Journal of Computational Physics 515 (2024) 113272

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Electromagnetic inverse wave scattering in anisotropic media via reduced order modeling

Liliana Borcea [a], Yiyang Liu [a], Jörn Zimmerling [b,*]

[a] Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA
[b] Uppsala Universitet, Department of Information Technology, Division of Scientific Computing, 75105 Uppsala, Sweden

ARTICLE INFO

MSC:
65M32
41A20

Keywords:
Inverse wave scattering
Imaging
Electromagnetic
Data driven
Reduced order modeling
Optimization

ABSTRACT

The inverse wave scattering problem seeks to estimate a heterogeneous, inaccessible medium, modeled by unknown variable coefficients in wave equations, from transient recordings of waves generated by probing signals. It is a widely studied inverse problem with important applications, that is usually formulated as a nonlinear least squares data fit optimization. For typical measurement setups and band-limited probing signals, the least squares objective function has spurious local minima far and near the true solution, so Newton-type optimization methods fail to obtain satisfactory answers. We introduce a different approach, for electromagnetic inverse wave scattering in lossless, anisotropic media. It is an extension of recently developed data driven reduced order modeling methods for the acoustic wave equation in isotropic media. Our reduced order model (ROM) is an algebraic, discrete time dynamical system derived from Maxwell's equations. It has four important properties: (1) It can be computed in a data driven way, without knowledge of the medium. (2) The data to ROM mapping is nonlinear and yet the ROM can be obtained in a non-iterative fashion, using numerical linear algebra methods. (3) The ROM has a special algebraic structure that captures the causal propagation of the wave field in the unknown medium. (4) It is an interpolation ROM i.e., it fits the data on a uniform time grid. We show how to obtain from the ROM an estimate of the wave field at inaccessible points inside the unknown medium. The use of this wave is twofold: First, it defines a computationally inexpensive imaging function designed to estimate the support of reflective structures in the medium, modeled by jump discontinuities of the matrix valued dielectric permittivity. Second, it gives an objective function for quantitative estimation of the dielectric permittivity, that has better behavior than the least squares data fitting objective function. The methodology introduced in this paper applies to Maxwell's equations in three dimensions. To avoid high computational costs, we limit the study to a cylindrical domain filled with an orthotropic medium, so the problem becomes two dimensional.

1. Introduction

The estimation of an inaccessible heterogeneous medium from transient recordings of waves generated by probing signals is important in radar imaging, remote sensing, medical diagnostics, nondestructive evaluation of materials and aging structures, exploration geophysics, underwater acoustics, and so on. Depending on the application, it is formulated as an inverse problem for the scalar

* Corresponding author.
E-mail addresses: borcea@umich.edu (L. Borcea), yiyangl@umich.edu (Y. Liu), jorn.zimmerling@it.uu.se (J. Zimmerling).

https://doi.org/10.1016/j.jcp.2024.113272
Received 6 March 2024; Received in revised form 6 July 2024; Accepted 8 July 2024

Available online 11 July 2024

(acoustic) wave equation, or for systems of equations that govern the propagation of vectorial (elastic or electromagnetic) waves. The medium is modeled by unknown variable coefficients of these equations, which are scalar valued in the isotropic case or matrix valued in the anisotropic case. The inverse wave scattering problem is to estimate these coefficients from the measurements.

We study inverse scattering with electromagnetic waves in lossless media, governed by Maxwell's equations. It is known that the dielectric permittivity and magnetic permeability of an isotropic medium occupying a bounded and simply connected domain are uniquely determined by the admittance map, which takes the tangential boundary trace of the electric field to the tangential boundary trace of the magnetic field [44,35]. The admittance map measured for time harmonic fields or transient (multi-frequency) fields is not sufficient for inverse scattering in anisotropic media. It determines the matrix valued coefficients only up to a diffeomorphic transformation that leaves the boundary unchanged. This behavior is typical of other inverse problems in anisotropic media [38] and has sparked interest in inversion algorithms that use prior information, such as the direction of anisotropy [31,1]. Interestingly, the shape of inclusions filled with anisotropic media is uniquely recoverable, as shown in [45,18] for measurements of the far field pattern of the magnetic field, with transverse electric polarization. In many applications, it is possible to get more information about the medium from data acquisitions that exploit different polarizations of the probing waves. The study [25] shows that, at least in perturbative regimes, the matrix valued dielectric permittivity of orthotropic media can be determined from measurements of the time harmonic electric field generated by two different incident polarizations.

We study inverse wave scattering with measurements gathered by an array of antennas that emit probing waves with two different polarizations and then measure the resulting electric wave field. Our approach applies in principle to Maxwell's equations in an arbitrary three dimensional domain. To reduce the computational cost, we work in a cylindrical domain $\Omega \times \mathbb{R}$, with bounded and simply connected cross-section $\Omega \subset \mathbb{R}^2$ that has Lipschitz continuous boundary $\partial\Omega$ [41, Definition 3.1].

Introduce the coordinate system $\vec{x} = (x, z)$, with $x \in \Omega$ and $z \in \mathbb{R}$ and assume that the medium does not vary in z. In most applications the medium is non-magnetic [21, Section 1.2], so we set the magnetic permeability equal to the scalar constant μ_o. Electromagnetic wave propagation in two dimensional isotropic media reduces to the acoustic wave equation. Thus, we consider an anisotropic, orthotropic medium [46,45,18], also known as a medium with monoclinic symmetry [52, Section 13.2.2], modeled by the piecewise smooth dielectric permittivity matrix

$$\epsilon(\vec{x}) = \begin{pmatrix} \underline{\underline{\varepsilon}}(x) & \mathbf{0} \\ \mathbf{0}^T & \varepsilon_z(x) \end{pmatrix}, \quad \underline{\underline{\varepsilon}}(x) = \begin{pmatrix} \varepsilon_{11}(x) & \varepsilon_{12}(x) \\ \varepsilon_{12}(x) & \varepsilon_{22}(x) \end{pmatrix}. \tag{1}$$

Here $\underline{\underline{\varepsilon}} \in \mathbb{R}^{2\times 2}$ is symmetric and positive definite, as it should be in passive and lossless media [30, Section 1.2.2], and ε_z is positive. We use consistently two underlines for 2×2 matrix valued fields.

The orthotropic assumption (1) implies that the z axis of our system of coordinates is aligned with an axis of anisotropy of the medium. Like in [46,45,18], we use it as a simplification that reduces the problem to a two dimensional one, which we can handle numerically.

We model the boundary $\partial\Omega \times \mathbb{R}$ of the cylindrical domain, with outer normal

$$\vec{n}(\vec{x}) = \begin{pmatrix} n(x) \\ 0 \end{pmatrix}, \quad \vec{x} = (x, z) \in \partial\Omega \times \mathbb{R}, \tag{2}$$

as perfectly conducting. This boundary may be physical, in which case we have a waveguide with open ends, or fictitious and justified by hyperbolicity of the problem: The waves propagate at finite speed, so they will not sense over the duration of the measurements a boundary that is sufficiently far from the source of waves. One can also use a non-reflecting boundary condition, implemented with a PML. The analysis is simpler for the perfectly conducting boundary condition, which is why we use it. However, the results should be the same as those obtained with a PML, as long as our $\partial\Omega$ is not felt by the waves received at the array during the finite duration of measurements.

The wave is defined by the electric field $\vec{E}$ and the magnetic induction $\vec{B}$. These satisfy Faraday's law and Ampère's circuital law in Maxwell's system of equations

$$\mathrm{curl}\vec{E}(t, \vec{x}) + \partial_t \vec{B}(t, \vec{x}) = 0, \tag{3}$$

$$-\mathrm{curl}\vec{B}(t, \vec{x}) + \mu_o \begin{pmatrix} \underline{\underline{\varepsilon}}(x) & \mathbf{0} \\ \mathbf{0}^T & \varepsilon_z(x) \end{pmatrix} \partial_t \vec{E}(t, \vec{x}) = -\mu_o \vec{J}(t, x), \tag{4}$$

for $t \in \mathbb{R}$ and $\vec{x} = (x, z) \in \Omega \times \mathbb{R}$, with boundary condition

$$\vec{n}(x) \times \vec{E}(t, \vec{x}) = \vec{0}, \quad t \in \mathbb{R}, \quad \vec{x} = (x, z) \in \partial\Omega \times \mathbb{R}, \tag{5}$$

and with initial conditions

$$\vec{E}(t, \vec{x}) \equiv \vec{0}, \quad \vec{B}(t, \vec{x}) \equiv \vec{0}, \quad t \ll 0, \quad \vec{x} = (x, z) \in \Omega \times \mathbb{R}. \tag{6}$$

The wave excitation in (4) is a z-independent transient source current density $\vec{J} = (J, J_z)$ supported away from $\partial\Omega$. In the initial conditions, $t \ll 0$ means prior to the excitation.

The assumption (1) on the dielectric permittivity, the choice of the domain and the z-independent source excitation, give that the initial boundary value problem (3)-(6) admits z independent solutions

$$\vec{\boldsymbol{E}}(t,\boldsymbol{x}) = \begin{pmatrix} \boldsymbol{E}(t,\boldsymbol{x}) \\ E_z(t,\boldsymbol{x}) \end{pmatrix}, \quad \vec{\boldsymbol{B}}(t,\boldsymbol{x}) = \begin{pmatrix} \boldsymbol{B}(t,\boldsymbol{x}) \\ B_z(t,\boldsymbol{x}) \end{pmatrix},$$

where $\boldsymbol{E}$ and $\boldsymbol{B}$ are vectors in the plane orthogonal to the z axis. To determine the dielectric permittivity (1), it suffices to work with the electric field, whose components satisfy the decoupled wave equations

$$\nabla^\perp \left[\nabla^\perp \cdot \boldsymbol{E}(t,\boldsymbol{x})\right] - \underline{\underline{c}}^{-2}(\boldsymbol{x})\partial_t^2 \boldsymbol{E}(t,\boldsymbol{x}) = \mu_o \partial_t \boldsymbol{J}(t,\boldsymbol{x}), \tag{7}$$

and

$$\Delta E_z(t,\boldsymbol{x}) - c_z^{-2}(\boldsymbol{x})\partial_t^2 E_z(t,\boldsymbol{x}) = \mu_o \partial_t J_z(t,\boldsymbol{x}), \tag{8}$$

for $t \in \mathbb{R}$ and $\boldsymbol{x} \in \Omega$. The wave speeds $\underline{\underline{c}}$ and c_z are defined by

$$\underline{\underline{c}}(\boldsymbol{x}) = \left[\mu_o \underline{\underline{\varepsilon}}(\boldsymbol{x})\right]^{-1/2}, \quad c_z(\boldsymbol{x}) = 1/\sqrt{\mu_o \varepsilon_z(\boldsymbol{x})}, \tag{9}$$

where we take the unique symmetric and positive definite matrix square root [33, Theorem 7.2.6]. The operator $\nabla^\perp = (-\partial_{x_2}, \partial_{x_1})$ is the two dimensional gradient $\nabla = (\partial_{x_1}, \partial_{x_2})$ rotated by 90^o and $\Delta = \partial_{x_1}^2 + \partial_{x_2}^2$ is the Laplacian. At $t \ll 0$ we have the homogeneous initial conditions (6) and we deduce from (2) and (5) the boundary conditions

$$\boldsymbol{n}^\perp(\boldsymbol{x}) \cdot \boldsymbol{E}(t,\boldsymbol{x}) = 0, \quad E_z(t,\boldsymbol{x}) = 0, \tag{10}$$

where $\boldsymbol{n}^\perp = (-n_2, n_1)$ is the vector $\boldsymbol{n} = (n_1, n_2)$ rotated by 90^o.

Note that (8) is the scalar wave equation in a medium with wave speed c_z. A data driven reduced order model methodology for estimating such a speed has been introduced recently in [14–16]. Thus, we study henceforth the estimation of $\underline{\underline{c}}$ using the vectorial wave equation (7).

The data for the inversion are gathered by an array of m antennas. They emit probing pulses f supported in the short time interval $(-T_f, T_f)$ and with Fourier transform $\widehat{f}$ that is non-negligible at frequencies ω satisfying $|\omega \mp \omega_o| < \pi b$. Here b is the bandwidth and ω_o is the central angular frequency. Typically, $\omega_o \gg b$ and in all imaging applications, the pulse satisfies

$$\widehat{f}(0) = \int_{-T_f}^{T_f} dt\, f(t) = 0. \tag{11}$$

The antennas are modeled by functions $F^{(s)} \in L^2(\Omega)$ supported around $\boldsymbol{x}_s \in \Omega$, in a set of small diameter with respect to the central wavelength λ_o, for $s = 1, \ldots, m$. Since the equation is linear, we can normalize $F^{(s)}$ to integrate to one and formally, we can view it as the Dirac delta at $\boldsymbol{x}_s$. The separation between the antennas is smaller than λ_o, so they behave like a collective entity, aka the array. The excitation from the s^{th} antenna has two polarizations and is given by the current density

$$\boldsymbol{J}^{(s,p)}(\boldsymbol{x}) = \mu_o^{-1} f(t) F^{(s)}(\boldsymbol{x}) \boldsymbol{e}_p, \quad s = 1, \ldots, m, \quad p = 1, 2, \tag{12}$$

where $\boldsymbol{e}_1, \boldsymbol{e}_2$ are the canonical basis vectors in $\mathbb{R}^2$. The resulting electric field, the solution of equation (7) with $\boldsymbol{J}$ replaced by $\boldsymbol{J}^{(s,p)}$, is denoted by $\boldsymbol{E}^{(s,p)}$. Its recordings at all the antennas define the (s,p) column of the $2m \times 2m$ time dependent array response matrix $\boldsymbol{\mathcal{W}}$, with entries

$$\mathcal{W}^{(s',p'),(s,p)}(t) = \boldsymbol{e}_{p'}^T \int_\Omega d\boldsymbol{x}\, F^{(s')}(\boldsymbol{x}) \boldsymbol{E}^{(s,p)}(\boldsymbol{x},t) \approx \boldsymbol{e}_{p'}^T \boldsymbol{E}^{(s,p)}(t, \boldsymbol{x}_{s'}), \tag{13}$$

for $s, s' = 1, \ldots, m$ and $p, p' = 1, 2$.

Note that we consider the idealized sources (12) to simplify the presentation of our methodology. These sources model electric dipoles that are commonly used in three dimensions, thus making the extension to three dimensional media straightforward. There are no such two dimensional electric dipoles in practice, but our theory can be extended to any $2m$ linearly independent source/receiver functions $\boldsymbol{J}^{(s,p)}$, using linear superposition.

The inverse problem is to estimate the 2×2 symmetric, positive definite matrix $\underline{\underline{c}}$ in Ω, from $\boldsymbol{\mathcal{W}}$ measured in some time interval $t \in (t_{\min}, t_{\max})$. The classic approach to solving it uses nonlinear least squares data fit optimization

$$\min_{\underline{\underline{\widetilde{c}}} \in \mathscr{C}} \mathcal{O}^{\text{FWI}}(\underline{\underline{\widetilde{c}}}) + \text{ regularization}, \quad \mathcal{O}^{\text{FWI}}(\underline{\underline{\widetilde{c}}}) = \int_{t_{\min}}^{t_{\max}} dt\, \|\boldsymbol{\mathcal{W}}(t) - \boldsymbol{\mathcal{W}}(t; \underline{\underline{\widetilde{c}}})\|_F^2. \tag{14}$$

Here $\|\cdot\|_F$ is the Frobenius norm, $\underline{\underline{\widetilde{c}}}$ denotes the search wave speed in some user defined space $\mathscr{C}$, and $\underline{\underline{\widetilde{c}}} \mapsto \boldsymbol{\mathcal{W}}(t; \underline{\underline{\widetilde{c}}})$ is the forward map, with output defined by the analogue of (13), calculated with $\underline{\underline{\widetilde{c}}}$ instead of the true $\underline{\underline{c}}$.

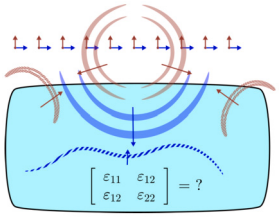

Fig. 1. Illustration of the setup and data acquisition with an active array of antennas that generate probing waves and measure the backscattered waves.

Note our notation convention: We omit the true speed $\underline{\underline{c}}$ from the list of arguments of the electric field and measurements. Thus, $\boldsymbol{E}^{(s,p)}(t,\boldsymbol{x})$ is the field in the true medium, with unknown $\underline{\underline{c}}$, whereas $\boldsymbol{E}^{(s,p)}(t,\boldsymbol{x};\widetilde{\underline{\underline{c}}})$, which defines $\boldsymbol{\mathcal{W}}(t;\widetilde{\underline{\underline{c}}})$, is the computable field for the search speed $\widetilde{\underline{\underline{c}}}$.

The data fit optimization formulation (14) is known as "full waveform inversion" (FWI). It has drawn a lot of attention especially in the geophysics community, where it is studied mostly for acoustic and elastic waves [54]. The FWI formulation is attractive because it is robust to Gaussian additive noise and applies to an arbitrary data acquisition geometry, although the result is dependent of it. In particular, it is easier to invert from measurements gathered all around Ω, so that both the transmitted and backscattered waves can be recorded. Having very low frequencies in the bandwidth of the probing pulse also helps [17,19]. In most applications, low frequencies are not available and the placement of the sensors is limited to one side of Ω, see Fig. 1, so only the backscattered waves can be measured. These restrictions and the complicated nonlinear forward map result in an FWI objective function (14) that has numerous spurious local minima near and far from the true $\underline{\underline{c}}$ [54,27,34]. This pathological behavior is known as "cycle skipping" and means that gradient based optimization methods, like Gauss-Newton, may not give good results even for a good initial guess. Cycle skipping has motivated the pursuit of better objective functions obtained, for example, by replacing the L^2 norm in (14) with the Wasserstein metric [26,27] or by introducing in a systematic way additional degrees of freedom in the optimization, known as "extended FWI" [34,39,49].

The linearization of the FWI objective function about the coefficients of a given reference medium, which is typically homogeneous, leads to a popular qualitative, aka imaging method, known as reverse time migration in the geophysics literature [5,6,48] or backprojection in radar [22,20]. This method succeeds in locating the rough part of the medium, called the reflectivity, if the smooth part of the medium, which determines the kinematics of wave propagation, is known. However, even with known kinematics, the images display "multiple artifacts", due to multiple scattering effects that are ignored by the linearization. The mitigation of such effects has been addressed partially in [53,55,7,32].

Our goal in this paper is to introduce a reduced order model (ROM) approach for both imaging and inverse wave scattering with electromagnetic waves in lossless, non-magnetic, anisotropic media. The ROM is an algebraic dynamical system that captures the wave propagation at discrete time steps $t_j = j\tau$, separated by a time interval τ that is chosen according to the Nyquist sampling criterium for the frequency content of f. It is defined via Galerkin projection of an exact time stepping scheme for equation (7), on the time grid $\{t_j\}_{j\geq 0}$. The projection space is spanned by the snapshots of the electric field at the first n time instants and for $\boldsymbol{x} \in \Omega$. These snapshots are not known, because the medium is not known inside Ω. Nevertheless, the ROM can be computed in a data driven way, from the array response matrix $\boldsymbol{\mathcal{W}}$ evaluated at $t = t_j$, with $j = 0, \ldots, 2n-1$. The number n is chosen so that over the duration $2n\tau$, the waves can travel to the desired depth in the medium, scatter along the way and then travel back to the array, where they are recorded.

The ROM computation is an extension of that introduced in [8,11,10,23] for the acoustic wave equation. We describe it in section 2. Then, we explain in section 3 how we can use the ROM to estimate the snapshots at points in Ω. We use this estimate in section 4 to formulate a qualitative i.e., imaging method, and in section 5 to formulate a quantitative i.e., inversion method. These two methods can be used in conjunction, as we explain in section 6. The performance of the methods is illustrated with numerical simulations in sections 4–6. We end with a summary in section 7.

2. ROM for electromagnetic wave propagation

The derivation of the ROM involves several steps. The first, described in section 2.1, introduces a transformation of the wave field, that leads to a convenient form of the wave equation, with a self-adjoint operator that has $\underline{\underline{c}}$ dependent coefficients. This operator has a nontrivial null space, but we explain in section 2.2 that the wave components in this space are short lived and can be removed from the measurements and the computation of the ROM. The definition of the ROM is based on a Galerkin projection of an exact time stepping scheme given in section 2.3. The data driven ROM computation is in section 2.4. We assume there ideal and noiseless measurements, but we explain in section 2.5 how to regularize the ROM computation so that noise and other factors are mitigated.

