
Reduced order modeling for inverse problems

Liliana Borcea

Mathematics, University of Michigan Ann Arbor

Joint work with:

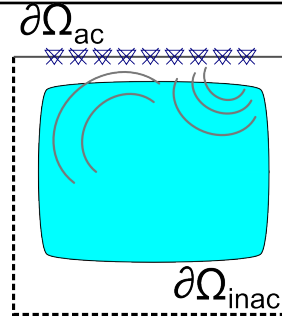
- **Vladimir Druskin** Worcester Polytechnic Institute
- **Alexander Mamonov** University of Houston
- **Mikhail Zaslavsky** Schlumberger
- **Jörn Zimmerling** University of Michigan

Support: U.S. Office of Naval Research, N00014-17-1-2057.

Goal of talk

- Describe two approaches for building data driven reduced order models (ROM) for linear PDE's.
 - Parabolic PDE and describe ROM construction in Laplace (frequency) domain.
 - Hyperbolic PDE and describe causal ROM construction in the time domain.
- Give some details on how we used the ROM for inversion.

Two inverse problems



- Find coefficient $q(\mathbf{x})$ of $A = -\nabla \cdot [q(\mathbf{x})\nabla]$ in parabolic PDE

$$\partial_t \mathbf{u}(t, \mathbf{x}) + A\mathbf{u}(t, \mathbf{x}) = 0$$

$$\mathbf{u}(0, \mathbf{x}) = \mathbf{b}(\mathbf{x}) = (b^{(1)}(\mathbf{x}), \dots, b^{(m)}(\mathbf{x}))$$

$$\partial_n \mathbf{u}(t, \mathbf{x}) = 0 \quad \text{on } \partial\Omega_{ac}, \quad \mathbf{u}(t, \mathbf{x}) = 0 \quad \text{on } \partial\Omega_{inac}$$

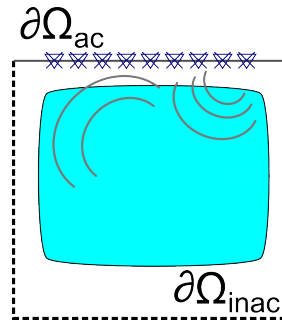
- “Sensor functions” $(b^{(s)}(\mathbf{x}))_{s=1}^m$ are supported near $\partial\Omega_{ac}$

Data: matrix $\mathbf{D}(t) = \int_{\Omega} d\mathbf{x} \mathbf{b}^T(\mathbf{x})\mathbf{u}(t, \mathbf{x}) = \langle \mathbf{b}, \mathbf{u}(t, \cdot) \rangle$

Work in frequency domain with $2n$ matrices:

$$\widehat{\mathbf{D}}(\omega_j) = \int_0^{\infty} dt e^{-\omega_j t} \mathbf{D}(t), \quad \partial_{\omega} \widehat{\mathbf{D}}(\omega_j), \quad j = 1, \dots, n$$

Two inverse problems



- Find coefficient $q(x)$ of symmetrized A in hyperbolic PDE*

$$\partial_t^2 \mathbf{u}(t, \mathbf{x}) + A\mathbf{u}(t, \mathbf{x}) = 0$$

$$\mathbf{u}(0, \mathbf{x}) = \mathbf{b}(\mathbf{x}), \quad \partial_t \mathbf{u}(0, \mathbf{x}) = 0$$

$$\partial_n \mathbf{u}(t, \mathbf{x}) = 0 \quad \text{on } \partial\Omega_{ac}, \quad \mathbf{u}(t, \mathbf{x}) = 0 \quad \text{on } \partial\Omega_{inac}$$

Data: $2n$ matrices $D(t_j) = \int_{\Omega} d\mathbf{x} \mathbf{b}^T(\mathbf{x}) \mathbf{u}(t_j, \mathbf{x}) = \langle \mathbf{b}, \mathbf{u}(t_j, \cdot) \rangle$

at time instants $t_j = j\tau$, for $j = 0, \dots, 2n - 1$

*Acoustics: $q = c = \text{wavespeed}$, $\mathbf{u} = \text{pressure}/c$ and $A = -c\Delta(c\cdot)$

Parabolic PDE: ROM in the frequency domain

- Laplace transform (dropping hats) of field at frequency ω_j is

$$\mathbf{u}_j(\mathbf{x}) := \mathbf{u}(\omega_j, \mathbf{x}) = (\omega_j I + A)^{-1} \mathbf{b}(\mathbf{x})$$

and we know

$$\begin{aligned} D(\omega_j) &= \langle \mathbf{b}, \mathbf{u}_j \rangle = \langle \mathbf{b}, (\omega_j I + A)^{-1} \mathbf{b} \rangle \\ \partial_\omega D(\omega_j) &= - \langle \mathbf{b}, (\omega_j I + A)^{-2} \mathbf{b} \rangle, \quad j = 1, \dots, n \end{aligned}$$

- **ROM** is a pair $\mathbf{A}^{\text{ROM}} \in \mathbb{R}^{nm \times nm}$ and $\mathbf{b}^{\text{ROM}} \in \mathbb{R}^{nm \times m}$ that interpolates data:

$$\begin{aligned} D(\omega_j) &= \mathbf{b}^{\text{ROM}T} (\omega_j I + \mathbf{A}^{\text{ROM}})^{-1} \mathbf{b}^{\text{ROM}} \\ \partial_\omega D(\omega_j) &= -\mathbf{b}^{\text{ROM}T} (\omega_j I + \mathbf{A}^{\text{ROM}})^{-2} \mathbf{b}^{\text{ROM}}, \quad j = 1, \dots, n \end{aligned}$$

ROM via Galerkin approximation

- Galerkin space: range $U(\mathbf{x})$, $U(\mathbf{x}) := (\mathbf{u}_1(\mathbf{x}), \dots, \mathbf{u}_n(\mathbf{x}))$

$$\mathbf{u}(\omega, \mathbf{x}) \approx U(\mathbf{x})\mathbf{g}(\omega) \quad \text{s.t.} \quad U^T [(\omega I + A)U\mathbf{g}(\omega) - \mathbf{b}] = 0$$

where $U^T U = (\langle \mathbf{u}_j, \mathbf{u}_l \rangle)_{1 \leq j, l \leq n}$

At $\omega = \omega_j$: $\mathbf{u}(\omega_j, \mathbf{x}) = \mathbf{u}_j(\mathbf{x}) = U(\mathbf{x})\mathbf{e}_j \rightsquigarrow \mathbf{g}(\omega_j) = \mathbf{e}_j$

Here $\mathbf{e}_j =$ matrix with j^{th} block equal to I_m and zero elsewhere.

