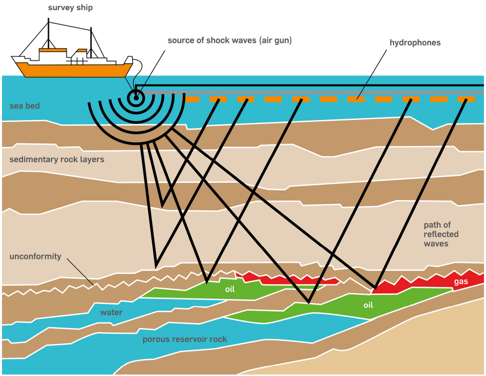
Seismic imaging and multiple removal via model order reduction

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

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

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

Motivation: seismic oil and gas exploration



Problems addressed:

1. **Imaging**: qualitative estimation of reflectors on top of known velocity model
2. **Multiple removal**: from measured data produce a new data set with only primary reflection events

- **Common framework**: data-driven 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

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

  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. **Imaging**: given $\mathbf{D}(t)$ estimate "reflectors", i.e. discontinuities of c
2. **Multiple removal**: given $\mathbf{D}(t)$ obtain "Born" data set $\mathbf{F}(t)$ with multiple reflection events removed

In both cases we are provided with a **kinematic model**, a smooth **non-reflective velocity** c_0

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** (Green's function over small time τ) is

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

- A **reduced order model** (ROM) $\widetilde{\mathbf{P}} \in \mathbb{R}^{mn\times mn}$, $\widetilde{\mathbf{B}} \in \mathbb{R}^{mn\times m}$ 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**

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

- We must project on **Krylov subspace**

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

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

ROM from measured data

- Wavefields in the whole domain $\mathbf{U}$ are **unknown**, thus $\mathbf{V}$ is unknown
- How to obtain ROM from just the data $\mathbf{D}_k$?
- Data does not give us $\mathbf{U}$, 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: 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

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

- As before, $\mathbf{U}$ and $\mathbf{V}$ are **unknown**
- We have an approximate **kinematic model**, i.e. the **travel times**
- Equivalent to knowing a **smooth velocity** c_0
- For known c_0 we can compute everything, including

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

ROM backprojection

- 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

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

- Approximate **perturbation** of the propagator

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

  is essentially the perturbation of the Green's function

$$\delta G(x, y) = G(x, y, \tau) - G_0(x, y, \tau)$$

- But $\delta G(x, y)$ depends on two variables $x, y \in \Omega$, how do we get a **single image**?

Backprojection imaging functional

- Take the **imaging functional** $\mathcal{I}$ to be

$$\mathcal{I}(x) \approx \delta G(x,x) = G(x,x,\tau) - G_0(x,x,\tau), \quad x \in \Omega$$

- In matrix form it means taking the **diagonal**

$$\mathcal{I} = \operatorname{diag}\left(\mathbf{V}_0(\widetilde{\mathbf{P}} - \widetilde{\mathbf{P}}_0)\mathbf{V}_0^T\right) \approx \operatorname{diag}(\mathbf{P} - \mathbf{P}_0)$$

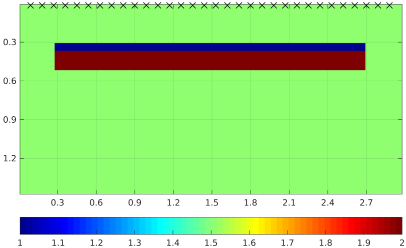
- Note that

$$\mathcal{I} = \operatorname{diag}\left([\mathbf{V}_0\mathbf{V}^T]\,\mathbf{P}\,[\mathbf{V}\mathbf{V}_0^T] - [\mathbf{V}_0\mathbf{V}_0^T]\,\mathbf{P}_0\,[\mathbf{V}_0\mathbf{V}_0^T]\right)$$

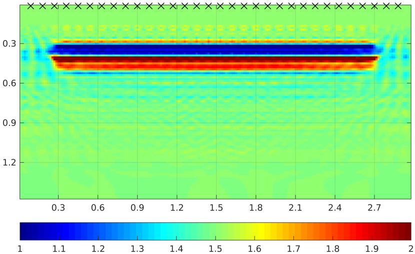
- Thus, approximation quality depends **only** on how well columns of $\mathbf{V}\mathbf{V}_0^T$ and $\mathbf{V}_0\mathbf{V}_0^T$ **approximate δ-functions**

