Deterministic Source Synthesis for Waveform Inversion

William Symes

The Rice Inversion Project

Annual Review, 2010

Agenda

Source Synthesis - Why & How

Deterministic source synthesis

Numerical Exploration

Conclusions and Prospects

Source Synthesis - Why & How

Main motivation for this work: More efficient inversion - use fewer sources (ideally, one for entire data set) in each iterative inversion step

- length-1 encoding (weighted zero-lag data stacks - Krebs et al. 2009)
- inversion using source blending, simultaneous shooting (Ayeni et al. 2009, Verschuur & Berkhout 2009)
- random filtering, incoherency

Explicit recovery of individual shots not primary goal - synthetic sources chosen to drive model towards optimal inversion solution

= model which best fits *any* data (so shots are *implicitly* recovered...)

Source Synthesis - Why & How

This talk explores deterministic source synthesis via optimality principle:

best source $\Leftrightarrow$ *worst residual*

- origin in other inversion/imaging technologies
- simple source selection algorithm for acoustic modeling
- a few examples suggest pitfalls, remedies
- many unanswered questions - notably, does it really work? (in FWI)

RICE

Agenda

Deterministic source synthesis

Introduced into biomedical Electrical Impedance Tomography (EIT) by Isaacson (1986) - similar ideas: array ultrasonics (Fink & Prada 2004), ocean acoustics (Roux & Kuperman 2005), SAR (Borcea & Papanicolaou 2007),...

David Isaacson

EIT: image anomalies interior to body by measuring voltage response to applied current on boundary.

Deterministic source synthesis

Acoustic Model: state $u =$ acoustic potential in model domain R (subsurface), model $m = $ (velocity v, density ρ),

$$\frac{1}{\rho v^2}\frac{\partial^2 u}{\partial t^2} - \nabla \cdot \frac{1}{\rho}\nabla u = f(x, y, t)\delta(z - z_s).$$

Synthetic source $f =$ divergence of force density, confined to source depth plane z_s - must be post-synthesized digitally.

Measured response: pressure at fixed spread receiver locations $\Lambda^d f = \{\partial u/\partial t(\mathbf{x}_r, t)\}$ - linear in f - synthesized from field (point) source data traces.

Predicted response for model m $(= (v, \rho))$: $\Lambda[m]f$, computed by FE or FD or...

Deterministic source synthesis

Isaacson's Distinguishability Principle: seek normalized f so that RMS difference is largest: given estimated model m,

$$\text{maximize } (\Lambda^d f - \Lambda[m]f)^T(\Lambda^d f - \Lambda[m]f) \text{ subject to } f^T f = 1$$

max value $\lambda[m]$ = largest eigenvalue (operator norm) of *distinguishability operator*

$$A[m] = (\Lambda^d - \Lambda[m])^T(\Lambda^d - \Lambda[m])$$

= *largest discrepancy* in response for *any* (normalized) source (applied current pattern).

Deterministic source synthesis

Isaacson's algorithm:

- initialize m, f
- while (not satisfied),
  - fixed m, update f: perform several power method steps: $f \leftarrow A[m]f$, $f \leftarrow (1/\sqrt{f^t f})f$
  - fixed f, update m: perform several quasi-Newton steps with objective function $f^t A[m] f$ (standard output least squares)

Deterministic source synthesis

A few practical points:

1. assuming field data wavelet w known (!), achievable synthetic sources are filters:

$$f(x,y,t) = \sum_{x_s,y_s} \int d\tau \, g(x_s,y_s,t-\tau) w(\tau) \delta(x-x_s) \delta(y-y_s)$$

   Possibilities for g (1) arbitrary length filters (random choice - Romero et al. 00); (2) length-1 filters (amplitude factor) - Krebs et al. 09.

2. transpose operator $\Lambda[m]^T = R\Lambda[m]R$, $R =$ time-reversal op
3. Isaacson's alternating algorithm: Each step of both types involves 2 or 3 simulations (forward and/or reverse time loops), *for single (array) source*

Agenda

Setup

2D numerical experiments, using FD modeling/inversion package (IWAVE++).

