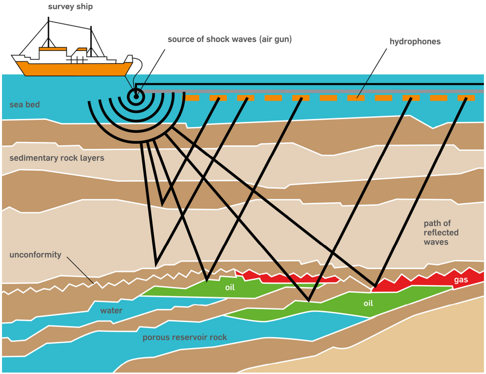
Waveform inversion via reduced order modeling

Alexander V. Mamonov[1],
Liliana Borcea[2], Josselin Garnier[3] and Jörn Zimmerling[4]

[1]University of Houston,
[2]University of Michigan Ann Arbor, [3]Ecole Polytechnique, [4]Uppsala University

Support: ONR N00014-21-1-2370

Motivation: seismic exploration



- **Reduced order model** (ROM) framework for acoustic velocity estimation:
  1. Construct a **data-driven** ROM from the data
  2. Formulate velocity estimation as **ROM misfit** optimization problem
- ROM misfit objective is much better behaved than conventional FWI least squares data misfit objective

Velocity estimation problem

- **Setting**: array of m sources/receivers (collocated at $\mathbf{x}_s$) drives pressure waves

$$\left[\partial_t^2 - c^2(\mathbf{x})\Delta\right] p^s(t,\mathbf{x}) = f'(t)\theta(\mathbf{x}-\mathbf{x}_s), \quad s = 1,\ldots,m,$$
$$p^s(t,\mathbf{x}) \equiv 0, \quad t \ll 0,$$

- Measured data $\mathcal{M}(t) \in \mathbb{R}^{m\times m}$ with entries

$$\mathcal{M}^{rs}(t) = \int_\Omega d\mathbf{x}\, \theta(\mathbf{x}-\mathbf{x}_r) p^s(t,\mathbf{x}), \quad r,s = 1,\ldots,m$$

- **Velocity estimation problem:** given $\mathcal{M}(t)$, estimate quantitatively acoustic velocity $c(\mathbf{x})$
- **Remark**: source/receiver collocation condition can be relaxed via data interpolation (numerical results available)

Symmetrized forward model

- Symmetrize the forward model, move source to initial condition (Duhamel-like argument), discretize in $\mathbf{x}$ on an N node grid

$$\partial_t^2 \mathbf{u} = \mathbf{A}\mathbf{u}, \quad t > 0,$$
$$\mathbf{u}(0) = \mathbf{b} \in \mathbb{R}^{N \times m}, \;\; \partial_t \mathbf{u}(0) = 0,$$

solved by

$$\mathbf{u}(t) = \cos\left(t\sqrt{\mathbf{A}}\right)\mathbf{b} \in \mathbb{R}^{N \times m}$$

- $\mathbf{A}$ is discretization of $-c(\mathbf{x})\Delta c(\mathbf{x})$
- Source/receiver matrix $\mathbf{b}$ depends on f, θ, c near $\mathbf{x}_s$
- Data becomes

$$\mathbf{D}(t) = \mathbf{b}^T \cos\left(t\sqrt{\mathbf{A}}\right)\mathbf{b} \in \mathbb{R}^{m \times m},$$

related to $\mathcal{M}(t)$ via

$$D^{rs}(t) = \frac{\mathcal{M}^{rs}(t) + \mathcal{M}^{rs}(-t)}{c(\mathbf{x}_r)c(\mathbf{x}_s)}, \quad t > 0$$

Projection based ROM

- Data is sampled discretely $\mathbf{D}_k = \mathbf{D}(k\tau)$, $k = 0, 1, \ldots, 2n-2$
- Define wavefield **snapshots** sampled at the same instants

$$\mathbf{u}_k = \mathbf{u}(k\tau) = \cos\left(k\tau\sqrt{\mathbf{A}}\right)\mathbf{b}$$

- Obtain ROM of **A** by **projecting** onto

$$\mathcal{K}_n = \text{colspan}(\mathbf{U}), \quad \mathbf{U} = [\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{n-1}] \in \mathbb{R}^{N \times mn}$$

