DOI:10.1190/image2022-3749692.1

Use of extended source inversion for estimating the noise level in seismic data

Huiyi Chen, The University of Texas at Dallas, William W. Symes, Rice University, Susan E. Minkoff, The University of Texas at Dallas*

SUMMARY

Modeled and recorded seismic data is contaminated by noise due to a variety of factors including geophones recording ambient noise unrelated to seismic exploration, equipment malfunction and limitations, and restrictive modeling assumptions. The level of noise in seismic data is not known a priori, hence being able to estimate the noise level in the data is a valuable tool for assessing the quality of mechanical Earth parameter estimates, especially resulting from inversion. Full waveform inversion (FWI) is a promising tool for estimating subsurface parameters such as wave velocity, but users of FWI encounter their own challenges. Specifically, commonly-used gradient-based local optimization techniques are well known to stall in geologically uninformative models if the starting guess for the optimization is not close enough to the desired global optimum. Extension-based methods relax physical constraints on model parameters to enlarge the search space of acceptable solutions, helping to reduce the impact of a poor initial guess and potentially convexifying the objective function. These extended inversion methods involve a penalty term that is added to the FWI least squares misfit. Then the challenge of performing the optimization shifts to adjusting the penalty weight to balance the reduction of data misfit with driving the penalty term towards physically-meaningful solutions. The source-extended objective function minimized using the discrepancy algorithm requires an estimate of the noise level in the data to proceed. In this work we illustrate an automated algorithm for simultaneously updating this noise estimate so that we bypass the cycle-skipping problem experienced by FWI and converge towards geologically meaningful parameter estimates.

INTRODUCTION

Measured seismic data contains noise that can be attributed to a wide variety of factors: vibrations from non-seismic events such as trucks driving over nearby roads in the vicinity of a survey, undersea rock slides, mechanical receiver failure, destruction of buried cables, etc. Moreover, no mathematical model ever captures all the physics of seismic energy generation and propagation.

We adopt the convention that "noise" means "unmodeled signal", whatever the origin of the modeling failure, and that "noise level" is the RMS size of the noise. Therefore finding the noise and its level is tantamount to finding the model parameters that best predict the data in the sense of least-squares (or Full Waveform Inversion (FWI)). While FWI is widely applied in both academic and industrial seismology, it remains a very active research topic due in large part to the tendency of gradient-based optimization to stagnate at uninformative models far from RMS error minimizers. As is well-known, this "cycle-skipping" phenomenon (Virieux and Operto (2009)) requires initial model choices that predict the arrival times of major data events within a half-wavelength at energetic data frequencies.

We describe a source extension algorithm which aims to simultaneously solve the FWI problem from an arbitrary initial guess and accurately estimate the noise level in the data. Source extension is one of numerous proposed modifications of FWI designed to suppress cycle-skipping (Symes, 2008; Huang et al., 2019; Pladys et al., 2021). The source extension algorithm relaxes physical constraints on the source (usually temporal or spatial localization) inherent in survey acquisition geometry, replacing these hard constraints by additive soft penalties. The penalty weights must be adjusted to ensure convergence to an approximate global FWI solution. We use an algorithm based on the discrepancy principle (Morozov (1984)) in which the penalty weight is dynamically updated as the model update progresses, to keep the data error within a prescribed range near a target data noise level (Fu and Symes, 2017). For best performance of the algorithm, this target noise level should be near the actual data noise level, which is *a priori* unknown.

We describe how to use (non-extended) FWI to update the target data noise level used in the discrepancy algorithm.This coupling of extended and non-extended inversion aims to make the combined algorithm independent of prior knowledge of the noise level, avoid cycle-skipping, and produce an estimate of noise level as an end product. We give a complete description of the algorithm and of an implementation for a very simple single-trace transmission problem. This simple model problem is used routinely to illustrate cycle-skipping (Virieux and Operto, 2009). In earlier work (Symes et al., 2020; Symes, 2022; Symes et al., 2022), the authors have given theoretical and numerical evidence that extended source inversion eliminates cycle-skipping for the single-trace transmission problem, and that the discrepancy algorithm appropriately controls the penalty weight, given a reasonable estimate of data noise. In this paper, we show that the noise estimation algorithm gives a reasonable estimate of true noise level along with an approximate global solution of the FWI problem.

