Nonlinear Differential Semblance: Concept and Implementation

Dong Sun

The Rice Inversion Project

Annual Review Meeting

Dec 9, 2010

Brief Intro

Problem:

produce well-focused subsurface image from reflection seismic data

Method: combine MVA and WI (WWS 08, 'MVA and WI')

- migration velocity analysis (MVA) – current industry approach
  - ✔ provides robust estimation of large scale structure (macro model)
  - × relies on Born approximation and neglects non-linear effects
- least squares waveform inversion (WI) – automatic data fitting
  - ✔ provides remarkably detailed structure and accounts for nonlinear effects
  - × not robust (depends on accurate initial guess, etc.)

Key idea :

using very low data frequencies as controls in waveform inversion

(just like using macro models as controls in MVA)

RICE

Brief Intro

Nonlinear Differential Semblance Optimization (nDSO):

- D. Sun (2008) introduced nDSO for layered media and demonstrated the smoothness and convexity of its objective
- Towards generalizing and implementing nDSO (D. Sun's Ph.D. project)
  - formed nDSO in a general form via Extended Modeling Concept
  - derived and constructed the gradient computation
  - implementing inversion framework (IWAVE++)
    - ✳ done: provided a platform for various imaging/inversion applications (RTM, standard LS waveform inversion, etc.)
    - ✳ to do: complement extended inversion block, documentation, ...

Agenda

1. *Motivation of nDSO*

2. *Nonlinear Differential Semblance (nDS) Strategy*
   - Reformulate WI as DSO
   - Implementation

3. *Summary and Future Work*

Waveform Inversion (WI)

WI: model-based data-fitting procedure often formulated as **least squares inversion/minimization**

$$\min_{m \in \mathcal{M}} J_{LS} := \frac{1}{2} \|\mathcal{F}[m] - d_o\|_{\mathcal{D}}^2 + \mathcal{R}(m)$$

- $\mathcal{F} \; : \mathcal{M} \rightarrow \mathcal{D}$: Forward Operator (defined via Wave Equation(s))
- $\mathcal{M}$: Model Space ($\{m(\mathbf{x})\}$: set of models of substructure, e.g., velocity)
- $\mathcal{D}$: Data Space ($\{d(\mathbf{x}_r, t; \mathbf{x}_s)\}$: set of traces)
- $d_o \in \mathcal{D}$: observed data, **highly redundant** & **band-limited**
- $\mathcal{R}(m)$: Regularization (e.g., TV, Tikhonov, ...)

RICE

Waveform Inversion (WI)

WI basic facts:

- large scale, nonlinear least-squares optimization driven by expensive simulation
- only gradient-related methods (local methods) computationally feasible (adjoint state method for gradient computation)
- hard for reflection data, easier for transmission but nontrivial
- key issue: interaction of spectral data incompleteness with strong nonlinearity (**local minima issue**)

How to address local minima issue?

- solve LS inversion with specific procedures

  ✳ multiscale inversion (Bunks 95, etc.)

  ✳ continuation in depth, time, frequency (Kolb et al 86, etc.)

  — **major observation: with very low frequencies, LS inversion is solvable**

- update background model and model perturbation alternatively - MVA - solution approach to Born FWI (lots of DS variants)

  — **can MVA's concepts (image gathering, semblance measuring) be imported into FWI?**

  $\Longrightarrow$ Nonlinear Differential Semblance Optimization — doing inversion over new control space with "convexified" objective

Agenda

1. Motivation of nDSO
2. Nonlinear Differential Semblance (nDS) Strategy
   - Reformulate WI as DSO
   - Implementation
3. Summary and Future Work

