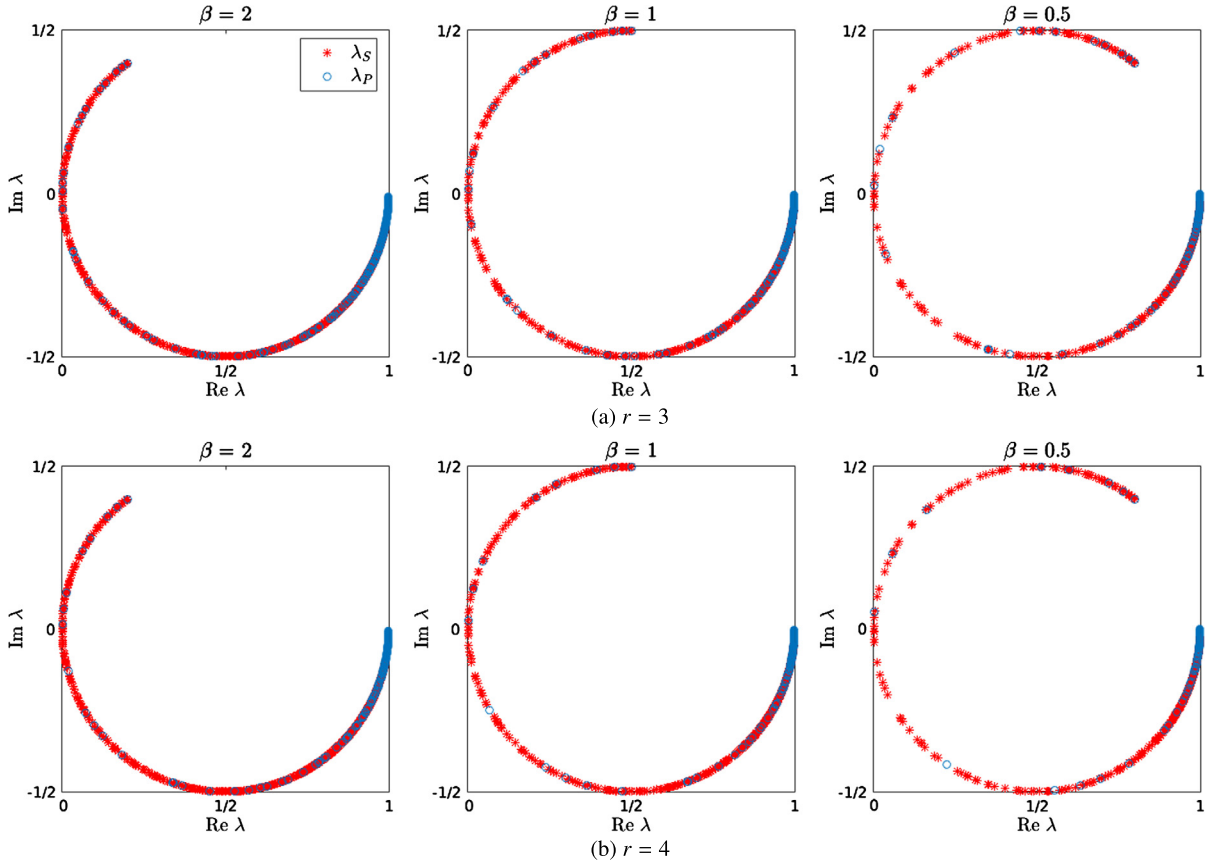


Fig. 3. Eigenvalue distribution $\lambda^h_{\mathrm{P},0}/\lambda^h_{\mathrm{P},\beta}$, $\lambda^h_{\mathrm{S},0}/\lambda^h_{\mathrm{S},\beta}$ of the preconditioned system (15) for different values of wave damping β and velocity ratio: (a) $r=3$, (b) $r=4$. The quantity $k_S h$ is fixed to the value $\pi/5$ which corresponds to a sampling of 10 points per wavelength. The plots here displayed correspond to the staggered-grid scheme, but an analogous behaviour can be observed for the Kelly's scheme. Note that for the staggered-grid scheme, the S-wave eigenvalues are independent of r.

decomposition of the discrete system (14) by the Fourier transform and the canonical P-wave and S-wave decomposition. One can easily find that, for the staggered-grid scheme, the P- and S-wave eigenvalues of the damped system $\mathbf{H}^h_\beta$ are

$$\begin{aligned}\lambda^h_{\mathrm{P},\beta} &= -k_S^2(1-\beta\,\mathrm{i}) + \frac{4r^2}{h^2}\left[\sin\left(\frac{\theta_x}{2}\right)^2 + \sin\left(\frac{\theta_z}{2}\right)^2\right],\\ \lambda^h_{\mathrm{S},\beta} &= -k_S^2(1-\beta\,\mathrm{i}) + \frac{4}{h^2}\left[\sin\left(\frac{\theta_x}{2}\right)^2 + \sin\left(\frac{\theta_z}{2}\right)^2\right],\end{aligned} \tag{16}$$

and, therefore, the preconditioning transforms the eigenvalues into

$$\frac{\lambda^h_{\mathrm{P},0}}{\lambda^h_{\mathrm{P},\beta}} = \frac{k_S^2h^2 - 4r^2\left[\sin\left(\frac{\theta_x}{2}\right)^2 + \sin\left(\frac{\theta_z}{2}\right)^2\right]}{k_S^2h^2(1-\beta\,\mathrm{i}) - 4r^2\left[\sin\left(\frac{\theta_x}{2}\right)^2 + \sin\left(\frac{\theta_z}{2}\right)^2\right]},$$

$$\frac{\lambda^h_{\mathrm{S},0}}{\lambda^h_{\mathrm{S},\beta}} = \frac{k_S^2h^2 - 4\left[\sin\left(\frac{\theta_x}{2}\right)^2 + \sin\left(\frac{\theta_z}{2}\right)^2\right]}{k_S^2h^2(1-\beta\,\mathrm{i}) - 4\left[\sin\left(\frac{\theta_x}{2}\right)^2 + \sin\left(\frac{\theta_z}{2}\right)^2\right]}.$$

Note that the second ratio matches the first in the limit for $r \downarrow 1$. These eigenvalue ratios are distributed on the complex plane along a circle with centre $1/2$ and radius $1/2$, exactly as in the acoustic case (Fig. 3). The same configuration arises with the Kelly's scheme.

Similarly to the acoustic case, as β gets smaller, the eigenvalues move away from 0 and cluster around 1, as can be seen in Fig. 3. A clustered spectrum is a favourable setting for Krylov subspace methods. However, as it will be shown later, finding the approximate inverse of $\mathbf{H}_\beta^h$ by multigrid is more efficient when β is large. Thus, the choice of β is a trade-off between the convergence rate of the multigrid solution of the damped system $\mathbf{H}_\beta^h$ and the desired spectral properties of the preconditioned system (15). In the elastic case, the multigrid efficiency will be also be affected by the velocity ratio r. We should point out that 'complex shifting' or added damping for the elastic system of equations has a similar effect on the corresponding spectra as with the acoustic equation. This fact provides a case for a successful application of multigrid preconditioning to the elastic wave equation and will be validated experimentally in Section 5. There, we will also consider the effect of the inexact multigrid approximation of the preconditioner. Note that our approach is mainly empirical. To the best of our knowledge, a rigorous theory for the convergence of shifted-Laplacian methods is presently lacking.

Recently, some research has been devoted to the combination of multigrid with deflation-based preconditioning [28,29]. The basic idea is to design a further preconditioner which clusters the eigenvalues in Fig. 3 towards 1. We refer to [30] and [31] for a discussion on this matter. We expect that the same idea can be applied successfully to the elastic case, but in this paper we will focus solely on the multigrid preconditioning (15).

4. Analysis of multigrid applied to the damped elastic wave equation

In the previous section, we studied preconditioning by the damped wave equation and assumed that the associated linear system could be solved exactly. Since we intend to obtain only an approximation of its solution by multigrid iterations, the previous analysis should be adapted accordingly. In this section, we will focus on the application of multigrid to the damped wave equation. The starting point is the previous work of [11,10], valid for the acoustic case. In order to assess the performance of multigrid, the classical tool of local-mode analysis was considered in [11,10]. For the elastic case, we will have to apply this method to a linear system, for which the theory can be found in [26, e.g.]. To validate the predictions of the following local-mode analysis, we will compare them to numerical experiments.

Multigrid is a well-known numerical technique for solving certain classes of partial differential equations [27,26, e.g.]. It exploits the multiscale behaviour of classical solvers by considering a hierarchy of discretization grids. Multigrid relies on two fundamental ingredients: smoothing and a coarse-grid correction. Given a linear system $\mathbf{H}^h$, which might arise from the discretization of a partial differential equation on the grid G^h, and starting from an approximation $\mathbf{u}_0^h$ of the solution $\mathbf{u}^h$, smoothing consists in reducing the high-frequency or short-wavelength components of the error $\mathbf{e}_0^h = \mathbf{u}^h - \mathbf{u}_0^h$. This is a typical trait of many elementary iterative methods, such as Jacobi and Gauss–Seidel, that can easily remove high-frequency solution errors but have difficulty with the low-frequency components. However, once a new approximation $\tilde{\mathbf{u}}_0^h$ is computed for which the error $\tilde{\mathbf{e}}_0^h$ is smooth and satisfies the error equation $\mathbf{H}^h \tilde{\mathbf{e}}_0^h = \tilde{\mathbf{r}}_0^h$ (where $\tilde{\mathbf{r}}_0^h$ is the corresponding residual), the fine-grid problem can be approximated on a coarser grid: $\mathbf{H}^{2h} \tilde{\mathbf{e}}_0^{2h} = \tilde{\mathbf{r}}_0^{2h}$. The solution of the coarse-grid problem is projected or prolongated back to the fine grid to give the so-called coarse-grid correction. Since 'low frequencies' on the fine grid become 'high frequencies' on the coarse grid after restriction, the same idea can be exploited recursively to solve for the coarse-grid problem.

4.1. Multigrid for the acoustic wave equation

Multigrid is typically effective for positive definite linear systems, for example those arising from the discretization of elliptic partial differential equations, and it is known to behave poorly for the undamped wave equation. The reasons are discussed in some detail in [14] and [10] and the same conclusions hold for the elastic case. Basically, classical smoothers like point-Jacobi will severely amplify the long wavelength components of the error on 'intermediate' grid levels, effectively hampering the performance of multigrid. Another issue concerns the coarse-grid correction: some eigenvalues corresponding to a certain discretization level h might undergo a change of sign on the coarser grid G^{2h}. It can be shown, as in [14], that the coarse-grid correction corresponding to this particular component will be updated in the 'wrong' direction. Multigrid is thus unable to solve for the space associated with these eigenvalues. When damping with β is introduced, the eigenvalues of $\mathbf{H}_\beta^h$ are shifted into the complex plane and multigrid can be effectively applied [10].