Simple example: layered model

True c

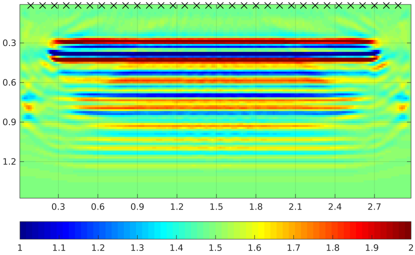


ROM backprojection image $\mathcal{I}$



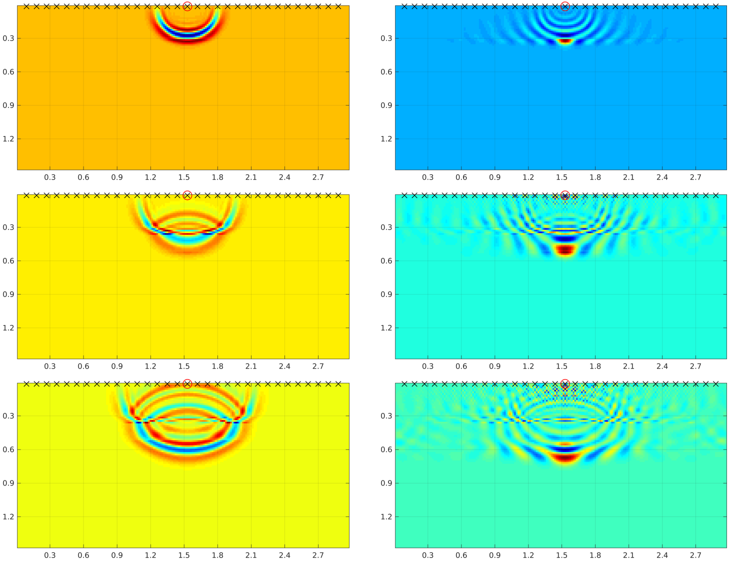
- A simple layered model, $p = 32$ sources/receivers (black $\times$)
- Constant velocity kinematic model $c_0 = 1500\ m/s$
- Multiple reflections from waves bouncing between layers and reflective top surface
- Each multiple creates an RTM artifact below actual layers

