



Fully scalable solver for frequency-domain visco-elastic wave equations in 3D heterogeneous media: A controllability approach

Jet Hoe Tang, Romain Brossier, Ludovic Métivier

► To cite this version:

Jet Hoe Tang, Romain Brossier, Ludovic Métivier. Fully scalable solver for frequency-domain visco-elastic wave equations in 3D heterogeneous media: A controllability approach. Journal of Computational Physics, 2022, 468, pp.111514. <10.1016/j.jcp.2022.111514>. <hal-03852539>

HAL Id: hal-03852539

<https://hal.science/hal-03852539v1>

Submitted on 15 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Fully scalable solver for frequency-domain visco-elastic wave equations in 3D heterogeneous media: a controllability approach

Jet Hoe Tang^a, Romain Brossier^a, Ludovic Métivier^{a,b}

^a*Univ. Grenoble Alpes, ISTERre, F-38000, Grenoble, France*

^b*Univ. Grenoble Alpes, CNRS, LJK, F-38000, Grenoble, France*

Abstract

We develop a controllability strategy for the computation of frequency-domain solutions of the 3D visco-elastic wave equation, in the perspective of seismic imaging applications. We generalize the controllability results for such equations beyond the sound-soft scattering (obstacle) problem. We detail the conjugate gradient implementation and show how an inner elliptic problem needs to be solved to compute the Riesz representative of the gradient at each iteration. We select a spectral-element spatial discretization and a fourth-order Runge-Kutta time discretization. We implement the controllability method in the framework of the SEM46 full waveform modeling and inversion software, to inherit for its excellent scalability which relies on an efficient domain decomposition algorithm. We perform a series of numerical experiments to validate the approach and illustrate its scalability up to more than fifteen hundred cores. In this case, with an elapsed time of less than 50 minutes, we solve a problem on a cubic domain containing up to 160 wavelengths in each direction, involving more than 1.7 billion unknowns.

1. Introduction

Developing robust and scalable solvers for computing frequency-domain solutions of the 3D visco-elastic wave equation is a long standing problem. The interest for such high performance computing (HPC) algorithm emanates from different applications, notably from the seismology and seismic imaging communities. In the frame of full waveform inversion (FWI), a high resolution seismic imaging technique, the solution of the 3D visco-elastic equations needs to be computed repeatedly in complex (heterogeneous) media, for possibly hundreds to thousands source positions [see 59, for a review].

First 3D exploration scale applications of FWI were led in the acoustic approximation, for the interpretation of marine data *i.e.* data acquired with air-guns as sources and hydrophones as receivers, in a water layer on top of the investigated medium [49, 42]. For targets involving the propagation of several tens of wavelengths in the three directions of space, conventional HPC facilities (for instance available in national computing centers) make it possible to use parallel direct solvers, relying on a LU factorization of the linear system resulting from the discretization of the 3D acoustic equations [42]. Such a LU

factorization is especially useful in this context of multiple source positions, each position corresponding to a different right hand side. If the LU factorization can be stored, the complexity for computing the solution for one right hand side is linear (two triangular solves). The limitation comes from the ability to perform the factorization, which requires to store $O(N^4)$ factors, where N is the number of grid points in one spatial direction. More precisely, the difficulty comes from the scalability of this operation, which remains limited even with state-of-the-art solver such as MUMPS [2, 33].

Investigating larger targets in the acoustic approximation, involving hundreds to thousands of propagated wavelengths in each spatial direction, can thus not rely on such direct solvers. Moving from an acoustic description to an elastic description of the wave propagation implies the same: for realistic size applications the number of propagated S-wavelength would be too large for direct solvers to be employed. Several motivations currently exist to employ such a more realistic description of the wave propagation: 1) when land data has to be interpreted, where sources (explosives, vibrators) and receivers (geophones) are deployed on land, recording data with a strong imprint of elastic effects (Love and Rayleigh surface waves, S-waves); 2) when node data has to be interpreted, where sources (airguns) are in the water layer, while 4-components receivers (hydrophones and geophones) are deployed on the seabed, recording data again with a strong imprint of elastic effects (Scholte waves at the sea bottom interface, S-waves and converted waves). From an imaging point of view, inverting land and node data is a challenge on its own. However, node acquisition systems are currently more and more deployed, the decoupling between sources and receivers making it possible to acquire wider offset/azimuth data and therefore increasing the illumination depth. Incorporating a visco-elastic modeling engine within the FWI algorithm also makes it possible to reconstruct mechanical parameters beyond P-wave velocities: S-wave velocities, attenuation factor, density, and anisotropy parameters, which are crucial to go beyond imaging towards reservoir characterization. Modern FWI algorithms are thus evolving from acoustic to fully anisotropic (visco-)elastic approximation.

Due to the limitation of direct solvers, FWI algorithms have evolved rapidly towards time-domain based modeling engine, for acoustic and elastic approximation. Time-domain implementation exhibit interesting properties: good scalability on HPC platform, thanks to a two-levels parallelism over sources (outer level) and domain decomposition (inner level), low memory requirement in $O(N^3)$, and the possibility to apply straightforwardly data-processing based on time/offset windowing. The vast majority of 3D FWI applications in seismic imaging and seismology are currently performed using such time-domain FWI algorithms [19, 57, 32, 24, 53, 47].

However, time-domain modeling still faces specific limitations. The most stringent is the CFL condition, which makes the time discretization step in $\left(\frac{v_{min}}{v_{max}} \times h\right)$ where v_{min} and v_{max} are respectively the minimum and maximum waves velocities, and h is the smallest element size in the spatial discretization mesh. For elastic media with shallow soft sediments, the S-wave velocity can

reach critically low values, making h and the ratio $\frac{v_{min}}{v_{max}}$ very low as well, resulting in a very small time discretization step. In addition, there exists no solution to treat efficiently multiple right-hand sides, aside from treating them in parallel. Finally, for inversion, attributing the same weight to each frequency within a frequency band requires data spectral whitening, which can be difficult to control on narrow frequency bands. This difficulty is not met by frequency-domain FWI algorithms. Such a frequency whitening is beneficial especially to invert for the lowest frequency band of the data when cycle skipping is to be avoided.

These difficulties prompt us to continue the investigation on HPC algorithms for computing frequency-domain solution of the visco-elastic equations, capable of overcoming the limitations of direct solvers. Iterative solvers are a natural alternative. Their matrix-free formalism enables low memory requirement and good scalability properties. However, the discretization of frequency-domain visco-elastic wave equation, as an instance of generalized Helmholtz equations, is known to yield ill-conditioned linear systems. Iterative solvers therefore need to be combined with efficient preconditioners to achieve satisfactory performances. Recent examples are for instance the CARP-CG approach [34], shifted Laplacian preconditioners [48], fast low-rank factorization/approximation based on a hierarchical approach [21], or more simple preconditioners based on a layered approximation of the subsurface [5]. However the convergence of these methods still deteriorates when tackling large scale problems as the efficiency of the preconditioners decreases. A nice analysis why generalized Helmholtz equations yield increasingly ill-conditioned linear systems as the size of the systems grows is provided in [17]. The same difficulty is faced with domain decomposition methods combining direct solvers on subdomains, with a preconditioned iterative solver for the unknowns at the interfaces [27, 51].

Other existing approaches rely instead on solving frequency-domain problems with time-domain solvers, to take advantage of the reliability and good scalability properties of the latter [41, 50]. Among these, the controllability method (CM), first proposed by Bristeau et al. for acoustic sound-soft scattering problems [7] has been shown particularly efficient. In this approach, a time-periodic solution of the time-domain wave equation is computed by minimizing the least-squares distance between the initial wavefield at time 0 and the wavefield after one period T corresponding to the frequency at which the frequency-domain solution is sought. The control variables are the initial value of the solution and its derivative in time. In their seminal paper, Bristeau et al. prove that the frequency domain solution can be expressed as a linear combination of these control variables at convergence.

Recent studies have extended the work of Bristeau et al. to general acoustic and electromagnetic wave propagation problems. In these studies, it is shown how **the CM makes** it possible to solve very high frequency problems using modern HPC devices with scalability observed over several thousands of computational units [28, 26, 55, 25, 10, 56]. For acoustic wave problems, Appelö et al. [3] have recently presented an alternative time-domain based approach called WaveHoltz method. It is based on fixed-point iterations. Time-dependent wave

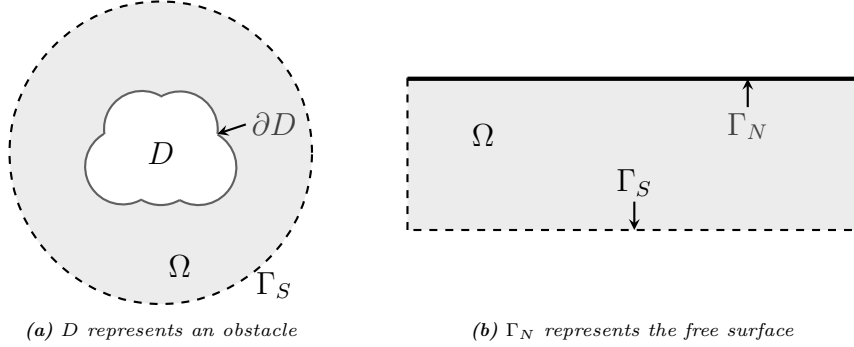


Figure 1: Illustration of the computational domain $\Omega \subset \mathbb{R}^d$, $\Gamma = \partial\Omega$, $d = 2, 3$, with physical boundaries $\partial D = \Gamma_D \cup \Gamma_N$ (—) and artificial boundaries Γ_S (- -).

equations are solved over a short time period at each iteration. The higher eigenmodes contained in the time-periodic solutions are filtered out by a Fourier-like filter procedure.

Regarding (visco-)elastic equations, the first extension of the CM to these equations has been performed by Mönkölä [38]. However in [her](#) study, the analysis is restricted to sound-soft scattering problems, and is performed in the context of fluid/solid coupling. It also assumes isotropic elastic media for the solid part, and the algorithm is illustrated in 2D. In this study, we are interested in an extension of [the CM](#) to general elastodynamics problem, beyond sound-soft scattering, considering fully anisotropic media and 3D implementation, with 3D seismic imaging as a perspective. We first prove a general controllability theorem showing how the frequency-domain solution can be retrieved from the initial conditions of the time-domain problem posed over a single period, considering a general set of boundary conditions, including the sound-soft scattering problem, but also configurations typically encountered in seismic imaging with a free surface on top and absorbing boundary conditions on the other sides of the computational domain. We implement this strategy for the solution of 3D frequency-domain visco-elastic equations using a spectral finite-element discretization, within the framework of the 3D full waveform modeling and inversion software SEM46 [58]. We analyze acceleration techniques based on the fast computation of an initial guess through a transient-phase approach. We illustrate the efficiency of the approach on 3D realistic examples: all computational steps of the algorithm benefit from the domain decomposition technique implemented in SEM46, making the algorithm fully scalable. The final example makes use of more than fifteen hundred cores with an elapsed time of 2851s for solving a problem on a cubic domain containing up to 160 wavelengths in each direction, involving more than 1.7 billion unknowns.

2. General elastic equations: expressing the frequency-domain solution from the time-domain solution

2.1. State of the art: sound-soft scattering problem

We consider the frequency-domain elastic wave equation in a bounded connected Lipschitz computational domain $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, whose boundary Γ is

composed in the most general settings as

$$\Gamma = \Gamma_D \cup \Gamma_N \cup \Gamma_S \quad (1)$$

where Γ_D (respectively Γ_N, Γ_S) stands for the part of the boundary where Dirichlet (respectively Neumann, absorbing) boundary conditions are imposed – see Figure 1.

Let \mathbf{u} denote the (complex-valued) wavefield which satisfies the frequency-domain elastic wave equation

$$-\omega^2 \rho(x) \mathbf{u}(x) - \nabla \cdot \sigma(x, \mathbf{u}) = \mathbf{f}(x), \quad x \in \Omega, \quad (2a)$$

$$-i\omega \rho(x) \mathbf{B}(x) \mathbf{u}(x) + \sigma(x, \mathbf{u}) \mathbf{n} = \mathbf{g}_S(x), \quad x \in \Gamma_S, \quad (2b)$$

$$\sigma(x, \mathbf{u}) \mathbf{n} = \mathbf{g}_N(x), \quad x \in \Gamma_N, \quad (2c)$$

$$\mathbf{u}(x) = \mathbf{g}_D(x), \quad x \in \Gamma_D, \quad (2d)$$

where $\omega > 0$ denotes the angular frequency, $\rho(x) \geq \rho_0 > 0$ the density, and $\mathbf{B}(x)$ is a real-valued symmetric positive definite matrix. The stress tensor σ is given by

$$\sigma(x, \mathbf{u}) = \mathbf{C}(x) : \varepsilon(x, \mathbf{u}), \quad (3)$$

with the fourth-order real-valued, symmetric positive definite, elastic tensor $\mathbf{C} = (C_{ijkl})_{i,j,k,l=1}^d$, and the linearized strain tensor by

$$\varepsilon(x, \mathbf{u}) = \frac{1}{2}(\nabla \mathbf{u}(x) + \nabla \mathbf{u}^\top(x)). \quad (4)$$

We always assume that the time-harmonic elastic wave equation (2) has a unique solution.

Remark 1.

- i. The one-dimensional elastic wave equation can be written as the acoustic wave equation, where the frequency-domain acoustic equation coincides with the classical Helmholtz equation, which is investigated in [7, 26, 25].
- ii. The linear operator \mathbf{B} is derived from the absorbing boundary conditions [12]. In the one-dimensional case with $\Omega = (0, L)$, the boundary conditions (2b) on Γ_S coincide with the first-order Sommerfeld radiation condition $-u'(x) = i\frac{\omega}{c(x)}u(x)$ on $x = 0$ and $u'(x) = i\frac{\omega}{c(x)}u(x)$ on $x = L$, where $B = \frac{1}{c(x)\rho(x)}$ is a positive scalar function. For higher dimensions ($d = 2, 3$), an explicit form of \mathbf{B} is given in Appendix A. For anisotropic media, we refer to [4].

Instead of solving Equation (2) directly, we consider the CM approach. Let us sketch briefly the method. It is well-known that the time-harmonic elastic wavefield \mathbf{y} of the solution \mathbf{u} of (2), given by

$$\mathbf{y}(x, t) = \operatorname{Re} \{ \mathbf{u}(x) e^{-i\omega t} \}, \quad (5)$$

solves the (real-valued) time-dependent elastic wave equation,

$$\rho(x) \frac{\partial^2 \mathbf{y}(x, t)}{\partial t^2} - \nabla \cdot \sigma(x, \mathbf{y}(\cdot, t)) = \operatorname{Re} \{ \mathbf{f}(x) e^{-i\omega t} \}, \quad x \in \Omega, \quad (6a)$$

$$\rho(x) \mathbf{B}(x) \frac{\partial}{\partial t} \mathbf{y}(x, t) + \sigma(x, \mathbf{y}(\cdot, t)) \mathbf{n} = \operatorname{Re} \{ \mathbf{g}_S(x) e^{-i\omega t} \}, \quad x \in \Gamma_S, \quad (6b)$$

$$\sigma(x, \mathbf{y}(\cdot, t)) \mathbf{n} = \operatorname{Re} \{ \mathbf{g}_N(x) e^{-i\omega t} \}, \quad x \in \Gamma_N, \quad (6c)$$

$$\mathbf{y}(x, t) = \operatorname{Re} \{ \mathbf{g}_D(x) e^{-i\omega t} \}, \quad x \in \Gamma_D, \quad (6d)$$

for $t > 0$, with the initial conditions

$$\mathbf{y}(x, 0) = \mathbf{y}_0(x), \quad x \in \Omega, \quad (6e)$$

$$\frac{\partial \mathbf{y}(x, 0)}{\partial t} = \mathbf{y}_1(x), \quad x \in \Omega, \quad (6f)$$

and

$$\mathbf{y}_0(x) = \operatorname{Re} \{ \mathbf{u}(x) \}, \quad \mathbf{y}_1(x) = \omega \operatorname{Im} \{ \mathbf{u}(x) \}. \quad (7)$$

The leading idea of CM is to seek for $(\mathbf{y}_0, \mathbf{y}_1)$ such that \mathbf{y} is time-harmonic (5). Once such solution \mathbf{y} is found, the solution \mathbf{u} of (2) is simply given by

$$\mathbf{u}(x) = \mathbf{y}(x, 0) + \frac{i}{\omega} \frac{\partial \mathbf{y}(x, 0)}{\partial t} = \mathbf{y}_0(x) + \frac{i}{\omega} \mathbf{y}_1(x). \quad (8)$$

The CM for the elastodynamics equations (2) was first proposed in [38], restricted to sound-soft elastodynamics scattering problems (*i.e.* obstacle problems), for which we have $\mathcal{H}^{d-1}(\Gamma_D), \mathcal{H}^{d-1}(\Gamma_S) > 0$ and $\mathcal{H}^{d-1}(\Gamma_N) = 0$ as illustrated in Figure 1a with $\Gamma_D = \partial D$ and $\Gamma_S = \Gamma \setminus \Gamma_D$. In this context, it is shown in [38] that it is sufficient to find a time-periodic solution \mathbf{y} , namely

$$|\mathbf{y}(\cdot, T) - \mathbf{y}_0|_{\mathbf{C}} = 0, \quad \left\| \frac{\partial}{\partial t} \mathbf{y}(\cdot, T) - \mathbf{y}_1 \right\|_{\rho} = 0, \quad (9)$$

where $|\mathbf{u}|_{\mathbf{C}} = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle_{\mathbf{C}}}$ and $\|\mathbf{v}\|_{\rho} = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle_{\rho}}$, $(\mathbf{u}, \mathbf{v}) \in \mathcal{W}$ with the bilinear forms $(\cdot, \cdot)_{\mathbf{C}} : (H^1(\Omega))^d \times (H^1(\Omega))^d \rightarrow \mathbb{R}$ and $(\cdot, \cdot)_{\rho} : (L^2(\Omega))^d \times (L^2(\Omega))^d \rightarrow \mathbb{R}$,

$$(\mathbf{u}_1, \mathbf{u}_2)_{\mathbf{C}} = \int_{\Omega} \mathbf{C}(x) : \varepsilon(x, \mathbf{u}_1) : \varepsilon(x, \mathbf{u}_2) \, dx, \quad \forall \mathbf{u}_1, \mathbf{u}_2 \in (H^1(\Omega))^d, \quad (10)$$

$$(\mathbf{v}_1, \mathbf{v}_2)_{\rho} = \int_{\Omega} \rho(x) \mathbf{v}_1(x) \cdot \mathbf{v}_2(x) \, dx, \quad \forall \mathbf{v}_1, \mathbf{v}_2 \in (L^2(\Omega))^d. \quad (11)$$

From (9) we conclude that \mathbf{y} also satisfies

$$\begin{aligned} \mathbf{y}(\cdot, T) &= \mathbf{y}_0, \\ \frac{\partial}{\partial t} \mathbf{y}(\cdot, T) &= \mathbf{y}_1, \end{aligned}$$

which is demonstrated later in Lemma 2. Then it is possible to use the fact that the solution of (6) in $[T, 2T]$ with the initial value $(\mathbf{y}(\cdot, T), \frac{\partial \mathbf{y}(\cdot, T)}{\partial t})$ is identically equal to \mathbf{y} in $[0, T]$, that is

$$\mathbf{y}(\cdot, T + t) = \mathbf{y}(\cdot, t), \quad t \in [0, T].$$

This directly follows from the well-posedness of (6) and the periodicity of all source terms. Consequently, \mathbf{y} with the same argument for $[mT, (m+1)T]$, $m \geq 1$, is time periodic in \mathbb{R} . The uniqueness of the time-periodic solution implies that \mathbf{y} satisfies (5) – see Theorem 1 presented in the next Section.