- With $nm \times nm$ matrices: $M = U^T U$ and $S = U^T A U \rightsquigarrow$

$$(\omega M + S)\mathbf{g}(\omega) = U^T \mathbf{b}, \quad \forall \omega. \quad (1)$$

We don't know approximation space but can get M and S and thus $\mathbf{g}(\omega)$ from data.

Proof: From data to M and S

- Diagonal of M is easy:

$$\begin{aligned} M_{jj} &= \langle \mathbf{u}_j, \mathbf{u}_j \rangle = \langle (\omega_j I + A)^{-1} \mathbf{b}, (\omega_j I + A)^{-1} \mathbf{b} \rangle \\ &= \langle \mathbf{b}, (\omega_j I + A)^{-2} \mathbf{b} \rangle = -\partial_{\omega} D(\omega_j) \end{aligned}$$

- Eq. (1) for $\omega = \omega_j$ s.t. $\mathbf{g}(\omega_j) = \mathbf{e}_j$ multiplied on left by $\mathbf{e}_l^T \rightsquigarrow$

$$\omega_j M_{lj} + S_{jl} = \mathbf{e}_l^T (\omega_j \mathbf{M} + \mathbf{S}) \mathbf{e}_j = \langle \mathbf{u}_l, \mathbf{b} \rangle = D(\omega_l)$$

- Taking $l = j \rightsquigarrow S_{jj} = D(\omega_j) + \omega_j \partial_{\omega} D(\omega_j)$

- If $l \neq j$, we obtain similarly, using symmetries,

$$\omega_l M_{lj} + S_{jl} = D(\omega_j) \rightsquigarrow$$

$$M_{jl} = \frac{D(\omega_l) - D(\omega_j)}{\omega_j - \omega_l}, \quad S_{jl} = \frac{\omega_j D(\omega_j) - \omega_l D(\omega_l)}{\omega_j - \omega_l}, \quad j \neq l$$

ROM via Galerkin projection

- Let columns in $V(x)$ be orthonormal basis of Galerkin space given by Gram-Schmidt

$$U(x) = V(x)R \rightsquigarrow M = U^T U = R^T R$$

Upper triangular R given by Cholesky factorization of M

- The ROM:

$$A^{\text{ROM}} = V^T A V = R^{-T} U^T A U R^{-1} = R^{-T} S R^{-1}$$

$$b^{\text{ROM}} = V^T b = R^{-T} U^T b = R^{-T} \begin{pmatrix} D(\omega_1) \\ \vdots \\ D(\omega_n) \end{pmatrix}$$

ROM equation

Galerkin eq. $(\omega M + S)g(\omega) = U^T b$ with $U = VR$ & $M = R^T R$

$$\rightsquigarrow R^T (\omega I + \underbrace{R^{-T} S R^{-1}}_{A^{\text{ROM}}}) R g(\omega) = R^T \underbrace{V^T b}_{b^{\text{ROM}}}$$

- ROM Galerkin equation:

$$(\omega I + A^{\text{ROM}}) \mathbf{u}^{\text{ROM}}(\omega) = b^{\text{ROM}}, \quad \mathbf{u}^{\text{ROM}}(\omega) = R g(\omega)$$

Data driven ROM satisfies discrete equivalent of the PDE

- Using $g(\omega_j) = e_j$, for $j = 1, \dots, n$,

$$\mathbf{u}_j^{\text{ROM}} := \mathbf{u}^{\text{ROM}}(\omega_j) = R e_j = V^T U e_j = V^T \mathbf{u}_j$$

and therefore

$$\mathbf{u}_j(\mathbf{x}) = V V^T \mathbf{u}_j(\mathbf{x}) = V(\mathbf{x}) \mathbf{u}_j^{\text{ROM}}$$

ROM data interpolation

- For $j = 1, \dots, n$ we have

$$D(\omega_j) = \langle \mathbf{b}, \mathbf{u}_j \rangle = \underbrace{\langle \mathbf{b}, \mathbf{V} \rangle}_{\mathbf{b}^{\text{ROM}T}} \underbrace{(\omega_j \mathbf{I} + \mathbf{A}^{\text{ROM}})^{-1} \mathbf{b}^{\text{ROM}}}_{\mathbf{u}_j^{\text{ROM}}}$$

- For derivative:

$$\begin{aligned} \partial_\omega D(\omega_j) &= - \langle \mathbf{b}, (\omega_j \mathbf{I} + \mathbf{A})^{-2} \mathbf{b} \rangle = - \langle \mathbf{u}_j, \mathbf{u}_j \rangle \\ &= - \langle \mathbf{V} \mathbf{u}_j^{\text{ROM}}, \mathbf{V} \mathbf{u}_j^{\text{ROM}} \rangle \\ &= - \mathbf{u}_j^{\text{ROM}T} \langle \mathbf{V}, \mathbf{V} \rangle \mathbf{u}_j^{\text{ROM}} \\ &= - \mathbf{u}_j^{\text{ROM}T} \mathbf{u}_j^{\text{ROM}} \end{aligned}$$

Which is the best ROM?

