Seismic inversion and imaging via model order reduction

Alexander V. Mamonov\textsuperscript{1},
Liliana Borcea\textsuperscript{2}, Vladimir Druskin\textsuperscript{3},
Andrew Thaler\textsuperscript{4} and Mikhail Zaslavsky\textsuperscript{3}

\textsuperscript{1}University of Houston,
\textsuperscript{2}University of Michigan Ann Arbor,
\textsuperscript{3}Schlumberger-Doll Research Center,
\textsuperscript{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**

\[ u_{tt} = Au \quad \text{in } \Omega, \quad t \in [0, T] \]

with initial conditions

\[ u|_{t=0} = B, \quad u_t|_{t=0} = 0, \]

sources are columns of \( B \in \mathbb{R}^{N \times m} \)

- The spatial operator \( A \in \mathbb{R}^{N \times N} \) is a (symmetrized) fine grid discretization of

\[ A = c^2 \Delta \]

with appropriate boundary conditions

- Wavefields for all sources are columns of

\[ u(t) = \cos(t\sqrt{A})B \in \mathbb{R}^{N \times m} \]
Data model and problem formulations

- For simplicity assume that sources and receivers are **collocated**, receiver matrix is also $B$
- The **data model** is
  \[
  D(t) = B^T u(t) = B^T \cos(t\sqrt{-A})B,
  \]
  an $m \times m$ matrix function of time

**Problem formulations:**

1. **Inversion**: given $D(t)$ estimate $c$
2. **Imaging**: given $D(t)$ and a smooth kinematic velocity model $c_0$, estimate “reflectors”, discontinuities of $c$
3. **Data preprocessing**: given $D(t)$ obtain $F(t)$ with multiple reflection events suppressed/removed
Data is always **discretely sampled**, say uniformly at \( t_k = k\tau \)

The choice of \( \tau \) is very important, optimally \( \tau \) around Nyquist rate

Discrete **data samples** are

\[
D_k = D(k\tau) = B^T \cos \left( k\tau \sqrt{-A} \right) B = B^T T_k(P)B,
\]

where \( T_k \) is Chebyshev polynomial and the **propagator** is

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

A **reduced order model** (ROM) \( \tilde{P}, \tilde{B} \) should fit the data

\[
D_k = B^T T_k(P)B = \tilde{B}^T T_k(\tilde{P})\tilde{B}, \quad k = 0, 1, \ldots, 2n - 1
\]
Projection ROMs

- Projection ROMs are of the form

\[ \tilde{P} = V^T P V, \quad \tilde{B} = V^T B, \]

where \( V \) is an orthonormal basis for some subspace

- What subspace to project on to fit the data?

- Consider a matrix of \textit{wavefield snapshots}

\[ U = [u_0, u_1, \ldots, u_{n-1}] \in \mathbb{R}^{N \times nm}, \quad u_k = u(k\tau) = T_k(P)B \]

- We must project on \textit{Krylov subspace}

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

- The data only knows about what \( P \) does to wavefield snapshots \( u_k \)
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 $u_k = T_k(\mathbf{P})\mathbf{B}$ and $\mathbf{D}_k = \mathbf{B}^T T_k(\mathbf{P})\mathbf{B}$ we get

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

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

- Suppose \( \mathbf{U} \) is orthogonalized by a block QR (Gram-Schmidt) procedure
  \[
  \mathbf{U} = \mathbf{V} \mathbf{L}^T,
  \]
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
  \[
  \mathbf{U}^T \mathbf{U} = \mathbf{LL}^T
  \]
- The projection is given by
  \[
  \tilde{\mathbf{P}} = \mathbf{V}^T \mathbf{P} \mathbf{V} = \mathbf{L}^{-1} \left( \mathbf{U}^T \mathbf{P} \mathbf{U} \right) \mathbf{L}^{-T},
  \]
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)
  \[\min_{c} \|D^{*} - D(\cdot; c)\|^{2}_{2}\]

- Replace the objective with a "nonlinearly preconditioned" functional
  \[\min_{c} \|\tilde{P}^{*} - \tilde{P}(c)\|^{2}_{F},\]
  where \(\tilde{P}^{*}\) is computed from the data \(D^{*}\) and \(\tilde{P}(c)\) is a (highly) nonlinear mapping
  \[\tilde{P} : c \rightarrow A(c) \rightarrow U \rightarrow V \rightarrow \tilde{P}\]

- Similar approach to **diffusive inversion** (parabolic PDE, CSEM) converges in **one Gauss-Newton iteration**
Conventional vs. ROM-preconditioned FWI in 1D

Conventional
CG iteration 1, $E_r = 0.278869$

ROM-preconditioned
CG iteration 1, $E_r = 0.272127$

Automatic removal of multiple reflections.
Conventional vs. ROM-preconditioned FWI in 1D