THEORY

We divide the model parameters into two groups: source parameters, w, and medium parameters, m, respectively. Here "medium" means anything not explicitly describing a source mechanism, for example p-wave velocity. Assuming the source parameters appear linearly, the modeling operator may be written as $F[m]w$, where $F[m]$ is a linear operator depending on medium parameters, acting on source parameters.

As mentioned above, the acquisition geometry and preprocessing constrain the source parameters. We use W_p to denote the

set of physical (constrained) source parameters.

Single trace transmission example

An isotropic time-dependent point source, $w(t)$, radiates causally in a homogeneous 3D fluid with slowness m. The resulting acoustic pressure field is recorded at a distance r from this source. The recorded trace (or data) is obtained by convolving w with the acoustic 3D Green's function (see Courant and Hilbert (1962), Ch. 5) and is given by:

$$F[m]w(t) = \frac{1}{4\pi r} w(t - mr). \tag{1}$$

Assuming efficient signature deconvolution (for example), the wavelet may be assumed non-zero for only a short time interval, $-\lambda \le t \le \lambda$. All such wavelets with maximum time lag λ constitute the set W_p of "physical" sources in this example. We will not attempt a full-fledged assessment of deconvolution efficiency in the context of this simple model problem but simply regard λ as given.

Extended inversion

The FWI problem can be posed as: given data d, find a model (m, w) with w belonging to W_p, to minimize the error in the L_2 norm, namely:

$$e[m, w; d] = \|F[m]w - d\| / \|d\|. \tag{2}$$

If the medium parameter set contains velocity parameters, this minimization (by means of a gradient-based descent method) is likely to cycle-skip, that is, stagnate far from the global minimizer.

Extended source inversion aims to avoid cycle-skipping by searching over a larger set W of extended sources. To compensate for dropping the hard constraint defining the physical source space, we introduce a penalty operator, A, that has large output for input extended sources that are far from physical. The penalty function formulation of extended source inversion combines the FWI objective e^2 with a quadratic penalty term g^2, with

$$g[w; d] = \|Aw\| / \|d\|. \tag{3}$$

In this formulation, the extended source inversion task is: given data d and the penalty parameter α, minimize

$$J_\alpha[m, w; d] = e[m, w; d]^2 + \alpha^2 g[w; d]^2 \tag{4}$$

over (m, w), with w in the set W of extended sources.

For the single-trace transmission example, the extended source set W consists of all square-integrable (finite energy) functions of time. Since the physical sources (W_p) are non-zero only in the short time interval $-\lambda \le t \le \lambda$, an appropriate choice for the penalty operator A is $Aw(t) = tw(t)$. (This penalty is also used in Adaptive Waveform Inversion, another source extension method (Warner and Guasch, 2016)). Note that we can always fit the data exactly by suitable choice of source in the extended set W, regardless of slowness m. In particular, for the special case of penalty weight $\alpha = 0$, J_α achieves the optimal minimum value of zero for any value of slowness. This ability to fit any data with extended models is a key characteristic of successful extension methods.