- If columns of $\mathbf{V} \in \mathbb{R}^{N \times mn}$ form **orthonormal basis** for $\mathcal{K}_n$, then

$$\widetilde{\mathbf{A}} = \mathbf{V}^T\mathbf{A}\mathbf{V} \in \mathbb{R}^{mn \times mn}, \quad \widetilde{\mathbf{b}} = \mathbf{V}^T\mathbf{b} \in \mathbb{R}^{mn \times m}$$

- **Difficulty: U** and **V** contain wavefields in the whole domain, hence they are **unknown!**

Data-driven ROM: mass matrix

- Define $mn \times mn$ **mass** matrix

$$\mathbf{M} = \mathbf{U}^T \mathbf{U}$$

- Use trigonometric identity

$$\cos\alpha \cos\beta = \frac{1}{2}\left(\cos(\alpha+\beta) + \cos(\alpha-\beta)\right)$$

to compute mass matrix **blocks** (using $\mathbf{A}^T = \mathbf{A}$)

$$\begin{aligned}
\mathbf{M}_{ij} &= \mathbf{u}_i^T \mathbf{u}_j \\
&= \mathbf{b}^T \cos\left(i\tau\sqrt{\mathbf{A}}\right) \cos\left(j\tau\sqrt{\mathbf{A}}\right) \mathbf{b} \\
&= \frac{1}{2}\mathbf{b}^T \left[\cos\left((i+j)\tau\sqrt{\mathbf{A}}\right) + \cos\left(|i-j|\tau\sqrt{\mathbf{A}}\right)\right] \mathbf{b} \\
&= \frac{1}{2}\left(\mathbf{D}_{i+j} + \mathbf{D}_{|i-j|}\right) \in \mathbb{R}^{m\times m},
\end{aligned}$$

for $i, j = 0, 1, \ldots, n-1$, **from data!**

Data-driven ROM: stiffness matrix

- Similarly to $\mathbf{M}$, define $mn \times mn$ **stiffness** matrix

$$\mathbf{S} = \mathbf{U}^T \mathbf{A} \mathbf{U}$$

- To compute $\mathbf{S}$ we need to know $\ddot{\mathbf{D}}_k = -\mathbf{b}^T \mathbf{A} \cos(k\tau\sqrt{\mathbf{A}})\mathbf{b}$ that can be obtained from $\mathbf{D}(t)$ via **Fourier** domain **differentiation**
- Given second derivative data $\ddot{\mathbf{D}}_k$, $k = 0, 1, \ldots, 2n-2$, compute

$$\begin{aligned}
\mathbf{S}_{ij} = \mathbf{u}_i^T \mathbf{A} \mathbf{u}_j &= \\
&= \mathbf{b}^T \cos\left(i\tau\sqrt{\mathbf{A}}\right) \mathbf{A} \cos\left(j\tau\sqrt{\mathbf{A}}\right) \mathbf{b} \\
&= \frac{1}{2}\mathbf{b}^T \left[\mathbf{A} \cos\left((i+j)\tau\sqrt{\mathbf{A}}\right) + \mathbf{A} \cos\left(|i-j|\tau\sqrt{\mathbf{A}}\right)\right] \mathbf{b} \\
&= -\frac{1}{2}\left(\ddot{\mathbf{D}}_{i+j} + \ddot{\mathbf{D}}_{|i-j|}\right) \in \mathbb{R}^{m \times m},
\end{aligned}$$

for $i, j = 0, 1, \ldots, n-1$, again **from data!**

Data-driven ROM: block Cholesky factorization

- Suppose **U** is orthogonalized by a **block QR** (block Gram-Schmidt) process

$$\mathbf{U} = \mathbf{V}\mathbf{R}, \text{ equivalently, } \mathbf{V} = \mathbf{U}\mathbf{R}^{-1},$$

where **R** is an upper-block-triangular **block Cholesky** factor of the **mass matrix** $\mathbf{M} = \mathbf{U}^T\mathbf{U}$ known from the data

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

- Projection ROM is given by

$$\widetilde{\mathbf{A}} = \mathbf{V}^T\mathbf{A}\mathbf{V} = \mathbf{R}^{-T}\left(\mathbf{U}^T\mathbf{A}\mathbf{U}\right)\mathbf{R}^{-1} = \mathbf{R}^{-T}\mathbf{S}\mathbf{R}^{-1},$$

where the **stiffness matrix** $\mathbf{S} = \mathbf{U}^T\mathbf{A}\mathbf{U}$ is also known from the data