4.2. Multigrid for the elastic wave equation

In this section, we will show that multigrid can as well be applied to the damped elastic wave equation. We start with a brief introduction on local-mode analysis — the main quantitative tool we will use to assess the numerical behaviour of multigrid. We then proceed by discussing the performance of point- and line-Jacobi smoothers and finally by comparing estimated convergence factors from n-grid analysis to the values obtained by numerical experiments.

4.2.1. Introduction on local-mode analysis for the elastic wave equation

Local-mode analysis is a predictive tool that can be used to effectively measure multigrid performance. In essence, it is the spectral decomposition of the linear operators involved on the space of scalar Fourier grid functions $\varphi(\boldsymbol{\theta})$ defined by

$$\varphi(\boldsymbol{\theta})_{m,n} = \exp[\,\mathrm{i}\,(\theta_x m + \theta_z n)\,], \qquad \boldsymbol{\theta} = (\theta_x, \theta_z) \in [-\pi, \pi]^2. \tag{17}$$

For simplicity, we will only treat the 2-D case. This decomposition is useful in estimating how the various Fourier components of the error of a particular estimate of the solution are reduced after each operation that is part of a multigrid cycle. Since linear operators like $\mathbf{H}_\beta^h$ are defined by a spatially compact stencil, and its action involves 'local' operations, the mode analysis is arguably still valid by setting homogeneous physical parameters in the definition of the discretized linear operators and considering unbounded grids G^h or periodic boundary conditions. The discussion that follows is made under these assumptions. We refer to [26] for a more extensive introduction on local-mode analysis of linear systems. In this section, we provide a succinct and formal definition of smoothing and n-grid analysis [32] with emphasis on the wave equation and its discretization by the staggered-grid scheme (11).

We start with a formal Fourier analysis of the elastic wave equation operator $\mathbf{H}_\beta^h$ acting on the space of grid function pairs $\mathbb{L}^2(\mathrm{G}^h) \times \mathbb{L}^2(\mathrm{G}^h)$, where $\mathbb{L}^2(\mathrm{G}^h)$ is the set of square-summable functions. Given a vector $\mathbf{u}_0 \in \mathbb{C}^2$, we define the Fourier symbol $\hat{\mathbf{H}}_\beta^h$ as:

$$\mathbf{H}_\beta^h \left[\varphi(\boldsymbol{\theta})\, \mathbf{u}_0\right] = \hat{\mathbf{H}}_\beta^h(\boldsymbol{\theta})\, \mathbf{u}_0, \tag{18}$$

which, in other words, is the 'eigenvalue matrix' of the Fourier grid function. Since $\mathbf{H}_\beta^h$ is a 2×2-block matrix, $\hat{\mathbf{H}}_\beta^h(\boldsymbol{\theta})$ is a 2×2 matrix. The eigenvalues of $\hat{\mathbf{H}}_\beta^h$ can be found by a standard P- and S-wave diagonal decomposition:

$$\hat{\mathbf{H}}_\beta^h(\boldsymbol{\theta}) = \hat{\mathbf{M}}^h(\boldsymbol{\theta})\, \hat{\boldsymbol{\Lambda}}_\beta^h(\boldsymbol{\theta})\, \hat{\mathbf{M}}^h(\boldsymbol{\theta})^*, \tag{19}$$

where the asterisk denotes conjugate transposition, $\hat{\mathbf{M}}^h(\boldsymbol{\theta}) = [\hat{\mathbf{u}}_{\mathrm{P}}^h(\boldsymbol{\theta}),\, \hat{\mathbf{u}}_{\mathrm{S}}^h(\boldsymbol{\theta})]$ is the matrix of P- and S-wave eigenvectors, and $\hat{\boldsymbol{\Lambda}}_\beta^h(\boldsymbol{\theta}) = \mathrm{diag}([\lambda_{\mathrm{P},\beta}^h(\boldsymbol{\theta}),\, \lambda_{\mathrm{S},\beta}^h(\boldsymbol{\theta})])$ is the eigenvalue matrix. As already noted in Section 2, the wave potentials of the staggered-grid scheme are defined by a discretization of the P- and S-wave Helmholtz equation, which is consistent with the stencils (11). Their Fourier transforms are

$$\hat{\partial}_{xx}^h(\boldsymbol{\theta}) = -\frac{4}{h^2} \sin\left(\frac{\theta_x}{2}\right)^2, \qquad \hat{\partial}_{xz;x}^h(\boldsymbol{\theta}) = -\frac{4}{h^2} e^{\mathrm{i}(-\theta_x+\theta_z)/2} \sin\left(\frac{\theta_x}{2}\right) \sin\left(\frac{\theta_z}{2}\right),$$

$$\hat{\partial}_{xz;z}^h(\boldsymbol{\theta}) = -\frac{4}{h^2} e^{\mathrm{i}(\theta_x-\theta_z)/2} \sin\left(\frac{\theta_x}{2}\right) \sin\left(\frac{\theta_z}{2}\right), \qquad \hat{\partial}_{zz}^h(\boldsymbol{\theta}) = -\frac{4}{h^2} \sin\left(\frac{\theta_z}{2}\right)^2.$$

It is easily found that the eigenvectors associated with the eigenvalues $\lambda_{\mathrm{P},\beta}^h$, $\lambda_{\mathrm{S},\beta}^h$ in (16), apart from scaling factors, are

$$\hat{\mathbf{u}}_{\mathrm{P}}^h(\boldsymbol{\theta}) = \frac{2\mathrm{i}}{h}\left[e^{-\mathrm{i}\theta_x/2} \sin\left(\frac{\theta_x}{2}\right),\, e^{-\mathrm{i}\theta_z/2} \sin\left(\frac{\theta_z}{2}\right)\right], \quad \hat{\mathbf{u}}_{\mathrm{S}}^h(\boldsymbol{\theta}) = \frac{2\mathrm{i}}{h}\left[-e^{\mathrm{i}\theta_z/2} \sin\left(\frac{\theta_z}{2}\right),\, e^{\mathrm{i}\theta_x/2} \sin\left(\frac{\theta_x}{2}\right)\right]. \tag{20}$$

In the expressions of $\hat{\mathbf{u}}_{\mathrm{P}}^h$ and $\hat{\mathbf{u}}_{\mathrm{S}}^h$, one may recognize the Fourier symbols of the gradient and curl operators, discretized by the centred schemes (12) underlying the stencils (11):

$$\hat{\mathbf{u}}_{\mathrm{P}}^h(\boldsymbol{\theta}) = \left[\hat{\partial}_{x-}^h(\boldsymbol{\theta}),\, \hat{\partial}_{z-}^h(\boldsymbol{\theta})\right] \approx \hat{\nabla}(\boldsymbol{\theta}), \qquad \hat{\mathbf{u}}_{\mathrm{S}}^h(\boldsymbol{\theta}) = \left[-\hat{\partial}_{z+}^h(\boldsymbol{\theta}),\, \hat{\partial}_{x+}^h(\boldsymbol{\theta})\right] \approx \hat{\mathrm{curl}}(\boldsymbol{\theta}). \tag{21}$$

The Fourier analysis of the operator $\mathbf{H}_\beta^h$ allows us to analyse the properties of many smoothing schemes. A simple form of smoothing, which proved to be effective for the damped acoustic wave equation, is given by the classical Jacobi scheme. Given a certain initial guess $\mathbf{u}_0^h$ of the solution of $\mathbf{H}_\beta^h \mathbf{u}^h = \mathbf{f}^h$, the Jacobi iteration consists of the update

$$\mathbf{u}_1^h = \mathbf{u}_0^h + \eta\, (\mathbf{D}_\beta^h)^{-1}\, \mathbf{r}_0^h, \qquad \mathbf{r}_0^h = \mathbf{f}^h - \mathbf{H}_\beta^h\, \mathbf{u}_0^h, \tag{22}$$

where $\mathbf{D}_\beta^h$ is the diagonal of $\mathbf{H}_\beta^h$ and η a relaxation parameter that needs to be adjusted for effective smoothing. To understand how the initial error $\mathbf{e}_0^h = \mathbf{u}_0^h - \mathbf{u}^h$ is reduced, we write

$$\mathbf{e}_1^h = \mathbf{S}^h\, \mathbf{e}_0^h, \qquad \mathbf{S}^h = \mathbf{I}^h - \eta\, (\mathbf{D}_\beta^h)^{-1}\, \mathbf{H}_\beta^h, \tag{23}$$

where $\mathbf{S}^h$ is the point-Jacobi iteration matrix. The local-mode analysis of $\mathbf{S}^h$ provides a theoretical estimate of how the different wavelength components of the error are damped by the smoothing. Since smoothing is designed to solve for the short-wavelength components, corresponding to

$$|\theta_x| > \frac{\pi}{2} \quad \text{or} \quad |\theta_z| > \frac{\pi}{2},$$

we define the smoothing factor as:

$$\mu_s^h = \max\left\{ |\lambda| : \lambda \text{ eigenvalue of } \hat{\mathbf{S}}^h(\boldsymbol{\theta}),\ |\theta_x| > \frac{\pi}{2} \quad \text{or} \quad |\theta_z| > \frac{\pi}{2} \right\}. \tag{24}$$

Clearly, besides smoothing, a comprehensive analysis of multigrid requires the study of the coarse-grid correction over many grid levels. Any number n of grid levels could be considered, but usually, $n \leq 3$ suffices to analyze multigrid schemes. For the elastic wave equation, we will consider $n = 4$. For ease of exposition, we will summarize the theory of n-grid analysis for the simple two-level case, $n = 2$. We refer to [32] for a more general treatment.