- For any orthogonal matrix $\mathbf{L} \in \mathbb{R}^{nm \times nm}$ the data is also interpolated by

$$\tilde{\mathbf{A}}^{\text{ROM}} := \mathbf{L}^T \mathbf{A}^{\text{ROM}} \mathbf{L} \quad \text{and} \quad \tilde{\mathbf{b}}^{\text{ROM}} := \mathbf{L}^T \mathbf{b}^{\text{ROM}}$$

- We use \mathbf{L} that makes $\tilde{\mathbf{A}}^{\text{ROM}}$ block tridiagonal and $\tilde{\mathbf{b}}^{\text{ROM}}$ zero except in the first block

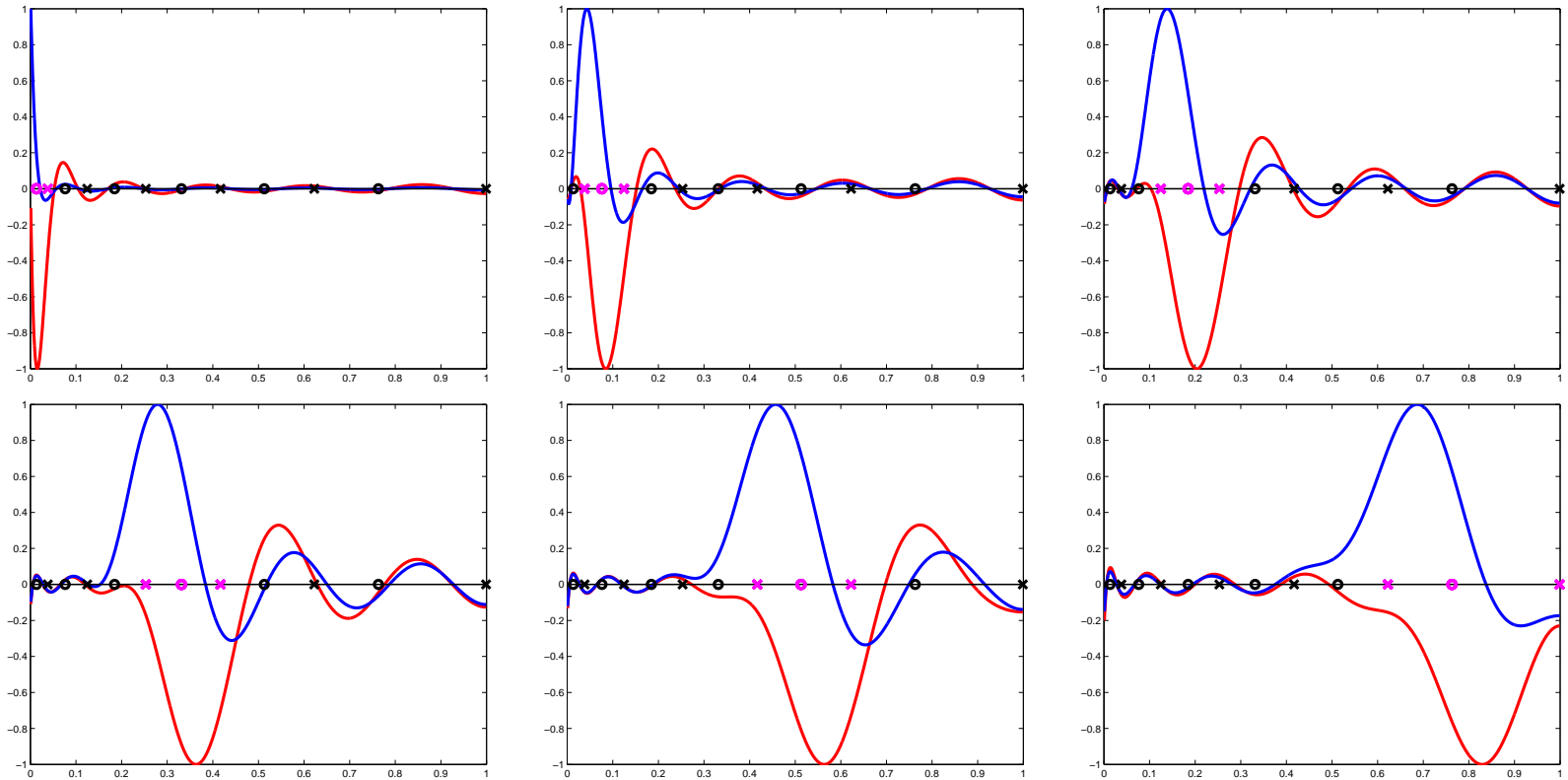
ROM eq. $(\omega \mathbf{I} + \tilde{\mathbf{A}}^{\text{ROM}}) \tilde{\mathbf{u}}^{\text{ROM}}(\omega) = \tilde{\mathbf{b}}^{\text{ROM}}$ is finite difference scheme for

$$(\omega \mathbf{I} + \mathbf{A})u(\omega, \mathbf{x}) = \mathbf{b}(\mathbf{x})$$

with 3–point stencil in depth

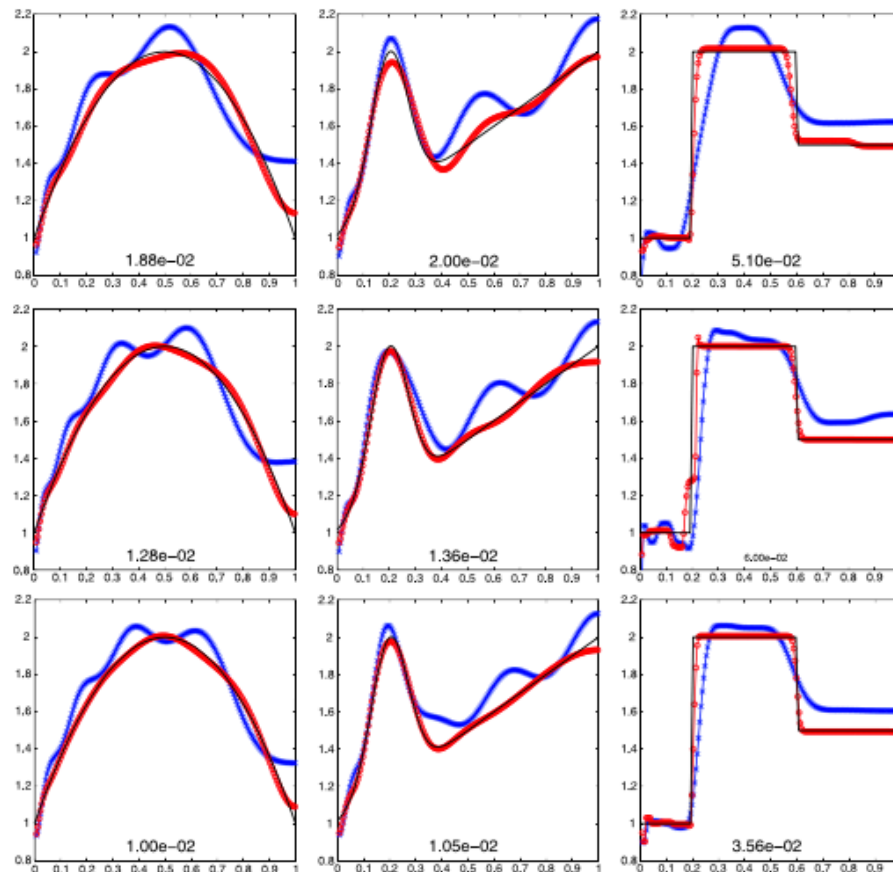
- Transformation to block tridiagonal is via the Lanczos iteration

