

Untangling the nonlinearity in inverse scattering with data-driven reduced order models

Liliana Borcea¹, Vladimir Druskin², Alexander V Mamonov³ 
and Mikhail Zaslavsky²

¹ Department of Mathematics, University of Michigan, Ann Arbor, MI 48109-1043, United States of America

² Schlumberger-Doll Research Center, 1 Hampshire St., Cambridge, MA 02139-1578, United States of America

³ Department of Mathematics, University of Houston, 3551 Cullen Blvd Houston, TX 77204-3008, United States of America

E-mail: borcea@umich.edu, druskin1@slb.com, mamonov@math.uh.edu
and mzaslavsky@slb.com

Received 9 November 2017, revised 18 February 2018

Accepted for publication 3 April 2018

Published 4 May 2018



Abstract

The motivation of this work is an inverse problem for the acoustic wave equation, where an array of sensors probes an unknown medium with pulses and measures the scattered waves. The goal of the inversion is to determine from these measurements the structure of the scattering medium, modeled by a spatially varying acoustic impedance function. Many inversion algorithms assume that the mapping from the unknown impedance to the scattered waves is approximately linear. The linearization, known as the Born approximation, is not accurate in strongly scattering media, where the waves undergo multiple reflections before they reach the sensors in the array. Thus, the reconstructions of the impedance have numerous artifacts. The main result of the paper is a novel, linear-algebraic algorithm that uses a reduced order model (ROM) to map the data to those corresponding to the single scattering (Born) model. The ROM construction is based only on the measurements at the sensors in the array. The ROM is a proxy for the wave propagator operator, that propagates the wave in the unknown medium over the duration of the time sampling interval. The output of the algorithm can be input into any off-the-shelf inversion software that incorporates state of the art linear inversion algorithms to reconstruct the unknown acoustic impedance.

Keywords: inverse scattering, model reduction, rational Krylov subspace projection, Born approximation

(Some figures may appear in colour only in the online journal)

1. Introduction

Let us formulate the problem in a general setting, for a hyperbolic system of equations of the form

$$\partial_t \begin{pmatrix} P(t, \mathbf{x}) \\ \mathbf{U}(t, \mathbf{x}) \end{pmatrix} = \begin{pmatrix} 0 & -L_q \\ L_q^T & 0 \end{pmatrix} \begin{pmatrix} P(t, \mathbf{x}) \\ \mathbf{U}(t, \mathbf{x}) \end{pmatrix}, \quad \mathbf{x} \in \Omega, \quad t > 0, \quad (1)$$

satisfied by a wave field with components $P(t, \mathbf{x})$ and $\mathbf{U}(t, \mathbf{x})$, in a simply connected domain $\Omega \subset \mathbb{R}^d$ with piecewise smooth boundary $\partial\Omega = \partial\Omega_a \cup \partial\Omega_i$; given by the union of two sets: the first set is the accessible boundary $\partial\Omega_a$, where the measurements are made, and the second set is the inaccessible boundary $\partial\Omega_i$. In this paper we consider sound waves, where $P(t, \mathbf{x}) \in \mathbb{R}$ corresponds to the acoustic pressure field and $\mathbf{U}(t, \mathbf{x}) \in \mathbb{R}^d$ to the velocity field, satisfying the boundary conditions

$$\mathbf{U}(t, \mathbf{x}) \cdot \boldsymbol{\nu}(\mathbf{x}) = 0, \quad \mathbf{x} \in \partial\Omega_a \quad \text{and} \quad P(t, \mathbf{x}) = 0, \quad \mathbf{x} \in \partial\Omega_i, \quad (2)$$

where $\boldsymbol{\nu}(\mathbf{x})$ is the outer unit normal at $\partial\Omega_a$. However, the results can be extended to other boundary conditions and to other waves that satisfy a system of form (1), such as electromagnetic and elastic waves.

Note that $\partial\Omega_i$ may be a true boundary or a fictitious one, for a truncation of an infinite medium, in which case we can use causality to set the condition (2) at $\partial\Omega_i$, without affecting the wave measured at the sensors on $\partial\Omega_a$, for the duration $t \leq t_{\max}$.

The wave evolves in time t starting from

$$P(0, \mathbf{x}) = b(\mathbf{x}), \quad \mathbf{U}(0, \mathbf{x}) = 0, \quad (3)$$

as described by the skew symmetric operator in (1), with L_q a first order partial differential operator in the \mathbf{x} variable, and L_q^T its adjoint. The coefficients in these operators depend linearly on a function $q(\mathbf{x})$, which is the unknown in the inverse problem, to be determined from measurements of the wave, ‘the data’. These are modeled by continuously differentiable measurement functions \mathbf{M}_j of the self-adjoint operator $L_q L_q^T$,

$$\mathbf{D}_j = \mathbf{M}_j(L_q L_q^T), \quad j = 0, 1, \dots, 2n - 1, \quad (4)$$

with \mathbf{D}_j scalar or matrix valued, depending on the dimension d . The index j counts the discrete time instances $t_j = j\tau$ of the measurements, at time step τ .

In the context of inverse scattering for sound waves considered in this paper, the data are gathered by a collection (aka array) of sensors at $\partial\Omega_a$. These sensors act as both sources and receivers that probe the medium with pulses and measure the reflected pressure field. They lie in a compact set (the array aperture) on the accessible surface $\partial\Omega_a$ and are spaced close together at distances comparable to the central wavelength of the probing pulses. The acoustic system of wave equations takes the form (1) after a Liouville transformation of the pressure and velocity fields, and the array measurements can be written in the form (4), as described in sections 2 and 4.

If the medium has constant density, there is a single unknown in the inverse problem, the wave speed $c(\mathbf{x})$. In general, the medium has variable density $\rho(\mathbf{x})$, so we have two unknowns: the wave speed $c(\mathbf{x})$ and the acoustic impedance $\sigma(\mathbf{x}) = \rho(\mathbf{x})c(\mathbf{x})$. We assume henceforth that the waves propagate through a medium with known wave speed⁴, and the unknown $q(\mathbf{x})$ in

⁴In applications, we can only know the smooth wave speed in the medium that contains scattering inhomogeneities. If these inhomogeneities have constant density, their impedance is a constant multiple of the wave speed, so by finding $\sigma(\mathbf{x})$ we can also determine the variations of $c(\mathbf{x})$. Even if the perturbations of $c(\mathbf{x})$ cannot be determined, their effect is mainly manifested in small travel time coordinate deformations, and our approach still suppresses multiple scattering artifacts, as illustrated in section 4.7.

the inverse problem is the logarithm of the acoustic impedance $\sigma(\mathbf{x})$. This formulation of the inverse scattering problem is motivated by the generic setup in imaging, where waves propagate in a reference medium with smooth wave speed $c(\mathbf{x})$, and the goal is to determine rough perturbations of the medium, the ‘reflectivity’. The setup reflects the separation of scales in the problem, where $c(\mathbf{x})$ determines the kinematics (travel time) of the waves, whereas scattering occurs at the rough variations in the medium [53], like boundaries of inclusions.

In some applications, such as ultrasonic non-destructive evaluation [20] or radar imaging [15], the reference medium is approximately homogeneous, like air, so $c(\mathbf{x})$ is constant and known. In other applications, like reflection seismology [6, 34, 53], $c(\mathbf{x})$ must be determined from the measurements. Velocity estimation is a difficult problem because the wave fields are oscillatory in time and small perturbations of $c(\mathbf{x})$ can result in travel time perturbations that exceed the period of oscillations, which is a major change of the wave. This is a serious issue for data fitting optimization methods that use successive linearizations, but there are effective approaches for estimating $c(\mathbf{x})$ [39, 50, 55].

In this paper we assume that $c(\mathbf{x})$ is known, and focus attention on imaging the reflectivity. Most of the imaging technology is based on the linearization (Born approximation) of the mapping of the reflectivity to the scattered wave [4, 5, 15, 47]. The so-called Kirchhoff formulas [52, chapter 6] show that if the aperture of the array is not too large, the Born approximation of the reflected waves depends to leading order only on the perturbations of the acoustic impedance $\sigma(\mathbf{x})$. This is the unknown in our setting.

While the linearization assumption has led to popular imaging methods known as Kirchhoff migration [6], matched filtering [54] or filtered back-projection [15], multiple scattering effects are present and may lead to significant image artifacts [19, 41]. There has been progress in the removal of multiple scattering effects in three different contexts:

- (1) For imaging point-like scatterers buried deep in media with small random variations of the wave speed on scales comparable to the wavelength [1, 2, 7, 12], and for imaging in strongly scattering layered media [8, 27].
- (2) For imaging in a smooth reference medium, in the context of reflection seismology [40, 41, 58] and in optics [45], using inverse Born series or Bremmer series.
- (3) For imaging general media using a group of related methods such as boundary control [3, 18], Marchenko redatuming [44, 57] and data-driven reduced order models (ROMs) [23, 25]. The latter works are the foundation of the algorithm in this paper.

Here we consider an arbitrary unknown acoustic impedance $\sigma(\mathbf{x})$ and seek to transform the scattering data (4) to measurements expected in the Born approximation. The transformation, called data to Born (DtB) mapping, is the main result of the paper. We define it using an ROM of the wave problem, which can be calculated from the measurements (4) for $q(\mathbf{x}) = \ln \sigma(\mathbf{x})$. The ROM is defined by a matrix $\tilde{\mathbf{L}}_q$ of special structure, constructed from the matching relations

$$\mathbf{D}_j = \tilde{\mathbf{M}}_j(\tilde{\mathbf{L}}_q \tilde{\mathbf{L}}_q^T), \quad j = 0, \dots, 2n - 1, \quad (5)$$

for continuously differentiable ROM measurement functions $\tilde{\mathbf{M}}_j$ that do not depend on q . These are consistent with the functions \mathbf{M}_j in (4), as explained in [23] and the next sections. The DtB maps the measured data set to a new data set of the same size. The one-to-one correspondence between the data $(\mathbf{D}_j)_{0 \leq j \leq 2n-1}$ and the non-zero entries in $\tilde{\mathbf{L}}_q$ plays a critical role in our algorithm.

The ROM construction is rooted in the theory of Stieltjes strings due to Krein [36]. An outgrowth of this theory, the spectrally matched grids, also called optimal grids, designed to give

spectrally accurate finite difference approximations of Dirichlet-to-Neumann maps [22], were used for discretizations of exterior and multi-scale problems in [21, 24], and for the numerical solution of the electrical impedance tomography problem in the model reduction framework in [10, 11]. A related approach, based on Krein's work and the theory of Marchenko, Gel'fand and Levitan [30, 37, 38, 43], has been used in inverse hyperbolic problems in layered media in [13, 14, 32, 33, 49, 51]. Recent extensions to higher dimensions can be found in [35, 56]. At the core of this theory is the reduction of the inverse scattering problem to a nonlinear Volterra integral equation, or a system of equations. In the discrete, linear algebra setting, this translates to the Lanczos and block Lanczos algorithms or, alternatively, the Stieltjes moment problems [26, 28, 29] and the Cholesky or block-Cholesky algorithms used in the construction of $\tilde{\mathbf{L}}_q$ [23, 25].

The linear-algebraic construction of $\tilde{\mathbf{L}}_q$ from the data (4) is among the features that differentiate our method from the recent work on Marchenko imaging and redatuming [17, 44, 57]. All these methods involve solving numerically integral equations, with some, often ad-hoc, discretization. We avoid this issue because we never compute any quantities in the physical domain Ω . The only discrete aspect of our method is due to the time sampling of the data $(\mathbf{D}_j)_{0 \leq j \leq 2n-1}$ at time intervals τ . In the absence of measurement noise and modeling errors, we have by the construction of the ROM an exact data fit (5).

We explain in sections 2 and 4 that the matrix $\tilde{\mathbf{L}}_q$ obtained from (5) is a Galerkin–Petrov approximation of the operator L_q , for carefully constructed bases of the spaces of approximation of the fields $P(t, \mathbf{x})$ and $\mathbf{U}(t, \mathbf{x})$. We also discuss in section 3 a related ROM, constructed from spectral measurements of the operator $L_q L_q^T$ in one dimension [9]. The analogy is useful for interpreting the entries of $\tilde{\mathbf{L}}_q$ in terms of averages of the unknown impedance $\sigma(\mathbf{x})$ on a special ‘spectrally matched’ grid.

While there are other choices of reduced order models, the ones considered in this paper have an important property: They are approximately linear in the unknown $q(\mathbf{x})$. This means that if we had a perturbation $q^\varepsilon(\mathbf{x})$ of a known $q^0(\mathbf{x})$, of the form

$$q^\varepsilon(\mathbf{x}) = q^0(\mathbf{x}) + \varepsilon[q(\mathbf{x}) - q^0(\mathbf{x})], \quad 0 < \varepsilon \ll 1, \quad (6)$$

the operator L_q , which is linear in q , would be perturbed as

$$L_{q^\varepsilon} = L_{q^0} + \varepsilon[L_q - L_{q^0}], \quad (7)$$

and the corresponding ROM would satisfy a similar relation

$$\tilde{\mathbf{L}}_{q^\varepsilon} \approx \tilde{\mathbf{L}}_{q^0} + \varepsilon[\tilde{\mathbf{L}}_q - \tilde{\mathbf{L}}_{q^0}]. \quad (8)$$

Here $\tilde{\mathbf{L}}_{q^0}$ is constructed the same way as $\tilde{\mathbf{L}}_q$, from the reference data $\mathbf{D}^0 = \{\mathbf{D}_j^0\}_{j=0}^{2n-1}$ calculated by solving equation (1) with the operator L_{q^0} .

We do not have access to the data \mathbf{D}_j^ε for coefficient (6). However, since the ROM is obtained from the matching conditions (5), we obtain from (8) the approximation

$$\mathbf{D}_j^\varepsilon \approx \tilde{\mathbf{M}}_j(\tilde{\mathbf{L}}_{q^\varepsilon} \tilde{\mathbf{L}}_{q^\varepsilon}^T) \approx \tilde{\mathbf{M}}_j([\tilde{\mathbf{L}}_{q^0} + \varepsilon(\tilde{\mathbf{L}}_q - \tilde{\mathbf{L}}_{q^0})][\tilde{\mathbf{L}}_{q^0} + \varepsilon(\tilde{\mathbf{L}}_q - \tilde{\mathbf{L}}_{q^0})]^T). \quad (9)$$

The Born data model is defined by

$$\mathbf{D}_j^{\text{Born}, \varepsilon} = \mathbf{D}_j^0 + \varepsilon \left[\frac{d}{d\varepsilon'} \mathbf{D}_j^{\varepsilon'} \Big|_{\varepsilon'=0} \right], \quad (10)$$

and using (9) we approximate it with the DtB mapping \mathcal{D} , which takes the measurements $\mathbf{D} = \{\mathbf{D}_j\}_{j=0}^{2n-1}$ with entries (5) for the unknown $q(\mathbf{x})$, and returns

$$\mathcal{D}[\mathbf{D}] = \left\{ \mathbf{D}_j^0 + \varepsilon \left[\frac{d}{d\varepsilon'} \tilde{\mathbf{M}}_j \left([\tilde{\mathbf{L}}_{q^0} + \varepsilon'(\tilde{\mathbf{L}}_q - \tilde{\mathbf{L}}_{q^0})] [\tilde{\mathbf{L}}_{q^0} + \varepsilon'(\tilde{\mathbf{L}}_q - \tilde{\mathbf{L}}_{q^0})]^T \right) \Big|_{\varepsilon'=0} \right] \right\}_{j=0}^{2n-1}. \quad (11)$$

Note that ε is an arbitrary scaling factor in this equation. We take it equal to 1 so that (6) equals q .

The DtB mapping (11) is described in sections 2 and 4. We define it from first principles in the one dimensional case $d = 1$ in section 2, and then extend the results to multi-dimensions in section 4. The related inverse spectral problem for the hyperbolic system (1) is discussed in section 3. We conclude with a summary and a brief discussion of directions of future research in section 5.

2. The DtB mapping in one dimension

We define here the mapping (11) in one dimension. We begin in section 2.1 with the derivation of the data model (4), starting from the acoustic wave equation. Then we introduce in section 2.2 the wave propagator operator, which we use in section 2.3 to construct the ROM. The matrix $\tilde{\mathbf{L}}_q$ that defines the ROM is a Galerkin–Petrov approximation of the operator L_q , as shown in section 2.4. The algorithm for computing the DtB mapping (11) is in section 2.5, and we illustrate its performance with numerical simulations in section 2.6.

2.1. Derivation of the data model

Let us consider sound waves modeled by the excess acoustic pressure denoted by $\mathbf{p}(t, x)$. We use the different script notation \mathbf{p} to distinguish this field from another, related pressure field defined below, in equation (25).

