Extended Inversion: when it works, and why

William W. Symes

symes@rice.edu

December 2020

Overview

Apply sophisticated methods to a simple inversion problem

Why:

- clear understanding of what works and what doesn't, and why
- all details "on paper"
- conclusions not justified only "in rear view mirror" by a few numerical examples
- road map for application to industrially relevant technology

"You can do this at home!"

Thanks to Prof. Sue Minkoff and GRA Huiyi Chen, UT Dallas

World's simplest waveform inversion problem

Single trace transmission in a fluid:

Point source at $\mathbf{x} = \mathbf{x}_s$ radiates with transient intensity $w(t)$ ("wavelet")

Resulting pressure wave propagates to point receiver at $\mathbf{x} = \mathbf{x}_r$, offset $|\mathbf{x}_r - \mathbf{x}_s| = h$

Pressure trace predicted by linear acoustics:

$$p(t) = \frac{1}{4\pi h} w(t - mh) = (F[m]w)(t)$$

where m = pressure wave slowness $(= 1/v_p)$, F = modeling operator - linear in w, nonlinear in m

World's simplest waveform inversion problem

Inverse problem: given pressure trace $d(t)$, wavelet $w(t)$, and offset h, find slowness m so that

$$F[m]w \approx d$$

Can solve by inspection! Use a ruler to measure time shift between $w(t)$ and $w(t - mh)$

BUT let's use the standard approach instead: turn it into an optimization - minimize over m the *mean square error*

$$\frac{1}{2}\int dt\,(F[m]w(t) - d(t))^2 = \frac{1}{2}\|F[m]w - d\|^2 = J_{\mathrm{FWI}}[m; w, d]$$

$=$ world's simplest FWI task

How to cycle-skip, in one easy lesson

Asymptotic transient-ness:

Mother wavelet w_1: $w_1(t) = 0, |t| > 1$ and $\int w_1^2 = 1$

Scaled wavelets $w_\lambda = \frac{1}{\sqrt{\lambda}} w_1(t/\lambda)$: $w_\lambda(t) = 0, |t| > \lambda$ and $\int w_\lambda^2 = 1$

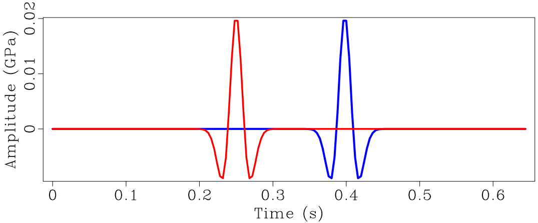
NB: if center frequency of $w_1 = \omega_1$, then center frequency of $w_\lambda = \omega_1/\lambda$

How to cycle-skip, in one easy lesson

Theorem Uh-Oh: If $d = F[m_*]w_\lambda$ (noise-free transient data) and $|m - m_*| > 2\lambda/h$, then

$$J_{\rm FWI}[m; w_\lambda, d] = \|d\|^2, \;\; \nabla_m J_{\rm FWI}[m; w_\lambda, d] = 0$$

Proof.

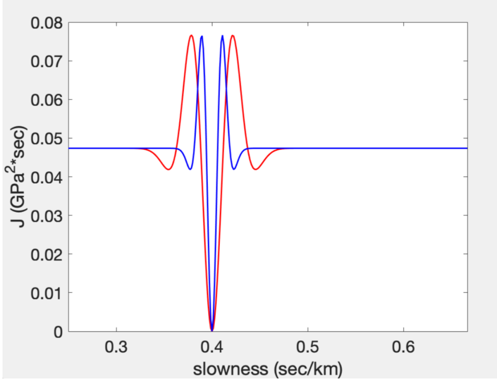


Data: same wavelet, $m = 0.25$ s/km (red) and $m_* = 0.4$ s/km (blue)

$F[m]w_\lambda$ orthogonal to $F[m_*]w_\lambda$ if $|m - m_*| > 2\lambda/h$ Q. E. D.

