Sparsity Promoting L1 Deconvolution: Application for a Global Seismology Parameter Search

G. Matharu, B. Postlethwaite[1]

Abstract. This report will outline an attempt to utilize the inherent sparsity in seismic receiver functions in order to provide better deconvolution results than the often utilized damped-L2 methods. By promoting sparsity with the wavelet transform the use of sparse solvers such as SPGL1 becomes possible. Synthetic tests are performed using the damped L2 convolution as well as the sparse wavelet transform L1 minimization using several wavelet basis. Real data is employed and the deconvolved results from the competing methods are used for an actual parameter grid search. The final results of the grid search are then compared. Further work on the sparse-in-wavelets L1 algorthim will be required if it is to beat the damped-L2 algorthim currently being employed.

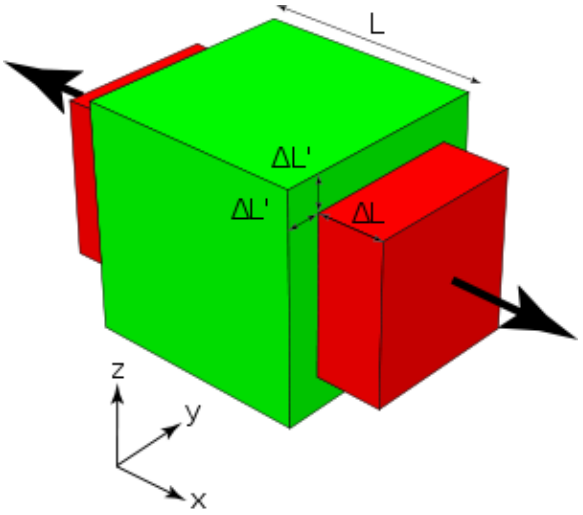

Figure 1. Poissons Ratio is the ratio of compression along one axis to extension along another

1. Introduction

Extracting material properties from seismic data spans industry and the research community. In the global seismic community material properties of the crust have been sought to validate existing geological models or provide clues for new ones. A good example of this is Poisson's ratio, which is the ratio between compression along one axis to extension on another when a material is stretched (see figure 1). Poisson's ratio is given by:

$$\sigma = \frac{1}{2}\frac{\left(\frac{V_p}{V_s}\right)^2 - 2}{\left(\frac{V_p}{V_s}\right)^2 - 1} \tag{1}$$

where the seismic velocities V_p and V_s determine the property entirely. Because of the direct link between seismic properties and Poisson's ratio this material property has been sought in numerous seismic studies such as Kanamori (2000) and Snyder (2010). All of these studies have performed a parameter search for $R = \frac{V_p}{V_s}$ and provided a background estimate for V_p. This allows for fast computation through a grid search and avoids the problem of the low-dependency of V_p in the grid search equations. Although Poisson's ratio can be estimated from R better geological constraints would be possible if V_p and V_s were resolved uniquely.

2. The Problem

In order to differentiate between V_p and V_s the travel times from seperate arrivals of energy must be compared. The first arrival a seismic station will pick up after the onset of an earthquake is from the higher velocity P-wave. The impulsive character and high amplitude of the P-wave arrival allow it to be windowed and filtered and used as an approximation for the source. There are three more arrivals of interest. The Ps arrival, which is P-energy converted to S-wave energy at the Moho boundary between the crust and the mantle. Two further arrivals follow which are the reflections from the free surface which have reflected again off the Moho and recorded at the station, figure (2). These reflected arrivals are usually of lower amplitude and often less well defined in the seismogram. If the three S-wave arrivals - the Ps, PpPs and PsPs - can be located in time this provides the data in which a model may be fit. Owing to the fact that the S-wave phases all travel within the same crustal thickess H, dividing the S-wave relected phases by the initial Ps phase will remove the dependence on crustal thickness. The traveltime ratios are then dependent on V_p and V_s only (Bostock, 2010). With this information a 2D grid search over the travel time equations with respect to V_p and V_s can be performed. The parameter V_p is not very sensitive to the traveltime differences between the S-wave phases, thus a very impulsive deconvolution result is needed to steepen the gradient in V_p space.