RICE

Reformulate WI as DSO

Linearized Case based on Born assumption $m := m_b + \delta m$

- data decomposition $\mathcal{D} = \prod_s \mathcal{D}_s$ (s, shot position, slowness, ...)
  D: primary reflections (without multiple)
- migration/linearized inversion $d_s \longrightarrow \delta m_s$ via $D\mathcal{F}[m_b]^T(d_s - \mathcal{F}[m_b])$ or $D\mathcal{F}[m_b]^\dagger(d_s - \mathcal{F}[m_b])$
- adjust m_b to reduce incoherence of $\delta\bar{m}(.,s)(:= \delta m_s)$ along s
- automatic MVA process – DSO, updating m_b to minimize perturbation incoherence

$$\min_{m_b} \quad J_{DS}[m_b] := \frac{1}{2}\left\|\frac{\partial \delta\bar{m}[m_b]}{\partial s}\right\|^2$$

$$\texttt{s.t.} \quad \delta\bar{m}[m_b] = D\mathcal{F}[m_b]^T(d_s - \mathcal{F}[m_b])$$

$$\texttt{or} \quad D\mathcal{F}[m_b]^\dagger(d_s - \mathcal{F}[m_b])$$

Lots of successful applications; but, nonlinear effects not included

Reformulate WI as DSO

Nonlinear Case

- data decomposition $\mathcal{D} = \prod_s \mathcal{D}_s$ (s, shot position, slowness, ...)
  D: contains all info
- nonlinear inversion $d_o \longrightarrow \bar{m}(.,s)$ ($:= m_s$ redundant models) via

$$\min_{\bar{m}\in\overline{\mathcal{M}}} \frac{1}{2}\sum_s \|\mathcal{F}[\bar{m}(.,s)] - (d_o)_s\|^2$$

- measure incoherence of $\bar{m}(.,s)$ along s via **DS functional**, e.g.,

$$J_{DS}[\bar{m}] := \frac{1}{2}\left\|\frac{\partial}{\partial s}\bar{m}\right\|^2$$

Note: $\bar{m}(\mathbf{x},s) := m_s(\mathbf{x})$ v.s. $m(\mathbf{x})$ $(\mathcal{M} \subset \overline{\mathcal{M}})$

RICE

Reformulate WI as nonlinear DSO

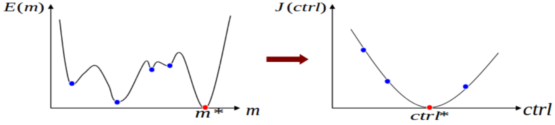
$$\min_{\bar{m} \in \overline{\mathcal{M}}} \quad J_{DS}[\bar{m}] := \frac{1}{2} \left\| \frac{\partial}{\partial s} \bar{m} \right\|^2 \qquad \texttt{(coherency)}$$

$$\texttt{s.t.} \quad \left\| \overline{\mathcal{F}}[\bar{m}] - d_o \right\|^2 \approx 0 \qquad \texttt{(data-fitting)}$$

$$\left\| \overline{\mathcal{F}}[\bar{m}] - d_o \right\|^2 := \sum_s \| \mathcal{F}[\bar{m}(., s)] - (d_o)_s \|^2$$

Key: need a proper control parameter (as m_b in linearized case), via updating which to navigate through the feasible model set

$$\left\{ \bar{m} \in \overline{\mathcal{M}} : E[\bar{m}] = \left\| \overline{\mathcal{F}}[\bar{m}] - d_o \right\|^2 \approx 0 \right\}$$



Formulate nDSO over Low-freq Control

Need new control, analogous to m_b (long-scale/low-frequency model structure) (DS 08, DS & WWS 09) :

- main observation: the solvability of the impulsive inverse problem:

  for source and data with full bandwidth down to 0 Hz (impulsive),

  least-squares inversion leads to "unique" model

  (WWS 86, Bunks et al 95, Shin & Min 2006, ...)

- **Key: use low-frequency data components as control**

Nonlinear DSO :

$$\min_{d_l \in \mathcal{D}_l} \quad J_{DS}[\bar{m}[d_l]] := \frac{1}{2} \left\| \frac{\partial}{\partial s} \bar{m}[d_l] \right\|^2$$

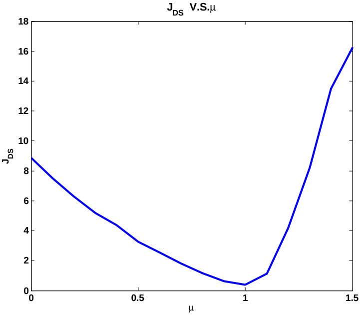
$$\texttt{s.t.} \quad \bar{m}[d_l] = \operatorname*{argmin}_{\bar{m} \in \overline{\mathcal{M}}} \mathcal{E}[\bar{m}],$$

