Use of Extended Source Inversion for Estimating the Noise Level in Seismic Data

Susan E. Minkoff[1], Huiyi Chen[1], and Bill Symes [2]

[1]Department of Mathematical Sciences, The University of Texas at Dallas

[2]Department of Computational and Applied Mathematics, Rice University

New Ideas in Computational Inverse Problems
Banff International Research Station (BIRS)
October 28, 2022

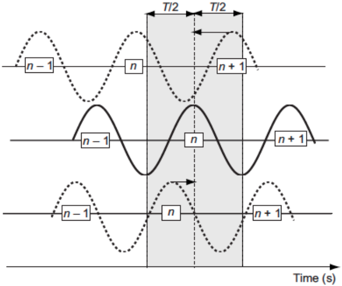
Acknowledgements: This research is partially supported by the sponsors of the UT Dallas "3D+4D Seismic FWI" research consortium.

Outline

- The problem with Full Waveform Inversion (FWI) (or least-squares): cycle-skipping!
- Simple problem setup and motivation
- Introduction of source extended objective function (ESI) and why it helps.
- Noise estimation algorithm
- Numerical examples:
  1. Noise level fixed
  2. Noise level updated: coherent noise
  3. Noise level updated: random noise
  4. Noise level updated: no noise
- Conclusions

Motivation Behind Extended Objective Functions

- Full Waveform Inversion (FWI) is now well-established as a useful tool for estimating parameters in the earth.
- Unfortunately, the FWI objective function is not convex. FWI stagnates at geologically uninformative earth models (local minima).



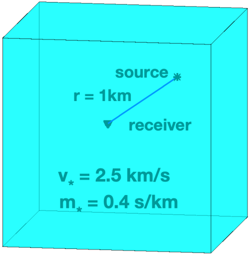
Schematic of cycle-skipping artifacts in FWI. Solid black line is seismogram of period T. Upper dashed line is seismogram with a time delay greater than $T/2$. Bottom example, has time delay less than $T/2$.[1]

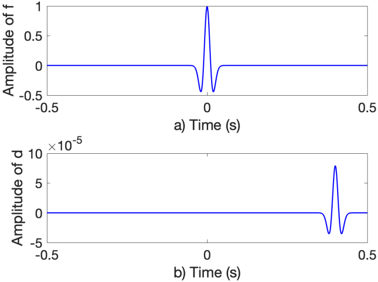
[1] Virieux, J., and S. Operto, 2009, "An overview of full-waveform inversion in exploration geophysics", *Geophysics*, **74**, WCC1–WCC26.

Motivation Behind Extended Objective Functions

- Extended inversion is one of the many ideas that have been advanced to overcome cycle-skipping. We will focus on "source extension".
- "Extended" signifies that additional degrees of freedom are provided to the modeling process.
- These extended degrees of freedom should be suppressed in the eventual solution since they are not physical.
- In the case of a very simple model problem, all computations can be done analytically. Results can be theoretically justified.
- Simple problem illustrates the same cycle skipping issues one encounters in FWI for more realistic problems.

Simple Experimental Setup





Left: single-trace experimental setup. Right top: the source wavelet (a 20 Hz Ricker). Right bottom: data.

Mathematical Background

- Assume small amplitude constant-density, acoustic wave propagation in $3D$:
$$m^2 \frac{\partial^2 P(x,t;x_s)}{\partial t^2} - \nabla^2 P(x,t;x_s) = \delta(x - x_s) w(t)$$
where w is wavelet, m is slowness.

- The pressure trace recorded at the receiver position is given by:
$$F[m]w(t) = \frac{1}{4\pi r} w(t - mr)$$
where $F[m]$ is the operator of convolution with acoustic $3D$ Green's function and r is the distance between the source and receiver.

- Inverse Problem : Given target noise level e_{tar}, maximum lag [2] $\lambda > 0$, find the slowness m and wavelet w so that:
  - $w(t) = 0$ if $|t| > \lambda$
  - $\frac{||F[m]w - d||}{||d||} < e_{tar}$

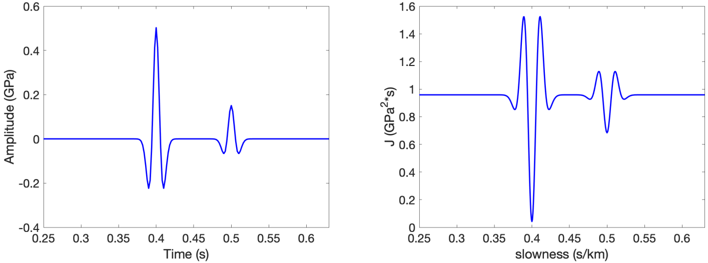
[2] Spiking deconvolution and other processing implies $w(t)$ is negligible for $|t| > \lambda$.