Variable projection

Because the predicted data has different sensitivities to the source and medium parameters, simultaneous estimation of these two parameter sets is difficult. Instead we use a nested approach, namely, the Variable Projection Method (Golub and Pereyra, 2003) and define the new objective function

$$\tilde{J}_\alpha[m; d] = \min_w J_\alpha[m, w; d] = J_\alpha[m, w_\alpha[m; d]; d] \tag{5}$$

which gives the wavelet as the solution of the normal equations (for fixed value of medium parameters, m) via

$$w_\alpha[m, d] = \left(F[m]^T F[m] + \alpha^2 A^T A\right)^{-1} F[m]^T d. \tag{6}$$

Equation 6 can be solved by an iterative linear solver such as the Conjugate Gradient method. However, for the single trace transmission problem, the solution is available in closed form (Symes, 2022). Local minimization of $\tilde{J}_\alpha$ for the slowness, m, can be accomplished by a gradient-based descent method. In some cases (Huang and Symes, 2015; Symes, 2022) theoretical proof is available that any local minimizer is a satisfactory solution of the FWI problem, with generally better quality for larger α.

Discrepancy algorithm

Control of α is equivalent to control of the noise level, via the discrepancy principle. In particular, for increasing α, the data error component e of $\tilde{J}_\alpha$ increases while the penalty term g decreases. Given an estimate e_{tar} of the global minimum value of $e[\cdot, \cdot; d]$ (equation 2), Fu and Symes (2017) show how to update α starting with $\alpha = 0$ to achieve a value of e within a target relative error of e_{tar}. The discrepancy algorithm alternates the α updates with m updates via a local optimization algorithm, terminating when the size of $\nabla_m \tilde{J}_\alpha$ falls below a preset tolerance, thus finding an approximate stationary point of J_α for a maximal value of α consistent with the assumed data noise level. Detailed descriptions of the discrepancy algorithm with examples may be found in Fu and Symes (2017) and Symes et al. (2022).

Noise estimation algorithm

Given medium parameters m, the minimum value of the error $e[m, w; d]$ (equation 2) over all physical sources w provides an (over) estimate of the noise level. Further, for any value of the medium parameters m, minimization of e over w in the physical source space is an "easy" linear problem, thereby resulting in an optimal noise level estimate for that value of m. Given a target noise level, the discrepancy algorithm produces a medium parameter vector m which can be used to estimate the noise level. Recall that the source w in the physical source space W_p is a constrained version of the source in the extended source space W. Therefore, if the target noise level is too small, the penalty weight will also be small, and the data fit with extended w will be much better than the fit with physical w (overfitting). If the target noise level is too large, then the penalty weight will be too large. Assuming that the estimated medium parameters m are at least somewhat accurate, the physical data fit will likely be better than the extended data fit (underfitting). This observation suggests iterative replacement of the target data noise level in the discrepancy algorithm with the estimated noise level obtained by minimizing

the FWI objective in Equation 2 over physical sources w, using the medium parameters m output by the discrepancy algorithm. This iteration should terminate when the two data noise levels (target and estimated) are within a specified tolerance of each other. A simple algorithm to accomplish this goal is:

- initialize the assumed target noise level $e_{\text{tar}} > 0$, estimated noise level $e_{\text{est}} = e_{\text{tar}}$, and $\delta > 0$
- do
  - run discrepancy algorithm to minimize extended objective $\tilde{J}_\alpha$ with target noise level e_{tar}, over medium parameter m and source vector w in extended source space W;
  - compute minimum value e_{est} of e over source vector w in physical source space W_p, with medium parameter m from discrepancy algorithm;
  - if $|e_{\text{tar}} - e_{\text{est}}| > \delta e_{\text{est}}$ then $e_{\text{tar}} = e_{\text{est}}$, else terminate.

EXPERIMENTS

In this section we demonstrate the success of the noise estimation algorithm with several numerical examples based on the single trace transmission problem. To start we define the noise-free data $F[m]w$ in which the slowness $m = 0.4$ s/km, the offset between source and receiver $r = 1$ km, and the source wavelet is a 40 Hz Ricker centered at $t = 0$. Equation 1 shows that this noise-free data is simply a scaled, shifted Ricker wavelet. In our experiments we will consider data with both 30% coherent and 30% random noise added to this noise-free data. For the inversion we take as our initial slowness $m = 0.343$ s/km. The discrepancy algorithm is implemented as described in (Symes et al., 2022), using Brent's method with a stopping tolerance of 10^{-4}. The method finds a near-stationary point in m and also gives a simple update formula for α. The 40 Hz Ricker wavelet $w(t)$ is effectively zero for $|t| > 0.082$, so we take $\lambda = 0.082$ in the definition of the physical source space W_p for these examples.