2.2. Generalization

The CM has not been extended to more general boundary conditions, especially for wave propagation problems in heterogeneous media in seismic imaging, with or without free surfaces and obstacles. The main difficulty to perform this generalization is the loss of uniqueness of the time-periodic solution when $\mathcal{H}^{d-1}(\Gamma_D) = 0$ or $\mathcal{H}^{d-1}(\Gamma_S) = 0$, in a similar way as what is shown in the acoustic case in [26, Theorem 1].

Our first contribution is to perform this generalization. We consider general boundary conditions on $\partial\Omega$, where each boundary Γ_D, Γ_N or Γ_S can be empty, and we prove the following theorem.

Theorem 1. *Let $\mathbf{u} \in (H^1(\Omega))^d$ be the unique solution of (2) and $\mathbf{y} \in C^0([0, T]; (H^1(\Omega))^d) \cap C^1([0, T]; (L^2(\Omega))^d)$ be the solution of (6) such that \mathbf{y} satisfies the time periodic condition (9). Then the following assertions hold true:*

(i) *The wavefield \mathbf{y} is given by the Fourier series expansion*

$$\mathbf{y}(x, t) = \operatorname{Re} \{ \mathbf{u}(x) e^{-i\omega t} \} + \frac{t}{T} \boldsymbol{\eta}(x) + \gamma_0(x) + \sum_{|\ell| > 1} \gamma_\ell(x) e^{-i\omega \ell t}, \quad (12)$$

where $\boldsymbol{\eta} = \mathbf{y}(\cdot, T) - \mathbf{y}(\cdot, 0)$ and $\gamma_\ell \in (H^1(\Omega))^d$, $|\ell| \neq 1$, solves the eigenvalue problem,

$$-\nabla \cdot \sigma(x, \gamma_\ell) = (\omega \ell)^2 \rho(x) \gamma_\ell(x), \quad x \in \Omega, \quad (13a)$$

$$-i\omega \ell \rho(x) \mathbf{B}(x) \gamma_\ell(x) + \sigma(x, \gamma_\ell) \mathbf{n} = \mathbf{0}, \quad x \in \Gamma_S, \quad (13b)$$

$$\sigma(x, \gamma_\ell) \mathbf{n} = \mathbf{0}, \quad x \in \Gamma_N, \quad (13c)$$

$$\gamma_\ell(x) = \mathbf{0}, \quad x \in \Gamma_D. \quad (13d)$$

Moreover, $\gamma_0, \boldsymbol{\eta} \in \mathcal{U}$ with

$$\mathcal{U} := \operatorname{Span} \{ \mathbf{e}_1, \dots, \mathbf{e}_d \} \oplus \operatorname{Span} \{ x_j \mathbf{e}_i - x_i \mathbf{e}_j \mid 1 \leq j < i \leq d \}. \quad (14)$$

Here \mathbf{e}_i denotes the i -th unit vector in \mathbb{R}^d .

(ii) *If $\mathcal{H}^{d-1}(\Gamma_D) > 0$, then $\boldsymbol{\eta} = \gamma_0 = \mathbf{0}$ a.e. in Ω .*

(iii) *If $\mathcal{H}^{d-1}(\Gamma_S) > 0$, then $\boldsymbol{\eta} = \gamma_\ell = \mathbf{0}$ a.e. in Ω , $|\ell| > 1$.*

Remark 2.

i. *Theorem 1 implicitly implies that sound-soft scatterings have a unique time-periodic solution which is identically equal to the time-harmonic solution.*

ii. *The elliptic solutions γ_0 and $\boldsymbol{\eta}$ both of (13) with $\ell = 0$ lie in the kernel \mathcal{U} of the linear operator $-\nabla \cdot \sigma$.*

For the proof of Theorem 1 we need the following lemmata:

Lemma 1. *Let $\boldsymbol{\tau} \in (H^1(\Omega))^d$ solve*

$$-\nabla \cdot \sigma(x, \boldsymbol{\tau}) = \mathbf{0}, \quad x \in \Omega, \quad (15a)$$

$$\sigma(x, \boldsymbol{\tau}) \mathbf{n} = \mathbf{0}, \quad x \in \Gamma_S \cup \Gamma_N, \quad (15b)$$

$$\boldsymbol{\tau}(x) = \mathbf{0}, \quad x \in \Gamma_D. \quad (15c)$$

Then $\boldsymbol{\tau} \in \mathcal{U}$. Moreover, $\boldsymbol{\tau} = \mathbf{0}$ a.e. in Ω when $\mathcal{H}^{d-1}(\Gamma_D) > 0$.

Proof of Lemma 1. Let $\boldsymbol{\tau} \in (H^1(\Omega))^d$ be a solution of (15). Then the divergence theorem, together with the boundary conditions, yields

$$\begin{aligned} 0 &= \int_{\Omega} (\nabla \cdot \sigma(\boldsymbol{\tau})) \cdot \boldsymbol{\tau} \, dx = \int_{\partial\Omega} (\sigma(\boldsymbol{\tau})\mathbf{n}) \cdot \boldsymbol{\tau} \, ds - \int_{\Omega} \sigma(\boldsymbol{\tau}) : \nabla \boldsymbol{\tau} \, dx \\ &= - \int_{\Omega} (\mathbf{C} : \varepsilon(\boldsymbol{\tau})) : \varepsilon(\boldsymbol{\tau}) \, dx \leq 0. \end{aligned}$$

From the symmetry and positive definiteness of the elastic tensor \mathbf{C} we have

$$\varepsilon(\boldsymbol{\tau}) = \mathbf{0} \quad \text{a.e. in } \Omega,$$

and consequently $\nabla \boldsymbol{\tau}$ is skew-symmetric,

$$\frac{\partial}{\partial x_i} \tau_i = 0, \quad i = 1, \dots, d, \quad (16)$$

$$\frac{\partial}{\partial x_j} \tau_i = -\frac{\partial}{\partial x_i} \tau_j \quad i, j = 1, \dots, d, \quad j \neq i. \quad (17)$$

This implies

$$\frac{\partial^2}{\partial^2 x_i} \tau_j = -\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} \tau_i = -\frac{\partial}{\partial x_j} \frac{\partial}{\partial x_i} \tau_i = 0, \quad i, j = 1, \dots, d. \quad (18)$$

We conclude that $\frac{\partial}{\partial x_i} \tau_j$ is independent of x_i , and moreover τ_j depends linearly on x_i , $i = 1, \dots, d$. Thus there exist $\alpha_i, \beta_{i,j} \in \mathbb{R}$ such that

$$\boldsymbol{\tau}(x) = \sum_{i=1}^d \left(\alpha_i + \sum_{j=1}^d \beta_{i,j} x_j \right) \mathbf{e}_i. \quad (19)$$

Equations (16) and (17) yield $\beta_{i,i} = 0$, $i = 1, \dots, d$, and

$$\beta_{i,j} = \frac{\partial}{\partial x_j} \tau_i = -\frac{\partial}{\partial x_i} \tau_j = -\beta_{j,i}, \quad i, j = 1, \dots, d. \quad (20)$$

This results

$$\boldsymbol{\tau}(x) = \sum_{i=1}^d \alpha_i \mathbf{e}_i + \sum_{1 \leq j < i \leq d} \beta_{i,j} (x_j \mathbf{e}_i - x_i \mathbf{e}_j) \in \mathcal{U},$$

which proves the first assertion.

Suppose that $\mathcal{H}^{d-1}(\Gamma_D) > 0$. Then $\tau_i = 0$ on Γ_D , $i = 1, \dots, d$, and we have $\tau_i = 0$ in Ω , since τ_i is a linear polynomial. This completes the proof. \square

Lemma 2. Let \mathbf{y} be the solution of (6) with $\mathcal{H}^{d-1}(\Gamma_D \cup \Gamma_S) > 0$ and satisfy (9). Then \mathbf{y} is $T = \frac{2\pi}{\omega}$ -periodic in the sense of

$$\mathbf{y}(x, T) = \mathbf{y}(x, 0), \quad \frac{\partial \mathbf{y}(x, T)}{\partial t} = \frac{\partial \mathbf{y}(x, 0)}{\partial t}, \quad x \in \Omega. \quad (21)$$

Proof of Lemma 2. Let $\boldsymbol{\eta} = \mathbf{y}(\cdot, T) - \mathbf{y}(\cdot, 0)$. Then the periodic condition (9), together with the symmetry and definiteness of the tensor \mathbf{C} , yields $\varepsilon(x, \boldsymbol{\eta}) = 0$

a.e. in Ω , and thus $\boldsymbol{\eta}$ obviously solves (15).

From Lemma 1 $\boldsymbol{\eta} \in \mathcal{U}$ and in the presence of Dirichlet boundary conditions $\boldsymbol{\eta}$ vanishes identically on the boundary Γ_D . Since $\boldsymbol{\eta}$ is a linear polynomial in Ω , $\boldsymbol{\eta} = \mathbf{0}$ a.e. in Ω , which shows Lemma 2 for $\mathcal{H}^{d-1}(\Gamma_D) > 0$.

Let Γ_D be empty and $\mathcal{H}^{d-1}(\Gamma_S) > 0$. Then, by integrating the elastic wave equation (6) multiplied with $\boldsymbol{\eta}$, the periodicity of the harmonic source terms and $\frac{\partial \mathbf{y}}{\partial t}$, and integration by parts, we obtain

$$\begin{aligned} 0 &= \int_0^T \int_{\Omega} \rho(x) \frac{\partial^2 \mathbf{y}(x, t)}{\partial^2 t} \cdot \boldsymbol{\eta}(x) \, dx \, dt - \int_0^T \int_{\Omega} \nabla \cdot \sigma(x, \mathbf{y}(\cdot, t)) \cdot \boldsymbol{\eta}(x) \, dx \, dt \\ &= \int_{\Omega} \rho(x) \left(\frac{\partial \mathbf{y}(x, T)}{\partial t} - \frac{\partial \mathbf{y}(x, 0)}{\partial t} \right) \cdot \boldsymbol{\eta}(x) \, dx + \int_0^T \int_{\Gamma_S} \sigma(x, \mathbf{y}(\cdot, t)) \mathbf{n} \cdot \boldsymbol{\eta}(x) \, ds \, dt \\ &= - \int_0^T \int_{\Gamma_S} \rho(x) \mathbf{B}(x) \frac{\partial \mathbf{y}(x, t)}{\partial t} \cdot \boldsymbol{\eta}(x) \, ds \, dt \\ &= - \int_{\Gamma_S} \rho(x) \mathbf{B}(x) \boldsymbol{\eta}(x) \cdot \boldsymbol{\eta}(x) \, ds. \end{aligned}$$

The symmetry and positive definiteness of the linear operator \mathbf{B} yields $\boldsymbol{\eta}(x) = \mathbf{0}$ a.e. on Γ_S , which again implies with Lemma 1 that $\boldsymbol{\eta} = \mathbf{0}$ a.e. in Ω and completes the proof. \square

Proof of Theorem 1. (i) We consider two situations: $\mathcal{H}^{d-1}(\Gamma_D \cup \Gamma_S) > 0$ and $\Gamma_N = \partial\Omega$.

Case 1: Let $\mathcal{H}^{d-1}(\Gamma_D \cup \Gamma_S) > 0$ and

$$\mathbf{z}(x, t) = \mathbf{y}(x, t) - \operatorname{Re} \{ \mathbf{u}(x) e^{-i\omega t} \}. \quad (22)$$

From Lemma 2 \mathbf{z} is clearly T -time periodic and solves the elastic wave equation (6) with homogeneous right hand sides and the initial conditions

$$\mathbf{z}_0(x) = \mathbf{y}_0(x) - \operatorname{Re} \{ \mathbf{u}(x) \}, \quad x \in \Omega, \quad (23)$$

$$\mathbf{z}_1(x) = \mathbf{y}_1(x) - \omega \operatorname{Im} \{ \mathbf{u}(x) \}, \quad x \in \Omega. \quad (24)$$

We can expand it in Fourier series as

$$\mathbf{z}(x, t) = \sum_{|\ell| \geq 0} \boldsymbol{\gamma}_{\ell}(x) e^{-i\omega \ell t} \quad (25)$$

with the complex-valued Fourier coefficients

$$\boldsymbol{\gamma}_{\ell} = \frac{1}{T} \int_0^T \mathbf{z}(\cdot, t) e^{i\omega \ell t} \, dt \in (H^1(\Omega))^d. \quad (26)$$

Next, we verify that $\boldsymbol{\gamma}_{\ell}$ solves (13), $\ell \in \mathbb{Z}$. The integration by parts, together with the periodicity of \mathbf{z} , the definition (26), and equation (6) with no external source term, yields

$$\begin{aligned} -(\omega \ell)^2 \rho \boldsymbol{\gamma}_{\ell} &= \frac{\rho}{T} \int_0^T \mathbf{z}(\cdot, t) (-\omega \ell)^2 e^{i\omega \ell t} \, dt = \frac{\rho}{T} \int_0^T \frac{\partial^2}{\partial^2 t} \mathbf{z}(\cdot, t) e^{i\omega \ell t} \, dt, \\ \nabla \cdot \sigma(\cdot, \boldsymbol{\gamma}_{\ell}) &= \frac{1}{T} \int_0^T \nabla \cdot \sigma(\cdot, \mathbf{z}(\cdot, t)) e^{i\omega \ell t} \, dt = \frac{\rho}{T} \int_0^T \frac{\partial^2}{\partial^2 t} \mathbf{z}(\cdot, t) e^{i\omega \ell t} \, dt. \end{aligned}$$

Consequently, we have

$$-(\omega\ell)^2 \rho(x) \gamma_\ell(x) - \nabla \cdot \sigma(x, \gamma_\ell) = \mathbf{0}. \quad (27)$$

Again, integration by parts in time and the boundary conditions (6b)–(6d) with $\mathbf{g}_D = \mathbf{g}_S = \mathbf{g}_N = \mathbf{0}$, together with the definition (26), imply that γ_ℓ satisfies the boundary conditions (13b)–(13d), whose fulfill (13) for $|\ell| \geq 1$. In particular, for $\ell = 1$, (13) coincides with (2) and without any source terms, and hence the uniqueness of \mathbf{u} implies that $\gamma_1 = \overline{\gamma_{-1}} = \mathbf{0}$ a.e. in Ω . For $\ell = 0$ and $\tau = \gamma_0$ we have

$$\begin{aligned} \nabla \cdot \sigma(x, \tau) &= \mathbf{0}, & x \in \Omega, \\ \sigma(x, \tau) \mathbf{n} &= \mathbf{0}, & x \in \Gamma_S \cup \Gamma_N, \\ \tau(x) &= \mathbf{0}, & x \in \Gamma_D. \end{aligned}$$

Lemma 1 ends the proof for $\mathcal{H}^{d-1}(\Gamma_D \cup \Gamma_S) > 0$.