The pressure field $\mathbf{p}(t, x)$ is defined in the domain $x > 0$, with sound hard boundary at $x = 0$,

$$\partial_x \mathbf{p}(t, 0) = 0. \quad (12)$$

For a finite duration $t < t_{\max}$, with

$$t_{\max} < T_\ell = \int_0^\ell \frac{dx}{c(x)}, \quad (13)$$

we can truncate the domain at $x = \ell$ without affecting the wave at $x = 0$, and set

$$\mathbf{p}(t, \ell) = 0. \quad (14)$$

Thus, $\mathbf{p}(t, x)$ satisfies the wave equation

$$(\partial_t^2 + A)\mathbf{p}(t, x) = \partial_t f(t)\delta(x - 0^+), \quad t \in \mathbb{R}, \quad x \in (0, \ell), \quad (15)$$

in the domain $\Omega = (0, \ell)$, with boundary condition (12) at the accessible boundary $\partial\Omega_a = \{0\}$ and (14) at the inaccessible boundary $\partial\Omega_i = \{\ell\}$. The operator A is given by

$$A = -\sigma(x)c(x)\partial_x \left[\frac{c(x)}{\sigma(x)} \partial_x \right]. \quad (16)$$

The medium is at equilibrium prior to the emission of the pulse $f(t)$ from a source located at $x = 0^+$,

$$\mathbf{p}(t, x) = 0, \quad t \ll 0. \quad (17)$$

For convenience in the derivation of the ROM, we take $f(t)$ real valued, with Fourier transform $\widehat{f}(\omega) \geq 0$. For example, $f(t)$ may be a modulated Gaussian with central frequency ω_0 and bandwidth B

$$f(t) = \frac{\cos(\omega_0 t)}{\sqrt{2\pi B}} e^{-\frac{(Bt)^2}{2}},$$

so that its Fourier transform is

$$\widehat{f}(\omega) = \int_{-\infty}^{\infty} dt e^{i\omega t} f(t) = \frac{1}{2} \left[e^{-\frac{(\omega - \omega_0)^2}{2B^2}} + e^{-\frac{(\omega + \omega_0)^2}{2B^2}} \right].$$

Note that A is self-adjoint in the Hilbert space $\mathcal{H}_{\frac{1}{\sigma c}} = L^2\left([0, \ell], \frac{1}{\sigma(x)c(x)} dx\right)$ with weighted inner product

$$\langle \varphi, \psi \rangle_{\frac{1}{\sigma c}} = \int_0^\ell dx \frac{\varphi(x)\psi(x)}{\sigma(x)c(x)}, \quad \forall \varphi, \psi \in \mathcal{H}_{\frac{1}{\sigma c}}, \quad (18)$$

on the domain of functions $\varphi(x) \in \mathcal{H}_{\frac{1}{\sigma c}}$, satisfying $\varphi'(0) = 0$, $\varphi(\ell) = 0$. It has simple and positive eigenvalues $\{\lambda_j\}_{j \geq 1}$ and the eigenfunctions $\{y_j(x)\}_{j \geq 1}$ form an orthonormal basis of $\mathcal{H}_{\frac{1}{\sigma c}}$. Expanding $\mathbf{p}(t, x)$ in this basis we obtain the separation of variables formula

$$\mathbf{p}(t, x) = f(t) \star \sum_{j=1}^{\infty} H(t) \cos(t\sqrt{\lambda_j}) \frac{y_j(0)y_j(x)}{\sigma(0)c(0)}, \quad (19)$$

where \star denotes convolution, $H(t)$ is the Heaviside step function, and the series is the causal Green's function of (15).

We work with the even time extension of $\mathbf{p}(t, x)$,

$$\mathbf{p}^e(t, x) = \mathbf{p}(t, x) + \mathbf{p}(-t, x), \quad (20)$$

because it has a simpler expression than (19),

$$\mathbf{p}^e(t, x) = \cos(t\sqrt{A}) \widehat{f}(\sqrt{A}) \delta(x) = \sum_{j=1}^{\infty} \cos(t\sqrt{\lambda_j}) \widehat{f}(\sqrt{\lambda_j}) \frac{y_j(0)y_j(x)}{\sigma(0)c(0)}. \quad (21)$$

This defines the data

$$D_j = \mathbf{p}^e(t_j, 0) = \mathbf{p}(t_j, 0) + \mathbf{p}(-t_j, 0), \quad j = 0, \dots, 2n-1, \quad (22)$$

for the inverse scattering problem with unknown impedance $\sigma(x)$. The instances $t_j = j\tau$ of measurement are equally spaced, at sufficiently small interval $\tau = t_{\max}/(2n-1)$, as explained in the next section. Since $\mathbf{p}(-t, 0) = 0$ for t exceeding the temporal support of the pulse $f(t)$, the second term in (22) plays a role only for the first few indexes j .

Using the expression (21) and the self-adjointness of A , we can rewrite (22) in the symmetric form

$$D_j = \sigma(0)c(0) \langle \delta(\cdot), \mathbf{p}^e(t, \cdot) \rangle_{\frac{1}{\sigma c}} = \left\langle \sqrt{\sigma} b, \cos(t_j \sqrt{A}) \sqrt{\sigma} b \right\rangle_{\frac{1}{\sigma c}}, \quad (23)$$

with the notation

$$\sqrt{\sigma(x)}b(x) = \sqrt{\sigma(0)c(0)}[\widehat{f}(\sqrt{A})]^{1/2}\delta(x) = \sum_{j=1}^{\infty} [\widehat{f}(\sqrt{\lambda_j})]^{1/2} \frac{y_j(0)y_j(x)}{\sqrt{\sigma(0)c(0)}}. \quad (24)$$

We call $b(x)$ the ‘sensor function’, because it is supported near $x = 0$ and appears in equation (23) as a model of the source and receiver⁵.

To arrive at the first order hyperbolic system (1), note that

$$p(t, x) = \cos(t\sqrt{A})\sqrt{\sigma(x)}b(x) \quad (25)$$

is the pressure field in the acoustic system of equations

$$\partial_t \begin{pmatrix} p(t, x) \\ -u(t, x) \end{pmatrix} = \begin{pmatrix} 0 & \sigma(x)c(x)\partial_x \\ \frac{c(x)}{\sigma(x)}\partial_x & 0 \end{pmatrix} \begin{pmatrix} p(t, x) \\ -u(t, x) \end{pmatrix}, \quad t > 0, \quad x \in (0, \ell), \quad (26)$$

with initial conditions

$$p(0, x) = \sqrt{\sigma(x)}b(x), \quad u(0, x) = 0, \quad (27)$$

and with boundary conditions

$$p(t, \ell) = 0, \quad u(t, 0) = 0. \quad (28)$$

Here $u(t, x)$ is the particle velocity.

The system (26) is not in the desired form for our purpose, because the unknown impedance $\sigma(x)$ appears in a nonlinear fashion in the coefficients of the differential operator. We show next how to transform (26) to the system (1), with operator L_q and its adjoint L_q^T depending linearly on $q(x) = \ln \sigma(x)$.

2.1.1. The Schrödinger system of equations. Consider the Liouville transformation

$$P(t, x) = \frac{p(t, x)}{\sqrt{\sigma(x)}}, \quad U(t, x) = -\sqrt{\sigma(x)}u(t, x), \quad (29)$$

which takes (26) to

$$\partial_t \begin{pmatrix} P(t, x) \\ U(t, x) \end{pmatrix} = \begin{pmatrix} 0 & -L_q \\ L_q^T & 0 \end{pmatrix} \begin{pmatrix} P(t, x) \\ U(t, x) \end{pmatrix}, \quad t > 0, \quad x \in (0, \ell). \quad (30)$$

This is the system (1) in the introduction, with

$$L_q = -c(x)\partial_x + \frac{1}{2}c(x)\partial_x q(x). \quad (31)$$

The adjoint of (31) with respect to the inner product $\langle \cdot, \cdot \rangle_{\frac{1}{c}}$ weighted by $1/c(x)$, is given by

$$L_q^T = c(x)\partial_x + \frac{1}{2}c(x)\partial_x q(x), \quad (32)$$

and we note that both L_q and L_q^T are first order Schrödinger operators with potentials that are linear in $q(x) = \ln \sigma(x)$.

⁵Our construction of the DtB mapping uses that $b(x)$ is supported near $x = 0$, but does not require knowing $b(x)$. In the case of a homogeneous medium we can calculate $b(x)$ in terms of $f(x/c)$, which is localized at $x = 0$. In a variable medium the eigenfunctions are not known, but they are oscillatory, and the right hand side in (24) is a generalized Fourier series of the smooth function $(\widehat{f})^{1/2}$. This series is localized near $x = 0$.

The transformed fields (29), called henceforth the ‘primary wave’ $P(t, x)$ and the ‘dual wave’ $U(t, x)$, satisfy the initial conditions

$$P(0, x) = b(x), \quad U(0, x) = 0, \quad (33)$$

and boundary conditions

$$U(t, 0) = 0, \quad P(t, \ell) = 0. \quad (34)$$

These are the conditions (2) and (3) stated in the introduction.

The data model follows from equations (23), (25) and (29)

$$D_j = \langle b, P(t_j, \cdot) \rangle_{\frac{1}{c(x)}} = \int_0^\ell \frac{dx}{c(x)} b(x) P(t_j, x), \quad j = 0, \dots, 2n - 1. \quad (35)$$

2.1.2. Travel time coordinates. In one dimension we can avoid dealing with weighted inner products, by changing coordinates in (30) from x to the travel time

$$T(x) = \int_0^x \frac{ds}{c(s)}, \quad x \in [0, \ell]. \quad (36)$$

This transformation is invertible for $T \in [0, T_\ell]$, with T_ℓ defined in (13) as the travel time from the accessible boundary at $x = 0$ to the inaccessible boundary at $x = \ell$. Thus, we can write $x = x(T)$, for $T \in [0, T_\ell]$.

We keep the same notation for the operator (31) in the travel time coordinates

$$L_q = -\partial_T + \frac{1}{2} \partial_T q(T), \quad (37)$$

and its adjoint with respect to the usual, Euclidean inner product $\langle \cdot, \cdot \rangle$

$$L_q^T = \partial_T + \frac{1}{2} \partial_T q(T). \quad (38)$$

We also let $q(T) = \ln \sigma(x(T))$ and $b(T) = b(x(T))$.

The measurements (35) are defined by the primary wave at the time instances $t_j = j\tau$, denoted by

$$P_j(T) = P(t_j, x(T)) = \cos\left(t_j \sqrt{L_q L_q^T}\right) b(T), \quad j = 0, \dots, 2n - 1. \quad (39)$$

We also use in the discussion the dual field $U(t, x(T))$ evaluated at the shifted time instances $t_j + \tau/2$. Solving equation (30) for U we get

$$U(t, x(T)) = \sin\left(t \sqrt{L_q^T L_q}\right) (L_q^T L_q)^{-1/2} L_q^T b(T), \quad (40)$$

and we denote

$$U_j(T) = U(t_j + \tau/2, x(T)), \quad j = 0, \dots, 2n - 1. \quad (41)$$

2.2. The propagator

The propagator of the primary wave is the operator

$$\mathcal{P} = \cos\left(\tau \sqrt{L_q L_q^T}\right), \quad (42)$$

that maps the initial condition $b(T)$ to $P(\tau, x(T))$. We use it in equation (39) to write

$$P_j(T) = \cos(j \arccos(\mathcal{P}))b(T) = \mathcal{T}_j(\mathcal{P})b(T), \quad j = 0, \dots, 2n-1, \quad (43)$$

where $\mathcal{T}_j(\mathcal{P}) = \cos(j \arccos(\mathcal{P}))$ are the Chebyshev polynomials of the first kind [48]. The data model (35) takes the form (4), with measurement functions M_j defined by⁶

$$D_j = M_j(L_q L_q^T) = \langle b, P_j \rangle = \langle b, \mathcal{T}_j(\mathcal{P})b \rangle, \quad j = 0, \dots, 2n-1. \quad (44)$$

The propagator of the dual wave is the operator

$$\mathcal{U} = \cos\left(\tau \sqrt{L_q^T L_q}\right), \quad (45)$$

and it is shown in [23, lemma 3.6] that

$$U_j(T) = \left[\mathcal{T}_j^{(2)}(\mathcal{U}) + \mathcal{T}_{j-1}^{(2)}(\mathcal{U})\right]U_0(T), \quad j = 0, \dots, 2n-1, \quad (46)$$

with U_0 obtained from (40) evaluated at $t = \tau/2$, and $\mathcal{T}_j^{(2)}$ the Chebyshev polynomials of the second kind [48].

2.2.1. Time stepping and factorization of the propagator. Because the Chebyshev polynomials satisfy the three term recurrence relation

$$\mathcal{T}_{j+1}(\mathcal{P}) + \mathcal{T}_{j-1}(\mathcal{P}) = 2\mathcal{P}\mathcal{T}_j(\mathcal{P}), \quad j \geq 1, \text{ and } \mathcal{T}_0(\mathcal{P}) = I, \quad \mathcal{T}_1(\mathcal{P}) = \mathcal{P}, \quad (47)$$

where I is the identity operator, we obtain from definition (43) that the primary wave satisfies the exact time stepping scheme

$$\frac{1}{\tau^2} \left[P_{j+1}(T) - 2P_j(T) + P_{j-1}(T) \right] = -\xi(\mathcal{P})P_j(T), \quad j = 0, \dots, 2n-1, \quad (48)$$

with initial conditions

$$P_0(T) = b(T), \quad P_{-1}(T) = P_1(T) = \mathcal{P}b(T). \quad (49)$$

Here we introduced the affine function

$$\xi(\mathcal{P}) = \frac{2}{\tau^2}(I - \mathcal{P}), \quad (50)$$

and the last relation in (49) is derived from

$$P_1(T) + P_{-1}(T) = \left[\mathcal{T}_1(\mathcal{P}) + \mathcal{T}_{-1}(\mathcal{P})\right]b(T) = 2\mathcal{P}b(T) = 2P_1(T).$$

Similarly, we obtain an explicit time stepping scheme for the dual wave, from equation (46) and the definition of the Chebyshev polynomials of the second kind

⁶ Note that in our formulation the sensor function $b(T)$ depends on q . We do not write this dependence explicitly in M_j because in the ROM construction given in section 2.3, $b(T)$ is mapped to the ‘ROM sensor vector’ $\tilde{\mathbf{b}} = D_0^{1/2}\mathbf{e}_1$, with $\mathbf{e}_1 = (1, 0, \dots, 0)^T \in \mathbb{R}^n$ and $n \ll N$. Thus, we can remove the dependence on q of the ROM measurement functions \tilde{M}_j by either normalizing the measurements with D_0 , or by assuming that $q(T)$ is known near the accessible boundary i.e. at $T \approx 0$.

$$\mathcal{T}_j^{(2)}(\mathcal{U}) = \begin{cases} 2 \sum_{i=1, i=\text{odd}}^j \mathcal{T}_i(\mathcal{U}), & j = \text{odd}, \\ 2 \sum_{i=0, i=\text{even}}^j \mathcal{T}_i(\mathcal{U}) - I, & j = \text{even}, \end{cases} \quad (51)$$

with $\mathcal{T}_{-2}^{(2)}(\mathcal{U}) = \mathcal{T}_{-1}^{(2)}(\mathcal{U}) = 0$. We have

$$\frac{1}{\tau^2} [U_{j+1}(T) - 2U_j(T) + U_{j-1}(T)] = -\xi(\mathcal{U})U_j(T), \quad j = 0, \dots, 2n-1, \quad (52)$$

with $U_0(T)$ defined by (40) at $t = \tau/2$ and

$$U_0(T) + U_{-1}(T) = 0, \quad (53)$$

derived from (51) and definition (46).

We can write these two schemes in first order system form, by factorizing the affine function of the propagators in the right hand side of (48) and (52). We obtain that

$$\xi(\mathcal{P}) = \frac{4}{\tau^2} \sin^2 \left(\frac{\tau}{2} \sqrt{L_q L_q^T} \right) = \mathcal{L}_q \mathcal{L}_q^T, \quad (54)$$

and

$$\xi(\mathcal{U}) = \frac{4}{\tau^2} \sin^2 \left(\frac{\tau}{2} \sqrt{L_q^T L_q} \right) = \mathcal{L}_q^T \mathcal{L}_q, \quad (55)$$

with operator

$$\mathcal{L}_q = \frac{2}{\tau} L_q (L_q^T L_q)^{-1/2} \sin \left(\frac{\tau}{2} \sqrt{L_q^T L_q} \right), \quad (56)$$

and its adjoint \mathcal{L}_q^T with respect to the Euclidean inner product. Then, equations (48) and (52) are equivalent to the first order time stepping scheme

$$\begin{aligned} \frac{P_{j+1}(T) - P_j(T)}{\tau} &= -\mathcal{L}_q U_j(T), \quad j = 0, \dots, 2n-2, \\ \frac{U_j(T) - U_{j-1}(T)}{\tau} &= \mathcal{L}_q^T P_j(T), \quad j = 1, \dots, 2n-1, \end{aligned} \quad (57)$$

with initial conditions

$$P_0(T) = b(T), \quad U_0(T) + U_{-1}(T) = 0. \quad (58)$$

This is the exact time discretization of the system (30), for time $t \in [0, t_{\max}]$ sampled at intervals τ .