Conventional FWI vs ROM inversion

- Conventional **full waveform inversion** (FWI): nonlinear least squares

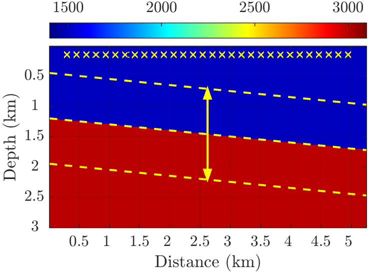
$$\underset{c(\mathbf{x})\in\mathcal{C}}{\text{minimize}} \sum_{k=0}^{2n-2} \|\mathbf{D}_k(c(\mathbf{x})) - \mathbf{D}_k^{\text{meas}}\|_F^2, \qquad (1)$$

  where $\mathbf{D}_k(c(\mathbf{x}))$ is the **forward map** and $\mathbf{D}_k^{\text{meas}}$ is **measured data**
- Objective of (1) is **notoriously non-convex**, optimization easily gets stuck in abundant local minima, especially when lacking **low-frequency** data (cycle skipping)
- Replace (1) with

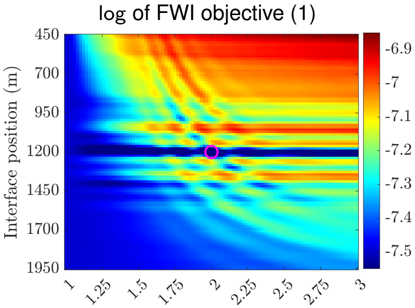
$$\underset{c(\mathbf{x})\in\mathcal{C}}{\text{minimize}} \left\| \widetilde{\mathbf{A}}(c(\mathbf{x})) - \widetilde{\mathbf{A}}^{\text{meas}} \right\|_F^2, \qquad (2)$$

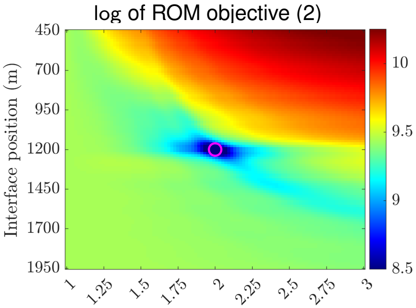
  where $\widetilde{\mathbf{A}}^{\text{meas}}$ is computed from $\mathbf{D}_k^{\text{meas}}$, $\ddot{\mathbf{D}}_k^{\text{meas}}$, $k = 0, 1, \ldots, 2n-2$
- Why objective (2) is better than (1)?

Objective topography: FWI vs ROM inversion



- Objective topography for a single interface model (left) with two parameters: **interface position** and **velocity contrast**
- Non-convexity of FWI objective (1): **cycle-skipping** results in horizontal stripes, also **local minima**
- ROM objective (2) has a **global minimum** at the true parameter values





Numerical experiments

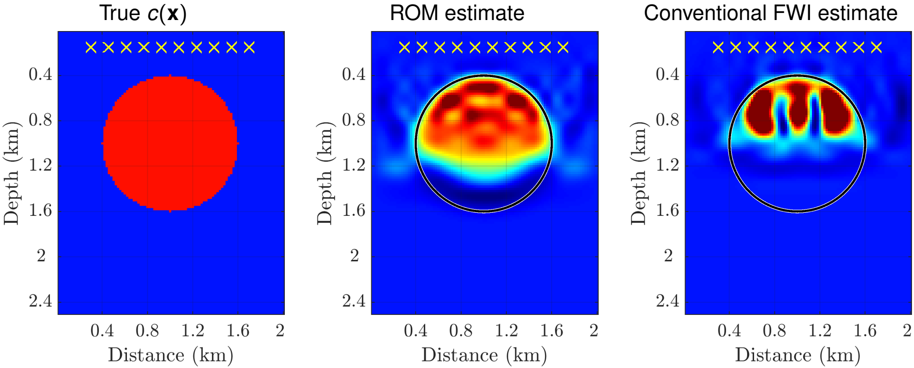
- Band-limited source wavelet

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

  with **central frequency** $\omega_0 = 2\pi(6Hz)$ and **bandwidth** $B_\omega = 2\pi(4Hz)$
- ROM based velocity estimation is solved via **Gauss-Newton** iteration regularized with **adaptive Tikhonov regularization**
- Construction of ROM $\widetilde{\mathbf{A}}$ is **causal**, optimization can be performed in a **layer-peeling** manner
- Two numerical examples:
  1. **"Camembert"** model with reflection data
  2. Section of the **Marmousi** model
- Marmousi velocity estimation is for **noisy data** (1% noise) using **regularized** ROM construction