Case 2: Let $\Gamma_N = \partial\Omega$ and $\Gamma_D = \Gamma_S = \emptyset$ and $\mathbf{z}(x, t) = \mathbf{y}(x, t) - \frac{t}{T} \boldsymbol{\eta}(x)$. Then, by the definition of $\boldsymbol{\eta}$ and \mathbf{z} and the periodicity of \mathbf{y} , we have

$$\begin{aligned} \mathbf{z}(x, T) - \mathbf{z}(x, 0) &= \mathbf{y}(x, T) - \mathbf{y}(x, 0) - \boldsymbol{\eta}(x) = \mathbf{0}, \\ \frac{\partial \mathbf{z}(x, T)}{\partial t} - \frac{\partial \mathbf{z}(x, 0)}{\partial t} &= \frac{\partial \mathbf{z}(x, T)}{\partial t} - \frac{1}{T} \boldsymbol{\eta}(x) - \frac{\partial \mathbf{z}(x, 0)}{\partial t} + \frac{1}{T} \boldsymbol{\eta}(x) = \mathbf{0}. \end{aligned}$$

From $\sigma(x, \boldsymbol{\eta}) = \mathbf{C} : \varepsilon(x, \boldsymbol{\eta}) = \mathbf{0}$ in Ω and $\sigma(x, \boldsymbol{\eta}) \mathbf{n} = \mathbf{0}$ on Γ_N , it is easy to verify that \mathbf{z} is a T -time periodic solution of (6) with the Neumann boundary conditions on the entire boundary $\partial\Omega$. Then the same argument in the first case with $\mathcal{H}^{d-1}(\Gamma_D \cup \Gamma_S) > 0$ implies that

$$\mathbf{y}(x, t) = \mathbf{z}(x, t) + \frac{t}{T} \boldsymbol{\eta}(x) = \operatorname{Re} \{ \mathbf{u}(x) e^{-i\omega t} \} + \frac{t}{T} \boldsymbol{\eta}(x) + \gamma_0(x) + \sum_{|\ell| > 1} \gamma_\ell e^{-i\omega \ell t}$$

which completes the first assertion.

(ii) The second assertion follows directly from (i) and Lemma 1.

(iii) From the first assertion and Lemma 1, we have $\boldsymbol{\eta} = \mathbf{0}$ a.e. in Ω . The boundary value problem (13) with $\mathcal{H}^{d-1}(\Gamma_S) > 0$ has only the trivial solution [16, 15, 14, 6], which implies that all γ_ℓ vanishes identically, $|\ell| > 1$. This completes the proof. \square

From Theorem 1, it is clear that the time-periodic solution of (6) is in general not unique, except for sound-soft scattering problems with $\mathcal{H}^{d-1}(\Gamma_D), \mathcal{H}^{d-1}(\Gamma_S) > 0$. To restore uniqueness, further post-processing are required: orthogonal projection and filtering method.

2.2.1. Orthogonal projection

In an unbounded domain, $\mathcal{H}^{d-1}(\Gamma_S) > 0$, the time-independent elastic wave equation (2) has a unique solution \mathbf{u} . However, in the absence of Dirichlet boundary conditions, $\Gamma_D = \emptyset$, Theorem 1 implies that for any solution \mathbf{y} of (6) satisfying the time-periodic condition (9) there is a spurious function $\gamma_0 \in \mathcal{U}$ such that

$$\mathbf{y}(x, t) = \operatorname{Re} \{ \mathbf{u}(x) \exp(-i\omega t) \} + \gamma_0(x),$$

in particular,

$$\mathbf{v}(x) := \mathbf{y}(x, 0) + \frac{i}{\omega} \frac{\partial}{\partial t} \mathbf{y}(x, 0) = \mathbf{u}(x) + \gamma_0(x). \quad (28)$$

In order to determine γ_0 , let $\{\psi_\ell\}_{\ell=1}^{\frac{d(d+1)}{2}}$ be an orthonormal basis of \mathcal{U} with respect to the L^2 -inner product (\cdot, \cdot) and $\alpha_\ell \in \mathbb{R}$ such that

$$\gamma_0(x) = \sum_{\ell=1}^{\frac{d(d+1)}{2}} \alpha_\ell \psi_\ell(x). \quad (29)$$

For example, a basis of \mathcal{U} is given as follows:

$$\begin{aligned} & \{\mathbf{e}_1, \mathbf{e}_2, (y \ -x)^\top\} \quad \text{for } d=2, \\ & \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, (y \ -x \ 0)^\top, (z \ 0 \ -x)^\top, (0 \ z \ -y)^\top\} \quad \text{for } d=3 \end{aligned} \quad (30)$$

Orthonormality is obtained by applying the modified Gram-Schmidt method to the basis (30). Then the variational formulation, derived directly from the boundary value problem (2), yields

$$\mathbf{0} = (\mathbf{f} + \omega^2 \mathbf{v} + \nabla \cdot \sigma(\mathbf{v}) - \omega^2 \gamma_0, \psi_\ell) = (\mathbf{f} + \omega^2 \mathbf{v} + \nabla \cdot \sigma(\mathbf{v}), \psi_\ell) - \omega^2 \alpha_\ell,$$

for $\ell = 1, \dots, \frac{d(d+1)}{2}$ and moreover

$$\alpha_\ell = \frac{(\mathbf{f} + \omega^2 \mathbf{v} + \nabla \cdot \sigma(\mathbf{v}), \psi_\ell)}{\omega^2}.$$

Finally, this gives, together with (28) and (29),

$$\mathbf{u} = \mathbf{v} - \gamma_0 = \mathbf{v} - \frac{1}{\omega^2} \sum_{\ell=1}^{\frac{d(d+1)}{2}} (\mathbf{f} + \omega^2 \mathbf{v} + \nabla \cdot \sigma(\mathbf{v}), \psi_\ell) \psi_\ell. \quad (31)$$

Remark 3. Here we refer (31) to the compatibility condition in [26, 25] for acoustic wave equations with instead $\mathcal{U} = \mathbb{P}^0$.

2.2.2. Filtering procedure

From Theorem 1 a time-periodic solution of the time-dependent elastic wave equation (6) may contain higher-order eigenmodes γ_ℓ , $|\ell| > 1$, i.e. $\Gamma_S = \emptyset$. The filtering procedure based on the Fourier transformation [55, 25, 10],

$$\hat{\mathbf{y}}(x) = \oint_0^T \left(\mathbf{y}(x, t) + \frac{i}{\omega} \frac{\partial}{\partial t} \mathbf{y}(x, t) \right) e^{i\omega t} dt = 2 \oint_0^T \mathbf{y}(x, t) e^{i\omega t} dt, \quad (32)$$

where

$$\oint_0^T f(t) dt = \frac{1}{T} \int_0^T f(t) dt \quad (33)$$

denotes the average of a function f over the interval $[0, T]$, is designed to filter out all higher modes γ_ℓ , $|\ell| \neq 1$, and the constant in the Fourier expansion (12).

Let \mathbf{y} be a T -time periodic solution of the elastic wave equation (6). Due to the mutually orthogonality of the different modes,

$$\oint_0^T \gamma_0(x) e^{i\omega t} dt = \gamma_0(x) \oint_0^T e^{i\omega t} dt = \mathbf{0} \quad (34)$$

and

$$\oint_0^T \gamma_\ell(x) e^{-i\omega \ell t} e^{i\omega t} dt = \gamma_\ell(x) \oint_0^T e^{-i\omega \ell t} e^{i\omega t} dt = \mathbf{0}, \quad |\ell| > 1. \quad (35)$$

Hence, by putting the expansion in (12) of \mathbf{y} into (32), the integration of the harmonic function γ_0 and the higher eigenmodes γ_ℓ , $|\ell| > 1$, vanish identically, and thus we have

$$\hat{\mathbf{y}} = \oint_0^T \left(\operatorname{Re}\{\mathbf{u} e^{-i\omega t}\} + \frac{i}{\omega} \frac{\partial}{\partial t} \operatorname{Re}\{\mathbf{u} e^{-i\omega t}\} \right) e^{i\omega t} dt = \mathbf{u} - \frac{i}{2\pi} \boldsymbol{\eta}.$$

Moreover, for $\mathcal{H}^{d-1}(\Gamma_D \cup \Gamma_S) > 0$, where $\boldsymbol{\eta} = \mathbf{0}$ a.e. in Ω , we have

$$\hat{\mathbf{y}}(x) = 2 \oint_0^T \mathbf{y}(x, t) e^{i\omega t} dt = \mathbf{u}(x), \quad x \in \Omega, \quad (36)$$

which restores the uniqueness.

Let $\partial\Omega = \Gamma_N$. Then we again consider as in Section 2.2.1 an orthonormal basis $\{\boldsymbol{\psi}_\ell\}$ of \mathcal{U} and we write

$$-\frac{i}{2\pi} \boldsymbol{\eta} = \sum_{\ell=1}^{\frac{d(d+1)}{2}} \alpha_\ell \boldsymbol{\psi}_\ell, \quad \alpha_\ell \in \mathbb{C}.$$

The variational formulation yields

$$\alpha_\ell = \frac{(\mathbf{f} + \omega^2 \hat{\mathbf{y}} + \nabla \cdot \sigma(\hat{\mathbf{y}}), \boldsymbol{\psi}_\ell)}{\omega^2}$$

and

$$\mathbf{u} = \hat{\mathbf{y}} - \frac{1}{\omega^2} \sum_{\ell=1}^{\frac{d(d+1)}{2}} (\mathbf{f} + \omega^2 \hat{\mathbf{y}} + \nabla \cdot \sigma(\hat{\mathbf{y}}), \boldsymbol{\psi}_\ell) \boldsymbol{\psi}_\ell. \quad (37)$$

We summarize the above results in the following propositions:

Proposition 1. *Let $\mathbf{y} \in C^0([0, T]; (H^1(\Omega))^d) \cap C^1([0, T]; (L^2(\Omega))^d)$ be a solution of (6) with $\mathcal{H}^{d-1}(\Gamma_S \cup \Gamma_D) > 0$ satisfying the time periodic condition (9). Then $\hat{\mathbf{y}}$ given by (32) solves (2).* \square

Proposition 2. *Let $\mathbf{y} \in C^0([0, T]; (H^1(\Omega))^d) \cap C^1([0, T]; (L^2(\Omega))^d)$ be a time-dependent solution of (6) with $\partial\Omega = \Gamma_N$, which satisfies the time periodic condition (9). Then the solution \mathbf{u} of (2) is given by (37).* \square

3. Exact controllability formulation

In order to find a T -time periodic solution of (6) to construct the frequency-domain solution of (2), we propose the controllability approach. This yields the PDE-constrained least-squares problem,

$$\min_{(\mathbf{y}_0, \mathbf{y}_1) \in \mathcal{V}_D} \mathcal{J}(\mathbf{y}_0, \mathbf{y}_1) \quad (38)$$

with $\mathcal{V}_D \subset \{(\mathbf{y}_0, \mathbf{y}_1) \in \mathcal{W} : \mathbf{y}_0 + \frac{i}{\omega} \mathbf{y}_1 = \mathbf{g}_D \text{ on } \Gamma_D\}$ and $\mathcal{W} = (H^1(\Omega))^d \times (L^2(\Omega))^d$.

We consider the cost functional $\mathcal{J} : \mathcal{V}_D \rightarrow \mathbb{R}_{\geq 0}$ defined by

$$\mathcal{J}(\mathbf{y}_0, \mathbf{y}_1) = \frac{1}{2} |\mathbf{y}(T) - \mathbf{y}_0|_{\mathbf{C}}^2 + \frac{1}{2} \|\dot{\mathbf{y}}(T) - \mathbf{y}_1\|_{\rho}^2, \quad (39)$$

where the state variable $\mathbf{y} = \mathbf{y}[(\mathbf{y}_0, \mathbf{y}_1)]$ solves the elastic wave equation (6) with the initial value, called control variable, $(\mathbf{y}_0, \mathbf{y}_1)$. For simplicity, we use the notation $\dot{\mathbf{y}}(x, t)$ for $\frac{\partial \mathbf{y}(x, t)}{\partial t}$ and we write $\mathbf{y}(t)$, $\dot{\mathbf{y}}(t)$ instead of $\mathbf{y}(x, t)$, $\dot{\mathbf{y}}(x, t)$, respectively, when there is no confusion from context.

To solve the optimization problem (38), we apply the conjugate gradient (CG) method, which requires the derivatives of the functional \mathcal{J} . This formulation includes the finite element (FE) or spectral element (SE) formulation, e.g. by choosing \mathcal{V}_D equal to the FE- or SE-space.

3.1. Fréchet Derivative

We first derive the Fréchet derivative of \mathcal{J} with the adjoint state technique. In doing so, we introduce $(\cdot, \cdot)_{\mathcal{E}} : \mathcal{W} \times \mathcal{W} \rightarrow \mathbb{R}$ and $|\cdot|_{\mathcal{E}} : \mathcal{W} \rightarrow \mathbb{R}_{\geq 0}$ with

$$((\mathbf{u}_1, \mathbf{v}_1), (\mathbf{u}_2, \mathbf{v}_2))_{\mathcal{E}} = (\mathbf{u}_1, \mathbf{u}_2)_{\mathbf{C}} + (\mathbf{v}_1, \mathbf{v}_2)_{\rho}, \quad (\mathbf{u}_1, \mathbf{v}_1), (\mathbf{u}_2, \mathbf{v}_2) \in \mathcal{W}, \quad (40a)$$

$$|(\mathbf{u}, \mathbf{v})|_{\mathcal{E}} = \sqrt{((\mathbf{u}, \mathbf{v}), (\mathbf{u}, \mathbf{v}))_{\mathcal{E}}}, \quad (\mathbf{u}, \mathbf{v}) \in \mathcal{W}. \quad (40b)$$

Then the cost functional \mathcal{J} in (39) is given by

$$\mathcal{J}(\mathbf{y}_0, \mathbf{y}_1) = \frac{1}{2} |(\mathbf{y}(T) - \mathbf{y}_0, \dot{\mathbf{y}}(T) - \mathbf{y}_1)|_{\mathcal{E}}^2.$$

For any perturbation $(\mathbf{w}_0, \mathbf{w}_1) \in \mathcal{V} \subset \{(\mathbf{v}_0, \mathbf{v}_1) \in \mathcal{W} : \mathbf{v}_0 = \mathbf{v}_1 = 0 \text{ on } \Gamma_D\}$ and $(\mathbf{y}_0, \mathbf{y}_1) \in \mathcal{V}_D$ we have

$$\begin{aligned} \langle \mathcal{J}'(\mathbf{y}_0, \mathbf{y}_1), (\mathbf{w}_0, \mathbf{w}_1) \rangle &= ((\mathbf{y}(T) - \mathbf{y}_0, \dot{\mathbf{y}}(T) - \mathbf{y}_1), (\mathbf{w}(T) - \mathbf{w}_0, \dot{\mathbf{w}}(T) - \mathbf{w}_1))_{\mathcal{E}} \\ &= -(\mathbf{y}(T) - \mathbf{y}_0, \mathbf{w}_0)_{\mathbf{C}} - (\dot{\mathbf{y}}(T) - \mathbf{y}_1, \mathbf{w}_1)_{\rho} \\ &\quad + (\mathbf{y}(T) - \mathbf{y}_0, \mathbf{w}(T))_{\mathbf{C}} + (\dot{\mathbf{y}}(T) - \mathbf{y}_1, \dot{\mathbf{w}}(T))_{\rho}, \end{aligned} \quad (41)$$

where the corresponding time-dependent solution \mathbf{w} of the homogeneous wave equation (6) with $\mathbf{f} = \mathbf{g}_S = \mathbf{g}_D = \mathbf{g}_N = 0$ and with the initial value $(\mathbf{w}_0, \mathbf{w}_1)$. We note that \mathcal{J}' lies in the dual space \mathcal{V}' of \mathcal{V} with respect to the dual pairing $\langle \cdot, \cdot \rangle$.

Next, let us introduce the solution \mathbf{z} of the adjoint or backward wave equation

$$\rho(x) \frac{\partial^2 \mathbf{z}(x, t)}{\partial^2 t} - \nabla \cdot \sigma(x, \mathbf{z}(\cdot, t)) = \mathbf{0}, \quad x \in \Omega, t > 0, \quad (42a)$$

$$\rho(x) \mathbf{B}(x) \frac{\partial}{\partial t} \mathbf{z}(x, t) + \sigma(x, \mathbf{z}(\cdot, t)) \mathbf{n} = \mathbf{0}, \quad x \in \Gamma_S, t > 0, \quad (42b)$$

$$\sigma(x, \mathbf{z}(\cdot, t)) \mathbf{n} = \mathbf{0}, \quad x \in \Gamma_N, t > 0, \quad (42c)$$

$$\mathbf{z}(x, t) = \mathbf{0}, \quad x \in \Gamma_D, t > 0, \quad (42d)$$

with the initial conditions

$$\mathbf{z}(x, 0) = \dot{\mathbf{y}}(x, T) - \mathbf{y}_1(x), \quad x \in \Omega, \quad (42e)$$

$$\begin{aligned} (\dot{\mathbf{z}}(\cdot, 0), \boldsymbol{\varphi})_\rho &= (\mathbf{y}(\cdot, T) - \mathbf{y}_0, \boldsymbol{\varphi})_{\mathbf{C}} \\ &\quad - (\rho \mathbf{B}(\dot{\mathbf{y}}(\cdot, T) - \mathbf{y}_1), \boldsymbol{\varphi})_{L^2(\Gamma_S)} \quad \forall \boldsymbol{\varphi} \in (H^1(\Omega))^d. \end{aligned} \quad (42f)$$

which again exactly coincides with the homogeneous elastic wave equation (6) with $\mathbf{f} = \mathbf{g}_D = \mathbf{g}_N = \mathbf{g}_S = \mathbf{0}$. Then, by applying (6a) on the perturbation \mathbf{w} , multiplied with the backward solution $\mathbf{z}(x, T-t)$ over $\Omega \times (0, T)$, the integration by parts results

$$\begin{aligned} 0 &= \int_0^T \int_\Omega \left(\rho(x) \frac{\partial^2 \mathbf{w}(x, t)}{\partial^2 t} - \nabla \cdot \sigma(x, \mathbf{w}(x, t)) \right) \cdot \mathbf{z}(x, T-t) \, dx \, dt \\ &= \int_{\Gamma_S} \rho(x) \left[(\mathbf{B} \mathbf{z}(x, T-t)) \cdot \mathbf{w}(x, t) \right] \Big|_{t=0}^{t=T} ds \\ &\quad + \int_\Omega \rho(x) \left[\frac{\partial \mathbf{w}(x, t)}{\partial t} \cdot \mathbf{z}(x, T-t) + \frac{\partial \mathbf{z}(x, T-t)}{\partial t} \cdot \mathbf{w}(x, t) \right] \Big|_{t=0}^{t=T} dx. \end{aligned}$$

This, together with the backward wave equation (42a) with the boundary conditions (42b)–(42d) and the initial conditions (42e)–(42f), gives

$$\begin{aligned} &(\dot{\mathbf{z}}(T), \mathbf{w}_0)_\rho + (\rho \mathbf{B} \mathbf{z}(T), \mathbf{w}_0)_{L^2(\Gamma_S)} + (\mathbf{z}(T), \mathbf{w}_1)_\rho \\ &= (\dot{\mathbf{z}}(0), \mathbf{w}(T))_\rho + (\rho \mathbf{B} \mathbf{z}(0), \mathbf{w}(T))_{L^2(\Gamma_S)} + (\mathbf{z}(0), \dot{\mathbf{w}}(T))_\rho \\ &= (\mathbf{y}(T) - \mathbf{y}_0, \mathbf{w}(T))_{\mathbf{C}} + (\dot{\mathbf{y}}(T) - \mathbf{y}_1, \dot{\mathbf{w}}(T))_\rho. \end{aligned}$$

This, Equation (41), and the initial conditions (42e)–(42f) immediately yield

$$\begin{aligned} \langle \mathcal{J}'(\mathbf{y}_0, \mathbf{y}_1), (\mathbf{w}_0, \mathbf{w}_1) \rangle &= (\dot{\mathbf{z}}(T) - \dot{\mathbf{z}}(0), \mathbf{w}_0)_\rho + (\rho \mathbf{B}(\mathbf{z}(T) - \mathbf{z}(0)), \mathbf{w}_0)_{L^2(\Gamma_S)} \\ &\quad + (\mathbf{z}(T) - \mathbf{z}(0), \mathbf{w}_1)_\rho. \end{aligned} \quad (43)$$

Interestingly, we note that the computation of the gradient for the **CM requires** the solution of one forward and one backward wave propagation problem, in a similar fashion as what is needed for the gradient computation in FWI [45]. The difference being that here the forward and backward wave equations are solved only over a single period of time, while they are computed over the whole time interval in FWI.