How to cycle-skip, in one easy lesson

$J_{\rm FWI}$: $d = F[m_*]w_\lambda$, $m_* = 0.4$ s/km, w_λ = Ricker ctr freq $= 1/\lambda$ Hz, $h = 1$ km



Red: $\lambda = 0.05$ s (20 Hz), Blue: $\lambda = 0.025$ s (40 Hz) [H. Chen et al. SEG 20]

How to cycle-skip, in one easy lesson

Postmortem:

- for most m, predicted data $F[m]w$ is not close to target data d
- for most m, gradient is not a useful search direction
- shape of objective function depends strongly on data frequency content - large residual, useless gradient more common for high frequencies (small λ)
- flip side: low frequency (big λ) $\Rightarrow$ useful updates for larger set of initial models

Wavefield Reconstruction Inversion

Extended inversion: *enlarge the model space* = add (nonphysical) degrees of freedom, converge to point in original model space

Rationale: more likely to be able to

- find low residual (extended) models ("hug your data"),
- explore useful update directions

WRI - van Leeuwen & Herrmann GJI 13, IP 16, many other papers - most studied extended inversion approach

Idea: view wave equation as *weak constraint* $\sim$ *add source parameters*

Wavefield Reconstruction Inversion

Cast:

- model parameters m (eg. m=slowness)
- Wave operator $L[m]$ (eg. constant density acoustic op $L[m] = m^2\partial_t^2 - \nabla^2$)
- Known (!) source field $q(\mathbf{x}, t)$ (eg. point source $w(t)\delta(\mathbf{x} - \mathbf{x}_s)$)
- dynamic fields $u(\mathbf{x}, t)$ solve $L[m]u = q$ (eg. $u =$ pressure)
- Sampling operator P extracts data trace(s) (pressure or ...) from u
- Modeling operator $F[m] = PL[m]^{-1}$

Wavefield Reconstruction Inversion

FWI: given d, q,

$$\min_m \|Pu - d\|^2 \text{ subj to } L[m]u = q$$

$$\Leftrightarrow \min_m \|F[m]q - d\|^2$$

WRI: given d, q,

$$\min_{m,u} \|Pu - d\|^2 + \alpha^2 \|L[m]u - q\|^2$$

(NB: α small $\Rightarrow$ emphasis on 1st term $\Rightarrow$ "hug your data")

Wavefield Reconstruction Inversion

equivalent (eg. Yingst & Wang SEG 16):

given d, q, set $g = L[m]u - q$ ("residual source"),

then $Pu - d = PL[m]^{-1}(g + q) - d = F[m]g + e[m]$

$e[m] = F[m]q - d =$ usual data residual

WRI: given q, d,

$$\min_{m,g} \|F[m]g + e[m]\|^2 + \alpha^2\|g\|^2$$

Wavefield Reconstruction Inversion

Simplification 1:

Variable Projection (Golub & Pereyra 73, 03): eliminate g by solving quadratic minimization

$$J_{\mathrm{WRI}}[m; q, d] = \min_{g} \|F[m]g + e[m]\|^2 + \alpha^2 \|g\|^2$$

Since $\min_{m,g}(\ldots) = \min_m(\min_g(\ldots))$, WRI problem $\equiv$ minimize J_{WRI}

Wavefield Reconstruction Inversion

Simplification 2:

Non-radiating sources $= \{g : F[m]g = 0\} =$ null space of $F[m]$

Rank-Nullity Theorem: $g = n + F[m]^T s$, $s \in$ range of $F =$ data, $n =$ non-radiating, *and* n and $F[m]^T s$ are orthogonal

$$\|F[m]g + e[m]\|^2 + \alpha^2\|g\|^2 = \|F[m](F[m]^T s + n) + e[m]\|^2 + \alpha^2\|F[m]^T s + n\|^2$$

$$= \|F[m]F[m]^T s + e[m]\|^2 + \alpha^2\|F[m]^T s\|^2 + \alpha^2\|n\|^2$$

so

$$J_{\mathrm{WRI}}[m; q, d] = \min_s \|F[m]F[m]^T s - e[m]\|^2 + \alpha^2\|F[m]^T s\|^2$$

NB: g is space-time field, s is data - vast memory reduction, practical 3D time domain (Yingst & Wang SEG 16, Rizzuti et al. SEG 19)

Wavefield Reconstruction Inversion

Simplification 3:

Minimization over s $\Leftrightarrow$ solution of normal equation

$$((F[m]F[m]^T)^2 + \alpha^2 F[m]F[m]^T)s = F[m]F[m]^T e[m]$$

Substitute solution into def of J_{WRI}, do a page of algebra (S., IP 20), then...

$$J_{\mathrm{WRI}} = \alpha^2 e[m]^T (F[m]F[m]^T + \alpha^2 I)^{-1} e[m]$$

compare:

$$J_{\mathrm{FWI}} = e[m]^T e[m]$$

J_{WRI} is weighted variant of J_{FWI}, weight operator $= \alpha^2 (F[m]F[m]^T + \alpha^2 I)^{-1}$

Wavefield Reconstruction Inversion

Apply to single-trace transmission:

$$F[m]g(t) = \int_{r \le R} dx \, \frac{g(t - mr)}{4\pi r}, \; F[m]^T s(\mathbf{x}, t) = \begin{cases} \frac{s(t+mr)}{4\pi r}, r \le R, \\ 0, \mathit{else} \end{cases}$$

