Data-to-Born transform for multiple removal, inversion and imaging with waves

Alexander V. Mamonov[1],
Liliana Borcea[2], Vladimir Druskin[3], and Mikhail Zaslavsky[4]

[1]University of Houston,
[2]University of Michigan Ann Arbor,
[3]Worcester Polytechnic Institute,
[4]Schlumberger-Doll Research Center

Support: NSF DMS-1619821, ONR N00014-17-1-2057

Introduction

- **Inversion with waves**: determine properties of a medium in the bulk from response measured at or near the surface
- **Highly nonlinear** problem due to, in part, **multiple scattering**
- Given the full waveform response, can we compute the response of the **same medium** if waves propagated in the **single scattering** regime, i.e. in **Born regime**?
- Turns out we can!
- A highly nonlinear transform takes full waveform data to single scattering data: **Data-to-Born (DtB) transform**
- Can use as preprocessing step and integrate into **existing workflows**

Forward model

- Generic wave equation: DtB works for both **acoustics** and **elasticity** (also **electromagnetics**):

$$\partial_t^2 \mathbf{P}(t,\mathbf{x}) + L_q L_q^T \mathbf{P}(t,\mathbf{x}) = 0, \quad \mathbf{x} \in \Omega, \quad t > 0,$$

  here L_q is a **first order** differential operator, q is the **reflectivity**
- Model m **shots** with corresponding wavefields in a single matrix

$$\mathbf{P}(t,\mathbf{x}) = \left[\mathbf{P}^{(1)}(t,\mathbf{x}), \ldots, \mathbf{P}^{(m)}(t,\mathbf{x})\right]$$

- Shots modeled by **initial conditions**

$$\mathbf{P}(0,\mathbf{x}) = \mathbf{b}(\mathbf{x}) = \left[\mathbf{b}^{(1)}(\mathbf{x}), \ldots, \mathbf{b}^{(m)}(\mathbf{x})\right], \quad \partial_t \mathbf{P}(0,\mathbf{x}) = 0$$

- Solution

$$\mathbf{P}(t,\mathbf{x}) = \cos\left(t\sqrt{L_q L_q^T}\right)\mathbf{b}(\mathbf{x})$$

Data model and wavefield snapshots

- **Collocated** sources and receivers: receiver matrix is also $\mathbf{b}(\mathbf{x})$
- Data is **sampled** in time at $2n$ instants $t_k = k\tau$, close to **Nyquist** rate
- **Data model** becomes

$$\mathbf{D}_k = \int_\Omega \mathbf{b}(\mathbf{x})^T \cos\left(t\sqrt{L_q L_q^T}\right)\mathbf{b}(\mathbf{x})d\mathbf{x} \in \mathbb{R}^{m\times m}, \quad k = 0, 1, \dots, 2n{-}1,$$

or simply

$$\mathbf{D}_k = \int_\Omega \mathbf{b}(\mathbf{x})^T \mathbf{P}_k(\mathbf{x})d\mathbf{x} \in \mathbb{R}^{m\times m},$$

where

$$\mathbf{P}_k(\mathbf{x}) = \mathbf{P}(t_k, \mathbf{x}) = \cos\left(k\tau\sqrt{L_q L_q^T}\right)\mathbf{b}(\mathbf{x})$$

are **wavefield snapshots**

The propagator

- Important object: **propagator** operator

$$\mathscr{P}_q = \cos\left(\tau\sqrt{L_q L_q^T}\right),$$

think of it as **Green's function**

- Using propagator, snapshots admit representation

$$\mathbf{P}_k = \mathcal{T}_k(\mathscr{P}_q)\mathbf{b}, \quad k = 0, 1, \dots, 2n-1,$$

via **Chebyshev polynomials** $\mathcal{T}_k$

- **Notation**: let T denote both transpose and $L_2(\Omega)$ inner product, then the **data model** becomes

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

Reduced order model (ROM)