Two models per experiment:

- Target model m^* - "measured" data $\Lambda^d f = \Lambda[m^*]f$
- Reference model m - "predicted" data $\Lambda[m]f$.

Measure progress in terms of *Rayleigh quotient* ("RQ"):

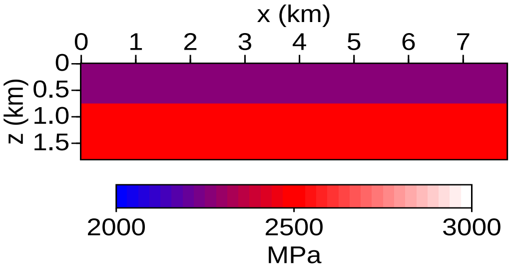
$$RQ = \frac{f^T A[m^*, m] f}{f^T f}$$

involves computing distinguishability operator $A[m^*, m] = (\Lambda[m^*] - \Lambda[m])^T (\Lambda[m^*] - \Lambda[m])$.

Setup

- staggered grid scheme for pressure, particle velocity
- source represented as *constitutive law defect* = RHS in pressure equation
- Models sampled at $\Delta x = \Delta z = 20$ m
- Absorbing BC on all sides of simulation domain
- Source, receiver depth 20 m - source = receiver locations
- 6 km fixed spread sampled at $\Delta x_s = 20$ m, 3 s recording interval
- 25 Hz high-cut imposed uniformly by filtering all sources, sources windowed to 0.0-0.4 s,

Layer over Half Space



Bulk modulus - 2.25 GPa to 0.75 km, 2.5 GPa below

Density is homogeneous = 1 gm/cc

Layer over Half Space

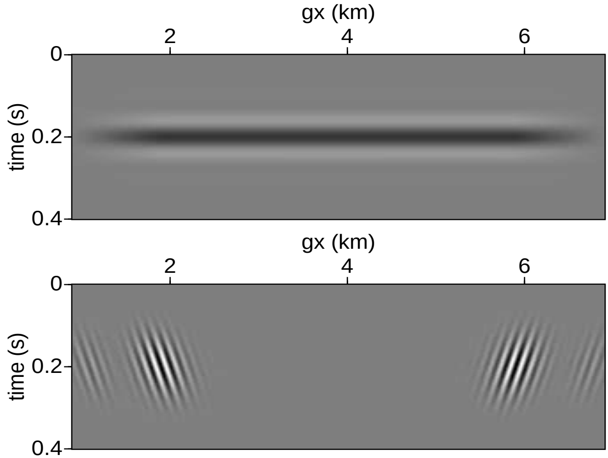
Initial source = truncated normal incidence plane wave

10 iterations of power method:

- initial Rayleigh quotient = 1.27
- final Raleigh quotient = 56.3

Looks great - however...

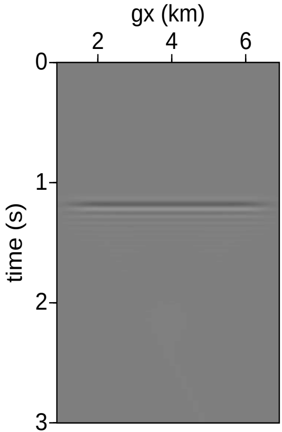
Layer over Half Space



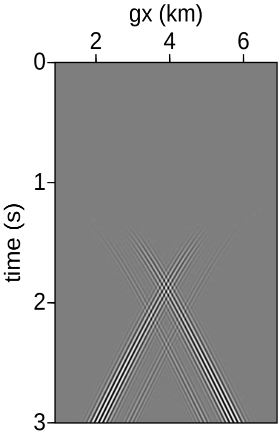
Initial (top), “optimal” (bottom) sources

Layer over Half Space

Data Difference $\Lambda[m^*]f - \Lambda[m]f$



it=0: RMS=1.1

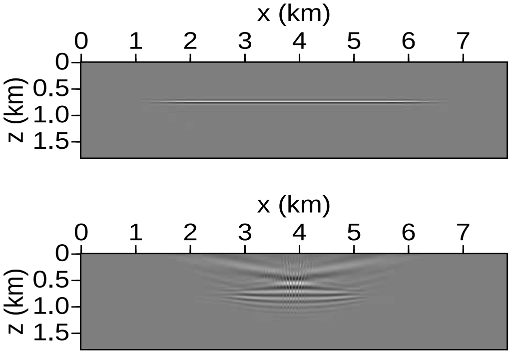


it=10: RMS=7.5

RICE

Layer over Half Space

RTM Image = Least Squares gradient



Amplitude of top (it=0) $10^{-2}\times$ amplitude of bottom (it=10).