The operations of two-level multigrid are

$$\mathbf{u}_1^h = \mathbf{u}_0^h + \eta\,(\mathbf{D}_\beta^h)^{-1}\,\mathbf{r}_0^h \qquad \text{(pre-smoothing, e.g., Jacobi)},$$

$$\mathbf{r}_1^{2h} = \mathbf{R}_h^{2h}\,\mathbf{r}_1^h \qquad \text{(restriction of the residual to the coarse grid)},$$

$$\mathbf{H}_\beta^{2h}\,\mathbf{e}_1^{2h} = \mathbf{r}_1^{2h} \qquad \text{(solution of the coarse-grid error equation)},$$

$$\tilde{\mathbf{e}}_1^h = \mathbf{P}_{2h}^h\,\mathbf{e}_1^{2h} \qquad \text{(interpolation of the solution error onto the fine grid)},$$

$$\mathbf{u}_2^h = \mathbf{u}_1^h + \tilde{\mathbf{e}}_1^h \qquad \text{(coarse-grid correction)},$$

$$\mathbf{u}_3^h = \mathbf{u}_2^h + \eta\,(\mathbf{D}_\beta^h)^{-1}\,\mathbf{r}_2^h \qquad \text{(post-smoothing, e.g., Jacobi)}.$$

The grid transfer of the functions of $\mathbb{L}^2(\mathrm{G}^h)$ and $\mathbb{L}^2(\mathrm{G}^{2h})$ is managed by the restriction operator $\mathbf{R}_h^{2h}$ and prolongation operator $\mathbf{P}_{2h}^h$. A natural choice would be

$$\mathbf{R}_h^{2h} = \begin{pmatrix} \mathrm{R}_h^{2h} & 0 \\ 0 & \mathrm{R}_h^{2h} \end{pmatrix}, \qquad \mathbf{P}_{2h}^h = \begin{pmatrix} \mathrm{P}_{2h}^h & 0 \\ 0 & \mathrm{P}_{2h}^h \end{pmatrix}, \tag{25}$$

where R_h^{2h} and P_{2h}^h correspond to scalar restriction and prolongation operators. We assumed standard coarsening. The coarse-grid linear system $\mathbf{H}_\beta^{2h}$ can be obtained from a direct discretization of the wave equation on the grid G^{2h} or from the Galerkin condition

$$\mathbf{H}_\beta^{2h} = \mathbf{R}_h^{2h}\,\mathbf{H}_\beta^h\,\mathbf{P}_{2h}^h. \tag{26}$$

When the restriction and prolongation operators are chosen as in (25), it corresponds to a coarse-grid block operator where each block satisfies the scalar Galerkin condition, e.g., $\mathrm{H}_{xx,\beta}^{2h} = \mathrm{R}_h^{2h}\,\mathrm{H}_{xx,\beta}^h\,\mathrm{P}_{2h}^h$ for the first block of $\mathbf{H}_\beta^{2h}$. Through the two-level multigrid algorithm, the error is transformed according to

$$\mathbf{e}_3^h = \mathbf{T}^{h,2h}\,\mathbf{e}_0^h, \qquad \mathbf{T}^{h,2h} = \left(\mathbf{S}^h\right)^{\nu_2}\mathbf{K}^{h,2h}\left(\mathbf{S}^h\right)^{\nu_1}, \quad \mathbf{K}^{h,2h} = \mathbf{I}^h - \mathbf{P}_{2h}^h\left(\mathbf{H}_\beta^{2h}\right)^{-1}\mathbf{R}_h^{2h}\,\mathbf{H}^h, \tag{27}$$

where ν_1 and ν_2 are, respectively, the number of pre and post-smoothing iterations. In the smoothing analysis, we exploit the diagonalization of the operator $\mathbf{S}^h$ over the Fourier grid functions. The complication of the two-level analysis stems from the fact that the coarse-grid operator $\mathbf{K}^{h,2h}$ mixes wavelength components and, therefore, does not diagonalize over the same basis. However, it admits a block diagonal decomposition over the direct sum of the following 8-dimensional subspaces

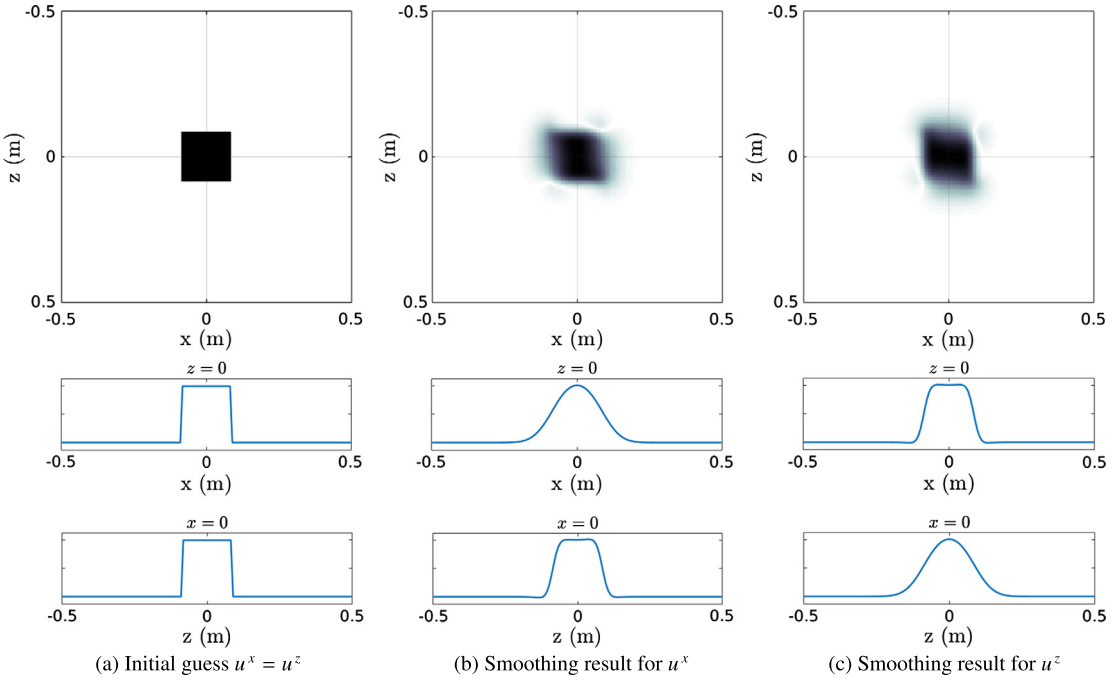
$$F_2[\![\boldsymbol{\theta}_0]\!] = \operatorname{span}\left\{ \varphi(\boldsymbol{\theta})\,\mathbf{u}_0 :\ \boldsymbol{\theta} \equiv \boldsymbol{\theta}_0 \bmod \pi,\ \mathbf{u}_0 \in \mathbb{C}^2 \right\}, \tag{28}$$

parametrized by the π modulo classes $[\![\boldsymbol{\theta}_0]\!]$, where $\boldsymbol{\theta}_0 \in [-\pi/2, \pi/2]^2$. For a 2-D analysis, these classes consist of four elements:

$$[\![\boldsymbol{\theta}_0]\!] = \left\{ \boldsymbol{\theta}_0^0 = \boldsymbol{\theta}_0,\ \boldsymbol{\theta}_0^1,\ \boldsymbol{\theta}_0^2,\ \boldsymbol{\theta}_0^3 \right\}.$$

It can be easily proven that $F_2[\![\boldsymbol{\theta}_0]\!]$ is an invariant subspace for the two-grid operator $\mathbf{T}^{h,2h}$, and the spectral analysis of two-grid multigrid boils down to the matrix representation:

$$\mathbf{T}^{h,2h}\,\varphi(\boldsymbol{\theta}_0^l)\,\mathbf{u}_0 = \sum_{k=0}^{3} \hat{\mathbf{T}}^{h,2h}(\boldsymbol{\theta}_0^k, \boldsymbol{\theta}_0^l)\,\varphi(\boldsymbol{\theta}_0^k)\,\mathbf{u}_0. \tag{29}$$



(a) Initial guess $u^x = u^z$ (b) Smoothing result for u^x (c) Smoothing result for u^z

Fig. 4. Effect of point-Jacobi smoothing applied to the elastic case for the homogeneous problem $\mathbf{H}^h_\beta\,\mathbf{u}^h = \mathbf{0}$ for $\beta = 2$, $k_S h = \pi/10$ (corresponding to 20 points per wavelength) and $r = 3$. The damping factor is $\eta = 0.5$. The initial guess $\mathbf{u}^h_0$, pictured in (a), consists of a unit function whose support is a square set, for both components. The results after 100 iterations show that point-Jacobi is not effectively smoothing in the vertical and horizontal directions, respectively, for the (b) horizontal and (c) vertical particle displacement components.

The 4×4 2×2-block matrix $\hat{\mathbf{T}}^{h,\,2h}[\![\boldsymbol{\theta}_0]\!]_{k,l} = \hat{\mathbf{T}}^{h,\,2h}(\boldsymbol{\theta}^k_0, \boldsymbol{\theta}^l_0)$ is obviously equivalent to an 8×8 matrix. Not only different wavelength components are intermixed by the two-grid operator, as in the acoustic (scalar) case, but also P- and S-wave modes. The two-grid convergence factor is then defined by

$$\mu^h_2 = \max\left\{ |\lambda| :\ \lambda \text{ eigenvalue of } \hat{\mathbf{T}}^{h,\,2h}[\![\boldsymbol{\theta}_0]\!],\ \ \boldsymbol{\theta}_0 \in [-\pi/2,\,\pi/2]^2 \right\}. \tag{30}$$

By induction, the same kind of analysis can be generalized to define n-grid convergence factors μ^h_n [32].

