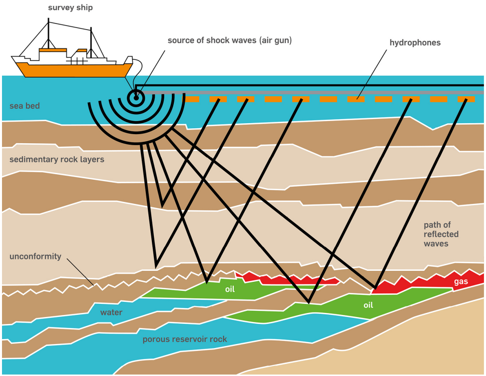
Seismic inversion and imaging via model order reduction

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

[1]University of Houston,
[2]University of Michigan Ann Arbor,
[3]Schlumberger-Doll Research Center,
[4]The Mathworks, Inc.

Motivation: seismic oil and gas exploration



Problems addressed:

1. **Inversion**: quantitative velocity estimation, FWI
2. **Imaging**: qualitative on top of velocity model
3. **Data preprocessing**: multiple suppression

- **Common framework**: Reduced Order Models (ROM)

Forward model: acoustic wave equation

- Acoustic wave equation in the **time domain**

$$\mathbf{u}_{tt} = \mathbf{A}\mathbf{u} \quad \text{in } \Omega, \quad t \in [0, T]$$

  with initial conditions

$$\mathbf{u}\big|_{t=0} = \mathbf{B}, \quad \mathbf{u}_t\big|_{t=0} = 0,$$

  sources are columns of $\mathbf{B} \in \mathbb{R}^{N \times m}$
- The spatial operator $\mathbf{A} \in \mathbb{R}^{N \times N}$ is a (symmetrized) fine grid discretization of

$$A = c^2 \Delta$$

  with appropriate boundary conditions
- Wavefields for all sources are columns of

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

Data model and problem formulations

- For simplicity assume that sources and receivers are **collocated**, receiver matrix is also $\mathbf{B}$
- The **data model** is

$$\mathbf{D}(t) = \mathbf{B}^T \mathbf{u}(t) = \mathbf{B}^T \cos(t\sqrt{-\mathbf{A}})\mathbf{B},$$

an $m \times m$ matrix function of time

Problem formulations:

1. **Inversion**: given $\mathbf{D}(t)$ estimate c
2. **Imaging**: given $\mathbf{D}(t)$ and a smooth kinematic velocity model c_0, estimate "reflectors", discontinuities of c
3. **Data preprocessing**: given $\mathbf{D}(t)$ obtain $\mathbf{F}(t)$ with multiple reflection events suppressed/removed

Reduced order models

- Data is always **discretely sampled**, say uniformly at $t_k = k\tau$
- The choice of τ is very important, optimally τ around Nyquist rate
- Discrete **data samples** are

$$\mathbf{D}_k = \mathbf{D}(k\tau) = \mathbf{B}^T \cos\left(k\tau\sqrt{-\mathbf{A}}\right)\mathbf{B} = \mathbf{B}^T T_k(\mathbf{P})\mathbf{B},$$

where T_k is Chebyshev polynomial and the **propagator** is

$$\mathbf{P} = \cos\left(\tau\sqrt{-\mathbf{A}}\right) \in \mathbb{R}^{N\times N}$$

- A **reduced order model** (ROM) $\widetilde{\mathbf{P}}$, $\widetilde{\mathbf{B}}$ should fit the data

$$\mathbf{D}_k = \mathbf{B}^T T_k(\mathbf{P})\mathbf{B} = \widetilde{\mathbf{B}}^T T_k(\widetilde{\mathbf{P}})\widetilde{\mathbf{B}}, \quad k = 0, 1, \ldots, 2n-1$$

Projection ROMs

- Projection ROMs are of the form

$$\widetilde{\mathbf{P}} = \mathbf{V}^T \mathbf{P} \mathbf{V}, \quad \widetilde{\mathbf{B}} = \mathbf{V}^T \mathbf{B},$$

  where **V** is an orthonormal basis for some subspace
- What subspace to project on to fit the data?
- Consider a matrix of **wavefield snapshots**

$$\mathbf{U} = [\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{n-1}] \in \mathbb{R}^{N \times nm}, \quad \mathbf{u}_k = \mathbf{u}(k\tau) = T_k(\mathbf{P})\mathbf{B}$$

- We must project on **Krylov subspace**

$$\mathcal{K}_n(\mathbf{P}, \mathbf{B}) = \text{colspan}[\mathbf{B}, \mathbf{P}\mathbf{B}, \ldots, \mathbf{P}^{n-1}\mathbf{B}] = \text{colspan } \mathbf{U}$$