RTM image



Snapshot orthogonalization

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



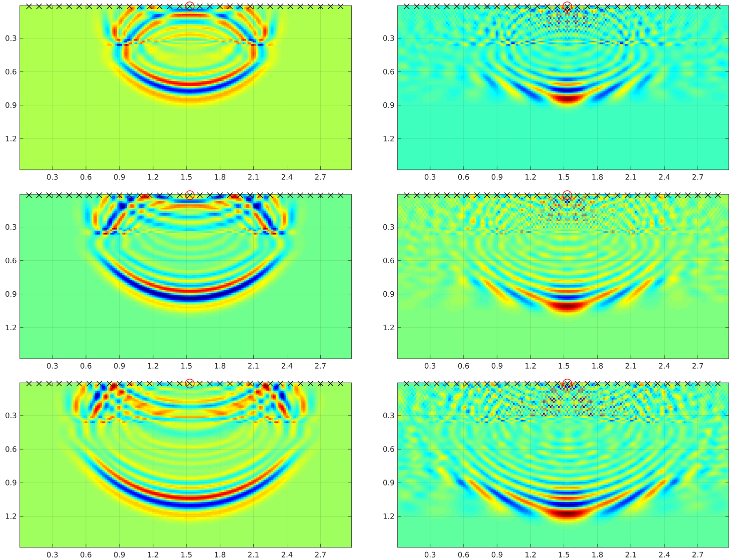
$t = 10\tau$

$t = 15\tau$

$t = 20\tau$

Snapshot orthogonalization

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



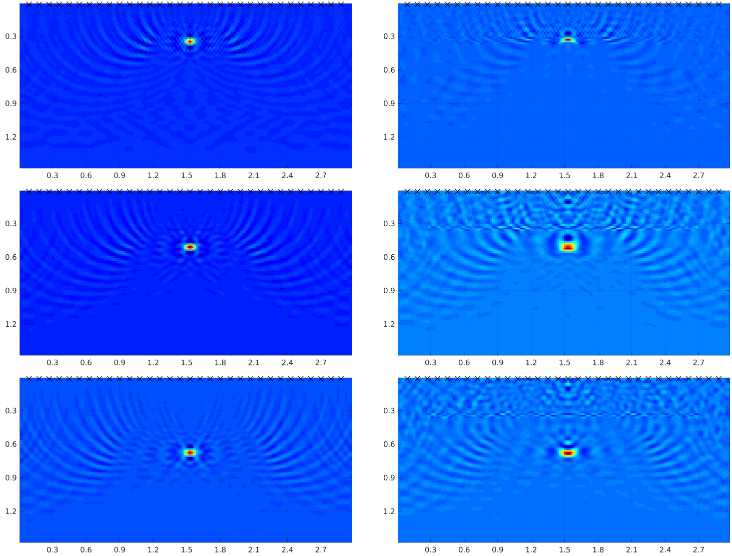
$t = 25\tau$

$t = 30\tau$

$t = 35\tau$

Approximation of δ-functions

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

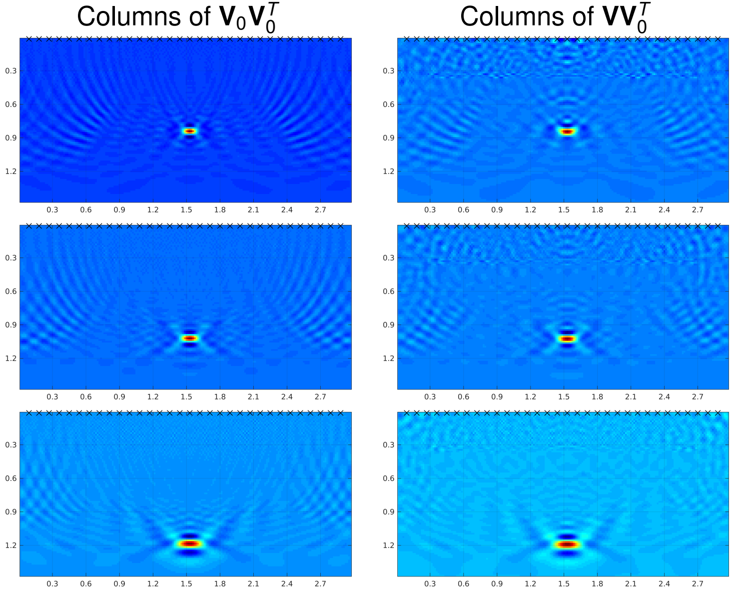


$y = 345\,m$

$y = 510\,m$

$y = 675\,m$

Approximation of δ-functions

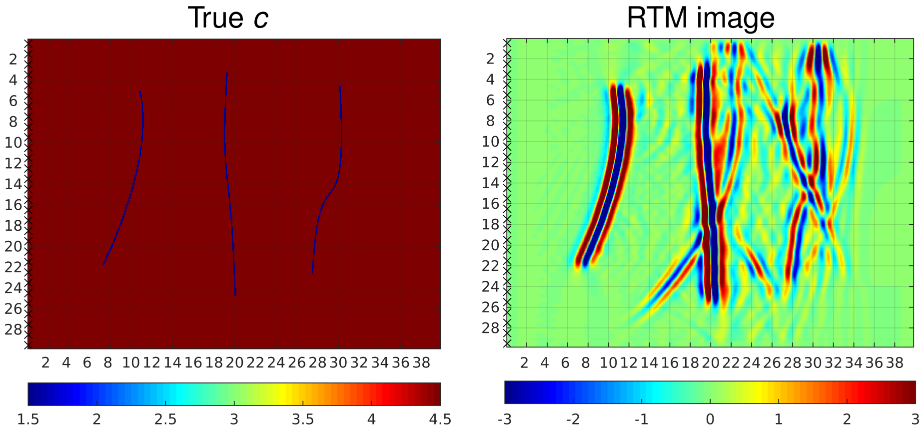


$y = 840\,m$

$y = 1020\,m$

$y = 1185\,m$

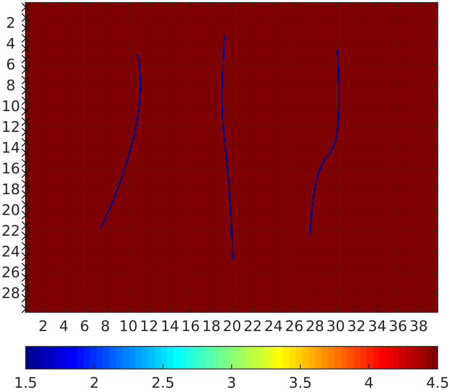
High contrast example: hydraulic fractures