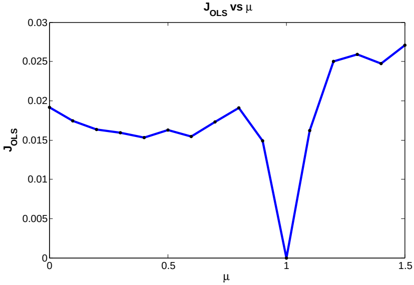
$$\mathcal{E}[\bar{m}] = \tfrac{1}{2} \left\| \overline{\mathcal{F}}[\bar{m}] - (d_o + d_l) \right\|^2 + \mathcal{R}(\bar{m})$$

Objective Scan in Layered Media

Scan J_{OLS} along: $c_\mu(z) = (1-\mu)\, c_{hom} + \mu\, c^*(z) \qquad (\mu \in [0, 1.5])$

Scan J_{DS} along : $\quad d_l(\mu) = (1-\mu)\, d_{l,pert} + \mu\, d_l^* \qquad (\mu \in [0, 1.5])$

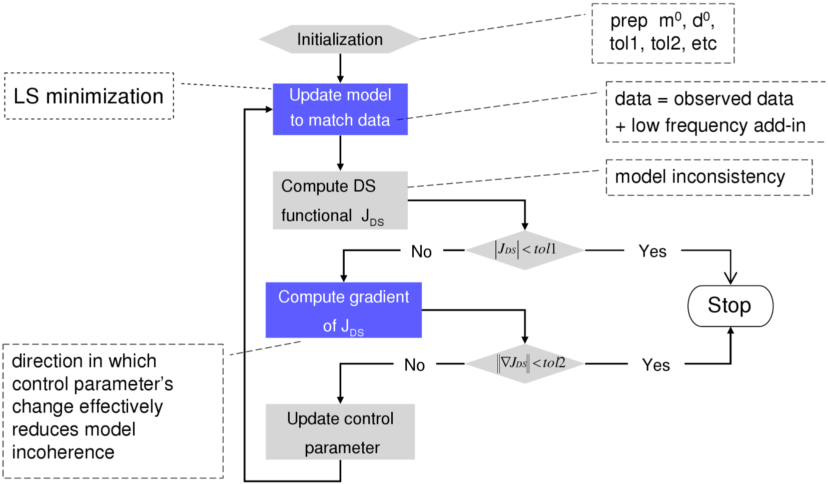




Favorable properties of the nDS Objective (Sun, 2008):

- **convex**
- **continuously differentiable** $\cdots$

nDSO flow



Initialization
prep m^0, d^0, tol1, tol2, etc
LS minimization
Update model to match data
data = observed data + low frequency add-in
Compute DS functional J_{DS}
model inconsistency
No
$|J_{DS}| < tol1$
Yes
Stop
Compute gradient of J_{DS}
No
$\|\nabla J_{DS}\| < tol2$
Yes
Update control parameter
direction in which control parameter's change effectively reduces model incoherence

Implementation

Main blocks:

- **Sub-LS Minimization (Inversion)**:

$$\min_{\bar{m}\in\overline{\mathcal{M}}} \mathcal{E}[\bar{m}] = \frac{1}{2}\left\|\overline{\mathcal{F}}[\bar{m}] - (d_l + d_o)\right\|^2 + \mathcal{R}(\bar{m})$$

Gradient and Gauss-Newton Hessian:

$$\nabla\mathcal{E}[\bar{m}] = D\overline{\mathcal{F}}[\bar{m}]^T\left(\overline{\mathcal{F}}[\bar{m}] - (d_o + d_l)\right) + \nabla\mathcal{R}(\bar{m})$$

$$H[\bar{m}] = D\overline{\mathcal{F}}[\bar{m}]^T\, D\overline{\mathcal{F}}[\bar{m}] + D^2\mathcal{R}(m)$$

- **nDSO Gradient**:

$$\nabla J_{DS}[d_l] = \Pi\, \underbrace{\overbrace{D\overline{\mathcal{F}}[\bar{m}[d_l]]\; \underbrace{H[\bar{m}[d_l]]^{-1}\left(\frac{\partial}{\partial s}\right)^T \frac{\partial}{\partial s}\,\bar{m}[d_l]}_{\in\overline{\mathcal{M}}}}^{\in\overline{\mathcal{D}}}}_{\in\mathcal{D}_l}$$