- The data only knows about what **P** does to wavefield snapshots $\mathbf{u}_k$

ROM from measured data

- Wavefields in the whole domain **U** are unknown, thus **V** is unknown
- How to obtain ROM from just the data $\mathbf{D}_k$?
- Data does not give us **U**, but it gives us **inner products**!
- Multiplicative property of Chebyshev polynomials

$$T_i(x)T_j(x) = \frac{1}{2}(T_{i+j}(x) + T_{|i-j|}(x))$$

- Since $\mathbf{u}_k = T_k(\mathbf{P})\mathbf{B}$ and $\mathbf{D}_k = \mathbf{B}^T T_k(\mathbf{P})\mathbf{B}$ we get

$$(\mathbf{U}^T\mathbf{U})_{i,j} = \mathbf{u}_i^T\mathbf{u}_j = \frac{1}{2}(\mathbf{D}_{i+j} + \mathbf{D}_{i-j}),$$

$$(\mathbf{U}^T\mathbf{P}\mathbf{U})_{i,j} = \mathbf{u}_i^T\mathbf{P}\mathbf{u}_j = \frac{1}{4}(\mathbf{D}_{j+i+1} + \mathbf{D}_{j-i+1} + \mathbf{D}_{j+i-1} + \mathbf{D}_{j-i-1})$$

ROM from measured data

- Suppose $\mathbf{U}$ is orthogonalized by a **block QR** (Gram-Schmidt) procedure
$$\mathbf{U} = \mathbf{V}\mathbf{L}^T, \text{ equivalently } \mathbf{V} = \mathbf{U}\mathbf{L}^{-T},$$
where $\mathbf{L}$ is a **block Cholesky** factor of the **Gramian** $\mathbf{U}^T\mathbf{U}$ known from the data
$$\mathbf{U}^T\mathbf{U} = \mathbf{L}\mathbf{L}^T$$
- The projection is given by
$$\widetilde{\mathbf{P}} = \mathbf{V}^T\mathbf{P}\mathbf{V} = \mathbf{L}^{-1}\left(\mathbf{U}^T\mathbf{P}\mathbf{U}\right)\mathbf{L}^{-T},$$
where $\mathbf{U}^T\mathbf{P}\mathbf{U}$ is also known from the data
- Cholesky factorization is essential, (block) lower triangular structure is the linear algebraic equivalent of **causality**

Problem 1: Inversion (FWI)

- Conventional FWI (OLS)

$$\underset{c}{\text{minimize}} \, \|\mathbf{D}^\star - \mathbf{D}(\,\cdot\,; c)\|_2^2$$

- Replace the objective with a **“nonlinearly preconditioned”** functional

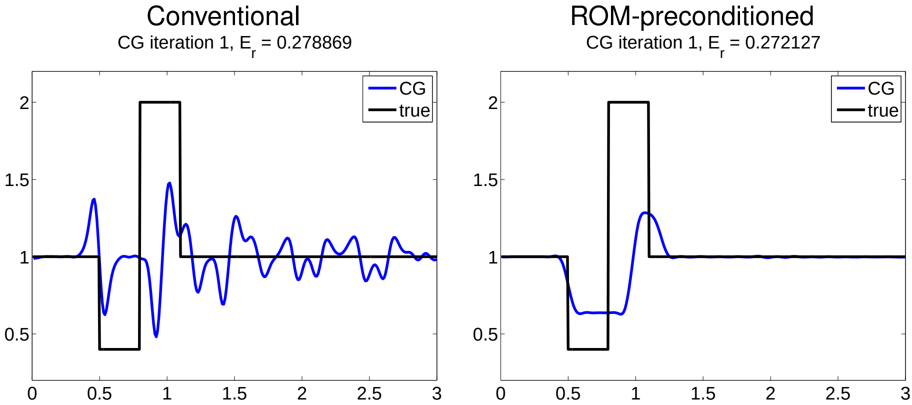
$$\underset{c}{\text{minimize}} \, \|\widetilde{\mathbf{P}}^\star - \widetilde{\mathbf{P}}(c)\|_F^2,$$

  where $\widetilde{\mathbf{P}}^\star$ is computed from the data $\mathbf{D}^\star$ and $\widetilde{\mathbf{P}}(c)$ is a (highly) nonlinear mapping

$$\widetilde{\mathbf{P}} : c \to \mathbf{A}(c) \to \mathbf{U} \to \mathbf{V} \to \widetilde{\mathbf{P}}$$