3.2. Controllability method combined with the conjugate gradient method

To solve the quadratic optimization problem (38) of the convex cost functional \mathcal{J} , we propose to use the conjugate gradient (CG) method [29, 35]. Each CG iteration requires to compute the gradient \mathcal{J}' given by (43), which needs the forward and backward solution of the wave equation (6). However, the gradient \mathcal{J}' only lies in the dual space \mathcal{V}' , and thus we need to find a (Riesz)-representative \mathbf{p} in \mathcal{V} ,

$$\langle \mathbf{p}, \varphi \rangle_{\mathcal{B}} = \langle \mathbf{g}, \varphi \rangle, \quad \forall \varphi \in \mathcal{V}, \quad (44)$$

where $\mathcal{B} : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ is a (suitable) coercive, bounded, and symmetric bilinear form. This leads to solve a further elliptic problem at each CG iteration.

The standard inner product, $\langle \cdot, \cdot \rangle_{\mathcal{B}} = (\cdot, \cdot)_{\mathcal{V}}$, refers to looking for $\mathbf{p} \in \mathcal{V}$ such that

$$(\mathbf{p}_0, \varphi)_{(H^1(\Omega))^d} = \langle \mathbf{g}_0, \varphi \rangle, \quad \varphi \in (H^1(\Omega))^d, \quad (45a)$$

$$\mathbf{p}_1 = \mathbf{g}_1. \quad (45b)$$

Any choice of scalar products in \mathcal{V} can also be used in (45). A natural choice is to consider the energy $(\cdot, \cdot)_{\mathcal{E}}$ in (40). In the presence of Dirichlet boundary conditions, $\mathcal{H}^{d-1}(\Gamma_D) > 0$, $(\cdot, \cdot)_{\mathcal{E}}$ is a coercive bilinear form on \mathcal{V} [38] so that the following elliptic problem has a unique solution:

$$-\nabla \cdot \sigma(x, \mathbf{p}_0) = \mathbf{g}_0(x), \quad x \in \Omega, \quad (46a)$$

$$\sigma(x, \mathbf{p}_0) \mathbf{n} = 0, \quad x \in \Gamma_S \cup \Gamma_N, \quad (46b)$$

$$\mathbf{p}_0(x) = 0, \quad x \in \Gamma_D, \quad (46c)$$

and $\mathbf{p}_1 = \mathbf{g}_1$. However, for $\Gamma_D = \emptyset$ the elliptic problem (46) is ill-posed. To recover the well-posedness, we need a compatibility condition such as

$$\begin{aligned} \int_{\Omega} \mathbf{g}_0(x) \cdot \psi(x) \, dx &= - \int_{\Omega} (\nabla \cdot \sigma(x, \mathbf{p}_0)) \cdot \psi(x) \, dx \\ &= \int_{\Omega} \sigma(x, \mathbf{p}_0) : \varepsilon(x, \psi) \, dx - \int_{\partial\Omega} \sigma(x, \mathbf{p}_0) \mathbf{n} \cdot \psi(x) \, ds = 0 \end{aligned}$$

for all $\psi \in \mathcal{U}$ in (14). Thus, we first project \mathbf{g}_0 into the range of $\nabla \cdot \sigma$ by eliminating each basis function ψ_{ℓ} in \mathcal{U} given in (30). By doing this, we apply the orthogonal projection,

$$\mathbf{g}_0 \leftarrow \mathbf{g}_0 - \frac{(\psi_{\ell}, \mathbf{g}_0)}{(\psi_{\ell}, \psi_{\ell})} \psi_{\ell}, \quad \ell = 1, 2, \dots, \frac{d(d+1)}{2}. \quad (48)$$

On top of the solution of one forward and one backward wave equation over a single period of time, an elliptic problem thus needs to be solved at each iteration of the CM technique, to find a (Riesz)-representative of the gradient. We propose to use the inner product $(\cdot, \cdot)_{\mathcal{E}}$ of (46)–(48), which is also used in the definition of the cost function \mathcal{J} and thus makes this choice more intuitive. This inner product can be considered as a preconditioner of the CG loop. Also, our numerical experiments indicate that the use of $(\cdot, \cdot)_{\mathcal{E}}$ can improve the convergence speed compared to using a standard Euclidean inner product, especially when considering heterogeneous media.

In order to avoid solving the elliptic problem, a mixed or first-order formulation of the acoustic wave equation was proposed instead by Glowinski et al. in [23] and later in [22, 25, 10]. For elastic equations, the conventional first-order velocity-stress formulation could be used in the same spirit. The advantage of this formulation is that the gradient of the corresponding function based on the first-order formulation is in a self-dual space, so no additional Riesz representative is needed. However, it requires working with a system of 9 equations instead of 3 in 3D when working with the second-order formulation. This is the reason why most if not all of the main elastodynamic codes are developed in the second-order formulation.

3.3. Acceleration procedure: initial guess computation

Except from applying preconditioners to the CG method, a suitable estimate of the solution as an initial guess for the CG method accelerates the convergence. Considering the long-time asymptotic behavior of (6) with an arbitrary initial value, running the time solver over a sufficiently long time interval $[0, mT]$, $m \geq 1$, should yield a “good” initial estimate. However, depending on the geometry of the computational domain and the boundary conditions, the time interval can be long, especially in the case of trapped modes [26, 25]. To overcome this difficulty, multiplying the external force term by a smooth transient function θ_{tr} ,

$$\theta_{tr}(t) = \begin{cases} \left(2 - \sin\left(\frac{\pi}{2} \frac{t}{T_{tr}}\right)\right) \sin\left(\frac{\pi}{2} \frac{t}{T_{tr}}\right), & t \leq T_{tr}, \\ 1, & t \geq T_{tr}, \end{cases} \quad (49)$$

as proposed by Mur in [40], and Glowinski et al. in [7], a proper initial estimate is computed for a more reasonable time interval $[0, T_{tr}]$, $T_{tr} \in T\mathbb{N}$ [25]; e.g.

$$T_{tr} = \left\lceil \frac{L}{T \min_{x \in \Omega} v_S(x)} \right\rceil, \quad (50)$$

where L is the longest dimensional length of Ω . This is the strategy we adopt in this study.

4. Implementation

We now discuss our implementation choices for the controllability method described in the two previous Sections.

4.1. Spatial Discretization: finite spectral element technique

We focus on a spatial discretization based on a finite spectral element technique. This method was first introduced by Patera [44] for fluid dynamics problem. The goal is to benefit from the flexibility of finite-element techniques while preserving the spectral convergence property of spectral methods. This method has been introduced to seismologists by Komatitsch & Vilotte [31], first in the context of elastic waveform modeling at the global scale. In the context of wave propagation, the method has shown excellent properties in terms of convergence, numerical dispersion, and scalability through domain decomposition.

The spectral element technique is based on the Galerkin principle (same approximation space for test functions and solution). The basis for the approximation space is composed of Lagrange polynomials of arbitrary degree n . The integration points and collocation points are chosen as the Gauss-Lobatto-Legendre (GLL) points, which ensure exact integration for polynomial to orders up to $2n-1$, continuity of the solution across elements, and spectral convergence properties. The fact that the collocation and integration points are chosen the same, together with the choice of Lagrange polynomial, ensure a diagonal mass matrix.

We apply this technique to discretize equations (6) and (42). Let $\Omega_h = \{x_j\}_{j=1}^{n_\Omega} \subset \Omega$, $\mathcal{V}_h \subset \mathcal{V}$, with $n_\Omega := \dim(\mathcal{V}_h) = \dim(\mathcal{W}_h) = |\Omega_h| < \infty$. For simplicity, we again write $\mathbf{y}(t)$, $\mathbf{f}(t)$, $\mathbf{g}_S(t)$, $\mathbf{g}_N(t)$, and $\mathbf{g}_D(t)$ instead of $\mathbf{y}(x, t)$, $\mathbf{f}(x, t)$, $\mathbf{g}_S(x, t)$, $\mathbf{g}_N(x, t)$, and $\mathbf{g}_D(x, t)$, respectively. Then, for fixed $t > 0$ we consider the variational formulation,

$$\begin{aligned} & \frac{\partial^2}{\partial t^2} \int_{\Omega} \rho \mathbf{y}(t) \cdot \boldsymbol{\varphi} \, dx + \int_{\Omega} (\mathbf{C} : \varepsilon(\mathbf{y}(t))) : \varepsilon(\boldsymbol{\varphi}) \, dx + \frac{\partial}{\partial t} \int_{\Gamma_S} \rho (\mathbf{B} \mathbf{y}(t)) \cdot \boldsymbol{\varphi} \, ds \\ &= \int_{\Omega} \mathbf{f}(t) \cdot \boldsymbol{\varphi} \, dx + \int_{\Gamma_S} \mathbf{g}_S(t) \cdot \boldsymbol{\varphi} \, ds + \int_{\Gamma_N} \mathbf{g}_N(t) \cdot \boldsymbol{\varphi} \, ds, \quad \boldsymbol{\varphi} \in \mathcal{V}_h. \end{aligned}$$

Then the stiffness matrix \mathcal{K} , the mass matrix \mathcal{M} , and the (absorbing) boundary matrix \mathcal{S} are given by

$$\mathcal{K} = \begin{pmatrix} \mathbf{K}_{11} & \cdots & \mathbf{K}_{1d} \\ \vdots & \ddots & \vdots \\ \mathbf{K}_{d1} & \cdots & \mathbf{K}_{dd} \end{pmatrix}, \quad \mathcal{M} = \begin{pmatrix} \mathbf{M} & & \\ & \ddots & \\ & & \mathbf{M} \end{pmatrix}, \quad \mathcal{S} = \begin{pmatrix} \mathbf{S}_{11} & \cdots & \mathbf{S}_{1d} \\ \vdots & \ddots & \vdots \\ \mathbf{S}_{d1} & \cdots & \mathbf{S}_{dd} \end{pmatrix},$$

where $\mathcal{K}, \mathcal{M}, \mathcal{S} \in \mathbb{R}^{(dn_\Omega) \times (dn_\Omega)}$ and

$$\begin{aligned} (\mathbf{K}_{kp})_{ij} &= \begin{cases} \int_{\Omega} \left((\lambda + \mu) \frac{\partial \varphi_i}{\partial x_k} \frac{\partial \varphi_j}{\partial x_k} + \mu \nabla \varphi_i \cdot \nabla \varphi_j \right) dx, & k = p, \\ \int_{\Omega} \left(\lambda \frac{\partial \varphi_i}{\partial x_k} \frac{\partial \varphi_j}{\partial x_p} + \mu \frac{\partial \varphi_i}{\partial x_p} \frac{\partial \varphi_j}{\partial x_k} \right) dx, & k \neq p, \end{cases} \\ (\mathbf{M})_{ij} &= \int_{\Omega} \rho \varphi_i \varphi_j \, dx, \\ (\mathbf{S}_{kp})_{ij} &= \int_{\Gamma_S} \rho B_{kp} \varphi_i \varphi_j \, ds = \int_{\Gamma_S} \rho ((v_P - v_S) n_k n_p + v_S \delta_{kp}) \varphi_i \varphi_j \, ds, \end{aligned}$$

and δ_{kp} denotes the Kronecker-delta function. In addition, we have the discrete right-hand side \mathbf{F} given by

$$(\mathbf{F})_i = \int_{\Omega} \mathbf{f} \varphi_i \, dx + \int_{\Gamma_N} \mathbf{g}_N \varphi_i \, ds + \int_{\Gamma_S} \mathbf{g}_S \varphi_i \, ds.$$

This yields a semi-discrete, linear, time-dependent (second-order) system of ordinary differential equations,

$$\frac{\partial^2}{\partial t^2} \mathcal{M} \vec{\mathbf{y}}(t) + \mathcal{K} \vec{\mathbf{y}}(t) + \frac{\partial}{\partial t} \mathcal{S} \vec{\mathbf{y}}(t) = \text{Re}\{\mathbf{F} e^{-i\omega t}\}, \quad (51)$$

where $\vec{\mathbf{y}}(t_m) = ((\vec{\mathbf{y}})_j(t_m))_{j=1}^{n_\Omega}$, $(\vec{\mathbf{y}})_j(t_m) \approx \mathbf{y}(x_j, t_m)$, denotes the semi-discrete solution of (51).

Remark 4. *Of course, other discretization methods such as finite differences (FD), finite elements, e.g., with mass lumps [11, 13], which avoid inversion of mass matrices, or discontinuous Galerkin (DG) with block diagonal mass matrices, also work. In particular, the latter method becomes more attractive for the elastic wave equation in the first-order formulation, since it overcomes the solution of the elliptic problems at each CG iteration [25, 10].*

4.2. Time Discretization

For most applications, the time discretization used together with the spectral element discretization in space is a second-order prediction-correction scheme named as the Newmark scheme [30, 18, 58]. We however have experienced that a higher order time-discretization scheme is desirable **for CM to improve** their convergence, especially when high order discretization in space is to be considered. In this study, we thus deploy a fourth-order explicit Runge Kutta scheme.

Let us consider the fully discrete elastic wave equation (6) with the discrete initial value $(\vec{\mathbf{y}}_0, \vec{\mathbf{y}}_1) \in \mathbb{R}^{2n_\Omega}$,

$$(\vec{\mathbf{y}}_0)_j = \mathbf{y}_0(x_j), \quad (\vec{\mathbf{y}}_1)_j = \mathbf{y}_1(x_j), \quad j = 1, \dots, n_\Omega,$$

in the finite number of time steps n_T , where $\Delta t = \frac{T}{n_T}$ and T is single time period. In order to apply classical higher-order one-step methods to (51), we write the second-order time integration of (51) as a first-order system of differential equations,

$$\frac{\partial}{\partial t} \vec{\mathbf{Y}}(t) = \Phi(t, \vec{\mathbf{Y}}(t)) := \mathcal{D} \vec{\mathbf{Y}}(t) + \text{Re}\{\mathcal{R} e^{-i\omega t}\}, \quad (52)$$

where

$$\vec{\mathbf{Y}}(t) = \begin{pmatrix} \mathbf{y}(t) \\ \frac{\partial}{\partial t} \mathbf{y}(t) \end{pmatrix}, \quad \mathcal{D} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathcal{M}^{-1} \mathcal{K} & -\mathcal{M}^{-1} \mathcal{S} \end{pmatrix}, \quad \mathcal{R} = \begin{pmatrix} \mathbf{0} \\ \mathcal{M}^{-1} \mathbf{F} \end{pmatrix}. \quad (53)$$

Then we apply the explicit Runge-Kutta (RK) to the first-order (52) formulation, which has been shown efficient. We denote an approximation $\vec{\mathbf{Y}}^m$ of $\vec{\mathbf{Y}}(t_m)$ and $\mathcal{R}^m = \text{Re}\{\mathcal{R} e^{-i\omega t_m}\}$. Then we consider

$$\begin{aligned} \vec{\mathbf{Y}}^{m+1} &= \vec{\mathbf{Y}}^m + \Delta t \sum_{p=1}^s b_p \mathbf{k}_p^m, \\ \mathbf{k}_p^m &= \Phi\left(t_m + c_p \Delta t, \vec{\mathbf{Y}}^m + \Delta t \sum_{q=1}^{p-1} a_{pq} \mathbf{k}_q^m\right), \end{aligned}$$

where $\mathbf{a} = (a_{pq})_{p,q=1}^s$ is the RK matrix and $\mathbf{b} = (b_p)_{p=1}^s$, $\mathbf{c} = (c_p)_{p=1}^s$ are the weights and nodes of the RK method.