2.3. The reduced order model

To avoid technical arguments, we work with the discretization of (57) on a very fine grid in the interval $(0, T_\ell)$, with $N \gg 1$ equidistant points at spacing $\Delta T = T_\ell/N$. Using a two point finite difference scheme on this grid, we obtain an $N \times N$ lower bidiagonal matrix \mathbf{L}_q , the discretization of the Schrödinger operator (37). The operator (56) is discretized by

$$\mathcal{L}_q = \frac{2}{\tau} L_q (L_q^T L_q)^{-1/2} \sin \left(\frac{\tau}{2} \sqrt{L_q^T L_q} \right) = L_q \left(\mathbf{I}_N - \frac{\tau^2}{24} L_q^T L_q + \dots \right), \quad (59)$$

where \mathbf{I}_N is the $N \times N$ identity matrix. Assuming a small time sampling interval τ , so that

$$\tau < \|\mathbf{L}_q^T \mathbf{L}_q\|^{-1/2} = O(\Delta T), \quad (60)$$

we obtain

$$\mathcal{L}_q \approx \mathbf{L}_q. \quad (61)$$

Then, the primary and dual propagator $N \times N$ matrices

$$\mathcal{P} = \mathbf{I}_N - \frac{\tau^2}{2} \mathcal{L}_q \mathcal{L}_q^T \approx \mathbf{I}_N - \frac{\tau^2}{2} \mathbf{L}_q \mathbf{L}_q^T, \quad (62)$$

and

$$\mathcal{U} = \mathbf{I}_N - \frac{\tau^2}{2} \mathcal{L}_q^T \mathcal{L}_q \approx \mathbf{I}_N - \frac{\tau^2}{2} \mathbf{L}_q^T \mathbf{L}_q, \quad (63)$$

are approximately tridiagonal. Here we used definition (50), the factorizations (54), (55) and the approximation (61).

We call the vectors \mathbf{P}_j and \mathbf{U}_j in \mathbb{R}^N , with entries approximating $P_j(T)$ and $U_j(T)$ on the fine grid, the primary and dual ‘solution snapshots’. They evolve from the initial values $\mathbf{P}_0 = \mathbf{b}$ and \mathbf{U}_0 according to the equations

$$\mathbf{P}_j = \mathcal{T}_j(\mathcal{P})\mathbf{b}, \quad \mathbf{U}_j = \left[\mathcal{T}_j^{(2)}(\mathcal{U}) + \mathcal{T}_{j-1}^{(2)}(\mathcal{U}) \right] \mathbf{U}_0, \quad (64)$$

for $j = 0, \dots, 2n - 1$. The data model (44) becomes

$$D_j = M_j(\mathbf{L}_q \mathbf{L}_q^T) \approx M_j(\mathcal{L}_q \mathcal{L}_q^T) = \mathbf{b}^T \mathbf{P}_j = \mathbf{b}^T \mathcal{T}_j(\mathcal{P})\mathbf{b}, \quad (65)$$

with small error of the approximation, of order $1/N$, for $N \gg 1$. Here \mathbf{b} is the ‘sensor vector’ in \mathbb{R}^N with entries defined by the values of the sensor function $b(T)$ on the grid, multiplied by $\sqrt{\Delta T}$, so that

$$D_0 = \mathbf{b}^T \mathbf{b} = \langle b, b \rangle + O(1/N). \quad (66)$$

We neglect henceforth the $O(1/N)$ error and treat (65) and (66) as equalities.

The ROM is defined by the $n \times n$ symmetric and tridiagonal (Jacobi) matrix $\tilde{\mathcal{P}}$, satisfying the data matching conditions

$$D_j = \tilde{\mathbf{b}}^T \mathcal{T}_j(\tilde{\mathcal{P}})\tilde{\mathbf{b}}, \quad \tilde{\mathbf{b}} = D_0^{1/2} \mathbf{e}_1, \quad j = 0, \dots, 2n - 1, \quad (67)$$

where $\mathbf{e}_1 = (1, 0, \dots, 0)^T \in \mathbb{R}^n$ and $n \ll N$. Comparing (67) with (65), we note that \mathbf{b} , which is supported in the first rows, is replaced by the ‘ROM sensor vector’ $\tilde{\mathbf{b}} = \|\mathbf{b}\| \mathbf{e}_1$, with $\|\mathbf{b}\| = D_0^{1/2}$ by (66). We refer to [23] for many details on the propagator $\tilde{\mathcal{P}}$. Here it suffices to obtain its factorization

$$\tilde{\mathcal{P}} = \mathbf{I}_n - \frac{\tau^2}{2} \tilde{\mathbf{L}}_q \tilde{\mathbf{L}}_q^T, \quad (68)$$

with \mathbf{I}_n the $n \times n$ identity, and lower bidiagonal $\tilde{\mathbf{L}}_q$. This is the matrix used in the DtB mapping (11), and we explain in the next section how to calculate it.

2.3.1. Projection ROM. It is shown in [23, lemma 4.5] that $\tilde{\mathcal{P}}$ can be constructed with an orthogonal projection of \mathcal{P} on the span of the first n primary snapshots $\{\mathbf{P}_j\}_{j=0}^{n-1}$, the range of the $N \times n$ matrix

$$\mathbf{P} = (\mathbf{P}_0, \dots, \mathbf{P}_{n-1}). \quad (69)$$

By equation (64), this is the Krylov subspace

$$\mathcal{K}_n(\mathbf{b}, \mathcal{P}) = \text{span}\{\mathbf{b}, \mathcal{P}\mathbf{b}, \dots, \mathcal{P}^{n-1}\mathbf{b}\} = \text{range}(\mathbf{P}). \quad (70)$$

It is intuitive that the projection space is determined only by the first n snapshots. The back-scattered wave measured at $x = 0$, for $t \leq t_{2n-1}$, cannot propagate farther than $x(t_{n-1})$ in the medium, before it reflects and turns back to $x = 0$. This means that we can image up to depth $x(t_{n-1})$, and all the information is contained in the subspace (70).

There are many ways to project on $\mathcal{K}_n(\mathbf{b}, \mathcal{P})$, depending on the choice of the basis. We use an orthonormal basis $\{\mathbf{V}_1, \dots, \mathbf{V}_n\}$ that makes the projection

$$\tilde{\mathcal{P}} = \mathbf{V}^T \mathcal{P} \mathbf{V}, \quad (71)$$

tridiagonal, where $\mathbf{V} = (\mathbf{V}_1, \dots, \mathbf{V}_n)$ is the $N \times n$ orthogonal matrix in the QR factorization

$$\mathbf{P} = \mathbf{V} \mathbf{R}, \quad (72)$$

with $\mathbf{R} \in \mathbb{R}^{n \times n}$ invertible and upper triangular [31]. Because of this triangular matrix we obtain from (72) that the basis satisfies the causality relations

$$\mathbf{V}_j \in \text{span}\{\mathbf{P}_0, \dots, \mathbf{P}_{j-1}\}, \quad j = 1, \dots, n. \quad (73)$$

This is important for at least two reasons: First, it ensures that $\tilde{\mathcal{P}}$ is tridiagonal, as shown in appendix A. Second, it concentrates the support of \mathbf{V}_j near the wavefront, at $T = t_j$, and makes the matrix \mathbf{V} almost independent of the unknown q .

There are two ways of explaining this last property of the vectors \mathbf{V}_j : One way is to start with \mathbf{V}_1 which equals $\mathbf{P}_0 = \mathbf{b}$, up to a normalization factor, and recall that \mathbf{b} is supported in the first rows, at travel time $T \approx t_0 = 0$. The support of the second snapshot \mathbf{P}_1 advances by the travel time $T = t_1$. Since $\mathbf{V}_2 \in \text{span}\{\mathbf{V}_1, \mathbf{P}_1\}$ and \mathbf{V}_2 is orthogonal to \mathbf{V}_1 , the entries in \mathbf{V}_2 must be large at the wavefront $T = t_1$. Arguing this way, with index j increased one by one, we see that the support of the orthonormal basis follows the wavefront of the wave. Depending on how oscillatory the pulse is, there are some reverberations behind the wavefront, but as shown in the numerical simulations, the entries in \mathbf{V}_j are larger around travel times $T = t_{j-1}$. This property is important in our context, because the travel times are determined by the known wave speed $c(x)$, and not the unknown impedance or, equivalently, $q(x)$. This means that \mathbf{V} is almost independent of $q(x)$, as illustrated in section 2.5.

The other way of explaining is algebraic: By causality, the matrix \mathbf{P} of the primary snapshots is approximately upper triangular. The approximation is because $\mathbf{P} \in \mathbb{R}^{N \times n}$ is a tall rectangular matrix and \mathbf{b} is not an exact delta-function, but an approximation. If there were no inhomogeneities in the medium, there would be no reflected waves and the matrix \mathbf{P} would be approximately diagonal. The inhomogeneities cause reflections, which fill-in the upper triangular part of \mathbf{P} . The QR orthogonalization (72) transforms the almost upper triangular matrix \mathbf{P} to the almost identity matrix \mathbf{V} , which is almost independent of q .

2.3.2. The calculation of the ROM. Although the QR factorization (72) is useful for understanding the ROM, we cannot use it directly to compute $\tilde{\mathcal{P}}$ because we do not know the matrix (69). We only know the inner products of its columns with \mathbf{b} , from (65). We now explain how to calculate $\tilde{\mathcal{P}}$ from the matching relations (67).

Let us begin with the calculation of the upper triangular matrix \mathbf{R} . We obtain from equations (64) and (72) that

$$(\mathbf{P}^T \mathbf{P})_{j,k} = \mathbf{b}^T \mathcal{T}_{j-1}(\mathcal{P}) \mathcal{T}_{k-1}(\mathcal{P}) \mathbf{b} = (\mathbf{R}^T \mathbf{R})_{j,k}, \quad j, k = 1, \dots, n, \quad (74)$$

where we used the symmetry of \mathcal{P} . The Chebyshev polynomials satisfy the relation

$$\mathcal{T}_j(\mathcal{P}) \mathcal{T}_k(\mathcal{P}) = \frac{1}{2} [\mathcal{T}_{j+k}(\mathcal{P}) + \mathcal{T}_{|j-k|}(\mathcal{P})], \quad (75)$$

so substituting in (74) and using (65), we get

$$(\mathbf{P}^T \mathbf{P})_{j,k} = (\mathbf{R}^T \mathbf{R})_{j,k} = \frac{1}{2} (D_{j+k-2} + D_{|j-k|}). \quad (76)$$

This shows that the $n \times n$ matrix $\mathbf{P}^T \mathbf{P}$ can be determined from the data, and that \mathbf{R} can be calculated from its Cholesky factorization [31]

$$\mathbf{P}^T \mathbf{P} = \mathbf{R}^T \mathbf{R}. \quad (77)$$

With the matrix \mathbf{R} calculated from (77), we solve for \mathbf{V} in (72) to obtain

$$\mathbf{V} = \mathbf{P} \mathbf{R}^{-1}, \quad (78)$$

and then rewrite (71) as

$$\widetilde{\mathcal{P}} = \mathbf{R}^{-T} (\mathbf{P}^T \mathcal{P} \mathbf{P}) \mathbf{R}^{-1}. \quad (79)$$

The matrix in parentheses has the entries

$$(\mathbf{P}^T \mathcal{P} \mathbf{P})_{j,k} = \mathbf{b}^T \mathcal{T}_{j-1}(\mathcal{P}) \mathcal{P} \mathcal{T}_{k-1}(\mathcal{P}) \mathbf{b}, \quad (80)$$

by definition (64). Then, relation (75), $\mathcal{T}_1(\mathcal{P}) = \mathcal{P}$, and definition (65) give

$$(\mathbf{P}^T \mathcal{P} \mathbf{P})_{j,k} = \frac{1}{4} (D_{j+k-1} + D_{|k-j+1|} + D_{|k-j-1|} + D_{|k+j-3|}), \quad (81)$$

for $j, k = 1, \dots, n$. This shows that $\mathbf{P}^T \mathcal{P} \mathbf{P}$ can be computed from the data, and the propagator \mathcal{P} follows from (79).

To obtain the factorization (68), we note from (71) that the spectral norm of the ROM propagator is bounded above by the spectral norm of $\mathcal{P} = \cos(\tau \sqrt{\mathbf{L}_q^T \mathbf{L}_q})$. With our choice (60) of τ this norm is strictly less than one, so $\mathbf{I}_n - \widetilde{\mathcal{P}}$ is positive definite. Therefore, we can calculate the matrix $\widetilde{\mathbf{L}}_q$ in (68) from another Cholesky factorization

$$\xi(\widetilde{\mathcal{P}}) = \frac{2}{\tau^2} (\mathbf{I}_n - \widetilde{\mathcal{P}}) = \widetilde{\mathbf{L}}_q \widetilde{\mathbf{L}}_q^T. \quad (82)$$

This is the ROM version of equation (54).

It remains to show that the vector \mathbf{b} in the data matching conditions has the form given in (67). We define $\widetilde{\mathbf{b}}$ as the projection of \mathbf{b} on the space (70), given by

$$\widetilde{\mathbf{b}} = \mathbf{V}^T \mathbf{b}. \quad (83)$$

Using equations (77), (78), (66) and the upper triangular structure of \mathbf{R} we get

$$\widetilde{\mathbf{b}} = \mathbf{R}^{-T} \mathbf{P}^T \mathbf{P} \mathbf{e}_1 = \mathbf{R} \mathbf{e}_1 = (\mathbf{e}_1^T \mathbf{P}^T \mathbf{P} \mathbf{e}_1)^{1/2} \mathbf{e}_1 = (\mathbf{b}^T \mathbf{b})^{1/2} \mathbf{e}_1 = D_0^{1/2} \mathbf{e}_1, \quad (84)$$

as stated in (67).

The ROM measurement functions \widetilde{M}_j in (5) are defined by

$$\widetilde{M}_j(\widetilde{\mathbf{L}}_q \widetilde{\mathbf{L}}_q^T) = D_0^{1/2} \mathbf{e}_1^T \mathcal{T}_j(\mathbf{I}_n - \frac{\tau^2}{2} \widetilde{\mathbf{L}}_q \widetilde{\mathbf{L}}_q^T) \mathbf{e}_1 D_0^{1/2}, \quad j = 0, \dots, 2n-1. \quad (85)$$

We can make them independent of q by normalizing the measurements with D_0 , which is strictly positive by (29) and (30). Alternatively, we may suppose that $q(T)$ is known near $T=0$, and conclude from the causality of the wave equation that D_0 is independent of the variations of $q(T)$ at larger T . We make this assumption henceforth, and treat D_0 as constant.

2.4. The Galerkin–Petrov approximation

Here we show that the lower bidiagonal matrix $\tilde{\mathbf{L}}_q$ computed above is a Galerkin–Petrov approximation of the operator \mathcal{L}_q in (59), which in turn is an approximation of the Schrödinger operator \mathbf{L}_q .

Multiplying (82) on the right with the inverse of $\tilde{\mathbf{L}}_q^T$, denoted by $\tilde{\mathbf{L}}_q^{-T}$, we have

$$\tilde{\mathbf{L}}_q = \xi(\tilde{\mathcal{P}})\tilde{\mathbf{L}}_q^{-T} = \mathbf{V}^T \left[\frac{2}{\tau^2} (\mathbf{I}_N - \mathcal{P}) \right] \mathbf{V} \tilde{\mathbf{L}}_q^{-T} = \mathbf{V}^T \mathcal{L}_q \mathcal{L}_q^T \mathbf{V} \tilde{\mathbf{L}}_q^{-T}, \quad (86)$$

where we used definitions (62), (71) and $\mathbf{V}^T \mathbf{V} = \mathbf{I}_n$. We rewrite the result as

$$\tilde{\mathbf{L}}_q = \mathbf{V}^T \mathcal{L}_q \mathbf{W}, \quad (87)$$

using the $N \times n$ matrix

$$\mathbf{W} = \mathcal{L}_q^T \mathbf{V} \tilde{\mathbf{L}}_q^{-T}, \quad (88)$$

which is orthogonal by equation (82),

$$\mathbf{W}^T \mathbf{W} = \tilde{\mathbf{L}}_q^{-1} \mathbf{V}^T \mathcal{L}_q \mathcal{L}_q^T \mathbf{V} \tilde{\mathbf{L}}_q^{-T} = \tilde{\mathbf{L}}_q^{-1} \xi(\tilde{\mathcal{P}}) \tilde{\mathbf{L}}_q^{-T} = \mathbf{I}_n. \quad (89)$$

Thus, we conclude that $\tilde{\mathbf{L}}_q$ is the Galerkin–Petrov approximation of the operator \mathcal{L}_q , with the primary field approximated in the space (70), and the dual field approximated in the range of \mathbf{W} . This is the same as the range of the matrix $\mathbf{U} = (\mathbf{U}_0, \dots, \mathbf{U}_{n-1})$ of the dual snapshots, as explained in section 3.1.

Remark 2.1. It follows from (61) and the linearity of \mathbf{L}_q with respect to q that \mathcal{L}_q is approximately linear in q . The discussion at the end of section 2.3.1, which is for the columns of matrix \mathbf{V} , but extends verbatim to matrix \mathbf{W} , explains that \mathbf{V} and \mathbf{W} are almost independent of q . Thus, equation (87) yields approximate linearity of the reduced order matrix $\tilde{\mathbf{L}}_q$ in q , as needed in the DtB mapping.

2.5. The DtB mapping

Let $\tilde{\mathbf{L}}_{q^0}$ be the ROM matrix in the reference medium with known impedance $\sigma^0(x)$ and Schrödinger potential $q^0(x) = \ln \sigma^0(x)$. Let also \mathbf{V}^0 and \mathbf{W}^0 be the projection matrices in this medium. As explained in section 2.3.1 and remark 2.1, these matrices change slowly with the potential $q(x)$, so for the perturbed $q^\varepsilon(x)$ defined in (6) we have

$$\mathbf{V}^\varepsilon \approx \mathbf{V}^0, \quad \mathbf{W}^\varepsilon \approx \mathbf{W}^0. \quad (90)$$

Equation (87) gives

$$\tilde{\mathbf{L}}_{q^\varepsilon} \approx \tilde{\mathbf{L}}_{q^0} + \mathbf{V}^{0T} (\mathcal{L}_{q^\varepsilon} - \mathcal{L}_{q^0}) \mathbf{W}^0, \quad (91)$$

and due to the approximation (61) and the linearity of \mathbf{L}_q in q , we have

$$\mathcal{L}_{q^\varepsilon} - \mathcal{L}_{q^0} \approx \tilde{\mathbf{L}}_{q^\varepsilon} - \tilde{\mathbf{L}}_{q^0} = \varepsilon(\tilde{\mathbf{L}}_q - \tilde{\mathbf{L}}_{q^0}). \quad (92)$$

Substituting (92) in (91) we get the approximate linearity relation (8), which makes the mapping (11) useful.