- Obviously, impossible to find $\mathscr{P}_q$ from **finite data** $\mathbf{D}_k \in \mathbb{R}^{m \times m}$, $k = 0, 1, \ldots, 2n-1$
- What can we find? **Reduced order model (ROM)** for $\mathscr{P}_q$!
- Specifically, **projection ROM**

$$\widetilde{\mathscr{P}}_q = \mathbf{V}^T \mathscr{P}_q \mathbf{V} \in \mathbb{R}^{nm \times nm}, \quad \widetilde{\mathbf{b}} = \mathbf{V}^T \mathbf{b} \in \mathbb{R}^{nm \times m},$$

  where "columns" of **V** form **orthonormal basis** for some subspace
- Of course, ROM must **fit the data**

$$\mathbf{D}_k = \mathbf{b}^T \mathcal{T}_k(\mathscr{P}_q)\mathbf{b} = \widetilde{\mathbf{b}}^T \mathcal{T}_k(\widetilde{\mathscr{P}}_q)\widetilde{\mathbf{b}}, \quad k = 0, 1, \ldots, 2n-1$$

- Data interpolation uniquely defines **projection (Krylov) subspace**

$$\mathrm{range}(\mathbf{\Pi}),$$

  spanned by snapshots, "columns" of **snapshot matrix**

$$\mathbf{\Pi} = \left[\mathbf{P}_0, \mathbf{P}_1, \ldots, \mathbf{P}_{n-1}\right]$$

Mass and stiffness matrices from data

- If we know **internal data**, snapshots $\mathbf{\Pi}$, we could **orthogonalize** them to find
$$\mathbf{V} = \left[\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{n-1}\right]$$
- **Multiplicative** property of Chebyshev polynomials to the rescue!
$$\mathcal{T}_j(x)\mathcal{T}_k(x) = \frac{1}{2}\left[\mathcal{T}_{j+k}(x) + \mathcal{T}_{|j-k|}(x)\right]$$
- Recall snapshots and data
$$\mathbf{P}_k = \mathcal{T}_k(\mathscr{P}_q)\mathbf{b}, \quad \mathbf{D}_k = \mathbf{b}^T\mathcal{T}_k(\mathscr{P}_q)\mathbf{b}$$
- Can find **inner products** from the **data**:
$$(\mathbf{\Pi}^T\mathbf{\Pi})_{j,k} = \mathbf{P}_j^T\mathbf{P}_k = \frac{1}{2}\left[\mathbf{D}_{j+k} + \mathbf{D}_{|j-k|}\right]$$
$$(\mathbf{\Pi}^T\mathscr{P}_q\mathbf{\Pi})_{j,k} = \mathbf{P}_j^T\mathscr{P}_q\mathbf{P}_k = \frac{1}{4}\left[\mathbf{D}_{j+k+1} + \mathbf{D}_{|j+k-1|} + \mathbf{D}_{|j-k+1|} + \mathbf{D}_{|j-k-1|}\right]$$

ROM from data

- Orthogonalized snapshots $\mathbf{V}$ can be related to $\mathbf{\Pi}$ via **block Gram-Schmidt** orthogonalization (**block QR factorization**)

$$\mathbf{\Pi} = \mathbf{V}\mathbf{R}, \quad \mathbf{V} = \mathbf{\Pi}\mathbf{R}^{-1},$$

  with block upper triangular $\mathbf{R}$ ($m \times m$ blocks)
- Then

$$\mathbf{\Pi}^T\mathbf{\Pi} = \mathbf{R}^T\mathbf{R}$$

  is **block Cholesky** factorization of **mass matrix** $\mathbf{\Pi}^T\mathbf{\Pi}$ known from the data
- Finally, **projection ROM** is given by

$$\widetilde{\mathscr{P}}_q = \mathbf{V}^T\mathscr{P}_q\mathbf{V} = \mathbf{R}^{-T}(\mathbf{\Pi}^T\mathscr{P}_q\mathbf{\Pi})\mathbf{R}^{-1},$$

  with both $\mathbf{R}$ and **stiffness matrix** $\mathbf{\Pi}^T\mathscr{P}_q\mathbf{\Pi}$ known from data

ROM properties

- ROM computation is entirely **data-driven**, no a priori information on continuum problem needed
- Gram-Schmidt orthogonalization (Cholesky) preserves **causality**: only looks backwards in time
- Reduced order propagator $\widetilde{\mathscr{P}}_q$ is **block tridiagonal**, blocks correspond to **layers of equal travel time** from the source array, can be seen as a (block) **second-order difference scheme**
- Orthogonalized snapshots $\mathbf{V}$ depend on the medium only **kinematically**, **reflections** are effectively **suppressed** in $\mathbf{V}$ (will see later in numerics)
- A version **robust** to noise and modeling errors exists: based on **spectral truncation** of the mass matrix $\mathbf{\Pi}^T\mathbf{\Pi}$, block Cholesky replaced with **block Lanczos**