Sensitivity of map $q \mapsto$ entries in $\widetilde{A}^{\text{ROM}}$ in $1 - D$



- Sensitivities are large in grid cells calculated for $q = 0$
- This grid tells us if we have good frequency samples

Inversion results (initial guess $q \equiv 1$)



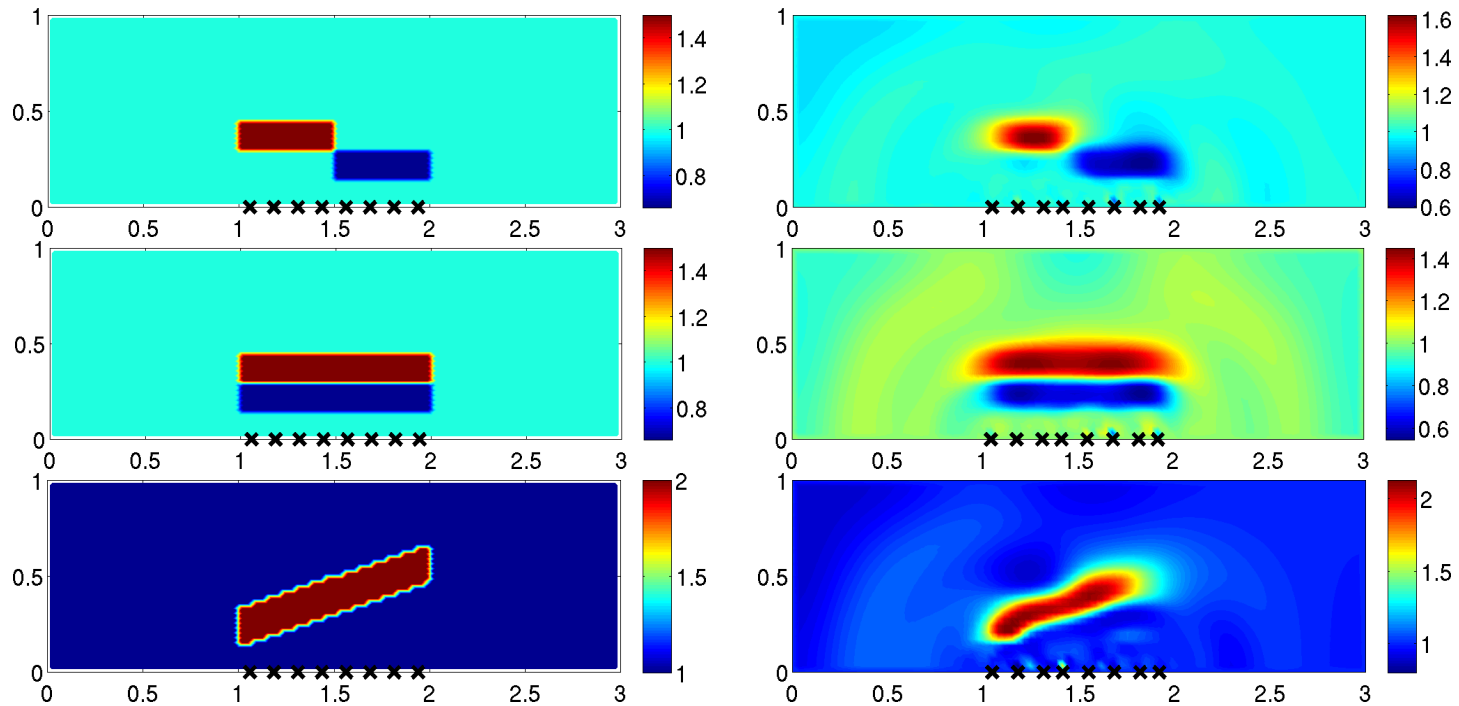
Black: True q
Blue: 1st iteration
Red: 5th iteration

ROM (top to bottom)
size is $n = 4, 5, 6$

Relative error displayed
in each plot

- Instead of Least Squares data fit, we find coefficient q by minimizing misfit between data driven $\widetilde{\mathbf{A}}^{\text{ROM}}$ and the ROM calculated for the trial q .

Inversion results 2-D (initial guess $q \equiv 1$)



- Other results based on estimates $u_j(\boldsymbol{x}) = V(\boldsymbol{x})\mathbf{u}_j^{\text{ROM}} \approx V_0(\boldsymbol{x})\mathbf{u}_j^{\text{ROM}}$ were discussed in Shari Moskow's talk.

Hyperbolic PDE: Causal ROM construction

- Find coefficient $q(\mathbf{x})$ of self-adjoint, positive definite A in

$$\partial_t^2 \mathbf{u}(t, \mathbf{x}) + A\mathbf{u}(t, \mathbf{x}) = 0$$

$$\mathbf{u}(0, \mathbf{x}) = \mathbf{b}(\mathbf{x}), \quad \partial_t \mathbf{u}(0, \mathbf{x}) = 0$$

with homogeneous boundary conditions.