Algorithm 2.1. The algorithm for computing the DtB mapping is as follows:

Input: data $\mathbf{D} = \{D_j\}_{j=0}^{2n-1}$.

1. Map the data to the ROM matrix $\tilde{\mathbf{L}}_q$ using equations (76), (79), (81) and the Cholesky factorizations (77) and (82).

2. Compute the data $\mathbf{D}^0 = \{D_j^0\}_{j=0}^{2n-1}$ in the reference medium with given q^0 , using formula

$$D_j^0 = \mathbf{b}^{0T} \cos\left(j\tau\sqrt{\tilde{\mathbf{L}}_{q^0}\tilde{\mathbf{L}}_{q^0}^T}\right)\mathbf{b}^0, \quad j = 0, \dots, 2n-1, \quad (93)$$

where \mathbf{b}^0 is the sensor vector in the reference medium and \mathbf{b}^{0T} is its transpose. Moreover, $\tilde{\mathbf{L}}_{q^0}$ is the $N \times N$ lower bidiagonal matrix, the discretization on the fine grid with N points of the operator (37) with reference potential q^0 .

3. Map the data \mathbf{D}^0 to the ROM matrix $\tilde{\mathbf{L}}_{q^0}$ using equations (76), (79), (81) and Cholesky factorizations (77), (82).

4. Use definition (85) and the chain rule given by algorithm 2.2 to calculate

$$\frac{d}{d\varepsilon} D_0^{-1/2} \tilde{M}_j(\tilde{\mathbf{L}}_\varepsilon \tilde{\mathbf{L}}_\varepsilon^T) D_0^{-1/2} \Big|_{\varepsilon=0} = \mathbf{e}_1^T \frac{d}{d\varepsilon} \mathcal{T}_j \left(\mathbf{I}_n - \frac{\tau^2}{2} \tilde{\mathbf{L}} \tilde{\mathbf{L}}^T \right) \Big|_{\varepsilon=0} \mathbf{e}_1 \quad (94)$$

where $\tilde{\mathbf{L}}^\varepsilon = \tilde{\mathbf{L}}_{q^0} + \varepsilon(\tilde{\mathbf{L}}_q - \tilde{\mathbf{L}}_{q^0})$ and $\tilde{\mathbf{L}}^{\varepsilon T}$ is its transpose. Recall that D_0 is the first measurement, at $t_0 = 0$.

Output: the transformed data $\mathcal{D}[\mathbf{D}]$ given by (11), with the derivative calculated from (94) and $\varepsilon = 1$.

The derivative in (94) must be calculated carefully, because the usual chain rule does not apply to matrix functions, unless the matrix commutes with its derivative. We obtain it in the next algorithm using the recursion relation (47).

Algorithm 2.2. The algorithm for computing (94) is as follows:

Input: $\tilde{\mathbf{L}}_q, \tilde{\mathbf{L}}_{q^0}$ and τ .

Denote $\tilde{\mathcal{P}}^\varepsilon = \mathbf{I}_n - \frac{\tau^2}{2} \tilde{\mathbf{L}}^\varepsilon \tilde{\mathbf{L}}^{\varepsilon T}$, $\tilde{\mathbf{P}}_j^\varepsilon = \mathcal{T}_j(\tilde{\mathcal{P}}^\varepsilon) \mathbf{e}_1$ and $\tilde{\mathbf{z}}_j = \frac{d}{d\varepsilon} \tilde{\mathbf{P}}_j^\varepsilon \Big|_{\varepsilon=0}$, $j = 0, \dots, 2n-1$.

We have $\tilde{\mathbf{z}}_0 = \mathbf{0}$ and $\tilde{\mathbf{z}}_1 = -\frac{\tau^2}{2} (\tilde{\mathbf{L}}_q \tilde{\mathbf{L}}_{q^0}^T + \tilde{\mathbf{L}}_{q^0} \tilde{\mathbf{L}}_q^T - 2\tilde{\mathbf{L}}_{q^0} \tilde{\mathbf{L}}_{q^0}^T) \mathbf{e}_1$, and for $j = 2, \dots, 2n-1$,

$$\tilde{\mathbf{z}}_j = 2\xi(\tilde{\mathcal{P}}^0) \tilde{\mathbf{z}}_{j-1} - \tilde{\mathbf{z}}_{j-2} - \tau^2 (\tilde{\mathbf{L}}_q \tilde{\mathbf{L}}_{q^0}^T + \tilde{\mathbf{L}}_{q^0} \tilde{\mathbf{L}}_q^T - 2\tilde{\mathbf{L}}_{q^0} \tilde{\mathbf{L}}_{q^0}^T) \tilde{\mathbf{P}}_{j-1}^0.$$

Output: The derivative (94) given by $\frac{d}{d\varepsilon} D_0^{-1/2} M_j(\tilde{\mathbf{L}}_\varepsilon \tilde{\mathbf{L}}_\varepsilon^T) D_0^{-1/2} \Big|_{\varepsilon=0} = \mathbf{e}_1^T \tilde{\mathbf{z}}_j$.

2.6. Numerical results

We present numerical results for a layered model, with relative acoustic impedance shown in figure 1. The relative impedance is defined as the ratio of the impedance and that of the homogeneous background. We display it as a function of the travel time, at steps τ chosen consistent

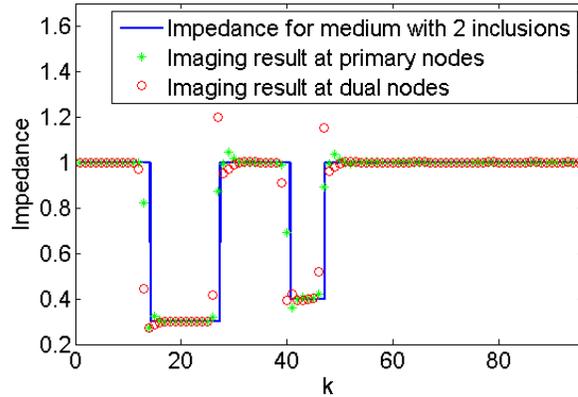


Figure 1. Layered relative acoustic impedance model $\sigma(x)$ and discrete inversion results. The discrete inversion is discussed in section 3.1. The abscissa is the spatial (primary) grid node index, sampled at interval τ .

with the Nyquist sampling rate of the Gaussian pulse used in the simulations. To avoid the ‘inverse crime’, the data are generated with a finite-difference time-domain algorithm, on an equidistant grid with steps much smaller than τ .

In figure 2 we show the primary snapshots, columns of \mathbf{P}^0 and \mathbf{P} , for the homogeneous background with relative impedance $\sigma^0 = 1$, and the layered model. The wave has crossed all the discontinuities of the impedance by the time $t = 80\tau$ of the snapshots displayed in the bottom row. Thus, we observe significant differences between \mathbf{P}_{80} and \mathbf{P}_{80}^0 . These consist of the decrease of the amplitude of the first arrival and the large multiple reflections.

In figure 2 we also show the columns of \mathbf{V}^0 and \mathbf{V} , for the homogeneous background and the layered model. We call these columns the primary orthonormal snapshots. We observe that they are almost independent of the medium, as discussed in remark 2.1. A similar behavior holds for the dual orthonormal snapshots, not shown here.

In figure 3 we show the raw scattering data, its Born approximation⁷ and the data obtained with the DtB algorithm. We observe that the strong multiples in the raw data are removed, and that the result is indistinguishable from the Born approximation.

Remark 2.2. Our experiments with different τ (not showed here) indicate that the l_∞ discrepancy between the true Born approximation and the output of the DtB algorithm decays as $O(\tau^2)$ for smooth $\sigma(x)$, in agreement with the approximation error in (59). We speculate that if we solved exactly for \mathbf{L}_q in (59), the discrepancy would decay exponentially in τ^{-1} .

3. A related inverse spectral problem

In this section we look in more detail at the entries of the ROM matrix $\tilde{\mathbf{L}}_q$, and compare it with another ROM obtained from spectral measurements of the operator (16) in the wave equation.

⁷The Born approximation cannot be computed in the inverse scattering problem, because it corresponds to solving the wave equation linearized with respect to the logarithm of the unknown impedance. This is why we need the DtB transform. We display the Born approximation only for comparison with the output of the DtB algorithm.

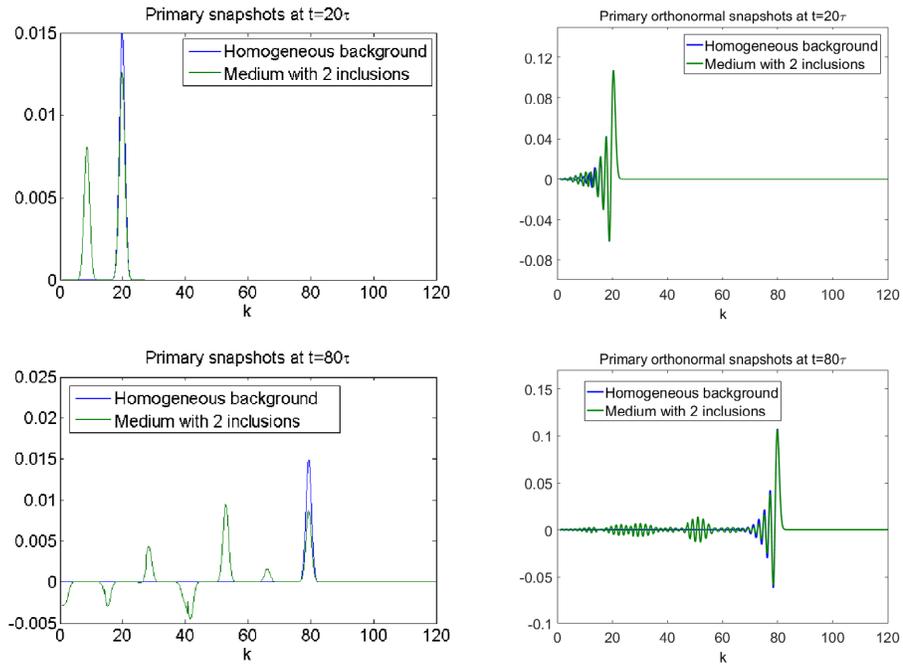


Figure 2. Primary snapshots P_j (left column) and primary orthonormal snapshots V_j (right column) at time index $j = 20$ (top row) and $j = 80$ (bottom row). The abscissa is the travel time index, sampled at interval τ .

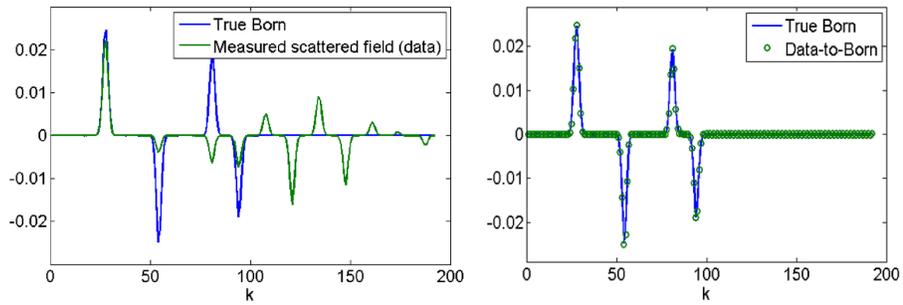


Figure 3. Left: the raw scattering data (green line) and the Born approximation (blue line). Right: the data transformed with the DtB mapping (green \circ) is indistinguishable from the Born approximation (blue line). The abscissa is the time index.

3.1. The orthonormal snapshots and the entries in $\tilde{\mathbf{L}}_q$

The first primary snapshots $\{\mathbf{P}_j\}_{j=0}^{n-1}$ span the Krylov space $\mathcal{K}_n(\mathbf{b}, \mathcal{P})$ defined in (70), and the dual snapshots $\{\mathbf{U}_j\}_{j=0}^{n-1}$ span the Krylov space $\mathcal{K}_n(\mathbf{U}_0, \mathcal{Q})$, as follows from definition (64).

The classical method for computing an orthonormal basis of a Krylov subspace is given by the Lanczos method [46], which is used in [23, algorithm 3.1] to calculate the orthogonal vectors⁸ $\{\bar{\mathbf{p}}_j\}_{j=1}^n$ and the orthogonal vectors $\{\bar{\mathbf{u}}_j\}_{j=1}^n$, satisfying the equations

$$\frac{\bar{\mathbf{p}}_{j+1} - \bar{\mathbf{p}}_j}{\gamma_j} = -\mathcal{L}_q \bar{\mathbf{u}}_j, \quad \frac{\bar{\mathbf{u}}_j - \bar{\mathbf{u}}_{j-1}}{\hat{\gamma}_j} = \mathcal{L}_q^T \bar{\mathbf{p}}_j, \quad (95)$$

for $j \geq 1$, with initial conditions $\bar{\mathbf{p}}_1 = \mathbf{b}$, $\bar{\mathbf{u}}_0 = 0$, and coefficients

$$\gamma_j = \frac{1}{\|\bar{\mathbf{u}}_j\|^2}, \quad \hat{\gamma}_j = \frac{1}{\|\bar{\mathbf{p}}_j\|^2}, \quad j = 1, \dots, n. \quad (96)$$

It is also shown in [23, sections 4.2 and 4.3] how the coefficients (96) enter the expression of the tridiagonal ROM propagator $\tilde{\mathcal{P}}$. Using the factorization (82), we obtain from those results the lower bidiagonal matrix $\tilde{\mathbf{L}}_q$ with entries

$$\begin{aligned} (\tilde{\mathbf{L}}_q)_{jj} &= -1/\sqrt{\gamma_j \hat{\gamma}_j}, \quad j = 1, \dots, n, \\ (\tilde{\mathbf{L}}_q)_{j+1,j} &= 1/\sqrt{\gamma_j \hat{\gamma}_{j+1}}, \quad j = 1, \dots, n-1. \end{aligned} \quad (97)$$

The columns of the projection matrix \mathbf{V} on the Krylov space $\mathcal{K}_n(\mathbf{b}, \mathcal{P})$, aka the orthonormal primary snapshots, are given by

$$\mathbf{V}_j = \sqrt{\hat{\gamma}_j} \bar{\mathbf{p}}_j, \quad j = 1, \dots, n, \quad (98)$$

and the projection matrix \mathbf{W} satisfies by definition (88)

$$\mathbf{W} \tilde{\mathbf{L}}_q^T = \mathcal{L}_q^T \mathbf{V}. \quad (99)$$

The j th column in this equation reads

$$\frac{1}{\sqrt{\hat{\gamma}_j}} \mathbf{W} \tilde{\mathbf{L}}_q^T \mathbf{e}_j = \mathcal{L}_q^T \bar{\mathbf{p}}_j = \frac{\bar{\mathbf{u}}_j - \bar{\mathbf{u}}_{j-1}}{\hat{\gamma}_j}, \quad (100)$$

where we used (95) and (98). The left hand side is a linear combination of the columns \mathbf{W}_j and \mathbf{W}_{j-1} of \mathbf{W} , because $\tilde{\mathbf{L}}_q$ is lower bidiagonal. Using the expression (97) of the entries in $\tilde{\mathbf{L}}_q$ and equation (98), we conclude that

$$\mathbf{W} = (\mathbf{W}_1, \dots, \mathbf{W}_n), \quad \mathbf{W}_j = \sqrt{\gamma_j} \bar{\mathbf{u}}_j. \quad (101)$$

This shows that \mathbf{W} is the matrix of orthonormal dual snapshots, as stated in the previous section.