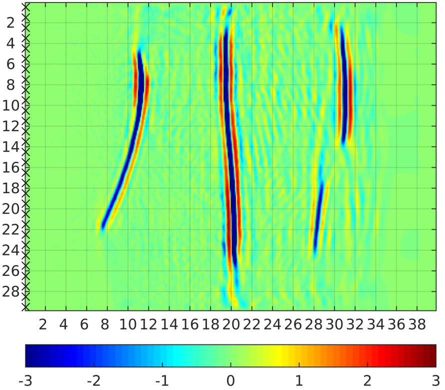
- Important application: hydraulic fracturing
- Three fractures 10 *cm* wide each
- Very high contrasts: $c = 4500\ m/s$ in the surrounding rock, $c = 1500\ m/s$ in the fluid inside fractures

High contrast example: hydraulic fractures

True c

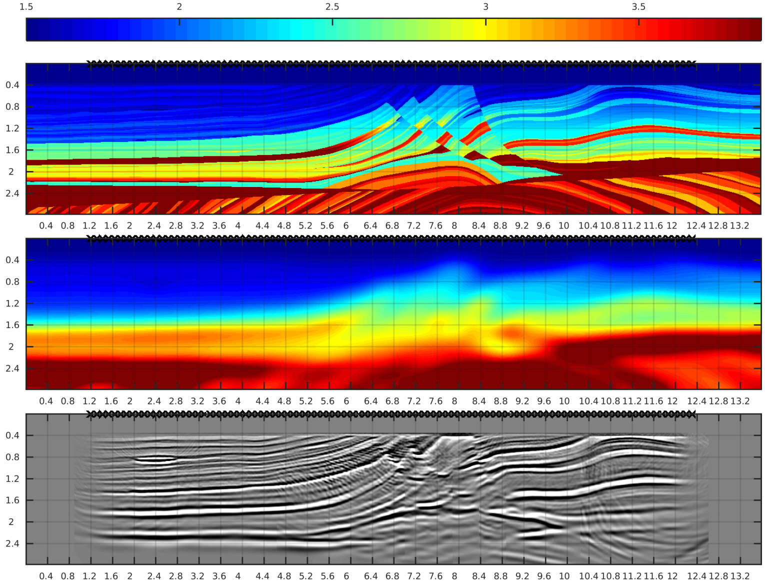


ROM backprojection image $\mathcal{I}$



- 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

Large scale example: Marmousi model



Problem 2: multiple removal

- Introduce **Data-to-Born** (DtB) transform: compute ROM from original data, then generate a new data set with primary reflection events only
- **Born** with respect to what?
- Consider wave equation in the form

$$u_{tt} = \sigma c \nabla \cdot \left(\frac{c}{\sigma} \nabla u \right),$$

  where **acoustic impedance** $\sigma = \rho c$
- Assume $c = c_0$ is a **known kinematic model**
- Only the impedance σ changes
- Above assumptions are for **derivation only**, the method works even if they are not satisfied

Born approximation

- Can show that
$$P \approx I - \frac{\tau^2}{2} L_q L_q^T,$$
where
$$L_q = -c\nabla \cdot \; + \frac{1}{2} c \nabla q \cdot, \quad L_q^T = c\nabla + \frac{1}{2} c \nabla q,$$
are **affine** in $q = \log \sigma$
- Consider **Born approximation** (linearization) with respect to q around known $c = c_0$
- Perform **second Cholesky factorization** on ROM
$$\frac{2}{\tau^2}(\widetilde{\mathbf{I}} - \widetilde{\mathbf{P}}) = \widetilde{\mathbf{L}}_q \widetilde{\mathbf{L}}_q^T$$
- Cholesky factors $\widetilde{\mathbf{L}}_q$, $\widetilde{\mathbf{L}}_q^T$ are **approximately affine** in q, thus the perturbation
$$\widetilde{\mathbf{L}}_q - \widetilde{\mathbf{L}}_0$$
is **approximately linear** in q

Data-to-Born transform

1. Compute $\widetilde{\mathbf{P}}$ from $\mathbf{D}$ and $\widetilde{\mathbf{P}}_0$ from $\mathbf{D}^0$ corresponding to $q \equiv 0$ ($\sigma \equiv 1$)
2. Perform **second Cholesky factorization**, find $\widetilde{\mathbf{L}}_q$ and $\widetilde{\mathbf{L}}_0$
3. Form the **perturbation**

$$\widetilde{\mathbf{L}}_\varepsilon = \widetilde{\mathbf{L}}_0 + \varepsilon(\widetilde{\mathbf{L}}_q - \widetilde{\mathbf{L}}_0), \quad \text{affine in } \varepsilon q$$