Here we explicitly list the classical fourth-order explicit Runge-Kutta (RK4) method in Algorithm 1, for $m \geq 0$, and $\vec{\mathbf{Y}}^0 = (\vec{\mathbf{y}}_0 \quad \vec{\mathbf{y}}_1)^\top$, where $\vec{\mathbf{y}}_0$ and $\vec{\mathbf{y}}_1$ are the

Algorithm 1: Classical explicit fourth-order Runge-Kutta method

$$\begin{aligned}
\mathbf{k}_1^m &= \Phi(t_m, \vec{\mathbf{Y}}^m); \\
\mathbf{k}_2^m &= \Phi(t_m + \frac{\Delta t}{2}, \vec{\mathbf{Y}}^m + \frac{\Delta t}{2} \mathbf{k}_1^m); \\
\mathbf{k}_3^m &= \Phi(t_m + \frac{\Delta t}{2}, \vec{\mathbf{Y}}^m + \frac{\Delta t}{2} \mathbf{k}_2^m); \\
\mathbf{k}_4^m &= \Phi(t_m + \Delta t, \vec{\mathbf{Y}}^m + \Delta t \mathbf{k}_3^m); \\
\vec{\mathbf{Y}}^{m+1} &= \vec{\mathbf{Y}}^m + \frac{\Delta t}{6} (\mathbf{k}_1^m + 2\mathbf{k}_2^m + 2\mathbf{k}_3^m + \mathbf{k}_4^m);
\end{aligned}$$

discrete initial displacement and velocity of the elastic wavefield \mathbf{y} . Thanks to the SE property, the evaluation of (53) requires only matrix matrix and matrix-vector multiplications and is thus fully explicit. The explicit RK4 is known to be efficient and robust, but not absolutely stable, and the time step Δt is bounded by the CFL constant and the smallest mesh size.

In other studies [7, 23, 46, 25, 10], it has been shown that methods of other order, such as Leap-Frog, explicit Euler, explicit Runge-Kutta second and third order time schemes for time integration (51) work as well as RK4, especially the higher order time scheme. However, lower order time schemes may not converge exactly to a time harmonic function. For better convergence, further treatment or slight modification of the time method is required, e.g., the method yields the exact time harmonic.

The analytical solution of the ODE, $y'(t) = f(t) := \text{Re} \{a \exp(-i\omega t)\}$, $a \in \mathbb{C}$, is given by $y(t) = \text{Re} \{ \frac{ia}{\omega} \exp(-i\omega t) \}$. The modification, e.g., of the explicit Euler method forced to converge with the time harmonic solution is

$$y_{n+1} = \alpha y_n + \beta \delta t f(t_n, y_n), \quad \alpha = \cos(\omega \delta t), \quad \beta = \frac{\sin(\omega \Delta t)}{\omega \Delta t},$$

with $y_0 = y(t_0)$. Then one can verify that

$$y_n = y(t_n), \quad \forall n \geq 0.$$

4.3. Filtering procedure

For general boundary conditions the main analysis in Theorem 1 shows that time-periodic solutions may not be unique. To recover the uniqueness shown in Propositions 1 and 2, we apply the filtering procedure (32) from Section 2.2.2,

$$\hat{\mathbf{y}}(x) = \frac{2}{T} \int_0^T \mathbf{y}(x, t) e^{i\omega t} dt \approx \frac{2}{T} \sum_{m=1}^{n_T} \int_{(m-1)\Delta t}^{m\Delta t} \mathbf{y}(x, t) e^{i\omega t} dt, \quad (54)$$

to the time-periodic solution $\mathbf{y}(t)$.

For the known wavefields $\vec{\mathbf{Y}}^{m-1} = (\mathbf{y}^{m-1} \quad \dot{\mathbf{y}}^{m-1})$ and $\vec{\mathbf{Y}}^m = (\mathbf{y}^m \quad \dot{\mathbf{y}}^m)$ at the m -th time step, $m \geq 1$, we first determine the Hermite interpolation of \mathbf{y} in $[(m-1)\Delta t, m\Delta t]$. Then we compute the partial integration using a higher order quadrature formula such as Gaussian quadrature over $[(m-1)\Delta t, m\Delta t]$ by replacing $\mathbf{y}(t)$ by its Hermite polynomial. Since only $\vec{\mathbf{Y}}^{m-1}$ and $\vec{\mathbf{Y}}^m$ are needed for each integration in the subinterval, the time integration can be computed efficiently on the fly and there is no need to store the entire history of $\vec{\mathbf{Y}}^m$.

4.4. Final CMCG Algorithm

We list the CMCG method in Algorithm 2 starting with the initial guess $\mathbf{Y}^{(0)}$ to find an initial value $\mathbf{Y} = (\mathbf{y}_0, \mathbf{y}_1)$ such that the corresponding solution \mathbf{y} of the time-dependent elastic wave equation (6) is time-periodic. Then we apply the filtering procedure (32) in Section 2.2.2 and 4.3 when the time-periodic solution is not unique. Consequently, the time-harmonic solution of (2) is given by the time-periodic solution (8).

Remark 5.

- i. In each CG iteration, we compute the gradient of the cost function \mathcal{J} , which requires the solution of the time-dependent forward and backward wave equation, but only over a single short time period $T = \frac{2\pi}{\omega}$. In practice, the time step Δt typically decreases antiproportionally to the frequency ω , so that the number of time steps n_T or the ratio between T and Δt is independent of ω and remains constant when the frequency is increased.
- ii. In order to compute the gradient in Step 2 and 5, we apply an inner CG loop to solve the (46). To accelerate the convergence, we additionally apply a Jacobi preconditioner to the CG method.
- iii. The outer CG loops stops when the relative CG residual in Step 11 achieves the outer CG tolerance tol . A typical choice is $\text{tol}=10^{-3}$ or 10^{-6} .
- iv. The inner product $\langle \cdot, \cdot \rangle_{\mathcal{B}}$ and its corresponding norm $\| \cdot \|_{\mathcal{B}}$ in Algorithm 2 depend on the choice of the coercive, bounded, and symmetric bilinear form $\mathcal{B} : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ used in (44).

4.5. Parallelization and integration in SEM46

Our implementation relies on the 3D (visco-)elastic full waveform modeling and inversion code SEM46, presented in Trinh et al. [58]. This package is developed in the perspective of seismic imaging by FWI. It makes use of a Cartesian-based mesh, and supports vertical deformations to conform to a spatially varying topography on top of the computational domain. Element size adaptation in the lateral directions is also implemented to adapt the discretization to the local wavelength given a frequency band and a local velocity. This choice of a Cartesian-based mesh is motivated by the will to avoid any dependence on a third-party mesher, and a greater flexibility and robustness: no need to adapt the mesh manually to remove too small or too deformed cells. This is motivated by the application to which the code is dedicated: FWI basically works with media where the only interface which is known *a priori* is the air/subsurface interface (free surface boundary condition).

SEM46 is developed in modern FORTRAN90 and has been ported on numerous high performance computing platforms, from local ones (CIMENT infrastructure, Univ. Grenoble Alpes), to national ones (IDRIS and TGCC, France) and European ones (Barcelona Supercomputing center, Spain). Its scalability has been tested satisfactorily up to several thousands of computational units. The parallelization relies on a simple message-passing-interface (MPI) based domain

Algorithm 2: CMCG algorithm

Result: Time-periodic solution \mathbf{Y}

- 1 Solve the forward and the backward wave equations (6) and (42) to determine the gradient of J , $\mathbf{g}^{(0)} = \mathcal{J}'(\mathbf{Y}^{(0)}) \in \mathcal{V}'$, defined by (43).
 - 2 Solve the coercive elliptic problem (44) with $\mathbf{g} = \mathbf{g}^{(0)}$ to determine the Riesz representative $\mathbf{p}^{(0)} \in \mathcal{V}$.
 - 3 Set $\mathbf{d}^{(0)} = -\mathbf{r}^{(0)} = \mathbf{p}^{(0)} \in \mathcal{V}$.
 - 4 **for** $\ell = 1, 2, \dots$, **do**
 - 5 Solve the homogeneous wave equations (6) ($\mathbf{f} = \mathbf{g}_D = \mathbf{g}_S = \mathbf{g}_N = 0$) with the initial values $\mathbf{d}^{(\ell)} = (\mathbf{d}_0^{(\ell)}, \mathbf{d}_1^{(\ell)})$ and (42).
 - 6 Compute the gradient $\mathbf{g}^{(\ell)} = \hat{\mathcal{J}}'(\mathbf{d}^{(\ell)})$ defined by (43). Solve the coercive elliptic problem (44) with $\mathbf{g} = \mathbf{g}^{(\ell)}$ to get $\mathbf{p}^{(\ell)}$.
 - 7 $\alpha_\ell = \frac{\|(\mathbf{r}_0^{(\ell)}, \mathbf{r}_1^{(\ell)})\|_{\mathcal{B}}^2}{\langle (\mathbf{p}_0^{(\ell)}, \mathbf{p}_1^{(\ell)}), (\mathbf{d}_0^{(\ell)}, \mathbf{d}_1^{(\ell)}) \rangle_{\mathcal{B}}} \in \mathbb{R}$
 - 8 $\mathbf{Y}^{(\ell+1)} = \mathbf{Y}^{(\ell)} + \alpha_\ell \mathbf{d}^{(\ell)} \in \mathcal{V}_D$
 - 9 $\mathbf{r}^{(\ell+1)} = \mathbf{r}^{(\ell)} - \alpha_\ell \mathbf{p}^{(\ell)} \in \mathcal{V}$
 - 10 $\beta_\ell = \frac{\|(\mathbf{r}_0^{(\ell+1)}, \mathbf{r}_1^{(\ell+1)})\|_{\mathcal{B}}^2}{\|(\mathbf{r}_0^{(\ell)}, \mathbf{r}_1^{(\ell)})\|_{\mathcal{B}}^2} \in \mathbb{R}$
 - 11 $\mathbf{d}^{(\ell+1)} = \mathbf{r}^{(\ell+1)} + \beta_\ell \mathbf{d}^{(\ell)} \in \mathcal{V}$
 - 12 Stop when the relative residual lies below the given tolerance tol ,
- $$\text{relres}^{(\ell)} = \frac{\|(\mathbf{r}_0^{(\ell+1)}, \mathbf{r}_1^{(\ell+1)})\|_{\mathcal{B}}}{\|(\mathbf{r}_0^{(0)}, \mathbf{r}_1^{(0)})\|_{\mathcal{B}}} \leq \text{tol}. \quad (55)$$
- 13 **end**
-

decomposition algorithm. Indeed, the spectral element method requires to exchange information only on the border of each domain, making it perfectly suitable for such an implementation.

This domain decomposition algorithm is used for mass and stiffness matrix-vector product. The mass matrix being diagonal, the computationally intensive part is related to the latter matrix-vector product. Of interest here, each part of the CM algorithm benefits from this domain decomposition algorithm. The outer CG loop requires the solution of two wave equations. The inner CG loop requires the solution of an elliptic problem involving stiffness matrix-vector products, which are already parallelized thanks to domain decomposition. Therefore, the whole CM algorithm is fully parallel, with no bottlenecks which might hamper the scalability of the algorithm.

5. Numerical experiments

In this Section we present a series of experiments to validate the CMCG approach and illustrate its scalability for solving 3D frequency-domain elastic wave problems (2). All the experiments are performed in the isotropic approximation where the stiffness tensor can be expressed with two parameters, however from

SEM46 it is possible to consider any kind of anisotropy up to fully triclinic media with 21 independent coefficient in the stiffness tensor [58].

We first investigate the convergence and the accuracy of the CMCG method at each (outer) CG iteration, applied to the plane-wave problems in a semi-open box. In order to speed up the convergence, we introduce an initial run-up phase by using the wave solver to find a proper initial estimate, which is already used in the acoustic case [26, 25]. Next, we consider a point source problem in a unbounded cuboid domain to compare the total computational time and memory requirement of the CMCG method with a direct solver based approach which solves (2) directly (with the same discretization). The direct solver is based on the LU multi-frontal decomposition method, implemented in the MUMPS library [1]. To illustrate the scalability of the CMCG method for larger problems, we measure the total computational time while we increase the number of cores. Then we repeat the previous experiment in a 3D heterogeneous elastic media, constructed from the 2D Marmousi II model [36]. Finally, in the perspective of FWI, we present a series of point source problems in several heterogeneous media corresponding to reconstructed media along FWI iterations. We show how we can use the solution from a given model to the subsequent model as an initial guess to speed-up the convergence of the CMCG method.

5.1. Convergence and accuracy tests on plane-wave solutions

The plane-wave \mathbf{u} in $\Omega = (0, 2000m) \times (0, 500m) \times (0, 500m)$ is given by

$$\mathbf{u}(\mathbf{x}) = \mathbf{A} \exp(i\mathbf{k} \cdot \mathbf{x}), \quad (56)$$

which solves (2) with the Dirichlet boundary conditions Γ_D everywhere except from the absorbing boundary conditions on $\Gamma_S = \{x = 2000m\}$. Here the frequency is $f = 10$ Hz, the velocities $v_p(x) = 5000\text{m/s}$, $v_s(x) = 2500\text{m/s}$, and the density $\rho(x) = 1 \text{ g/m}^3$ are all constant, and moreover, $\mathbf{A} \parallel \mathbf{k}$ and $|\mathbf{k}| = \frac{\omega}{v_P}$ with the angle of 5° in the xz -plane and $|\mathbf{A}| = 10^{-6}$. We apply the CMCG method \mathcal{P}^5 -SEM with $h = 50m$. Figure 2a monitors the relative CG residual (55) at CG iteration and stops at 162 CG iterations to the CG tolerance $= 10^{-6}$, which coincides with solving the time-dependent wave equation (6) over 324 time periods. Moreover, the cost functional \mathcal{J} in (39), evaluated at each CG iteration, decreases monotonically, where the relative numerical L^2 -error to the exact solution (56) decreases as well to $4.6 \cdot 10^{-5}$, which is shown in Figures 2b and 2c, respectively.

According to the limiting amplitude principle [39, 60], the solution $\mathbf{y}(\cdot, t)$ of (6) converges asymptotically to the time-harmonic solution given in (5) when $t = mT$, $m \in \mathbb{N}$, tends to infinity, which requires only the solution of (6) for a long time, where no optimization and elliptic solution is needed. To compare this "simple" wave solver with the CMCG method for finding a time-harmonic solution, Figure 3a shows the history of the relative L^2 error of both methods for the same number of periods. We can see that solving only the wave equation (6) may result in a slow convergence or even no convergence at all.

Next, we investigate the efficiency of the initial run-up described in Section 3.3 and try to find what a suitable run-up time would be. For $m = 80$ time

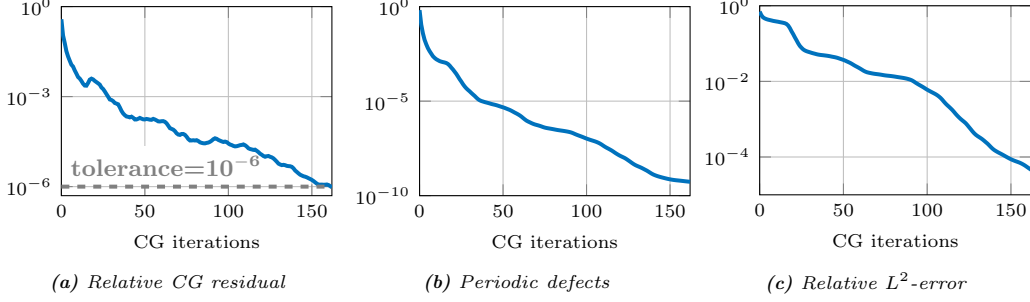


Figure 2: 3D plane-wave: History of the relative CG residual (a), the periodic defects (b), and the relative L^2 -error, obtained from the CMCG method at each CG iteration.

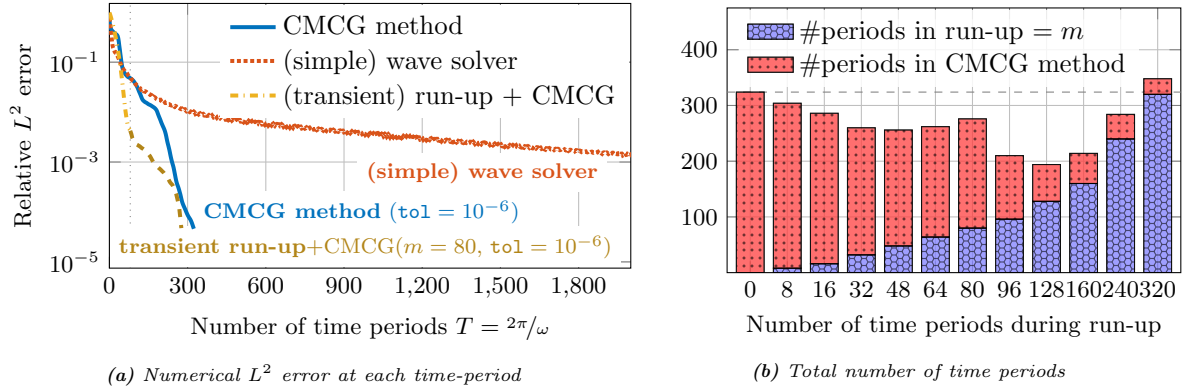


Figure 3: 3D plane-wave: (a) Numerical L^2 error at each time-period, obtained with the CMCG method, the “simple” wave solver, and the initial run-up of the length $m = 80$ combined with the CMCG method. (b) Total number of forward and backward elastic wave equations (6) solved over one period T until convergence, started with an initial run-up of length of m periods.

periods, the relative L^2 error decreases faster than with the CMCG method, but does not reach the same accuracy with a relative L^2 error of $4 \cdot 10^{-3}$. Maintaining the state with the CMCG method, it subsequently converges at 98 iterations. Consequently, the CMCG method solves for a total of 276 time periods with the initial estimate, and unlike 324, it saves about 15% time periods.