4.2.2. Point- and line-relaxation schemes

In the previous section, we discussed smoothing analysis by considering the elementary point-Jacobi smoothing. As already mentioned, this is an effective choice in the context of shifted-Laplacian multigrid for the acoustic case. However, the simple numerical example in Fig. 4 shows that point-Jacobi is an anisotropic smoother for the elastic case. This observation can be confirmed theoretically by applying the techniques of the previous section. First, we denote the diagonal operator $\mathbf{D}^h_\beta = \operatorname{diag}\mathbf{H}^h_\beta$, in stencil notation, by

$$\mathrm{D}^h_{xx,\,\beta} \triangleq \begin{bmatrix} 0 & 0 & 0 \\ 0 & \mathrm{H}^h_{xx,\,\beta}(0,0) & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad D^h_{zz,\,\beta} \triangleq \begin{bmatrix} 0 & 0 & 0 \\ 0 & \mathrm{H}^h_{zz,\,\beta}(0,0) & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{31}$$

and $\mathrm{D}^h_{xz,\,\beta} = \mathrm{D}^h_{zx,\,\beta} \triangleq \mathbf{0}$, where $\mathrm{H}^h_{xx,\,\beta}(0,0) = \mathrm{H}^h_{zz,\,\beta}(0,0) = -k_S^2(1-\beta\,\mathrm{i}) + 2(r^2+1)/h^2$. We then apply the P- and S-wave decomposition to the smoothing operator $\hat{\mathbf{S}}^h$ in (23). Its eigenvalues are

$$\begin{aligned} \lambda^{s,\,h}_{\mathrm{P},\,\beta}(\boldsymbol{\theta}) &= 1 - \eta\,\frac{k_S^2 h^2(1-\beta\,\mathrm{i}) - 4r^2\left[\sin\left(\frac{\theta_x}{2}\right)^2 + \sin\left(\frac{\theta_z}{2}\right)^2\right]}{k_S^2 h^2(1-\beta\,\mathrm{i}) - 2(r^2+1)}, \\ \lambda^{s,\,h}_{\mathrm{S},\,\beta}(\boldsymbol{\theta}) &= 1 - \eta\,\frac{k_S^2 h^2(1-\beta\,\mathrm{i}) - 4\left[\sin\left(\frac{\theta_x}{2}\right)^2 + \sin\left(\frac{\theta_z}{2}\right)^2\right]}{k_S^2 h^2(1-\beta\,\mathrm{i}) - 2(r^2+1)}. \end{aligned} \tag{32}$$

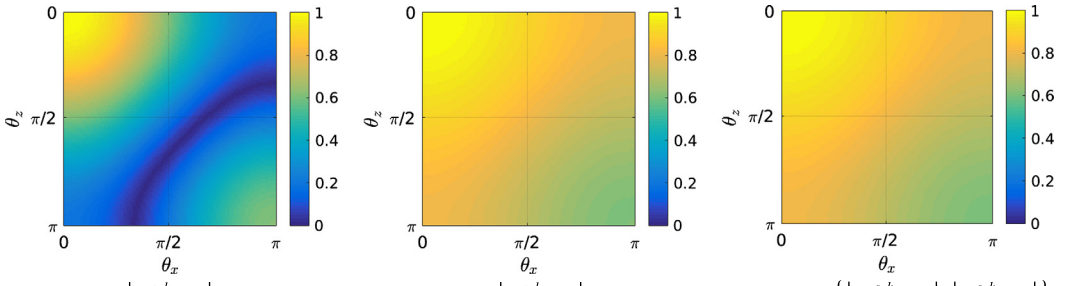


Fig. 5. Smoothing factors of point-Jacobi for the elastic wave equation as a function of the Fourier component relative to $\boldsymbol{\theta}$: (a) $\left|\lambda^{s,h}_{\mathrm{P},\beta}(\boldsymbol{\theta})\right|$, (b) $\left|\lambda^{s,h}_{\mathrm{S},\beta}(\boldsymbol{\theta})\right|$, (c) $\mu^h_s = \max\left\{\left|\lambda^{s,h}_{\mathrm{P},\beta}(\boldsymbol{\theta})\right|,\left|\lambda^{s,h}_{\mathrm{S},\beta}(\boldsymbol{\theta})\right|\right\}$, for $\beta = 2$, $k_S h = \pi/10$ (corresponding to 20 points per wavelength), $r = 2$ and $\eta = 0.5$.

In Fig. 5, we depict these factors as a function of the Fourier component relative to $\boldsymbol{\theta}$. From (32), it is easily found that

$$\lambda^{s,h}_{\mathrm{P},\beta}(\boldsymbol{\theta}) \approx 1 - 2\eta\left[\sin\left(\frac{\theta_x}{2}\right)^2 + \sin\left(\frac{\theta_z}{2}\right)^2\right], \quad \lambda^{s,h}_{\mathrm{S},\beta}(\boldsymbol{\theta}) \approx 1, \qquad \text{for } r \to +\infty. \tag{33}$$

As evident from (33), for large values of the velocity ratio, η can be tuned to only smooth P-wave modes, but the S-wave eigenvalue remains equal to 1 regardless of $\boldsymbol{\theta}$ and the relaxation factor η. Recalling the identity (21) and considering an initial error decomposed over the P- and S-wave eigenbasis (20) as $\hat{\mathbf{e}}^h_0(\boldsymbol{\theta}) = a\,\hat{\mathbf{u}}^h_P(\boldsymbol{\theta}) + b\,\hat{\mathbf{u}}^h_S(\boldsymbol{\theta})$, point-Jacobi transforms the error into

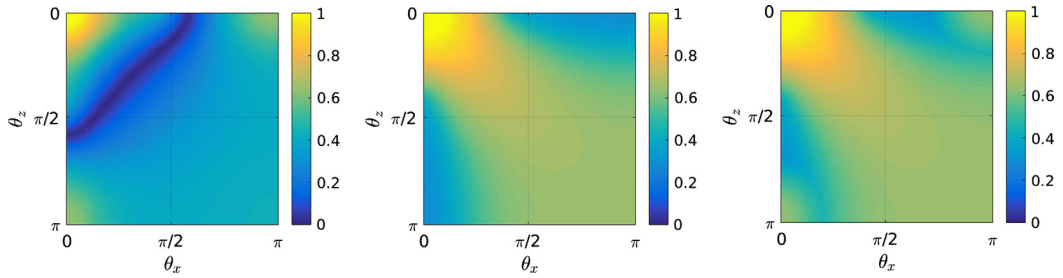
$$\hat{\mathbf{e}}^h_1(\boldsymbol{\theta}) = [\hat{\mathbf{S}}^h(\boldsymbol{\theta})]^\nu\,\hat{\mathbf{e}}^h_0(\boldsymbol{\theta}) = a\,[\lambda^{s,h}_{\mathrm{P},\beta}(\boldsymbol{\theta})]^\nu\,\hat{\mathbf{u}}^h_{\mathrm{P}}(\boldsymbol{\theta}) + b\,[\lambda^{s,h}_{\mathrm{S},\beta}(\boldsymbol{\theta})]^\nu\,\hat{\mathbf{u}}^h_{\mathrm{S}}(\boldsymbol{\theta}) \approx b\,\hat{\mathbf{u}}^h_{\mathrm{S}}(\boldsymbol{\theta}),$$

for a sufficiently large number of smoothing steps ν. Because of the inadequate smoothing of the S-wave components for high velocity ratios displayed by equation (33), only the S-wave mode component of $\hat{\mathbf{e}}^h_1(\boldsymbol{\theta})$ survives. From the expression of $\hat{\mathbf{u}}^h_{\mathrm{S}}(\boldsymbol{\theta})$ in (20), we readily observe that the horizontal component of $\hat{\mathbf{e}}^h_1(\boldsymbol{\theta})$ is effectively damped for $\theta_z \approx 0$, and the vertical component for $\theta_x \approx 0$. Thus, the horizontal and vertical components of $\hat{\mathbf{e}}^h_1(\boldsymbol{\theta})$ will show a rough profile along, respectively, the vertical and horizontal directions (as in Fig. 4).

We showed that point-Jacobi fails to solve for the short-wavelength components of the S-wave mode of the error, and this results in smoothing anisotropy. This is a well-known issue when, for example, dealing with regular grids defined by different horizontal and vertical step-lengths: $\mathrm{G}^{h_x,h_z} = h_x\mathbb{Z} \times h_z\mathbb{Z}$. Operators like anisotropic Helmholtz, (e.g., each of the blocks along the diagonal of $\mathbf{H}_\beta$ in (4)) are equivalent to the Helmholtz equation defined on an anisotropic grid, and therefore display the same behaviour. In the literature, the standard solutions for this type of shortcomings are: (*i*) semi-coarsening, (*ii*) line smoothing. Semi-coarsening, for example along the horizontal direction x, consists of the discretization hierarchy: $\mathrm{G}^{h_x,h_z} \to \mathrm{G}^{2h_x,h_z} \to \mathrm{G}^{4h_x,h_z} \to \ldots$. Semi-coarsening can also be applied simultaneously in multiple directions [33]. Even though this is, in principle, a valid approach, in this paper we stick to the second option of line smoothing.

A line smoother for the elastic wave equation. Line smoothing is based on the idea of treating the unknowns of the problem as a set of 'block variables', each block corresponding to the grid points positioned along a row or column of the grid G^h. The operators involved are then reinterpreted as acting on these variable sets as a block, rather than on individual grid points. This point of view concerns the way the inverse of an operator is approximated. For example, when the interaction between variables is completely neglected, namely when each blocks corresponds to a single grid-point, it leads to the approximation of the inverse of a linear system $\mathbf{H}^h_\beta$ by its diagonal inverse $(\mathbf{D}^h_\beta)^{-1}$, as in point-Jacobi smoothing (23). Based on the observations made about the anisotropy of point-Jacobi, we treat the wavefield variables $\mathbf{u}$ in the following way: the horizontal component u^x is considered as a collection of columns, while the vertical component u^z as a set of rows. The resulting block-diagonal operator of $\mathbf{H}^h_\beta$ can then be expressed in stencil notation as

$$\mathrm{D}^h_{xx,\beta} \triangleq \begin{bmatrix} 0 & 0 & 0 \\ \mathrm{H}^h_{xx,\beta}(-1,0) & \mathrm{H}^h_{xx,\beta}(0,0) & \mathrm{H}^h_{xx,\beta}(1,0) \\ 0 & 0 & 0 \end{bmatrix}, \qquad D^h_{zz,\beta} \triangleq \begin{bmatrix} 0 & \mathrm{H}^h_{zz,\beta}(0,-1) & 0 \\ 0 & \mathrm{H}^h_{zz,\beta}(0,\ 0) & 0 \\ 0 & \mathrm{H}^h_{zz,\beta}(0,\ 1) & 0 \end{bmatrix}, \tag{34}$$



(a) $\left|\lambda^{s,h}_{1,\beta}(\boldsymbol{\theta})\right|$ (b) $\left|\lambda^{s,h}_{2,\beta}(\boldsymbol{\theta})\right|$ (c) $\max\left\{\left|\lambda^{s,h}_{1,\beta}(\boldsymbol{\theta})\right|, \left|\lambda^{s,h}_{2,\beta}(\boldsymbol{\theta})\right|\right\}$

Fig. 6. Smoothing factors of line-Jacobi for the elastic wave equation as a function of the Fourier component relative to $\boldsymbol{\theta}$: (a) $\left|\lambda^{s\,h}_{1,\beta}(\boldsymbol{\theta})\right|$, (b) $\left|\lambda^{s\,h}_{2,\beta}(\boldsymbol{\theta})\right|$, (c) $\mu^{h}_{s} = \max\left\{\left|\lambda^{s\,h}_{1,\beta}(\boldsymbol{\theta})\right|, \left|\lambda^{s\,h}_{2,\beta}(\boldsymbol{\theta})\right|\right\}$, for $\beta = 2$, $k_S h = \pi/10$ (corresponding to 20 points per wavelength), $r = 2$ and $\eta = 0.8$.