2.1. Transformation of the wave field

We assume that the medium is known, isotropic and homogeneous near $\partial\Omega$ and in the vicinity of the array, with dielectric permittivity $\varepsilon_o \underline{\underline{I}}$, which defines the reference, scalar wave speed $c_o = 1/\sqrt{\mu_o \varepsilon_o}$. Here vicinity means within a distance $2c_o T_f$ traveled

by the waves over the duration of the probing pulse f. At larger distance the medium is anisotropic and heterogeneous, with variable wave speed $\underline{\underline{c}}$.

To derive the expression of the ROM, we will use functional calculus on the operator $\underline{\underline{c}}^2\nabla^\perp[\nabla^\perp\cdot]$ with domain

$$\boldsymbol{H}_0(\nabla^\perp;\Omega)=\left\{\boldsymbol{\psi}(\boldsymbol{x})\in\boldsymbol{L}^2(\Omega):\ \nabla^\perp\cdot\boldsymbol{\psi}(\boldsymbol{x})\in L^2(\Omega),\ \ \boldsymbol{n}^\perp\cdot\boldsymbol{\psi}|_{\partial\Omega}=0\right\}. \tag{15}$$

This operator is the two dimensional version of the operator $\underline{\underline{c}}^2\mathrm{curl}\,\mathrm{curl}$ with domain $\boldsymbol{H}_0(\mathrm{curl};\Omega)$ [41, Theorem 3.33].

Note that $\underline{\underline{c}}^2\nabla^\perp[\nabla^\perp\cdot]$ is symmetric in the inner product weighted by $\underline{\underline{c}}^{-2}$. We prefer to work with the Euclidian inner product, so we define

$$\boldsymbol{U}^{(s,p)}(t,\boldsymbol{x})=c_o\underline{\underline{c}}^{-1}(\boldsymbol{x})\boldsymbol{E}^{(s,p)}(t,\boldsymbol{x}),\quad t\in\mathbb{R},\ \boldsymbol{x}\in\Omega, \tag{16}$$

which equals the electric field in the vicinity of the array, because $\underline{\underline{c}}=c_o\underline{\underline{I}}$ there. We derive from (6), (7) and (10), with excitation (12), the equation

$$\left(A+\partial_t^2\right)\boldsymbol{U}^{(s,p)}(t,\boldsymbol{x})=-c_o^2f'(t)F^{(s)}(\boldsymbol{x})\boldsymbol{e}_p,\quad t>0,\ \boldsymbol{x}\in\Omega, \tag{17}$$

with homogeneous initial and boundary conditions

$$\boldsymbol{U}^{(s,p)}(t,\boldsymbol{x})=\boldsymbol{0},\quad t<-T_f,\ \boldsymbol{x}\in\Omega, \tag{18}$$

$$\boldsymbol{n}^\perp\cdot\boldsymbol{U}^{(s,p)}(t,\boldsymbol{x})=0,\quad t\in\mathbb{R},\ \boldsymbol{x}\in\partial\Omega. \tag{19}$$

Here we introduced the operator

$$A=-\underline{\underline{c}}(\boldsymbol{x})\nabla^\perp\left[\nabla^\perp\cdot\left(\underline{\underline{c}}(\boldsymbol{x})\cdot\right)\right], \tag{20}$$

with domain $\boldsymbol{\mathcal{D}}$ consisting of vector valued functions whose multiplication by $\underline{\underline{c}}$ lies in the space $\boldsymbol{H}_0(\nabla^\perp;\Omega)$.

2.2. Wave decomposition and data transformation

The operator A defined in (20) is self-adjoint and positive semidefinite. Its null space is a closed subspace of $\boldsymbol{\mathcal{D}}$ and consists of functions $\underline{\underline{c}}^{-1}\nabla N$, where N is scalar valued in $H^1(\Omega)$, with constant trace at $\partial\Omega$ [41, Lemma 4.5].

The solution of (17)–(19) has the Helmholtz decomposition [41, Lemma 4.5]

$$\boldsymbol{U}^{(s,p)}(t,\boldsymbol{x})=\mathbb{U}^{(s,p)}(t,\boldsymbol{x})+\underline{\underline{c}}^{-1}(\boldsymbol{x})\nabla N^{(s,p)}(t,\boldsymbol{x}), \tag{21}$$

where $\underline{\underline{c}}^{-1}\mathbb{U}^{(s,p)}$ is divergence free, in the weak sense. The first term in this decomposition solves the wave equation

$$\left(A+\partial_t^2\right)\mathbb{U}^{(s,p)}(t,\boldsymbol{x})=-c_o^2f'(t)\mathbb{F}^{(s,p)}(\boldsymbol{x}),\quad t\in\mathbb{R},\ \boldsymbol{x}\in\Omega, \tag{22}$$

driven by the divergence free part $\mathbb{F}^{(s,p)}$ of the excitation, the projection of $F^{(s)}(\boldsymbol{x})\boldsymbol{e}_p$ on range(A). This range is a closed subspace of $\boldsymbol{L}^2(\Omega)$, consisting of functions $\boldsymbol{\psi}$ such that $\underline{\underline{c}}^{-1}\boldsymbol{\psi}$ is divergence free, in the weak sense. According to [40, Section 6.6] or by direct calculation, $\left[\mathrm{range}(A)\right]^\perp=\mathrm{null}(A)$.

The second term in (21) lies in null(A) and solves

$$\partial_t^2\nabla\cdot\left[\underline{\underline{c}}^{-2}(\boldsymbol{x})\nabla N^{(s,p)}(t,\boldsymbol{x})\right]=-c_of'(t)\partial_{x_p}F^{(s)}(\boldsymbol{x}),\quad t\in\mathbb{R},\ \boldsymbol{x}\in\Omega, \tag{23}$$

with homogeneous initial condition at $t<-T_f$. This term is short lived, because integrating (23) in time and using the initial condition, we obtain

$$\nabla\cdot\left[\underline{\underline{c}}^{-2}(\boldsymbol{x})\nabla N^{(s,p)}(t,\boldsymbol{x})\right]=-c_o\int_{-T_f}^{t}dt'\,f(t')\partial_{x_p}F^{(s)}(\boldsymbol{x}). \tag{24}$$

The right hand side of this equation vanishes for $t>T_f$, due to the property (11) of the probing signal f.

Since the medium is known and homogeneous within a distance $2c_oT_f$ from the array, we can compute $\boldsymbol{U}^{(s,p)}$ at $t\in(-T_f,T_f)$ and remove its component in the null space. The array response matrix is then transformed to the matrix $\mathbb{W}(t)\in\mathbb{R}^{2m\times2m}$, with entries

$$\mathbb{W}^{(s',p'),(s,p)}(t)=\boldsymbol{e}_{p'}^T\int_\Omega d\boldsymbol{x}\,F^{(s')}(\boldsymbol{x})\mathbb{U}^{(s,p)}(t,\boldsymbol{x})\approx\boldsymbol{e}_{p'}^T\mathbb{U}^{(s,p)}(t,\boldsymbol{x}_{s'}), \tag{25}$$

for $s,s'=1,\ldots,m$ and $p,p'=1,2$.

Once we have removed the null space components of the wave, we can work with the operator $\mathbb{A}$, the restriction of A to $\boldsymbol{\mathcal{D}}\setminus\mathrm{null}(A)$. This operator is positive definite, with compact and self-adjoint resolvent [41, Section 4.7]. It has a countable infinite set of positive

eigenvalues $(\theta_n)_{n\geq 1}$ sorted in increasing order, with $\theta_n \to \infty$ as $n \to \infty$, and the eigenfunctions $(\boldsymbol{\varphi}_n)_{n\geq 1}$ form an orthonormal basis of range(A). Functional calculus on $\mathbb{A}$ is defined as usual: If $\Psi : \mathbb{C} \mapsto \mathbb{C}$ is a continuous function, then $\Psi(\mathbb{A})$ is the operator with the same eigenfunctions as $\mathbb{A}$ and the eigenvalues $\Psi(\theta_n)$, for $n \geq 1$.

Let us write the solution of equation (22) as a series,[1] using the spectral decomposition of $\mathbb{A}$,

$$\mathbb{U}^{(s,p)}(t,\boldsymbol{x}) = -c_o^2 f(t) \star_t 1_{[0,\infty)}(t) \sum_{j=1}^{\infty} \cos(t\sqrt{\theta_j}) \Big[e_p^T \int_{\Omega} d\boldsymbol{x}'\, \boldsymbol{F}^{(s)}(\boldsymbol{x}') \boldsymbol{\varphi}_j(\boldsymbol{x}') \Big] \boldsymbol{\varphi}_j(\boldsymbol{x}),$$

where $1_{[0,\infty)}$ is the indicator function of the interval $[0,\infty)$ equal to 1 if $t \geq 0$ and zero otherwise and $\star_t$ denotes convolution in t. For the derivation of the ROM, it is convenient to work with the wave

$$\begin{aligned} \mathbb{U}_e^{(s,p)}(t,\boldsymbol{x}) &= -c_o^{-2} \left[f(-t) \star_t \mathbb{U}^{(s,p)}(t,\boldsymbol{x}) + f(t) \star_t \mathbb{U}^{(s,p)}(-t,\boldsymbol{x}) \right] \\ &= \sum_{j=1}^{\infty} \cos(t\sqrt{\theta_j}) |\widehat{f}(\sqrt{\theta_j})|^2 \Big[e_p^T \int_{\Omega} d\boldsymbol{x}'\, \boldsymbol{F}^{(s)}(\boldsymbol{x}') \boldsymbol{\varphi}_j(\boldsymbol{x}') \Big] \boldsymbol{\varphi}_j(\boldsymbol{x}). \end{aligned} \tag{26}$$

The last equality is obtained by writing the time convolutions in the Fourier domain, where $|\widehat{f}\,|$ is even, and we have the Fourier transform formula

$$\int_{-\infty}^{\infty} dt\, 1_{[0,\infty)}(t) \cos(t\sqrt{\theta_j}) e^{i\omega t} = \frac{\pi}{2} \left[\delta(\omega - \sqrt{\theta_j}) + \delta(\omega + \sqrt{\theta_j}) \right] + \frac{i\omega}{\theta_j - \omega^2}.$$

The purpose of introducing the even in time wave $\mathbb{U}_e^{(s,p)}$ is twofold: First, the transformation in (26) is like a Duhamel principle which maps the source excitation to an initial condition. Indeed, the right hand side in (26) equals

$$\mathbb{U}_e^{(s,p)}(t,\boldsymbol{x}) = |\widehat{f}(\sqrt{\mathbb{A}})| \boldsymbol{u}^{(s,p)}(t,\boldsymbol{x}), \tag{27}$$

where

$$\boldsymbol{u}^{(s,p)}(t,\boldsymbol{x}) = \cos(t\sqrt{\mathbb{A}}) \boldsymbol{u}_0^{(s,p)}(\boldsymbol{x}), \tag{28}$$

solves the homogeneous wave equation

$$[\mathbb{A} + \partial_t^2] \boldsymbol{u}^{(s,p)}(t,\boldsymbol{x}) = 0, \quad t > 0, \quad \boldsymbol{x} \in \Omega, \tag{29}$$

with initial condition

$$\boldsymbol{u}^{(s,p)}(0,\boldsymbol{x}) = \boldsymbol{u}_0^{(s,p)}(\boldsymbol{x}), \quad \partial_t \boldsymbol{u}^{(s,p)}(0,\boldsymbol{x}) = 0, \quad \boldsymbol{x} \in \Omega, \tag{30}$$

defined by

$$\boldsymbol{u}_0^{(s,p)}(\boldsymbol{x}) = \sum_{j=1}^{\infty} |\widehat{f}(\sqrt{\theta_j})| \Big[e_p^T \int_{\Omega} d\boldsymbol{x}'\, \boldsymbol{F}^{(s)}(\boldsymbol{x}') \boldsymbol{\varphi}_j(\boldsymbol{x}') \Big] \boldsymbol{\varphi}_j(\boldsymbol{x}). \tag{31}$$

The second purpose of (26) is that the measurements (25), which are approximately point evaluations of the wave fields, can be mapped to a new data matrix

$$\mathbb{D}(t) = -c_o^{-2} \left[f(-t) \star_t \mathbb{W}(t) + f(t) \star_t \mathbb{W}(-t) \right], \tag{32}$$

whose entries can be expressed as the inner products

$$\begin{aligned} \mathbb{D}^{(s',p'),(s,p)}(t) &= \int_{\Omega} d\boldsymbol{x}\, \boldsymbol{e}_{p'}^T \boldsymbol{F}^{(s')}(\boldsymbol{x}) \mathbb{U}_e^{(s,p)}(t,\boldsymbol{x}) \\ &\overset{(27),(31)}{=} \int_{\Omega} d\boldsymbol{x} \left[\boldsymbol{u}_0^{(s',p')}(\boldsymbol{x}) \right]^T \boldsymbol{u}^{(s,p)}(t,\boldsymbol{x}), \end{aligned} \tag{33}$$

for $s, s' = 1, \ldots, m$ and $p, p' = 1, 2$. As we shall see, the matrices arising in the Galerkin projection scheme that define our ROM also involve such inner products, which is why we can compute the ROM in a data driven way.

Note that the second term in definition (32) contributes only at $t < 2T_f$ and it can be determined from the measurements if the recordings start at $t = -T_F$. Otherwise, we can compute it because the medium is known and homogeneous near the array. We work henceforth with the data matrix $\mathbb{D}$ and show below how we can compute the ROM from it.

[1] The square bracket equals $\int_{\Omega} d\boldsymbol{x}'\, \mathbb{F}^{(s,p)}(\boldsymbol{x}') \cdot \boldsymbol{\varphi}_j(\boldsymbol{x}')$, because $\boldsymbol{F}^{(s)}(\boldsymbol{x}) \boldsymbol{e}_p - \mathbb{F}^{(s,p)} \in \big[\text{range}(A)\big]^{\perp}$ and $\boldsymbol{\varphi}_j \in \text{range}(A)$.

2.3. Wave snapshots and time stepping

We are interested in the evolution of the wave fields (28) on a uniform time grid $\{t_j = j\tau, j \geq 0\}$, with step chosen according to the Nyquist criterium $\tau \lesssim \pi/\omega_o$. To avoid heavy index notation, we use block algebra and define the $2 \times 2m$ dimensional field, called the wave snapshot at time t_j,

$$\begin{aligned} \boldsymbol{u}_j(\boldsymbol{x}) &= \left(\boldsymbol{u}^{(1,1)}(t_j,\boldsymbol{x}), \boldsymbol{u}^{(1,2)}(t_j,\boldsymbol{x}), \ldots, \boldsymbol{u}^{(m,1)}(t_j,\boldsymbol{x}), \boldsymbol{u}^{(m,2)}(t_j,\boldsymbol{x})\right) \\ &\overset{(28)}{=} \cos(j\tau\sqrt{\mathbb{A}})\boldsymbol{u}_0(\boldsymbol{x}), \quad \boldsymbol{x} \in \Omega, \quad j \geq 0. \end{aligned} \tag{34}$$

We can now see the advantage of the transformations in the previous section: The trigonometric identity

$$\cos[(j+1)\alpha] + \cos[|j-1|\alpha] = 2\cos(\alpha)\cos(j\alpha),$$

for $\alpha \in \mathbb{R}$, and functional calculus give that the snapshots satisfy the exact time stepping scheme

$$\boldsymbol{u}_{j+1}(\boldsymbol{x}) = 2\mathcal{P}\boldsymbol{u}_j(\boldsymbol{x}) - \boldsymbol{u}_{|j-1|}(\boldsymbol{x}), \quad j \geq 0, \ \boldsymbol{x} \in \Omega, \tag{35}$$

driven by the "propagator operator"

$$\mathcal{P} = \cos(\tau\sqrt{\mathbb{A}}). \tag{36}$$

Equation (35) defines a discrete time dynamical system, with initial state $\boldsymbol{u}_0$. The ROM derived next is an algebraic discrete dynamical system, obtained via Galerkin projection of (35) on the space

$$\mathcal{S} = \text{span}\{\boldsymbol{\mathcal{U}}(\boldsymbol{x})\}, \quad \boldsymbol{\mathcal{U}}(\boldsymbol{x}) = \left(\boldsymbol{u}_0(\boldsymbol{x}), \ldots, \boldsymbol{u}_{n-1}(\boldsymbol{x})\right), \quad \boldsymbol{x} \in \Omega. \tag{37}$$

The end time index n is chosen according to the distance from the array at which we wish to image. If this distance is L, then we should have $nc_o\tau \gtrsim L$.

Remark 1. If the data are noiseless, the time step τ is chosen according to the Nyquist criterium and the antenna separation is of order λ_o, the dimension of $\mathcal{S}$ is typically $2nm$. We assume that this is so for now, but we discuss later, in section 2.5, how to deal with the case of linearly dependent snapshots and noisy data.

2.4. Data driven ROM

The Galerkin approximation of the snapshots (34) is defined in a standard way

$$\boldsymbol{u}_j^{\text{Gal}}(\boldsymbol{x}) = \boldsymbol{\mathcal{U}}(\boldsymbol{x})\mathfrak{g}_j, \quad j \geq 0, \ \boldsymbol{x} \in \Omega, \tag{38}$$

where $\mathfrak{g}_j \in \mathbb{R}^{2mn \times 2m}$ are the Galerkin coefficients, calculated so that when substituting (38) in (35), the residual is orthogonal to the space $\mathcal{S}$.

To write the Galerkin equation, we use an orthonormal basis stored in the $2 \times 2nm$ dimensional field $\boldsymbol{V}(\boldsymbol{x}) = \left(\boldsymbol{v}_0(\boldsymbol{x}), \ldots, \boldsymbol{v}_{n-1}(\boldsymbol{x})\right)$, whose $2 \times 2m$ components satisfy

$$\int_\Omega d\boldsymbol{x}\, \boldsymbol{v}_j^T(\boldsymbol{x})\boldsymbol{v}_k(\boldsymbol{x}) = \boldsymbol{I}_{2m}\delta_{jk}, \quad j,k = 0, \ldots, n-1. \tag{39}$$

Here $\boldsymbol{I}_{2m}$ denotes the $2m \times 2m$ identity matrix and δ_{jk} is the Kronecker delta symbol. We index the components $\boldsymbol{v}_j$ of $\boldsymbol{V}$ the same way as the snapshots, because our basis is causal i.e.,

$$\boldsymbol{v}_j(\boldsymbol{x}) \in \text{span}\left\{\boldsymbol{u}_0(\boldsymbol{x}), \ldots, \boldsymbol{u}_j(\boldsymbol{x})\right\}, \quad j = 0, \ldots, n-1. \tag{40}$$

Explicitly, $\boldsymbol{V}$ is defined via the Gram-Schmidt orthogonalization of $\boldsymbol{\mathcal{U}}$,

$$\boldsymbol{\mathcal{U}}(\boldsymbol{x}) = \boldsymbol{V}(\boldsymbol{x})\boldsymbol{R}, \quad \boldsymbol{x} \in \Omega, \tag{41}$$

where $\boldsymbol{R}$ is block upper triangular, with $2m \times 2m$ blocks.

The equation for the Galerkin coefficients can now be written as

$$\begin{aligned} \boldsymbol{0} &= \int_\Omega d\boldsymbol{x}\, \boldsymbol{V}^T(\boldsymbol{x}) \left[\boldsymbol{\mathcal{U}}(\boldsymbol{x})\left(\mathfrak{g}_{j+1} + \mathfrak{g}_{|j-1|}\right) - 2\mathcal{P}\boldsymbol{\mathcal{U}}(\boldsymbol{x})\mathfrak{g}_j\right] \\ &= \boldsymbol{R}^{-T}\left[\mathbb{M}\left(\mathfrak{g}_{j+1} + \mathfrak{g}_{|j-1|}\right) - 2\mathbb{S}\mathfrak{g}_j\right], \quad j \geq 0, \end{aligned} \tag{42}$$

with initial condition

$$\mathfrak{g}_0 = i_0, \text{ such that } \boldsymbol{u}_0^{\text{Gal}}(\boldsymbol{x}) = \boldsymbol{u}_0(\boldsymbol{x}), \quad \boldsymbol{x} \in \Omega. \tag{43}$$

Here $\boldsymbol{R}^{-T}$ is the transpose of the inverse of $\boldsymbol{R}$ and we introduced the Gramian, also known as "mass" matrix

$$\mathbb{M} = \int_{\Omega} d\boldsymbol{x}\, \boldsymbol{\mathcal{U}}^T(\boldsymbol{x})\boldsymbol{\mathcal{U}}(\boldsymbol{x}) \in \mathbb{R}^{2nm\times 2nm}, \tag{44}$$

and the "stiffness matrix"

$$\mathbb{S} = \int_{\Omega} d\boldsymbol{x}\, \boldsymbol{\mathcal{U}}^T(\boldsymbol{x})\mathcal{P}\boldsymbol{\mathcal{U}}(\boldsymbol{x}) \in \mathbb{R}^{2nm\times 2nm}, \tag{45}$$

of the Galerkin scheme. We also denote by $\boldsymbol{i}_j$ the $(j+1)^{\text{th}}$ column block of size $2nm\times 2m$, of the $2nm\times 2nm$ identity matrix $\boldsymbol{I}_{2nm}$.