To study the convergence rate as a function of the run-up length m , we repeat the initial run-up with different lengths m of the initial phase in combination with the CMCG method. We first consider $m = 8$ as suggested in (50), which reduces the total number of wave equations solved in one period to 304. Increasing the length m of the initial run-up accelerates the convergence, but stops accelerating at $m \geq 160$.

We conclude that the initial run-up does indeed provide a good estimate for the CMCG method and accelerates convergence, but the optimal number of time periods is generally unknown and must be set in advance.

Remark 6. *The discretization of the sixth-order SE scheme and the fourth-order time scheme is not order-preserving. However, the total error between*

	#Cores	Elapsed Timing	Max alloc. Mem.	rel. L^2 error		
				u_x	u_y	u_z
MUMPS	256	855s	1423 GB	4.6%	8.7%	4.6%
CMCG	8	378s	0.8 GB	3.6%	7.3%	3.6%

Table 1: Comparison between the CMCG method and MUMPS direct solver with respect to the memory consumption, computational time, and numerical error.

f in Hz	n_{DOFs} in Mio.	# {Cores}	Memory in GB	Initial runup #{periods}	CMCG #{iterations}
10	2	1– 18	0.2	13	22
20	8	8– 64	0.8	25	17
40	42	32– 512	4	50	21
80	257	256–1536	24	100	23
160	1784	800–1440	166	200	27

Table 2: 3D-Point source: estimated allocated memory of the CMCG method by increasing the number of cores for a fixed frequency f , total number of CG iterations in the outer loop and number of degrees of freedoms.

numerical and analytical solution contains the spatial and the time discretization error as well as the residual of the CG method, where the spatial discretization dominates. In order to reduce the total error, i.e., by increasing the number of grid points, one can use order-preserving time schemes instead and reduce the CG tolerance.

5.2. Comparison with a direct solver for point source problem in a homogeneous medium

For the known analytical solution $\mathbf{u}(x)$,

$$\begin{aligned}
u_1(x) &= \frac{(x_1 - p_1)(x_3 - p_3)}{4\pi\rho\omega^2 r^5} ((r^2 k_p^2 - 3 + 3irk_p) e^{ik_p r} - (r^2 k_s^2 - 3 + 3irk_s) e^{ik_s r}) \\
u_2(x) &= \frac{(x_2 - p_2)(x_3 - p_3)}{4\pi\rho\omega^2 r^5} ((r^2 k_p^2 - 3 + 3irk_p) e^{ik_p r} - (r^2 k_s^2 - 3 + 3irk_s) e^{ik_s r}) \\
u_3(x) &= \frac{e^{ik_p r}}{4\pi\rho\omega^2 r^5} ((x_3 - p_3)^2 r^2 k_p^2 + (r^2 - 3(x_3 - p_3)^2)(1 - irk_p)) \\
&\quad + \frac{e^{ik_s r}}{4\pi\rho\omega^2 r^5} ((r^2 - (x_3 - p_3)^2) r^2 k_s^2 - (r^2 - 3(x_3 - p_3)^2)(1 - irk_s))
\end{aligned}$$

where $r = |x - p|$ and p is the position of the point source, we apply the CMCG method to the point source problem in a homogeneous unbounded domain $\Omega = (0, 2.5\text{km})^3$ with $\Gamma_S = \partial\Omega$, $\rho = 1\frac{\text{g}}{\text{m}^3}$, $v_P = 5000\frac{\text{m}}{\text{s}}$, and $v_S = 2500\frac{\text{m}}{\text{s}}$.

We first compare the CMCG method with the strategy presented in Li et al. [33] where the frequency-domain equations (2) are discretized by the same spectral element technique, within the frame of SEM46, and a LU based direct solver (MUMPS, [1]) is used to solve the associated linear system. The computational domain consists of $28 \times 28 \times 28$ \mathcal{P}^5 -SE with 4 sponge elements in each

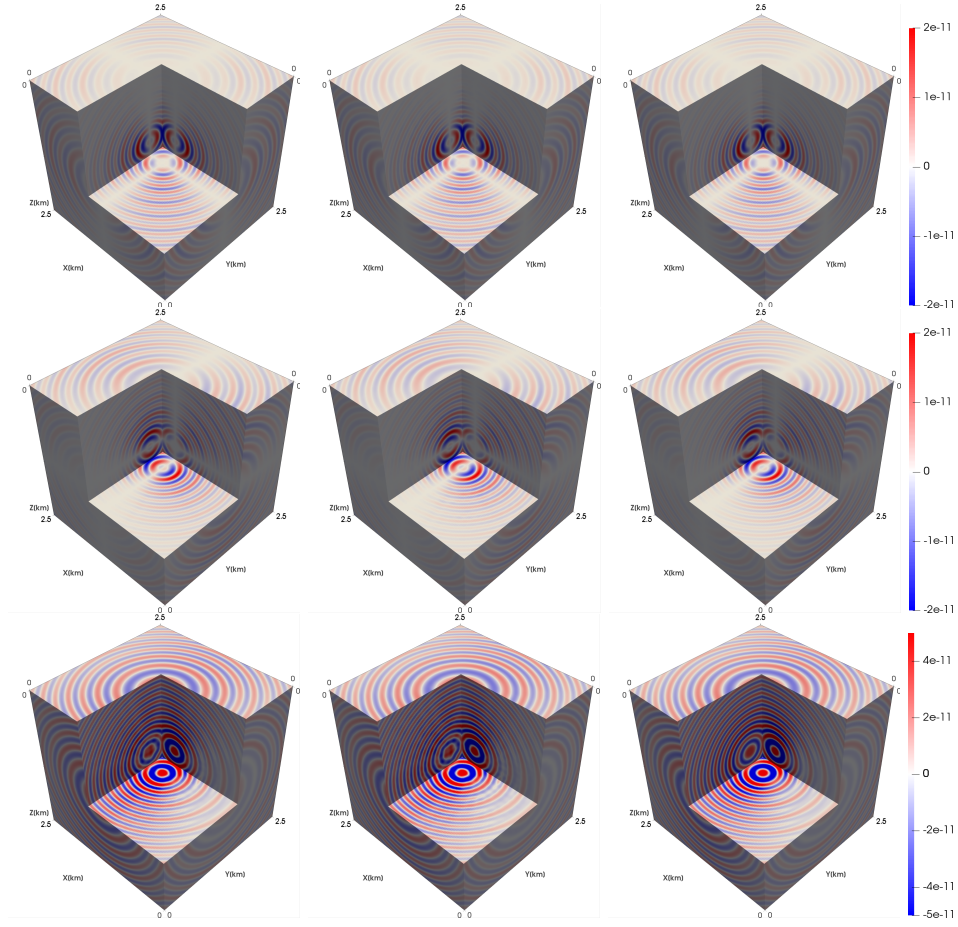


Figure 4: wavefields u_x (1st row), u_y (2nd row), and u_z (3rd row) of the point source problem of 20 Hz, obtained with the CMCG method (1st column) and the direct solver (2nd column) and the analytical solutions (3rd column).

direction.

In Table 1, we present the comparison between the numerical solutions, obtained with the CMCG method, combined with the initial run-up with a length of 25 periods, and the direct solver approach using the MUMPS package. The RK4 method listed in Algorithm 1 is applied to solve the time integration of (6) with a total number of time steps $n_T = 48$. The CG tolerance of the inner and outer CG loops for the stopping criterion is set to 10^{-4} . We see that the CMCG method requires less computation time and particularly about 1700 times less memory for a comparably accurate solution. Figure 4 illustrates the CMCG solution (1st column), the MUMPS solution (2nd column), and the difference between both methods (3rd column).

In [33], it is very challenging to adopt the direct linear solver and MUMPS for solving (2) with frequencies higher than 20 Hz. Especially, the memory requirement quickly reaches limits of conventional computational resources in

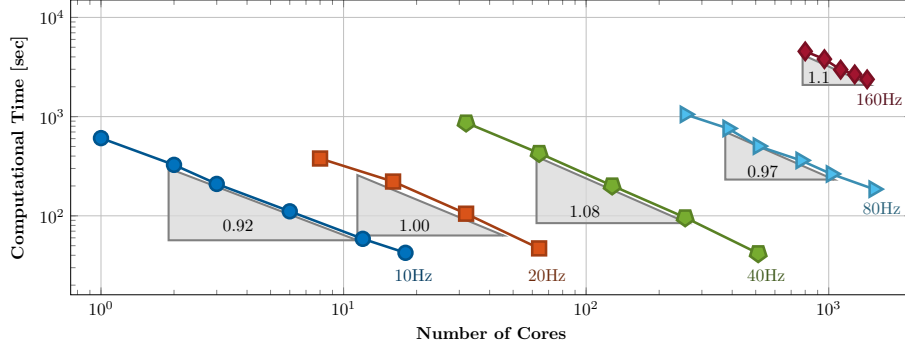


Figure 5: Point source at $f = 10, 20, 40, 80$, and 160 Hz: total elapsed timing spent in the CMCG method by increasing the total number of cores.

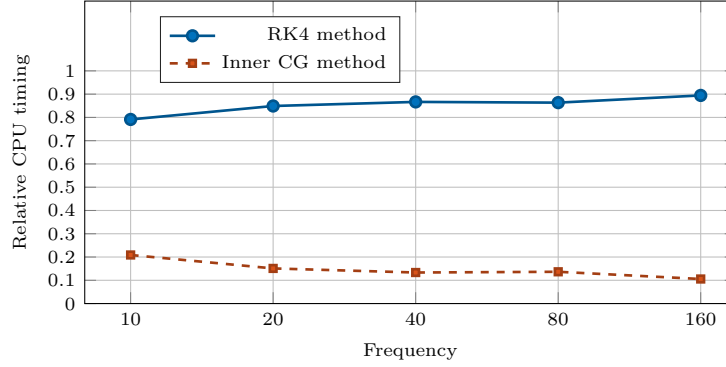


Figure 6: Pointsource at 10, 20, 40, 80 and 160 Hz model: relative average elapsed timing spent in the time-dependent solver (RK4 method) and the elliptic solver (inner CG method).

national HPC centers. In contrast, the CMCG method only requires modest amount of memory, which is shown in Table 2. To show the efficiency of the CMCG method by solving a larger problem, we increase the frequency from 20 Hz to 40 Hz, 80 Hz and 160 Hz, and respectively the mesh from $18 \times 18 \times 18$, $28 \times 28 \times 28$ to $48 \times 48 \times 48$, $88 \times 88 \times 88$, and $168 \times 168 \times 168$ \mathcal{P}^5 -SE, with 4 sponge elements in each direction, as well as the length of the initial phase from 13 to 25, 50, 100, and 200 periods.

While the number of grid points increases linearly with the frequency ω , the time periods $T = \frac{2\pi}{\omega}$ and the time steps Δt decrease in an antiproportional fashion with respect to ω , so that the number of time steps n_T for each ω remains equal to 48. As expected the CMCG method, which only consists of matrix-vector multiplications, is inherently parallel. A linear scalability is observed up to more than fifteen hundred cores, as illustrated in Figure 5. The ratio of computational effort between the time-dependent and elliptic solvers is 80% to 20% at low frequencies (e.g. 10 Hz). However, this changes at higher frequencies, as most of the computational effort shifts to the initial run-up to some extent (90% time-dependent and 10% elliptic at 160 Hz), which can be seen in Figure 6 and Table 2, where the total number of CG iterations increases only slightly as we increase the frequency.

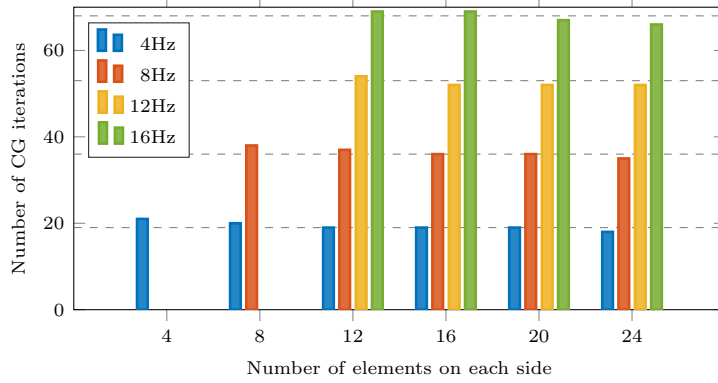


Figure 7: *Pointsource at 4, 8, 12, and 16Hz Model: number of CG iterations required in CM until convergence reaches tolerance 10^{-3} .*

f in Hz	n_{DOFs} in Mio.	$\#\{\text{Cores}\}$	Memory in GB	Initial runup $\#\{\text{periods}\}$	CMCG $\#\{\text{iterations}\}$
10	5	1– 8	0.5	30	25
20	26	16– 256	3	60	17
40	155	64– 512	16	120	10
80	1061	256–1024	103	240	21

Table 3: *3D-Marmousi: total computational time of the CMCG method with various number of cores, total number of CG iterations in the outer loop, as well as the number of degrees of freedoms.*

Further, it is assumed that the convergence of the CG method in CM depends only on the problem given by the geometry of the domain, the homogeneity or inhomogeneity of the media, the frequency, and the external source. In particular, the convergence is only slightly affected by the discretization of the mesh, e.g., the mesh refinement. To verify this, we solve the point source problem again for frequencies $\omega = 4, 8, 12, 16$ Hz using the CMCG method as above with \mathcal{P}^5 -SEM and RK4, but without sponge layer and initial excitation (run-up). At fixed frequency and source, the CMCG method is repeatedly applied to the elastic wave equation (6), increasing the number of elements from 4 to 24 elements in each direction. Figure 7 shows that the number of CG iterations (18–21, 35–38, 52–54, and 66–69) is nearly constant for a fixed setting (4, 8, 12, and 16 Hz)], confirming independence of mesh size. Interestingly, we observe that the number of CG iterations increases almost linearly in this low frequency settings (from 4 to 16 Hz).

5.3. CMCG method for a 3D heterogeneous elastic medium from the Marmousi II model

We consider now a 3D heterogeneous medium built as a 3D extension of a sub-target of the 2D Marmousi II model [36], where the 2D medium is extended constantly along the y direction. Now both velocities v_P and v_S and the density ρ vary over the entire computational domain $\Omega = (0, 3900m) \times (0, 1000m) \times (0, 1200m)$ (see Fig. 8).

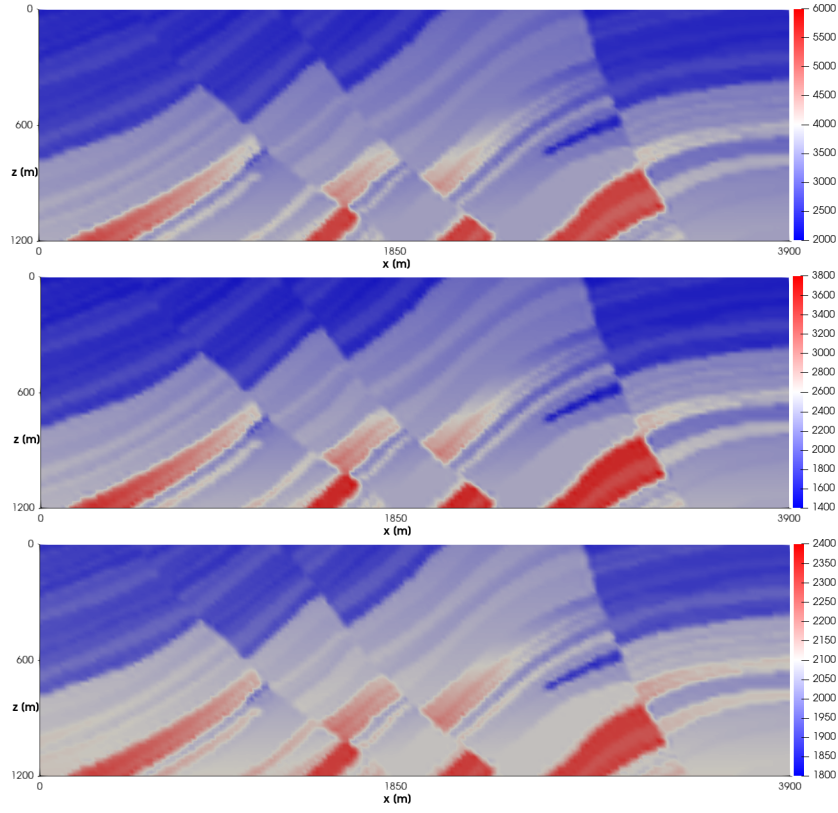


Figure 8: 3D-Marmousi model: 2D front view of the Marmousi model. Velocities v_P and v_S in $\frac{\text{m}}{\text{s}}$ in the 1st and 2nd rows and density ρ in $\frac{\text{mg}}{\text{m}^3}$ in the 3rd row.

We apply the CMCG method to compute a solution in this medium for a point source at $(1950\text{m}, 500\text{m}, 100\text{m})$, with a free surface condition on top $\Gamma_N = \{z = 0\text{m}\}$ and absorbing boundary conditions everywhere $\Gamma_S = \partial\Omega \setminus \Gamma_N$. To further dampen the outgoing waves on Γ_S , we also add 4 sponge elements in each direction. In order to show the parallel performances of the CMCG method, we increase the number of cores for a fixed frequency and fixed mesh, which is listed in Table 3, and measure the total elapsed time.

Here, the computational domain consists of $47 \times 18 \times 16$, $86 \times 28 \times 28$, $164 \times 48 \times 52$, and $320 \times 88 \times 100$ \mathcal{P}^5 -SE at frequencies of 10, 20, 40, and 80Hz, respectively, with memory requirements ranging from 500MB to 103GB. The time integration of (6) is solved in $n_T = 118$ time steps. In Figure 9 we observe a doubling of speed of the CMCG method for each frequency, indicating linear scalability, while increasing the number of cores to 1024. Finally, Figure 10 illustrates a cross section from the numerical solution of 80Hz to Marmousi problem, obtained with the CMCG method, where the background corresponds to the P-Wave v_P shown in Figure 8.