$\Pi : \overline{\mathcal{D}} \to \mathcal{D}_l$ be the orthogonal projector (low-pass filter)

RICE

Implementation

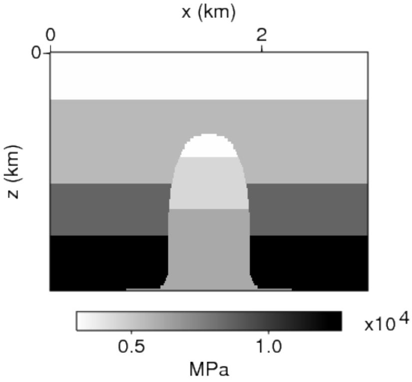
A framework implementation of inversion (IWAVE++) based on IWAVE, RVL, TSOPT, UMIN:

- Modeling operator (`IWaveOp`), with methods to compute
  - forward modeling $\mathcal{F}[m]$
  - derivative (action of Born Map $D\mathcal{F}[m]$)
  - adjoint derivative (adjoint action of Born Map $D\mathcal{F}[m]^T$)

  makes IWAVE capabilities - large-scale, parallel, extensible FD modeling - available for inversion/migration applications
- Inversion operator (`IWaveInvOp`), with methods to
  - maps d to m via optimization process
  - apply derivative map ($\delta d \rightarrow \delta m$)
  - apply adjoint derivative ($\lambda m \rightarrow \lambda d$)

  wraps inversion procedure and accommodates various performance-improving techniques
- Annihilator, e.g., differentiation operator (`GridDiffOp`)

Dome Model:



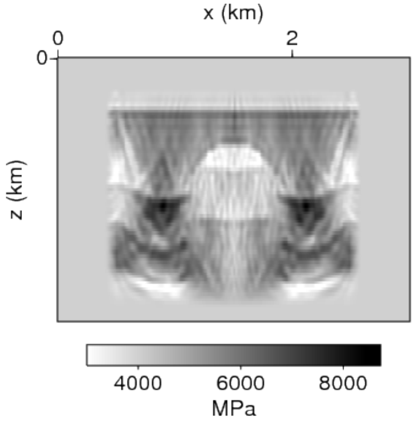
- 1 km $\times$ 3 km domain
- background: v $=$ 2 km/s, $\rho = 1$ g/cm^3
- plane-wave src (0 - 30 Hz)

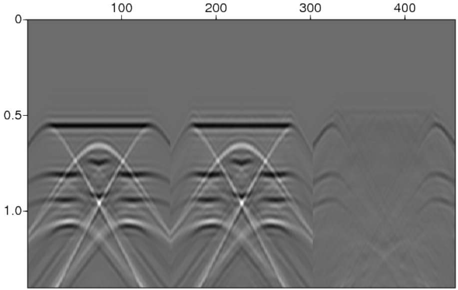
- 10 m grid $\sim$ 7 gridpts / (minimum) wavelength
- absorbing BCs on all sides
- src depth 10 m
- receiver locations $(i * 20, 30)$ m, $i = 0, \ldots, 150$

Numerical Experiments — Inversion Operator

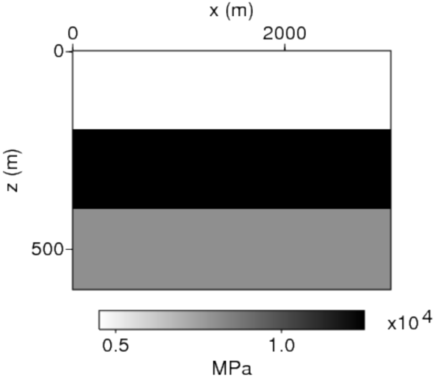
Inversion — 3 sub-runs with increasing freq bands – each with 10 LBFGS iter