- Data are

$$D(t_j) = \langle \mathbf{b}, u(j\tau, \cdot) \rangle = \langle \mathbf{b}, \cos(j\tau\sqrt{A})\mathbf{b} \rangle, \quad j = 0, \dots, 2n - 1.$$

- **ROM** is not for A but for **wave propagator** $\mathcal{P} = \cos(\tau\sqrt{A})$

Note: $\mathbf{u}_j(\mathbf{x}) := \mathbf{u}(j\tau, \mathbf{x}) = \cos(j\tau\sqrt{A})\mathbf{b}(\mathbf{x}) = \cos(j \arccos \mathcal{P})\mathbf{b}(\mathbf{x})$

$\mathcal{T}_j(\mathcal{P}) = \cos(j \arccos \mathcal{P}) =$ Tchebyshev polynomial of first kind

Dynamical system for propagation and Galerkin projection

- Recurrence relation of Chebyshev polynomials \rightsquigarrow

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

$$\mathbf{u}_0(\mathbf{x}) = \mathbf{b}(\mathbf{x})$$

$$\mathbf{u}_1(\mathbf{x}) = \mathbf{u}_{-1}(\mathbf{x})$$

- Galerkin space: $\text{range } U(\mathbf{x}), \quad U(\mathbf{x}) := (\mathbf{u}_0(\mathbf{x}), \dots, \mathbf{u}_{n-1}(\mathbf{x}))$

$\mathbf{u}_j(\mathbf{x}) \approx U(\mathbf{x})\mathbf{g}_j$, with $\mathbf{g}_j \in \mathbb{R}^{nm \times m}$ satisfying

$$U^T(U\mathbf{g}_{j+1} - 2\mathcal{P}U\mathbf{g}_j + U\mathbf{g}_{j-1}) = 0, \quad U^T U = (\langle \mathbf{u}_j, \mathbf{u}_l \rangle)_{0 \leq j, l \leq n-1}$$

First n coefficients: $\mathbf{g}_0 = \mathbf{e}_1, \dots, \mathbf{g}_{n-1} = \mathbf{e}_n$

From data to Galerkin projection

- Galerkin eq. $\mathbf{M}\mathbf{g}_{j+1} = 2\mathbf{S}\mathbf{g}_j - \mathbf{M}\mathbf{g}_{j-1}, \quad j \geq 0$
 $\mathbf{g}_j = \mathbf{e}_{j+1}, \quad 0 \leq j \leq n-1$

- Data $\mapsto nm \times nm$ matrices $\mathbf{M} = \mathbf{U}^T \mathbf{U}$ and $\mathbf{S} = \mathbf{U}^T \mathcal{P} \mathbf{U}$:

$$\begin{aligned} \mathbf{M}_{lj} &= \langle \mathbf{u}_l, \mathbf{u}_j \rangle = \langle \mathcal{T}_l(\mathcal{P})\mathbf{b}, \mathcal{T}_j(\mathcal{P})\mathbf{b} \rangle \\ &= \langle \mathbf{b}, \mathcal{T}_l(\mathcal{P})\mathcal{T}_j(\mathcal{P})\mathbf{b} \rangle = \frac{1}{2} \langle \mathbf{b}, [\mathcal{T}_{l+j}(\mathcal{P}) + \mathcal{T}_{|l-j|}(\mathcal{P})]\mathbf{b} \rangle \\ &= \frac{1}{2} (\mathbf{D}_{l+j} + \mathbf{D}_{|l-j|}) \end{aligned}$$

Similarly,

$$\mathbf{S}_{lj} = \langle \mathbf{u}_l, \mathcal{P}\mathbf{u}_j \rangle = \frac{1}{4} (\mathbf{D}_{l+j+1} + \mathbf{D}_{|j-l+1|} + \mathbf{D}_{|j-l-1|} + \mathbf{D}_{|l+j-1|})$$

ROM dynamical system

- Square root $M = R^T R$ (Cholesky) $\rightsquigarrow R = \begin{pmatrix} R_{11} & R_{12} & \dots & R_{1n} \\ 0 & R_{22} & \dots & R_{2n} \\ \vdots & & & \\ 0 & 0 & \dots & R_{nn} \end{pmatrix}$
- Multiply Galerkin eq. $Mg_{j+1} = 2Sg_j - Mg_{j-1}$ by R^{-T} on left

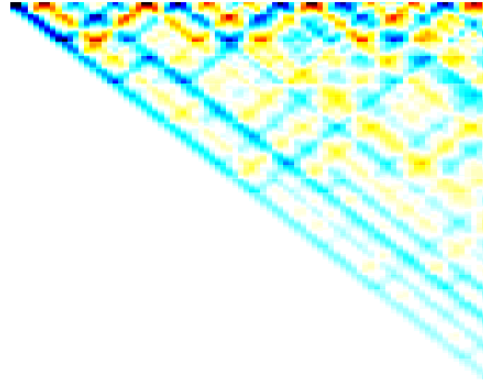
$$\begin{aligned} \mathbf{u}_{j+1}^{\text{ROM}} &= 2\mathcal{P}^{\text{ROM}} \mathbf{u}_j^{\text{ROM}} - \mathbf{u}_{j-1}^{\text{ROM}} = 0, \quad j \geq 0 \\ \mathbf{u}_0^{\text{ROM}} &= \mathbf{b}^{\text{ROM}} \\ \mathbf{u}_1^{\text{ROM}} &= \mathbf{u}_{-1}^{\text{ROM}} \end{aligned}$$

- ROM snapshots: $\mathbf{u}_j^{\text{ROM}} = Rg_j$ and propagator $\mathcal{P}^{\text{ROM}} = R^{-T}SR^{-1}$
- Gram-Schmidt: $U(x) = V(x)R \rightsquigarrow \mathbf{u}_j^{\text{ROM}} = Rg_j = V^T \underbrace{Ug_j}_{\approx \mathbf{u}_j}$ and

$$\mathcal{P}^{\text{ROM}} = R^{-T}SR^{-1} = R^{-T}U^T\mathcal{P}UR^{-1} = V^T\mathcal{P}V$$

Preserving causality is important

- 1-D illustration (1 source/receiver): snapshots $\mathbf{u}(j\tau, \mathbf{x})$ plotted vs. $j = 0, 1, \dots$ (horizontal axis) and x (vertical axis)



- ROM snapshots: $(\mathbf{u}_0^{\text{ROM}}, \dots, \mathbf{u}_{n-1}^{\text{ROM}}) = \mathbf{R} = \begin{pmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} & \dots & \mathbf{R}_{1n} \\ \mathbf{0} & \mathbf{R}_{22} & \dots & \mathbf{R}_{2n} \\ \vdots & & & \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{R}_{nn} \end{pmatrix}$

– In higher dimensions we have $m \times m$ blocks (m sources/receivers). Rows of blocks model advancement of wavefront.