Camembert model

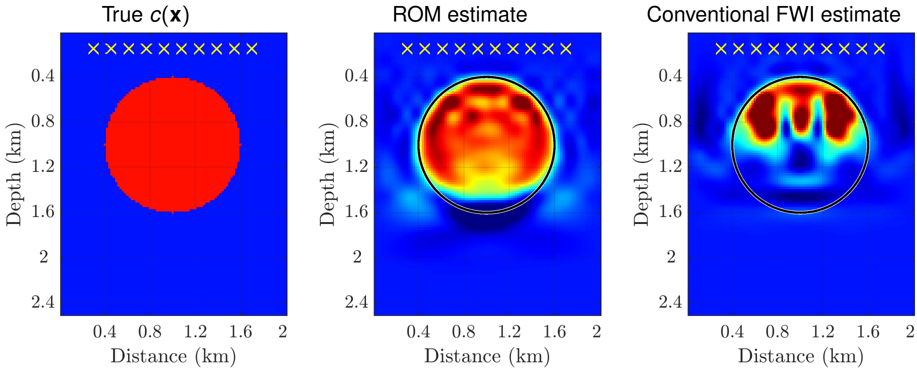
- Conventional FWI (1) vs. ROM estimation (2) after 10 GN iterations



- Camembert model with **reflection data**
- Circular inclusion ($c(\mathbf{x}) = 4000$m/s) of radius 600m in a homogeneous background ($c(\mathbf{x}) = 3000$m/s), data collected at $m = 10$ sensors
- Very challenging for FWI, difficult to fill in the inclusion

Camembert model

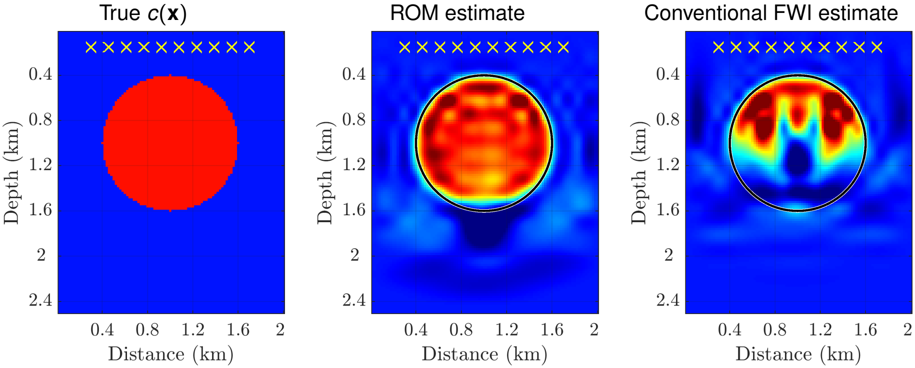
- Conventional FWI (1) vs. ROM estimation (2) after 20 GN iterations



- Camembert model with **reflection data**
- Circular inclusion ($c(\mathbf{x}) = 4000$m/s) of radius 600m in a homogeneous background ($c(\mathbf{x}) = 3000$m/s), data collected at $m = 10$ sensors
- Very challenging for FWI, difficult to fill in the inclusion

Camembert model

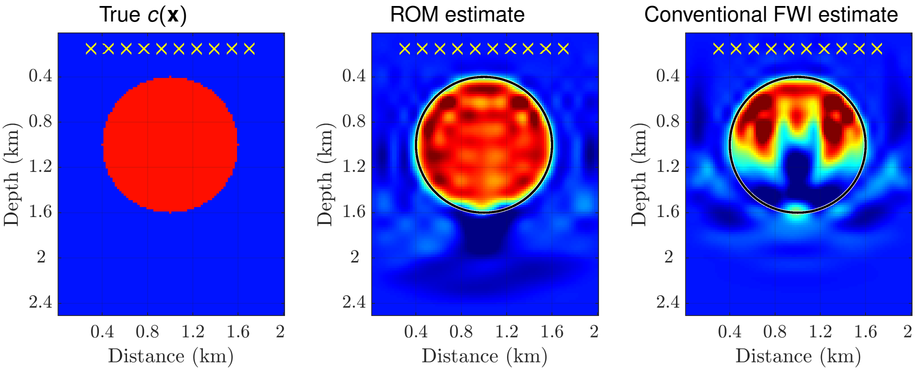
- Conventional FWI (1) vs. ROM estimation (2) after 40 GN iterations