Let us take the constant reference impedance $\sigma^0 = 1$, corresponding to the potential $q^0 = 0$, and define the coefficients

$$\sigma_j = \frac{\hat{\gamma}_j^0}{\hat{\gamma}_j} = \frac{\|\bar{\mathbf{p}}_j\|^2}{\|\bar{\mathbf{p}}_j^0\|^2}, \quad \hat{\sigma}_j = \frac{\gamma_j}{\gamma_j^0} = \frac{\|\bar{\mathbf{u}}_j^0\|^2}{\|\bar{\mathbf{u}}_j\|^2}. \quad (102)$$

With these coefficients we introduce the discrete Liouville transform

⁸ The vectors $\bar{\mathbf{p}}_j$ and $\bar{\mathbf{u}}_j$ are called orthogonalized primary and dual snapshots in [23].

$$\bar{\mathbf{P}}_j = \frac{\bar{\mathbf{P}}_j}{\sqrt{\hat{\sigma}_j}}, \quad \bar{\mathbf{U}}_j = \sqrt{\hat{\sigma}_j} \bar{\mathbf{u}}_j, \quad (103)$$

substitute it in (95) and obtain after straightforward algebraic manipulations the system

$$\frac{\bar{\mathbf{P}}_{j+1} - \bar{\mathbf{P}}_j}{\hat{\gamma}_j^0} + \bar{\mathbf{P}}_{j+1} \left(\frac{\sqrt{\hat{\sigma}_{j+1}} - \sqrt{\hat{\sigma}_j}}{\hat{\gamma}_j^0 \sqrt{\hat{\sigma}_j}} \right) + \bar{\mathbf{P}}_j \left(\frac{\sqrt{\hat{\sigma}_j} - \sqrt{\hat{\sigma}_j}}{\hat{\gamma}_j^0 \sqrt{\hat{\sigma}_j}} \right) = -\mathcal{L}_q \bar{\mathbf{U}}_j, \quad (104)$$

$$\frac{\bar{\mathbf{U}}_j - \bar{\mathbf{U}}_{j-1}}{\hat{\gamma}_j^0} - \bar{\mathbf{U}}_j \left(\frac{\sqrt{\hat{\sigma}_j} - \sqrt{\hat{\sigma}_j}}{\hat{\gamma}_j^0 \sqrt{\hat{\sigma}_j}} \right) - \bar{\mathbf{U}}_{j-1} \left(\frac{\sqrt{\hat{\sigma}_j} - \sqrt{\hat{\sigma}_{j-1}}}{\hat{\gamma}_j^0 \sqrt{\hat{\sigma}_{j-1}}} \right) = \mathcal{L}_q^T \bar{\mathbf{P}}_j, \quad (105)$$

for $j = 1, \dots, n$. Recalling the approximation (61), we see that the finite difference operators in the left hand sides of equations (104) and (105) can be interpreted as discretizations of $L_q^T = \partial_T + \frac{1}{2} \partial_T q(T)$ and $-L_q = \partial_T - \frac{1}{2} \partial_T q(T)$, with $q(T) = \ln \sigma(T)$. The discretization is on a special grid with primary points spaced at $\hat{\gamma}_j^0$, and dual points spaced at $\hat{\gamma}_j^0$. In our case these equal τ [23]. In figure 2 the primary grid corresponds to the integer values in the abscissae, and the dual grid points are offset by $\tau/2$. We note that the peaks of the primary orthonormal snapshots \mathbf{V}_j are approximately aligned with the j th primary grid point, which is the location of the wavefront.

The coefficients σ_j and $\hat{\sigma}_j$ are approximations of the impedance at the primary and dual grid points, and the terms in the parentheses in (104) and (105) are discretizations of

$$\frac{1}{2} \partial_T q(T) = \frac{\partial_T \sqrt{\sigma(T)}}{\sqrt{\sigma(T)}}.$$

We display in figure 1 the values σ_j and $\hat{\sigma}_j$ computed for the impedance model considered in section 2.5. They give a reasonable approximation of the discontinuous impedance, but better results can be obtained by inverting the data processed by the DtB algorithm, which is almost indistinguishable from the Born approximation.

We show next that the same approximation formulas (102) arise for another ROM, constructed from different measurement functions of the operator $L_q L_q^T$ than in (65). The values of the ROM coefficients $\{\gamma_j, \hat{\gamma}_j\}_{j=1}^n$ are different, but the same ratios $\{\hat{\gamma}_j^0 / \hat{\gamma}_j, \gamma_j / \gamma_j^0\}_{j=1}^n$ define approximations of $\sigma(T)$ on the ROM dependent grids with primary and dual point spacings defined by $\{\gamma_j^0, \hat{\gamma}_j^0\}_{j=1}^n$. With this ROM described below, it is proved in [9] that the approximations (102) converge to the unknown impedance function $\sigma(T)$ in the limit $n \rightarrow \infty$.

3.2. The spectrally matched ROM

In this section we draw an analogy between the ROM constructed from the data matching conditions (5) and the ‘spectrally matched’ ROM introduced and analyzed in [9]. Spectrally matched means that the ROM defines a three point finite difference scheme in x for the wave equation satisfied by the pressure field $p(t, x)$ in (26), modeled with an $n \times n$ tridiagonal matrix constructed from the truncated spectral measure of the differential operator A in (16).

The Laplace transform of $p(t, x)$ with respect to time t , written in the travel time coordinates (36),

$$\check{p}(s, T) = \int_0^\infty p(t, x(T)) e^{-st} ds, \quad (106)$$

satisfies the boundary value problem

$$(A_q + s^2)\check{p}(s, T) = s\sqrt{\sigma(T)}b(T) \text{ for } T \in (0, T_\ell), \quad \partial_T \check{p}(s, 0) = \check{p}(s, T_\ell) = 0. \quad (107)$$

Here we wrote the operator in (16) in the travel time coordinates

$$A_q = -\sigma(T)\partial_T \left(\frac{1}{\sigma(T)} \partial_T \right) = -\partial_T^2 + \partial_T q(T) \partial_T, \quad (108)$$

with $q(T) = \ln \sigma(T)$. This is the formulation considered in [9], and the spectral measure of A_q is defined by its eigenvalues $\{\lambda_j > 0\}_{j \geq 1}$ and $\{\zeta_j = y_j^2(0)\}_{j \geq 1}$, where $\{y_j(T)\}_{j \geq 1}$ are the eigenfunctions, orthonormal with respect to the weighted inner product $\langle \cdot, \cdot \rangle_{1/\sigma}$.

The spectrally matched ROM is defined in [9] by an $n \times n$ tridiagonal matrix with spectral measure defined by $\{\lambda_j, \zeta_j\}_{j=1}^n$. To compare it with the ROM defined in section 2.3, let us consider the Liouville transform

$$\check{P}(s, T) = \frac{\check{p}(s, T)}{\sqrt{\sigma(T)}}, \quad (109)$$

and suppose that $\sigma(T)$ is constant in a vicinity of $T = 0$. Then, $\check{P}(s, T)$ satisfies

$$(L_q L_q^T + s^2)\check{P}(s, T) = sb(T) \text{ for } T \in (0, T_\ell), \quad \partial_T \check{P}(s, 0) = \check{P}(s, T_\ell) = 0, \quad (110)$$

with L_q and L_q^T defined in (37) and (38). Note that $L_q L_q^T$ is related to A_q by a similarity transformation,

$$L_q L_q^T = \sigma(T)^{-1/2} A_q \sigma(T)^{1/2}, \quad (111)$$

so it has the same eigenvalues λ_j . The eigenfunctions $z_j(T) = \sigma(T)^{-1/2} y_j(T)$ are orthonormal with respect to the Euclidian $L^2((0, T_\ell))$ inner product. Thus, the spectral measure of $L_q L_q^T$ is the same as that of A_q , up to the multiplicative constant $1/\sigma(0)$, which we take equal to 1.

The measurement functions in the data model (4) are now

$$M_j(L_q L_q^T) = \begin{cases} \lambda_{j+1}, & j = 0, \dots, n-1 \\ z_{j-n+1}^2(0), & j = n, \dots, 2n-1. \end{cases} \quad (112)$$

The ROM is defined by the $n \times n$ symmetric, positive definite and tridiagonal matrix $\tilde{\mathcal{A}}_q$ in the finite difference discretization of (110) on a special grid with n points in $[0, T_\ell]$,

$$(\tilde{\mathcal{A}}_q + s^2)\tilde{P}(s) = s\|b\|\mathbf{e}_1. \quad (113)$$

Let

$$\tilde{\mathcal{A}}_q = \tilde{L}_q \tilde{L}_q^T, \quad (114)$$

be the Cholesky factorization of this matrix, with lower bidiagonal \tilde{L}_q . Let also $\tilde{\lambda}_j$ and \tilde{z}_j be the eigenvectors of $\tilde{\mathcal{A}}_q$, with Euclidean norm $\|\tilde{z}_j\| = 1$. The matrix $\tilde{\mathcal{A}}_q$ is obtained from the matching conditions

$$M_j(L_q L_q^T) = \tilde{M}_j(\tilde{L}_q \tilde{L}_q^T) = \begin{cases} \tilde{\lambda}_{j+1}, & j = 0, \dots, n-1 \\ (\mathbf{e}_1^T \tilde{z}_j)^2, & j = n, \dots, 2n-1, \end{cases} \quad (115)$$

using the Lanczos algorithm [16]. The resulting lower bidiagonal matrix \tilde{L}_q has the entries

$$(\tilde{\mathbf{L}}_q)_{jj} = -\frac{1}{\sqrt{\gamma_j \hat{\gamma}_j}}, \quad 1 \leq j \leq n, \quad (\tilde{\mathbf{L}}_q)_{j+1,j} = \frac{1}{\sqrt{\gamma_j \hat{\gamma}_{j+1}}}, \quad 1 \leq j \leq n-1, \quad (116)$$

that have the same expression as in (97), but the values of $\{\gamma_j, \hat{\gamma}_j\}_{j=1}^n$ are different.

3.3. Inversion on the spectrally matched grid

With the factorization (114) we can rewrite (113) as the first order system

$$s\tilde{\mathbf{U}}(s) = \tilde{\mathbf{L}}_q^T \tilde{\mathbf{P}}(s), \quad s\tilde{\mathbf{P}}(s) = -\tilde{\mathbf{L}}_q \tilde{\mathbf{U}}(s) + \|b\| \mathbf{e}_1, \quad (117)$$

for the ROM primary and dual vectors $\tilde{\mathbf{P}}(s)$ and $\tilde{\mathbf{U}}(s)$. We now show that the entries in these vectors represent discretizations of the Laplace transforms $\check{\mathbf{P}}(s, T)$ and $\check{\mathbf{U}}(s, T)$ of the primary and dual fields in equation (30), rewritten in travel time coordinates. The discretization grid is defined by the spectrally matched ROM coefficients $\{\gamma_j^0, \hat{\gamma}_j^0\}_{j=1}^n$ calculated in the reference medium with constant impedance $\sigma^0 = 1$ i.e. potential $q^0 = 0$. It is proved in [9, lemma 3.2] that these coefficients define a staggered grid in the interval $[0, T_\ell]$,

$$0 = \hat{T}_0 = T_1 < \hat{T}_1 < T_2 < \hat{T}_2 < \dots < \hat{T}_n < T_{n+1}, \quad (118)$$

with T_{n+1} approaching T_ℓ from below in the limit $n \rightarrow \infty$. The primary field $\check{\mathbf{P}}(s, T)$ is discretized on the grid with points $\{T_j\}_{j=1}^{n+1}$, spaced at intervals $h_j = \gamma_j^0 = T_{j+1} - T_j$, and the dual field $\check{\mathbf{U}}(s, T)$ is discretized on the grid with points $\{\hat{T}_j\}_{j=0}^n$, spaced at intervals $\hat{h}_j = \hat{\gamma}_j^0 = \hat{T}_j - \hat{T}_{j-1}$, for $j = 1, \dots, n$.

Let us define the diagonal matrices

$$H^{1/2} = \text{diag}(h_1^{1/2}, \dots, h_n^{1/2}), \quad \hat{H}^{1/2} = \text{diag}(\hat{h}_1^{1/2}, \dots, \hat{h}_n^{1/2}),$$

and write the vectors in (117) in the form

$$\tilde{\mathbf{P}}(s) = \hat{H}^{1/2} \begin{pmatrix} \check{P}_{\hat{\tau}_1}(s) \\ \vdots \\ \check{P}_{\hat{\tau}_n}(s) \end{pmatrix}, \quad \tilde{\mathbf{U}}(s) = H^{1/2} \begin{pmatrix} \check{U}_{\hat{\tau}_1}(s) \\ \vdots \\ \check{U}_{\hat{\tau}_n}(s) \end{pmatrix}, \quad (119)$$

so that the first equation in the system (117) becomes

$$s \begin{pmatrix} \check{U}_{\hat{\tau}_1}(s) \\ \vdots \\ \check{U}_{\hat{\tau}_n}(s) \end{pmatrix} = H^{-1/2} \tilde{\mathbf{L}}_q^T \hat{H}^{1/2} \begin{pmatrix} \check{P}_{\hat{\tau}_1}(s) \\ \vdots \\ \check{P}_{\hat{\tau}_n}(s) \end{pmatrix}. \quad (120)$$

Let also

$$\sigma_j = \frac{\hat{h}_j}{\hat{\gamma}_j} = \frac{\hat{\gamma}_j^0}{\hat{\gamma}_j}, \quad \hat{\sigma}_j = \frac{\gamma_j}{h_j} = \frac{\gamma_j}{\gamma_j^0}, \quad j = 1, \dots, n, \quad (121)$$

and write explicitly the j th equation in (120)

$$s \check{U}_{\hat{\tau}_j}(s) = \frac{\check{P}_{\hat{\tau}_{j+1}}(s) - \check{P}_{\hat{\tau}_j}(s)}{h_j} + \check{P}_{\hat{\tau}_{j+1}}(s) \left(\frac{\sqrt{\sigma_{j+1}} - \sqrt{\hat{\sigma}_j}}{h_j \sqrt{\hat{\sigma}_j}} \right) + \check{P}_{\hat{\tau}_j}(s) \left(\frac{\sqrt{\hat{\sigma}_j} - \sqrt{\sigma_j}}{h_j \sqrt{\sigma_j}} \right).$$

This is the discretization of equation

$$s\check{U}(s, T) = L_q^T \check{P}(s, T) = \partial_T \check{P}(s, T) + \partial_T \ln \sqrt{\sigma(T)} \check{P}(s, T),$$

on the spectrally matched grid. A similar result applies to the second equation in (117), which is the discretization of

$$s\check{P}(s, T) = -L_q \check{U}(s, T) + b(T).$$

The ratios (121) are of the same form as in (102) and represent approximations of the impedance function $\sigma(T)$ on the spectrally matched grid (118). Specifically, it is proved in [9, theorem 6.1] that if $\sigma(T)$ is smooth, then the piecewise constant function

$$\sigma^n(T) = \begin{cases} \sigma_j & \text{for } T \in [T_j, \widehat{T}_j), j = 1, \dots, n, \\ \widehat{\sigma}_j & \text{for } T \in [\widehat{T}_j, T_{j+1}), j = 1, \dots, n, \\ \widehat{\sigma}_n & \text{for } T \in [T_{n+1}, T_\ell], \end{cases}$$

converges to $\sigma(T)$ pointwise and in $L^1[0, T_\ell]$, as $n \rightarrow \infty$.

4. The multi dimensional case

In this section we generalize the DtB mapping from one dimension, as described in section 2, to \mathbb{R}^d with $d > 1$. The derivation follows the same strategy as in section 2, with certain modifications described below.

4.1. Data model for an array of sensors

In the multi-dimensional case we consider an array of m sensors on the accessible boundary $\partial\Omega_a$, located at points $\mathbf{x}_s \in \partial\Omega_a$, for $s = 1, \dots, m$. Each sensor excites a pressure field, the solution of the wave equation

$$(\partial_t^2 + A)\mathbf{p}_s(t, \mathbf{x}) = \partial_t f(t) \delta(\mathbf{x} - \mathbf{x}_s), \quad -\infty < t < \infty, \quad \mathbf{x} \in \Omega, \quad (122)$$

where the index s denotes the source and the operator A is now defined as

$$A = -\sigma(\mathbf{x})c(\mathbf{x})\nabla \cdot \left[\frac{c(\mathbf{x})}{\sigma(\mathbf{x})} \nabla \right], \quad (123)$$

with ∇ denoting the gradient and $\nabla \cdot$ the divergence operator. For simplicity, we assume that the same pulse $\partial_t f(t)$ is emitted from all the sensors. The boundary conditions at $\partial\Omega_a$ are $\nabla \mathbf{p}_s(t, \mathbf{x}) \cdot \boldsymbol{\nu}(\mathbf{x}) = 0$, where $\boldsymbol{\nu}(\mathbf{x})$ is the outer unit normal, and on the inaccessible boundary $\partial\Omega_i$ we let $\mathbf{p}_s(t, \mathbf{x}) = 0$. The medium is at rest initially, so we set $\mathbf{p}_s(t, \mathbf{x}) = 0$, for $t \ll 0$.

Following the same argument that lead to equation (22), we define the matrix-valued data $\mathbf{D}_k \in \mathbb{R}^{m \times m}$ with entries defined by

$$(\mathbf{D}_k)_{ij} = \mathbf{p}_i(t_k, \mathbf{x}_j) + \mathbf{p}_j(-t_k, \mathbf{x}_i), \quad i, j = 1, \dots, m, \quad (124)$$

in terms of the measurements at the instances $t_k = k\tau$ for $k = 0, 1, \dots, 2n - 1$. For each k the matrix \mathbf{D}_k is symmetric due to the source-receiver reciprocity.