- Gram-Schmidt $\mathbf{U}(\mathbf{x}) = \mathbf{V}(\mathbf{x})\mathbf{R}$ maps nearly triangular $\mathbf{U}(\mathbf{x})$ to $\mathbf{R} \rightsquigarrow \mathbf{V}(\mathbf{x})$ weakly dependent of reflectivity.

Illustration for sound waves in 1-D

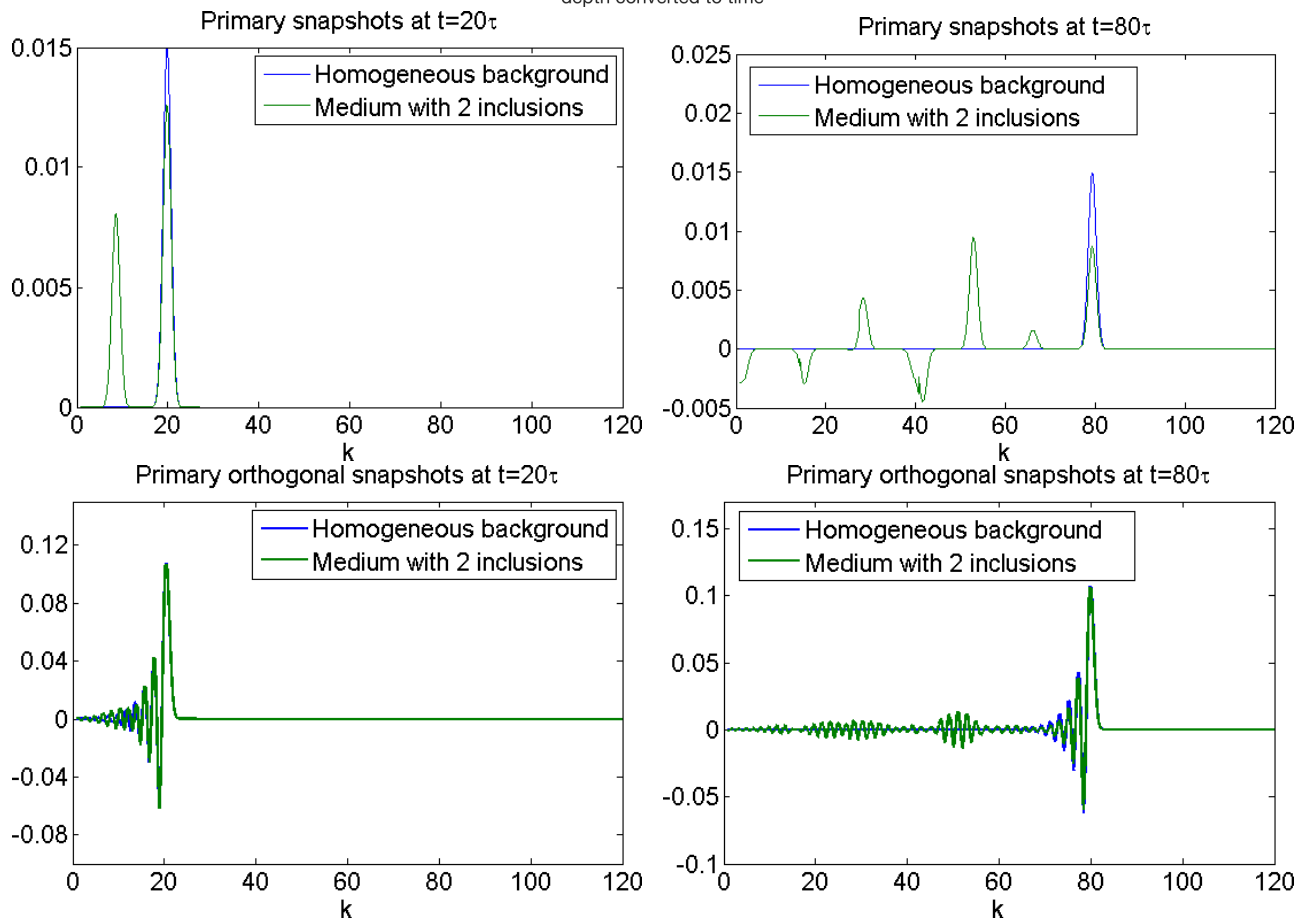
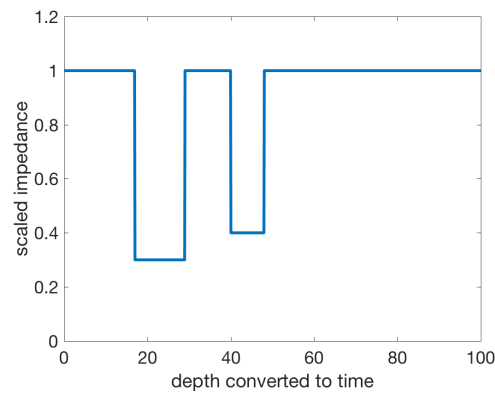
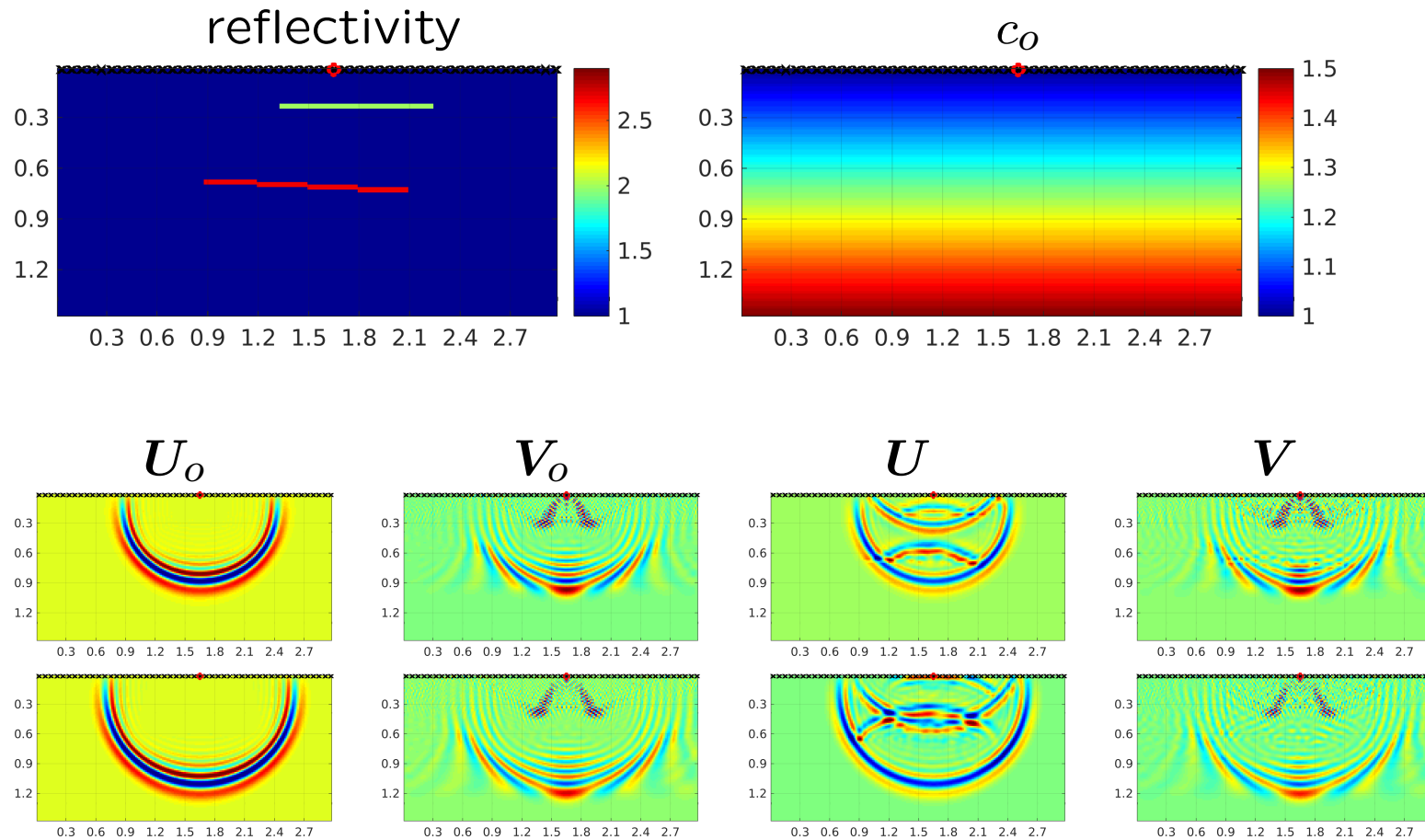