- Camembert model with **reflection data**
- Circular inclusion ($c(\mathbf{x}) = 4000$m/s) of radius 600m in a homogeneous background ($c(\mathbf{x}) = 3000$m/s), data collected at $m = 10$ sensors
- Very challenging for FWI, difficult to fill in the inclusion

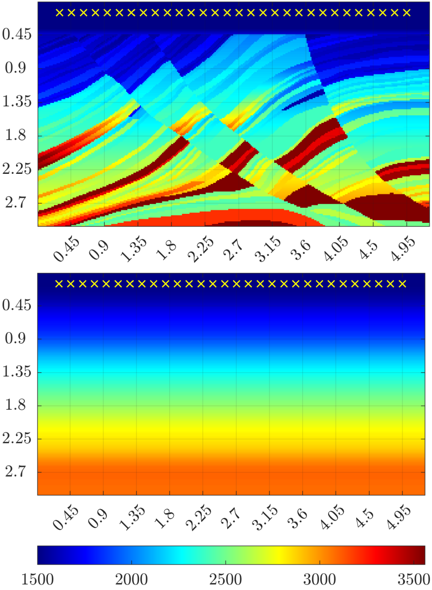
Camembert model

- Conventional FWI (1) vs. ROM estimation (2) after 60 GN iterations



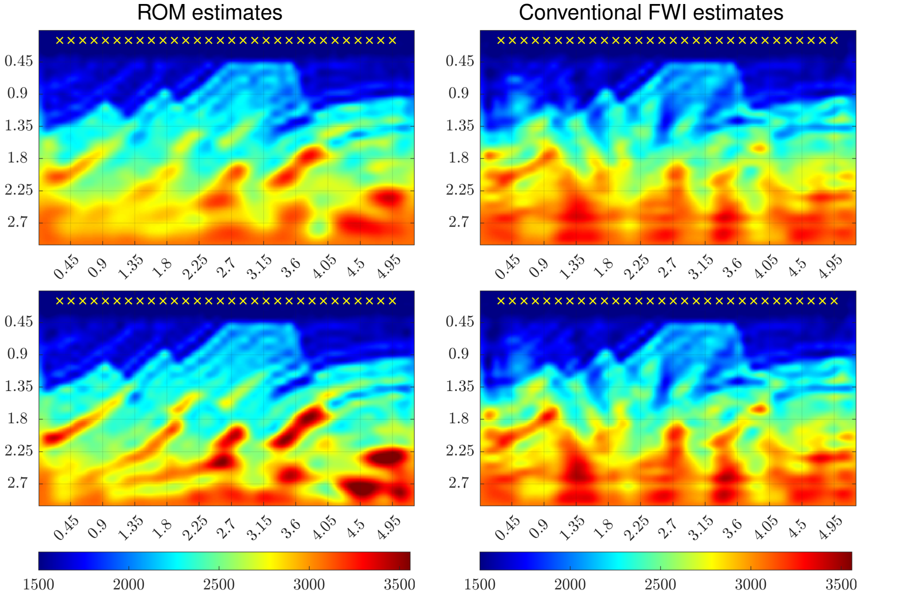
- Camembert model with **reflection data**
- Circular inclusion ($c(\mathbf{x}) = 4000$m/s) of radius 600m in a homogeneous background ($c(\mathbf{x}) = 3000$m/s), data collected at $m = 10$ sensors
- Very challenging for FWI, difficult to fill in the inclusion

Marmousi model



- **Top:** section of Marmousi model 5.25*km* $\times$ 3*km*
- **Bottom:** initial guess is a 1D gradient in depth
- Data collected at $m = 30$ sensors
- Perform 18 regularized Gauss-Newton iterations
- Compare to conventional FWI: it gets stuck in a low quality solution, likely not enough low-frequency information

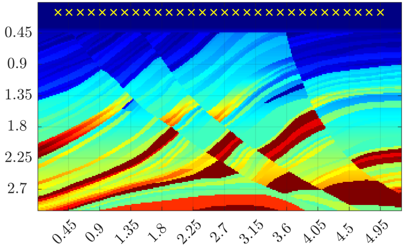
Marmousi model: FWI vs. ROM iterations 12 & 18



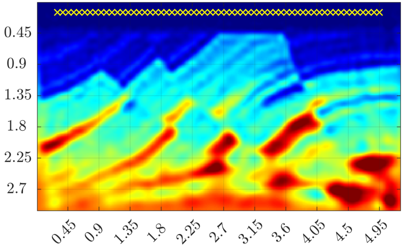
Marmousi model: velocity refinement

Refine velocity using data from $m = 60$ sensors. Same frequency content!