The following theorem is a straightforward generalization of the results in [8,11,10,23], obtained for the scalar wave equation. We include its proof for the convenience of the reader. The result says that even though we do not know the snapshots stored in $\boldsymbol{\mathcal{U}}$ and therefore the approximation space (37), all the terms in the Galerkin equation (42) are data driven.

Theorem 1. *The $2m\times 2m$ blocks of the mass and stiffness matrices (44)-(45) are determined by the data matrices (32) as follows*

$$\mathbb{M}_{j,l} = \frac{1}{2}\left[\mathbb{D}(t_{j+l}) + \mathbb{D}(t_{|j-l|})\right], \tag{46}$$

$$\mathbb{S}_{j,l} = \frac{1}{4}\left[\mathbb{D}(t_{j+l+1}) + \mathbb{D}(t_{|j-l-1|}) + \mathbb{D}(t_{|j+l-1|}) + \mathbb{D}(t_{|j-l+1|})\right], \tag{47}$$

for $j,l = 0,\dots,n-1$. The block upper triangular matrix $\boldsymbol{R}$ in the Gram-Schmidt orthogonalization is the block Cholesky square root of the mass matrix[2]

$$\mathbb{M} = \boldsymbol{R}^T\boldsymbol{R}. \tag{48}$$

The first n Galerkin coefficients are the $2nm\times 2m$ block columns of $\boldsymbol{I}_{2nm}$,

$$\mathfrak{g}_j = \boldsymbol{i}_j, \quad j = 0,\dots,n-1. \tag{49}$$

Proof. Note that equation (49) combined with definition (38) give that the Galerkin approximation of the snapshots is exact for the first n time instants. This seems natural because the first n snapshots span the approximation space. However, for the result to hold, it is essential that the time stepping scheme (35) is exact, so that with the coefficients (49) we have a zero residual. If we used an approximate time stepping, obtained via some finite difference approximation of ∂_t^2 in the wave equation, (49) would not be satisfied.

The proof of (49) is as follows: According to Remark 1, $\dim(\mathcal{S}) = 2nm$. This means that the matrix $\boldsymbol{R}$ in the Gram-Schmidt orthogonalization (41) and the mass matrix are full rank. Then, equation (42) with initial condition (43) has a unique solution $\mathfrak{g}_j$, for $j \geq 1$. But $\{\mathfrak{g}_j = \boldsymbol{i}_j\}_{j=0}^{n-1}$ solve (42), because

$$\boldsymbol{\mathcal{U}}(\boldsymbol{x})(\boldsymbol{i}_{j+1} + \boldsymbol{i}_{|j-1|}) - 2\mathcal{P}\boldsymbol{\mathcal{U}}(\boldsymbol{x})\boldsymbol{i}_j = \boldsymbol{u}_{j+1} + \boldsymbol{u}_{|j-1|} - 2\mathcal{P}\boldsymbol{u}_j \overset{(35)}{=} \boldsymbol{0},$$

for $j = 0,\dots,n-1$. This proves (49).

The block Cholesky factorization (48) is deduced from the definition (44) of the mass matrix and the Gram-Schmidt equation (41)

$$\mathbb{M} = \boldsymbol{R}^T\int_{\Omega} d\boldsymbol{x}\, \boldsymbol{V}^T(\boldsymbol{x})\boldsymbol{V}(\boldsymbol{x})\boldsymbol{R} \overset{(39)}{=} \boldsymbol{R}^T\boldsymbol{R}.$$

The expression (46) of the blocks of $\mathbb{M}$ is derived using the definition of the snapshots and the symmetry of $\mathbb{A}$, as follows

$$\begin{aligned}
\mathbb{M}_{j,l} &\overset{(44)}{=} \int_{\Omega} d\boldsymbol{x}\, \boldsymbol{u}_j^T(\boldsymbol{x})\boldsymbol{u}_l(\boldsymbol{x}) \\
&\overset{(34)}{=} \int_{\Omega} d\boldsymbol{x}\, \boldsymbol{u}_0^T(\boldsymbol{x})\cos(j\tau\sqrt{\mathbb{A}})\cos(l\tau\sqrt{\mathbb{A}})\boldsymbol{u}_0(\boldsymbol{x}) \\
&= \int_{\Omega} d\boldsymbol{x}\, \boldsymbol{u}_0^T(\boldsymbol{x})\frac{1}{2}\left[\cos((j+l)\tau\sqrt{\mathbb{A}}) + \cos(|j-l|\tau\sqrt{\mathbb{A}})\right]\boldsymbol{u}_0(\boldsymbol{x}) \\
&\overset{(34),(33)}{=} \frac{1}{2}\left[\mathbb{D}(t_{j+l}) + \mathbb{D}(t_{|j-l|})\right], \quad j,l = 0,\dots,n-1.
\end{aligned}$$

Finally, the blocks of the stiffness matrix are

[2] To ensure a unique block Cholesky factorization, we take the symmetric, positive definite square root of the diagonal blocks.

$$\mathbb{S}_{j,l} \overset{(44)}{=} \int_\Omega d\boldsymbol{x}\, \boldsymbol{u}_j^T(\boldsymbol{x}) \mathcal{P} \boldsymbol{u}_l(\boldsymbol{x}) \overset{(35)}{=} \int_\Omega d\boldsymbol{x}\, \boldsymbol{u}_j^T(\boldsymbol{x}) \frac{1}{2} \left[\boldsymbol{u}_{l+1}(\boldsymbol{x}) + \boldsymbol{u}_{|l-1|}(\boldsymbol{x}) \right],$$

and the result (47) follows from the calculation above. □

2.4.1. The ROM

We define the ROM as an algebraic, discrete time dynamical system analogue of (35), with states

$$\boldsymbol{u}_j^{\text{ROM}} = \boldsymbol{R} \mathfrak{g}_j, \quad j \geq 0, \tag{50}$$

called the ROM snapshots. The first n such snapshots are the $2nm \times 2m$ block columns of $\boldsymbol{R}$, as follows from equation (49),

$$\left(\boldsymbol{u}_0^{\text{ROM}}, \ldots, \boldsymbol{u}_{n-1}^{\text{ROM}} \right) = \boldsymbol{R}. \tag{51}$$

The evolution of the ROM snapshots is dictated by the equation

$$\boldsymbol{u}_{j+1}^{\text{ROM}} = 2\mathcal{P}^{\text{ROM}} \boldsymbol{u}_j^{\text{ROM}} - \boldsymbol{u}_{|j-1|}^{\text{ROM}}, \quad j \geq 0, \tag{52}$$

derived from (42) and the Cholesky factorization (48), where

$$\mathcal{P}^{\text{ROM}} = \boldsymbol{R}^{-T} \mathbb{S} \boldsymbol{R}^{-1} \tag{53}$$

is the $2nm \times 2nm$ ROM propagator matrix.

By Theorem 1, $\boldsymbol{R}$ and $\mathbb{S}$ are data driven, so equations (50) and (53) give that the ROM can be computed directly from $\mathbb{D}(t_j)$, for $j = 0, \ldots, 2n-1$.

2.4.2. Properties of the ROM

If we solve for $\boldsymbol{R}$ in the Gram-Schmidt orthogonalization equation (41) and then use the result in the definition (50) of the ROM snapshots, we obtain

$$\boldsymbol{u}_j^{\text{ROM}} = \int_\Omega d\boldsymbol{x}\, \boldsymbol{V}^T(\boldsymbol{x}) \boldsymbol{U}(\boldsymbol{x}) \mathfrak{g}_j \overset{(38)}{=} \int_\Omega d\boldsymbol{x}\, \boldsymbol{V}^T(\boldsymbol{x}) \boldsymbol{u}_j^{\text{Gal}}(\boldsymbol{x}), \quad j \geq 0. \tag{54}$$

Thus, the ROM snapshots are the orthogonal projection of the Galerkin approximation of the snapshots on the space $\mathcal{S}$ defined in (37). Since this approximation is exact at the first n time instants, we deduce from this equation and (41) that

$$\boldsymbol{u}_j^{\text{ROM}} = \int_\Omega d\boldsymbol{x}\, \boldsymbol{V}^T(\boldsymbol{x}) \boldsymbol{u}_j(\boldsymbol{x}) \quad \& \quad \boldsymbol{u}_j(\boldsymbol{x}) = \boldsymbol{V}(\boldsymbol{x}) \boldsymbol{u}_j^{\text{ROM}}, \quad j = 0, \ldots, n-1. \tag{55}$$

The ROM propagator matrix defined in (53) equals the orthogonal projection of the propagator operator (36) on $\mathcal{S}$,

$$\mathcal{P}^{\text{ROM}} = \boldsymbol{R}^{-T} \mathbb{S} \boldsymbol{R}^{-1} \overset{(45)}{=} \int_\Omega d\boldsymbol{x}\, \boldsymbol{R}^{-T} \boldsymbol{U}^T(\boldsymbol{x}) \mathcal{P} \boldsymbol{U}(\boldsymbol{x}) \boldsymbol{R}^{-1}$$
$$\overset{(41)}{=} \int_\Omega d\boldsymbol{x}\, \boldsymbol{V}^T(\boldsymbol{x}) \mathcal{P} \boldsymbol{V}(\boldsymbol{x}). \tag{56}$$

It is a symmetric $2nm \times 2nm$ matrix, with block tridiagonal structure. One way to deduce this structure is to follow the proof in [8, Appendix C]. Alternatively, the result follows by iterating equation (52) for $j = 0, \ldots, n-2$ and using (51): If we count the $2m \times 2m$ blocks of $\mathcal{P}^{\text{ROM}}$ as $\mathcal{P}^{\text{ROM}}_{j,l}$, with $j, l = 0, \ldots, n-1$, and we let $\boldsymbol{R}_{j,l}$ be the blocks of $\boldsymbol{R}$, which are non-zero for $0 \leq j \leq l \leq n-1$, we obtain from (52) evaluated at $j = 0$ that

$$2\mathcal{P}^{\text{ROM}} \boldsymbol{R} \boldsymbol{i}_0 = 2 \begin{pmatrix} \mathcal{P}^{\text{ROM}}_{0,0} \boldsymbol{R}_{0,0} \\ \mathcal{P}^{\text{ROM}}_{1,0} \boldsymbol{R}_{0,0} \\ \vdots \\ \mathcal{P}^{\text{ROM}}_{n-1,0} \boldsymbol{R}_{0,0} \end{pmatrix} = 2\boldsymbol{R} \boldsymbol{i}_1 = 2 \begin{pmatrix} \boldsymbol{R}_{0,1} \\ \boldsymbol{R}_{1,1} \\ \boldsymbol{0} \\ \vdots \\ \boldsymbol{0} \end{pmatrix}.$$

The block $\boldsymbol{R}_{0,0}$ is invertible, because $\boldsymbol{R}$ is invertible, so $\mathcal{P}^{\text{ROM}}_{j,0} = \mathcal{P}^{\text{ROM}}_{0,j} = \boldsymbol{0}$, for $j = 2, \ldots, n-1$. The next step, obtained from (52) evaluated at $j = 1$, gives

$$2\mathcal{P}^{\text{ROM}} \boldsymbol{R} \boldsymbol{i}_1 = \boldsymbol{R} \boldsymbol{i}_2 + \boldsymbol{R} \boldsymbol{i}_0 \quad \Longrightarrow \quad \mathcal{P}^{\text{ROM}}_{j,1} \boldsymbol{R}_{1,1} = \boldsymbol{0}, \quad j = 3, \ldots, n-1,$$

and since $\boldsymbol{R}_{1,1}$ is invertible, $\mathcal{P}^{\text{ROM}}_{j,1} = \mathcal{P}^{\text{ROM}}_{1,j} = \boldsymbol{0}$, for $j = 3, \ldots, n-1$. Proceeding this way, until we reach $j = n-2$ in equation (52), we obtain that $\mathcal{P}^{\text{ROM}}$ is indeed block tridiagonal.

The ROM interpolates the time dependent data matrix $\mathbb{D}$ from which it is computed. Specifically, for $j = 0, \ldots, n-1$, we have

$$\mathbb{D}(t_j) \overset{(33)}{=} \int_\Omega d\boldsymbol{x}\, \boldsymbol{u}_0^T(\boldsymbol{x}) \boldsymbol{u}_j(\boldsymbol{x}) \overset{(55)}{=} (\boldsymbol{u}_0^{\text{ROM}})^T \int_\Omega d\boldsymbol{x} \boldsymbol{V}^T(\boldsymbol{x}) \boldsymbol{V}(\boldsymbol{x}) \boldsymbol{u}_j^{\text{ROM}}$$
$$\overset{(39)}{=} (\boldsymbol{u}_0^{\text{ROM}})^T \boldsymbol{u}_j^{\text{ROM}}, \tag{57}$$

and for $j = 1, \ldots, n-1$ we have

$$\mathbb{D}(t_{n-1+j}) \overset{(46)}{=} 2\mathbb{M}_{n-1,j} - \mathbb{D}(t_{n-1-j}). \tag{58}$$

The second term in the right hand side of this equation is matched as in (57) and the first term is

$$\mathbb{M}_{n-1,j} = \int_\Omega d\boldsymbol{x}\, \boldsymbol{u}_{n-1}^T(\boldsymbol{x})\boldsymbol{u}_j(\boldsymbol{x}) \overset{(55)}{=} (\boldsymbol{u}_{n-1}^{\text{ROM}})^T \int_\Omega d\boldsymbol{x} \boldsymbol{V}^T(\boldsymbol{x})\boldsymbol{V}(\boldsymbol{x})\boldsymbol{u}_j^{\text{ROM}}$$
$$\overset{(39)}{=} (\boldsymbol{u}_{n-1}^{\text{ROM}})^T \boldsymbol{u}_j^{\text{ROM}}. \tag{59}$$

Equations (57)–(59) show that the ROM snapshots give exactly the matrices $\mathbb{D}(t_j)$ at $t_j = j\tau$, for $j = 0, \ldots, 2n-2$. In fact, with a bit more work, one can show that this result extends one more step, to $j = 2n-1$ [8, Appendix B].

2.5. Regularization of the ROM computation

There are two critical and nonlinear steps in the computation of the ROM: The Cholesky factorization (48) of $\mathbb{M}$ and the inversion of the block Cholesky square root $\boldsymbol{R}$. Both steps require a symmetric positive definite $\mathbb{M}$ computed from the data as in equation (46).

In the absence of noise, the time dependent matrix $\mathbb{D}$, which is obtained from the array response matrix as in equation (32), is symmetric due to source/receiver reciprocity. This ensures the symmetry of $\mathbb{M}$. The matrix $\mathbb{D}$ is not symmetric for noisy measurements, but once we compute $\mathbb{M}$ from equation (46), we can just take $(\mathbb{M} + \mathbb{M}^T)/2$ to obtain a symmetric mass matrix.

If the time step τ and/or the separation between the antennas are chosen too small, then $\mathbb{M}$ is singular in finite precision arithmetic. For noisy measurements, even after the symmetrization, $\mathbb{M}$ is likely indefinite. Thus, the ROM construction requires a regularization procedure, which maps $\mathbb{M}$ to a symmetric, positive definite $\mathbb{M}^{\text{reg}}$.

There are various ways to carry out the regularization. Each one must ensure the correct causal structure of the ROM, which is manifested algebraically in the block upper triangular square root $\boldsymbol{R}$ of $\mathbb{M}$ and the tridiagonal structure of the ROM propagator matrix $\mathcal{P}^{\text{ROM}}$. We have tried the following two regularization approaches:

1. Replace $\mathbb{D}(0)$ by $(1+2\alpha)\mathbb{D}(0)$, which according to (46), amounts to boosting the diagonal block of $\mathbb{M}$ by $\alpha\mathbb{D}(0)$, where $0 < \alpha \ll 1$.
2. Project $\mathbb{M}$ on the space spanned by the eigenvectors corresponding to the eigenvalues that exceed a user defined positive threshold.

The advantage of the first regularization approach is that it is simple and preserves the causal algebraic structure of the ROM. The matrix $\mathbb{D}(0)$ can be easily computed because it depends only on the medium near the array, which is known, isotropic and homogeneous. It is a positive definite $2m \times 2m$ matrix, so by adding $\alpha\mathbb{D}(0)$ to the diagonal of $\mathbb{M}$, we can get a positive definite $2nm \times 2nm$ regularized $\mathbb{M}^{\text{reg}}$. The disadvantage is that the eigenvectors of $\mathbb{M}^{\text{reg}}$ corresponding to the smallest eigenvalues are significantly affected by noise, which can cause inversion artifacts.

The second regularization approach is more complicated, but it has the advantage that it removes the subspace spanned by the "noisy eigenvectors". Let $\{\boldsymbol{y}_j\}_{j=1}^{2nm}$ be the orthonormal eigenfunctions of $\mathbb{M}$, corresponding to the eigenvalues counted in descending order. For regularization, we wish to keep the largest $2rm$ eigenvalues and associated eigenvectors, for $1 \le r < n$. The multiple $2m$ is needed here to carry out the block algebra calculations, with $2m \times 2m$ blocks. The projection of $\mathbb{M}$ on the space spanned by the leading eigenvectors is $\boldsymbol{\Lambda} = \boldsymbol{Y}^T\mathbb{M}\boldsymbol{Y}$, where $\boldsymbol{Y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_{2rm})$. The matrix $\boldsymbol{\Lambda}$ is positive definite, with spectrum that is weakly affected by noise. However, $\boldsymbol{\Lambda}$ is diagonal, while the mass matrix should have block Hankel plus Toeplitz structure (recall (46)), in order to get a ROM that preserves causality at the algebraic level. This causality is manifested in the block tridiagonal structure of the ROM propagator matrix. To obtain the desired $\mathbb{M}^{\text{reg}}$, we compute first

$$\boldsymbol{\Pi} = \boldsymbol{\Lambda}^{-1/2}\boldsymbol{Y}^T\mathbb{S}\boldsymbol{Y}\boldsymbol{\Lambda}^{-1/2}, \tag{60}$$

and then use the block Lanczos algorithm [29] to obtain the orthogonal matrix $\mathbb{Q} \in \mathbb{R}^{2rm \times 2rm}$ that gives a block tridiagonal $\mathbb{Q}^T\boldsymbol{\Pi}\mathbb{Q}$. This is the regularized ROM propagator. The regularized mass matrix is

$$\mathbb{M}^{\text{reg}} = \mathbb{Q}^T\boldsymbol{\Lambda}\mathbb{Q}. \tag{61}$$

3. Estimation of the internal wave

It is known that knowledge of the internal wave field i.e., the snapshots $\boldsymbol{u}_j$ at all points in Ω, would simplify considerably the inverse problem. This is the whole point of multi-physics "hybrid" imaging modalities like photo-acoustic tomography, transient elastography, etc. [42,2]. These modalities are limited to medical applications because they involve delicate and accurate apparatus for transmitting several types of waves and measuring all around the body. It was shown recently in [15,16], in the context of inverse scattering with acoustic waves, that the ROM can be used to estimate the internal wave. Here we extend those results to the vectorial problem, governed by equation (29).

We begin in section 3.1 with the definition of the estimated snapshots $\boldsymbol{u}_j^{\text{est}}$, followed by a numerical illustration in section 3.2. Then, we describe in section 3.3 the relation between $\boldsymbol{u}_j^{\text{est}}$ and the dyadic Green's function. This relation sheds light on the approximation of $\boldsymbol{u}_j$ by $\boldsymbol{u}_j^{\text{est}}$ and we use it in the analysis of the imaging method introduced in section 4.

3.1. Estimated wave snapshots

The estimation of the snapshots $\boldsymbol{u}_j$ is based on equation (55), which has two factors: The ROM snapshots $\{\boldsymbol{u}_j^{\text{ROM}} = \boldsymbol{R}\boldsymbol{i}_j\}_{j=0}^{n-1}$, which are computed from the data, and the $2\times 2nm$ field $\boldsymbol{V}$ that stores the orthonormal basis of the space $\mathcal{S}$. This field cannot be computed because $\mathcal{S}$ is not known, so we estimate the snapshots using the basis computed with the guess wave speed $\underline{\underline{\widetilde{c}}}$,

$$\boldsymbol{u}_j^{\text{est}}(\boldsymbol{x};\underline{\underline{\widetilde{c}}}) = \boldsymbol{V}(\boldsymbol{x};\underline{\underline{\widetilde{c}}})\boldsymbol{u}_j^{\text{ROM}} = \boldsymbol{V}(\boldsymbol{x};\underline{\underline{\widetilde{c}}})\boldsymbol{R}\boldsymbol{i}_j, \quad j=0,\ldots,n-1, \quad \boldsymbol{x}\in\Omega. \tag{62}$$

We call the medium with wave speed $\underline{\underline{\widetilde{c}}}$ the "reference medium", but note that $\underline{\underline{\widetilde{c}}}$ will change during our iterative approach to inversion, introduced in section 5.