($r = |\mathbf{x} - \mathbf{x}_r|$ and $g = 0$ if $r > R$), so

$$F[m]F[m]^T s(t) = \frac{R}{4\pi} s(t), \; \alpha^2 (F[m]F[m]^2 + \alpha^2 I)^{-1} = \frac{4\pi\alpha^2}{R + 4\pi\alpha^2} I$$

Wavefield Reconstruction Inversion

Theorem Uh-Oh # 2:

$$J_{\mathrm{WRI}}[m; w(t)\delta(\mathbf{x} - \mathbf{x}_s), d] = \frac{4\pi\alpha^2}{R + 4\pi\alpha^2} J_{\mathrm{FWI}}[m; w, d]$$

Postmortem:

- WRI just as skippy as FWI!!!
- enlarging the search space not enough
- "hugging your data" also not enough: WRI cycle-skips for small α
- conclusion not limited to single trace transmission: similar for other transmission IPs, eg. diving wave inversion (Fang & Demanet, SEG 20)

Matched Source Inversion

Back to FWI setting:

$$F[m]w(t) = \frac{1}{4\pi h} w(t - mh)$$

A different extended inversion: add wavelet to unknowns

However, $F[m]$ is invertible for every m, so need constraint on w

Idea: after signature decon, wavelet should be *compact* $\sim$ nonzero only near $t = 0$ - inspires *Matched Source Objective*

$$\min_{m,w} \|F[m]w - d\|^2 + \alpha^2 \|Aw\|^2, \; Aw(t) = tw(t)$$

(S. 94, Plessix et al. 99, similar to Adaptive Waveform Inversion (AWI) Warner & Guatsch 14)

Matched Source Inversion

Variable Projection:

$$J_{\mathrm{MSI}}[m;d] = \min_{w} \|F[m]w - d\|^2 + \alpha^2 \|Aw\|^2$$

Minimizer $w[m;d]$ solves normal equation

$$(F[m]^T F[m] + \alpha^2 A^T A)w = F[m]^T d$$

Explicit calculation:

$$F[m]^T F[m] + \alpha^2 A^T A = \left(\frac{1}{(4\pi h)^2} + \alpha^2 t^2 \right) I, \; F[m]^T d(t) = \frac{1}{4\pi h} d(t + mh)$$

Matched Source Inversion

Plug it in, get explicit formulas for everything - in particular, if data is noise-free, $d = F[m_*]w_\lambda$, then

$$\nabla_m J_{\mathrm{MSI}}[m; d] = -h\alpha^2 \int dt \, \frac{t - (m - m_*)h}{(1 + (4\pi h\alpha(t - (m - m_*)h))^2)^2} w_\lambda(t)^2$$

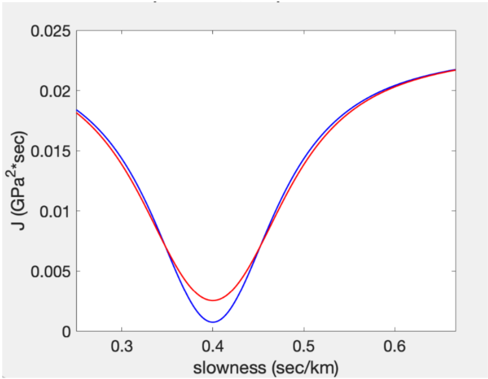
Theorem Ah-Ha! If $d = F[m_*]w_\lambda$,

$$\nabla_m J_{\mathrm{MSI}}[m; d] \begin{cases} < 0 \text{ if } m < m_* - \lambda/h \\ > 0 \text{ if } m > m_* + \lambda/h \end{cases}$$

$\Rightarrow$ if m is stationary point of $J_{\mathrm{MSI}}[\cdot; d]$, then $|m - m_*| \leq \lambda$.