Table 1
Smoothing factors μ^{h}_{s}, as defined in equation (24), for: (a) point- and (b) line-Jacobi smoothing with a staggered-grid discretization. The comparison is made for different grid levels, velocity ratios and direct discretization as well as Galerkin coarse-grid operators (defined by full weighting as restriction and bilinear interpolation as prolongation operator). For this numerical comparison, we let the damping factor $\beta = 2$ and S-wave wavenumber $k_S h = \pi/10$ (corresponding to 20 points per wavelength on the finest grid). For each of the four experiments, the relaxation parameters η were tuned to minimize the smoothing factors for the range of velocity ratios $r \in [1, 5]$ on each grid level h, $2h$, $4h$, $8h$, $16h$.

Smoothing factors

	Direct					Galerkin			
	$r=2$	$r=3$	$r=4$	$r=5$		$r=2$	$r=3$	$r=4$	$r=5$
h	0.90	0.95	0.97	0.98	h	0.90	0.95	0.97	0.98
$2h$	0.91	0.96	0.97	0.98	$2h$	0.82	0.91	0.95	0.97
$4h$	0.93	0.98	0.99	0.99	$4h$	0.87	0.94	0.97	0.98
$8h$	1.00	1.00	1.00	1.00	$8h$	1.00	1.00	1.00	1.00
$16h$	0.71	0.77	0.83	0.84	$16h$	0.93	0.91	0.83	0.94

(a) Point-Jacobi

	Direct					Galerkin			
	$r=2$	$r=3$	$r=4$	$r=5$		$r=2$	$r=3$	$r=4$	$r=5$
h	0.84	0.84	0.90	0.93	h	0.84	0.84	0.90	0.93
$2h$	0.92	0.92	0.92	0.94	$2h$	0.76	0.76	0.80	0.83
$4h$	0.75	0.87	0.93	0.95	$4h$	0.81	0.81	0.82	0.84
$8h$	0.80	0.94	0.98	1.00	$8h$	0.70	0.71	0.79	0.82
$16h$	0.58	0.66	0.86	0.97	$16h$	0.56	0.56	0.61	0.74

(b) Line-Jacobi

and $\mathrm{D}^{h}_{xz,\beta} = \mathrm{D}^{h}_{zx,\beta} \triangleq \mathbf{0}$. Line-Jacobi smoothing is then defined exactly as in (22), by replacing the point-wise diagonal operator with its line-wise counterpart (34). For the staggered-grid scheme, its Fourier symbol reads

$$\hat{\mathrm{D}}^{h}_{xx,\beta}(\boldsymbol{\theta}) = -k_S^2(1-\beta\,\mathrm{i}) + \frac{4r^2}{h^2}\sin\left(\frac{\theta_x}{2}\right)^2 + \frac{2}{h^2}, \qquad \hat{\mathrm{D}}^{h}_{zz,\beta}(\boldsymbol{\theta}) = -k_S^2(1-\beta\,\mathrm{i}) + \frac{4r^2}{h^2}\sin\left(\frac{\theta_z}{2}\right)^2 + \frac{2}{h^2}. \tag{35}$$

We should observe that the line-smoothing operator does not diagonalize over the same basis of gradient and curl operators (20), as in the case of the point smoother, meaning that the P- and S-wave components of the error are mixed. The components relative to this new eigenbasis are damped according to the eigenvalues $\lambda_{1,\beta}$, $\lambda_{2,\beta}$ shown in Fig. 6. The expressions for these eigenvalues and eigenvectors are lengthy and therefore omitted. A comparison of the smoothing factors in Fig. 5c and Fig. 6c demonstrates the effectiveness of this scheme: indeed, the smoothing factors of point-Jacobi take values close to 1 on the short wavelength region $\{|\theta_x| > \pi/2 \text{ or } |\theta_z| > \pi/2\}$ (Fig. 5c), while it is reasonably below 1 for line-Jacobi (Fig. 6c). For simplicity, we only have analyzed the behaviour of smoothing by direct discretization of the elastic wave equation on a certain grid level G^{h}. In Table 1, we report the results of our smoothing analysis for point and line-Jacobi and we observe that line smoothing is also a good choice for the Galerkin coarse-grid discretization (26).

From a computational point of view, line-Jacobi requires the solution of the linear system:

$$\mathbf{D}^{h}_{\beta}\,\mathbf{v}^{h} = \mathbf{f}^{h}, \qquad \mathbf{D}^{h}_{\beta} = \begin{pmatrix} \mathrm{D}^{h}_{xx,\beta} & 0 \\ 0 & \mathrm{D}^{h}_{zz,\beta} \end{pmatrix},$$

following the definition (34). This system can be decomposed in a series of independent tri-diagonal 1-D problems — horizontal-wise for $\mathrm{D}^h_{xx,\beta}$ or vertical-wise for $\mathrm{D}^h_{zz,\beta}$ — that can be solved efficiently with linear complexity (see Thomas algorithm in [34]), resulting in a well-suited scheme for parallel implementation.

Analysis of line-Jacobi for the limit case $r \to +\infty$. Even though we demonstrated the effectiveness of line smoothing over the point-wise counterpart for the staggered-grid scheme, its smoothing quality deteriorates when $r \to +\infty$, showing a behaviour similar to equation (33). Indeed, it can be proven that

$$\lambda^{s,h}_{1,\beta}(\boldsymbol{\theta}) \approx 1 - 2\eta, \quad \lambda^{s,h}_{2,\beta}(\boldsymbol{\theta}) \approx 1, \qquad \text{for } r \to +\infty. \tag{36}$$

The same analysis, however, shows more favourable properties for Kelly's finite-difference scheme (9). The eigenvalues related to point-Jacobi, in this case, are

$$\lambda^{s,h}_{\mathrm{P},\beta}(\boldsymbol{\theta}) \approx 1 - 2\eta \left[\sin\left(\frac{\theta_x}{2}\right)^2 + \sin\left(\frac{\theta_z}{2}\right)^2 - \sin\left(\frac{\theta_x}{2}\right)^2 \sin\left(\frac{\theta_z}{2}\right)^2 \right],$$

$$\lambda^{s,h}_{\mathrm{S},\beta}(\boldsymbol{\theta}) \approx 1 - 2\eta \sin\left(\frac{\theta_x}{2}\right)^2 \sin\left(\frac{\theta_z}{2}\right)^2$$

for $r \to +\infty$, with eigenvectors

$$\hat{\mathbf{u}}^h_{\mathrm{P}}(\boldsymbol{\theta}) \approx \left[\sin\left(\frac{\theta_x}{2}\right) \cos\left(\frac{\theta_z}{2}\right), \cos\left(\frac{\theta_x}{2}\right) \sin\left(\frac{\theta_z}{2}\right) \right],$$

$$\hat{\mathbf{u}}^h_{\mathrm{S}}(\boldsymbol{\theta}) \approx \left[-\cos\left(\frac{\theta_x}{2}\right) \sin\left(\frac{\theta_z}{2}\right), \sin\left(\frac{\theta_x}{2}\right) \cos\left(\frac{\theta_z}{2}\right) \right].$$

So far, we experience the same difficulties seen for point-Jacobi applied to the Virieux's scheme, namely the inefficient smoothing of the S-wave error modes (this is true whenever θ_x or θ_z are close to zero). For the same reasons discussed for the staggered-grid scheme, this produces anisotropic smoothing that can be eased by a line solver. A computation analogous to (36) shows that, for the line smoother, we have

$$\lambda^{s,h}_{\cdot,\beta}(\boldsymbol{\theta}) \approx 1 - \eta \left[1 \pm \cos\left(\frac{\theta_x}{2}\right) \cos\left(\frac{\theta_z}{2}\right) \right],$$

meaning that $1 - \eta \le \lambda^{s,h}_{1,\beta}(\boldsymbol{\theta}) \le 1 - \eta\,(1 - \sqrt{2}/2)$ and $1 - \eta\,(1 + \sqrt{2}/2) \le \lambda^{s,h}_{2,\beta}(\boldsymbol{\theta}) \le 1 - \eta$ for short-wavelength components and $\eta > 0$. Therefore, the relaxation parameter η can be chosen such that the smoothing factor is strictly less than 1, contrary to what is the case with the staggered-grid scheme. Now, the smoothing properties of the line relaxation scheme are acceptable, as will appear in the numerical experiments of Section 5.

Line relaxation for acoustic regions ($\mu = 0$). As a final remark, we analyse line-Jacobi for the acoustic case $\mu = 0$. Here, we will specifically consider the elastic formulation (5), adapted to the acoustic case. We point out that the smoothing schemes just presented for the staggered-grid finite-difference scheme cannot effectively handle acoustic regions inside the medium. Indeed, similarly to (36), the eigenvalues of the line-wise iteration operator for the acoustic case are

$$\lambda^{s,h}_{\cdot,\beta}(\boldsymbol{\theta}) = 1 - \eta \pm \eta \frac{4 \left| \sin\left(\frac{\theta_x}{2}\right) \right| \left| \sin\left(\frac{\theta_z}{2}\right) \right|}{\sqrt{\left[k_{\mathrm{P}}^2 h^2 (1 - \beta\,\mathrm{i}) - 4 \sin\left(\frac{\theta_x}{2}\right)^2 \right] \left[k_{\mathrm{P}}^2 h^2 (1 - \beta\,\mathrm{i}) - 4 \sin\left(\frac{\theta_z}{2}\right)^2 \right]}},$$

which, for small P-wave wavenumbers k_{P}, reduces to the limit case for S-waves (36). The expression for Kelly's scheme is fairly similar, where, in the numerator of the fraction, the arguments $\theta_x/2$, $\theta_z/2$ of the sine function are replaced by θ_x, θ_z. Therefore, the line smoother is able to properly reduce the short-wavelength components of the error. In this paper, we avoid the issue altogether by considering only strictly elastic media with $c_s \gg 0$.

4.2.3. Comparison of n-grid and numerical convergence factors

We conclude the section on elastic multigrid by testing the theoretical findings on the line smoother in combination with a complete multigrid cycle. We refer to [27] for an overview on the different types of multigrid scheduling strategies. In our numerical study, we will consider the traditional V- and W-cycles.