Conventional
CG iteration 5, $E_r = 0.265722$

ROM-preconditioned
CG iteration 5, $E_r = 0.197026$

Automatic removal of multiple reflections.
Conventional vs. ROM-preconditioned FWI in 1D

Conventional

CG iteration 10, $E_r = 0.273922$

ROM-preconditioned

CG iteration 10, $E_r = 0.157774$

Automatic removal of multiple reflections.
Automatic removal of multiple reflections.
Conventional vs. ROM-preconditioned FWI in 1D

Conventional
CG iteration 1, $E_r = 0.173770$

ROM-preconditioned
CG iteration 1, $E_r = 0.147049$

Avoiding the cycle skipping.
Conventional vs. ROM-preconditioned FWI in 1D

Conventional
CG iteration 5, $E_r = 0.174695$

ROM-preconditioned
CG iteration 5, $E_r = 0.105966$

Avoiding the cycle skipping.
Conventional vs. ROM-preconditioned FWI in 1D

Avoiding the cycle skipping.

Conventional
CG iteration 10, $E_r = 0.174688$

ROM-preconditioned
CG iteration 10, $E_r = 0.095547$
Conventional vs. ROM-preconditioned FWI in 1D

Avoiding the cycle skipping.

Conventional
CG iteration 15, $E_r = 0.174689$

ROM-preconditioned
CG iteration 15, $E_r = 0.086519$
Problem 2: Imaging

- ROM is a projection, we can use **backprojection**

- If $\text{span}(\mathbf{U})$ is sufficiently rich, then columns of $\mathbf{V}\mathbf{V}^T$ should be good approximations of $\delta$-functions, hence

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

- 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

$$
\mathbf{U}_0, \quad \mathbf{V}_0, \quad \mathbf{\tilde{P}}_0
$$
Backprojection imaging functional

- Take backprojection $P \approx \tilde{V}PV^T$ and make another approximation: replace unknown $V$ with $V_0$

$$P \approx V_0\tilde{P}V_0^T$$

- For the kinematic model we know $V_0$ exactly

$$P_0 \approx V_0\tilde{P}_0V_0^T$$

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

$$G(\cdot, \cdot, \tau) - G_0(\cdot, \cdot, \tau) = \text{diag}(P - P_0) \approx \text{diag} \left( V_0(\tilde{P} - \tilde{P}_0)V_0^T \right) = I$$

- Approximation quality depends **only** on how well columns of $VV_0^T$ and $V_0V_0^T$ approximate $\delta$-functions
A simple layered model, \( p = 32 \) sources/receivers (black ×)

Constant velocity kinematic model \( c_0 = 1500 \text{ m/s} \)

Multiple reflections from waves bouncing between layers and surface

Each multiple creates an RTM artifact below actual layers
Snapshot orthogonalization

**Snapshots U**

**Orthogonalized snapshots V**

\[
\begin{align*}
t &= 10\tau \\
t &= 15\tau \\
t &= 20\tau
\end{align*}
\]
Snapshot orthogonalization

Snapshots $\mathbf{U}$

Orthogonalized snapshots $\mathbf{V}$

$t = 25\tau$

$t = 30\tau$

$t = 35\tau$
Approximation of $\delta$-functions

Columns of $V_0V_0^T$

$y = 345 \text{ m}$

$y = 510 \text{ m}$

$y = 675 \text{ m}$

Columns of $VV_0^T$
Approximation of $\delta$-functions

Columns of $V_0V_0^T$

Columns of $VV_0^T$

$y = 840 \text{ m}$

$y = 1020 \text{ m}$

$y = 1185 \text{ m}$
High contrast example: hydraulic fractures

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

- Important application: hydraulic fracturing
- Three fractures 10 cm wide each
- Very high contrasts: $c = 4500 \text{ m/s}$ in the surrounding rock, $c = 1500 \text{ m/s}$ in the fluid inside fractures
Backprojection imaging: Marmousi model
Problem 3: Data preprocessing

- Use **multiple-suppression** properties of ROM to preprocess data
- Compute $\tilde{P}$ from $D$ and $\tilde{P}_0$ from $D_0$ corresponding to $c_0$
- Propagator perturbation

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

- Propagate the perturbation

$$D_{\epsilon,k} = \tilde{B}^T T_k (\tilde{P}_\epsilon) \tilde{B}$$

- Generate filtered data

$$F_k = D_{0,k} + \left. \frac{dD_{\epsilon,k}}{d\epsilon} \right|_{\epsilon=0}$$

- Can show that $F_k$ 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

\[ D_k - D_{0,k} \]

\[ F_k - D_{0,k} \]
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