A very popular imaging method, known as reverse time migration in the geophysics community [5,6,48] or backprojection in radar [22,20], is based on the linearization of the forward map, known as the Born approximation. Imaging is carried out by back-propagating the measurements to the imaging points in Ω, using the internal wave computed in the reference medium. The snapshots of this wave satisfy

$$\boldsymbol{u}_j(\boldsymbol{x};\underline{\underline{\widetilde{c}}}) = \boldsymbol{V}(\boldsymbol{x};\underline{\underline{\widetilde{c}}})\boldsymbol{R}(\underline{\underline{\widetilde{c}}})\boldsymbol{i}_j, \quad j=0,\ldots,n-1, \quad \boldsymbol{x}\in\Omega. \tag{63}$$

Equation (63) may look similar to (62), but there is a big difference: The matrix $\boldsymbol{R}$, which stores the first n ROM snapshots, is for the true and unknown medium, whereas in (63) we have $\boldsymbol{R}(\underline{\underline{\widetilde{c}}})$ computed in the reference medium. This difference is important, because the estimated snapshots (62) are consistent with the measurements, whereas those in (63) are not. Indeed, if we substitute the snapshots $\boldsymbol{u}_j$ in equations (57)–(59) with the estimated ones in (62), we obtain that the data interpolation still holds

$$\int_\Omega d\boldsymbol{x}\left[\boldsymbol{u}_0^{\text{est}}(\boldsymbol{x};\underline{\underline{\widetilde{c}}})\right]^T \boldsymbol{u}_j^{\text{est}}(\boldsymbol{x};\underline{\underline{\widetilde{c}}}) = (\boldsymbol{u}_0^{\text{ROM}})^T \underbrace{\int_\Omega d\boldsymbol{x}\boldsymbol{V}^T(\boldsymbol{x};\underline{\underline{\widetilde{c}}})\boldsymbol{V}(\boldsymbol{x};\underline{\underline{\widetilde{c}}})}_{\boldsymbol{I}_{2nm}}\boldsymbol{u}_j^{\text{ROM}} \overset{(57)}{=} \mathbb{D}(t_j),$$

for $j=0,\ldots,n-1$. Moreover, we have

$$\int_\Omega d\boldsymbol{x}\left[\boldsymbol{u}_{n-1}^{\text{est}}(\boldsymbol{x};\underline{\underline{\widetilde{c}}})\right]^T \boldsymbol{u}_j^{\text{est}}(\boldsymbol{x};\underline{\underline{\widetilde{c}}}) = (\boldsymbol{u}_{n-1}^{\text{ROM}})^T \underbrace{\int_\Omega d\boldsymbol{x}\boldsymbol{V}^T(\boldsymbol{x};\underline{\underline{\widetilde{c}}})\boldsymbol{V}(\boldsymbol{x};\underline{\underline{\widetilde{c}}})}_{\boldsymbol{I}_{2nm}}\boldsymbol{u}_j^{\text{ROM}} = \mathbb{M}_{n-1,j},$$

which in light of (58) gives the interpolation of $\mathbb{D}(t_{n-1+j})$, for $j=1,\ldots,n-1$. If we replace $\boldsymbol{u}_j^{\text{est}}$ by (63) in the left hand side of these equations, we deduce the interpolation of $\mathbb{D}(t;\underline{\underline{\widetilde{c}}})$, not $\mathbb{D}(t)$.

3.2. Numerical illustration of the estimated internal wave

The exact data fit equations above show that all the information about the waves recorded at the array is contained in $\boldsymbol{R}$. The orthonormal basis stored in $\boldsymbol{V}$ plays no role in the data interpolation. Its purpose is to map the information in $\boldsymbol{R}$ from the algebraic space, to the physical space, at points $\boldsymbol{x}\in\Omega$. For acoustic waves, extensive numerical simulations [15,14,24,10,23] and analysis in some special media [13, Appendix A] have shown that the orthonormal basis stored in $\boldsymbol{V}$ depends mostly on the smooth part of the medium, which controls the kinematics of wave propagator. The analysis and conclusion in [13, Appendix A] extend to the vector valued wave field $\boldsymbol{u}$. For brevity, we do not include it here. Instead, we show the results of a numerical simulation that illustrate the typical behavior of the estimated internal wave.

In Fig. 2 we compare the components along $\boldsymbol{e}_2$ of the true wave $\boldsymbol{u}_j$, the estimated wave $\boldsymbol{u}_j^{\text{est}}(\cdot;\underline{\underline{\widetilde{c}}})$ and the wave $\boldsymbol{u}_j(\cdot;\underline{\underline{\widetilde{c}}})$ calculated at the reference wave speed $\underline{\underline{\widetilde{c}}} = c_o\underline{\underline{\boldsymbol{I}}}$, at the point $\boldsymbol{x}$ indicated with a red cross. The true medium is isotropic, with wave speed $c\underline{\underline{\boldsymbol{I}}}$ displayed in the top left plot. The excitation is along $\boldsymbol{e}_2$. We indicate with arrows a few arrival events: The event denoted by a is the direct arrival from the source. The event denoted by b is the wave that traveled from the source to the boundary, scattered there and then arrived at $\boldsymbol{x}$. These two events are seen in all the plots. The events c and d are arrivals of waves that scattered in the medium. These are absent in the reference snapshots $\boldsymbol{u}_j(\cdot;\underline{\underline{\widetilde{c}}})$ but are seen in both the true wave snapshots and the estimated ones. Since in this example the kinematics is perturbed slightly, the estimated snapshots are a good approximation of the true ones.

3.3. Estimated internal wave in terms of the dyadic Green's function

The causal dyadic Green's function for the wave equation (7) is the 2×2 matrix valued solution $\underline{\underline{\boldsymbol{G}}}$ of

$$-\nabla^\perp\left[\nabla^\perp\cdot\underline{\underline{\boldsymbol{G}}}(t,\boldsymbol{x},\boldsymbol{y})\right] + \underline{\underline{c}}^{-2}(\boldsymbol{x})\partial_t^2\underline{\underline{\boldsymbol{G}}}(t,\boldsymbol{x},\boldsymbol{y}) = \delta(t)\underline{\underline{\boldsymbol{I}}}\delta(\boldsymbol{x}-\boldsymbol{y}), \tag{64}$$

for $t\in\mathbb{R}$ and $\boldsymbol{x}\in\Omega$, with initial and boundary conditions

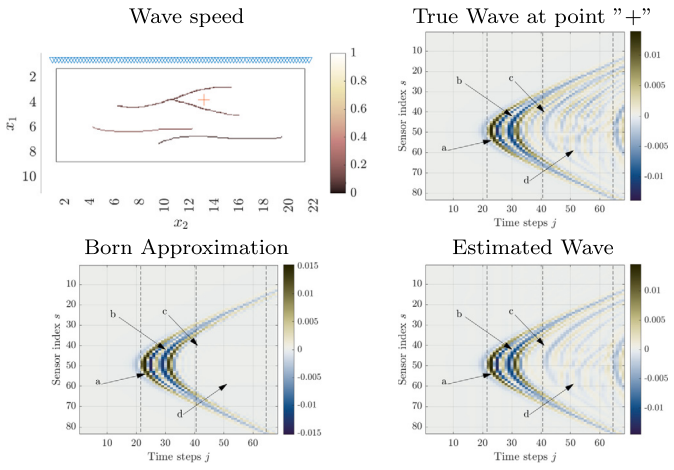

Fig. 2. Top, from left to right: Plot of $c(\boldsymbol{x})/c_o$ and the wave $\boldsymbol{e}_2^T\boldsymbol{u}_j(\boldsymbol{x})$ at the point $\boldsymbol{x}$ indicated with a red cross. Bottom, from left to right: The waves $\boldsymbol{e}_2^T\boldsymbol{u}_j(\boldsymbol{x}, c_o\underline{\underline{\boldsymbol{I}}})$ and $\boldsymbol{e}_2^T\boldsymbol{u}_j^{\rm est}(\boldsymbol{x}, c_o\underline{\underline{\boldsymbol{I}}})$. The source is polarized along $\boldsymbol{e}_2$. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

$$\underline{\underline{\boldsymbol{G}}}(t,\boldsymbol{x},\boldsymbol{y}) \equiv \underline{\underline{\boldsymbol{0}}}, \quad t<0, \quad \boldsymbol{x}\in\Omega, \tag{65}$$

$$\boldsymbol{n}^{\perp}(\boldsymbol{x})\cdot\underline{\underline{\boldsymbol{G}}}(t,\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{0}, \quad t\in\mathbb{R}, \quad \boldsymbol{x}\in\partial\Omega, \tag{66}$$

where the operators are understood to act columnwise. In light of our transformation of the wave operator in (64) to $A+\partial_t^2$, and the fact that in the right hand side of (17) we have a time derivative, we also use $\underline{\underline{\boldsymbol{\mathcal{G}}}}$, satisfying

$$A\underline{\underline{\boldsymbol{\mathcal{G}}}}(t,\boldsymbol{x},\boldsymbol{y}) + \partial_t^2\underline{\underline{\boldsymbol{\mathcal{G}}}}(t,\boldsymbol{x},\boldsymbol{y}) = \delta'(t)\underline{\underline{\boldsymbol{I}}}\delta(\boldsymbol{x}-\boldsymbol{y}), \quad t\in\mathbb{R}, \quad \boldsymbol{x}\in\Omega, \tag{67}$$

$$\underline{\underline{\boldsymbol{\mathcal{G}}}}(t,\boldsymbol{x},\boldsymbol{y}) \equiv \underline{\underline{\boldsymbol{0}}}, \quad t<0, \quad \boldsymbol{x}\in\Omega, \tag{68}$$

$$\boldsymbol{n}^{\perp}\cdot\underline{\underline{\boldsymbol{\mathcal{G}}}}(t,\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{0}, \quad t\in\mathbb{R}, \quad \boldsymbol{x}\in\Omega. \tag{69}$$

The next lemma, proved in Appendix A, gives the connection between these dyadics:

Lemma 1. *The Green dyadics $\underline{\underline{\boldsymbol{\mathcal{G}}}}$ and $\underline{\underline{\boldsymbol{G}}}$ are related by the identity*

$$\partial_t\underline{\underline{\boldsymbol{G}}}(t,\boldsymbol{x},\boldsymbol{y}) = \underline{\underline{\boldsymbol{c}}}(\boldsymbol{x})\underline{\underline{\boldsymbol{\mathcal{G}}}}(t,\boldsymbol{x},\boldsymbol{y})\underline{\underline{\boldsymbol{c}}}(\boldsymbol{y}), \quad t\in\mathbb{R}, \quad \boldsymbol{x},\boldsymbol{y}\in\Omega. \tag{70}$$

The approximation of the forcing $\underline{\underline{\boldsymbol{I}}}\delta(\boldsymbol{x}-\boldsymbol{y})$ in our space $\mathcal{S}$, with orthonormal basis stored in $\boldsymbol{V}$, is

$$\underline{\underline{\boldsymbol{\delta}}}(\boldsymbol{x},\boldsymbol{y}) = \sum_{j=0}^{n-1}\boldsymbol{v}_j(\boldsymbol{x})\boldsymbol{v}_j^T(\boldsymbol{y}) = \boldsymbol{V}(\boldsymbol{x})\boldsymbol{V}^T(\boldsymbol{y}). \tag{71}$$

The next lemma, proved in Appendix B, shows that the columns of the wave snapshots are wave fields with source determined by (71). The components of the estimated internal wave (62) have a similar expression, except that the source is determined by

$$\underline{\underline{\boldsymbol{\delta}}}^{\rm est}(\boldsymbol{x},\boldsymbol{y};\widetilde{\underline{\underline{\boldsymbol{c}}}}) = \sum_{j=1}^{n-1}\boldsymbol{v}_j(\boldsymbol{x})\boldsymbol{v}_j^T(\boldsymbol{y};\widetilde{\underline{\underline{\boldsymbol{c}}}}) = \boldsymbol{V}(\boldsymbol{x})\boldsymbol{V}^T(\boldsymbol{y};\widetilde{\underline{\underline{\boldsymbol{c}}}}). \tag{72}$$

Lemma 2. *Introduce the 2×2 matrix valued fields*

$$\underline{\underline{\boldsymbol{g}}}(t,\boldsymbol{x},\boldsymbol{y}) = \cos(t\sqrt{\mathbb{A}})\left|\widehat{f}(\sqrt{\mathbb{A}})\right|\underline{\underline{\boldsymbol{\delta}}}(\boldsymbol{x},\boldsymbol{y}), \tag{73}$$

and

$$\underline{\underline{\boldsymbol{g}}}^{\rm est}(t,\boldsymbol{x},\boldsymbol{y};\widetilde{\underline{\underline{\boldsymbol{c}}}}) = \cos(t\sqrt{\mathbb{A}})\left|\widehat{f}(\sqrt{\mathbb{A}})\right|\underline{\underline{\boldsymbol{\delta}}}^{\rm est}(\boldsymbol{x},\boldsymbol{y};\widetilde{\underline{\underline{\boldsymbol{c}}}}). \tag{74}$$

Let $j=0,\ldots,n-1$. The wave snapshots can be written as

$$\boldsymbol{u}_j(\boldsymbol{y}) \approx \left(\underline{\underline{\boldsymbol{g}}}(t_j,\boldsymbol{x}_1,\boldsymbol{y}),\ldots,\underline{\underline{\boldsymbol{g}}}(t_j,\boldsymbol{x}_m,\boldsymbol{y})\right), \quad \boldsymbol{y}\in\Omega, \tag{75}$$

and the estimated snaphots satisfy

$$\boldsymbol{u}_j^{\rm est}(\boldsymbol{y};\underline{\underline{\widetilde{c}}}) \approx \left(\underline{\underline{\boldsymbol{g}}}^{\rm est}(t_j,\boldsymbol{x}_1,\boldsymbol{y};\underline{\underline{\widetilde{c}}}),\ldots,\underline{\underline{\boldsymbol{g}}}^{\rm est}(t_j,\boldsymbol{x}_m,\boldsymbol{y};\underline{\underline{\widetilde{c}}})\right), \quad \boldsymbol{y}\in\Omega. \tag{76}$$

The expression (75) of the wave snapshots says that its (s,p) columns are wave vector fields that originate from the vicinity of the location $\boldsymbol{x}_s$ of the $s^{\rm th}$ antenna, with polarization along $\boldsymbol{e}_p$, for $s=1,\ldots,m$ and $p=1,2$. Vicinity means within the support[3] of $\underline{\underline{\boldsymbol{\delta}}}$ defined in (71). The estimated snapshots (76) have a similar interpretation, except that their initial support is dictated by $\underline{\underline{\boldsymbol{\delta}}}^{\rm est}$ defined in (72). The support of $\underline{\underline{\boldsymbol{\delta}}}$ depends on how rich the approximation space $\mathcal{S}$ is, while the support of $\underline{\underline{\boldsymbol{\delta}}}^{\rm est}$ also depends on the approximation of $\boldsymbol{V}$ by $\boldsymbol{V}(\cdot;\underline{\underline{\widetilde{c}}})$. This approximation seems to be mostly dependent on the kinematics i.e., the smooth part of $\underline{\underline{c}}$.

We can now state the connection between the estimated internal wave and the Green dyadic. The proof is in appendix C and the result is used in the analysis of the imaging function introduced in the next section.

Proposition 1. *Let $\check{f}$ be the even time signal with Fourier transform $|\widehat{f}|$,*

$$\check{f}(t)=\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}|\widehat{f}(\omega)|\cos(\omega t). \tag{77}$$

Suppose that $\underline{\underline{\boldsymbol{\delta}}}^{\rm est}(\boldsymbol{x},\boldsymbol{y};\underline{\underline{\widetilde{c}}})$ defined in (72) peaks near $\boldsymbol{y}$. Then, the components of the estimated snapshots in equation (76), evaluated at $\boldsymbol{y}\in\Omega$ satisfying $|\boldsymbol{y}-\boldsymbol{x}_s|>c_o t_F$, are approximated by

$$\underline{\underline{\boldsymbol{g}}}^{\rm est}(t,\boldsymbol{x}_s,\boldsymbol{y};\underline{\underline{\widetilde{c}}}) \approx c_o^{-1}\check{f}'(t)\star_t\int_\Omega d\boldsymbol{\xi}\,\underline{\underline{\boldsymbol{G}}}(t,\boldsymbol{x}_s,\boldsymbol{\xi})\underline{\underline{c}}^{-1}(\boldsymbol{\xi})\underline{\underline{\boldsymbol{\delta}}}^{\rm est}(\boldsymbol{\xi},\boldsymbol{y};\underline{\underline{\widetilde{c}}}). \tag{78}$$

4. Imaging

The goal of imaging is to obtain a qualitative picture of the heterogeneous medium, which localizes reflective structures, modeled by the "rough features" of $\underline{\underline{\boldsymbol{\varepsilon}}}$ and thus $\underline{\underline{c}}$, like jump discontinuities [48]. The smooth part of the medium, which determines the kinematics of wave propagation, is assumed known. In this section we model this smooth part by the constant and isotropic reference wave speed $\underline{\underline{\widetilde{c}}}=c_o\underline{\underline{\boldsymbol{I}}}$. This is typical in applications like radar imaging or nondestructive evaluation [20,22], although in seismic imaging the smooth part is variable [5,48].

The most popular imaging method is reverse time migration [5], also known as backprojection [20]. We call it henceforth the "traditional method" and its imaging function is defined by

$$\boldsymbol{\mathcal{I}}^{\rm RTM}(\boldsymbol{y})=\sum_{s,s'=1}^{m}\sum_{p,p'=1}^{2}\boldsymbol{e}_{p'}^T\underline{\underline{\boldsymbol{G}}}(-t,\boldsymbol{x}_{s'},\boldsymbol{y};\underline{\underline{\widetilde{c}}})\star_t\underline{\underline{\boldsymbol{G}}}(-t,\boldsymbol{y},\boldsymbol{x}_s;\underline{\underline{\widetilde{c}}})\boldsymbol{e}_p\star_t\mathcal{W}^{(s',p'),(s,p)}(t)\Big|_{t=0}. \tag{79}$$

This definition is based on the linearization of the forward mapping about the reference medium. The linearization, known as the Born approximation, assumes that the waves propagate from the source at $\boldsymbol{x}_s$ to a point $\boldsymbol{y}\in\Omega$, where they scatter once, and then return to the array, at the receiver location $\boldsymbol{x}_{s'}$. The image is formed by back-propagating the recorded waves to search points $\boldsymbol{y}$ in the imaging domain $\Omega_{\rm im}\subset\Omega$. The backpropagation uses the time reversibility of the wave equation and it is modeled in (79) by the convolution with the time reversed Green's dyadic in the reference medium.

Traditional imaging works well if the medium is a slight perturbation of the reference one. This holds when the reflectivity is weak and/or the support of the unknown features is small. If this is true, then the imaging functions (79) peaks near the reflective structures. If the reference medium has the wrong kinematics, the images are unfocused and if the reflectivity is strong, the images display ghost features, which are multiple scattering artifacts.

We introduce next an imaging method that uses the estimated internal wave (62). As we have discussed in the previous section, this wave contains all the scattering events recorded at the array. However, these events are mapped to wrong locations in Ω if the kinematics is strongly perturbed. Thus, it is not surprising that the imaging is successful when the reference medium gives a good approximation of the kinematics. The interesting property of the new imaging function is that it mitigates multiple scattering artifacts.

4.1. Imaging with the estimated internal wave

Similar to equation (34), let us write the estimated internal wave (62) in terms of its components

$$\boldsymbol{u}_j^{\rm est}(\boldsymbol{y};\underline{\underline{\widetilde{c}}})=\left(\boldsymbol{u}_j^{{\rm est}(1,1)}(\boldsymbol{y};\underline{\underline{\widetilde{c}}}),\boldsymbol{u}_j^{{\rm est}(1,2)}(\boldsymbol{y};\underline{\underline{\widetilde{c}}}),\ldots,\boldsymbol{u}_j^{{\rm est}(m,2)}(\boldsymbol{y};\underline{\underline{\widetilde{c}}})\right). \tag{80}$$

[3] By support we mean the set of points $\boldsymbol{y}$ around $\boldsymbol{x}_s$, where $\underline{\underline{\boldsymbol{\delta}}}(\boldsymbol{x}_s,\boldsymbol{y})$ is significant.