As pointed out in Section 4.2.1, it is possible to analytically determine the convergence factors of multigrid under the simplifications discussed earlier. The analysis describes the error reduction by multigrid over a limited number of grid levels on a problem with constant coefficients. On the coarsest grid, the problem is assumed to be solved exactly. In the literature,

Table 2
Comparison of the convergence factors, μ_n, computed analytically by n-grid analysis for $n=2$, $n=3$, and $n=4$ as in (30) to the numerical convergence factors, μ_{num}, determined by the power method, using the staggered-grid finite-difference scheme. Here we considered a W-cycle with a single pre-smoothing and post-smoothing step. The coarse-grid discretization is obtained by Galerkin coarsening, with full weighting as restriction and bilinear interpolation as prolongation operator. The relaxation parameters are optimized to minimize the smoothing factors. The comparison is made for different values of the velocity ratio r, damping factor β and wavenumber: (a) $k_S h = \pi/5$, (b) $k_S h = \pi/10$.

W(1, 1)-cycle convergence factors

	$\beta = 1.5$				$\beta = 2$			
	$r=2$	$r=3$	$r=4$	$r=5$	$r=2$	$r=3$	$r=4$	$r=5$
$\mu_s^{\nu_1+\nu_2}$	0.51	0.72	0.82	0.88	0.50	0.71	0.82	0.88
μ_2	0.50	0.71	0.82	0.88	0.50	0.71	0.82	0.88
μ_3	0.50	0.72	0.82	0.88	0.50	0.71	0.82	0.88
μ_4	0.50	0.72	0.91	**1.37**	0.50	0.71	0.82	0.88
μ_{num}	0.50	0.72	0.84	**1.26**	0.49	0.70	0.82	0.88

(a) $k_S h = \pi/5$

	$\beta = 1.5$				$\beta = 2$			
	$r=2$	$r=3$	$r=4$	$r=5$	$r=2$	$r=3$	$r=4$	$r=5$
$\mu_s^{\nu_1+\nu_2}$	0.47	0.68	0.80	0.86	0.47	0.68	0.80	0.86
μ_2	0.46	0.68	0.80	0.86	0.46	0.68	0.80	0.86
μ_3	0.46	0.68	0.80	0.86	0.46	0.68	0.80	0.86
μ_4	0.46	0.68	0.80	0.86	0.46	0.68	0.80	0.86
μ_{num}	0.46	0.68	0.80	0.86	0.46	0.67	0.80	0.86

(b) $k_S h = \pi/10$

usually two or three levels are considered. In general, the smaller the number of coarse-grid levels, the more optimistic the estimate. Here, we needed $n=4$ levels to address the complications of the elastic wave equation.

The comparison is made with numerically computed convergence factors. This requires the spectral radius of

$$\mathbf{I} - \mathbf{M}^h \mathbf{H}_\beta^h, \qquad \mathbf{M}^h \approx \left(\mathbf{H}_\beta^h\right)^{-1},$$

where $\mathbf{M}^h$ is the approximate inverse of $\mathbf{H}_\beta^h$ after a multigrid cycle. This computation can be carried out by simply using the power method. In Table 2, we list the results of the n-grid analysis for the staggered-grid scheme using a multigrid W-cycle, for different values of the velocity ratio, damping factor and S-wave wavenumber. We observe that we need damping values β that are larger than the one used in [10,11] for the acoustic case to obtain a convergent scheme, highlighting a more problematic near-null subspace of modes, caused by the coarse-grid discretization.

As an example, Table 2 compares the results of the n-grid analysis to the numerically computed convergence factors using a W-cycle. The numerical results are in good agreement with the theoretical estimates, demonstrating that local-mode analysis can be used as a guiding tool to choose and optimize the different components of multigrid.

4.2.4. Extension of the smoothing scheme to 3-D

The generalization of the proposed line-smoother to 3-D is of paramount importance for seismic imaging applications. As previously noted for the 2-D case, line solvers can benefit from parallelization. The most obvious generalization of line-smoothing to 3-D corresponds to plane-wise smoothing, which amounts to the solution of many 2-D problems. For large problems, however, this might result in a very expensive scheme. To ease the computation in 3-D, we foresee the following alternatives: alternating line-smoothing (x, y and z directions), semi-coarsening, or a combination of both. We might expect better smoothing factors with plane-smoothing than with 3-D alternating line-smoothing. These alternatives can be assessed by local-mode analysis, similar to the 2-D case presented here, and will be the subject of future study.

5. Numerical examples

So far, we discussed the performance of different smoothers and their combination with multigrid for the damped elastic wave equation. We validated our theoretical understanding by local-mode analysis and found that these results agree with the actual numerical behaviour of the scheme. Now that we have addressed the accuracy and effectiveness of the preconditioner in the system (15), estimated by multigrid, we are ready to study the numerical performance of the iterative solver applied to (15).

In Section 3, we studied the effect of preconditioning by the complex-shifted operator assuming its exact inversion. Clearly, the actual numerical behaviour of a chosen iterative scheme should account for the inexact inversion of the complex-shifted wave operator by multigrid. Consequently, the spectrum configuration as seen in Fig. 3 and the convergence of the

iterative scheme will be affected. The purpose of this section is to perform a series of numerical tests of the preconditioning scheme (15) using the actual multigrid implementation described in Section 4.

We start with a brief comment on the specific numerical implementation of the various multigrid ingredients introduced in Section 4.2.2 as well as the discrete representation of the problem at hand. We perform two types of experiments: one to test the numerical scheme for the homogeneous case and the other for an arbitrarily heterogeneous model, based on the well-known benchmark acoustic Marmousi model: the Marmousi2 model [35]. Each experiment is accompanied by a discussion of the results, which will show a behaviour — including shortcomings — similar to the acoustic shifted-Laplacian multigrid preconditioning.

5.1. Additional settings and multigrid implementation

In the following numerical experiments, we test the finite-difference schemes of Kelly and Virieux, discussed in Section 2. The choice of the discretization of the computational model is made such that the quantity $k_S^{\max} h$, where $k_S^{\max}$ is the maximum wavenumber, is kept constant. This condition is equivalent to a discretization that keeps the number of points per minimum wavelength constant. This means that for a given frequency, the size of the computational grid is adjusted accordingly.

As for boundary conditions, we set up sponge layers to mimic absorption at the edges of the model. Of course, perfectly matched layers can be also considered [36]. Theoretical and numerical studies on the effect of this type of boundary conditions for acoustic multigrid can be found in [37] and [38].

For the multigrid implementation, we stick to traditional choices for the various components. For the coarse-grid operator we can consider: (*i*) direct discretization of the wave operator on the corresponding grid level, or (*ii*) Galerkin condition (26). Local-mode analysis suggests that (*ii*) is a more robust approach. For restriction and prolongation operators, we use, respectively, full weighting and bilinear interpolations. In stencil notation,

$$R_h^{2h} \triangleq \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}, \qquad P_{2h}^{h} \triangleq \frac{1}{4}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix},$$

which symbolize the action $(R_h^{2h} u^h)_{2i,2j} = (R_h^{2h} \star u^h)_{2i,2j}$, defined in (6), and $(P_{2h}^{h} u^{2h})_{i,j} = (P_{2h}^{h} \star I_{2h}^{h} u^{2h})_{i,j}$, where $I_{2h}^{h} u^{2h}$ is the injection of u^{2h} on the grid G^h, determined by the values of u^{2h} on $\mathrm{G}^{2h} \subseteq \mathrm{G}^h$ and vanishing on every other grid point. Operator dependent grid transfer restrictions and prolongations, as studied by [39] and extended by [10] to complex matrices, are also an option. As of multigrid grid-level scheduling cycles, we will consider V(ν_1, ν_2)-cycles, for computational efficiency. The number of pre and post-smoothing steps will be set to $\nu_1 = \nu_2 = 2$. The relaxation parameters η in (22), for the line-Jacobi smoother, are optimized for the range of velocity ratios displayed by the medium, and for all the wavenumbers up to the maximum value of the model. We should point out that the chosen amount of damping, β, is based on numerical investigations rather than rigorous theoretical analysis. As discussed in Section 3, large values of β lead to better multigrid convergence of the preconditioner while small values generate a more favourable spectrum for the preconditioned system if that system could be solved exactly.

Lastly, a number of options is available for the outer iterative method. For non-symmetric problems, we could consider any Krylov subspace method such as GMRES, Bi-CGSTAB or IDR(s). In our experience, Bi-CGSTAB performs the best for the problem at hand and will be the choice for the following numerical study.

5.2. The homogeneous case

We start with testing how the preconditioning scheme of Section 3 behaves in the homogeneous case. The computational domain is the 2-D square $[-1/2, 1/2]^2$ and we solve for the wavefield generated by a point source located at the centre $(0, 0)$ of the domain. The type of source considered is a vertical force. We study how the number of iteration increases as a function of the frequency. For each of the experiments, the quantity $k_S h$ will be kept fixed at a value of $\pi/5$ and we vary the wavenumber k_S. Therefore, the grid spacing h is adjusted accordingly, as mentioned before. Also, different velocity ratios r and damping factors β are considered in the analysis.

5.2.1. Results

Table 3 lists the results, which highlight a reasonable convergence rate for the considered frequency range. As was the case with the scalar Helmholtz equation, we found experimentally that the required number of iterations increases linearly with the wavenumber k_S. Since the multigrid convergence factors get worse for increasing velocity ratio r, as discussed in Section 4, the iterative method will perform worse for high values of r.

5.3. A highly heterogeneous problem: the elastic Marmousi model

Now that we have validated the scheme for the homogeneous problem, we test its behaviour for the heterogeneous case. We will adopt an elastic model based on the well-known Marmousi model, originally designed by the Institut Français du

Table 3
Numerical experiment for a homogeneous model. The computational domain is the square $[-1/2, 1/2]^2$. A vertical force is positioned at the centre as a delta source. For each of the performed experiments, we fixed the quantity $k_S h = \pi/5$, corresponding to 10 samples per wavelength. We list the number of iterations required to decrease the residual norm by 6 decades. The selected values of k_S increase linearly and the grid spacing h varies along. We employ a V(2, 2) multigrid cycle, where each of the components are optimized according to the specific velocity ratio r. The experiments consistently show a linear growth of the number of iterations with respect to the wavenumber k_S. We also observe that, as the multigrid convergence factors decay for increasing velocity ratios r, more iterations are required. A comparison is made between the performance of the (a) Virieux's scheme and (b) Kelly's scheme.