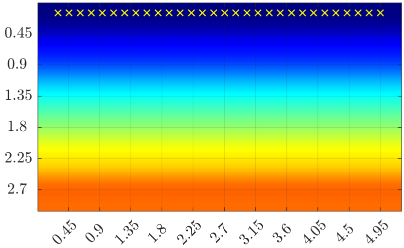
True $c(\mathbf{x})$



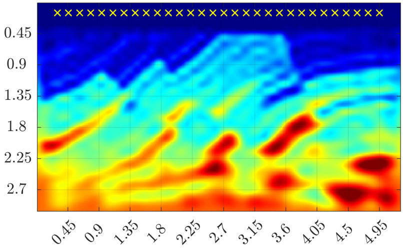
ROM refined velocity after 4 GN iterations



Initial guess

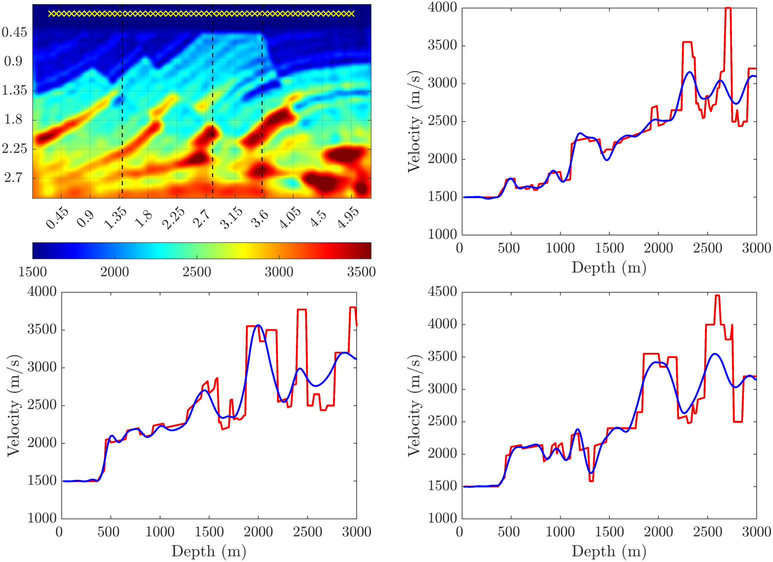


ROM velocity estimate after 18 GN iterations



Marmousi model: vertical velocity slices

True velocity (**red**) and the best ROM estimate (**blue**)



Conclusions and future work

- We introduced **ROM** framework for acoustic velocity estimation
- **Time domain** formulation is essential, linear algebraic analogues of **causality**: Gram-Schmidt, Cholesky
- Separate velocity estimation problem into **two steps**:
  1. Construct wave equation operator ROM from **data**
  2. Use **ROM misfit** as optimization objective
- Much better behaved than conventional FWI least squares data misfit even for **band-limited** sources: ROM misfit optimization objective is very close to **convex**
- **Robust** version exists for **noisy** and/or **incomplete data**, requires non-trivial regularization of ROM construction process

Future work:

- Extend to **vectorial** problems, e.g., electromagnetics, elasticity

References

- *Waveform inversion via reduced order modeling*, L. Borcea, J. Garnier, A.V. Mamonov and J. Zimmerling, to appear in **Geophysics**, preprint **arXiv:2202.01824 [math.NA]**

Related work:

1. *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**
2. *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**
3. *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**
4. *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**
5. *Reduced Order Model Approach to Inverse Scattering*, L. Borcea, V. Druskin, A.V. Mamonov, M. Zaslavsky, J. Zimmerling, **SIAM Journal on Imaging Sciences 13(2):685-723, 2020**
6. *Reduced order model approach for imaging with waves*, L. Borcea, J. Garnier, A.V. Mamonov, J. Zimmerling, **Inverse Problems, 38(2):025004, 2022**