The imaging function for incident polarization along the direction e_p and measurements along $e_{p'}$ is

$$\mathcal{I}^{(p',p)}(\boldsymbol{y}) = \sum_{j=0}^{n-1}\sum_{s=1}^{m}\left|e_{p'}^T \boldsymbol{u}_j^{\mathrm{est}(s,p)}(\boldsymbol{y};\underline{\underline{\widetilde{c}}})\right|^2, \qquad \boldsymbol{y}\in\Omega_{\mathrm{im}}\subset\Omega. \tag{81}$$

This is an extension of the imaging function introduced in [13] for acoustic waves. As we show next, its expression is connected to the time reversal point spread function [28], which describes mathematically the following experiment: Suppose that there is a point source at $\boldsymbol{y}\in\Omega$, which emits a wave. If we record this wave at the array, time reverse the recordings and re-emit the wave back into the medium, then the wave will focus near $\boldsymbol{y}$, due to the time reversibility of the wave equation. The sharpness of the focusing depends on the frequency content of the wave, the aperture of the array and the medium. The striking experiments in [28] show that the more scattering in the medium, the better the focusing.

The analysis below shows that our imaging function (81) is determined by the time reversed point spread function averaged around $\boldsymbol{y}$, in the support of $\underline{\underline{\boldsymbol{\delta}}}^{\mathrm{est}}$ defined in (72). Thus, it is designed to be sensitive to changes in the medium around $\boldsymbol{y}$.

4.1.1. Connection to the time reversal experiment

To analyze (81), we recall the result (76) in Lemma 2 and approximate the sum over j by a time integral

$$\mathcal{I}^{(p',p)}(\boldsymbol{y}) \approx \frac{1}{\tau}\int_0^{n\tau} dt \sum_{s=1}^{m}\left|e_{p'}^T \boldsymbol{g}^{\mathrm{est}}(t,\boldsymbol{x}_s,\boldsymbol{y})e_p\right|^2. \tag{82}$$

Now use the result (78) of Proposition 1 to write

$$\mathcal{I}^{(p',p)}(\boldsymbol{y}) \approx \int_\Omega d\boldsymbol{\xi}\int_\Omega d\boldsymbol{\xi}'\left[\boldsymbol{\eta}^{(p)}(\boldsymbol{\xi},\boldsymbol{y};\underline{\underline{\widetilde{c}}})\right]^T \underline{\underline{\boldsymbol{K}}}^{(p')}(\boldsymbol{\xi},\boldsymbol{\xi}')\boldsymbol{\eta}^{(p)}(\boldsymbol{\xi}',\boldsymbol{y};\underline{\underline{\widetilde{c}}}), \tag{83}$$

in terms of the vector

$$\boldsymbol{\eta}^{(p)}(\boldsymbol{\xi},\boldsymbol{y};\underline{\underline{\widetilde{c}}}) = \underline{\underline{c}}^{-1}(\boldsymbol{\xi})\underline{\underline{\boldsymbol{\delta}}}^{\mathrm{est}}(\boldsymbol{\xi},\boldsymbol{y};\underline{\underline{\widetilde{c}}})e_p, \tag{84}$$

and the matrix $\underline{\underline{\boldsymbol{K}}}^{(p')} = \left(K_{j,j'}^{(p')}\right)_{j,j'=1,2}$, with entries

$$\begin{aligned} K_{j,j'}^{(p')}(\boldsymbol{\xi},\boldsymbol{\xi}') &= \sum_{s=1}^{m}\int_0^{n\tau}\frac{dt}{c_o^2\tau}\,\check{f}'(t)\star_t G_{p',j}(t,\boldsymbol{x}_s,\boldsymbol{\xi})\check{f}'(t)\star_t G_{p',j'}(t,\boldsymbol{x}_s,\boldsymbol{\xi}') \\ &= \sum_{s=1}^{m}\int_0^{n\tau}\frac{dt}{c_o^2\tau}\,\check{f}'(t)\star_t G_{j,p'}(t,\boldsymbol{\xi},\boldsymbol{x}_s)\check{f}'(t)\star_t G_{p',j'}(t,\boldsymbol{x}_s,\boldsymbol{\xi}'). \end{aligned} \tag{85}$$

The second equality is due to the reciprocity of Green's dyadic [51,47]

$$\underline{\underline{\boldsymbol{G}}}(t,\boldsymbol{x}_s,\boldsymbol{\xi}) = \underline{\underline{\boldsymbol{G}}}^T(t,\boldsymbol{\xi},\boldsymbol{x}_s). \tag{86}$$

We relate next the expression (83) to the time reversal experiment, using the following two steps:

Step 1: Let $\boldsymbol{\psi}^{(p)}$ be the wave field satisfying the wave equation

$$-\nabla^\perp\left[\nabla^\perp\cdot\boldsymbol{\psi}^{(p)}(t,\boldsymbol{x},\boldsymbol{y})\right] + \underline{\underline{c}}^{-2}(\boldsymbol{x})\partial_t^2\boldsymbol{\psi}^{(p)}(t,\boldsymbol{x},\boldsymbol{y}) = \check{f}'(t)\boldsymbol{\eta}^{(p)}(\boldsymbol{x},\boldsymbol{y};\underline{\underline{\widetilde{c}}}),$$

for $t\in\mathbb{R}$, $\boldsymbol{x}\in\Omega$, with initial and boundary conditions

$$\begin{aligned} &\boldsymbol{\psi}^{(p)}(t,\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{0}, \quad t\ll 0,\ \boldsymbol{x}\in\Omega, \\ &\boldsymbol{n}^\perp(\boldsymbol{x})\cdot\boldsymbol{\psi}^{(p)}(t,\boldsymbol{x},\boldsymbol{y}) = 0 \quad t\in\mathbb{R},\ \boldsymbol{x}\in\partial\Omega. \end{aligned}$$

This wave can be written in terms of Green's dyadic $\underline{\underline{\boldsymbol{G}}}$, the solution of (64)-(66), using linear superposition

$$\boldsymbol{\psi}^{(p)}(t,\boldsymbol{x},\boldsymbol{y}) = \check{f}'(t)\star_t\int_\Omega d\boldsymbol{\xi}'\,\underline{\underline{\boldsymbol{G}}}(t,\boldsymbol{x},\boldsymbol{\xi}')\boldsymbol{\eta}^{(p)}(\boldsymbol{\xi}',\boldsymbol{y};\underline{\underline{\widetilde{c}}}).$$

We are interested in the p' component of this wave, evaluated at $\boldsymbol{x}=\boldsymbol{x}_s$. Substituting it in equations (83) and (85), we get

$$\begin{aligned} \mathcal{I}^{(p',p)}(\boldsymbol{y}) \approx &\int_\Omega d\boldsymbol{\xi}\left[\boldsymbol{\eta}^{(p)}(\boldsymbol{\xi},\boldsymbol{y};\underline{\underline{\widetilde{c}}})\right]^T \\ &\cdot\int_0^{n\tau}\frac{dt}{c_o^2\tau}\,\check{f}'(t)\star_t\sum_{s=1}^{m}\underline{\underline{\boldsymbol{G}}}(t,\boldsymbol{\xi},\boldsymbol{x}_s)e_{p'}\left[e_{p'}^T\boldsymbol{\psi}^{(p)}(t,\boldsymbol{x}_s,\boldsymbol{y})\right]. \end{aligned}$$

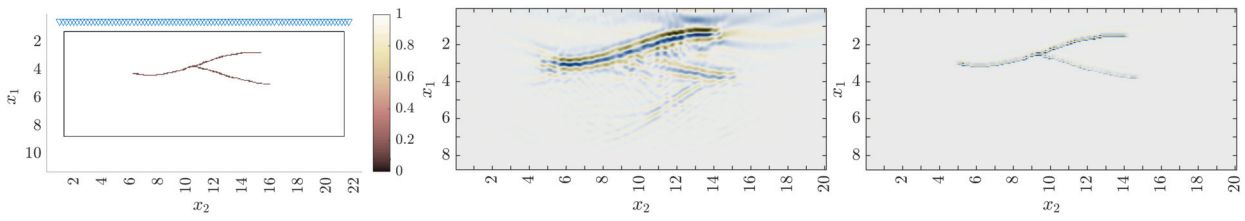

Fig. 3. Left: Display of the isotropic medium with a crack feature. The medium is modeled by the dielectric permittivity $\epsilon(\boldsymbol{x})\underline{\underline{\boldsymbol{I}}}$ and the colorbar shows the contrast $\epsilon(\boldsymbol{x})/\epsilon_o$. The antennas are drawn as triangles near the top boundary and the imaging domain $\Omega_{\rm im}$ is inside the black rectangle. Middle: The traditional imaging 80 antennas and $\tau = 0.3\pi/w_c$.

Step 2: Change variables in the time integral from t to $-t$ and recall that $\check{f}$ is even, which means that $\check{f}'$ is odd. Writing explicitly the convolution we get

$$\mathcal{I}^{(p',p)}(\boldsymbol{y}) \approx -\frac{1}{c_o^2\tau}\int_{\Omega} d\boldsymbol{\xi}\left[\boldsymbol{\eta}^{(p)}(\boldsymbol{\xi},\boldsymbol{y};\underline{\underline{\widetilde{c}}})\right]^T \boldsymbol{\beta}^{(p',p)}(\boldsymbol{\xi},\boldsymbol{y}), \tag{87}$$

in terms of the vector field

$$\boldsymbol{\beta}^{(p',p)}(\boldsymbol{\xi},\boldsymbol{y}) = \check{f}'(t') \star_t \int_{-n\tau}^{0} dt \sum_{s=1}^{m} \underline{\underline{\boldsymbol{G}}}(t'-t,\boldsymbol{\xi},\boldsymbol{x}_s)\boldsymbol{e}_{p'}\left[\boldsymbol{e}_{p'}^T\boldsymbol{\psi}^{(p)}(-t,\boldsymbol{x}_s,\boldsymbol{y})\right]\Big|_{t'=0}. \tag{88}$$

The right hand side in (88) is obtained by emitting from all the antennas, over the time interval $t \in [0, n\tau]$, the time reversed wave $\boldsymbol{e}_{p'}^T\boldsymbol{\psi}^{(p)}$, polarized along $\boldsymbol{e}_{p'}$. The result is then convolved with $\check{f}'$ and it is evaluated at $t = 0$, when we expect from the time reversibility of the wave equation the refocusing to occur, at points $\boldsymbol{\xi}$ near $\boldsymbol{y}$ [3].

4.1.2. Expected focusing of the imaging function

Recall definition (84) of $\boldsymbol{\eta}^{(p)}$ and write the expression (87) more explicitly

$$\mathcal{I}^{(p',p)}(\boldsymbol{y}) \approx -\frac{1}{c_o^2\tau}\int_{\Omega} d\boldsymbol{\xi}\,\boldsymbol{e}_p^T\left[\underline{\underline{\boldsymbol{\delta}}}^{\rm est}(\boldsymbol{\xi},\boldsymbol{y};\underline{\underline{\widetilde{c}}})\right]^T \underline{\underline{c}}^{-1}(\boldsymbol{\xi})\boldsymbol{\beta}^{(p',p)}(\boldsymbol{\xi},\boldsymbol{y}). \tag{89}$$

We know that the time reversal experiment gives a field $\boldsymbol{\beta}^{(p,p')}$ that is peaked at $\boldsymbol{\xi} \approx \boldsymbol{y}$, with resolution that depends on the signal $\check{f}$, the aperture of the array and the medium through which the waves propagate [3]. The shorter the signal and the larger the aperture, the better the refocusing. Our imaging function averages $\boldsymbol{\beta}^{(p,p')}$ in the support of $\underline{\underline{\boldsymbol{\delta}}}^{\rm est}(\cdot,\boldsymbol{y};\underline{\underline{\widetilde{c}}})\boldsymbol{e}_p$ and it is sensitive to the changes of the wave speed $\underline{\underline{c}}$ there. Thus, as long as the components in the $p^{\rm th}$ column of $\underline{\underline{\boldsymbol{\delta}}}^{\rm est}(\cdot,\boldsymbol{y};\underline{\underline{\widetilde{c}}})$ are peaked near $\boldsymbol{y}$, the imaging function computed from the internal wave as in equation (81) gives good estimates of the support of the reflectivity of the medium.

4.2. Numerical results

We refer to Appendix D for the setup of the numerical simulations and some comments on the computational cost. The simulations are carried out in a rectangular Ω with perfectly conducting $\partial\Omega$. The array of antennas is near the top boundary and the system of coordinates is $\boldsymbol{x} = (x_1, x_2)$, with the "range coordinate" x_1 pointing downwards and the "cross-range" coordinate x_2 orthogonal to it. The length unit in the plots is λ_c, the wavelength calculated at the highest frequency ω_c in the bandwidth of the probing signal f.

We show imaging results for isotropic and anisotropic media with crack like features, where the kinematics is slightly perturbed, but the reflectivity is strong enough to cause visible multiple scattering effects in the traditional imaging function (79). To explore the role of polarization in the measurements, we display the imaging function (81) for $p, p' = 1, 2$, instead of showing the sum over all polarizations. Note that we plot the range derivatives of the imaging functions, in order to emphasize the large changes which are related to the jump discontinuities of the wave speed. The array aperture, antenna separation and time step τ are given in the captions.

4.2.1. Imaging in isotropic media

The first example considers an isotropic medium, with dielectric permittivity displayed in the left plot of Fig. 3. The traditional imaging function (79) is shown in the middle plot. It does find the crack, but it has ghost features below it, induced by multiple scattering. The ideal image, the analogue of $\mathcal{I}^{(2,2)}$, defined for the uncomputable true internal wave, is shown for reference in the right plot. It has no ghost features and the resolution is excellent, as expected from the expression (89) with $\underline{\underline{\boldsymbol{\delta}}}^{\rm est}$ replaced by $\underline{\underline{\boldsymbol{\delta}}}$.

The computable ROM based imaging functions (81) are displayed in Fig. 4 for all $p, p' = 1, 2$. The images identify the crack better than the traditional approach (top middle plot in Fig. 3). The best image is $\mathcal{I}^{(2,2)}$. This is expected because we have a transversal wave that propagates mainly along the range direction $\boldsymbol{e}_1$. The dominant component of this wave is therefore along the cross-range axis $\boldsymbol{e}_2$.

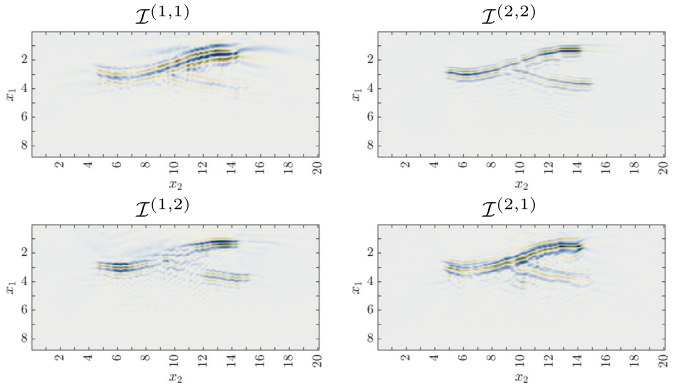

Fig. 4. Imaging function (81). The setup and axes are as in Fig. 4.

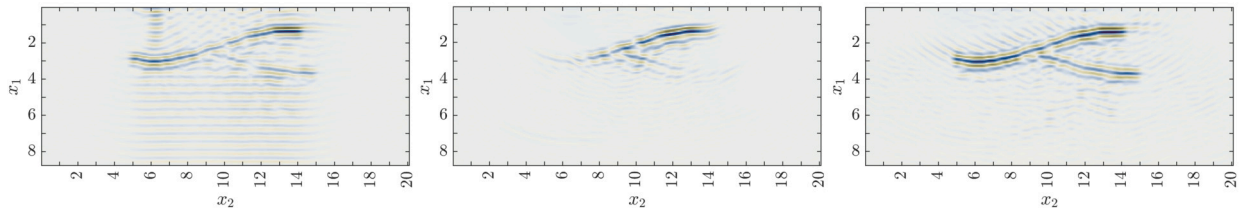

Fig. 5. Imaging function $\mathcal{I}^{(2,2)}$ for: $\tau = 0.9\pi/\omega_c$ (left), array aperture $5\lambda_c$ (middle) and sensor separation λ_c (right). The remaining parameters are as in Fig. 4. Compare the results displayed here with the top right plot in Fig. 4.

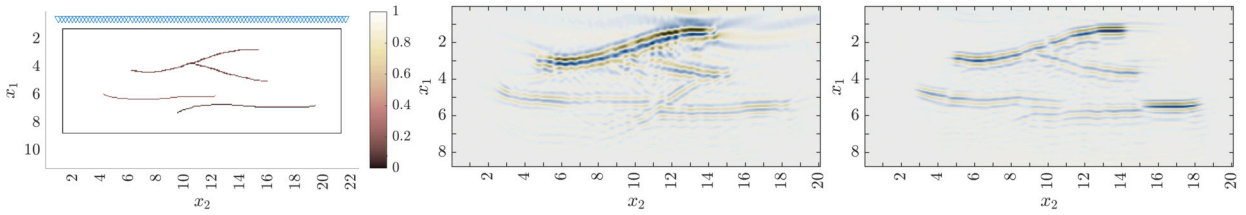

Fig. 6. Left: Isotropic medium with multiple cracks, modeled by the dielectric permittivity $\varepsilon(\boldsymbol{x})\underline{\underline{\boldsymbol{I}}}$. The colorbar shows the contrast $\varepsilon(\boldsymbol{x})/\varepsilon_o$. The antennas are drawn as triangles near the top boundary and the imaging domain is inside the black rectangle. Middle: The traditional imaging function. Right: The imaging function $\mathcal{I}^{(2,2)}$. The axes are in units of λ_c. The array aperture is $20\lambda_c$, there are 80 antennas and $\tau = 0.3\pi/w_c$.

We expect that, as is the case with any imaging method, a larger aperture should give a better cross-range resolution and improved view of the crack feature. We also expect that at the minimum, the time step τ should give two grid points in the period $2\pi/w_c$ of oscillation of the waves, and the antennas should be at least $\lambda_c/2$ apart, so they behave as an array. The results in Fig. 3-4 are obtained for a large aperture, a conservative τ that gives about 7 points per period of oscillation, and 80 antennas spaced by the distance $\lambda_c/4$. We illustrate in Fig. 5 the effect of less conservative choices. Note how undersampling in time and space causes unwanted ripples in the image. The reduced aperture leads to poor imaging of the ends of the crack.

The next example, shown in Fig. 6, is for a more complicated medium, with multiple cracks. Note the ghost feature induced by the multiple scattering in the traditional image (middle plot) that makes it difficult to tell if there is a single long crack or two disconnected ones. Our imaging function (right plot) is clearly superior.

4.2.2. *Imaging in anisotropic media*

Imaging methods are not designed to give quantitative estimates of the components of $\underline{\underline{\boldsymbol{c}}}$. The definition of the initial state (84) in the time reversal experiment shows that even if we have a jump in a single component of $\underline{\underline{\boldsymbol{c}}}$, that will be mapped to variations in $\mathcal{I}^{(p,p')}$ for all $p, p' = 1, 2$. Thus, we expect our images to indicate the location of all the jumps of $\underline{\underline{\boldsymbol{c}}}$. Indeed, this is the case, as illustrated in Fig. 8, for the medium shown in Fig. 7. We do not display the traditional image for this case, because it does not bring additional insight from what is shown in the figures above.

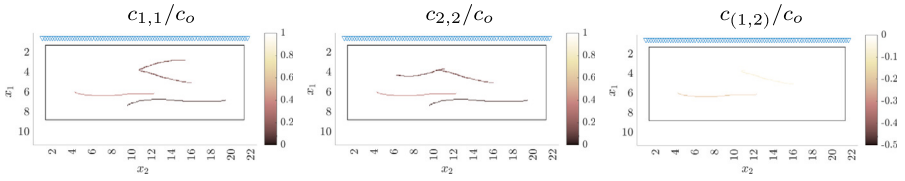

Fig. 7. The anisotropic medium. The colorbar is the same in all the plots.