4.2. First order system form and Liouville transformation

Similar to the one-dimensional case, we introduce the sensor functions

$$b_s(\mathbf{x}) = \sqrt{\frac{\sigma(\mathbf{x}_s)c(\mathbf{x}_s)}{\sigma(\mathbf{x})}} [\widehat{f}(\sqrt{A})]^{1/2} \delta(\mathbf{x} - \mathbf{x}_s), \quad \mathbf{x} \in \Omega, \quad s = 1, \dots, m, \quad (125)$$

and define the analogue of (25)

$$p_s(t, \mathbf{x}) = \cos(t\sqrt{A}) \sqrt{\sigma(\mathbf{x})} b_s(\mathbf{x}), \quad s = 1, \dots, m. \quad (126)$$

Here $p_s(t, \mathbf{x})$ is the pressure field in the first order system

$$\partial_t \begin{pmatrix} p_s(t, \mathbf{x}) \\ -\mathbf{u}_s(t, \mathbf{x}) \end{pmatrix} = \begin{pmatrix} 0 & \sigma(\mathbf{x})c(\mathbf{x})\nabla \cdot \\ \frac{c(\mathbf{x})}{\sigma(\mathbf{x})}\nabla & 0 \end{pmatrix} \begin{pmatrix} p_s(t, \mathbf{x}) \\ -\mathbf{u}_s(t, \mathbf{x}) \end{pmatrix}, \quad t > 0, \quad \mathbf{x} \in \Omega, \quad (127)$$

with initial conditions

$$p_s(0, \mathbf{x}) = \sqrt{\sigma(\mathbf{x})} b_s(\mathbf{x}) \quad \text{and} \quad \mathbf{u}_s(0, \mathbf{x}) = 0, \quad \mathbf{x} \in \Omega, \quad (128)$$

and boundary conditions

$$p_s(t, \mathbf{x})|_{\mathbf{x} \in \partial\Omega_i} = 0 \quad \text{and} \quad \boldsymbol{\nu}(\mathbf{x}) \cdot \mathbf{u}_s(t, \mathbf{x})|_{\mathbf{x} \in \partial\Omega_a} = 0, \quad t > 0, \quad (129)$$

where we recall that $\boldsymbol{\nu}(\mathbf{x})$ is the outer unit normal at $\partial\Omega_a$. This system is the analogue of (26) and the vector field $\mathbf{u}_s(t, \mathbf{x})$ is the particle velocity.

Using the Liouville transformation

$$P(t, \mathbf{x}; \mathbf{x}_s) = \frac{p(t, \mathbf{x}; \mathbf{x}_s)}{\sqrt{\sigma(\mathbf{x})}}, \quad \mathbf{U}(t, \mathbf{x}; \mathbf{x}_s) = -\sqrt{\sigma(\mathbf{x})} \mathbf{u}(t, \mathbf{x}; \mathbf{x}_s), \quad (130)$$

we rewrite (127) as a Schrödinger system (1) with the operators L_q and L_q^T given by

$$L_q = -c(\mathbf{x})\nabla \cdot + \frac{1}{2}c(\mathbf{x})\nabla q(\mathbf{x}) \cdot \quad \text{and} \quad L_q^T = c(\mathbf{x})\nabla + \frac{1}{2}c(\mathbf{x})\nabla q(\mathbf{x}), \quad (131)$$

and the same $q(\mathbf{x}) = \ln \sigma(\mathbf{x})$. The transformed boundary and initial conditions take the form (2) and (3), and the entries of the data matrix $\mathbf{D}_k \in \mathbb{R}^{m \times m}$, for $k = 0, \dots, 2n - 1$, are expressed in terms of the primary wave $P_i(t, \mathbf{x})$ as

$$(\mathbf{D}_k)_{ij} = \langle b_j, P_i(t_k, \cdot) \rangle_{\frac{1}{c(\mathbf{x})}} = \int_{\Omega} \frac{d\mathbf{x}}{c(\mathbf{x})} b_j(\mathbf{x}) P_i(t_k, \mathbf{x}), \quad i, j = 1, \dots, m. \quad (132)$$

4.3. Symmetrized data model, propagator and measurement function

In one dimension we used travel time coordinates to write the data model in the symmetrized form. Because such a transformation is not available in higher dimensions, we follow a different approach to symmetrize (132) and thus obtain an analogue of (44).

Combining (132) with (126) and (130) we write

$$(\mathbf{D}_k)_{ij} = \int_{\Omega} \frac{d\mathbf{x}}{c(\mathbf{x})} b_j(\mathbf{x}) \sigma^{-1/2}(\mathbf{x}) \cos(t_k \sqrt{A}) \sigma^{1/2}(\mathbf{x}) b_i(\mathbf{x}), \quad (133)$$

where we assume for the remainder of this section $i, j = 1, \dots, m$, and $k = 0, \dots, 2n - 1$. From the definition (131) of L_q it follows that

$$A = \sigma^{1/2}(\mathbf{x}) L_q L_q^T \sigma^{-1/2}(\mathbf{x}). \quad (134)$$

Hence,

$$(\mathbf{D}_k)_{i,j} = \int_{\Omega} \frac{d\mathbf{x}}{c(\mathbf{x})} b_j(\mathbf{x}) \cos\left(t_k \sqrt{L_q L_q^T}\right) b_i(\mathbf{x}), \quad (135)$$

where we used that analytic matrix functions commute with similarity transformations. We use another similarity transformation to rewrite (135) as

$$(\mathbf{D}_k)_{i,j} = \int_{\Omega} d\mathbf{x} b_j^c(\mathbf{x}) \cos\left(t_k \sqrt{c^{-1/2} L_q L_q^T c^{1/2}}\right) b_i^c(\mathbf{x}), \quad (136)$$

with the rescaled sensor functions

$$b_s^c(\mathbf{x}) = c^{-1/2}(\mathbf{x}) b_s(\mathbf{x}), \quad s = 1, \dots, m. \quad (137)$$

We also define the rescaled operators

$$L_q^c = c^{-1/2}(\mathbf{x}) L_q c^{1/2}(\mathbf{x}) = c^{1/2}(\mathbf{x}) \left(-\nabla \cdot + \frac{1}{2} \nabla q(\mathbf{x}) \cdot \right) c^{1/2}(\mathbf{x}), \quad (138)$$

$$L_q^{cT} = c^{-1/2}(\mathbf{x}) L_q^T c^{1/2}(\mathbf{x}) = c^{1/2}(\mathbf{x}) \left(\nabla + \frac{1}{2} \nabla q(\mathbf{x}) \right) c^{1/2}(\mathbf{x}), \quad (139)$$

which are adjoint to each other with respect to the standard $L^2(\Omega)$ inner product. These operators define the propagator for the multi-dimensional problem as in (42),

$$\mathcal{P} = \cos\left(\tau \sqrt{L_q^c L_q^{cT}}\right). \quad (140)$$

The data model (136) is now in symmetric form, and the measurement functions $\mathbf{M}_k(L_q L_q^T) \in \mathbb{R}^{m \times m}$ are defined component-wise by

$$(\mathbf{D}_k)_{i,j} = [\mathbf{M}_k(L_q L_q^T)]_{i,j} = \langle b_j^c(\mathbf{x}), \mathcal{T}_k(\mathcal{P}) b_i^c(\mathbf{x}) \rangle. \quad (141)$$

Similar to the one-dimensional case, the propagator \mathcal{P} can be used to define an exact time stepping scheme for

$$P_{k,s}(\mathbf{x}) = \mathcal{T}_k(\mathcal{P}) b_s^c(\mathbf{x}), \quad k = 0, \dots, 2n-1, \quad s = 1, \dots, m. \quad (142)$$

Here we use the convention that the first index denotes the discrete time instance and the second index denotes the source. We obtain the same second-order time stepping scheme (48),

$$\frac{1}{\tau^2} \left[P_{k+1,s}(\mathbf{x}) - 2P_{k,s}(\mathbf{x}) + P_{k-1,s}(\mathbf{x}) \right] = -\xi(\mathcal{P}) P_{k,s}(\mathbf{x}), \quad k = 0, \dots, 2n-1, \quad (143)$$

with the affine function ξ given by (50) and the initial conditions

$$P_{0,s}(\mathbf{x}) = b_s^c(\mathbf{x}), \quad P_{-1,s}(\mathbf{x}) = P_{1,s}(\mathbf{x}) = \mathcal{P} b_s^c(\mathbf{x}). \quad (144)$$

We also have the same factorization (54) of $\xi(\mathcal{P})$ in terms of the operator \mathcal{L}_q defined as in (56), with L_q and L_q^T replaced by L_q^c and L_q^{cT} .

4.4. Multi-input, multi-output reduced order model

The main difference between one and multi dimensions is the type of ROM that we use. In one dimension we had a single-input, single output (SISO) projection ROM (79), obtained from matching the ROM output (85) to the scalar valued data (67). In multi dimensions we need a multi-input, multi-output (MIMO) ROM that matches the matrix valued data (141).

As in section 2.3, let us introduce the matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$, a discretization of the operator (123) on a very fine, uniform grid with a total of N nodes and step size h . Note that the operator (123) is related to the Schrödinger operators (138) and (139) as

$$\mathbf{A} = \sigma^{1/2}(\mathbf{x}) \mathbf{L}_q \mathbf{L}_q^T \sigma^{-1/2}(\mathbf{x}) = [\sigma(\mathbf{x})c(\mathbf{x})]^{1/2} \mathbf{L}_q^c \mathbf{L}_q^{cT} [\sigma(\mathbf{x})c(\mathbf{x})]^{-1/2}. \quad (145)$$

Let $\mathbf{\Sigma} \in \mathbb{R}^{N \times N}$ and $\mathbf{C} \in \mathbb{R}^{N \times N}$ be the diagonal matrices with entries given by $\sigma(\mathbf{x})$ and $c(\mathbf{x})$ evaluated at the fine grid nodes. Then we can set \mathbf{L}_q , a fine grid approximation of \mathbf{L}_q^c , to be a Cholesky factor of

$$\mathbf{\Sigma}^{-1/2} \mathbf{C}^{-1/2} \mathbf{A} \mathbf{\Sigma}^{1/2} \mathbf{C}^{1/2} = \mathbf{L}_q \mathbf{L}_q^T, \quad (146)$$

where we drop the index c on \mathbf{L}_q to simplify notation. We assume as in one dimension that the discretized propagator $\mathcal{P} \in \mathbb{R}^{N \times N}$ is well approximated by (62) on the fine grid.

As in one dimension, we call the fine grid discretization of the field P at the measurement instances ‘the primary snapshots’. It is convenient to arrange these into matrices $\mathbf{P}_k \in \mathbb{R}^{N \times m}$, for $k = 0, \dots, 2n - 1$, with each column corresponding to a different sensor. These matrices satisfy a fine grid analogue of the time stepping scheme (143),

$$\frac{1}{\tau^2} (\mathbf{P}_{k+1} - 2\mathbf{P}_k + \mathbf{P}_{k-1}) = -\xi(\mathcal{P})\mathbf{P}_k, \quad k = 0, \dots, 2n - 1, \quad (147)$$

with initial conditions $\mathbf{P}_0 = \mathbf{b}$ and $\mathbf{P}_{-1} = \mathbf{P}_1 = \mathcal{P}\mathbf{b}$. The sensor matrix $\mathbf{b} \in \mathbb{R}^{N \times m}$ is

$$\mathbf{b} = [\mathbf{b}_1, \dots, \mathbf{b}_m], \quad (148)$$

where the entries of each column $\mathbf{b}_s \in \mathbb{R}^N$ are the values of the rescaled sensor function $b_s^c(\mathbf{x})$ evaluated on the fine grid, multiplied by $h^{d/2}$.

The discretized data model and the measurement functions are given by

$$\mathbf{D}_k = \mathbf{M}_k(\mathbf{L}_q \mathbf{L}_q^T) \approx \mathbf{M}_k(\mathbf{L}_q \mathbf{L}_q^T) = \mathbf{b}^T \mathbf{P}_k = \mathbf{b}^T \mathcal{T}_k(\mathcal{P})\mathbf{b} \in \mathbb{R}^{m \times m}, \quad (149)$$

for $k = 0, \dots, 2n - 1$. The analogue of equation (66), which relates the data \mathbf{D}_0 at the first time instant to the sensor matrix is

$$\mathbf{D}_0 = \begin{bmatrix} \langle b_1^c(\mathbf{x}), b_1^c(\mathbf{x}) \rangle & \langle b_1^c(\mathbf{x}), b_2^c(\mathbf{x}) \rangle & \dots & \langle b_1^c(\mathbf{x}), b_m^c(\mathbf{x}) \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle b_m^c(\mathbf{x}), b_1^c(\mathbf{x}) \rangle & \langle b_m^c(\mathbf{x}), b_2^c(\mathbf{x}) \rangle & \dots & \langle b_m^c(\mathbf{x}), b_m^c(\mathbf{x}) \rangle \end{bmatrix} \approx \mathbf{b}^T \mathbf{b}. \quad (150)$$

As we did in one dimension, we neglect henceforth the fine grid discretization errors and treat the approximate relations (149) and (150) as equalities.

The MIMO ROM consists of the symmetric matrix $\tilde{\mathcal{P}} \in \mathbb{R}^{nm \times nm}$, called the ROM propagator, and the ROM sensor matrix $\tilde{\mathbf{b}} \in \mathbb{R}^{nm \times m}$, satisfying the data matching conditions

$$\mathbf{D}_k = \tilde{\mathbf{b}}^T \mathcal{T}_k(\tilde{\mathcal{P}})\tilde{\mathbf{b}}, \quad k = 0, \dots, 2n - 1. \quad (151)$$

The matrix $\tilde{\mathcal{P}}$ is block tridiagonal, with $m \times m$ blocks, while $\tilde{\mathbf{b}}$ has all zeros except for the uppermost $m \times m$ block. Using the $m \times m$ identity matrix \mathbf{I}_m and zero matrix $\mathbf{0}_m$, we write

$$\tilde{\mathbf{b}} = \mathbf{E}_1(\mathbf{D}_0)^{1/2}, \quad \text{with } \mathbf{E}_1 = \begin{bmatrix} \mathbf{I}_m \\ \mathbf{0}_m \\ \vdots \\ \mathbf{0}_m \end{bmatrix} \in \mathbb{R}^{nm \times m}. \quad (152)$$

4.5. Calculation of the projection MIMO ROM

The MIMO ROM satisfying the data matching conditions (151) is given by the orthogonal projection of \mathcal{P} on the block Krylov subspace, spanned by the columns of

$$\mathbf{P} = [\mathbf{P}_0, \dots, \mathbf{P}_{n-1}] \in \mathbb{R}^{N \times nm}.$$

We follow the notation in section 2.3.1, and let $\mathbf{V} = [\mathbf{V}_1, \dots, \mathbf{V}_n]$ be the matrix containing the orthonormal basis for $\text{range}(\mathbf{P})$. Here $\mathbf{V} \in \mathbb{R}^{N \times nm}$ and each \mathbf{V}_k is an $N \times m$ matrix.

To compute the matrices $\mathbf{P}^T \mathbf{P}$ and $\mathbf{P}^T \mathcal{P} \mathbf{P}$ from the data we can still use the formulas (76) and (81), however, the indexing is understood block-wise. Thus, when we write

$$(\mathbf{P}^T \mathbf{P})_{i,j} = \frac{1}{2} (\mathbf{D}_{i+j-2} + \mathbf{D}_{|i-j|}) \in \mathbb{R}^{m \times m}, \quad (153)$$

$$(\mathbf{P}^T \mathcal{P} \mathbf{P})_{i,j} = \frac{1}{4} (\mathbf{D}_{i+j-1} + \mathbf{D}_{|j-i+1|} + \mathbf{D}_{|j-i-1|} + \mathbf{D}_{|j+i-3|}) \in \mathbb{R}^{m \times m}, \quad (154)$$

for $i, j = 1, \dots, n$, we use the notation $(\mathbf{P}^T \mathbf{P})_{i,j}$ for the $m \times m$ block of $\mathbf{P}^T \mathbf{P} \in \mathbb{R}^{nm \times nm}$, at the intersection of rows $(i-1)m+1, \dots, im$ and columns $(j-1)m+1, \dots, jm$. We use this notation for $\mathbf{P}^T \mathcal{P} \mathbf{P} \in \mathbb{R}^{nm \times nm}$ and all other block matrices with block size $m \times m$.

Note that (153) and (154) give that $\mathbf{P}^T \mathbf{P}$ and $\mathbf{P}^T \mathcal{P} \mathbf{P}$ and their blocks are symmetric.

4.6. The DtB map

The main difference in the calculation of the DtB mapping is that the Cholesky factorizations (77) and (82) at step 1 of algorithm 2.1 are replaced with their block Cholesky counterparts given below.

Algorithm 4.1 (Block Cholesky factorization).

Input: the symmetric block matrix $\mathbf{X} \in \mathbb{R}^{nm \times nm}$ with $m \times m$ blocks.

To obtain the block Cholesky factorization of \mathbf{X} perform the following steps:

$$\text{For } k = 0, 1, \dots, n-1 \text{ compute : } \mathbf{R}_{k,k} = \mathbf{Q}_k \left(\mathbf{X}_{k,k} - \sum_{i=0}^{k-1} \mathbf{R}_{i,k}^T \mathbf{R}_{i,k} \right)^{1/2}, \quad (155)$$

where $\mathbf{Q}_k \in \mathbb{R}^{m \times m}$ is an arbitrary orthogonal matrix.

$$\text{For } j = k+1, \dots, n-1 \text{ compute : } \mathbf{R}_{k,j} = \mathbf{R}_{k,k}^{-1} \left(\mathbf{X}_{k,j} - \sum_{i=0}^{k-1} \mathbf{R}_{i,k}^T \mathbf{R}_{i,j} \right). \quad (156)$$

Output: the block matrix $\mathbf{R} \in \mathbb{R}^{nm \times nm}$ with $m \times m$ blocks, satisfying $\mathbf{X} = \mathbf{R}^T \mathbf{R}$.

While the regular Cholesky factorization is defined uniquely (assuming it uses the principal value of the square root), there is an ambiguity in defining the block Cholesky factorization, which comes from the computation of the diagonal blocks $\mathbf{R}_{k,k}$ in (155). An optimal choice of this matrix is still an open question. We obtained good results with $\mathbf{Q}_k = \mathbf{I}$, but here we present another choice (yielding equally good numerical results). This choice allows us to extend the Galerkin–Petrov reasoning of section 2.4 to the MIMO case. Explicitly, we choose the factor \mathbf{Q}_k consistent with the MIMO analogue of recursion (95)–(97) for computing the primary and dual orthogonalized block snapshots $\bar{\mathbf{p}}_j \in \mathbb{R}^{N \times m}$ and $\bar{\mathbf{u}}_j \in \mathbb{R}^{N \times m}$, for $j = 1, \dots, n$.