Initial RMS resid $= 5.5 \times 10^8$; Final RMS resid $= 7.9 \times 10^7$



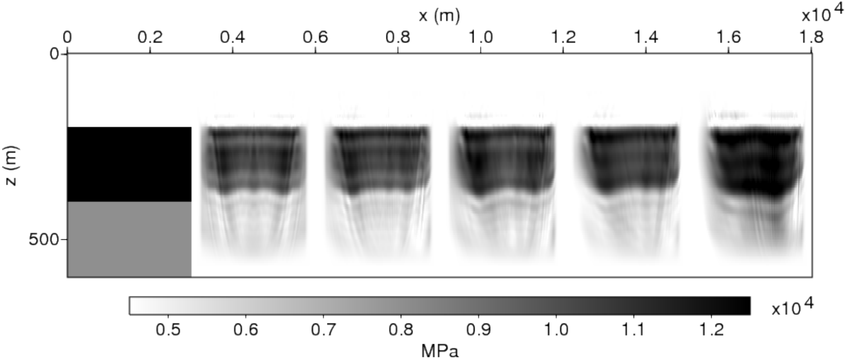


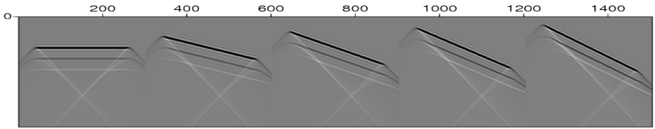
Layered Model:



- 0.6 km $\times$ 3 km domain
- v $=$ 1.5, 2.5, 2 km/s, $\rho =$ 1 g/cm^3
- plane-wave src (0 - 30 Hz)
- slowness 0, 0.15, 0.21, 0.26, 0.3
- absorbing (next test with free surface) BCs

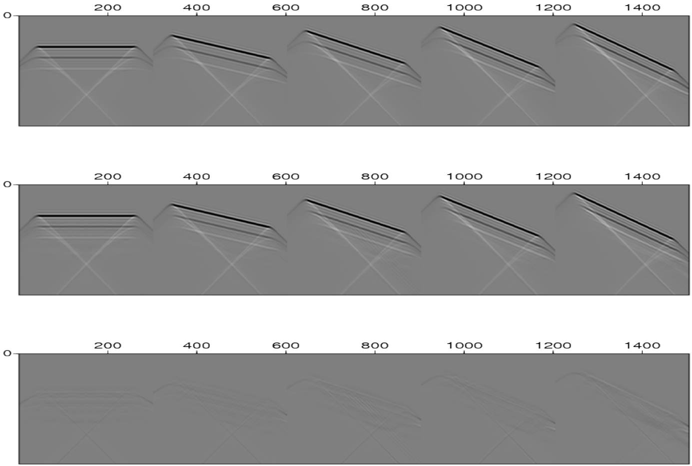
Inversion for data with correct low-frequency components



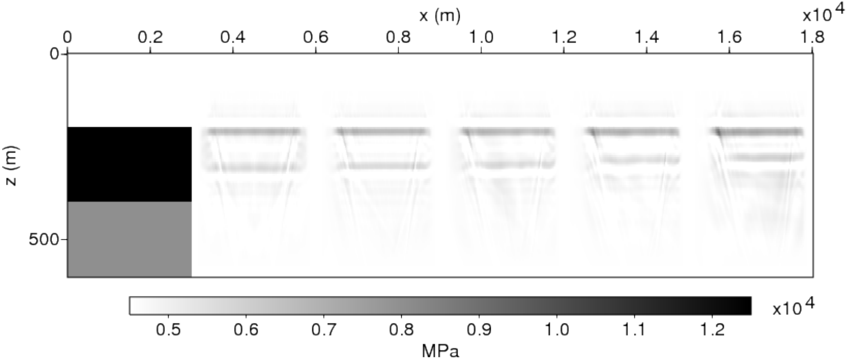


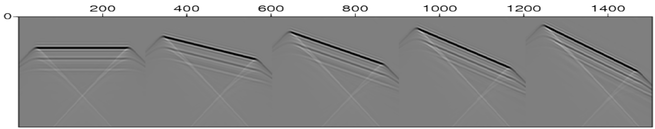
Numerical Experiments — layered test with absorbing surface

Initial RMS resid $= 4.9 \times 10^9;\ 6.2 \times 10^9;\ 8.0 \times 10^9;\ 1.1 \times 10^{10};\ 1.5 \times 10^{10};$
Final RMS resid $= 4.1 \times 10^8;\ 3.7 \times 10^8;\ 6.6 \times 10^8;\ 5.0 \times 10^8;\ \ 9.1 \times 10^8$