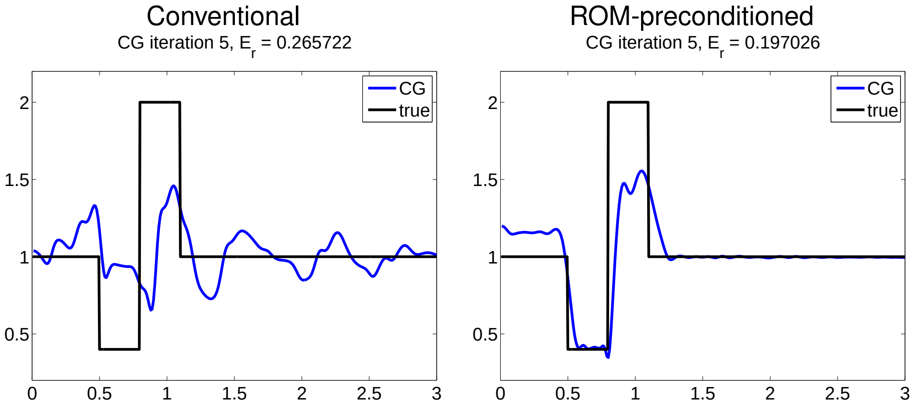
- Similar approach to **diffusive inversion** (parabolic PDE, CSEM) converges in **one Gauss-Newton iteration**

Conventional vs. ROM-preconditioned FWI in 1D



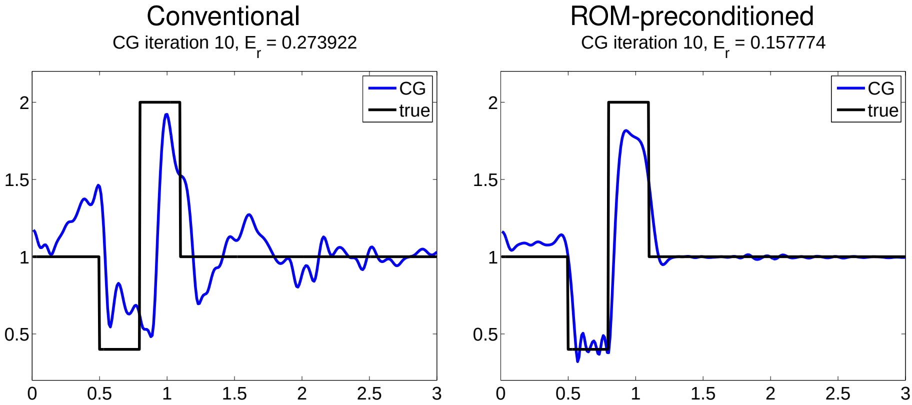
Automatic removal of multiple reflections.

Conventional vs. ROM-preconditioned FWI in 1D



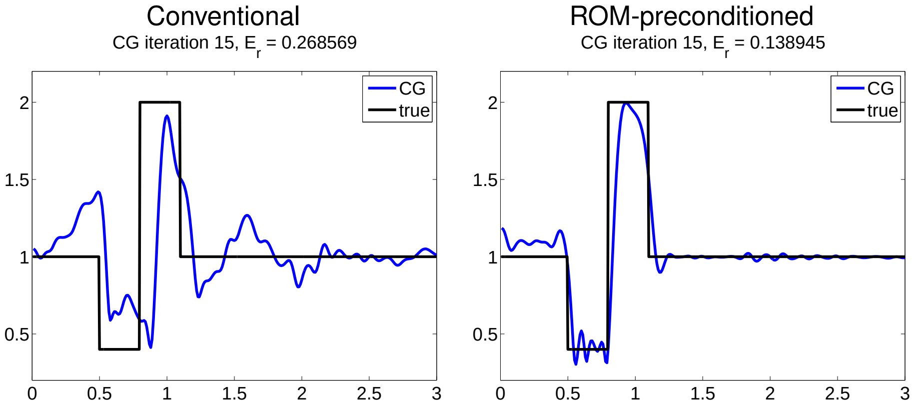
Automatic removal of multiple reflections.

Conventional vs. ROM-preconditioned FWI in 1D



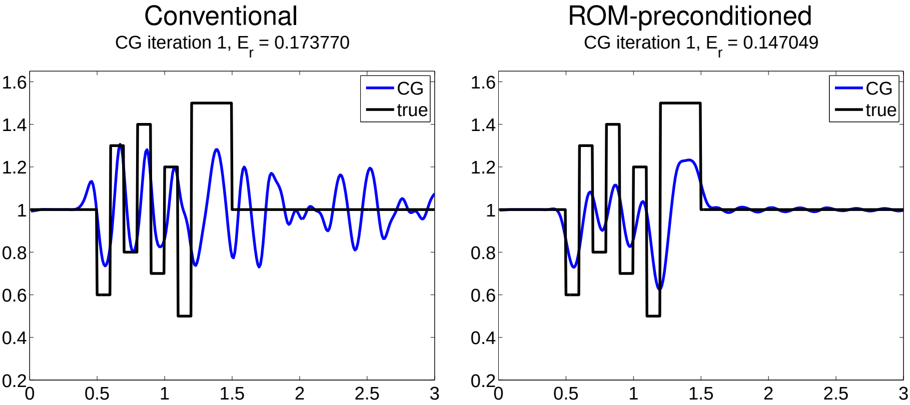
Automatic removal of multiple reflections.