Experiment 1: Discrepancy algorithm, coherent noise

Figure 1 shows the data with coherent noise added. This coherent noise is a copy of the noise-free data shifted so it is centered at $t = 0.5$ s and scaled by 0.3. The best estimated noise level is 0.287253, obtained by minimizing e (equation 2) over physical wavelets w with $m = 0.4$ s/km. In Experiment 1 we solve the inverse problem by the discrepancy algorithm with various target noise levels. We estimate the noise level after each inversion, by minimizing e (equation 2) over wavelets w in the physical source space W_p, fixing m at the value returned by the discrepancy algorithm. We repeat this experiment with various target noise levels ranging from $e_{tar} = 0.1$ to 0.6. The results are shown in Table 1. Note that for $e_{tar} = 0.1$ to 0.3 (rows 1 to 3 in Table 1), the slowness is inverted accurately, and the estimated noise levels in column 2 are all close to the best estimate 0.287253. For $e_{tar} = 0.4$ to 0.6 (rows 4 to 6 in Table 1), the large values of α impact the shape of the reduced extended objective function. As α increases, the reduced extended objective function behaves more and more like the FWI objective, with more than one local minimum. There is a local minimum near $m = 0.5$ s/km in the reduced-extended objective function for large α (see Figure 19 in (Symes et al., 2022)). Therefore, we would expect an incorrect slowness value of $m \approx 0.5$ s/km which occurs when e_{tar} is set to values above the true noise level (in the range of 0.4-0.6), with a correspondingly large value of e_{est}.

Experiment 2: Noise estimation algorithm, coherent noise

In Experiment 2 we use the same data as in Experiment 1 (with 30% coherent noise). We solve the inverse problem using the noise estimation algorithm. We repeat the experiment with different initial target noise levels ranging from $e_{tar} = 0.1$ to 0.6. For both Experiments 2 and 3 we choose the stopping tolerance δ for the noise estimation algorithm to be 0.1. Table 2 gives the final estimated noise levels in column 2. Note that all the noise levels estimated by this algorithm are close to the best estimated noise level of 0.287253, regardless of our starting guess for the noise level. The inverted slowness is also estimated accurately. Thus, a direct comparison between Tables 1 and 2 shows that we can correctly estimate both the noise level in the data and the global minimizer of the inverse problem if we allow the discrepancy algorithm to automatically update the target noise level during the inversion.

Experiment 3: Noise estimation algorithm, filtered random noise

Figure 2 shows the data with 30% random filtered noise added to the noise-free data. Table 3 shows the resulting estimates of noise level, penalty parameter α, and slowness m. The inverted estimates of the slowness are all accurate estimates of the true slowness of 0.4. The final estimated noise level e_{est} is very close to the the best estimated noise level of 0.256468, regardless of the initial target noise level.

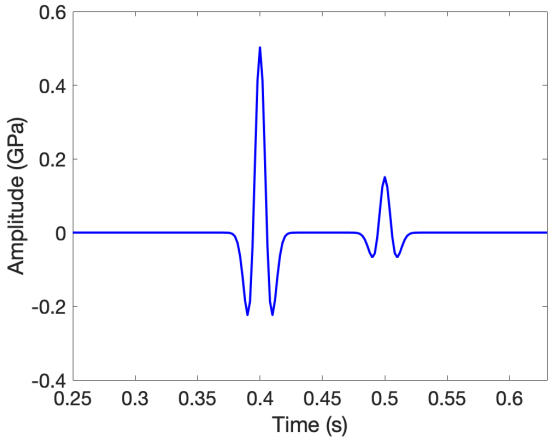

Figure 1: Data for experiments 1 and 2 produced by adding to the noise-free data a shifted, scaled multiple of itself.