The challenge with full waveform inversion: cycle skipping

The FWI objective function is defined as:

$$J_{FWI}[m, w] = \frac{1}{2}||F[m]w - d||^2,$$

We solve for slowness m and wavelet w, for maximum lag λ.



Left: Data produced by adding to the noise-free data a shifted, scaled multiple of itself.
Right: the FWI objective function plotted as function of slowness, for fixed w.

- There are entire intervals of local minimizers far from the global minimizer m_*.
- Initial guess for slowness m must be within $2\lambda/r$ of the global minimizer m_*, or we fail to solve the inverse problem.[3] **"cycle-skipping"!!**

[3] Symes, W. W., 2022, "Error bounds for extended source inversion applied to an acoustic transmission inverse problem", *Inverse Problems*, **38**, 115002.

Extended Source Inversion

- Add degrees of freedom to F to avoid local minima.
- By including the source wavelet as one of the modeling parameters and dropping the support constraint on w, we extend space of possible solutions.

Define the extended source inversion for the minimization over $\{m, w\}$ as

$$J^{\alpha}_{ESI}[m, w; d] = \frac{1}{2}\left(\underbrace{\frac{||F[m]w - d||^2}{||d||^2}}_{= e^2} + \frac{\alpha^2||Aw||^2}{||d||^2}\right),$$

A is an annihilator to penalize energy away from $t = 0$:

$$Aw(t) = tw(t).$$

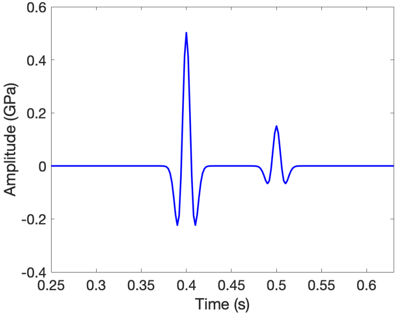
- The value of J^{α}_{ESI} will be small at the minimizer of J_{FWI}.

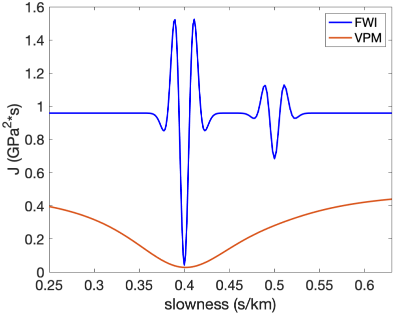
The Variable Projection Method (VPM) produces a reduced objective function of m alone:

$$J^{\alpha}_{VPM}[m; d] = \inf_{w} J^{\alpha}_{ESI}[m, w; d]$$

Extended Inversion vs FWI: Data with Coherent Noise

- ESI objective hard to minimize for both *m* and *w* simultaneously.
- Use Variable Projection Method[4] with inner minimization over *w* then an outer minimization over *m*.
- In this case, wavelet solution given analytically by the normal equations.



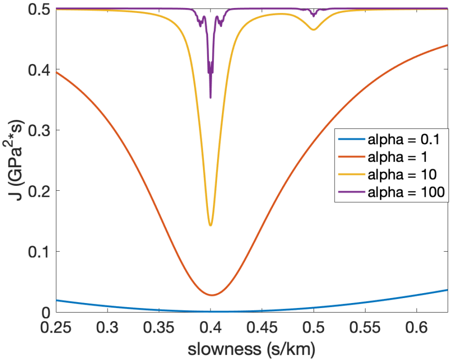


Left: Data with coherent noise. **Right:** The VPM (red curve, with $\alpha = 1$) and FWI (blue curve) objective functions plotted as functions of slowness.

[4] Golub, G., and V. Pereyra, 2003, "Separable nonlinear least squares: the variable projection method and its applications", *Inverse Problems*, **19**, R1-R26.

VPM objective functions with different values of α

- α has a big impact on the rate of convergence of the algorithm[5].
- If α can increase dynamically during run we see improved performance of the algorithm.



VPM objective functions plotted with different values of α for the data with coherent noise.

- As α goes to 0, the VPM objective function goes to 0, and the error also goes to 0.
- As α increases, the region of convexity gets smaller, and the resolution of the slowness *m* gets better.
- The error increases as α increases.

[5] W. W. Symes, H. Chen, and S. E. Minkoff, 2022, "Solution of an Acoustic Transmission Inverse Problem by Extended Inversion", *Inverse Problems*, 38, 115002.