Theory: Why this happens, what to do

Wave packed data:

$$f(x, y, t) = A(x, y, t)e^{i(k_x x + k_y y + \omega t)}$$

Guess: solution of wave equation

$$\frac{1}{\rho v^2}\frac{\partial^2 u}{\partial t^2} - \nabla \cdot \frac{1}{\rho}\nabla u = f(x, y, t)\delta(z - z_s)$$

takes form for $\pm z > 0$

$$u \simeq B_\pm e^{i(k_x x + k_y y \pm k_z z + \omega t)},$$

where $k_z = \pm\left(\frac{\omega^2}{v^2} - k_x^2 - k_y^2\right)^{\frac{1}{2}}$ and $B_\pm$ solves transport equation.

Theory: Why this happens, what to do

Causality: $\pm k_z > 0$ if $\pm z > 0$. Choose test function $\phi(x,y,z,t)$, then integration by parts gives

$$\int\int\int\int dxdydzdt\, p(x,y,z,t)\left(\frac{1}{\rho v^2}\frac{\partial^2\phi}{\partial t^2} - \nabla\cdot\frac{1}{\rho}\nabla\phi\right)$$

$$= \int\int\int dxdydt\, f(x,y,t)\phi(x,y,t)$$

Since both p and $\partial p/\partial z$ are continuous (normal stress, displacement), can split first integral into $z < 0$ and $z > 0$ pieces and integrate by parts again. Because of wave equation for p, only boundary terms left:

$$= \int\int\int dxdydt\left\{[p]\frac{\partial\phi}{\partial z} - \left[\frac{\partial p}{\partial z}\right]\phi\right\}$$

RICE

Theory: Why this happens, what to do

This identity must hold for any test function (smooth, vanishing for large $\mathbf{x}, t$) - in particular, can choose ϕ to be $= 0$ on $z = 0$ whilst $\partial\phi/\partial z$ takes on arbitrary values. Hence [p]=0. Since ϕ can also take arbitrary values, follows that

$$f(x, y, t) = -\left[\frac{\partial p}{\partial z}\right](x, y, t)$$

First condition implies that $B_- = B_+$ on $z = 0$; second, that

$$f(x, y, t) = -2ik_z B_\pm e^{i(k_x x + k_y y + \omega t)},$$

Thus

$$u \simeq \frac{\tilde{A}}{k_z} e^{i(k_x x + k_y y + k_z z + \omega t)},$$

where $\tilde{A}|_{z=0} = \frac{i}{2}A$, and $\tilde{A}$ solves transport eqns.

Theory: Why this happens, what to do

Upshot: k_z small $\Rightarrow$ energy transfer to acoustic field *extremely efficient* per RMS unit f.

k_z small $\Rightarrow$ most energy propagates near-horizontally - limits imaging aperture, vertical resolution.

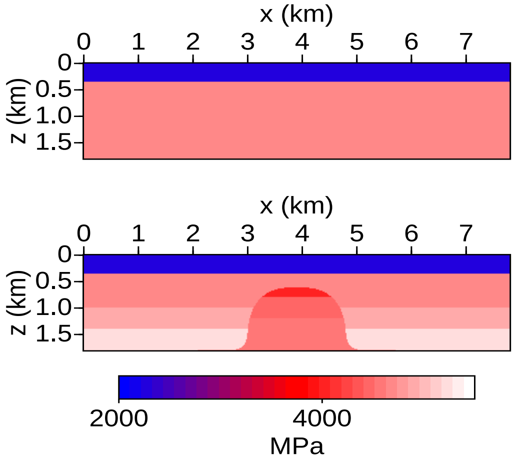
Solution: depress part of spectrum of $A[m]$ corresponding to small k_z by composing $\Lambda^d - \Lambda[m]$ with dip filter.