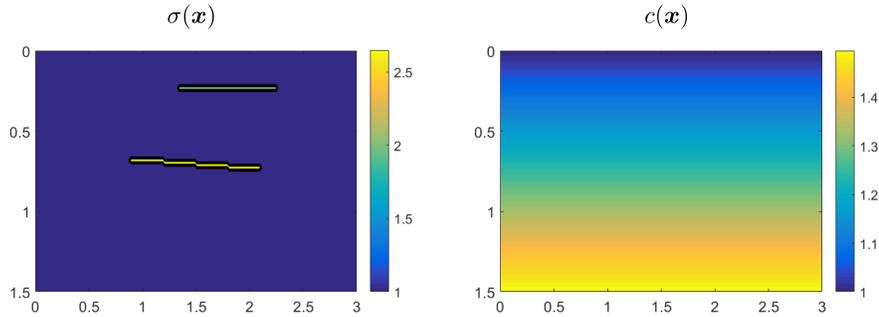


Figure 4. Model $\sigma(\mathbf{x})$ with two inclusions (left) and linear $c(\mathbf{x})$ (right). The axes are in km units.

Here the coefficients are no longer scalar, but symmetric positive definite matrices $\gamma_j \in \mathbb{R}^{m \times m}$ and $\hat{\gamma}_j \in \mathbb{R}^{m \times m}$. These matrices give the estimates of $\sigma(\mathbf{x})$ in the multidimensional case [23, section 7.3], so we can extend the reasoning of section 3 to the MIMO case, relating the block-bidiagonal $\tilde{\mathbf{L}}_q$ to a discretization of the Schrodinger equation, via the solution of the discrete inverse problem. We describe the computation of the block-bidiagonal $\tilde{\mathbf{L}}_q$ in appendix B.

After the factor $\tilde{\mathbf{L}}_q$ is found, the DtB mapping is computed using the block versions of algorithms 2.1 and 2.2, with \mathbf{b} given by (148) and \mathbf{e}_1 replaced by \mathbf{E}_1 from (152). We can also rewrite the MIMO counterpart of (87) in block form. The validity of the MIMO DtB mapping is based on the assumption of weak dependence of the primary and dual block-QR orthonormal snapshots \mathbf{V} and \mathbf{W} on $q = \ln \sigma$. Because of the consistency with the discrete inverse problem discussed above, this weak dependence can be understood using the same reasoning as in the SISO case. In particular, similar to the SISO case, the orthonormal snapshots approximate columns of the identity, as shown in figure 5. We also refer for more details to [25].

Remark 4.1. Unlike in the one dimensional case, in multi dimensions the DtB algorithm becomes mildly ill-posed, even for the space-time sampling close to the Nyquist rate. A simple regularization algorithm presented in [25] makes the DtB mapping practically insensitive to a reasonable (order of few per cent) level of noise in the measured data for the problem sizes considered in the numerical simulations shown in the next section.

4.7. Numerical results

We begin with numerical results for a two dimensional impedance model with two inclusions and a linear velocity model, shown in figure 4. We display the relative impedance and wave speed, normalized by their constant values at the sensors. Both the time sampling τ and the distances between the $m = 50$ sensors in the array are chosen close to the Nyquist sampling rate for the Gaussian pulse used in the experiments. As in the one dimensional case, the scattering data and the true Born approximation are computed using a fine grid finite difference time domain scheme, with grid steps much smaller than τ .

In figure 5 we plot the primary snapshots at two time instances $t_k = k\tau$, with $k = 25$ and $k = 30$, for the reference medium with $\sigma^0 \equiv 1$ and the scattering medium displayed in figure 4. We also display the orthonormal snapshots. The primary snapshots in the reference medium (first column) show the wavefront. In the scattering medium (second column) there

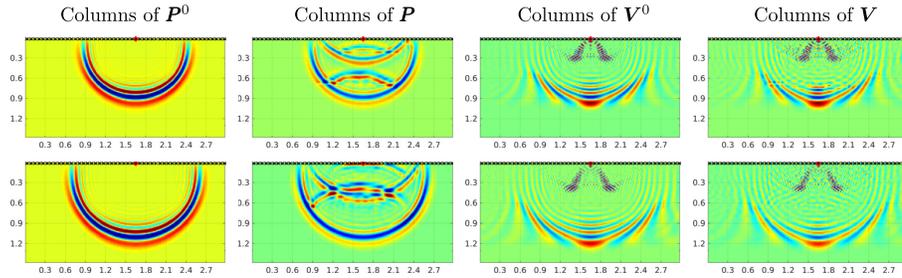


Figure 5. Primary snapshots and their orthonormal counterparts for the reference medium with $\sigma^0(\mathbf{x}) \equiv 1$ and the true medium. For every snapshot we plot a single column of the $N \times mn$ matrices, corresponding to the source $j = 28$ (red \circ) out of $m = 50$ sensors in the array (black \times). The rows (top to bottom) correspond to times $t_k = k\tau$, for $k = 25, 30$. The axes are in km units.

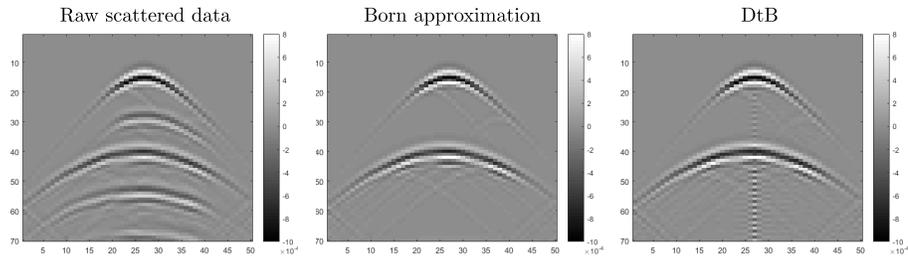


Figure 6. From left to right: raw scattered data, Born approximation and the data transformed by the DtB algorithm. The abscissa is the sensor index, and the ordinate is the index k of the discrete time instants $t_k = \tau k$, $\tau = 0.034s$.

are multiple reflections behind the wavefront. The orthonormal snapshots for both media (third and fourth columns) have a ‘smile-like’ shape with a ‘thick lower lip’. They can be viewed as approximations of delta functions. The reflections are suppressed in the orthonormal snapshots, and we note that they are almost the same in both media i.e. they are almost independent of σ . A similar result holds for the dual snapshots, not shown here.

In the left plot of figure 6 we display the raw scattered data at the sensors, due to the excitation from the sensor at the center of the linear array, lying just below the top boundary. The Born approximation and the data transformed by the DtB algorithm are in the middle and right plots. The results are almost the same. To illustrate better the agreement between the Born approximation and the output of the DtB algorithm, we display in figure 7 a comparison of several traces (signals at certain receivers) from figure 6.

To illustrate the benefit of the DtB transformation on imaging, we display in figure 8 the reverse time migration images⁹ obtained with the raw data and the transformed data shown in the left and right plots of figure 6. The artifacts due to multiple scattering are evident in the image shown on the left, which displays multiple ghost reflectors. The image obtained with the transformed data, shown on the right, does not have multiple artifacts and localizes well the two reflectors.

⁹We refer to [6, 52, 53] for details on the reverse time migration. It amounts to taking the data, time reversing it and backpropagating it in the reference (nonscattering) medium, with velocity $c(\mathbf{x})$. The image displays the resulting wave field evaluated at the travel time to points in the imaging region.

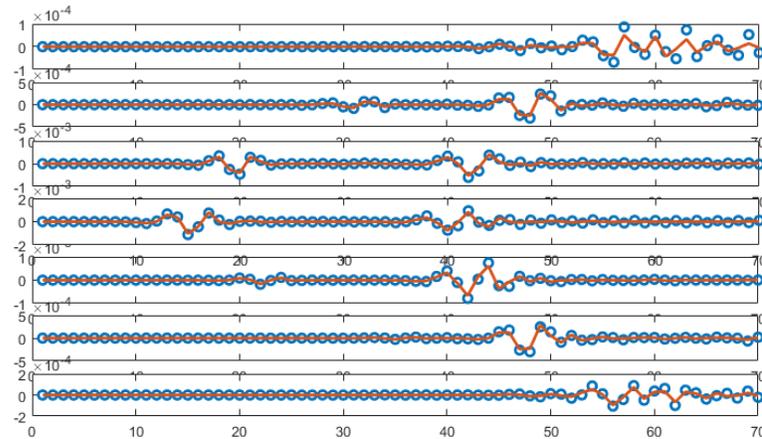


Figure 7. Comparison of a few traces of the Born approximation (red line) and the DtB algorithm output (blue \circ) from figure 6. The abscissa is the time index.

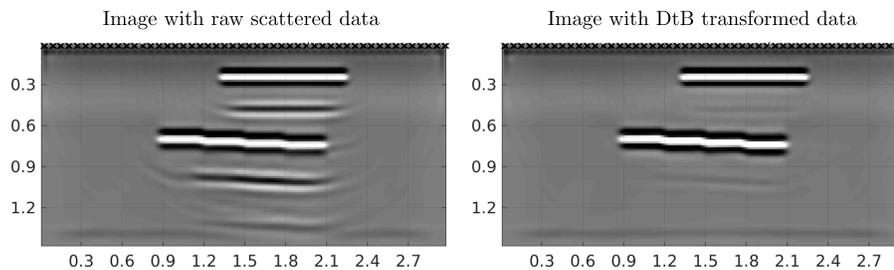


Figure 8. Reverse time migration image obtained with the raw scattered data (left) and the data transformed with the DtB algorithm (right). The axes are in km units, as in figure 4.

As we mentioned in the introduction, in practice only the smooth part of $c(\mathbf{x})$ may be known. To illustrate that the DtB algorithm can deal with perturbations of the sound speed, we present in figure 9 numerical results for three inclusions embedded in a medium with constant wave speed $c_0 = 1\text{km/s}$ and constant density ρ . The inclusions are modeled by the variation of $c(\mathbf{x})$ displayed in the left plot, but the density is kept constant i.e. $\sigma(\mathbf{x}) = \rho c(\mathbf{x})$. Only c_0 is assumed known in the DtB algorithm, meaning that we used the incorrect speed c_0 instead of the true $c(\mathbf{x})$. The data gathered by the array, for the excitation from the source shown with a red circle in the left plot, are displayed in the middle plot. They contain the primary reflections from each inclusion and multiply scattered reflections between the inclusions. The output of the DtB algorithm is displayed in the right plot of figure 9. The multiply scattered echoes are removed and there are three, clearly separated echoes, corresponding to each inclusion. Note the unmasking of the second echo, due to the smaller inclusion, that was mixed with a multiply scattered echo in the middle plot.

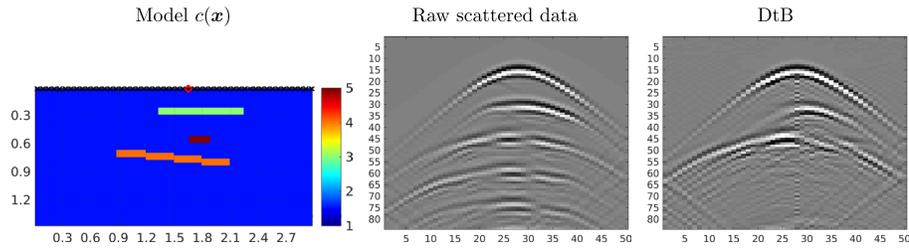


Figure 9. Left: model of a medium with three inclusions. The value of the wave speed $c(x)$ is shown in the color bar, in units of km s^{-1} . The axes are in km units and the array sits on the top, as in the previous simulations. Middle: data gathered for a single source shown as red \circ in the left figure. The abscissa is the sensor receiver index and the ordinate is the discrete time sample with step $\tau = 0.0225\text{s}$. Right: the output of the DtB algorithm.

5. Summary and future work

This paper is motivated by the inverse scattering problem for the wave equation, where an array of sensors probes an unknown scattering medium with pulses and measures the reflected waves. The goal of the inversion is to estimate the perturbations of the acoustic impedance in the medium, which cause wave scattering. We introduced a direct, linear-algebraic method, called the data to Born (DtB) algorithm, for transforming the data collected by the array to data corresponding to the single scattering (Born) approximation. These data can then be used by any off-the-shelf algorithms that incorporate state of the art linear inversion methodologies. The key ingredient in the DtB algorithm is a data driven, reduced order model (ROM), that approximates the wave propagator operator.

Data driven reduced order models have proved useful in inversion and imaging with waves [23, 25], but there are many questions for future research, such as:

- Since the DtB algorithm involves only linear algebra operations, like matrix–matrix multiplications and block Cholesky factorizations, the cost of the algorithm is $O((mn)^3)$, where m is the number of sensors and $2n$ is the number of time samples of the measurements. However, due to the Toeplitz-plus-Hankel structure of the mass and stiffness matrices, this cost can be reduced, most likely to $O(m^3n^2 \log n)$. A possibility of a fast Cholesky factorization specialized for Toeplitz-plus-Hankel matrices is currently being investigated.
- The DtB algorithm does not rely on the particular structure of the spatial differential operator L_q of the hyperbolic system (1). It only uses that L_q depends linearly on q . Thus, the algorithm works for any hyperbolic problem that can be written in the form (1). We are currently working on extensions to elastodynamics and Maxwell’s equations.
- Data-driven ROM construction in general and the DtB algorithm in particular can be susceptible to instability due to noise in the data, modeling errors, oversampling relative to Nyquist rate, etc. A simple, Tikhonov-like regularization approach to ROM construction was proposed in [25]. Alternatively, one may employ truncated spectral projections of the mass and stiffness matrices. This technique is essential to the elastodynamic case mentioned above as it does not admit a single Nyquist sampling rate due to different propagation speeds of P- and S-waves. Therefore, oversampling may cause instability even in the absence of noise, and regularization should become an integral part of the DtB algorithm.

- The DtB algorithm can be extended to map the measurements to second or higher order scattering data. This can be useful when the primary reflections do not provide an adequate illumination of the object of interest. For example, it is shown in [42] that when properly taken into account, multiply reflected waves greatly increase the imaging quality of bottom surfaces of salt bodies in reflection seismology.

Acknowledgments

This material is based upon research supported in part by the US Office of Naval Research under award number N00014-17-1-2057 to L Borcea and A V Mamonov. A V Mamonov was also partially supported by the National Science Foundation grant DMS-1619821. This material is based upon work supported by the National Science Foundation under Grant No. DMS-1439786 while the authors were in residence at the Institute for Computational and Experimental Research in Mathematics in Providence, RI, during the Fall 2017 semester.

Appendix A. The tridiagonal structure of the ROM propagator

In this appendix we show that the ROM propagator given by the projection (71) is a tridiagonal matrix. Obviously, \mathcal{P} is symmetric, so it suffices to show that its entries

$$(\widetilde{\mathcal{P}})_{j+l,j} = \mathbf{V}_{j+l}^T \mathcal{P} \mathbf{V}_j \quad (\text{A.1})$$

are zero when $l \geq 2$.

We obtain from equation (78) that $\mathbf{V}_j = \sum_{i=0}^{j-1} (\mathbf{R}^{-1})_{i+1,j} \mathbf{P}_i$, where we used that the inverse \mathbf{R}^{-1} of the upper triangular matrix \mathbf{R} is upper triangular. The relation (75) satisfied by the Chebyshev polynomials and definition (64) give

$$\mathcal{P} \mathbf{P}_i = \mathcal{T}_1(\mathcal{P}) \mathcal{T}_i(\mathcal{P}) \mathbf{b} = \frac{1}{2} [\mathcal{T}_{i+1}(\mathcal{P}) + \mathcal{T}_{|i-1|}(\mathcal{P})] \mathbf{b} = \frac{1}{2} (\mathbf{P}_{i+1} + \mathbf{P}_{|i-1|}), \quad (\text{A.2})$$

so equation (A.1) becomes

$$(\widetilde{\mathcal{P}})_{j+l,j} = \frac{1}{2} \sum_{i=0}^{j-1} (\mathbf{R}^{-1})_{i+1,j} (\mathbf{V}_{j+l}^T \mathbf{P}_{i+1} + \mathbf{V}_{j+l}^T \mathbf{P}_{|i-1|}). \quad (\text{A.3})$$

Each term in this sum can be calculated from (72) as $\mathbf{V}_{j+l}^T \mathbf{P}_i = (\mathbf{R})_{j+l,i+1}$, so we obtain

$$(\widetilde{\mathcal{P}})_{j+l,j} = \frac{1}{2} \sum_{i=0}^{j-1} (\mathbf{R}^{-1})_{i+1,j} [(\mathbf{R})_{j+l,i+2} + (\mathbf{R})_{j+l,|i-1|+1}]. \quad (\text{A.4})$$

Note that $j+l > \max\{i+2, |i-1|+1\}$ when $l \geq 2$ and $i = 0, \dots, j-1$, so the right hand side in (A.4) is zero by the upper triangular structure of \mathbf{R} . This means that $\widetilde{\mathcal{P}}$ is tridiagonal.