Motivation behind the noise estimation algorithm

- We define "noise" as the unmodeled signal, which is the minimum of the FWI objective function over physical models (that is, the wavelets w with maximum lag constraint).
- We have to select α to use the ESI/VPM approach to estimate slowness.
- The error increases monotonically with α for the VPM solution, which is equivalent to selecting target noise level e_{tar} and updating error (by updating α) until it's near e_{tar}.
- If the difference between the estimated noise level e_{est} and the target noise level e_{tar} is too big, then we can replace e_{tar} by e_{est}, repeat the calculation, and obtain a better α hence a better estimate of m, which in turn gives a better e_{est}. If e_{est} is about the same as e_{tar}, the process stops.[6]

[6] H. Chen, W. W. Symes, and S. E. Minkoff, "Use of extended source inversion for estimating the noise level in seismic data," *Proceedings of the Second International Meeting for Applied Geoscience & Energy,* Houston, TX., pp 887-891, 2022.

Noise estimation algorithm

The noise level as estimated in the algorithm is computed by the following formula:

$$e_{est} = \min_{w} \frac{||F[m]w - d||}{||d||}.$$

Note that there is maximum lag constraint on w.

- Start with arbitrary m, $\alpha = 0$, $\gamma < 1$, $\delta > 0$, gradient tolerance β. Initialize the assumed target noise level $e_{tar} > 0$ and a range of minimum and maximum allowable errors $0 < (1-\gamma)e_{tar} < (1+\gamma)e_{tar}$.
- Do:

  Start discrepancy algorithm:[7]

  - first fix m, update α so that $(1-\gamma)e_{tar} \leq e \leq (1+\gamma)e_{tar}$
  - then fix α, update m so that $|\nabla J^{\alpha}_{VPM}| < \beta$ (use local descent method)
  - repeat until $(1-\gamma)e_{tar} \leq e \leq (1+\gamma)e_{tar}$ AND $|\nabla J_{VPM}| < \beta$

  End discrepancy algorithm
- compute minimum value e_{est} of e over physical source vector w, with the updated medium parameter m;
- if $|e_{\mathrm{tar}} - e_{\mathrm{est}}| > \delta e_{\mathrm{est}}$ then $e_{\mathrm{tar}} = e_{\mathrm{est}}$, repeat, else terminate.

[7] L. Fu and W. W. Symes, 2017, "A discrepancy-based penalty method for extended waveform inversion", *Geophysics*, **82**, no. 5, R287-R298.

Experiment 1: data with 30% coherent noise, no noise level update

- noise-to-signal ratio of 30%
- initial $m = 0.343$. Initial $\alpha = 0$.

initial noise level: e_{tar}	**estimated noise level:** e_{est}	**penalty weight:** α	**estimated slowness:** m (s/km)
0.1	0.285395	0.626626	0.403532
0.2	0.287649	1.253252	0.400797
0.3	0.287910	1.409909	0.400575
0.4	0.959087	5.013009	0.498857
0.5	0.961294	7.832827	0.499765
0.6	0.963973	11.279270	0.499937

Table: Noise estimation algorithm with initial target noise levels ranging from 0.1 to 0.6.

- Best $e_{est} = 0.287253$, which is computed by

$$e_{best} = \min_{w} \frac{||F[m_*]w - d||}{||d||},$$

  where $m_* = 0.4$ s/km is the true slowness.
- Note: local minimum near m = 0.5 s/km which we find when target noise level is too large.

Experiment 2: noise estimation for data with coherent noise

- noise-to-signal ratio of 30%
- initial $m = 0.343$. Initial $\alpha = 0$.

initial noise level: e_{tar}	estimated noise level: e_{est}	penalty weight: α	estimated slowness: m (s/km)
0.1	0.287253	3.878713	0.400021
0.2	0.287253	4.167532	0.400017
0.3	0.287253	2.136271	0.400156
0.4	0.287253	4.118976	0.400017
0.5	0.287253	4.118293	0.400017
0.6	0.287253	4.118268	0.400017

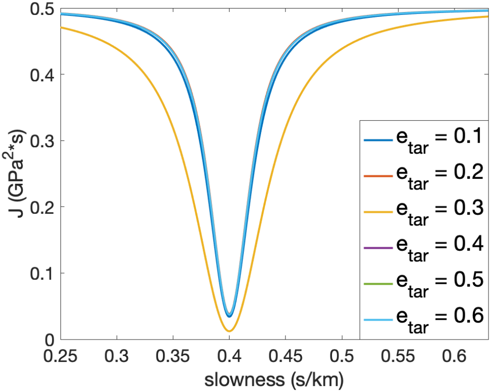
Table: Noise estimation algorithm starting with initial target noise levels ranging from 0.1 to 0.6.