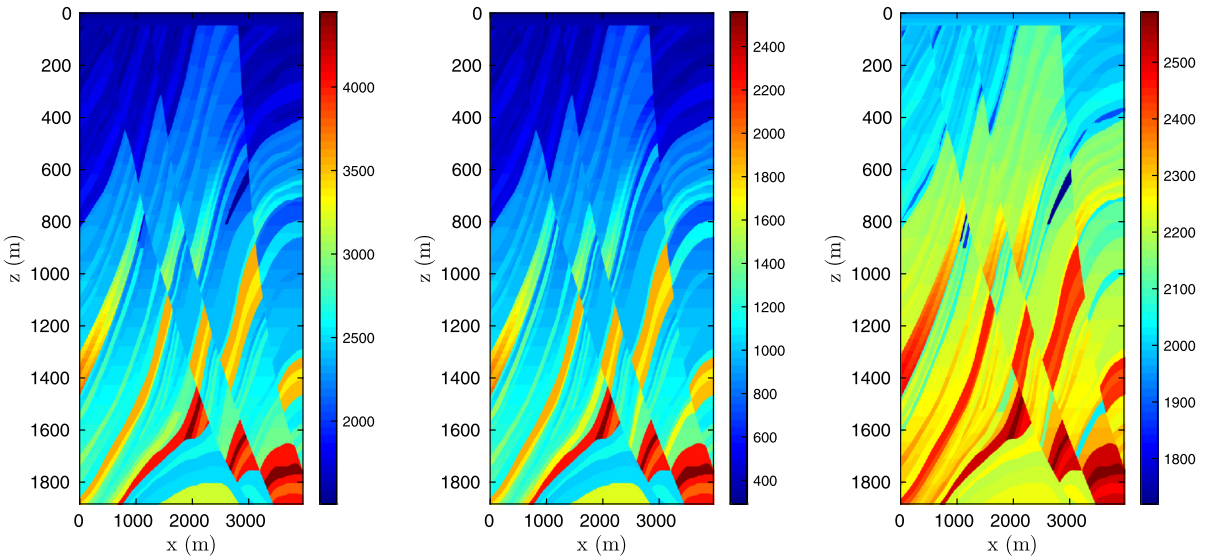
Results for the homogeneous problem: iteration count

	$r = 2$					
	$k_S = 100.5$	$k_S = 181.0$	$k_S = 261.4$	$k_S = 341.8$	$k_S = 422.2$	$k_S = 502.7$
$\beta = 1.5$	106	159	214	268	324	382
$\beta = 2$	139	226	312	402	443	501
	$r = 3$					
	$k_S = 100.5$	$k_S = 181.0$	$k_S = 261.4$	$k_S = 341.8$	$k_S = 422.2$	$k_S = 502.7$
$\beta = 1.5$	121	203	286	391	467	562
$\beta = 2$	140	212	285	352	446	522

(a) Virieux's scheme

	$r = 2$					
	$k_S = 100.5$	$k_S = 181.0$	$k_S = 261.4$	$k_S = 341.8$	$k_S = 422.2$	$k_S = 502.7$
$\beta = 1.5$	108	150	204	256	327	375
$\beta = 2$	135	208	289	325	430	486
	$r = 3$					
	$k_S = 100.5$	$k_S = 181.0$	$k_S = 261.4$	$k_S = 341.8$	$k_S = 422.2$	$k_S = 502.7$
$\beta = 1.5$	108	174	240	299	392	467
$\beta = 2$	124	188	245	320	391	480

(b) Kelly's scheme



(a) P-wave velocity c_P (m s^{-1}) (b) S-wave velocity c_S (m s^{-1}) (c) Density ρ (kg m^{-3})

Fig. 7. Subset of the elastic Marmousi2 model: (a) P-wave velocity c_P, (b) S-wave velocity c_S and (c) density ρ, used in the numerical experiments in Section 5. The water column on the top of the model has been replaced by an elastic layer.

Pétrole. This acoustic version has been extended to the elastic case by [35]. Here, we will only consider the top-central portion of the full model. We discard the water layer to avoid the complications discussed in Section 4.2.2 that will affect the performance of multigrid for the staggered-grid scheme. The corresponding elastic parameters are depicted in Fig. 7. For this model, the maximum value of the velocity ratio r is about 5.2.

In the numerical experiments, we are going to model the frequencies 1.45 Hz, 2.9 Hz, 5.8 Hz and 11.6 Hz with different grid spacings. Fig. 8 shows the real part of the vertical particle displacement generated by a point source.

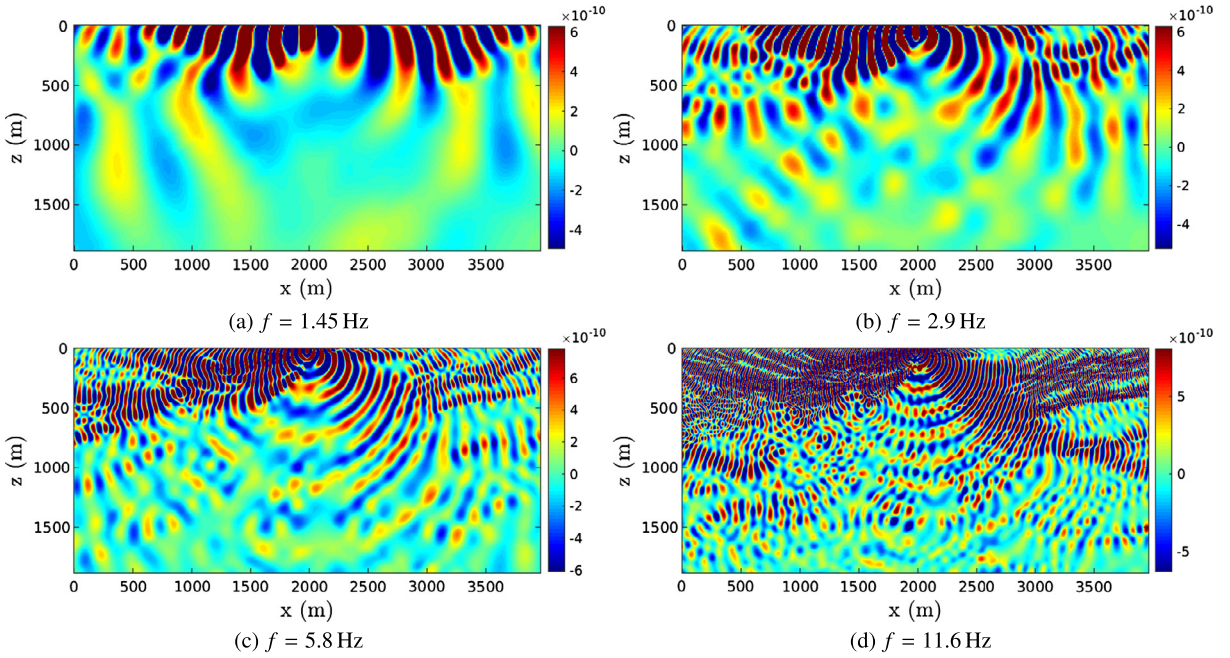


Fig. 8. Real part of the vertical component of the wavefield propagating through the Marmousi2 model (Fig. 7), for the frequencies (a) 1.45 Hz, (b) 2.9 Hz, (c) 5.8 Hz and (d) 11.6 Hz, generated by a point source.

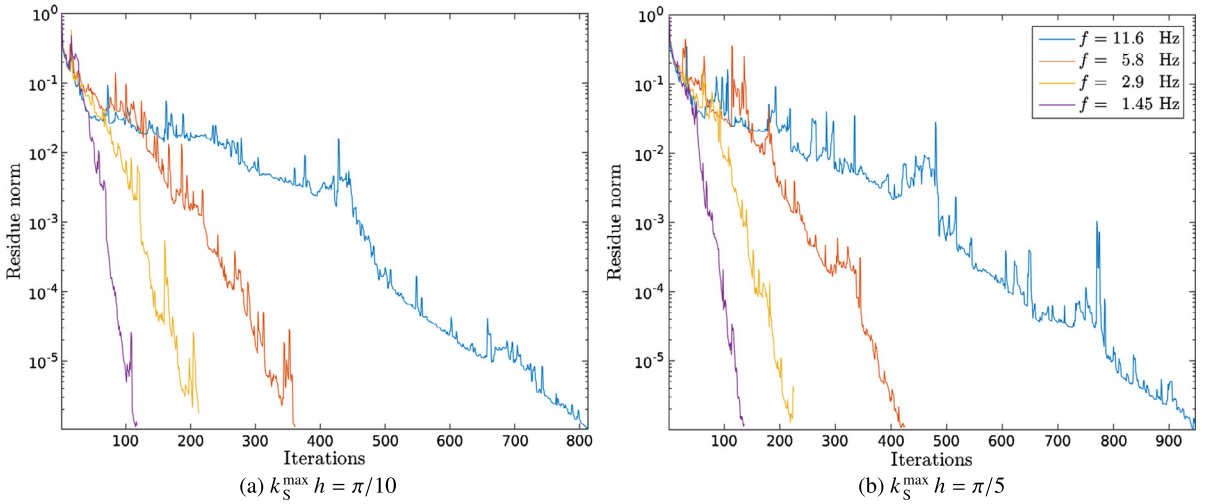


Fig. 9. Convergence history of the iterative solver applied to the Marmousi2 problem for a number of frequencies and wavenumbers. Here, the wave equation is discretized by the staggered-grid scheme. The discretization for the various frequencies is chosen in such a way that the quantities (a) $k_S^{max} h = \pi/10$ (equivalent to a discretization of 20 points per minimum wavelength) and (b) $k_S^{max} h = \pi/5$ (10 points per minimum wavelength) remain constant. In other words, the size of the model is adjusted to the relative frequency as in Table 4. The coarse-grid operator is obtained by Galerkin coarsening. We use the damping factor $\beta = 2$ in each experiment. The multigrid cycle is V(2, 2).

5.3.1. Results

We tested the convergence behaviour of the iterative solver for different frequencies and values of $k_S^{max} h$. The numerical experiments performed show good convergence properties. The convergence history of Bi-CGSTAB for the staggered-grid scheme is reported in Fig. 9. However, the number of iterations grows linearly with increasing frequency, as was the case for the acoustic wave equation.

Table 4 summarizes the numerical experiments. A comparison between the performance of Virieux's and Kelly's scheme is included. As highlighted by the local-mode analysis in Section 4.2.1, the multigrid performance for Kelly's scheme are better than for Virieux's. However, this comes at the cost of inferior accuracy (Fig. 2).