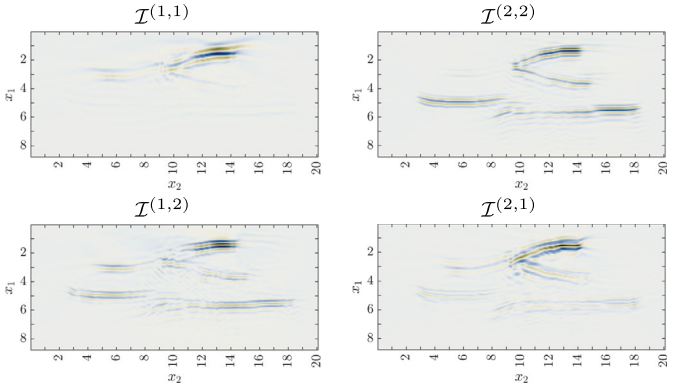

Fig. 8. Imaging function (81) for all polarization pairs. The setup and axes are as in Fig. 4 and the medium has the dielectric permittivity shown in Fig. 7.

5. Inversion

The numerical results in the previous section illustrate that for imaging purposes, it is enough to use a single polarization i.e., work with the imaging function $\mathcal{I}^{(2,2)}$. This is in agreement with the results in [45,18], which show that the shape of inclusions filled with orthotropic media can be recovered from the admittance map. However, since this map does not determine uniquely the anisotropic medium [44,35], for the quantitative estimation of the wave speed $\underline{\underline{c}}$ described in this section, we need the two polarizations.

The classic approach to isotropic or anisotropic medium estimation is via data fit optimization, as described in the introduction and in equation (14). In light of our data mapping described in section 2.2, we reformulate the traditional (FWI) approach as

$$\min_{\underline{\underline{\widetilde{c}}}\in\mathscr{C}} \mathcal{O}^{\text{FWI}}(\underline{\underline{\widetilde{c}}}) + \text{regularization}, \quad \mathcal{O}^{\text{FWI}}(\underline{\underline{\widetilde{c}}}) = \tau \sum_{j=0}^{2n-1} \left\| \mathbb{D}(j\tau) - \mathbb{D}(j\tau; \underline{\underline{\widetilde{c}}}) \right\|_F^2. \tag{90}$$

We introduce next a different inversion method, using our estimated internal wave (62). As explained in section 3.1, this wave fits the data by construction. However, due to the wrong orthonormal basis stored in $\boldsymbol{V}(\cdot; \underline{\underline{\widetilde{c}}})$, the wave (62) is not a solution of the wave equation. If it was a solution, then $\underline{\underline{\widetilde{c}}}$ would be close to the true wave speed, assuming uniqueness and stability of the inverse wave scattering problem, which holds for proper regularization and an appropriate search space $\mathscr{C}$.

The solution of the wave equation with wave speed $\underline{\underline{\widetilde{c}}}$ is given by (63). Thus, we can minimize the internal wave misfit

$$\int_\Omega d\boldsymbol{x} \sum_{j=0}^{n-1} \left\| \boldsymbol{u}_j^{\text{est}}(\boldsymbol{x}; \underline{\underline{\widetilde{c}}}) - \boldsymbol{u}_j(\boldsymbol{x}; \underline{\underline{\widetilde{c}}}) \right\|_F^2 \overset{(62),(63)}{=} \int_\Omega d\boldsymbol{x} \left\| \boldsymbol{V}(\boldsymbol{x}, \underline{\underline{\widetilde{c}}}) \left[\boldsymbol{R} - \boldsymbol{R}(\underline{\underline{\widetilde{c}}}) \right] \right\|_F^2$$

$$= \left\| \boldsymbol{R} - \boldsymbol{R}(\underline{\underline{\widetilde{c}}}) \right\|_F^2, \tag{91}$$

where the last equality is because the basis in $\boldsymbol{V}(\cdot; \underline{\underline{\widetilde{c}}})$ is orthonormal.

We estimate the wave speed using the optimization

$$\min_{\underline{\underline{\widetilde{c}}}\in\mathscr{C}} \mathcal{O}(\underline{\underline{\widetilde{c}}}) + \text{regularization}, \quad \mathcal{O}(\underline{\underline{\widetilde{c}}}) = \left\| \boldsymbol{R}(\underline{\underline{\widetilde{c}}}) \boldsymbol{R}^{-1} - \boldsymbol{I}_{2nm} \right\|_F^2, \tag{92}$$

where the objective function is a slight modification of (91), motivated by the following: It is known that variations of $\underline{\underline{\widetilde{c}}}$ that are far from the array can produce weak reflections that are masked from reflections from shallower depth. These weak reflections appear as small entries in the data matrices and, according to Theorem 1, the Gramian $\mathbb{M}$. These, in turn, result in small eigenvalues of the

square root $\boldsymbol{R}$ of $\mathbb{M}$. By using the inverse $\boldsymbol{R}^{-1}$ in the definition of the objective function,[4] we expect to recover the hidden reflections that are distinguishable from noise (recall the discussion in section 2.5) and thus image better.

5.1. The inversion algorithm

There are two user defined components of the inversion algorithm: The first is the choice of the search space $\mathscr{C}$ and the second is the choice of the regularization in the optimization (92).

We define $\mathscr{C}$ as follows: First, we parametrize the Cholesky square root of the squared inverse wave speed tensor

$$\mu_o \widetilde{\underline{\underline{\varepsilon}}}(\boldsymbol{x}) \overset{(9)}{=} \underline{\underline{c}}^{-2}(\boldsymbol{x}) = \underline{\underline{\boldsymbol{\gamma}}}^T(\boldsymbol{x})\underline{\underline{\boldsymbol{\gamma}}}(\boldsymbol{x}), \qquad \underline{\underline{\boldsymbol{\gamma}}}(\boldsymbol{x}) = \begin{pmatrix} \gamma_1(\boldsymbol{x}) & \gamma_3(\boldsymbol{x}) \\ 0 & \gamma_2(\boldsymbol{x}) \end{pmatrix}, \tag{93}$$

to ensure that $\widetilde{\underline{\underline{\varepsilon}}}$ remains symmetric and positive definite. The parametrization is done using the basis functions $\{\phi_j\}_{j=1}^N$ described in Appendix D,

$$\gamma_l(\boldsymbol{x}) = (1-\delta_{l,3})c_o^{-1} + \sum_{j=1}^{N} \alpha_{l,j}\phi_j(\boldsymbol{x}), \tag{94}$$

with $\boldsymbol{\alpha} = \left(\alpha_{1,1}, \ldots, \alpha_{3,N}\right) \in \mathbb{R}^{3N}$ to be determined by optimization.

We use the simplest Tikhonov type regularization penalty $\nu\|\boldsymbol{\alpha}\|^2$ in (92), with parameter ν chosen as explained in Appendix D. However, other choices of regularization, that incorporate prior information about $\underline{\underline{c}}$, can be used and would likely lead to sharper estimates.

Algorithm 1 *(Inversion algorithm).*
Input: Data matrices $\mathbb{D}(j\tau)$, for $j = 0, \ldots, 2n-1$.

1. Compute $\mathbb{M}$ with block entries given in equation (46). If $\mathbb{M}$ is indefinite or poorly conditioned, use regularization as described in section 2.5.
2. Compute the block Cholesky square root $\boldsymbol{R}$ of $\mathbb{M}$.
3. Starting with $\boldsymbol{\alpha}^{(0)} = \boldsymbol{0}$ proceed:

- For update index $j \geq 1$ calculate $\underline{\underline{\boldsymbol{\gamma}}}(\boldsymbol{x})$ as in equation (94), with $\boldsymbol{\alpha} = \boldsymbol{\alpha}^{(j-1)}$. Compute $\widetilde{\underline{\underline{c}}}$ from equation (93).
- Calculate $\boldsymbol{R}(\widetilde{\underline{\underline{c}}})$ following the same procedure of calculating $\boldsymbol{R}$.
- Compute $\boldsymbol{\alpha}^{(j)}$ as a Gauss-Newton update for minimizing the objective function (92) with the user's choice of the regularization parameter ν.
- Go to the next iteration or stop when the user defined convergence criterion has been met.

Output: The estimate of $\underline{\underline{c}}$ given by (93)-(94) with $\boldsymbol{\alpha}$ calculated at step 3.

Note that since the ROM construction is causal, the inversion can be carried out in a layer peeling fashion, by time windowing the data. If we use measurements up to time instant $t_{n'}$, for $n' < n$, then the ROM contains information up to the distance of order $c_o t_{n'}/2$ from the array. Thus, we can estimate the medium near the array, and then increase n' to obtain estimates at further distance. This can be very useful in speeding up the optimization.

5.2. Numerical results

We illustrate the performance of the inversion algorithm for the medium shown in Fig. 9. There are three anisotropic inclusions, modeled by the piecewise constant wave speed $\underline{\underline{c}}$ with components plotted in the figure.

The inversion results are shown in Fig. 10. The ROM approach (top plots) recovers well the shape of the inclusions, especially the left and middle ones. The components of $\underline{\underline{c}}$ there are also well estimated. There is some leakage in the off-diagonal component of $\underline{\underline{c}}$ for the right, thin inclusion. This inclusion is very difficult to determine, because the data recorded at the array is dominated by the reflection at its top corner.

As is typical of FWI inversion [54], the results shown in the bottom plots of Fig. 10 do not determine the shape of the inclusions and the wave speed there. For instance, FWI sees the top and bottom of the leftmost inclusion, which cause two strong reflections registered at the array. However, the bottom is misplaced because the wave speed inside the inclusion is not correctly identified.

Both methods tend to overshoot the value of the wave speed near the edges of the inclusions. This is a normal Gibbs phenomenon seen with Tikhonov regularization and it can be mitigated with a different regularization that is better suited for a piecewise constant $\underline{\underline{c}}$.

[4] Once we regularize the ROM construction, the mass matrix and its block Cholesky square root $\boldsymbol{R}$ are well conditioned, so inverting $\boldsymbol{R}$ is not an issue.

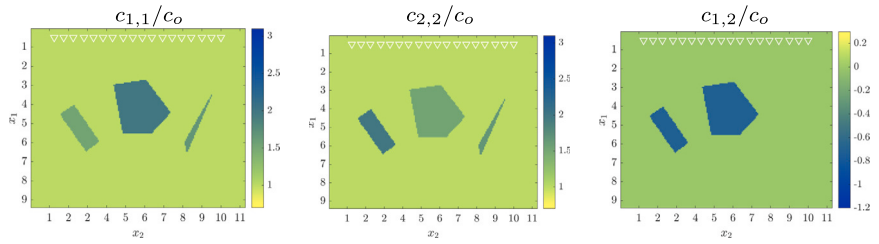

Fig. 9. The anisotropic medium with three inclusions. The axes are in units of λ_c. The contour of the inclusions is superposed on these plots.

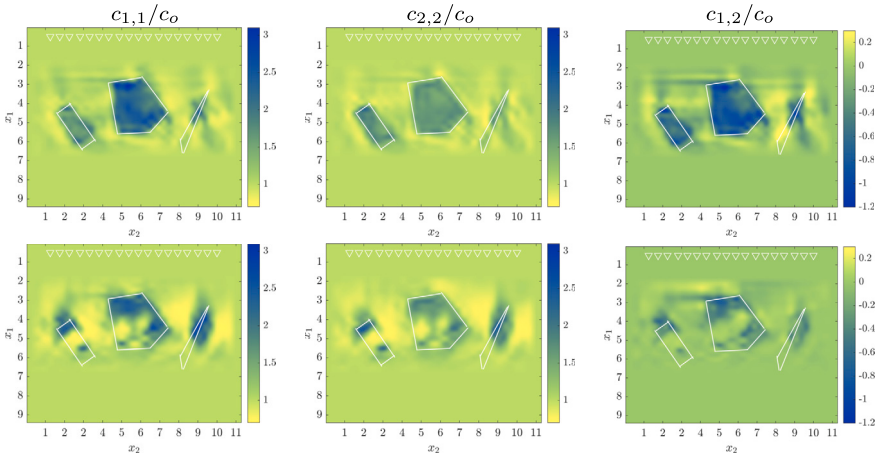

Fig. 10. The inversion results for the medium plotted in Fig. 9. Top row: results obtained with Algorithm 1. Bottom row: Results given by the FWI approach. The axes are in units of λ_c. The contour of the inclusions is superposed on these plots.

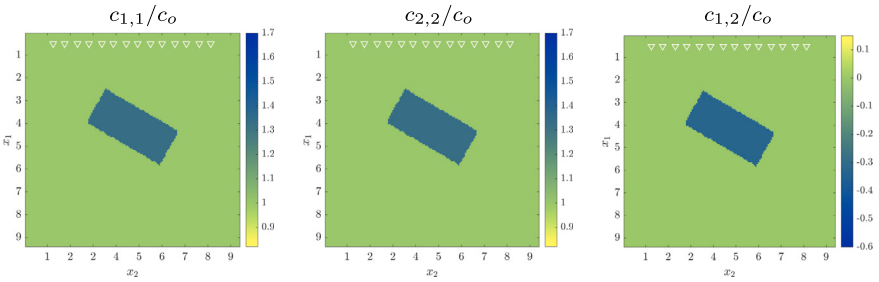

Fig. 11. Anisotropic medium with rectangle shaped inclusion. The axes are in units of λ_c. The contour of the inclusion is superposed on these plots.

6. Noisy data and imaging with the estimated kinematics

In this section, we explain how one may combine the imaging and inversion methods described in sections 4.1 and 5.1. This may be useful for distinguishing better the contour of inclusions buried in the medium, if the regularization of the inversion is not designed to do so. We also illustrate the performance of the inversion for noisy data.

We use for illustration the example in Fig. 11, where the medium contains a rectangular inclusion filled with an anisotropic medium. The inclusion is large enough to affect significantly the kinematics. Consequently, our imaging function (81), displayed in the left plot of Fig. 12, gives the wrong estimate of the bottom right corner of the inclusion. The traditional image shown in the middle plot behaves similarly, although it shows a ghost feature at the bottom of the domain, due to multiple scattering. It also has spurious oscillations near the top corner of the inclusion.

To correct the image in Fig. 12, we need a better estimate of the kinematics. One way to get such an estimate is to use the result of the inversion with Algorithm 1. The result of this inversion is shown in the top row of plots in Fig. 13. For comparison, we also show in the bottom row of plots the result of the traditional (FWI) inversion. As we saw before, FWI identifies the top of the inclusion, but it does not give a good approximation of the smooth part of the wave speed, which determines the kinematics.

If we compute the imaging function (81) using the reference wave speed obtained by our inversion Algorithm 1 (top plots in Fig. 13), we obtain the result shown in the right plot of Fig. 12. This is an improvement because it identifies correctly the entire

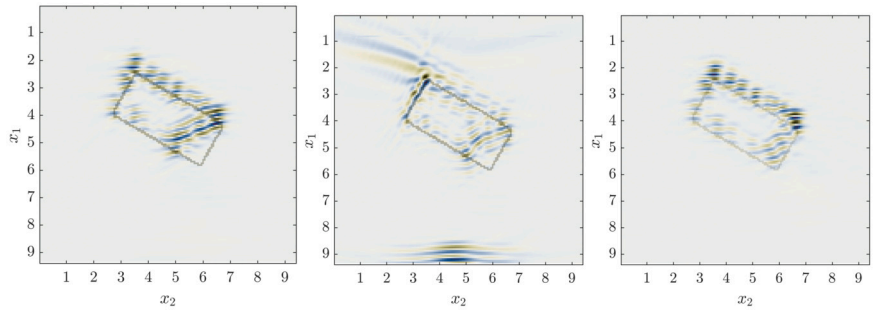

Fig. 12. First two plots: images of the medium in Fig. 11, calculated using the constant reference speed $c_o\underline{\underline{\boldsymbol{I}}}$. Left plot: the range derivative of the imaging function $\mathcal{I}^{(2,2)}$. Middle plot the traditional imaging function. The right plot: the imaging function $\mathcal{I}^{(2,2)}$ computed in the medium estimated by our inversion approach. The axes are in units of λ_c. The contour of the inclusion is superposed on these plots.

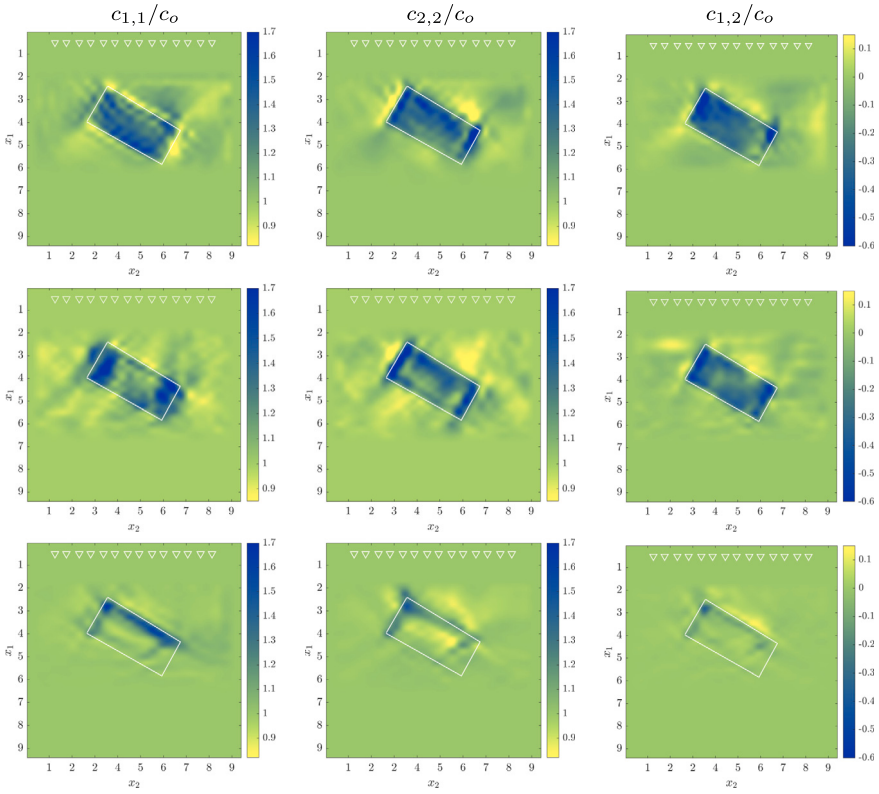

Fig. 13. The inversion results for the medium plotted in Fig. 11. Top and middle rows: results obtained with Algorithm 1. Noiseless data (top row) and noisy data (middle row). Bottom row: FWI results for noiseless data. The axes are in units of λ_c. The contour of the inclusion is superposed on these plots.

contour of the inclusion. It also complements the inversion results in Fig. 13, which do not give a precise estimation of this contour and also display some small oscillations that are typical of the Tikhonov regularization procedure.

To illustrate that our inversion algorithm works with noisy data, we plot in the middle row of Fig. 13 the results for 1% Gaussian, uncorrelated noise added to the matrices $\boldsymbol{D}_j$, for $j = 1, \ldots, 2n-1$. The matrix $\boldsymbol{D}_0$ is noiseless because it can be computed using the known medium near the array of antennas. The regularization of the ROM is via "option 2" described in section 2.5, where we project the noisy mass matrix $\mathbb{M}$ onto the space spanned by its $2rm$ leading eigenvectors of $\mathbb{M}$. The threshold $r = n\left\lfloor \frac{p}{2nm} \right\rfloor - 1$ is determined using the index p of the smallest, non-negative eigenvalue of $\mathbb{M}$. The eigenvalues are assumed sorted in descending order, as explained in section 2.5.

7. Summary

We introduced a novel approach for inverse scattering with electromagnetic waves in lossless, anisotropic heterogeneous media. It is an extension of results obtained recently for the acoustic wave equation in lossless, isotropic media. The approach uses a reduced

order model (ROM) which is an algebraic dynamical system that captures the propagation of the electric wave field inside the inaccessible medium, on a uniform time grid with appropriately chosen time step. The ROM is data driven, meaning that it can be computed just from measurements gathered by an active array of antennas that emit probing signals and measure the generated electric field. We showed how to use the ROM to obtain an estimate of the electric field inside the medium, called the internal wave. This wave is consistent by construction with the data measured at the antennas. We use it to formulate a novel imaging method, designed to locate the rough part of the medium, the reflectivity, in a known smooth background. The imaging function is easy and inexpensive to compute and it is superior to the traditional imaging approach because it does not display multiple scattering artifacts. We also introduced a novel inversion algorithm that uses the estimated internal wave to determine the matrix valued wave speed. Both methods are assessed with numerical simulations and are shown to perform better than the traditional approaches.

Our ROM methodology relies on having a self-adjoint wave operator. This holds in lossless media. If the medium is lossy i.e., there is a nonzero isotropic or anisotropic conductivity, the problem is very different and requires a new ROM methodology. First steps have been made in one-dimensional (layered) media in [12], but the higher dimensional case is open, so far.