3. Deconvolution - L2

Currently simultaneous deconvolution with generalized cross validation is being employed as it performs well and is very fast. Since division in the fourier basis is deconvolution in time we can write:

$$\mathbf{R} = \mathbf{F}^{-1}[\mathbf{G}(\omega)]$$

where

$$\mathbf{G}(\omega) = \left[\frac{\sum_n^N \mathbf{S}_n\mathbf{P}_n^*}{\sum_n^N \mathbf{P}_n\mathbf{P}_n^* + \delta}\right]. \tag{2}$$

[1]Department of Earth and Ocean Sciences, University of British Columbia, Vancouver, B.C. Canada.

0148-0227/12/$9.00

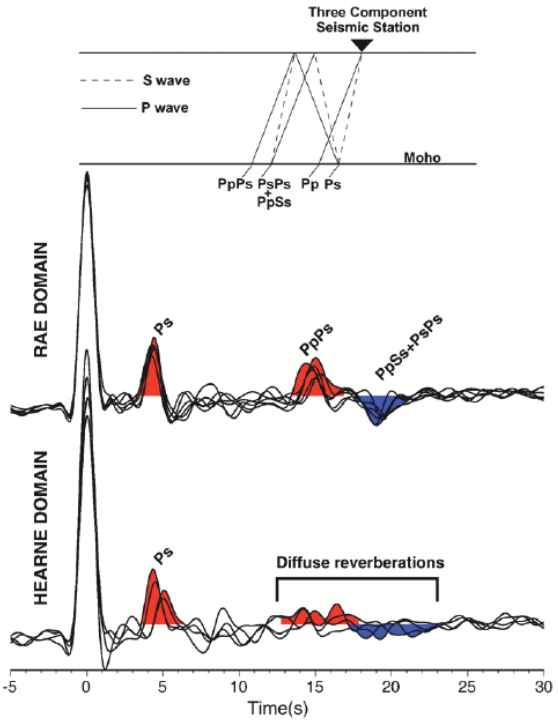

Figure 2. Top: Shows direct phases Pp and Ps from a seismic event hitting receiver. Also the reflected phases PpPs and PsPs are shown. The phases ending with 's' are the phases hitting the station as S-waves. These S-wave phases are used as the primary data, the direct Pp phase is used as the source, and the two are deconvolved to attain the receiver functions shown at Bottom: The receiver functions attained from deconvolution. Depending on the quality of the data and the sharpness of the deconvolution, the results may look like those given for the Rae Domain -better- or the Hearne Domain -worse.

Here $\mathbf{F}$ is the Fourier transform, $\mathbf{G}$ is the model, $\mathbf{S}$ is a matrix of S-wave seismograms, $\mathbf{P}$ is a matrix of P-wave source coda, $*$ is the complex conjugate and δ is the regularization parameter. δ is chosen by minimizing the Generalized Cross Validation function with respect to δ:

$$\min \frac{\left(\sum_n^N \sum_m^M \left(\overbrace{\mathbf{S}_n(\omega_m) - \mathbf{P}(\omega_m) \left[\frac{\sum_n^N \mathbf{S}_n \mathbf{P}_n^*}{\sum_n^N \mathbf{P}_n \mathbf{P}_n^* + \delta} \right]}^{\text{Misfit: more accurate with } \delta \to 0. \text{ Less Stable}} \right) \right)^2}{\left(\underbrace{NM - \sum_m^M \frac{\sum_n^N \mathbf{P}_n \mathbf{P}_n^*}{\sum_n^N \mathbf{P}_n \mathbf{P}_n^* + \delta}}_{\text{As } \delta \text{ increases denominator increases, promotes stability.}} \right)^2} \tag{3}$$

Since the sharpness of the deconvolution helps determine the confidence level contours on the final parameter estimation the prospect of a better deconvolution methods is very appealing. As the earth model being used in this parameter search is a simple layered model and the data is piecewise smooth the wavelet transform could be effectively used to promote sparsity. If we could leverage this fact and use an L1 technique which promotes sparse solutions we may be able to increase the accuracy of the deconvolved receiver functions.

4. Method

We now outline the framework under which the deconvolution problem was performed. The general convolution problem in seismology can be written as

$$u(t) = s(t) * g(t) + n \tag{4}$$

where $u(t)$ is the ground displacement recorded at seismograms, $s(t)$ is the source function, n is noise and $g(t)$ is the receiver function which we aim to obtain through a deconvolution of the source function from recorded seismograms. In this problem recorded seismograms were arranged into bins separated by slowness. Recorded seismograms with similar slownesses should provide receiver functions which are approximately the same, therefore binning them allows for the development of the deconvolution problem as an optimization problem. In an ideal case the receiver function would be represented by some type of spike train, however in the real Earth receiver functions will be noisier and blurred. As a result the receiver function can be considered to be piecewise smooth, this motivates the use of wavelets as a sparsifying transform.