Matched Source Inversion

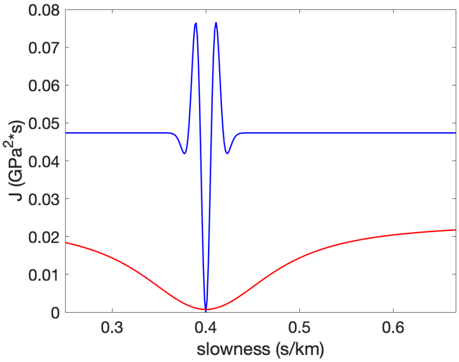
J_{MSI}: $d = F[m_*]w_\lambda$, $m_* = 0.4$ s/km, w_λ = Ricker ctr freq $= 1/\lambda$ Hz, $h = 1$ km, $\alpha = 1 s^{-1} km^{-1}$



Red: $\lambda = 0.05$ s (20 Hz), Blue: $\lambda = 0.025$ s (40 Hz) [H. Chen et al. SEG 20]

Matched Source Inversion

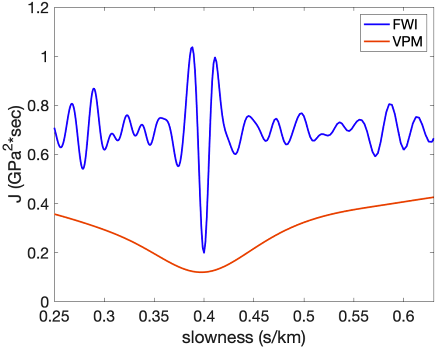
J_{MSI} *vs.* J_{FWI}: $d = F[m_*]w_\lambda$, $m_* = 0.4$ s/km, $w_\lambda =$ Ricker ctr freq $= 40$ Hz ($\lambda = 0.025$ s), $h = 1$ km, $\alpha = 1 s^{-1} km^{-1}$



Red: MSI, Blue: FWI [H. Chen et al. SEG 20]

Matched Source Inversion

$J_{\rm MSI}$ *vs.* $J_{\rm FWI}$: $d = F[m_*]w_\lambda + n$, $n =$ 100% RMS filtered stationary random noise, otherwise same



Red: MSI, Blue: FWI [H. Chen]

Matched Source Inversion

Verdict: MSI

- appears to avoid cycle-skipping
- “hugs data”, expands search space - but effectively
- Lots of multi-D relatives show good cycle-skip busting examples: matched source extension (Huang & S. 16), volume and space-time source extensions (Huang et al GEO 18), surface source extension (Huang et al. SEG 19), receiver shift extension (Métivier SEG 20), AWI (Warner & Guatsch GEO 16 - field scale 3D, commercial)

So why?

Common penalty form for extended source inversion methods:

$$\min_{m,w} \|F[m]w - d\|^2 + \alpha^2 \|Aw\|^2$$

- m = vector of parameter *fields* incl. velocities
- $F[m]$ = modeling operator
- d = multichannel data
- w = extended source, A = annihilator: physical source satisfy $Aw \approx 0$

Example: (AWI/MSI) point radiator $w(t)\delta(\mathbf{x} - \mathbf{x}_s)$: physical source = impulsive, focused at $t = 0$; extended source: spread in t; $Aw(t) = tw(t)$.

So why?

Necessary condition: reduced objective (VPM) must be regular (differentiable) in m, d *jointly*

For penalty functions as above: A *must* be (pseudo)differential (Stolk & S 03)

Re-write of FWI, WRI reduced objectives: analogous op *not* pseudodifferential

- FWI, WRI *fail*, MSI *passes* (S. IP 20)

So why?

Sufficient condition: gradient related to traveltime tomography gradient

- Structure of wave-based modeling operators:

$$D_m(F[m]w)\delta m = Q[m, \delta m]F[m]w$$

$Q =$ order 1 approx. skew-symm. (pseudo)differential op

Example: (AWI/MSI) $F[m]w(\mathbf{x}_r, t) \approx a(\mathbf{x}_s, \mathbf{x}_r)w(t - \tau[m](\mathbf{x}_s, \mathbf{x}_r))$, so

$$D_m(F[m]w)\delta m \approx -D\tau(\mathbf{x}_s, \mathbf{x}_r)\delta m \frac{\partial}{\partial t} F[m]w$$

(chain rule!)

So why?

(...bunch of algebra...) $\Rightarrow$

$$\nabla J_{\mathrm{MSI}} \propto \nabla_m \|\tau[m] - \tau_{\mathrm{data}}\|^2$$

so MSI $\approx$ travel time inversion

And that's why.

(Chen et al. SEG 20, Huang & S. SEG 16)

Thanks to...

- Prof. Susan. E. Minkoff, GRA Huiyi Chen, UT Dallas
- Guanghui Huang (PGS), Rami Nammour, Mohamed Dolliazal, Paul Williamson (Total)
- many other students, postdocs, and colleagues
- former sponsors of The Rice Inversion Project (1992-2019)
- you!

Key References

G. Huang et al., "Waveform Inversion by Source Extension": SEG 2019

H. Chen et al., "Full waveform inversion by source extension: why it works", SEG 2020

W. Symes, "Wavefield reconstruction inversion: an example", *Inverse Problems* v. 36(10), 2020