Second order difference formulation

- We computed ROM propagator $\widetilde{\mathscr{P}}_q$, can we find reduced model for L_q itself?
- Wavefield snapshots **satisfy exactly** the **second order difference scheme**

$$\frac{\mathbf{P}_{k+1} - 2\mathbf{P}_k + \mathbf{P}_{k-1}}{\tau^2} + \mathcal{L}_q \mathcal{L}_q^T \mathbf{P}_k = 0, \quad k \geq 0,$$
$$\mathbf{P}_0 = \mathbf{b}, \quad \mathbf{P}_{-1} = \mathbf{P}_1,$$

with

$$\frac{2}{\tau^2}(\mathcal{I} - \mathscr{P}_q) = \mathcal{L}_q \mathcal{L}_q^T$$

- Can show

$$\mathcal{L}_q = L_q + O(\tau^2)$$

- This construction has a **reduced order analogue**

ROM propagator factorization

- **Reduced order snapshots** $\widetilde{\mathbf{P}}_k = \mathcal{T}_k(\widetilde{\mathscr{P}}_q)\widetilde{\mathbf{b}}$ also satisfy a second order scheme

$$\frac{\widetilde{\mathbf{P}}_{k+1} - 2\widetilde{\mathbf{P}}_k + \widetilde{\mathbf{P}}_{k-1}}{\tau^2} + \widetilde{\mathbf{L}}_q \widetilde{\mathbf{L}}_q^T \widetilde{\mathbf{P}}_k = 0, \quad k \geq 0,$$
$$\widetilde{\mathbf{P}}_0 = \widetilde{\mathbf{b}} = \mathbf{R}\mathbf{E}_1, \quad \widetilde{\mathbf{P}}_{-1} = \widetilde{\mathbf{P}}_1,$$

- To compute $\widetilde{\mathbf{L}}_q$ perform **second block Cholesky factorization**

$$\frac{2}{\tau^2}(\mathbf{I} - \widetilde{\mathscr{P}}_q) = \widetilde{\mathbf{L}}_q \widetilde{\mathbf{L}}_q^T$$

- So we have $\widetilde{\mathbf{L}}_q \in \mathbb{R}^{nm \times nm}$, a finite dimensional approximation of L_q
- Since $\widetilde{\mathscr{P}}_q$ is block tridiagonal, $\widetilde{\mathbf{L}}_q$ is **block lower bi-diagonal**
- Why is $\widetilde{\mathbf{L}}_q$ useful?

Example: acoustic wave equation

- Consider acoustic wave equation for pressure $p(t, \mathbf{x})$ in the form

$$\partial_t^2 p(t,\mathbf{x}) - \sigma(\mathbf{x}) c(\mathbf{x}) \nabla \cdot \left[\frac{c(\mathbf{x})}{\sigma(\mathbf{x})} \nabla p(t,\mathbf{x}) \right] = 0,$$

  with **velocity** $c(\mathbf{x})$ and **impedance** $\sigma(\mathbf{x})$
- Assume **kinematics** is known, seek Born approximation with respect to perturbation of $\sigma(\mathbf{x})$
- **Liouville transform** converts wave equation to **first order system**

$$\partial_t \begin{pmatrix} \mathbf{P}(t,\mathbf{x}) \\ \widehat{\mathbf{P}}(t,\mathbf{x}) \end{pmatrix} = \begin{pmatrix} 0 & -L_q \\ L_q^T & 0 \end{pmatrix} \begin{pmatrix} \mathbf{P}(t,\mathbf{x}) \\ \widehat{\mathbf{P}}(t,\mathbf{x}) \end{pmatrix},$$

  with corresponding second order form

$$\partial_t^2 \mathbf{P}(t,\mathbf{x}) + L_q L_q^T \mathbf{P}(t,\mathbf{x}) = 0$$