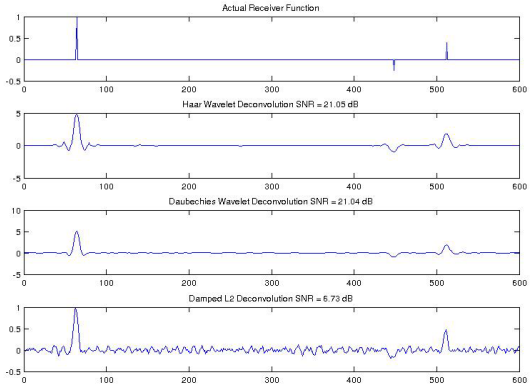

Figure 3. Synthetic Results: Actual receiver shown in first plot followed by the solution recovered by solving a wavelet based L1 optimization with Haar and Daubechies wavelets. Final plot is solution from an L2 based simulataneous deconvolution.

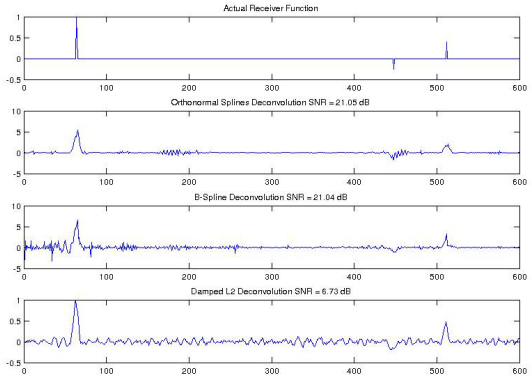

Figure 4. Synthetic Results II: Same as above but 2nd and 3rd plots are solutions recovered using orthonormal and B spline wavelets respectively.

Having identified the sparsifying transform to be used we represent the forward problem as a linear system.

$$\begin{aligned} \mathbf{y} &= \mathbf{A}\mathbf{x} \\ \mathbf{y} &= [\mathcal{F}^T\mathcal{F}(s)\mathcal{F}(g)]\mathcal{W}^T\mathbf{x} \end{aligned} \tag{5}$$

where $\mathbf{y}$ is a vector which holds n seismograms which are contained in any given p bin. $\mathbf{A}$ is a matrix which is defined in Eq. 5. and $\mathbf{x}$ is a vector which represents the wavelet coefficients of the receiver function $g(t)$. The action of the $\mathbf{A}$ matrix is to first perform the inverse wavelet transform on $\mathbf{x}$, and then perform a convolution between $s(t)$ and $g(t)$ in the Fourier domain. The final result would be the recorded seismograms $\mathbf{y}$ in the time domain. Rather than forming explicit matrices spot operators are used and combined using kronecker products.

The convex optimization problem is constructed as a basis persuit denoise problem,

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \|\mathbf{x}\|_1 \\ \text{subject to} \quad & \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 \leq \sigma \end{aligned} \tag{6}$$

where σ is an estimate of the required data misfit and is taken to be the norm of the noise in the case. We solve an optimization problem which promotes a sparse representation of the receiver function in the wavelet domain.

5. Synthetic Trials

Before conducting the deconvolution problem on a complete data set we created a synthetic example and tested our L1 based algorithm. The problem used windowed P wave arrivals from 5 seismograms as the source functions which were convolved with a synthetic receiver function to produce synthetic seismograms. 10% gaussian noise was added to the signal. All L1 optimization problems in this paper were solved using SPGL1, for the synthetic cases the solver was allowed to run to completion. We tested a range of different wavelet types during the trials namely Haar and Daubechies wavelets, for which shift invariant versions were used. Orthonormal and B spline wavelets were also used.

Ideally our synthetic receiver function should have been a piecewise smooth function rather than the current spike representation which is sparse in the dirac basis, and not necessarily sparse in the wavelet domain.

We used SNR $= -20log\frac{\|\mathbf{y}-w^T\mathbf{x}\|}{\|\mathbf{y}\|}$ as an initial measure of quality for the recovered receiver function, however this penalized solutions which were less sparse and can be seen in the low SNR for the L2 deconvolution. Whilst having a sparse receiver function was important, a more important requirement was that the recovered peaks be as sharp as possible, a property which allows for better constrained results when performing the grid search for V_P/V_S and depth to the moho at later stages. We did a simple comparison of the peak recovery for all wavelet families to determine how the performance was in this scenario.