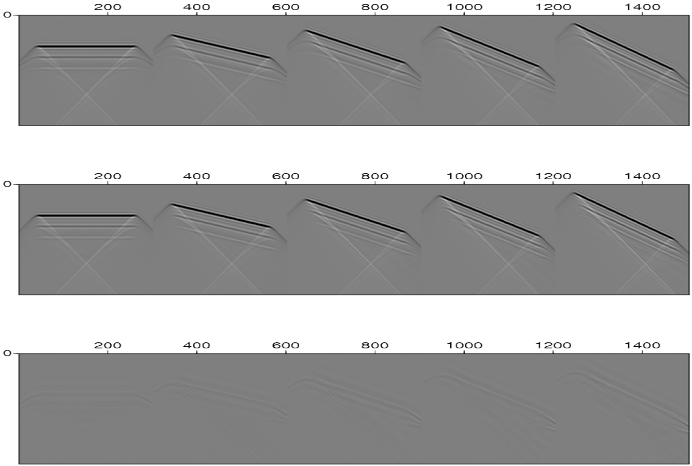
Inversion for data with wrong low-frequency components (driven from homogeneous model)



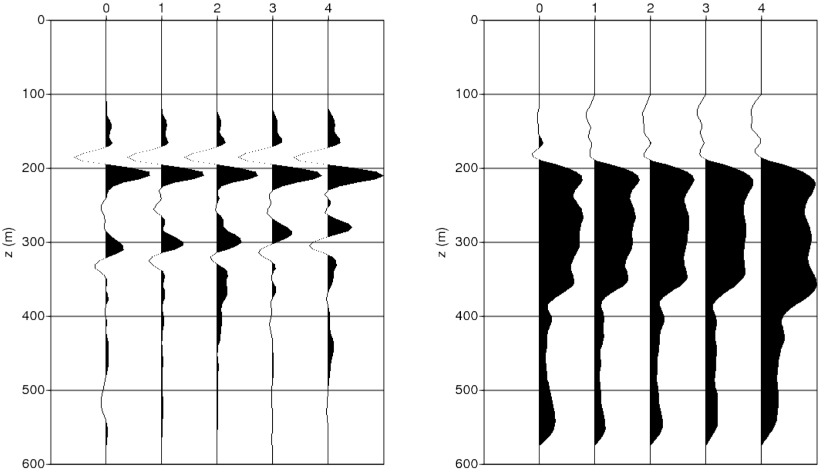


Numerical Experiments — layered test with absorbing surface

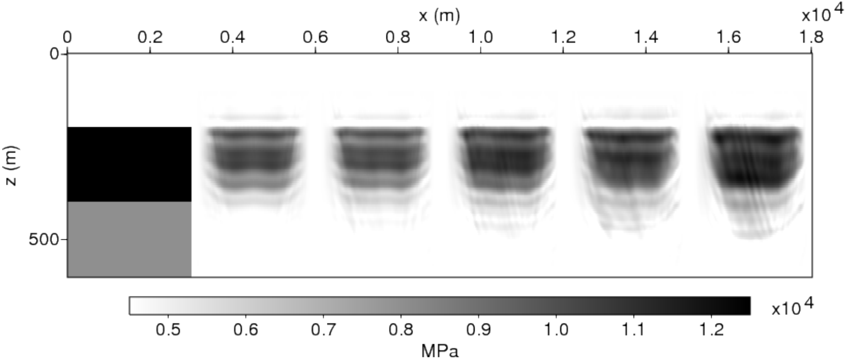
Initial RMS resid $= 3.4 \times 10^9;\ 4.2 \times 10^9;\ 5.3 \times 10^9;\ 7.0 \times 10^9;\ 9.4 \times 10^9;$
Final RMS resid $= 2.0 \times 10^8;\ 3.1 \times 10^8;\ 4.5 \times 10^8;\ 4.3 \times 10^8;\ 7.9 \times 10^8$