The reflectivity

- The operators L_q and L_q^T are given by

$$L_q = -\sqrt{c(\mathbf{x})}\nabla \cdot \sqrt{c(\mathbf{x})} + \frac{c(\mathbf{x})}{2}[\nabla q(\mathbf{x})]\cdot,$$
$$L_q^T = \sqrt{c(\mathbf{x})}\nabla \sqrt{c(\mathbf{x})} + \frac{c(\mathbf{x})}{2}[\nabla q(\mathbf{x})],$$

  with **reflectivity** $q(\mathbf{x}) = \ln \sigma(\mathbf{x})$
- If $c(\mathbf{x})$ is known and fixed, then L_q and L_q^T are **affine** in $q(\mathbf{x})$
- Since

$$\widetilde{\mathbf{L}}_q \approx L_q,$$

  then $\widetilde{\mathbf{L}}_q$ is **approximately affine in reflectivity** $q(\mathbf{x})$!
- **Perturbing** with respect to $q(\mathbf{x})$ becomes easy!

First order reduced order system

- **Reduced order** analogue of the **first order system**

$$\frac{\widetilde{\mathbf{P}}_{k+1} - \widetilde{\mathbf{P}}_k}{\tau} = -\widetilde{\mathbf{L}}_q \widehat{\widetilde{\mathbf{P}}}_k, \quad k = 0, \ldots, 2n-2,$$
$$\frac{\widehat{\widetilde{\mathbf{P}}}_k - \widehat{\widetilde{\mathbf{P}}}_{k-1}}{\tau} = \widetilde{\mathbf{L}}_q^T \widetilde{\mathbf{P}}_k, \quad k = 1, \ldots, 2n-1,$$

  with initial conditions

$$\widetilde{\mathbf{P}}_0 = \widetilde{\mathbf{b}}, \quad \widehat{\widetilde{\mathbf{P}}}_0 + \widehat{\widetilde{\mathbf{P}}}_{-1} = \mathbf{0}$$

- The right hand side is **approximately affine** in $q(\mathbf{x})$
- Perturbing $\widetilde{\mathbf{L}}_q$ with respect to q simply gives

$$\delta\widetilde{\mathbf{L}} = \widetilde{\mathbf{L}}_q - \widetilde{\mathbf{L}}_0,$$

  where $\widetilde{\mathbf{L}}_0$ is computed in **reference medium** with $q \equiv 0$

Data-to-Born transform

- Born approximation is a linearized perturbation
- **Perturbed** reduced order first order system

$$\frac{\delta\widetilde{\mathbf{P}}_{k+1} - \delta\widetilde{\mathbf{P}}_k}{\tau} = -\widetilde{\mathbf{L}}_0 \delta\widehat{\widetilde{\mathbf{P}}}_k - (\widetilde{\mathbf{L}}_q - \widetilde{\mathbf{L}}_0)\widehat{\widetilde{\mathbf{P}}}_{0,k}, \quad k = 0, \ldots, 2n-2,$$

$$\frac{\delta\widehat{\widetilde{\mathbf{P}}}_k - \delta\widehat{\widetilde{\mathbf{P}}}_{k-1}}{\tau} = \widetilde{\mathbf{L}}_0^T \delta\widetilde{\mathbf{P}}_k + (\widetilde{\mathbf{L}}_q^T - \widetilde{\mathbf{L}}_0^T)\widetilde{\mathbf{P}}_{0,k}, \quad k = 1, \ldots, 2n-1,$$

with initial conditions

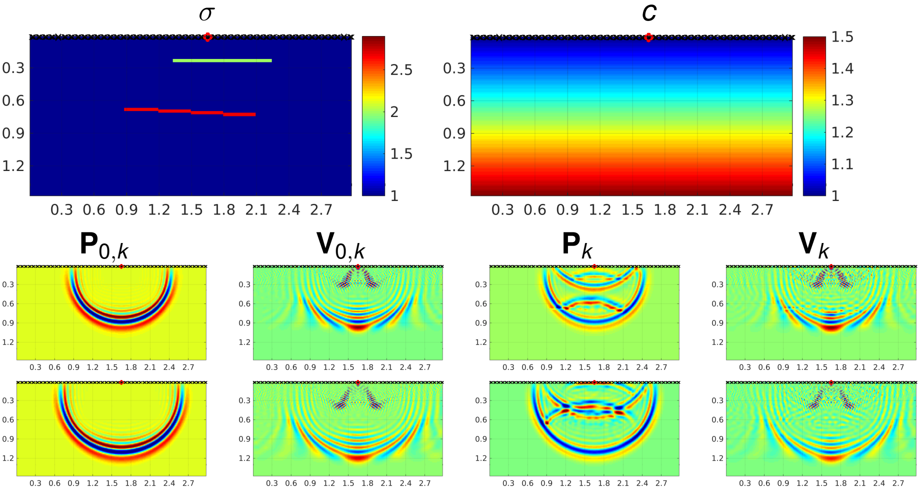
$$\delta\widetilde{\mathbf{P}}_0 = \mathbf{0}, \quad \delta\widehat{\widetilde{\mathbf{P}}}_0 + \delta\widehat{\widetilde{\mathbf{P}}}_{-1} = \mathbf{0}$$