Appendix B. Computation of the block-bidiagonal $\widetilde{\mathbf{L}}_q$

We describe the computation of the block-Cholesky factor $\widetilde{\mathbf{L}}_q$ using an approach outlined in [24]. As mentioned in section 4.5, the block Cholesky factorization is not uniquely defined. Clearly, if $\mathbf{Q}_k = \mathbf{I}_m$, the diagonal blocks $\mathbf{R}_{k,k}$, are symmetric, for $k = 1, \dots, n$. We denote the corresponding MIMO ROM matrix given by (79) as $\widetilde{\mathcal{P}}^T$.

For non-trivial orthogonal matrices \mathbf{Q}_k , the MIMO ROM matrix $\tilde{\mathcal{P}}^{\mathcal{Q}}$ is given by $\mathbf{Q}\tilde{\mathcal{P}}^{\mathcal{I}}\mathbf{Q}^T$ where \mathbf{Q} is the block-diagonal matrix with orthogonal blocks \mathbf{Q}_k , $k = 1, \dots, n$. However, the block bidiagonal factor $\tilde{\mathbf{L}}_q$ in (82) has to be consistent with the matrix analogue of the recursion (95)–(97) for computing the primary and dual orthogonalized block snapshots $\tilde{\mathbf{p}}_j \in \mathbb{R}^{N \times m}$ and $\tilde{\mathbf{u}}_j \in \mathbb{R}^{N \times m}$,

$$\gamma_j^{-1}(\tilde{\mathbf{p}}_{j+1} - \tilde{\mathbf{p}}_j) = -\mathcal{L}_q \tilde{\mathbf{u}}_j, \quad \hat{\gamma}_j^{-1}(\tilde{\mathbf{u}}_j - \tilde{\mathbf{u}}_{j-1}) = \mathcal{L}_q^T \tilde{\mathbf{p}}_j, \quad j \geq 1, \quad (\text{B.1})$$

with initial conditions $\tilde{\mathbf{p}}_1 = \mathbf{b}$ and $\tilde{\mathbf{u}}_0 = \mathbf{0}$, and symmetric positive definite matrix coefficients

$$0 < \gamma_j = \gamma_j^T = (\tilde{\mathbf{u}}_j^T \tilde{\mathbf{u}}_j)^{-1} \in \mathbb{R}^{m \times m}, \quad 0 < \hat{\gamma}_j = \hat{\gamma}_j^T = (\tilde{\mathbf{p}}_j^T \tilde{\mathbf{p}}_j)^{-1} \in \mathbb{R}^{m \times m}, \quad (\text{B.2})$$

for $j = 1, \dots, n$. Then,

$$(\tilde{\mathbf{L}}_q)_{jj} = -\sqrt{\hat{\gamma}_j^{-1}} \sqrt{\gamma_j^{-1}}, \quad 1 \leq j \leq n, \quad (\tilde{\mathbf{L}}_q)_{j+1,j} = \sqrt{\hat{\gamma}_{j+1}^{-1}} \sqrt{\gamma_j^{-1}} \quad 1 \leq j \leq n-1. \quad (\text{B.3})$$

We now determine the matrix \mathbf{Q} such that the factorization

$$\xi(\mathbf{Q}\tilde{\mathcal{P}}^{\mathcal{I}}\mathbf{Q}^T) = \mathbf{Q}\xi(\tilde{\mathcal{P}}^{\mathcal{I}})\mathbf{Q}^T = \tilde{\mathbf{L}}_q \tilde{\mathbf{L}}_q^T, \quad (\text{B.4})$$

corresponds to $\tilde{\mathbf{L}}_q$ of the form (B.3) with symmetric positive definite γ_j and $\hat{\gamma}_j$, $j = 1, \dots, n$.

Denote the diagonal and the off-diagonal blocks of $\xi(\tilde{\mathcal{P}}^{\mathcal{I}})$ by $\alpha_j \in \mathbb{R}^{m \times m}$, for $j = 1, \dots, n$ and $\beta_j \in \mathbb{R}^{m \times m}$, for $j = 2, \dots, n$. By definition,

$$\hat{\gamma}_1 = (\mathbf{b}^T \mathbf{b})^{-1}. \quad (\text{B.5})$$

The remaining matrix coefficients γ_j and $\hat{\gamma}_j$ are obtained from (B.4) block-wise. From the first diagonal block we obtain that $\sqrt{\hat{\gamma}_1^{-1}} \gamma_1^{-1} \sqrt{\hat{\gamma}_1^{-1}} = \mathbf{Q}_1 \alpha_1 \mathbf{Q}_1^T$ or, equivalently,

$$\gamma_1 = \left(\sqrt{\hat{\gamma}_1} \mathbf{Q}_1 \alpha_1 \mathbf{Q}_1^T \sqrt{\hat{\gamma}_1} \right)^{-1}. \quad (\text{B.6})$$

Clearly, $\gamma_1 = \gamma_1^T$ for any matrix \mathbf{Q}_1 , so for simplicity we set $\mathbf{Q}_1 = \mathbf{I}_m$. Then, from the off-diagonal blocks for $1 \leq j \leq n-1$ we have $\sqrt{\hat{\gamma}_j^{-1}} \gamma_j^{-1} \sqrt{\hat{\gamma}_{j+1}^{-1}} = \mathbf{Q}_j \beta_{j+1} \mathbf{Q}_{j+1}^T$. Hence, $\sqrt{\hat{\gamma}_{j+1}^{-1}} \mathbf{Q}_{j+1} = \gamma_j \sqrt{\hat{\gamma}_j} \mathbf{Q}_j \beta_{j+1}$. That is to say, the pair of matrices $\sqrt{\hat{\gamma}_{j+1}^{-1}} \mathbf{Q}_{j+1}$ and \mathbf{Q}_{j+1} is a (left) polar decomposition of $\mathbf{M}_j = \gamma_j \sqrt{\hat{\gamma}_j} \mathbf{Q}_j \beta_{j+1}$. Its solution is

$$\hat{\gamma}_{j+1} = (\mathbf{M}_j \mathbf{M}_j^T)^{-1}, \quad \mathbf{Q}_{j+1} = \sqrt{\hat{\gamma}_{j+1}} \mathbf{M}_j. \quad (\text{B.7})$$

Finally, considering the diagonal blocks for $1 \leq j \leq n-1$, we obtain

$$\sqrt{\hat{\gamma}_{j+1}^{-1}} (\gamma_j^{-1} + \gamma_{j+1}^{-1}) \sqrt{\hat{\gamma}_{j+1}^{-1}} = \mathbf{Q}_{j+1} \alpha_{j+1} \mathbf{Q}_{j+1}^T$$

and therefore

$$\gamma_{j+1} = \left(\sqrt{\hat{\gamma}_{j+1}} \mathbf{Q}_{j+1} \alpha_{j+1} \mathbf{Q}_{j+1}^T \sqrt{\hat{\gamma}_{j+1}} - \gamma_j^{-1} \right)^{-1}. \quad (\text{B.8})$$

Algorithm B.1 (Computation of \tilde{L}_q).

Input: the block tridiagonal matrix $\tilde{\mathcal{P}}^J \in \mathbb{R}^{nm \times nm}$ with $m \times m$ blocks and $\mathbf{b}^T \mathbf{b}$.

To find a block diagonal \mathbf{Q} such that the block Cholesky factorization $\tilde{L}_q \tilde{L}_q^T$ of $\xi(\tilde{\mathcal{P}}^Q)$ has factors \tilde{L}_q in the form (B.3), perform the following steps:

1. Compute $\hat{\gamma}_1$ via (B.5) and γ_1 via (B.6) for arbitrary \mathbf{Q}_1 (for simplicity, we set $\mathbf{Q}_1 = \mathbf{I}_m$)
2. For $j = 1, \dots, n - 1$: Compute $\hat{\gamma}_{j+1}$ and \mathbf{Q}_{j+1} via (B.7) and γ_{j+1} via (B.8).
3. Compute \tilde{L}_q via (B.3)

Output: the block diagonal orthogonal matrix \mathbf{Q} , the block tridiagonal propagator matrix $\tilde{\mathcal{P}}^Q = \mathbf{Q} \tilde{\mathcal{P}}^J \mathbf{Q}^T$ and the block lower bidiagonal factor \tilde{L}_q of $\xi(\tilde{\mathcal{P}}^Q)$ consistent with (B.3).

ORCID iDs

Alexander V Mamonov  <https://orcid.org/0000-0002-1270-7535>

References

- [1] Alonso R, Borcea L, Papanicolaou G and Tsogka C 2011 Detection and imaging in strongly backscattering randomly layered media *Inverse Problems* **27** 025004
- [2] Aubry A and Derode A 2009 Detection and imaging in a random medium: a matrix method to overcome multiple scattering and aberration *J. Appl. Phys.* **106** 044903
- [3] Belishev M I 2007 Recent progress in the boundary control method *Inverse Problems* **23** R1
- [4] Beylkin G 1985 Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform *J. Math. Phys.* **26** 99–108
- [5] Beylkin G and Burridge R 1990 Linearized inverse scattering problems in acoustics and elasticity *Wave Motion* **12** 15–52
- [6] Biondi B 2006 *3D Seismic Imaging* (Tulsa, OK: Society of Exploration Geophysicists)
- [7] Borcea L, Del Cueto F G, Papanicolaou G and Tsogka C 2010 Filtering random layering effects in imaging *Multiscale Model. Simul.* **8** 751–81
- [8] Borcea L, del Cueto F G, Papanicolaou G and Tsogka C 2012 Filtering deterministic layer effects in imaging *SIAM Rev.* **54** 757–98
- [9] Borcea L, Druskin V and Knizhnerman L 2005 On the continuum limit of a discrete inverse spectral problem on optimal finite difference grids *Commun. Pure Appl. Math.* **58** 1231–79
- [10] Borcea L, Druskin V, Mamonov A V and Zaslavsky M 2014 A model reduction approach to numerical inversion for a parabolic partial differential equation *Inverse Problems* **30** 125011
- [11] Borcea L, Druskin V, Vasquez F G and Mamonov A 2011 Resistor network approaches to electrical impedance tomography *Inverse Problems Appl.* **60** 55–118
- [12] Borcea L, Papanicolaou G and Tsogka C 2011 Adaptive time-frequency detection and filtering for imaging in heavy clutter *SIAM J. Imaging Sci.* **4** 827–49
- [13] Bube K P and Burridge R 1983 The one-dimensional inverse problem of reflection seismology *SIAM Rev.* **25** 497–559
- [14] Burridge R 1980 The Gelfand–Levitan, the Marchenko, and the Gopinath–Sondhi integral equations of inverse scattering theory, regarded in the context of inverse impulse-response problems *Wave Motion* **2** 305–23
- [15] Cheney M and Borden B 2009 *Fundamentals of Radar Imaging* vol 79 (Philadelphia, PA: SIAM)
- [16] Chu M T and Golub G H 2002 Structured inverse eigenvalue problems *Acta Numer.* **11** 1
- [17] da Costa Filho C A, Ravasi M, Curtis A and Meles G A 2014 Elastodynamic green's function retrieval through single-sided marchenko inverse scattering *Phys. Rev. E* **90** 063201

- [18] de Hoop M, Kepley P and Oksanen L 2017 Recovery of a smooth metric via wave field and coordinate transformation reconstruction (arXiv:1710.02749)
- [19] Delprat-Jannaud F and Lailly P 2005 A fundamental limitation for the reconstruction of impedance profiles from seismic data *Geophysics* **70** R1–14
- [20] Drinkwater B W and Wilcox P D 2006 Ultrasonic arrays for non-destructive evaluation: a review *NDT E Int.* **39** 525–41
- [21] Druskin V, Guttel S and Knizhnerman L 2016 Near-optimal perfectly matched layers for indefinite helmholtz problems *SIAM Rev.* **58** 90–116
- [22] Druskin V and Knizhnerman L 1999 Gaussian spectral rules for the three-point second differences: I. A two-point positive definite problem in a semi-infinite domain *SIAM J. Numer. Anal.* **37** 403–22
- [23] Druskin V, Mamonov A V, Thaler A E and Zaslavsky M 2016 Direct, nonlinear inversion algorithm for hyperbolic problems via projection-based model reduction *SIAM J. Imaging Sci.* **9** 684–747
- [24] Druskin V, Mamonov A V and Zaslavsky M 2017 Multiscale S-fraction reduced-order models for massive wavefield simulations *Multiscale Model. Simul.* **15** 445–75
- [25] Druskin V, Mamonov A V and Zaslavsky M 2018 A nonlinear method for imaging with acoustic waves via reduced order model backprojection *SIAM J. Imaging Sci.* **11** 164–96
- [26] Dyukarev Y M 2004 Indeterminacy criteria for the stieltjes matrix moment problem *Math. Notes* **75** 66–82
- [27] Fomel S, Landa E and Taner M T 2007 Poststack velocity analysis by separation and imaging of seismic diffractions *Geophysics* **72** U89–94
- [28] Gallivan K, Grimme E, Sorensen D and Van Dooren P 1996 On some modifications of the Lanczos algorithm and the relation with Padé approximations *Math. Res.* **87** 87–116
- [29] Gallivan K, Grimme G and Van Dooren P 1996 A rational Lanczos algorithm for model reduction *Numer. Algorithms* **12** 33–63
- [30] Gel'fand I M and Levitan B M 1955 *On the Determination of a Differential Equation from its Spectral Function (American Mathematical Society Translations)* (Providence, RI: American Mathematical Society) pp 253–304
- [31] Golub H G and Van Loan C F 1996 *Matrix Computations* 3rd edn (Baltimore, MD: The Johns Hopkins University Press)
- [32] Gopinath B and Sondhi M 1971 Inversion of the telegraph equation and the synthesis of nonuniform lines *Proc. IEEE* **59** 383–92
- [33] Habashy T M 1991 A generalized Gel'fand–Levitan–Marchenko integral equation *Inverse Problems* **7** 703
- [34] Jannane M *et al* 1989 Wavelengths of earth structures that can be resolved from seismic reflection data *Geophysics* **54** 906–10
- [35] Kabanikhin S I and Shishlenin M A 2011 Numerical algorithm for two-dimensional inverse acoustic problem based on Gel'fand–Levitan–Krein equation *J. Inverse Ill-Posed Problems* **18** 979–95
- [36] Kac I and Krein M 1974 On the spectral functions of the string *Am. Math. Soc. Transl.* **103** 19–102
- [37] Krein M G 1951 Solution of the inverse Sturm–Liouville problem *Dokl. Akad. Nauk SSSR* **76** 21–4
- [38] Krein M G 1953 On the transfer function of a one-dimensional boundary problem of second order *Dokl. Akad. Nauk SSSR* **88** 405–8
- [39] Liu Z and Bleistein N 1995 Migration velocity analysis: theory and an iterative algorithm *Geophysics* **60** 142–53
- [40] Malcolm A E and de Hoop M V 2005 A method for inverse scattering based on the generalized bremer coupling series *Inverse Problems* **21** 1137
- [41] Malcolm A E, de Hoop M V and Calandra H 2007 Identification of image artifacts from internal multiples *Geophysics* **72** S123–32
- [42] Malcolm A E, Ursin B and De Hoop M V 2009 Seismic imaging and illumination with internal multiples *Geophys. J. Int.* **176** 847–64
- [43] Marchenko V A 1950 Some problems in the theory of one-dimensional second-order differential operators *Dokl. Akad. Nauk. SSSR* **2** 457–560
- [44] Meles G A, Wapenaar K and Curtis A 2016 Reconstructing the primary reflections in seismic data by marchenko redatuming and convolutional interferometry *Geophysics* **81** Q15–26
- [45] Moskow S, Kilgore K and Schotland J C 2012 Inverse Born series for scalar waves *J. Comput. Math.* **30** 601–14
- [46] Parlett B N 1998 *The Symmetric Eigenvalue Problem* (Philadelphia, PA: SIAM)
- [47] Rakesh 1988 A linearised inverse problem for the wave equation *Commun. PDE* **13** 573–601

- [48] Rivlin T J 1990 *Chebyshev Polynomials* (New York: Wiley)
- [49] Santosa F 1982 Numerical scheme for the inversion of acoustical impedance profile based on the Gelfand–Levitan method *Geophys. J. Int.* **70** 229–43
- [50] Stolk C C and Symes W W 2002 Smooth objective functionals for seismic velocity inversion *Inverse Problems* **19** 73
- [51] Symes W 1979 Inverse boundary value problems and a theorem of Gel’fand and Levitan *J. Math. Anal. Appl.* **71** 379–402
- [52] Symes W W 1995 Mathematics of reflection seismology. The Rice Inversion Project, Department of Computational and Applied Mathematics, Rice University <http://trip.rice.edu/downloads/preamble.pdf>
- [53] Symes W W 2009 The seismic reflection inverse problem *Inverse Problems* **25** 123008
- [54] Therrien C W 1992 *Discrete Random Signals and Statistical Signal Processing* (Englewood Cliffs, NJ: Prentice Hall)
- [55] Uhlmann G 2001 Travel time tomography *J. Korean Math. Soc.* **38** 711–22
- [56] Wapenaar K, Broggini F, Slob E and Snieder R 2013 Three-dimensional single-sided Marchenko inverse scattering, data-driven focusing, Greens function retrieval, and their mutual relations *Phys. Rev. Lett.* **110** 084301
- [57] Wapenaar K, Thorbecke J, Van Der Neut J, Broggini F, Slob E and Snieder R 2014 Marchenko imaging *Geophysics* **79** WA39–57
- [58] Weglein A B 1999 Multiple attenuation: an overview of recent advances and the road ahead *Leading Edge* **18** 40–4