Conventional vs. ROM-preconditioned FWI in 1D



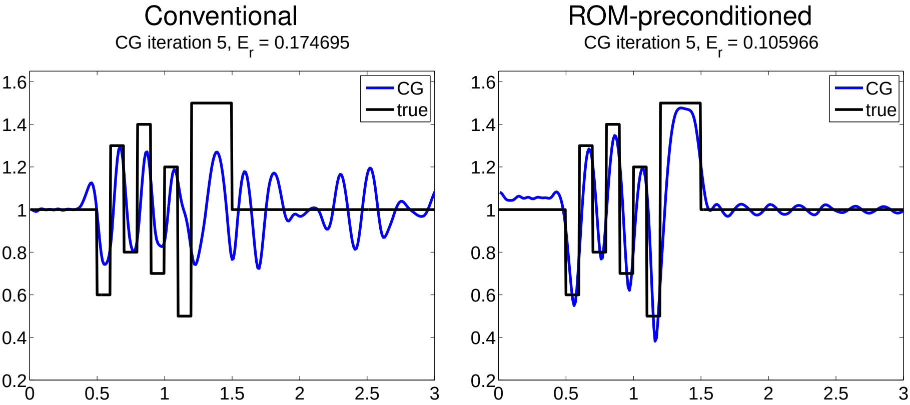
Automatic removal of multiple reflections.

Conventional vs. ROM-preconditioned FWI in 1D



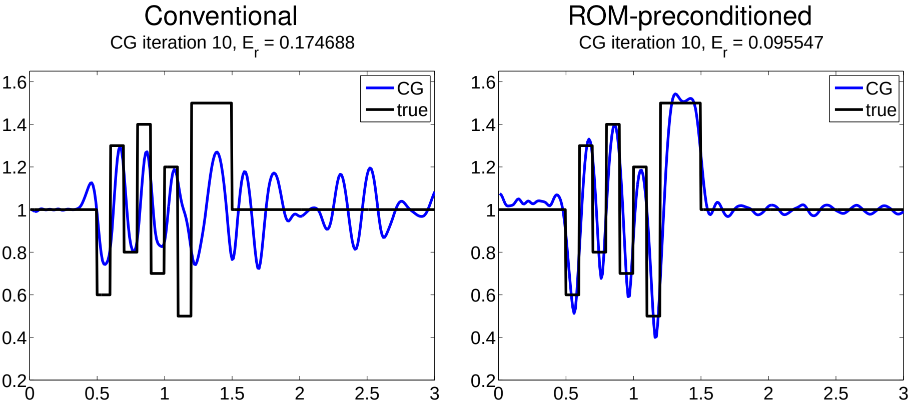
Avoiding the cycle skipping.

Conventional vs. ROM-preconditioned FWI in 1D



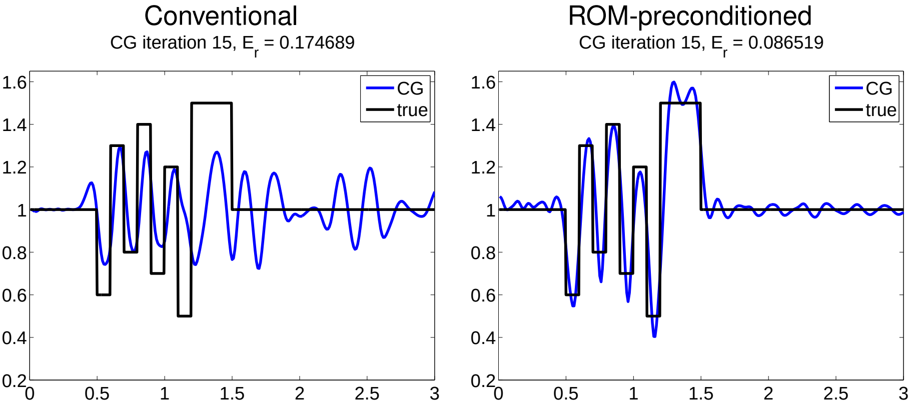
Avoiding the cycle skipping.