- Here $\widetilde{\mathbf{P}}_{0,k}, \widehat{\widetilde{\mathbf{P}}}_{0,k}$ are reduced order snapshots in reference media
- **Data-to-Born transform** is

$$\mathbf{D}_k^{DtB} = \mathbf{D}_{0,k} + \widetilde{\mathbf{b}}^T \delta\widetilde{\mathbf{P}}_k, \quad k = 0, 1, \ldots, 2n-1,$$

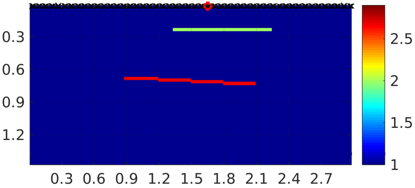
compare to full waveform data $\mathbf{D}_k = \widetilde{\mathbf{b}}^T \widetilde{\mathbf{P}}_k$

Numerical results: Acoustic snapshots



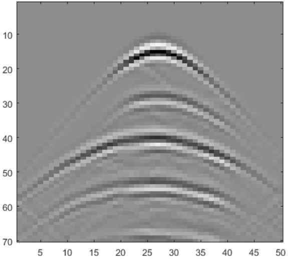
- Array with $m = 50$ sensors $\times$
- Snapshots plotted for a single source $\circ$

Numerical results: Acoustic true Born vs. DtB

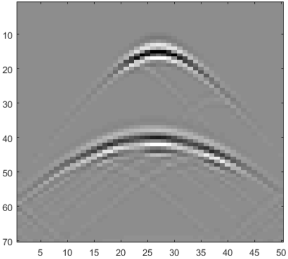


- Single row of data matrix corresponding to source ○
- **Vertical: time** (in units of τ)
- **Horizontal: receiver index** (out of $m = 50$)

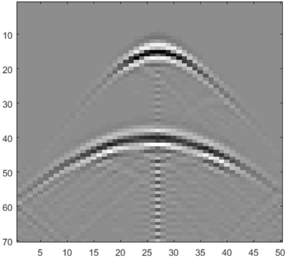
Full waveform data

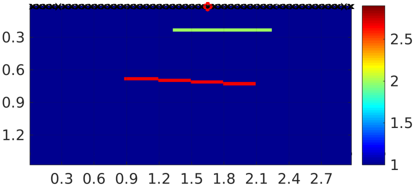


True Born data



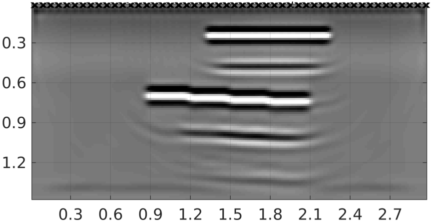
DtB



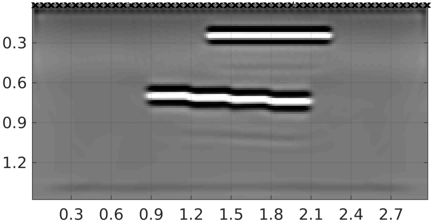


- **Reverse time migration** (RTM) image computed from both measured full waveform data and DtB transformed data