The case of dispersive media also requires special treatment. In the time domain, dispersion introduces fractional time derivatives, which are difficult to handle. Thus, frequency domain ROMs are better suited for such media. Their construction differs from that of our time domain ROM, which uses the causality of wave propagation. An interesting case of orthotropic and dispersive media corresponds to having ferrite materials in a strong $z-$ directed magnetic field. Such materials are non-reciprocal and require a radically different ROM based approach to inversion and imaging.

CRediT authorship contribution statement

Liliana Borcea: Writing – original draft, Supervision, Project administration, Validation, Funding acquisition, Formal analysis, Methodology, Conceptualization. **Yiyang Liu:** Writing – review & editing, Visualization, Validation, Software, Formal analysis, Methodology, Investigation. **Jörn Zimmerling:** Writing – review & editing, Visualization, Validation, Supervision, Software, Formal analysis, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgements

This material is based upon research supported in part by the AFOSR award number FA9550-22-1-0077 and by the ONR award number N00014-21-1-2370. The computations were partially enabled by resources in project NAISS 2023/22-1157 and SNIC 2022/22-1060 provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS) at UPPMAX, funded by the Swedish Research Council through grant agreement no. 2022-06725. We thank Vladimir Druskin, Josselin Garnier, Alexander Mamonov and Mikhail Zaslavsky for the feedback and their work on data driven ROM imaging and inversion with acoustic waves.

Appendix A. Proof of Lemma 1

Take the time derivative in (64) and introduce the new 2×2 field

$$\underline{\underline{\boldsymbol{H}}}(t,\boldsymbol{x},\boldsymbol{y}) = \underline{\underline{\boldsymbol{c}}}^{-1}(\boldsymbol{x})\partial_t \underline{\underline{\boldsymbol{G}}}(t,\boldsymbol{x},\boldsymbol{y}), \tag{95}$$

which satisfies

$$\begin{aligned} -\nabla^\perp\big[\nabla^\perp\cdot(\underline{\underline{\boldsymbol{c}}}(\boldsymbol{x})\underline{\underline{\boldsymbol{H}}}(t,\boldsymbol{x},\boldsymbol{y}))\big] + \underline{\underline{\boldsymbol{c}}}^{-1}(\boldsymbol{x})\partial_t^2\underline{\underline{\boldsymbol{H}}}(t,\boldsymbol{x},\boldsymbol{y}) &= \delta'(t)\underline{\underline{\boldsymbol{I}}}\delta(\boldsymbol{x}-\boldsymbol{y}),\ t\in\mathbb{R},\ \boldsymbol{x}\in\Omega,\\ \underline{\underline{\boldsymbol{H}}}(t,\boldsymbol{x},\boldsymbol{y}) &\equiv \boldsymbol{0},\quad t<0,\quad \boldsymbol{x}\in\Omega,\\ \boldsymbol{n}^\perp(\boldsymbol{x})\cdot\underline{\underline{\boldsymbol{H}}}(t,\boldsymbol{x},\boldsymbol{y}) &= 0,\quad t\in\mathbb{R},\quad \boldsymbol{x}\in\partial\Omega. \end{aligned}$$

Multiply this equation on the left by $\underline{\underline{\boldsymbol{c}}}$ and recall the definition (20) of the operator A. The equation for $\underline{\underline{\boldsymbol{H}}}$ becomes

$$A\underline{\underline{\boldsymbol{H}}}(t,\boldsymbol{x},\boldsymbol{y}) + \partial_t^2\underline{\underline{\boldsymbol{H}}}(t,\boldsymbol{x},\boldsymbol{y}) = \delta'(t)\underline{\underline{\boldsymbol{c}}}(\boldsymbol{y})\delta(\boldsymbol{x}-\boldsymbol{y}),\quad t\in\mathbb{R},\quad \boldsymbol{x}\in\Omega.$$

It remains to show that

$$\underline{\underline{\boldsymbol{H}}}(t,\boldsymbol{x},\boldsymbol{y}) = \underline{\underline{\boldsymbol{\mathcal{G}}}}(t,\boldsymbol{x},\boldsymbol{y})\underline{\underline{\boldsymbol{c}}}(\boldsymbol{y}). \tag{96}$$

We check columnwise: For $p = 1, 2$ we have

$$(A + \partial_t^2)\left[\underline{\underline{\boldsymbol{G}}}(t, \boldsymbol{x}, \boldsymbol{y})\underline{\underline{\boldsymbol{c}}}(\boldsymbol{y})\boldsymbol{e}_p\right] = \sum_{j=1}^{2} c_{j,p}(\boldsymbol{y})(A + \partial_t^2)\left[\underline{\underline{\boldsymbol{G}}}(t, \boldsymbol{x}, \boldsymbol{y})\boldsymbol{e}_j\right],$$

and using equation (67) we get

$$(A + \partial_t^2)\left[\underline{\underline{\boldsymbol{G}}}(t, \boldsymbol{x}, \boldsymbol{y})\underline{\underline{\boldsymbol{c}}}(\boldsymbol{y})\boldsymbol{e}_p\right] = \sum_{j=1}^{2} c_{j,p}(\boldsymbol{y})\delta'(t)\boldsymbol{e}_j\delta(\boldsymbol{x} - \boldsymbol{y}) = \delta'(t)\underline{\underline{\boldsymbol{c}}}(\boldsymbol{y})\boldsymbol{e}_p\delta(\boldsymbol{x} - \boldsymbol{y}).$$

This shows that $\underline{\underline{\boldsymbol{G}}}(t, \boldsymbol{x}, \boldsymbol{y})\underline{\underline{\boldsymbol{c}}}(\boldsymbol{y})$ solves the same equation as $\underline{\underline{\boldsymbol{H}}}(t, \boldsymbol{x}, \boldsymbol{y})$. They both vanish at $t < 0$ and they satisfy the same homogeneous boundary condition, due to the assumption that $\underline{\underline{\boldsymbol{c}}} = c_o\underline{\underline{\boldsymbol{I}}}$ at $\partial\Omega$. Since the wave equation has a unique solution, the result (96) holds. □

Appendix B. Proof of Lemma 2

Introduce the $2m \times 2$ auxiliary fields

$$\boldsymbol{\sigma}_l(\boldsymbol{y}) = \boldsymbol{i}_l^T\boldsymbol{R}^{-1}\boldsymbol{V}^T(\boldsymbol{y}; \widetilde{\underline{\underline{\boldsymbol{c}}}}), \qquad l = 0, \ldots, n-1, \tag{97}$$

and observe that they satisfy

$$\sum_{l=0}^{n-1}\boldsymbol{i}_l\boldsymbol{\sigma}_l(\boldsymbol{y}) = \underbrace{\sum_{l=0}^{n-1}\boldsymbol{i}_l\boldsymbol{i}_l^T}_{\boldsymbol{I}_{2nm}}\boldsymbol{R}^{-1}\boldsymbol{V}^T(\boldsymbol{y}; \widetilde{\underline{\underline{\boldsymbol{c}}}}) = \boldsymbol{R}^{-1}\boldsymbol{V}^T(\boldsymbol{y}; \widetilde{\underline{\underline{\boldsymbol{c}}}}).$$

We deduce that

$$\begin{aligned}\sum_{l=0}^{n-1}\boldsymbol{u}_l(\boldsymbol{x})\boldsymbol{\sigma}_l(\boldsymbol{y}) &= \boldsymbol{\mathcal{U}}(\boldsymbol{x})\sum_{l=0}^{n-1}\boldsymbol{i}_l\boldsymbol{\sigma}_l(\boldsymbol{y}) = \boldsymbol{\mathcal{U}}(\boldsymbol{x})\boldsymbol{R}^{-1}\boldsymbol{V}^T(\boldsymbol{y}; \underline{\underline{\boldsymbol{c}}}) \\ &\overset{(41)}{=} \boldsymbol{V}(\boldsymbol{x})\boldsymbol{V}^T(\boldsymbol{y}; \widetilde{\underline{\underline{\boldsymbol{c}}}}) \overset{(72)}{=} \underline{\underline{\boldsymbol{\delta}}}^{\mathrm{est}}(\boldsymbol{x}, \boldsymbol{y}; \widetilde{\underline{\underline{\boldsymbol{c}}}}).\end{aligned} \tag{98}$$

Therefore,

$$\begin{aligned}\cos(t_j\sqrt{\mathbb{A}})\underline{\underline{\boldsymbol{\delta}}}^{\mathrm{est}}(\boldsymbol{x}, \boldsymbol{y}; \widetilde{\underline{\underline{\boldsymbol{c}}}}) &\overset{(98)}{=} \cos(t_j\sqrt{\mathbb{A}})\sum_{l=0}^{n-1}\boldsymbol{u}_l(\boldsymbol{x})\boldsymbol{\sigma}_l(\boldsymbol{y}) \\ &\overset{(34)}{=} \cos(t_j\sqrt{\mathbb{A}})\sum_{l=0}^{n-1}\cos(t_l\sqrt{\mathbb{A}})\boldsymbol{u}_0(\boldsymbol{x})\boldsymbol{\sigma}_l(\boldsymbol{y}) \\ &= \frac{1}{2}\sum_{l=0}^{n-1}\left[\cos(t_{j+l}\sqrt{\mathbb{A}}) + \cos(t_{|j-l|}\sqrt{\mathbb{A}})\right]\boldsymbol{u}_0(\boldsymbol{x})\boldsymbol{\sigma}_l(\boldsymbol{y}) \\ &\overset{(34)}{=} \frac{1}{2}\sum_{l=0}^{n-1}\left[\boldsymbol{u}_{j+l}(\boldsymbol{x}) + \boldsymbol{u}_{|j-l|}(\boldsymbol{x})\right]\boldsymbol{\sigma}_l(\boldsymbol{y}).\end{aligned} \tag{99}$$

With this result and recalling the expression (33) of the entries of the data matrices,

$$\begin{aligned}\int_\Omega d\boldsymbol{x}\,\boldsymbol{u}_0^T(\boldsymbol{x})\cos(t_j\sqrt{\mathbb{A}})\underline{\underline{\boldsymbol{\delta}}}^{\mathrm{est}}(\boldsymbol{x}, \boldsymbol{y}; \widetilde{\underline{\underline{\boldsymbol{c}}}}) &= \frac{1}{2}\sum_{l=0}^{n-1}\left[\mathbb{D}_{j+l} + \mathbb{D}_{|j-l|}\right]\boldsymbol{\sigma}_l(\boldsymbol{y}) \\ &\overset{(46)}{=} \sum_{l=0}^{n-1}\mathbb{M}_{j,l}\boldsymbol{\sigma}_l(\boldsymbol{y}),\end{aligned}$$

which is a $2m \times 2$ field. Moreover, using the definition (97) of $\boldsymbol{\sigma}_l$ and that

$$\sum_{l=0}^{n-1}\mathbb{M}_{j,l}\boldsymbol{i}_l^T = \sum_{l=0}^{n-1}\underbrace{\boldsymbol{i}_j^T\mathbb{M}\boldsymbol{i}_l}_{\boldsymbol{M}_{j,l}}\boldsymbol{i}_l^T = \boldsymbol{i}_j^T\mathbb{M},$$

we obtain that

$$\int_\Omega d\boldsymbol{x}\, \boldsymbol{u}_0^T(\boldsymbol{x})\cos(t_j\sqrt{\mathbb{A}})\underline{\underline{\boldsymbol{\delta}}}^{\rm est}(\boldsymbol{x},\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}) = \boldsymbol{i}_j^T\mathbb{M}\boldsymbol{R}^{-1}\boldsymbol{V}^T(\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}})$$

$$\overset{(48)}{=} \boldsymbol{i}_j^T\boldsymbol{R}^T\boldsymbol{V}^T(\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}) \overset{(62)}{=} \left[\boldsymbol{u}_j^{\rm est}(\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}})\right]^T. \tag{100}$$

Now, recall definition (31) of $\boldsymbol{u}_0$, which says that its (s,p) column is

$$\boldsymbol{u}_0^{(s,p)}(\boldsymbol{x}) = \left|\widehat{f}(\sqrt{\mathbb{A}})\right| \boldsymbol{e}_p F^{(s)}(\boldsymbol{x}).$$

Using this in (100) and recalling that $\mathbb{A}$ is symmetric and functions of $\mathbb{A}$ commute, we get that the (s,p) row in the left hand side of (100) is

$$\int_\Omega d\boldsymbol{x}\, \boldsymbol{e}_p^T F^{(s)}(\boldsymbol{x})\cos(t_j\sqrt{\mathbb{A}})\left|\widehat{f}(\sqrt{\mathbb{A}})\right|\underline{\underline{\boldsymbol{\delta}}}^{\rm est}(\boldsymbol{x},\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}) \overset{(74)}{=} \int_\Omega d\boldsymbol{x}\, \boldsymbol{e}_p^T F^{(s)}(\boldsymbol{x})\underline{\underline{\boldsymbol{g}}}^{\rm est}(t_j,\boldsymbol{x},\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}})$$

$$\approx \boldsymbol{e}_p^T\underline{\underline{\boldsymbol{g}}}^{\rm est}(t_j,\boldsymbol{x}_s,\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}). \tag{101}$$

Here the approximation is due to the assumption of point-like support of the antenna, modeled by $F^{(s)}$, and also that functions in the range of $|\widehat{f}(\sqrt{\mathbb{A}})|$ are smooth.

The result (76) follows from (100)-(101). The proof of (75) is the same, once we replace $\boldsymbol{V}(\cdot;\underline{\underline{\widetilde{\boldsymbol{c}}}})$ by the true $\boldsymbol{V}$. □

Appendix C. Proof of Proposition 1

Using the jump conditions induced by the derivative of the Dirac delta in (67), we can write equations (67)–(68) in the equivalent form

$$(A+\partial_t^2)\underline{\underline{\boldsymbol{\mathcal{G}}}}(t,\boldsymbol{x},\boldsymbol{y}) = \underline{\underline{\boldsymbol{0}}}, \qquad t>0,\quad \boldsymbol{x}\in\Omega,$$

$$\underline{\underline{\boldsymbol{\mathcal{G}}}}(0+,\boldsymbol{x},\boldsymbol{y}) = \underline{\underline{\boldsymbol{I}}}\delta(\boldsymbol{x}-\boldsymbol{y}),\quad \partial_t\underline{\underline{\boldsymbol{\mathcal{G}}}}(0,\boldsymbol{x};\boldsymbol{y}) = \underline{\underline{\boldsymbol{0}}}, \qquad \boldsymbol{x}\in\Omega.$$

By definition (74), the field $\underline{\underline{\boldsymbol{g}}}^{\rm est}$ satisfies the same equation as $\underline{\underline{\boldsymbol{\mathcal{G}}}}$, but has a different initial condition

$$(A+\partial_t^2)\underline{\underline{\boldsymbol{g}}}^{\rm est}(t,\boldsymbol{x},\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}) = \underline{\underline{\boldsymbol{0}}}, \qquad t>0,\quad \boldsymbol{x}\in\Omega,$$

$$\underline{\underline{\boldsymbol{g}}}^{\rm est}(0,\boldsymbol{x},\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}) = |\widehat{f}(\sqrt{\mathbb{A}})|\underline{\underline{\boldsymbol{\delta}}}^{\rm est}(\boldsymbol{x},\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}),\quad \partial_t\underline{\underline{\boldsymbol{g}}}^{\rm est}(0,\boldsymbol{x},\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}) = \boldsymbol{0}, \qquad \boldsymbol{x}\in\Omega.$$

Here we used that $\mathbb{A}$ is the restriction of A on $\boldsymbol{\mathcal{D}}\setminus{\rm null}(A)$ and also that the columns of $\underline{\underline{\boldsymbol{\delta}}}^{\rm est}$ do not have components in null(A). By the principle of linear superposition,

$$\underline{\underline{\boldsymbol{g}}}^{\rm est}(t,\boldsymbol{x},\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}) = \int_\Omega d\boldsymbol{\xi}\,\underline{\underline{\boldsymbol{\mathcal{G}}}}(t,\boldsymbol{x},\boldsymbol{\xi})|\widehat{f}(\sqrt{\mathbb{A}})|\underline{\underline{\boldsymbol{\delta}}}^{\rm est}(\boldsymbol{\xi},\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}). \tag{102}$$

Similar to what we have done in equation (21), we decompose $\underline{\underline{\boldsymbol{\mathcal{G}}}}$ as follows

$$\underline{\underline{\boldsymbol{\mathcal{G}}}}(t,\boldsymbol{x};\boldsymbol{y}) = \underline{\underline{\mathbb{G}}}(t,\boldsymbol{x};\boldsymbol{y}) + \underline{\underline{\boldsymbol{\Gamma}}}(t,\boldsymbol{x};\boldsymbol{y}). \tag{103}$$

The first term solves the same equation as $\underline{\underline{\boldsymbol{\mathcal{G}}}}$, but its initial state is the approximation of $\underline{\underline{\boldsymbol{I}}}\delta(\boldsymbol{x}-\boldsymbol{y})$ in the range of A and therefore $\mathbb{A}$,

$$\underline{\underline{\boldsymbol{\delta}}}^{\mathbb{A}}(\boldsymbol{x},\boldsymbol{y}) = \sum_{j=1}^{\infty}\boldsymbol{\varphi}_j(\boldsymbol{x})\boldsymbol{\varphi}_j^T(\boldsymbol{y}). \tag{104}$$

It can be written in compact form as

$$\underline{\underline{\mathbb{G}}}(t,\boldsymbol{x};\boldsymbol{y}) = 1_{[0,\infty)}(t)\cos(t\sqrt{\mathbb{A}})\underline{\underline{\boldsymbol{\delta}}}^{\mathbb{A}}(\boldsymbol{x},\boldsymbol{y}), \tag{105}$$

and has a series expansion in the eigenfunctions of $\mathbb{A}$, which form an orthonormal basis of range$(\mathbb{A})$. The second term in (103) has columns that lie in the null space of A. It solves the equation

$$\partial_t^2\underline{\underline{\boldsymbol{\Gamma}}}(t,\boldsymbol{x},\boldsymbol{y}) = \delta'(t)\left[\underline{\underline{\boldsymbol{I}}}\delta(\boldsymbol{x}-\boldsymbol{y}) - \underline{\underline{\boldsymbol{\delta}}}^{\mathbb{A}}(\boldsymbol{x},\boldsymbol{y})\right], \tag{106}$$

and it is given by

$$\underline{\underline{\boldsymbol{\Gamma}}}(t,\boldsymbol{x};\boldsymbol{y}) = 1_{[0,\infty)}(t)\left[\underline{\underline{\boldsymbol{I}}}\delta_{\boldsymbol{y}}(\boldsymbol{x}) - \underline{\underline{\boldsymbol{\delta}}}(\boldsymbol{x};\boldsymbol{y})\right]. \tag{107}$$

Note that since $\text{null}(A) = [\text{range}(\mathbb{A})]^{\perp}$, $\underline{\underline{\boldsymbol{\Gamma}}}$ has no contribution in (102) i.e.,

$$\underline{\underline{\boldsymbol{g}}}^{\text{est}}(t,\boldsymbol{x},\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}) = \int_{\Omega} d\boldsymbol{\xi}\, \underline{\underline{\mathbb{G}}}(t,\boldsymbol{x},\boldsymbol{\xi})|\widehat{f}(\sqrt{\mathbb{A}})|\underline{\underline{\boldsymbol{\delta}}}^{\text{est}}(\boldsymbol{\xi},\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}). \tag{108}$$

Using the expression (105) of $\underline{\underline{\mathbb{G}}}$ in equation (108) and recalling the spectral decomposition of $\mathbb{A}$, we get

$$\begin{aligned}
\underline{\underline{\boldsymbol{g}}}^{\text{est}}(t,\boldsymbol{x},\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}) &= \int_{\Omega} d\boldsymbol{y}\, \cos(t\sqrt{\mathbb{A}})\underline{\underline{\boldsymbol{\delta}}}^{\mathbb{A}}(\boldsymbol{x},\boldsymbol{\xi})|\widehat{f}(\sqrt{\mathbb{A}})|\underline{\underline{\boldsymbol{\delta}}}^{\text{est}}(\boldsymbol{\xi},\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}) \\
&= \sum_{q=1}^{\infty} \cos(t\sqrt{\theta_q})\boldsymbol{\varphi}_q(\boldsymbol{x}) \int_{\Omega} d\boldsymbol{\xi}\, \boldsymbol{\varphi}_q^T(\boldsymbol{\xi})|\widehat{f}(\sqrt{\mathbb{A}})|\underline{\underline{\boldsymbol{\delta}}}^{\text{est}}(\boldsymbol{\xi},\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}) \\
&= \sum_{q=1}^{\infty} \cos(t\sqrt{\theta_q})|\widehat{f}(\sqrt{\theta_q})|\boldsymbol{\varphi}_q(\boldsymbol{x}) \int_{\Omega} d\boldsymbol{\xi}\, \boldsymbol{\varphi}_q^T(\boldsymbol{\xi})\underline{\underline{\boldsymbol{\delta}}}^{\text{est}}(\boldsymbol{\xi},\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}).
\end{aligned} \tag{109}$$