Conventional vs. ROM-preconditioned FWI in 1D



Avoiding the cycle skipping.

Conventional vs. ROM-preconditioned FWI in 1D



Avoiding the cycle skipping.

Problem 2: Imaging

- ROM is a projection, we can use **backprojection**
- If *span*($\mathbf{U}$) is suffiently rich, then columns of $\mathbf{V}\mathbf{V}^T$ should be good approximations of δ**-functions**, hence

$$\mathbf{P} \approx \mathbf{V}\mathbf{V}^T\mathbf{P}\mathbf{V}\mathbf{V}^T = \mathbf{V}\widetilde{\mathbf{P}}\mathbf{V}^T$$

- Problem: $\mathbf{U}$ and $\mathbf{V}$ are unknown
- We have a rough idea of **kinematics**, i.e. the **travel times**
- Equivalent to knowing a smooth **kinematic velocity model** c_0
- For known c_0 we can compute

$$\mathbf{U}_0, \quad \mathbf{V}_0, \quad \widetilde{\mathbf{P}}_0$$

Backprojection imaging functional

- Take backprojection $\mathbf{P} \approx \mathbf{V}\widetilde{\mathbf{P}}\mathbf{V}^T$ and make another approximation: replace unknown $\mathbf{V}$ with $\mathbf{V}_0$

$$\mathbf{P} \approx \mathbf{V}_0\widetilde{\mathbf{P}}\mathbf{V}_0^T$$

- For the kinematic model we know $\mathbf{V}_0$ exactly

$$\mathbf{P}_0 \approx \mathbf{V}_0\widetilde{\mathbf{P}}_0\mathbf{V}_0^T$$

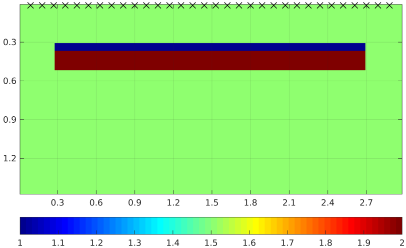
- Take the **diagonals** of backprojections to extract **approximate** Green's functions

$$G(\cdot\,,\,\cdot\,,\tau) - G_0(\cdot\,,\,\cdot\,,\tau) = \text{diag}(\mathbf{P} - \mathbf{P}_0) \approx \text{diag}\left(\mathbf{V}_0(\widetilde{\mathbf{P}} - \widetilde{\mathbf{P}}_0)\mathbf{V}_0^T\right) = \mathcal{I}$$

- Approximation quality depends **only** on how well columns of $\mathbf{V}\mathbf{V}_0^T$ and $\mathbf{V}_0\mathbf{V}_0^T$ **approximate δ-functions**

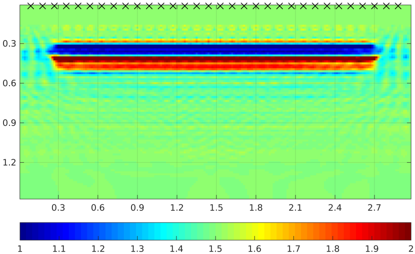
Simple example: layered model

True sound speed c

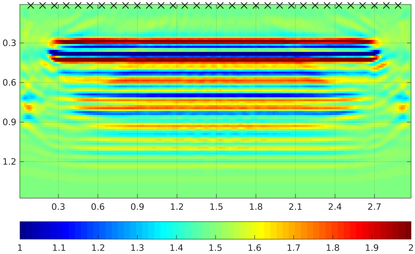


- A simple layered model, $p = 32$ sources/receivers (black $\times$)
- Constant velocity kinematic model $c_0 = 1500\ m/s$
- Multiple reflections from waves bouncing between layers and surface
- Each multiple creates an RTM artifact below actual layers