Table 4
Summary of the results of the numerical experiments on the Marmousi2 model. We compare Virieux's and Kelly's finite-difference schemes. For the reasons discussed in Section 4.2.2, Kelly's discretization leads to a slightly better performance, even though the accuracy of the scheme is worse than the staggered-grid's (see Fig. 2). For each of the finite-difference schemes, we choose the Galerkin coarse-grid operator, $\beta = 2$ as damping factor, and a V(2, 2) multigrid cycle.

Results for the Marmousi2 model: iteration count

f (Hz)	$k_S^{\max} h = \pi/10$			$k_S^{\max} h = \pi/5$		
	Size	Virieux	Kelly	Size	Virieux	Kelly
1.45	385×593	118	102	193×297	137	98
2.9	577×993	214	165	289×497	226	160
5.8	961×1793	362	291	481×897	426	297
11.6	1665×329	814	694	833×1665	948	647

6. Conclusions

This paper comprises a first assessment of the applicability of the preconditioning technique by Laplacian shifting, using multigrid, applied to the time-harmonic elastic wave equation. In order to substantiate this analysis, we made use of the classical theoretical tool of local-mode analysis.

By means of local-mode analysis, we were able to point out the shortcomings of the natural choice of point-Jacobi as a smoother, otherwise effective in the acoustic case. This is due to a defective solution of the S-wave modes in the case of large P-to-S velocity ratios. In terms of particle displacement wavefield, this results in smoothing anisotropy. We therefore introduced and validated a suitable line-wise variant of the Jacobi scheme.

In the present experience, different numerical discretization schemes of the wave equation might lead to a different multigrid performance, as it was the case with Kelly's and Virieux's scheme. For Virieux's scheme, in particular, handling acoustic subsets of the model is problematic and should be dealt with in a special way. This leaves the question on how to find the optimal trade-off between accuracy and efficiency.

The results obtained on highly heterogeneous models suggest that the proposed method provides reasonable convergence, in line with what was obtained for the acoustic case. With the elementary multigrid component used in this paper, we observed that the damping factors should have higher values than in the acoustic case, highlighting a more problematic near-null kernel. The deflation techniques discussed by [30] and [31] might be especially helpful in the elastic case.

The elastic case suffers from the same shortcomings as the acoustic shifted-Laplacian multigrid, namely the linear growth of the number of iterations for increasing frequency at fixed number of points per wavelength. As noted in the introduction, this means that frequency-domain imaging methods can only obtain the same computational complexity as their time-domain counterparts, from a theoretical point of view. In practice, the constants involved in the complexity analysis play an important role for the actual performance and the iterative solver here discussed might be a efficient alternative to the time domain, as motivated by [4] for acoustic imaging. While in 2-D a direct LU decomposition with nested dissection will be the preferred choice, if many shots have to be computed in the same model, in 3-D this is no longer the case, and the present study provides guidelines as to how to proceed.

References

[1] R.G. Pratt, Frequency-domain elastic wave modeling by finite differences: a tool for crosshole seismic imaging, Geophysics 55 (5) (1990) 626–632.
[2] W.A. Mulder, R.E. Plessix, How to choose a subset of frequencies in frequency-domain finite-difference migration, Geophys. J. Int. 158 (2004) 801–812.
[3] A. George, J.W.H. Liu, Computer Solution of Large Sparse Positive Definite Systems, Prentice Hall, 1981.
[4] H. Knibbe, W.A. Mulder, C.W. Oosterlee, C. Vuik, Closing the performance gap between an iterative frequency-domain solver and an explicit time-domain scheme for 3D migration on parallel architectures, Geophysics 79 (2) (2014) 547–561.
[5] J. Xia, S. Chandrasekaran, M. Gu, X.S. Li, Superfast multifrontal method for large structured linear systems of equations, SIAM J. Matrix Anal. Appl. 31 (3) (2009) 1382–1411.
[6] S. Wang, M.V. de Hoop, J. Xia, On 3D modeling of seismic wave propagation via a structured parallel multifrontal direct Helmholtz solver, Geophys. Prospect. 59 (2011) 857–873.
[7] S. Wang, M.V. de Hoop, J. Xia, X.S. Li, Massively parallel structured multifrontal solver for time-harmonic elastic waves in 3-D anisotropic media, Geophys. J. Int. 191 (2012) 346–366.
[8] C. Stolk, A rapidly converging domain decomposition method for the Helmholtz equation, J. Comput. Phys. 241 (2013) 240–252, http://dx.doi.org/10.1016/j.jcp.2013.01.039.
[9] L. Zepeda-Núñez, L. Demanet, Nested domain decomposition with polarized traces for the 2D Helmholtz equation, URL http://arxiv.org/abs/1510.01831.
[10] Y.A. Erlangga, C.W. Oosterlee, C. Vuik, A novel multigrid based preconditioner for heterogeneous Helmholtz problems, J. Sci. Comput. 27 (4) (2006) 1471–1492.
[11] Y.A. Erlangga, C. Vuik, C.W. Oosterlee, On a class of preconditioners for solving the Helmholtz equation, Appl. Numer. Math. 50 (2004) 409–425.
[12] A. Bayliss, C.I. Goldstein, E. Turkel, An iterative method for Helmholtz equation, J. Comput. Phys. 49 (1983) 443–457.
[13] A.L. Laird, M.B. Giles, Preconditioned iterative solution of the 2D Helmholtz equation, Report 02/12, Oxford University Computing Laboratoryo, 2002.
[14] H.C. Elman, G.O. Ernst, D.P. O'Leary, A multigrid method enhanced by Krylov subspace iteration for discrete Helmholtz equations, SIAM J. Sci. Comput. 23 (4) (2001) 1291–1315.
[15] P. Tsuji, R. Tuminaro, Augmented AMG-shifted Laplacian preconditioners for indefinite Helmholtz problems, Numer. Linear Algebra Appl. 22 (6) (2015) 1077, http://dx.doi.org/10.1002/nla.
[16] Y. Saad, Iterative Methods for Sparse Linear Systems, 2nd edition, SIAM, 2003.

[17] H.A. van der Vorst, Bi-CGSTAB: a fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems, SIAM J. Sci. Stat. Comput. 13 (2) (1992) 631–644.
[18] P. Sonneveld, M.B. van Gijzen, IDR(s): a family of simple and fast algorithms for solving large nonsymmetric systems of linear equations, SIAM J. Sci. Comput. 31 (2) (2008) 1035–1062, http://dx.doi.org/10.1137/070685804.
[19] K.R. Kelly, R.W. Ward, S. Treitel, R.M. Alford, Synthetic seismograms: a finite-difference approach, Geophysics 41 (1) (1976) 2–27.
[20] J. Virieux, P-SV wave propagation in heterogeneous media: velocity-stress finite-difference method, Geophysics 51 (4) (1986) 889–901.
[21] I. Štekl, R.G. Pratt, Accurate viscoelastic modeling by frequency-domain finite differences using rotated operators, Geophysics 63 (5) (1998) 1779–1794.
[22] D.-J. Min, C. Shin, B.-D. Kwon, S. Chung, Improved frequency-domain elastic wave modeling using weighted-averaging difference operators, Geophysics 65 (3) (2000) 884–895.
[23] D.-J. Min, C. Shin, R.G. Pratt, H.S. Yoo, Weighted-averaging finite-element method for 2D elastic wave equations in the frequency domain, Bull. Seismol. Soc. Am. 93 (2) (2003) 904–921.
[24] D.-J. Min, C. Shin, H.S. Yoo, Free-surface boundary condition in finite-difference elastic wave modeling, Bull. Seismol. Soc. Am. 94 (1) (2004) 237–250.
[25] A. Brandt, O.E. Livne, Multigrid Techniques: 1984 Guide with Applications to Fluid Dynamics, SIAM, 2011.
[26] U. Trottenberg, C.W. Oosterlee, A. Schüller, Multigrid, Academic Press, 2001.
[27] W.L. Briggs, V.E. Henson, S. McCormick, A Multigrid Tutorial, SIAM, 2000.
[28] R.A. Nicolaides, Deflation of conjugate gradients with applications to boundary value problems, SIAM J. Numer. Anal. 24 (2) (1987) 355–365.
[29] J. Frank, C. Vuik, On the construction of deflation-based preconditioners, SIAM J. Sci. Comput. 23 (2) (2001) 442–462.
[30] Y.A. Erlangga, R. Nabben, Multilevel projection-based nested Krylov iteration for boundary value problems, J. Sci. Comput. 30 (3) (2008) 1572–1595.
[31] A.H. Sheik, C. Vuik, D. Lahaye, On the convergence of shifted Laplace preconditioner combined with multilevel deflation, Numer. Linear Algebra Appl. 20 (2013) 645–662.
[32] R. Wienands, C.W. Oosterlee, On three-grid Fourier analysis for multigrid, SIAM J. Sci. Comput. 23 (2) (2001) 651–671.
[33] W.A. Mulder, A new multigrid approach to convection problems, J. Comput. Phys. 83 (2) (1989) 303–323, http://dx.doi.org/10.1016/0021-9991(89)90121-6.
[34] W. Press, S. Teukolsky, W. Vetterling, B. Flannery, Numerical Recipes 3rd Edition: The Art of Scientific Computing, Cambridge University Press, 2007.
[35] G.S. Martin, R. Wiley, K.J. Marfurt, Marmousi2: an elastic upgrade for Marmousi, Lead. Edge 25 (2) (2006) 156–166.
[36] J.P. Berenger, A perfectly matched layer for the absorption of electromagnetic waves, J. Comput. Phys. 114 (2) (1994) 185–200, http://dx.doi.org/10.1006/jcph.1994.1159.
[37] B. Reps, W. Vanroose, Analyzing the wave number dependency of the convergence rate of a multigrid preconditioned Krylov method for the Helmholtz equation with an absorbing layer, Numer. Linear Algebra Appl. 19 (2012) 232–252, http://dx.doi.org/10.1002/nla.1806.
[38] G. Pan, A. Abubakar, Iterative solution of 3D acoustic wave equation with perfectly matched layer boundary condition and multigrid preconditioner, Geophysics 78 (2013) T133–T140, http://dx.doi.org/10.1111/j.1365-246X.2011.05244.x.
[39] P.M. de Zeeuw, Matrix-dependent prolongations and restrictions in a blackbox multigrid solver, J. Comput. Appl. Math. 33 (1) (1990) 1–27, http://dx.doi.org/10.1016/0377-0427(90)90252-U.