All methods provided a good recovery of the receiver function, the two shift invariant wavelet transforms provided the sparsest recovery for the receiver function, however they also recovered broad peaks. Both spline wavelets recoverd narrower peaks but had a less sparse recovery.

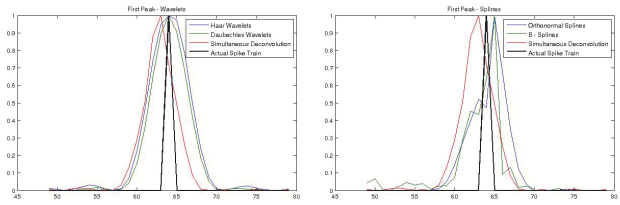

Figure 5. Comparison of peak recovery. Haar and Daubechies wavelets recover broader peaks, a less desirable result. B-Spline wavelets appear to provide the best recovery by recovering a narrower, better located peak when compared to ortho-spline wavelets.

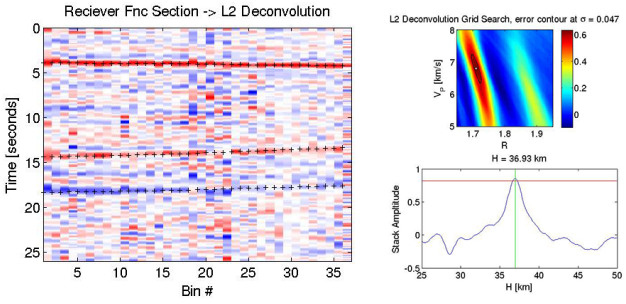

Figure 6. Receiver function section for L2 deconvolution and error contours for V_P v V_P/V_S. Receiver functions traces are plotted for all 36 bins with the result showm here. Also shown is a plot of V_P v R where $R = V_P$ v V_P/V_S. Line search results for depth are also plotted shown with a depth estimate obtained at H = 36.93km.

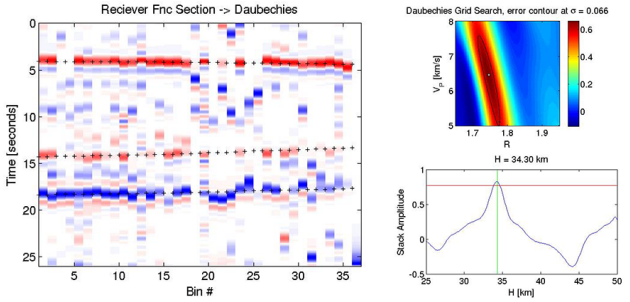

Figure 7. Receiver function section for L1 deconvolution using Daubechies wavelets. Receiver function section is noticeabley sparser than the L2 section, however the recovered peaks are also broader which leads to poorly constrained V_P v V_P/V_S in comparison. H = 34.30km

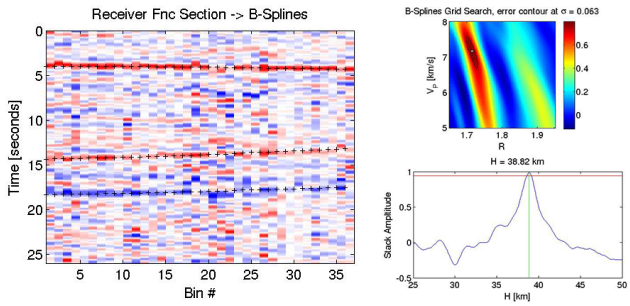

Figure 8. Receiver function section for L1 deconvolution using B-spline wavelets. Receiver function section provides narrower peak recovery and improves the error contour constraints. The depth estimate is also closer to that from L2 at H = 38.82km

Through a comparison of peak recovery, shown in Fig. 5, it was deemed that B-spline wavelets provided the most desirable properties in that the solution was sparse whilst recovering a narrow and well located peak.

6. Results

The data consisted of seismograms split into 36 p bins with the number of seismograms per bin varying between 1-8. For each p bin the L1 optimization problem defined in Eq. 6 was solved using 100 iterations of SPGL1. The initial guess provided to SPGL1 was the wavelet transformed solution obtained from the L2 simulataneous deconvolution. Each solution from SPGL1 provided the wavelet coefficients of a receiver function for the corresponding p bin from which the receiver function in the time domain was obtained. 3 different wavelet types were tested on the entire data set, these consisted of shift invariant Daubechies wavelets along with orthonormal and B spline wavelets. Once all 36 receiver functions were collected for a single wavelet type, the traces were collected by bin and a grid search was conducted to determine V_P ans V_S values and a line search was conducted to determine an estimate for the depth to the moho.