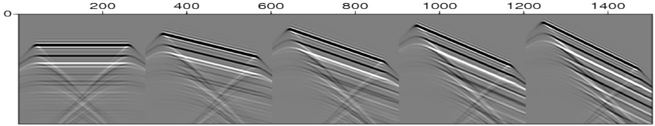


Inversion gathers



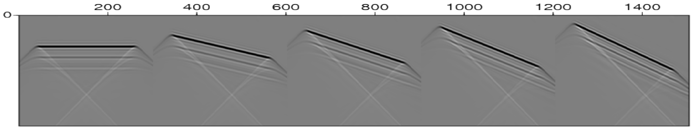
Inversion for data with correct low-frequency components

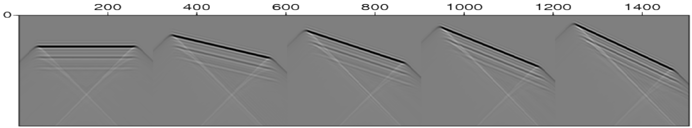


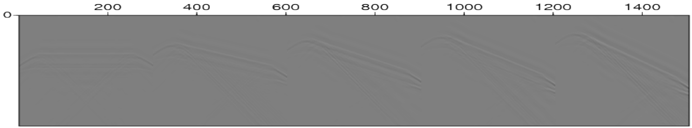


Numerical Experiments — layered tests with free surface

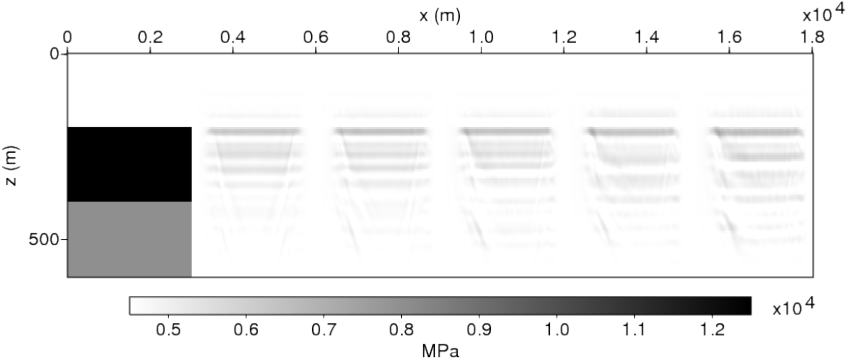
Initial RMS resid $= 1.1 \times 10^{11};\ 1.2 \times 10^{11};\ 1.5 \times 10^{11};\ 2.2 \times 10^{11};\ 3.8 \times 10^{11};$
Final RMS resid $= 7.9 \times 10^{9};\ \ 8.8 \times 10^{9};\ \ 9.8 \times 10^{9};\ 1.7 \times 10^{10};\ 2.8 \times 10^{10}$

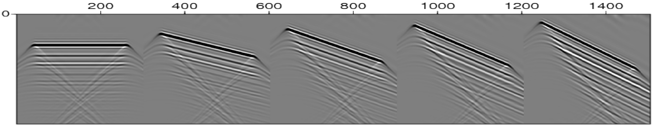






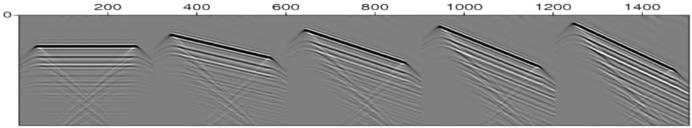
Inversion for data with wrong low-frequency components (driven from homogeneous model)

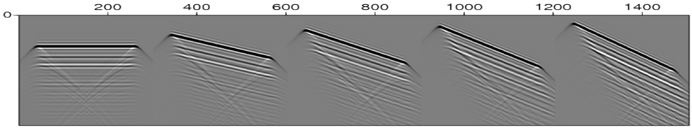


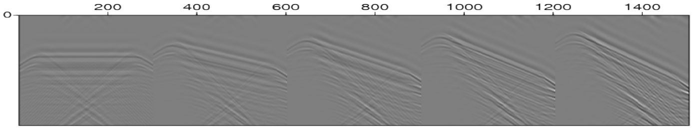


Numerical Experiments — layered tests with free surface

Initial RMS resid $= 6.6 \times 10^{10};\ 7.1 \times 10^{10};\ 7.8 \times 10^{10};\ 1.1 \times 10^{11};\ 2.1 \times 10^{11};$
Final RMS resid $= 5.8 \times 10^{9};\ \ 6.1 \times 10^{9};\ \ 9.6 \times 10^{9};\ 1.3 \times 10^{10};\ 2.4 \times 10^{10}$