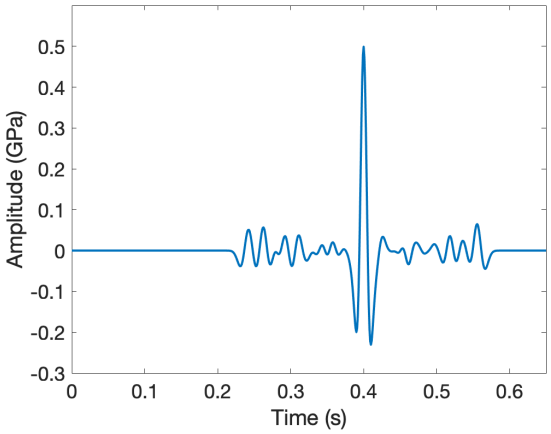

Figure 2: Data for experiment 3, containing the noise-free data contaminated with 30% random filtered noise.

e_{tar}	e_{est}	α	m
0.1	0.285162	0.573794	0.403991
0.2	0.287115	1.388210	0.400599
0.3	0.287253	2.136137	0.400156
0.4	0.957051	5.713097	0.499285
0.5	0.957510	16.235973	0.499981
0.6	0.957510	28.737005	0.499996

Table 1: Experiment 1, discrepancy algorithm with a range of target noise levels, applied to data of Figure 1. Target noise level shown in column 1, estimated noise level in column 2, penalty parameter in column 3, and estimated slowness in Column 4.

e_{tar}	e_{est}	α	m
0.1	0.287253	1.937027	0.400215
0.2	0.287253	2.786099	0.400065
0.3	0.286229	0.731463	0.402755
0.4	0.287253	1.551564	0.400431
0.5	0.287253	1.601533	0.400391
0.6	0.280798	0.352477	0.406161

Table 2: Experiment 2, noise estimation algorithm starting with initial target noise levels ranging from 0.1 to 0.6, applied to data of Figure 1. Note that all estimates e_{est} of data noise level are close to the best estimate 0.287253.

e_{tar}	e_{est}	α	m
0.1	0.256468	1.258596	0.400490
0.2	0.256468	1.313458	0.400494
0.3	0.256526	1.519694	0.400504
0.4	0.256468	1.979759	0.400500
0.5	0.256468	2.854779	0.400460
0.6	0.256468	0.384910	0.400040

Table 3: Experiment 3, noise estimation algorithm with initial target noise levels ranging from 0.1 to 0.6, applied to data of Figure 2. Note that all estimates e_{est} of data noise level are close to the best estimate 0.256468.

CONCLUSION

We have described an approach to noise estimation based on extended inversion. We demonstrate this approach in the context of a simple single-trace transmission problem - possibly the simplest wave inverse problem to exhibit cycle-skipping.

The "noise" concept underlying this approach is extremely simple: noise is unmodeled signal, and its size (noise level) is the global L^2 norm of the unmodeled signal. Thus finding noise is equivalent to solving the basic FWI problem, and our algorithm is based on the presumption that extended inversion can provide a good approximate FWI solution. This presumption has been given full mathematical justification in only a few cases, including the single trace transmission problem used here (Symes, 2022), although numerical evidence suggests its broader validity. Many noise characteristics other than misfit appear to influence the effectiveness of extended inversion and noise estimation (for example: spectral distribution, coherence, etc), even for the simplest examples (Symes et al., 2022). Our use of the discrepancy principle to control the penalty weight has a good theoretical foundation (Symes (2021), Appendix A) and has been applied in much more computationally complex settings (Fu and Symes, 2017). However it is not the only possible route to controlling penalties. For example, Aghamiry et al. (2019) has used the Alternating Direction Method of Multipliers to control penalization in the context of a different source extension. Finally, extension methods are only one class of potential cycle-skipping remedy (for example, (Métivier et al., 2018; Yang et al., 2018)). Presumably any of these alternative approaches to FWI could serve as the basis for noise estimation in the sense employed here.

