Extended Inversion: when it works, and why

William W. Symes

Rice University

December 2020

How to Cycle-Skip, in one easy lesson

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

How to Cycle-Skip, in one easy lesson

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

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

Standard approach: turn it into an optimization - minimize over m the *mean square error*

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

$=$ 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 $\int w_\lambda^2 = 1$

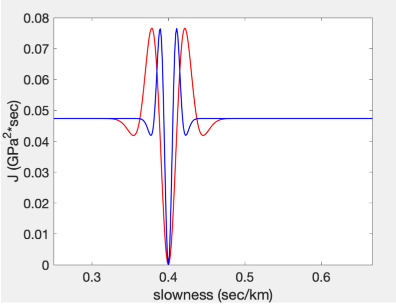
NB: if center frequency of $w_1 = \omega_1$, then center frequency of $w_\lambda = \omega_1/\lambda$

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

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

How to Cycle-Skip, in one easy lesson

$J_{\rm FWI}[m; w_\lambda, d]$, $d = F[m_*]w_\lambda$, w_λ = Ricker wavelet center frequency $= 1/\lambda$, $m_* = 0.4$ s/km, $h = 1km$



Red: $\lambda = 0.05$ (20 Hz), Blue: $\lambda = 0.25$ (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

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*

Rationale: more likely to be able to

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

Wavefield Reconstruction Inversion

Cast:

- model parameters m (eg. m=slowness)
- Wave operator $L[m]$ (eg. constant density acoustic op

$$L[m] = m^2 \frac{\partial^2}{\partial t^2} - \nabla^2 \Big)$$

- 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 $g = L[m]u - q$,

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

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

minimizer $= g[m; q, d]$, then

$$= \|F[m]g[m; q, d] + e[m]\|^2 + \alpha^2 \|g[m; q, d]\|^2$$

m minimizes $J_{\mathrm{WRI}} \Leftrightarrow (m, g[m; q, d])$ solves WRI problem

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

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

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 (IP, v. 36 no. 10 (2020)), 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:

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

$(r = |\mathbf{x} - \mathbf{x}_r|$ 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

Matched Source Extension

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*

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

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