Backprojection: $c_0 + \alpha\mathcal{I}$

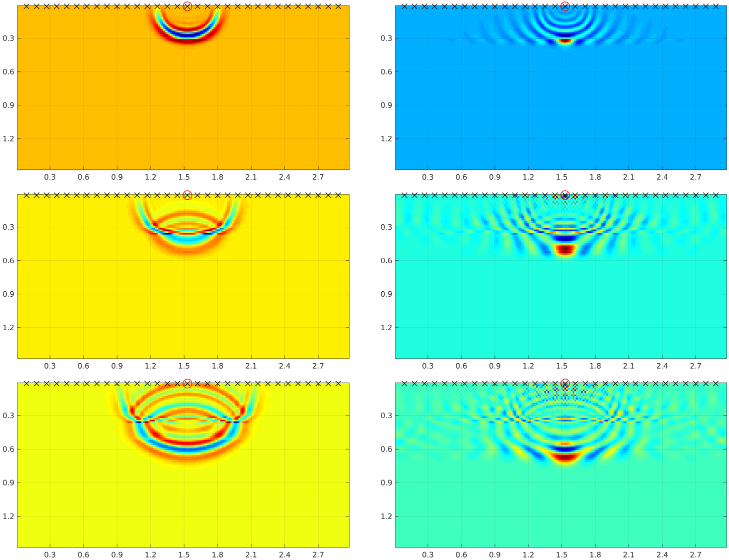


RTM image



Snapshot orthogonalization

Snapshots **U** | Orthogonalized snapshots **V**



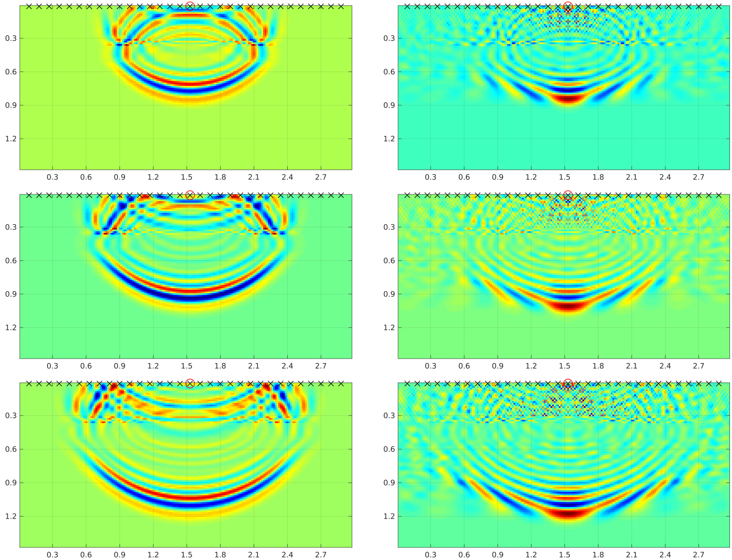
$t = 10\tau$

$t = 15\tau$

$t = 20\tau$

Snapshot orthogonalization

Snapshots **U** | Orthogonalized snapshots **V**



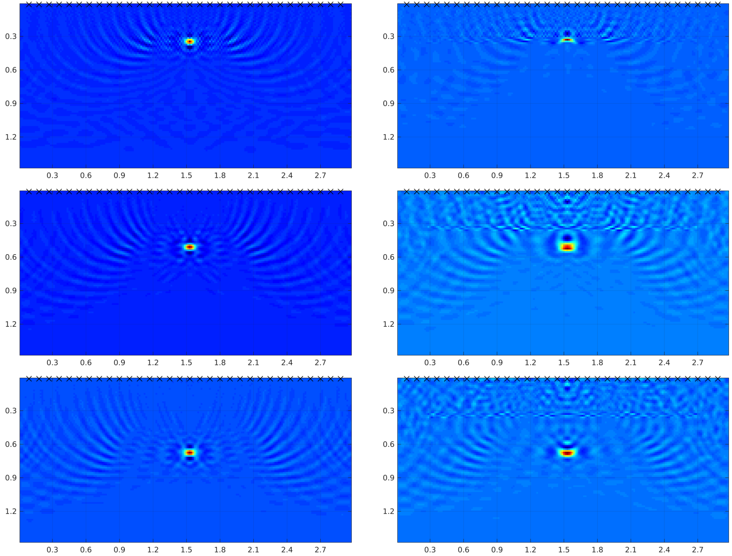
$t = 25\tau$

$t = 30\tau$

$t = 35\tau$

Approximation of δ-functions

Columns of $\mathbf{V}_0\mathbf{V}_0^T$ Columns of $\mathbf{V}\mathbf{V}_0^T$



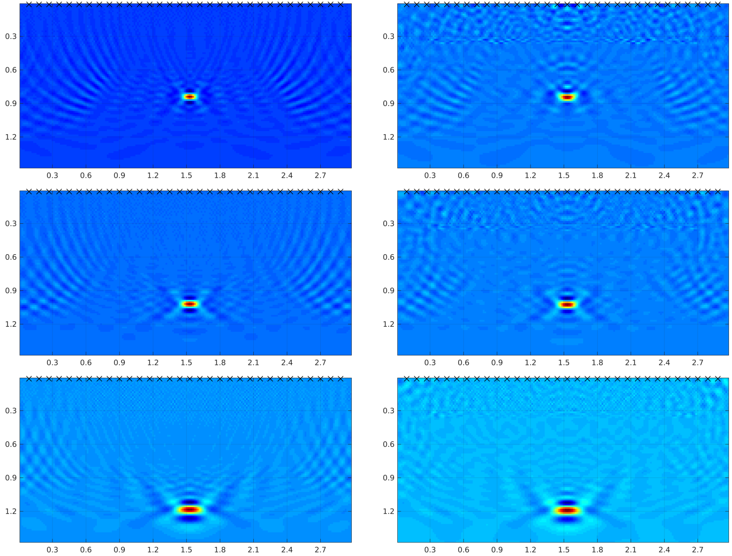
$y = 345\,m$

$y = 510\,m$

$y = 675\,m$

Approximation of δ-functions