Fig. 6 shows the receiver function section obtained through the L2 based deconvolution method and acts as a benchmark for our comparisons. The results obtained through L2 deconvolution provide relatively sharp spikes in the recovered receiver functions. The error contours on V_P v V_P/V_S are also quite tightly constrained with a small standard error.

The first wavelet based L1 trial that we ran used shift invariant Daubechies wavelets. The receiver function section shows results consistent with our earlier synthetic trials showing a continuation of two properties, the first being that the recovered receiver functions are noticeabley sparser than the L2 deconvolution. However the peak recovery is not as sharp, just as in the synthetic trials, and this leads to poorly constrained V_P as shown by the larger error contours in Fig. 7. The recovered depth to the moho, $H(km)$, was significantly different from that obtained via the L2 method.

Ortho splines produced results similar to B-splines but with broader error contours for V_P v V_P/V_S and also provided a depth estimate less than that from L2 deconvolution. Both ortho and B splines (Fig. 8) improved upon Daubechies wavelets by providing sharper peak recoveries in the receiver functions and better constraints on the V_P v V_P/V_S error contours. Both also provided depth estimates closer to the L2 depth estimate. The B-spline depth estimate was the closest to the L2 estimate at H = 38.82km.

At present time the current wavelet framework has failed to produce results which are as tightly constrained as the L2 deconvolution method, however early results show promise and can likely be improved with some more sophisticated input.

7. Conclusion

We aimed to obtain a receiver function by deconvolving multiple source functions from multiple recorded seismograms which were grouped based on their slowness p. We developed a deconvolution framework based on L1 optimization and compared results to previous work using an L2 based method. We performed deconvolution by performing an inverse problem with an L1 regularization, promoting sparsity of the receiver function in the wavelet domain. SPGL1 was used to solve the convex optimization problem and was conducted for a range of different wavelet types including Daubechies, ortho-spline and B-spline wavelets. Receiver function solutions were plotted by bin and used to obtain V_P and V_s along with a depth estimate for the moho. In general Daubechies wavelets provided a very sparse representation of the receiver functions but failed to recover a sharp peak which lead to large error contours in V_P v V_P/V_S. Spline wavelets improved by producing narrower recovered peaks and improved error estimates however at present, all L1 wavelet methods failed to produce V_P v V_P/V_S constraints which were as tightly bound as the results derived using the L2 method.

8. Future Work

Even with a fairly preliminary L1 minimization setup the results are encouraging. Further work could focus on the 2D receiver function plots and usage of curvelets to interpolate between poor data. This would significantly improve the estimates made by the grid search. While redundant wavelet dictionaries were employed in the non-spline cases these authors are currently unsure if the B-spline methods were shift-invariant or not. This will require further investigation. Also during a recent presentation of the work contained in this report Felix Herrmann suggested looking at Radon Transformations. Both authors are unsure of the utility of these transformations and this will also require further research. However, with a sophisticated tool-chain in place we believe that L1 minimization and the utilization of sparsity inherent in this problem will result in better deconvolution than the current L2 method being employed. Since there is a large body of global seismological work which has been built on results from L2 deconvolution there is a lot of potential to apply these methods to other problems. Along the same lines, once sparsity is being utilized, the larger compressive sensing framework should surely find a place in the global seismic research community. There is vast potential in this arena and likely many exciting avenues for phd research. Both the authors of this report are interested in exploring this new paradigm further in the context of global seismology.

References

Herrmann, F. J., (2012), Class material from EOS 513: Imaging and Estimation with Wavelets, *University of British Columbia*

Bostock, M. G., Kumar, M. R., (2010), Bias in seismic estimates of crustal properties, *Geophysical Journal International, 182*, 403–407.

Thompson, D. A., Bastow, I.D., Helffrich, G., Kendall, J-M., Wookey, J., Snyder, D.B., Eaton, D.W., (2010), Precambrian crustal evolution: Seismic constraints from the Canadian Shield, *Earth and Planetary Science Letters, 297*, 655–666.

Kanamori, H. (2000), Moho depth variation in southern California from teleseismic receiver functions, *Geophysical Research, 105*, 2969–2980.