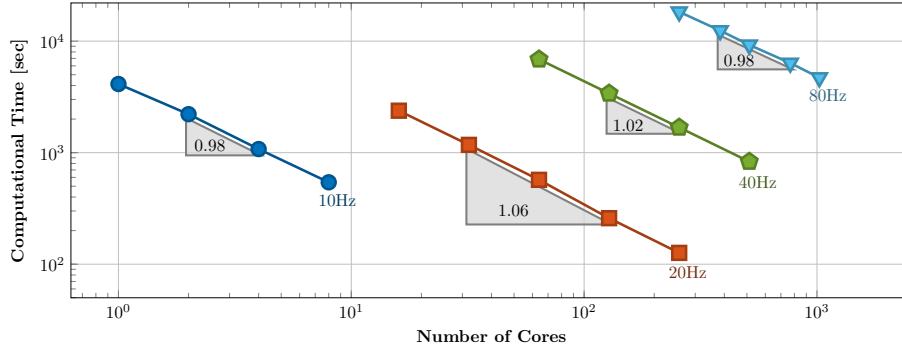


Figure 9: 3D-Marmousi: total computational time of the CMCG method by increasing the number of cores for a fixed frequency, 10Hz, 20Hz, 40Hz, and 80Hz.

5.4. CMCG method on successive 3D heterogeneous media corresponding to FWI iterations

To complete this series of tests, in the perspective of FWI, we investigate the behavior of the CMCG method when it is initialized with a solution computed in a medium close from the one in which the computation is done. The motivation is that in FWI, the subsurface model is iteratively updated from an initial to a final estimation. Wave propagation problems are to be solved in each of these models, which, along the FWI iterations, are relatively close one from each other.

To perform this preliminary study, we design a schematic synthetic FWI experiment on a subtarget of the Marmousi II model extended in 3D, similar to the one considered in the previous experiment. Its dimensions are $\Omega = (0, 3900) \times (0, 1000) \times (0, 1400)$ m³. The acquisition is composed of a single shot line in the middle of the x-y plane with 24 sources each 160m along x-axis, and 5795 receivers sampling the x-y plane, with a spacing of 12.5m in x-direction and 50 m in y-direction. The sources are vertical forces and the receivers vertical geophones. The source signal is a Ricker with 4 Hz peak frequency, generating a synthetic data set in the frequency band 0 - 10Hz.

The principle of FWI is to minimize iteratively the least-squares distance between a reference dataset and data computed numerically. Describing FWI precisely is beyond the scope of this study, and the reader is referred to [59] for a review. Here for this schematic application, the data is simply inverted in a single step (no multi-scale strategy), within 60 iterations of the *l*-BFGS algorithm implemented in the SEISCOPE toolbox [37]. The exact, initial, and final models are presented in Figure 11, denoted respectively by (v_P^*, v_S^*) , (v_P^0, v_S^0) , and (v_P^{60}, v_S^{60}) . The misfit function evolution along the 60 iterations is also provided in Figure 12.

Along the convergence path, a series of (v_P^ℓ, v_S^ℓ) is generated, $\ell = 1, \dots, 60$. We consider two situations. For the first, we consider two models in the early iterations, namely (v_P^1, v_S^1) and (v_P^2, v_S^2) where the model differences are significant, and localized in the upper part of the model, in the vicinity of the source (Figure 13a,b). For the second, we consider two models in the final iterations, namely (v_P^{58}, v_S^{58}) and (v_P^{59}, v_S^{59}) , for which the model differences are

less pronounced, and distributed in the whole model (Figure 13c,d). The relative L^2 difference between the models $(v_P^{(1)}, v_S^{(1)})$, $(v_P^{(58)}, v_S^{(58)})$ and $(v_P^{(2)}, v_S^{(2)})$, $(v_P^{(59)}, v_S^{(59)})$ are (0.012%, 0.043%) and (0.006%, 0.015%), respectively. We denote by $\mathbf{u}^{(1)}$ and $\mathbf{u}^{(58)}$ the solutions to (2) on the 1st and 58th models $(v_P^{(1)}, v_S^{(1)})$ and $(v_P^{(58)}, v_S^{(58)})$.

We apply the CMCG method, combined with the RK4 and \mathcal{P}^4 -SE method, to (2) for the 2nd and 59th model, where Ω_h consists of $47 \times 18 \times 18$ elements. Figures 14a and 14b show the history of the relative CG residual at each CG iteration, obtained with the CMCG method applied to (2) for the corresponding model. We observe that the CMCG method converges for the 2nd model with and without the initial estimate $\mathbf{u}^{(1)}$ as the initial guess at 175 and 347 CG iterations (which implies the solution of the time-dependent equation (6) over 350 and 694 time periods, respectively). The computational time with the initial guess thus represents 50% of the computational time without the initial guess, yielding a factor 2 acceleration. Applying the CMCG method to the 59th model with or without the $\mathbf{u}^{(58)}$ solution in the 58th model, we observe that the CMCG method converges in 46 and 336 CG iterations respectively (which implies solving the wave equation over 92 and 672 time periods respectively). Here, using the initial guess reduces the computational time to 13% of the computational cost without initial guess, corresponding to an acceleration by a factor larger than 7. We also observe that the CMCG method without initial estimation seems to converge consistently in about 350 CG iterations for different models built along the convergence path of a given full waveform inversion application.

From this preliminary test, we observe that the CMCG method is able to exploit efficiently the information from solutions computed in similar models to accelerate its convergence, a very interesting feature in the perspective of FWI, which is not shared by purely time-domain solvers.

6. Conclusion and perspectives

The controllability method (CM), combined with the conjugate gradient (CG) method, is proposed to find a time-periodic solution of the elastic wave equation, which yields the frequency-domain elastic solution. Although the time-periodic solution may contain additional numerical errors involved by the mass-lumping and numerical errors in the time integration and deficiency from inexact periodicity, the results is still comparably accurate to the solution obtained with a direct linear solver.

The main analysis proves that the filtering procedure extends the original CMCG method in [38] from elastic sound-soft scattering problems to more general elastic problems.

Numerical experiments illustrate the usefulness, efficiency, and good parallel scalability of the CMCG method. We observe that most of the computational efforts are spent either in the forward and backward wave propagation solver or in the inner CG loop, which are both inherently parallel through domain

decomposition.

An appropriate initial guess of the frequency-domain solution is also significant to speed up the convergence of the outer loop so that an initial run-up process as proposed in [7, 26] is considered. On models extracted from a full waveform inversion, it is shown how starting the CMCG method from a solution computed in the model from previous FWI iteration is beneficial to the convergence. To generalize this approach, a scattered field formulation could also be used [54, 43]. This approach makes it possible to take benefit from an already known solution in a medium close from the one investigated in a more general way. The known solution could be an analytical solution in a simple medium, a solution based on travel-times computed with an eikonal solver, or, as investigated here, a full waveform frequency-domain solution computed in a previous medium in the frame of full waveform inversion. This will be the matter of future investigations.

The extension of the CMCG approach to fluid-solid coupling problems, as has been already explored in the time-domain within the frame of the SEM46 package by Cao et al. [8], should be also the matter of future investigations.

Finally, we note that there is a current trend, at least in the exploration industry, to perform FWI on the whole usable frequency band, up to 100Hz and above. We believe that the design of efficient 3D modeling engine which, as it seems to be the case for the CMCG approach, have, on top of an excellent scalability and low memory requirement, the capability to exploit the information from previous iterations to build the wave equation solution in an efficient way, is of particular interest.

7. Acknowledgments

The authors would like to thank gratefully Jian Cao for his insightful comments and his help for generating the series of velocity models from full waveform inversion. This study was partially funded by the SEISCOPE consortium (<http://seiscope2.osug.fr>), sponsored by AKERBP, CGG, CHEVRON, EQUINOR, EXXON-MOBIL, JGI, SHELL, SINOPEC, SISPROBE and TOTAL, and by the HIWAI ANR project (ANR-16-CE31-0022-01). This study was granted access to the HPC resources of the Froggy platform of the CIMENT infrastructure (<https://ciment.ujf-grenoble.fr>), which is supported by the Rhône-Alpes region (GRANT CPER07_13 CIRA), the OSUG@2020 labex (reference ANR10 LABX56) and the Equip@Meso project (reference ANR-10-EQPX-29-01) of the programme Investissements d’Avenir supervised by the Agence Nationale pour la Recherche, and the HPC resources of CINES/IDRIS/TGCC under the allocation 046091 made by GENCI.

Appendix A. Operator B in the isotropic case for dimension $d > 1$

In the isotropic approximation, we write the stress tensor σ as

$$\sigma(x, \mathbf{u}) = \lambda(x)(\nabla \cdot \mathbf{u}(x))\mathbf{I}_d + 2\mu(x)\varepsilon(x, \mathbf{u}) \quad (\text{A.1})$$

where λ and μ are the Lamé parameters, or

$$\sigma(x, \mathbf{u}) = \rho(x)(v_P^2(x) - 2v_S^2(x))(\nabla \cdot \mathbf{u}(x))\mathbf{I}_d + 2\rho(x)v_S^2(x)\varepsilon(x, \mathbf{u}), \quad (\text{A.2})$$

where v_P is the pressure wave (P-wave) velocity and v_S is the shear wave (S-wave) velocity with

$$v_P = \sqrt{\frac{\lambda + 2\mu}{\rho}}, \quad v_S = \sqrt{\frac{\mu}{\rho}}. \quad (\text{A.3})$$

For physical media, we assume $v_P > v_S > 0$ and $\rho > 0$ such that $\lambda(x) \geq \lambda_0 > 0$ and $\mu(x) \geq \mu_0 > 0$.

To simulate the outgoing waves by imposing the absorbing boundary conditions on Γ_S , we introduce the (coercive) linear operator $\mathbf{B} \in \mathbb{R}^{d \times d}$ given by

$$\mathbf{B} = v_P \mathbf{nn}^\top + v_S(\mathbf{I}_d - \mathbf{nn}^\top) = (v_P - v_S)\mathbf{nn}^\top + v_S \mathbf{I}_d, \quad (\text{A.4})$$

where \mathbf{n} denotes the outward unit normal vector [52, 12]. Note that

$$\mathbf{B}\mathbf{w} = v_P \mathbf{w}_\perp + v_S \mathbf{w}_\parallel \quad (\text{A.5})$$

with the normal component of \mathbf{w} ,

$$\mathbf{w}_\perp = (\mathbf{nn}^\top)\mathbf{w} = (\mathbf{w} \cdot \mathbf{n})\mathbf{n}, \quad (\text{A.6})$$

and the tangential component of \mathbf{w} ,

$$\mathbf{w}_\parallel = \mathbf{w} - \mathbf{w}_\perp = (\mathbf{I}_d - \mathbf{nn}^\top)\mathbf{w} = \mathbf{w} - (\mathbf{w} \cdot \mathbf{n})\mathbf{n}. \quad (\text{A.7})$$

In order to attenuate the reflected waves at the artificial boundaries Γ_S , we additionally expand the computational domain surrounded by the sponge layer [9, 20].

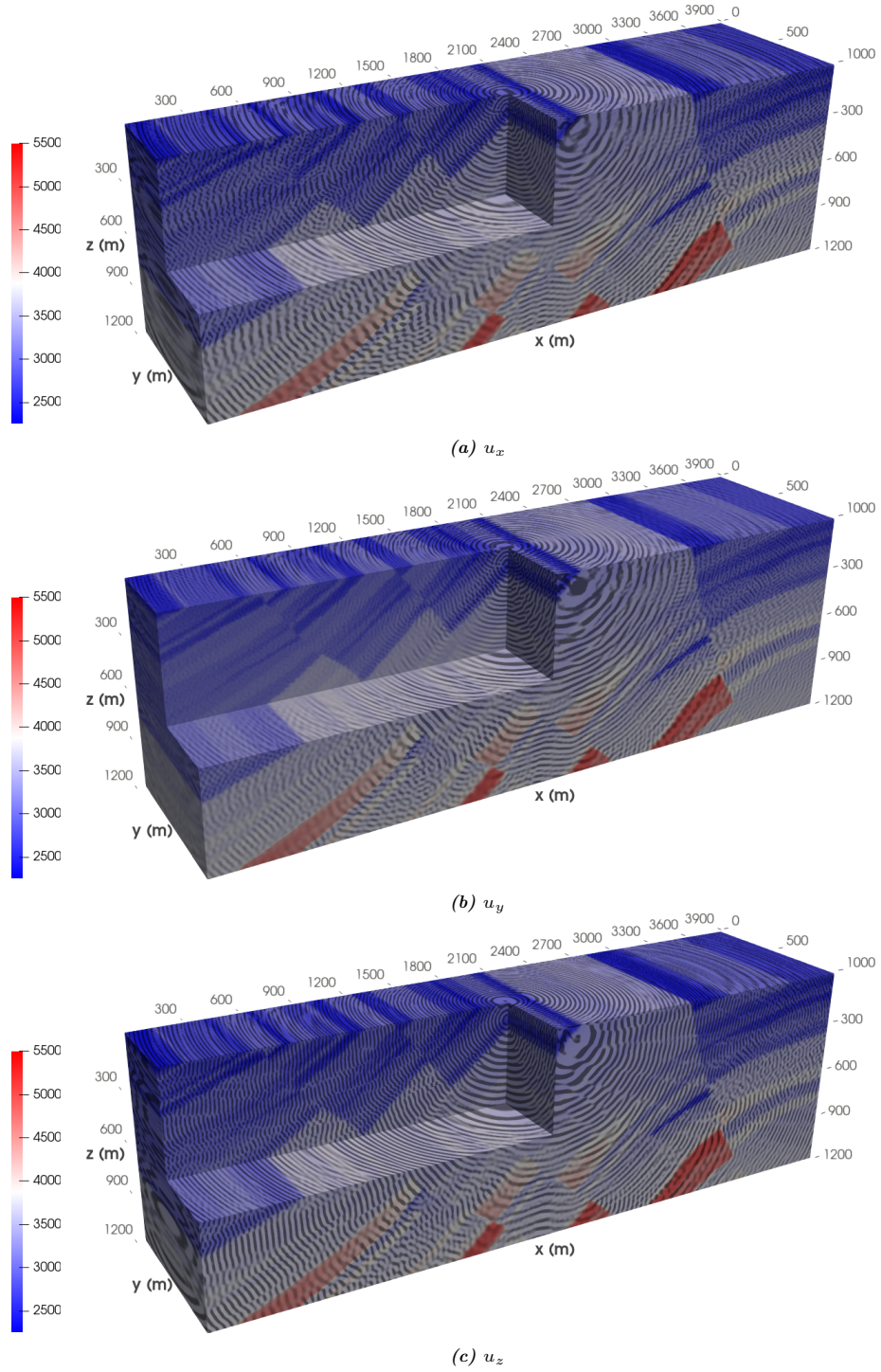


Figure 10: 3D-Marmousi solutions, u_x (1st row), u_y (2nd row), and u_z (3rd row) of the Marmousi problem of 80Hz, obtained with the CMCG method, where the background corresponds to the P-wave v_P shown in Figure 8.