For water layer near surface: k_z small when $|k_x| \simeq 0.67$ s/km.

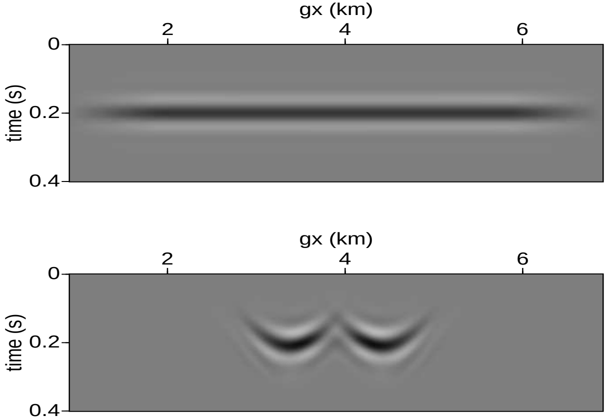
Example: for LOHS example, choose dip filter with corner slope of 0.3 s/km, cut slope of 0.5 s/km - then optimal source is small modification of plane wave source.

Laterally Heterogeneous Example

Bulk moduli - reference (top), target (bottom)



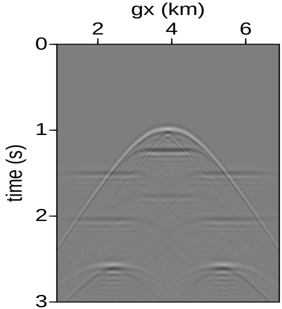
Laterally Heterogeneous Example



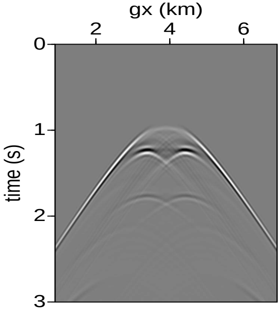
Initial (top), “optimal” (bottom) sources

Laterally Heterogeneous Example

Data Difference $\Lambda[m^*]f - \Lambda[m]f$



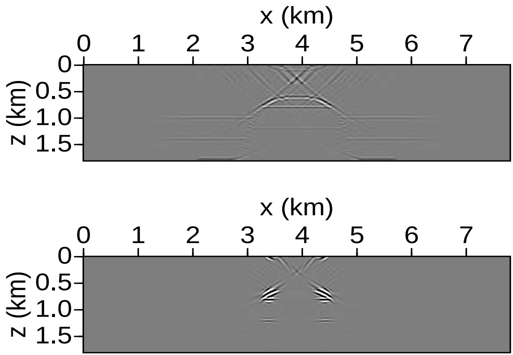
it=0: RMS=2.0, RQ=5.8



it=10: RMS=3.3, RQ=11.2

Laterally Heterogeneous Example

RTM Image = Least Squares gradient



Top (it=0) and bottom (it=10, 10× clip)

Agenda

Source Synthesis - Why & How

Deterministic source synthesis

Numerical Exploration

Conclusions and Prospects

Conclusions and Prospects

Many numerical experiments suggest

- "optimal" source emphasizes largest features in residual data, as intended
- dip filtering effectively controls tendency to produce horizontally traveling energy
- selective illuminates features in gradient (RTM residual image)

Conclusions and Prospects

If anything, illumination is *too* selective - a single source is likely not sufficient

Gao et al. 2010: find all eigenpairs of distinguishability operator above a threshhold, use these collectivley - still much smaller than number of source positions in typical survey (?)

Natural method: Lanczos algorithm - finds segment of spectrum, rather than merely largest eigenvalue.

Next step: use Lanczos implementation in RVL to explore time-domain version of Gao et al. proposal.

Thanks to...

Dong Sun, Igor Terentyev, Tanya Vdovina, Rami Nammour, Marco Enriquez, Xin Wang

Joakim Blanch and John E. Anderson for dome model

SEAM Project - development of IWAVE

Fuchun Gao and Paul Williamson for sharing their work

Sponsors of The Rice Inversion Project

RICE