Columns of $\mathbf{V}_0\mathbf{V}_0^T$ | Columns of $\mathbf{V}\mathbf{V}_0^T$



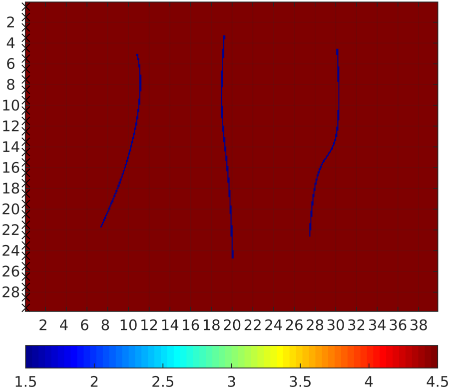
$y = 840\, m$

$y = 1020\, m$

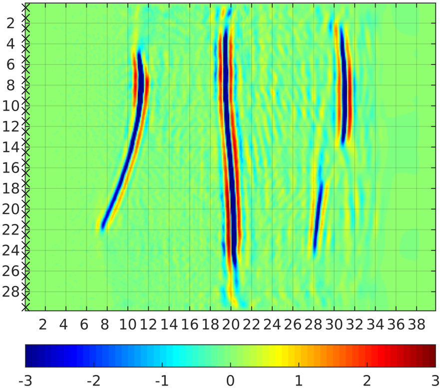
$y = 1185\, m$

High contrast example: hydraulic fractures

True c

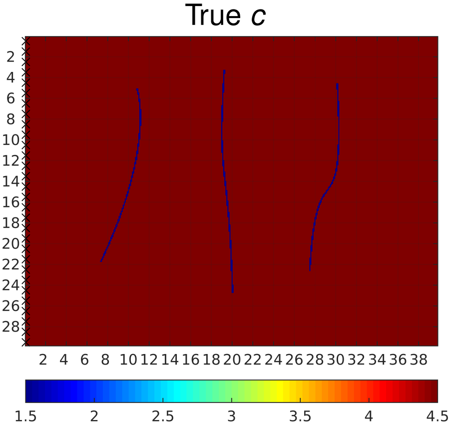


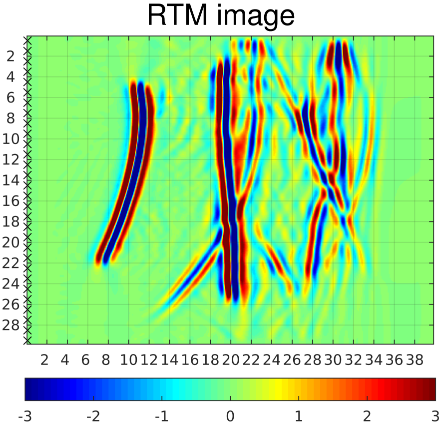
Backprojection image $\mathcal{I}$



- Important application: hydraulic fracturing
- Three fractures 10 *cm* wide each
- Very high contrasts: $c = 4500\ m/s$ in the surrounding rock, $c = 1500\ m/s$ in the fluid inside fractures

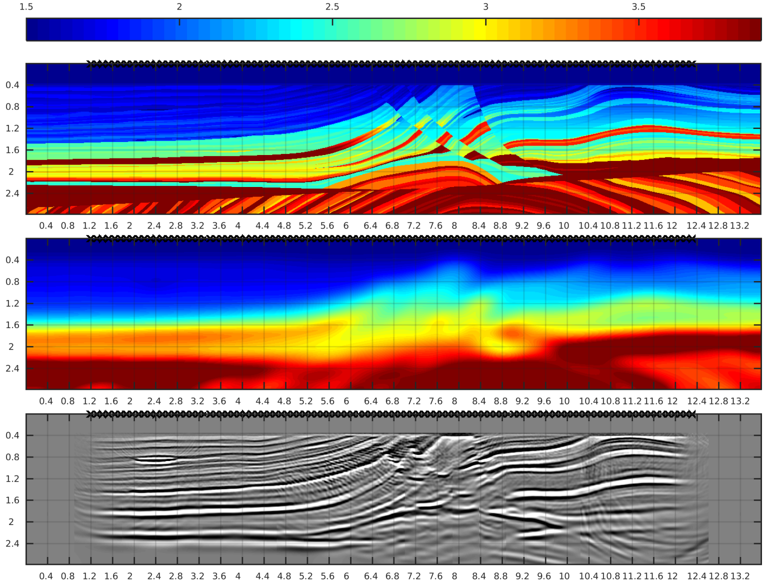
High contrast example: hydraulic fractures