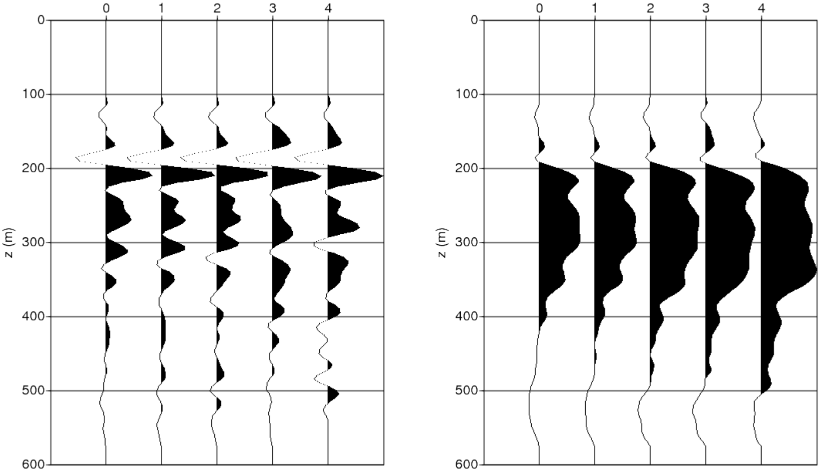




RICE

Numerical Experiments — layered tests with free surface

Inversion gathers



Agenda

1. Motivation of nDSO
2. Nonlinear Differential Semblance (nDS) Strategy
   - Reformulate WI as DSO
   - Implementation
3. Summary and Future Work

Summary and Future Work

nDS:

- imports the concepts from MVA into FWI (nonlinear MVA for FWI)
- fits into a general inversion framework
- addresses the spectral data incompleteness and local-minima
  - more feasible to gradient-related approaches
  - at least a good strategy to find initial model for FWI

Future Work:

- **add extended inversion functionality to IWAVE++ (variants of DS)**
- **explore techniques to improve the effectiveness of nDS**
- explore efficient solution to normal equation
- explore different optimization methods

RICE

Summary and Future Work

Future Work:

- **add extended inversion functionality to IWAVE++ (variants of DS)**

  Extended least-squares inversion:

$$\min_{\bar{m}\in\overline{\mathcal{M}}} \mathcal{E}[\bar{m}] = \frac{1}{2}\left\|\overline{\mathcal{F}}[\bar{m}] - \bar{d}_o\right\|^2 + \overline{\mathcal{R}}(\bar{m})$$

  Standard least-squares inversion:

$$\min_{m\in\mathcal{M}} J_{LS} := \frac{1}{2}\left\|\mathcal{F}[m] - d_o\right\|^2 + \mathcal{R}(m)$$

$\bar{m}(\mathbf{x}, s) := m_s(\mathbf{x})$ v.s. $m(\mathbf{x})$ $(\mathcal{M} \subset \overline{\mathcal{M}})$

Future Work:

- explore techniques to improve the effectiveness of nDS
  - add extra pre-ops to DS operator (tapering op, weighting ops to emphasize moveouts, ...)
  - understand effects of acquisition parameter range and sampling, ...
  - understand how iteratively solving LS and computing gradient affects DSO
  - improve LS inversion via regularization (DS-type), adaptive scaling, ...

$$\min_{\bar{m}\in\overline{\mathcal{M}}} \mathcal{E}[\bar{m}] = \frac{1}{2}\left\|\overline{\mathcal{F}}[\bar{m}] - \bar{d}_o\right\|^2 + \frac{1}{2}\alpha^2\left\|\frac{\partial}{\partial s}\bar{m}\right\|^2 + \overline{\mathcal{R}}(\bar{m})$$

RICE

Acknowledgements

Great thanks to

- my Ph.D. committee:
  William Symes, Matthias Heinkenschloss, Yin Zhang, Colin Zelt
- Present and former TRIP team members
- ExxonMobil URC FWI team
  especially Dave Hinkley, Jerry Krebs
- Sponsors of The Rice Inversion Project
- NSF DMS 0620821

Thank you!