- All estimates e_{est} of data noise level are close to the best estimate 0.287253, which is computed by

$$e_{best} = \min_{w} \frac{||F[m_*]w - d||}{||d||},$$

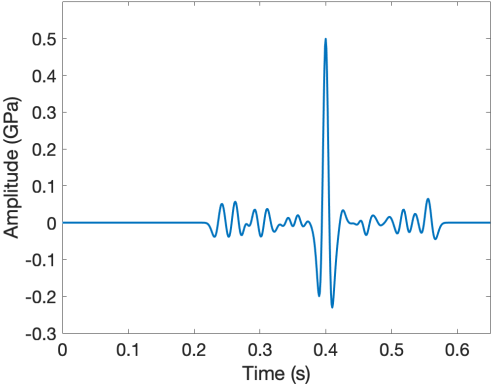
  where $m_* = 0.4$ s/km is the true slowness.
- All estimated slownesses m are close to the true slowness of 0.4 s/km.

VPM objective functions for the α selected by the noise estimation algorithm



Regardless of the values of the initial noise level, the penalty weights α from the noise estimation algorithm can help to convexify the VPM objective function.

Experiment 3: noise estimation for data with random noise



Data for experiment 3 which contains 30% random filtered noise.

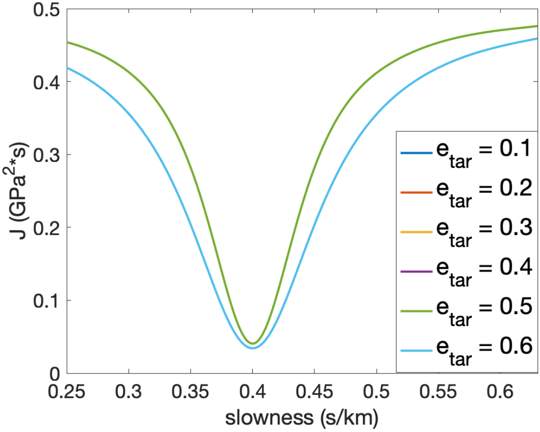
Experiment 3: noise estimation for data with random noise

initial noise level: e_{tar}	estimated noise level: e_{est}	penalty weight: α	estimated slowness: m (s/km)
0.1	0.256468	1.246196	0.400486
0.2	0.256468	1.245998	0.400492
0.3	0.256468	1.245833	0.400489
0.4	0.256468	1.246121	0.400488
0.5	0.256526	1.749055	0.400507
0.6	0.256468	1.246053	0.400489

Table: Noise estimation algorithm with initial target noise levels ranging from 0.1 to 0.6.

- All estimates e_{est} of data noise level are close to the best estimate 0.256468.
- All estimated slownesses m are close to the true slowness of 0.4 s/km.

VPM objective functions for the α selected by the noise estimation algorithm



Regardless of the values of the initial noise level, the penalty weights α from the noise estimation algorithm can help to convexify the VPM objective function.

Experiment 4: noise estimation for noise-free data

initial noise level: e_{tar}	estimated noise level: e_{est}	penalty weight: α	estimated slowness: m (s/km)
0.1	0.000000	3.854345	0.400000
0.2	0.000000	5.139127	0.400000
0.3	0.000000	8.672277	0.400000
0.4	0.000000	15.417382	0.400000
0.5	0.000000	24.089659	0.400000
0.6	0.000000	34.689109	0.400000

Table: Noise estimation algorithm with initial target noise levels ranging from 0.1 to 0.6.

- All estimates e_{est} of data noise level are close to the best estimate of 0.0.
- All estimated slownesses m are close to the true slowness of 0.4 s/km.

Conclusions

- Even this very simple single-trace transmission problem exhibits cycle skipping, so FWI can fail without a good enough initial guess.
- The ESI objective function can be efficiently solved using the Discrepancy Algorithm which maintains the data misfit within a reasonable range while also increasing the penalty parameter.
- ESI avoids cycle-skipping, allowing us to solve the inverse problem using standard local optimization.
- The discrepancy algorithm dynamically updates the penalty parameter α to more efficiently solve the inverse problem, but it requires an estimate of the data noise level.
- We can simultaneously update our estimate of the noise level in the data while solving the inverse problem.
- Goal is to test how noise estimation combined with ESI works on more realistic problems.