- Important application: hydraulic fracturing
- Three fractures 10 *cm* wide each
- Very high contrasts: $c = 4500\ m/s$ in the surrounding rock, $c = 1500\ m/s$ in the fluid inside fractures

Backprojection imaging: Marmousi model



Problem 3: Data preprocessing

- Use **multiple-suppression** properties of ROM to preprocess data
- Compute $\widetilde{\mathbf{P}}$ from $\mathbf{D}$ and $\widetilde{\mathbf{P}}_0$ from $\mathbf{D}_0$ corresponding to c_0
- Propagator perturbation

$$\widetilde{\mathbf{P}}_\epsilon = \widetilde{\mathbf{P}}_0 + \epsilon(\widetilde{\mathbf{P}} - \widetilde{\mathbf{P}}_0)$$

- Propagate the perturbation

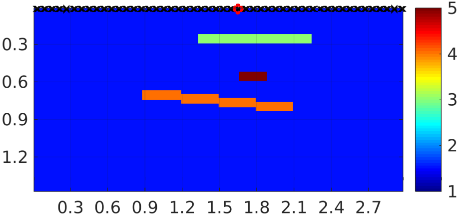
$$\mathbf{D}_{\epsilon,k} = \widetilde{\mathbf{B}}^T T_k(\widetilde{\mathbf{P}}_\epsilon)\widetilde{\mathbf{B}}$$

- Generate filtered data

$$\mathbf{F}_k = \mathbf{D}_{0,k} + \left.\frac{d\mathbf{D}_{\epsilon,k}}{d\epsilon}\right|_{\epsilon=0}$$

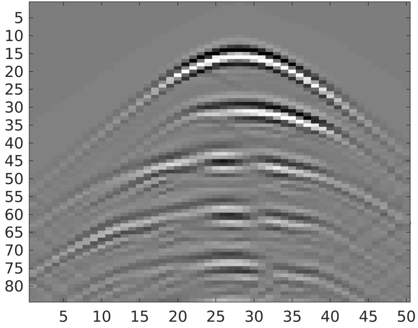
- Can show that $\mathbf{F}_k$ corresponds to data that a **Born** forward model will generate

Example: seismogram comparison

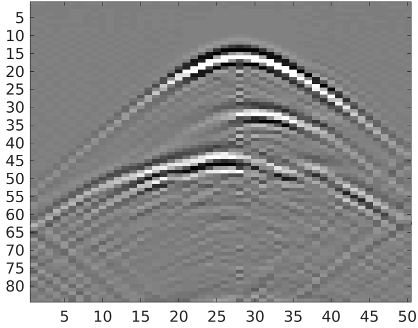


- Three direct arrivals + three multiples
- Direct arrival from small scatterer masked by the first multiple

$\mathbf{D}_k - \mathbf{D}_{0,k}$



$\mathbf{F}_k - \mathbf{D}_{0,k}$



Conclusions and future work

- **ROMs** for inversion, imaging, data preprocessing
- **Time domain** formulation is essential, linear algebraic analogues of **causality**: Gram-Schmidt, Cholesky
- Implicit **orthogonalization** of wavefield snapshots: **suppression of multiples** in backprojection imaging and data preprocessing
- Accelerated convergence, alleviated cycle-skipping in **ROM-preconditioned FWI**

Future work:

- Non-symmetric ROM for non-collocated sources/receivers
- Noise effects and stability
- ROM-preconditioned FWI in 2D/3D

References

1. *Nonlinear seismic imaging via reduced order model backprojection*, A.V. Mamonov, V. Druskin, M. Zaslavsky, **SEG Technical Program Expanded Abstracts 2015: pp. 4375–4379**.
2. *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.
3. *A nonlinear method for imaging with acoustic waves via reduced order model backprojection*, V. Druskin, A.V. Mamonov, M. Zaslavsky, 2017, **arXiv:1704.06974 [math.NA]**
4. *Untangling nonlinearity in inverse scattering with data-driven reduced order models*, L. Borcea, V. Druskin, A.V. Mamonov, M. Zaslavsky, 2017, **arXiv:1704.08375 [math.NA]**