4. **Propagate** the perturbation

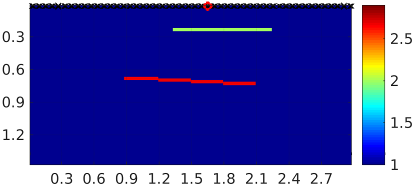
$$\mathbf{D}_k^\varepsilon = \widetilde{\mathbf{B}}^T T_k \left(\widetilde{\mathbf{I}} - \frac{\tau^2}{2} \widetilde{\mathbf{L}}_\varepsilon \widetilde{\mathbf{L}}_\varepsilon^T \right) \widetilde{\mathbf{B}}$$

5. **Differentiate** to obtain DtB transformed data

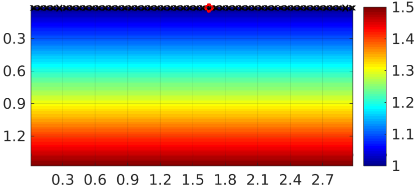
$$\mathbf{F}_k = \mathbf{D}_k^0 + \left. \frac{d\mathbf{D}_k^\varepsilon}{d\varepsilon} \right|_{\varepsilon=0}, \quad k = 0, 1, \ldots, 2n-1$$

Example: DtB seismogram comparison

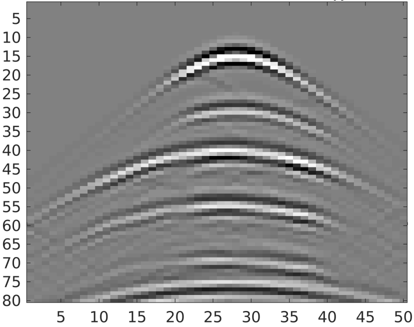
Impedance $\sigma = \rho c$



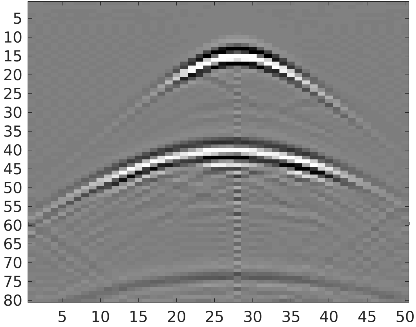
Velocity c



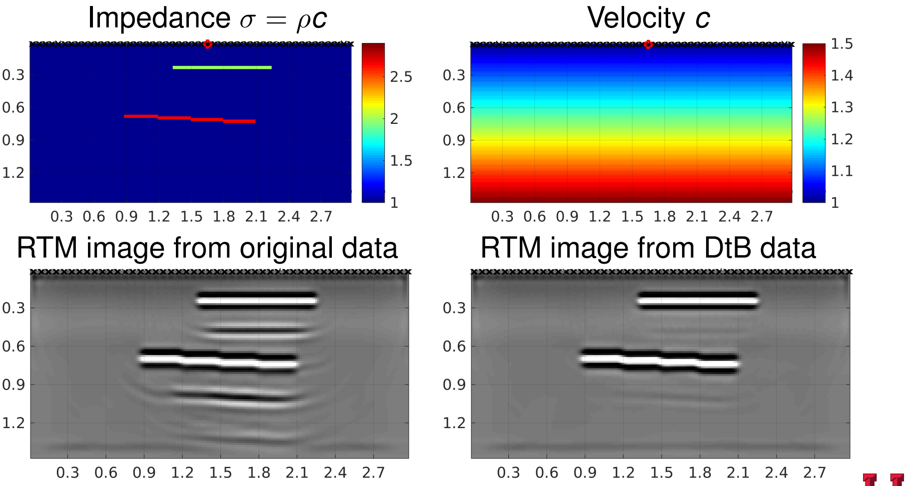
Original data $\mathbf{D}_k - \mathbf{D}_k^0$



DtB transformed data $\mathbf{F}_k - \mathbf{D}_k^0$



Example: DtB+RTM imaging



Conclusions and future work

- **ROMs** for imaging and multiple removal (DtB)
- **Time domain** formulation is essential, linear algebraic analogues of **causality**: Gram-Schmidt, Cholesky
- Implicit **orthogonalization** of wavefield snapshots: **removal of multiples** in backprojection imaging and DtB transform
- Existing linearized imaging (RTM) and inversion (LS-RTM) methods can be applied to DtB transformed data

Future work:

- **Data completion** for partial data (including monostatic, aka backscattering measurements)
- **Elasticity**: promising preliminary results
- **Stability** and noise effects (SVD truncation of the Gramian, etc.)
- **Frequency domain** analogue (data-driven PML)

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 the 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]**