While many open questions and avenues of investigation beckon, the results we have reported here suggests that noise estimation via inversion may be a feasible addition to the geophysicist's toolkit.

ACKNOWLEDGEMENTS

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

REFERENCES

Aghamiry, H., A. Gholami, and S. Operto, 2019, Improving full-waveform inversion by wavefield reconstruction with the alternating direction method of multipliers: Geophysics, **84**, no. 1, R125–R148, doi: https://doi.org/10.1190/geo2018-0093.1.

Courant, R., and D. Hilbert, 1962, Methods of mathematical physics, volume II: Wiley-Interscience.

Fu, L., and W. W. Symes, 2017, A discrepancy-based penalty method for extended waveform inversion: Geophysics, **82**, no. 5, R287–R298, doi: https://doi.org/10.1190/geo2016-0326.1.

Golub, G., and V. Pereyra, 2003, Separable nonlinear least squares: the variable projection method and its applications: Inverse Problems, **19**, R1–R26, doi: https://doi.org/10.1088/0266-5611/19/2/201.

Huang, G., and W. W. Symes, 2015, Full waveform inversion via matched source extension: 85th Annual International Meeting, SEG, Expanded Abstracts, 1320–1325, doi: https://doi.org/10.1190/segam2015-5872566.1.

Huang, G., R. Nammour, W. Symes, and M. Dollizal, 2019, Waveform inversion by source extension: 89th Annual International Meeting, SEG, Expanded Abstracts, 4761–4765, doi: https://doi.org/10.1190/segam2019-3216338.1.

Metivier, L., A. Allain, B. Brossier, Q. Merigot, E. Oudet, and J. Virieux, 2018, Optimal transport for mitigating cycle skipping in full-waveform inversion: A graph-space transform approach: Geophysics, **83**, no. 5, R515–R540, doi: https://doi.org/10.1190/geo2017-0807.1.

Morozov, V., 1984, Methods for solving incorrectly posed problems: Springer Verlag.

Pladys, A., R. Brossier, Y. Li, and L. Metivier, 2021, On cycle-skipping and misfit function modification for fullwave inversion: Comparison of five recent approaches: Geophysics, **86**, no. 4, R563–R587, doi: https://doi.org/10.1190/geo2020-0851.1.

Symes, W., 2008, Migration velocity analysis and waveform inversion: Geophysical Prospecting, **56**, 765–790, doi: https://doi.org/10.1111/j.1365-2478.2008.00698.x.

Symes, W. W., 2021, Solution of an acoustic transmission inverse problem by extended inversion: Theory: arXiv:2110.15494v1.

Symes, W., 2022, Error bounds for extended source inversion applied to an acoustic transmission inverse problem: arXiv:2110.15494v2.

Symes, W., H. Chen, and S. Minkoff, 2020, Full waveform inversion by source extension: Why it works: 90th Annual International Meeting, SEG, Expanded Abstracts, 765–769, doi: https://doi.org/10.1190/segam2020-3424509.1.

Symes, W., H. Chen, and S. Minkoff, 2022, Solution of an acoustic transmission inverse problem by extended inversion: arXiv:2201.08891.

Virieux, J., and S. Operto, 2009, An overview of full waveform inversion in exploration geophysics: Geophysics, **74**, no. 6, WCC127–WCC152, doi: https://doi.org/10.1190/1.3238367.

Warner, M., and L. Guasch, 2016, Adaptive waveform inversion: Theory: Geophysics, **81**, no. 6, R429–R445, doi: https://doi.org/10.1190/geo2015-0387.1.

Yang, Y., B. Engquist, J. Sun, and B. Hamfeldt, 2018, Application of optimal transport and the quadratic Wasserstein metric to full-waveform inversion: Geophysics, **83**, no. 1, R43–R62, doi: https://doi.org/10.1190/geo2016-0663.1.