Next, we use the expression (77) of the signal $\check{f}$ to get the convolution. First, we deduce that

$$\begin{aligned}
\int_{-\infty}^{\infty} dt\, \check{f}(t)\cos(\omega t) &\overset{(77)}{=} \int_{-\infty}^{\infty} \frac{d\omega'}{2\pi}|\widehat{f}(\omega')| \int_{-\infty}^{\infty} dt\, \cos(\omega' t)\cos(\omega t) \\
&= \int_{-\infty}^{\infty} \frac{d\omega'}{2}|\widehat{f}(\omega')|\left[\delta(\omega-\omega') + \delta(\omega+\omega')\right] \\
&= \frac{1}{2}\left[|\widehat{f}(\omega)| + |\widehat{f}(-\omega)|\right] = |\widehat{f}(\omega)|.
\end{aligned}$$

Therefore, we can write

$$\begin{aligned}
|\widehat{f}(\sqrt{\theta_q})|\cos(t\sqrt{\theta_q}) &= \int_{-\infty}^{\infty} dt'\, \check{f}(t')\cos(t'\sqrt{\theta_q})\cos(t\sqrt{\theta_q}) \\
&= \frac{1}{2}\int_{-\infty}^{\infty} dt'\, \check{f}(t')\cos[(t-t')\sqrt{\theta_q}] + \frac{1}{2}\int_{-\infty}^{\infty} dt'\, \check{f}(t')\cos[(t+t')\sqrt{\theta_q}].
\end{aligned}$$

Changing $t' \mapsto -t'$ in the last term and using that $\check{f}$ is even, we get

$$|\widehat{f}(\sqrt{\theta_q})|\cos(t\sqrt{\theta_q})| = \int_{-\infty}^{\infty} dt'\, \check{f}(t')\cos[(t-t')\sqrt{\theta_q}] = \check{f}(t) \star_t \cos\left(t\sqrt{\theta_q}\right). \tag{110}$$

Substituting (110) in (109) evaluated at $\boldsymbol{x} = \boldsymbol{x}_s$ and switching the convolution with the series (justified via the dominated convergence theorem), we get

$$\begin{aligned}
\underline{\underline{\boldsymbol{g}}}^{\text{est}}(t,\boldsymbol{x}_s,\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}) &= \check{f}(t) \star_t \sum_{q=1}^{\infty} \cos(t\sqrt{\theta_q})\boldsymbol{\varphi}_q(\boldsymbol{x}_s) \int_{\Omega} d\boldsymbol{\xi}\, \boldsymbol{\varphi}_q^T(\boldsymbol{\xi})\underline{\underline{\boldsymbol{\delta}}}^{\text{est}}(\boldsymbol{\xi},\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}) \\
&= \check{f}(t) \star_t \int_{\Omega} d\boldsymbol{\xi}\, \left[\cos(t\sqrt{\mathbb{A}})\underline{\underline{\boldsymbol{\delta}}}^{\mathbb{A}}(\boldsymbol{x}_s,\boldsymbol{\xi})\right]\underline{\underline{\boldsymbol{\delta}}}^{\text{est}}(\boldsymbol{\xi},\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}) \\
&\overset{(105)}{=} \check{f}(t) \star_t \int_{\Omega} d\boldsymbol{\xi}\, \underline{\underline{\mathbb{G}}}(t,\boldsymbol{x}_s;\boldsymbol{\xi})\underline{\underline{\boldsymbol{\delta}}}^{\text{est}}(\boldsymbol{\xi},\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}) \\
&= \check{f}(t) \star_t \int_{\Omega} d\boldsymbol{\xi}\, \underline{\underline{\boldsymbol{G}}}(t,\boldsymbol{x}_s;\boldsymbol{\xi})\underline{\underline{\boldsymbol{\delta}}}^{\text{est}}(\boldsymbol{\xi},\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}).
\end{aligned} \tag{111}$$

Here the second equality is by the definition of $\cos(t\sqrt{\mathbb{A}})$ and equation (104) and the last equality follows from the decomposition (103). The third equality requires more explanation. Indeed, when taking the convolution, the cosine will be evaluated at negative times if $t = O(t_F)$, so why can we use the causal Green's function? This is where we use the assumption that $\underline{\underline{\boldsymbol{\delta}}}^{\text{est}}(\boldsymbol{\xi},\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}})$ peaks at points $\boldsymbol{\xi}$ in the vicinity of $\boldsymbol{y}$ and that $\boldsymbol{x}_s$ is far from $\boldsymbol{y}$. By causality, we have $\underline{\underline{\boldsymbol{g}}}^{\text{est}}(t,\boldsymbol{x}_s,\boldsymbol{y};\underline{\underline{\widetilde{\boldsymbol{c}}}}) \approx 0$ for $t = O(t_F)$, so we are not interested in such early times.

To complete the proof, we use Lemma 1. We get

$$\underline{\underline{g}}^{\text{est}}(t,\boldsymbol{x}_s,\boldsymbol{y};\underline{\underline{\widetilde{c}}}) = \check{f}(t) \star_t \int_{\Omega} d\boldsymbol{\xi}\, \underline{\underline{c}}^{-1}(\boldsymbol{x}_s)\partial_t \underline{\underline{\boldsymbol{G}}}(t,\boldsymbol{x}_s,\boldsymbol{\xi})\underline{\underline{c}}^{-1}(\boldsymbol{\xi})\underline{\underline{\boldsymbol{\delta}}}^{\text{est}}(\boldsymbol{\xi},\boldsymbol{y};\underline{\underline{\widetilde{c}}}).$$

Switching the time derivative in the convolution to $\check{f}$ and recalling that $\underline{\underline{c}}(\boldsymbol{x}_s) = c_o \underline{\underline{\boldsymbol{I}}}$, we get the result. □

Appendix D. Numerical simulations

Here we describe briefly the setup of the numerical simulations and the computational cost.

D.1. Setup

To compute the data matrices (32), we solve the wave equation (29). We use a finite difference time domain numerical method [50], with discretization of the operator A on a Lebedev grid with step size ℓ. This consists of two Yee grids, shifted by $\ell/2$, on which we discretize the first and second component of the electric field, respectively [37,36,56,43]. The time derivative is discretized with a second order, centered difference scheme, on a uniform time grid with small step Δt satisfying the CFL condition. The time step τ used in the computation of the ROM is an integer multiple of Δt.

Computing the CFL condition in heterogeneous anisotropic media is difficult, because one needs to find the largest eigenvalue of a large sparse matrix. In practice, choosing the timestep as $\Delta t < \frac{1}{2}\frac{\ell}{c_{\max}}$, where $c_{\max}$ is the largest eigenvalue of the variable $\underline{\underline{c}}$, for all points in Ω, is sufficient to satisfy the CFL condition. In our simulations we pick first the time step τ and then chose Δt such that it satisfies this CFL condition and $\tau/\Delta t$ is an integer.

The initial condition (31) is computed using the spectral decomposition of the discretization of the operator A. Equations (16), (21), (23) and (27) and the finite wave speed imply that the initial condition depends only on the medium near the array. Thus, it can be computed using the discretization of A in a smaller domain around the array.

The wave equation (29) has homogeneous Dirichlet boundary conditions, corresponding to a perfectly conducting boundary $\partial\Omega$. This was used in the derivations to define functions of the operator A via the spectral decomposition. Our finite difference solution of the forward problem, with such boundary conditions, is implemented on a large enough domain Ω such that the top and bottom boundary reflections are not observed during the duration of measurements. For long durations, this increases the computational domain. Thus, it may be advantageous to use absorbing (PML) boundary conditions [4] to solve the forward problem. This should give the same data at the sensors as our approach, for large enough Ω and proper time gating.

The probing signal has the Fourier transform

$$\widehat{f}(\omega) = \frac{\omega^2}{2}\left[e^{-\frac{(\omega-\omega_o)^2}{2\omega_b^2}} + e^{-\frac{(\omega+\omega_o)^2}{2\omega_b^2}}\right], \tag{112}$$

with standard deviation ω_b chosen so that the highest frequency at -25 dB cut-off is $\omega_c = 5/3\omega_o$. The central wavelength is $\lambda_o = 26.7\ell$ and the wavelength at the cut-off frequency is $\lambda_c = 16\ell$.

The array of antennas is located at a distance 8ℓ from the top boundary. Unless stated otherwise, for imaging, the antennas are separated by the distance $\lambda_c/4$ and the data are sampled at time step $\tau = 0.3\pi/\omega_c$. For the inversion, we used the antenna separation $\lambda_c/2$ and the time step $\tau = 0.45\pi/\omega_c$.

The parametrization (93)-(94) of the search dielectric permittivity tensor is done using the Gaussian basis functions

$$\phi_j(\boldsymbol{x}) = \exp\left[-\frac{(x_1 - X_{1,j})^2}{2\sigma_1^2} - \frac{(x_2 - X_{2,j})^2}{2\sigma_2^2}\right], \tag{113}$$

where $\{(X_{1,j}, X_{2,j})\}_{j=1}^N$ is a rectangular lattice in the inversion domain, with spacing $\lambda_c/4$ along the x_1 axis and $5\lambda_c/16$ along the x_2 axis. The standard deviations are chosen as $\sigma_1 = 2.3\ell$ and $\sigma_2 = 2.9\ell$ to be able to represent a wide range of media and partition the search domain well.

The Tikhonov regularization parameter ν is chosen adaptively, based on the eigenvalues of the "Hessian" in the Gauss-Newton method. For N basis functions as in equation (94), we see from (93) that we have $3N$ unknowns. If $\{\lambda_j\}_{j=1}^{3N}$ are the eigenvalues of the Hessian, sorted in descending order, then in all our simulations we set $\nu = \lambda_{\text{round}(0.9N)}$.

D.2. Computational cost

We begin with the computational cost of solving the forward problem, and thus computing the data matrices $\mathbb{D}(t)$. If we have a mesh size with N_1 points in the x_1 direction and N_2 points in the x_2 direction, and we use $2m$ excitations, then at each time step we multiply a sparse matrix of size $4(N_1N_2)\times 4(N_1N_2)$ with a vector of size $4N_1N_2\times 2m$ at a cost of $O(mN_1N_2)$. The cost of computing $\mathbb{D}(t)$, at $t\in(0,T)$, with $T = N_t\Delta t$, is therefore, $O(mN_1N_2N_t)$.

The cost of each Gauss-Newton iteration for optimizing our objective function or the FWI objective function, is dominated by the computation of the Jacobian. There is an efficient way of computing the Jacobian for FWI, using the adjoint formula [54]. The extension of this formula to our objective function can be made, but we have not done so. With such a formula, the Jacobian

could be computed very efficiently. At the moment, the Jacobian is computed using finite differences and it is the bottleneck of our computations.

The extra cost of computing our objective function vs. FWI is due to the block Cholesky algorithm, which takes $O(m^3 n^3)$ operations.

If $mn \gg 1$, which is likely the case in three dimensions, the dominant computational cost should be in solving the normal equations for the Gauss-Newton updates. This can be handled using iterative methods. Moreover, one can reduce the computational cost by working with a collection of sub-arrays. This idea has been used for a different problem in [9].

Appendix. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jcp.2024.113272.

References

[1] J.F.P. Abascal, W.R. Lionheart, S.R. Arridge, M. Schweiger, D. Atkinson, D.S. Holder, Electrical impedance tomography in anisotropic media with known eigenvectors, Inverse Probl. 27 (6) (2011) 065004.
[2] G. Bal, Hybrid inverse problems and internal functionals, in: Inverse Problems and Applications: Inside Out, II, vol. 60, 2013, pp. 325–368.
[3] G. Bal, L. Ryzhik, Time reversal and refocusing in random media, SIAM J. Appl. Math. 63 (5) (2003) 1475–1498.
[4] J.P. Berenger, A perfectly matched layer for the absorption of electromagnetic waves, J. Comput. Phys. 114 (2) (1994) 185–200.
[5] B. Biondi, 3D Seismic Imaging, Society of Exploration Geophysicists, 2006.
[6] N. Bleistein, J. Stockwell, J. Cohen, Multidimensional Seismic Inversion, Springer, 2001.
[7] L. Borcea, F.G. del Cueto, G. Papanicolaou, C. Tsogka, Filtering deterministic layer effects in imaging, SIAM Rev. 54 (4) (2012) 757–798.
[8] L. Borcea, V. Druskin, A. Mamonov, M. Zaslavsky, J. Zimmerling, Reduced order model approach to inverse scattering, SIAM J. Imaging Sci. 13 (2) (2020) 685–723.
[9] L. Borcea, V. Druskin, A.V. Mamonov, M. Zaslavsky, A model reduction approach to numerical inversion for a parabolic partial differential equation, Inverse Probl. 30 (12) (2014) 125011.
[10] L. Borcea, V. Druskin, A.V. Mamonov, M. Zaslavsky, Untangling the nonlinearity in inverse scattering with data-driven reduced order models, Inverse Probl. 34 (2018) 065008, http://iopscience.iop.org/10.1088/1361-6420/aabb16.
[11] L. Borcea, V. Druskin, A.V. Mamonov, M. Zaslavsky, Robust nonlinear processing of active array data in inverse scattering via truncated reduced order models, J. Comput. Phys. 381 (2019) 1–26.
[12] L. Borcea, V. Druskin, J. Zimmerling, A reduced order model approach to inverse scattering in lossy layered media, J. Sci. Comput. 89 (1) (2021) 1.
[13] L. Borcea, J. Garnier, A. Mamonov, J. Zimmerling, Reduced order model approach for imaging with waves, Inverse Probl. 38 (2) (2021) 025004.
[14] L. Borcea, J. Garnier, A. Mamonov, J. Zimmerling, Waveform inversion via reduced order modeling, Geophysics 88 (2) (2022) 1–91.
[15] L. Borcea, J. Garnier, A. Mamonov, J. Zimmerling, Waveform inversion with a data driven estimate of the internal wave, SIAM J. Imaging Sci. 16 (1) (2023) 280–312.
[16] L. Borcea, J. Garnier, A.V. Mamonov, J. Zimmerling, When data driven reduced order modeling meets full waveform inversion, SIAM Rev. (2024), in press, preprint, arXiv:2302.05988, 2023.
[17] C. Bunks, F. Saleck, S. Zaleski, G. Chavent, Multiscale seismic waveform inversion, Geophysics 60 (5) (1995) 1457–1473.
[18] F. Cakoni, D. Colton, P. Monk, J. Sun, The inverse electromagnetic scattering problem for anisotropic media, Inverse Probl. 26 (7) (2010) 074004.
[19] Y. Chen, Inverse scattering via Heisenberg's uncertainty principle, Inverse Probl. 13 (2) (1997) 253.
[20] M. Cheney, B. Borden, Fundamentals of Radar Imaging, SIAM, 2009.
[21] S. Cloude, Polarisation: Applications in Remote Sensing, OUP, Oxford, 2009.
[22] J. Curlander, R. McDonough, Synthetic Aperture Radar, vol. 11, Wiley, 1991.
[23] V. Druskin, A.V. Mamonov, A.E. Thaler, M. Zaslavsky, Direct, nonlinear inversion algorithm for hyperbolic problems via projection-based model reduction, SIAM J. Imaging Sci. 9 (2) (2016) 684–747.
[24] V. Druskin, A.V. Mamonov, M. Zaslavsky, A nonlinear method for imaging with acoustic waves via reduced order model backprojection, SIAM J. Imaging Sci. 11 (1) (2018) 164–196.
[25] P. Elbau, L. Mindrinos, O. Scherzer, The inverse scattering problem for orthotropic media in polarization-sensitive optical coherence tomography, GEM Int. J. Geomath. 9 (2018) 145–165.
[26] B. Engquist, B. Froese, Application of the Wasserstein metric to seismic signals, Commun. Math. Sci. 12 (5) (2014) 979–988.
[27] B. Engquist, Y. Yang, Optimal transport based seismic inversion: beyond cycle skipping, Commun. Pure Appl. Math. 75 (10) (2022) 2201–2244.
[28] M. Fink, Time-reversed acoustics, Sci. Am. 281 (1999) 91–97.
[29] G.H. Golub, C.F. Van Loan, Matrix Computations, John Hopkins University Press, Baltimore, 2013.
[30] G.W. Hanson, A.B. Yakovlev, Operator Theory for Electromagnetics: an Introduction, Springer Science & Business Media, 2013.
[31] J. Heino, E. Somersalo, Estimation of optical absorption in anisotropic background, Inverse Probl. 18 (3) (2002) 559.
[32] F.J. Herrmann, U. Böniger, D.J. Verschuur, Non-linear primary-multiple separation with directional curvelet frames, Geophys. J. Int. 170 (2) (2007) 781–799.
[33] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, 2012.
[34] G. Huang, R. Nammour, W. Symes, Source-independent extended waveform inversion based on space-time source extension: frequency-domain implementation, Geophysics 83 (5) (2018) R449–R461.
[35] C.E. Kenig, M. Salo, G. Uhlmann, Inverse problems for the anisotropic Maxwell equations, Duke Math. J. 157 (2) (2011) 369–419.
[36] V.I. Lebedev, Difference analogues of orthogonal decompositions, basic differential operators and some boundary problems of mathematical physics, I, USSR Comput. Math. Math. Phys. 3 (4) (1964) 69–92.
[37] V.I. Lebedev, Difference analogues of orthogonal decompositions, basic differential operators and some boundary problems of mathematical physics. II, USSR Comput. Math. Math. Phys. 4 (4) (1964) 36–50.
[38] J.M. Lee, G. Uhlmann, Determining anisotropic real-analytic conductivities by boundary measurements, Commun. Pure Appl. Math. 42 (8) (1989) 1097–1112.
[39] T. van Leeuwen, F. Herrmann, Mitigating local minima in full-waveform inversion by expanding the search space, Geophys. J. Int. 195 (1) (2013) 661–667.
[40] R.C. McOwen, Partial Differential Equations: Methods and Applications, 2 edn, Pearson Education Inc., Upper Saddle River, NJ, 2003, 07458.
[41] P. Monk, Finite Element Methods for Maxwell's Equations, Oxford University Press, 2003.
[42] A. Nachman, A. Tamasan, A. Timonov, Recovering the conductivity from a single measurement of interior data, Inverse Probl. 25 (3) (2009) 035014.
[43] M. Nauta, M. Okoniewski, M. Potter, FDTD method on a Lebedev grid for anisotropic materials, https://doi.org/10.1109/TAP.2013.2247373, 2013.
[44] P. Ola, L. Päivärinta, E. Somersalo, An inverse boundary value problem in electrodynamics, Duke Math. J. 70 (1993) 617–653.
[45] M. Piana, On uniqueness for anisotropic inhomogeneous inverse scattering problems, Inverse Probl. 14 (6) (1998) 1565.
[46] R. Potthast, Electromagnetic scattering from an orthotropic medium, J. Integral Equ. Appl. (1999) 197–215.

[47] M. Strunc, Constitutive relations and conditions for reciprocity in bianisotropic media (macroscopic approach), WSEAS Trans. Electron 4 (2007) 208–212.
[48] W. Symes, Migration velocity analysis and waveform inversion, Geophys. Prospect. 56 (6) (2008) 765–790.
[49] W. Symes, Error bounds for extended source inversion applied to an acoustic transmission inverse problem, Inverse Probl. 38 (11) (2022) 115002.
[50] A. Taflove, S. Hagness, Computational Electrodynamics: The Finite-Difference Time-Domain Method, 3rd edition, vol. 2062, 2005.
[51] C.T. Tai, Complementary reciprocity theorems in electromagnetic theory, IEEE Trans. Antennas Propag. 40 (6) (1992) 675–681.
[52] S. Torquato, Random Heterogeneous Materials Microstructure and Macroscopic Properties, Springer, 2000.
[53] D. Verschuur, Seismic Multiple Removal Techniques: Past, Present and Future, EAGE Publications Houten, The Netherlands, 2013.
[54] J. Virieux, S. Operto, An overview of full-waveform inversion in exploration geophysics, Geophysics 74 (6) (2009) WCC1–WCC26.
[55] A. Weglein, F. Gasparotto, P. Carvalho, R. Stolt, An inverse-scattering series method for attenuating multiples in seismic reflection data, Geophysics 62 (6) (1997) 1975–1989.
[56] K. Yee, Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media, IEEE Trans. Antennas Propag. 14 (3) (1966) 302–307, https://doi.org/10.1109/TAP.1966.1138693.