Illustration for sound waves in 2-D



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

Qualitative imaging*

- Consider internal “fictitious source” at $\mathbf{y} \in \Omega$:

$$\delta_{\mathbf{y}}(\mathbf{x}) := \mathbf{V}\mathbf{V}^T \delta(\mathbf{x} - \mathbf{y}) = \sum_{j=0}^{n-1} v_j(\mathbf{x}) \langle v_j, \delta(\cdot - \mathbf{y}) \rangle = \mathbf{V}(\mathbf{x})\mathbf{V}^T(\mathbf{y})$$

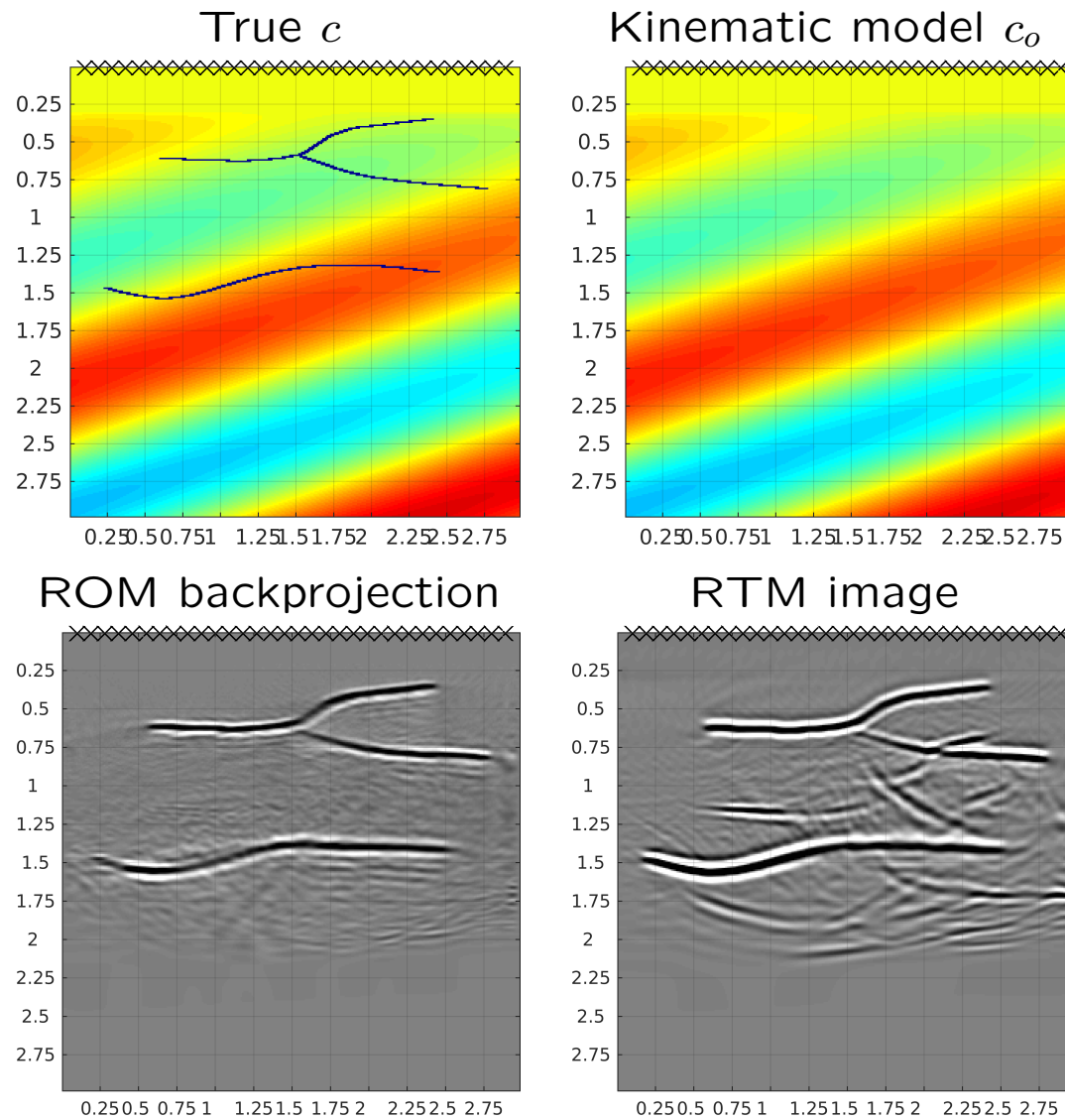
- Generated wave after one time step τ :