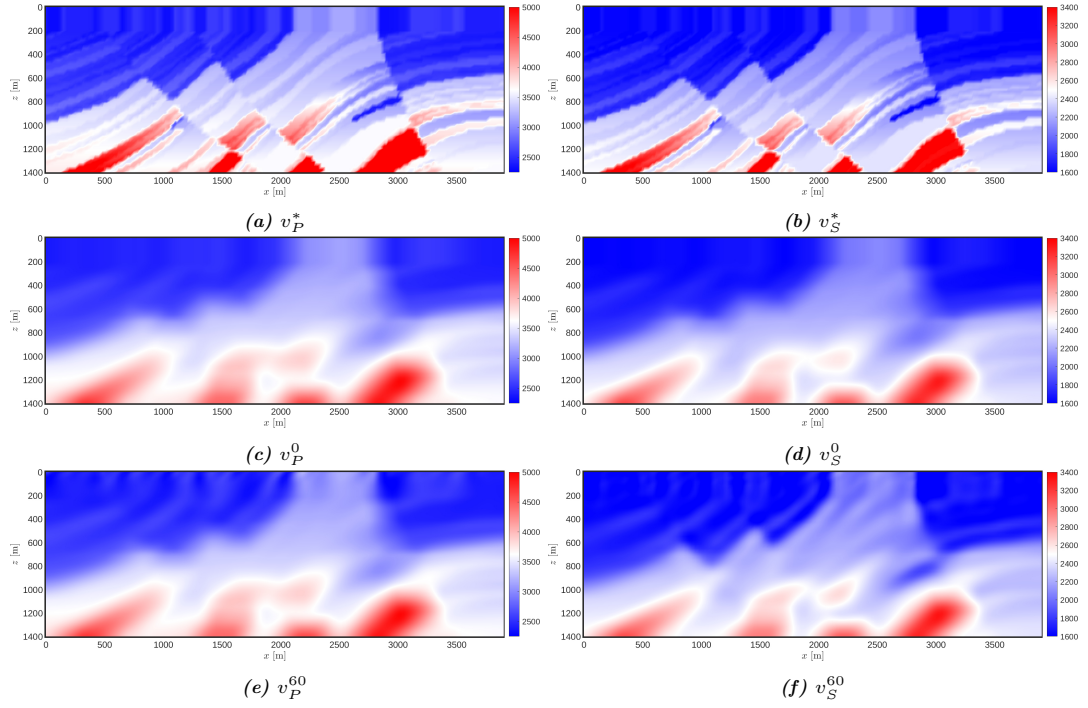


Figure 11: 3D-Marmousi model: 2D vertical slices of the exact Marmousi models v_P^* and v_S^* , initial models v_P^0 and v_S^0 , final models v_P^{60} and v_S^{60}

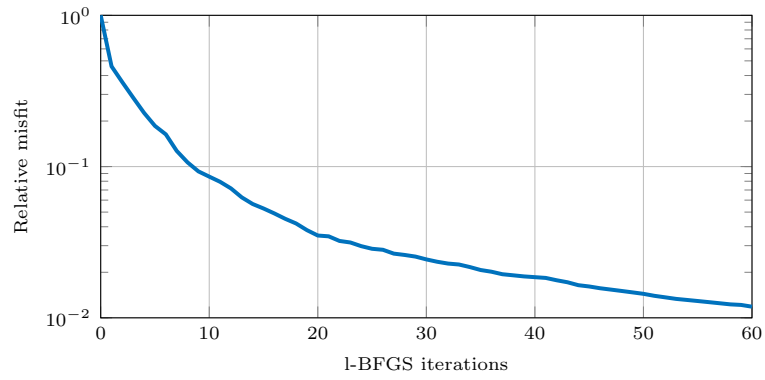


Figure 12: FWI convergence: reduction of the misfit function along the iterations

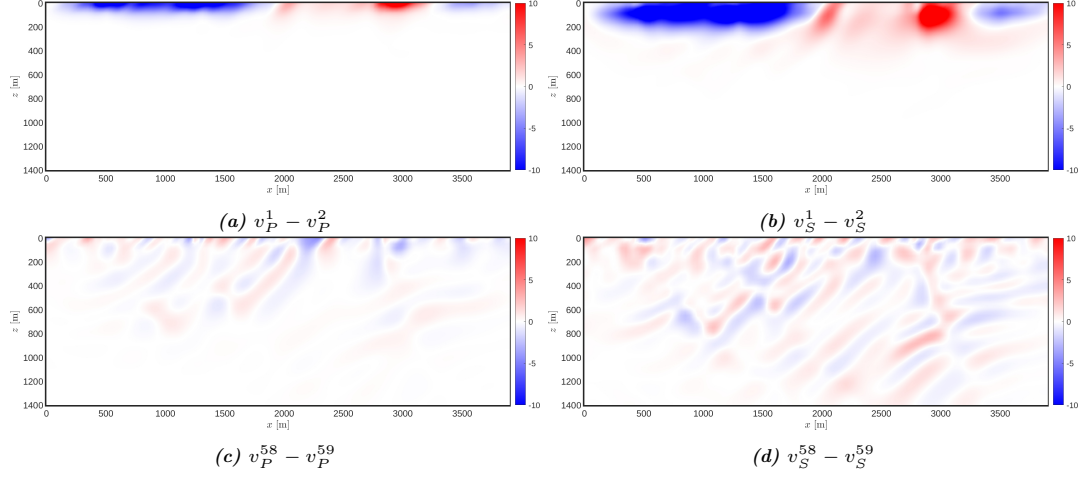


Figure 13: 3D-Marmousi model: 2D vertical slices of v_P and v_S model differences at iterations 2 and 1 (a,b) and 59 and 58 (c,d) respectively.

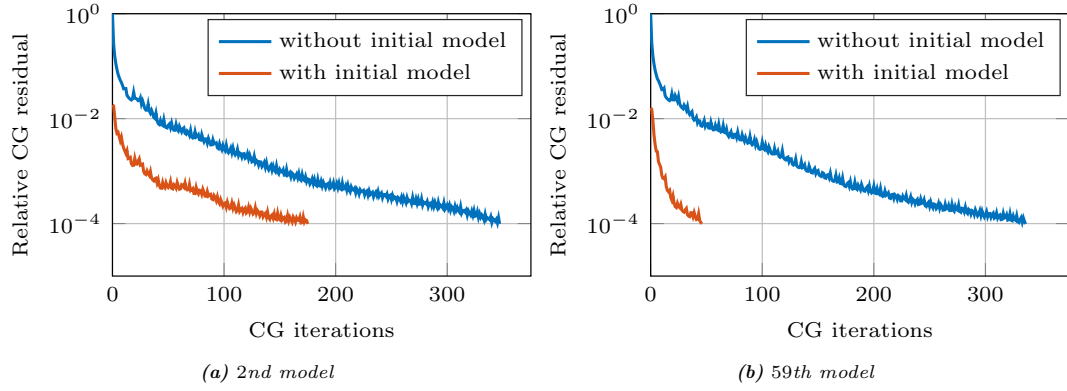


Figure 14: 3D-Marmousi model: history of the relative CG residual at each CG iteration, obtained with the CMCG method applied to (2) for the 2nd and 59th model.

References

- [1] Amestoy, P., Buttari, A., Duff, I., Guermouche, A., L’Excellent, J.-Y., and Uçar, B. (2011). *Mumps*, pages 1232–1238. Springer US, Boston, MA.
- [2] Amestoy, P., Duff, I. S., and L’Excellent, J. Y. (2000). Multifrontal parallel distributed symmetric and unsymmetric solvers. *Computer Methods in Applied Mechanics and Engineering*, 184(2-4):501–520.
- [3] Appelö, D., Garcia, F., and Runborg, O. (2020). WaveHoltz: Iterative solution of the Helmholtz equation via the wave equation. *SIAM J. Sci. Comput.*
- [4] Barucq, H., Boillot, L., Calandra, H., and Diaz, J. (2012). Absorbing boundary conditions for anisotropic elastodynamic media. In d’Acoustique, S. F., editor, *Acoustics 2012*, pages –, Nantes, France.
- [5] Belonosov, M., Kostin, V., Neklyudov, D., and Tcheverda, V. (2018). 3D numerical simulation of elastic waves with a frequency-domain iterative solver. *Geophysics*, 83(6):T333–T344.
- [6] Bramble, J. H. and Pasciak, J. E. (2008). A note on the existence and uniqueness of solutions of frequency domain elastic wave problems: A priori estimates in h^1 . *Journal of Mathematical Analysis and Applications*, 345(1):396 – 404.
- [7] Bristeau, M. O., Glowinski, R., and Périaux, J. (1998). Controllability Methods for the Computation of Time-Periodic Solutions; Application to Scattering. *Journal of Computational Physics*, 147(2):265–292.
- [8] Cao, J., Brossier, R., Górszczyk, A., Métivier, L., and Virieux, J. (2021). 3D multi-parameter full-waveform inversion for ocean-bottom seismic data using an efficient fluid-solid coupled spectral-element solver. *Geophysical Journal International*.
- [9] Cerjan, C., Kosloff, D., Kosloff, R., and Reshef, M. (1985). A nonreflecting boundary condition for discrete acoustic and elastic wave equations. *Geophysics*, 50(4):2117–2131.
- [10] Chaumont-Frelet, T., Grote, M. J., Lanteri, S., and Tang, J. H. (in press). A controllability method for maxwell’s equations. *SIAM Journal on Scientific Computing*.
- [11] Chin-Joe-Kong, M., Mulder, W., and van Veldhuizen, M. (1999). Higher-order triangular and tetrahedral finite elements with mass lumping for solving the wave equation. *Journal of Engineering Mathematics*, 35(4):405–426.
- [12] Clayton, R. and Engquist, B. (1977). Absorbing boundary conditions for acoustic and elastic wave equations. *Bulletin of the Seismological Society of America*, 67:1529–1540.
- [13] Cohen, G., Joly, P., Roberts, J. E., and Tordjman, N. (2001). Higher order triangular finite elements with mass lumping for the wave equation. *SIAM Journal on Numerical Analysis*, 38(6):2047–2078.

- [14] Cortés-Vega, L., Fernández, C., and Menzala, G. (2006). On outgoing solutions for a system of time-harmonic elastic wave in the exterior of a star-shaped domain. *Proyecciones*, 25:205–229.
- [15] Cortés-Vega, L. A. (2003). Resonant frequencies for a system of time-harmonic elastic wave. *Journal of Mathematical Analysis and Applications*, 279(1):43–55.
- [16] Cummings, P. and Feng, X. (1999). Domain decomposition methods for a system of coupled acoustic and elastic Helmholtz equations. In *ELEVENTH INTERNATIONAL CONFERENCE ON DOMAIN DECOMPOSITION METHODS*, pages 203–210. Domain Decomposition Press.
- [17] Ernst, O. G. and Gander, M. J. (2012). Why it is Difficult to Solve Helmholtz Problems with Classical Iterative Methods. In Graham, I. G., Hou, T. Y., Lakkis, O., and Scheichl, R., editors, *Numerical Analysis of Multiscale Problems*, volume 83 of *Lecture Notes in Computational Science and Engineering*, pages 325–363. Springer Berlin Heidelberg.
- [18] Festa, G. and Vilotte, J.-P. (2005). The newmark scheme as a velocity-stress time staggering: An efficient PML for spectral element simulations of elastodynamics. *Geophysical Journal International*, 161(3):789–812.
- [19] Fichtner, A., Kennett, B. L. N., Igel, H., and Bunge, H. P. (2009). Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods. *Geophysical Journal International*, 179(3):1703–1725.
- [20] Fletcher, W. R., George, J., Kubichek, R., Smithson, S., and Sochacki, J. (1987). Absorbing boundary conditions and surface waves. *GEOPHYSICS*, 52(1):60–71.
- [21] Gatto, P. and Hesthaven, J. S. (2017). Efficient preconditioning of hp-FEM matrices by hierarchical low-rank approximations. *JOURNAL OF SCIENTIFIC COMPUTING*, 72(1):49–80.
- [22] Glowinski, R., Kähkönen, S., Rossi, T., and Mäkinen, R. (2011). Solution of time-periodic wave equation using mixed finite elements and controllability techniques. *Journal of Computational Acoustics*, 19(4):335–352.
- [23] Glowinski, R. and Rossi, T. (2006). A mixed formulation and exact controllability approach for the computation of the periodic solutions of the scalar wave equation. (I): Controllability problem formulation and related iterative solution. *Comptes Rendus Mathématique*, 343(7):493–498.
- [24] Górszczyk, A., Brossier, R., and Métivier, L. (2021). Graph-space optimal transport concept for time-domain full-waveform inversion of ocean-bottom seismometer data: Nankai trough velocity structure reconstructed from a 1d model. *Journal of Geophysical Research: Solid Earth*, 126(5):e2020JB021504. e2020JB021504 2020JB021504.
- [25] Grote, M. J., Nataf, F., Tang, J. H., and Tournier, P.-H. (2020). Parallel controllability methods for the Helmholtz equation. *Computer Methods in Applied Mechanics and Engineering*, 362:112846.

- [26] Grote, M. J. and Tang, J. H. (2019). On controllability methods for the Helmholtz equation. *Journal of Computational and Applied Mathematics*, 358:306–326.
- [27] Haidar, A. (2008). *On the parallel scalability of hybrid linear solvers for large 3D problems*. PhD thesis, Institut National Polytechnique de Toulouse - CERFACS TH/PA/08/57.
- [28] Heikkola, E., Mönkölä, S., Pennanen, A., and Rossi, T. (2007). Controllability method for the Helmholtz equation with higher-order discretizations. *J. Comput. Phys.*, 225:1553–1576.
- [29] Hestenes, M. R. and Stiefel, E. (1952). Methods of conjugate gradient for solving linear systems. *Journal of Research of the National Bureau of Standards*, 49:409–436.
- [30] Komatitsch, D., Ritsema, J., and Tromp, J. (2002). The spectral-element method, Beowulf computing, and global seismology. *Science*, 298(5599):1737–1742.
- [31] Komatitsch, D. and Vilotte, J. P. (1998). The spectral element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures. *Bulletin of the Seismological Society of America*, 88:368–392.
- [32] Lei, W., Ruan, Y., Bozdağ, E., Peter, D., Lefebvre, M., Komatitsch, D., Tromp, J., Hill, J., Podhorszki, N., and Pugmire, D. (2020). Global adjoint tomography—model glad-m25. *Geophysical Journal International*, 223(1):1–21.
- [33] Li, Y., Brossier, R., and Métivier, L. (2020). 3d frequency-domain elastic wave modeling with the spectral element method using a massively parallel direct solver. *GEOPHYSICS*, 85(2):T71–T88.
- [34] Li, Y., Métivier, L., Brossier, R., Han, B., and Virieux, J. (2015). 2D and 3D frequency-domain elastic wave modeling in complex media with a parallel iterative solver. *Geophysics*, 80(3):T101–T118.
- [35] Málek, J. and Strakoš, Z. (2014). *Preconditioning and the Conjugate Gradient Method in the Context of Solving PDEs*. Society for Industrial and Applied Mathematics.
- [36] Martin, G. S., Wiley, R., and Marfurt, K. J. (2006). Marmousi2: An elastic upgrade for Marmousi. *The Leading Edge*, 25(2):156–166.
- [37] Métivier, L. and Brossier, R. (2016). The SEISCOPE optimization toolbox: A large-scale nonlinear optimization library based on reverse communication. *Geophysics*, 81(2):F11–F25.
- [38] Mönkölä, S. (2010). Time-harmonic solution for acousto-elastic interaction with controllability and spectral elements. *Journal of Computational and Applied Mathematics*, 234(6):1904–1911. Eighth International Conference on Mathematical and Numerical Aspects of Waves (Waves 2007).
- [39] Morawetz, C. (1962). The limiting amplitude principle. *Comm. Pure Appl. Math.*, XV:349–361.

- [40] Mur, G. (1993). The finite-element modeling of three-dimensional electromagnetic fields using edge and nodal elements. *IEEE Transactions on Antennas and Propagation*, 41(7):948–953.
- [41] Nihei, K. T. and Li, X. (2007). Frequency response modelling of seismic waves using finite difference time domain with phase sensitive detection (TD-PSD). *Geophysical Journal International*, 169:1069–1078.
- [42] Operto, S., Miniussi, A., Brossier, R., Combe, L., Métivier, L., Monteiller, V., Ribodetti, A., and Virieux, J. (2015). Efficient 3-D frequency-domain mono-parameter full-waveform inversion of ocean-bottom cable data: application to Valhall in the visco-acoustic vertical transverse isotropic approximation. *Geophysical Journal International*, 202(2):1362–1391.
- [43] Pageot, D., Operto, S., Vallée, M., Brossier, R., and Virieux, J. (2013). A parametric analysis of two-dimensional elastic full waveform inversion of teleseismic data for lithospheric imaging. *Geophysical Journal International*, 193(3):1479–1505.
- [44] Patera, A. T. (1984). A spectral element method for fluid dynamics: laminar flow in a channel expansion. *Journal of Computational Physics*, 54:468–488.
- [45] Plessix, R. E. (2006). A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, 167(2):495–503.
- [46] Råbinä, J., Mönkölä, S., and Rossi, T. (2015). Efficient time integration of maxwell’s equations with generalized finite differences. *SIAM Journal on Scientific Computing*, 37(6):B834–B854.
- [47] Raknes, E. B., Arntsen, B., and Weibull, W. (2015). Three-dimensional elastic full waveform inversion using seismic data from the sleipner area. *Geophysical Journal International*, 202(3):1877–1894.
- [48] Rizzuti, G. and Mulder, W. (2016). Multigrid-based shifted-laplacian preconditioning for the time-harmonic elastic wave equation. *Journal of Computational Physics*, 317:47–65.
- [49] Sirgue, L., Barkved, O. I., Dellinger, J., Etgen, J., Albertin, U., and Kommedal, J. H. (2010). Full waveform inversion: the next leap forward in imaging at Valhall. *First Break*, 28:65–70.
- [50] Sirgue, L., Etgen, J. T., and Albertin, U. (2008). 3D Frequency Domain Waveform Inversion using Time Domain Finite Difference Methods. In *Proceedings 70th EAGE, Conference and Exhibition, Roma, Italy*, page F022.
- [51] Sourbier, F., Haiddar, A., Giraud, L., Ben-Hadj-Ali, H., Operto, S., and Virieux, J. (2011). Three-dimensional parallel frequency-domain visco-acoustic wave modelling based on a hybrid direct/iterative solver. *Geophysical Prospecting*, 59(5):834–856.
- [52] Stacey, R. (1988). Improved transparent boundary formulations for the elastic-wave equation. *Bulletin of the Seismological Society of America*, 78(6):2089–2097.

- [53] Stopin, A., Plessix, R.-E., and Al Abri, S. (2014). Multiparameter waveform inversion of a large wide-azimuth low-frequency land data set in Oman. *Geophysics*, 79(3):WA69–WA77.
- [54] Taflove, A. and Hagness, S. C. (2005). *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Artech House, 3rd edition.
- [55] Tang, J. H. (2019). *Solving forward and inverse Helmholtz equations via controllability methods*. dissertation, University of Basel.
- [56] Tang, J. H., Brossier, R., and Métivier, L. (2021). *Solving frequency-domain elastic wave equations via parallel controllability methods*, pages 2470–2474.
- [57] Tape, C., Liu, Q., Maggi, A., and Tromp, J. (2010). Seismic tomography of the southern California crust based on spectral-element and adjoint methods. *Geophysical Journal International*, 180:433–462.
- [58] Trinh, P. T., Brossier, R., Métivier, L., Tavaré, L., and Virieux, J. (2019). Efficient 3D time-domain elastic and viscoelastic Full Waveform Inversion using a spectral-element method on flexible Cartesian-based mesh. *Geophysics*, 84(1):R75–R97.
- [59] Virieux, J., Asnaashari, A., Brossier, R., Métivier, L., Ribodetti, A., and Zhou, W. (2017). An introduction to Full Waveform Inversion. In Grechka, V. and Wapenaar, K., editors, *Encyclopedia of Exploration Geophysics*, pages R1–1–R1–40. Society of Exploration Geophysics.
- [60] Vullierme-Ledard, M. (1987). The limiting amplitude principle applied to the motion of floating bodies. *ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique*, 21(1):125–170.