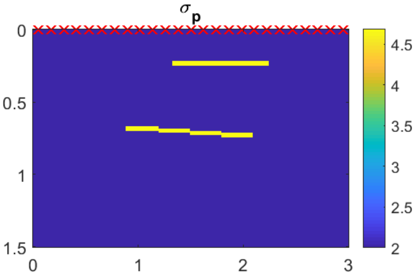
RTM from full waveform data



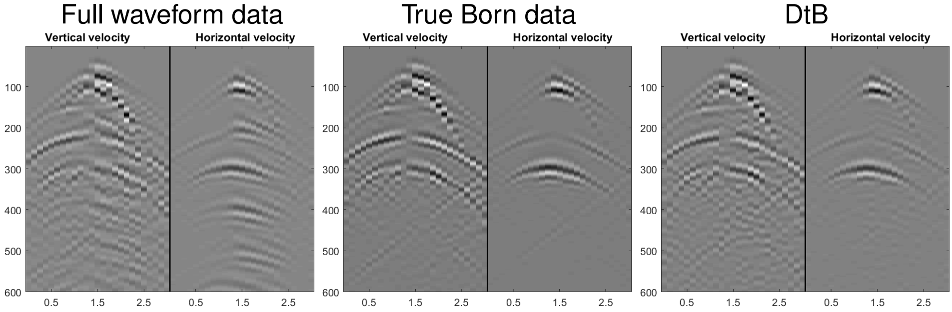
RTM from DtB

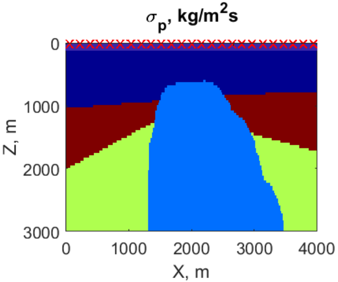


Numerical results: Elasticity, two cracks

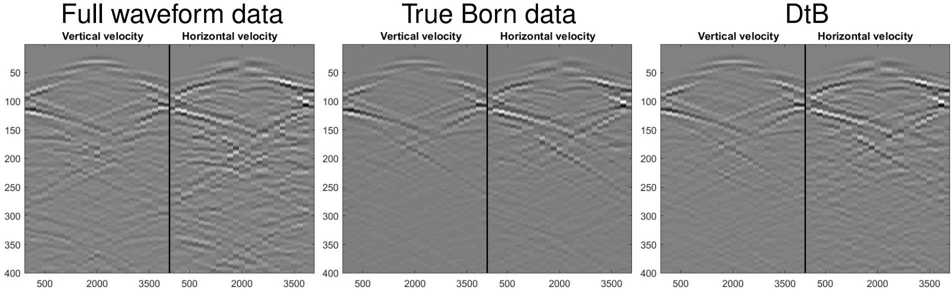


- Transform elasticity problem to first order form: **Liouville transform**
- If both velocities are fixed (here $c_p = 2c_s$), there is only **one independent impedance** σ_p
- Source: **horizontal force**, $m = 25$





- Transform elasticity problem to first order form: **Liouville transform**
- If both velocities are fixed (here $c_p = 2c_s$), there is only **one independent impedance** σ_p
- Source: **horizontal force**, $m = 25$



Conclusions and future work

- **Data-to-Born**: transform full waveform data to single scattered Born data for the same medium
- Based on techniques of **model order reduction**
- **Data-driven** approach relying on classical **linear algebra** algorithms (Cholesky, Lanczos), no computations in the continuum
- Works for all linear waves: **acoustic, elastic, electromagnetic**
- Easy to integrate into **existing workflows** as a preprocessing step
- Enables the use of **linearized inversion** algorithms

Future work:

- Test linearized inversion (e.g. **LS-RTM**) on DtB data
- Extend to **frequency domain** wave equation (Helmholtz)
- Use DtB-like approach to extract **higher orders of scattering** from full waveform data

References

- *Robust nonlinear processing of active array data in inverse scattering via truncated reduced order models*, L. Borcea, V. Druskin, A.V. Mamonov, M. Zaslavsky, **Journal of Computational Physics 381:1-26**, 2019.
- *Untangling the nonlinearity in inverse scattering with data-driven reduced order models*, L. Borcea, V. Druskin, A.V. Mamonov, M. Zaslavsky, **Inverse Problems 34(6):065008**, 2018.

Related work:

- *A nonlinear method for imaging with acoustic waves via reduced order model backprojection*, V. Druskin, A.V. Mamonov, M. Zaslavsky, **SIAM Journal on Imaging Sciences, 11(1):164-196**, 2018.
- *Direct, nonlinear inversion algorithm for hyperbolic problems via projection-based model reduction*, V. Druskin, A. Mamonov, A.E. Thaler and M. Zaslavsky, **SIAM Journal on Imaging Sciences 9(2):684–747**, 2016.