$$\mathcal{P}\delta_{\mathbf{y}}(\mathbf{x}) \approx \mathbf{V}\mathbf{V}^T \mathcal{P}\delta_{\mathbf{y}}(\mathbf{x}) = \mathbf{V}\mathbf{V}^T \mathcal{P}\mathbf{V}\mathbf{V}^T \delta(\mathbf{x} - \mathbf{y}) = \mathbf{V}(\mathbf{x})\mathcal{P}^{\text{ROM}}\mathbf{V}^T(\mathbf{y})$$

↪ Backprojection imaging function:

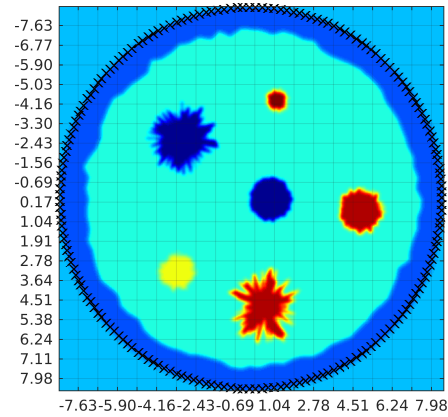
$$\mathcal{I}(\mathbf{y}) := \mathbf{V}_o(\mathbf{y}) \left(\mathcal{P}^{\text{ROM}} - \mathcal{P}_o^{\text{ROM}} \right) \mathbf{V}_o^T(\mathbf{y}), \quad \forall \mathbf{y} \in \Omega.$$

Backprojection vs Reverse Time Migration

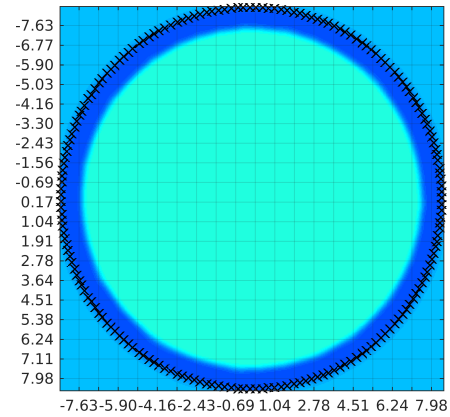


Backprojection vs Reverse Time Migration

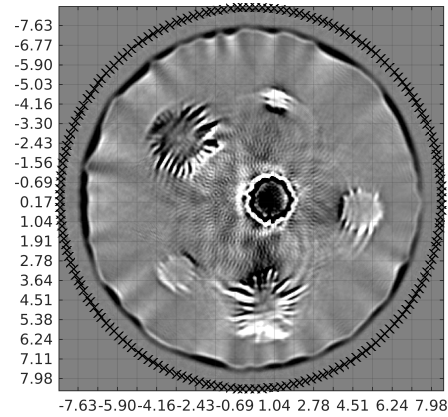
True c



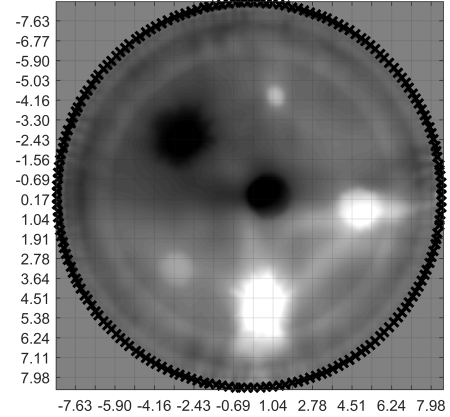
Kinematic model c_0



ROM backprojection



RTM image



Other inversion results

- Quantitative inversion for reflectivity (rough part of wave speed):

B., Druskin, Mamonov, Zaslavsky, Zimmerling, *Reduced Order Model Approach to Inverse Scattering*, SIAM Imaging Sciences 13 (2), 2020, p. 685-723.

- ROM used to approx. derivative of reflectivity $\mapsto \{\mathbf{D}_j\}_{0 \leq j \leq 2n-1}$ i.e., linearized (Born) model assumed in conventional imaging:

B., Druskin, Mamonov, Zaslavsky, *Untangling the nonlinearity in inverse scattering with data-driven reduced order models*, Inverse Problems 34 (6), 2018, p. 065008

B., Druskin, Mamonov, Zaslavsky, *Robust nonlinear processing of active array data in inverse scattering via truncated reduced order models*, Journal Comp. Physics 381, 2019, p. 1-26.

- With Josselin Garnier we are working on other applications and on velocity estimation.