Using SVD for improved interferometric Green's function recovery.

Gabriela Melo[*1], *Alison Malcolm*[1], *Dylan Mikessel*[2], *and Kasper van Wijk*[2]
[1]*Earth Resources Laboratory - Earth, Atmospheric, and Planetary Sciences Department - Massachusetts Institute of Technology*
[2]*Physical Acoustic Lab - Department of Geosciences - Boise State University*

SUMMARY

Seismic interferometry is a technique used to estimate the Green's function (GF) between two receiver locations, as if there were a source at one of the locations. By crosscorrelating the recorded seismic signals at the two locations we generate a crosscorrelogram. Stacking the crosscorrelogram over sources generates an estimate of the inter-receiver GF. However, in most applications, the requirements to recover the exact GF are not satisfied and stacking the crosscorrelograms yields a poor estimate of the GF. For these non-ideal cases, we enhance the real events in the virtual shot gathers by applying Singular Value Decomposition (SVD) to the crosscorelograms before stacking. The SVD approach preserves energy that is stationary in the crosscorrelogram, thus enhancing energy from sources in stationary positions, which interfere constructively, and attenuating energy from non-stationary sources that interfere distructively. We apply this method to virtual gathers containing the virtual refraction artifact and find that using SVD enhances physical arrivals. We also find that SVD is quite robust in recovering physical arrivals from noisy data when these arrivals are obscured by or even lost in the noise in the standard seismic interferometry technique.

INTRODUCTION

Seismic Interferometry (SI), first envisioned by Claerbout (1968), estimates the Green's function (GF) between two receivers. This technique, often referred to as the virtual source method in exploration seismology (Bakulin and Calvert, 2006), requires that the sources completely surround the receivers (Wapenaar and Fokkema, 2006). This assumption is rarely met in practice and results in a degradation in the quality of the recovered GF. To overcome this problem, we study the collection of crosscorrelated traces from many sources for a pair of receivers, referred to as a crosscorrelogram. Crosscorrelograms contain pre-stack information, e.g., high frequency data that is often lost during stacking. Therefore, it is natural to expect that we can extract parameters of interest from the crosscorrelogram.

We filter the crosscorrelograms using the singular value decomposition (SVD) (see e.g. Hansen (1999)), a numerical technique commonly used in seismic data processing (Ulrych et al., 1988; Sacchi et al., 1998), with the goal of enhancing events that are coherent across multiple sources, i.e. stationary energy. Hansen et al. (2006) showed the relationship between singular values and frequency – larger singular values correspond to lower frequencies and smaller singular values correspond to higher frequencies. We apply SVD to crosscorrelograms, relying on the fact that signal due to stationary sources occurs at low spatial frequency (in the source coordinate) and signal from non-stationary sources (and random noise) generally occurs at higher spatial frequencies. After decomposing the crosscorrelogram into components that depend on spatial frequency, we expect the signal from stationary sources to correspond to large singular values (i.e., low spatial frequency). Thus, we decompose the crosscorrelogram using SVD, construct a lower-rank approximation of the crosscorrelograms, using only the largest singular values, and stack the lower-rank crosscorrelogram to estimate the GF. This approach recovers the correct GF using fewer sources than required in the standard SI technique.

Mikesell et al. (2009) showed that in shot gathers estimated using SI (i.e., virtual shot gathers) in a two-layered model with head waves, an artifact arises from the crosscorrelation of head waves recorded at the two receivers. They call this artifact the virtual refraction. Rather than trying to suppress this artifact, they showed that it can be used to estimate the depth to the interface and the velocity in the model layers. The real head wave, also present in the virtual shot gather, contains this information as well. However, they showed that its amplitude is considerably weaker than the virtual refraction, making the latter easier to interpret. We apply the SVD technique described above to the same synthetic data set used by Mikesell et al. (2009) in an attempt to further enhance the virtual refraction. We find instead that SVD enhances the real head wave. We also find that applying SVD to a noisy version of this data set enhances the reflected wave, even when it is almost completely obscured by noise in the standard virtual shot gather.

THEORY

Wapenaar and Fokkema (2006) showed that the exact Green's function between two receivers at locations $\boldsymbol{x}_A$ and $\boldsymbol{x}_B$ can be represented by the integral over sources distributed throughout a closed surface surrounding the two receivers,

$$\hat{G}(\boldsymbol{x}_A,\boldsymbol{x}_B,\omega)+\hat{G}^*(\boldsymbol{x}_A,\boldsymbol{x}_B,\omega) = \oint_S \frac{-1}{i\omega\rho(\boldsymbol{x})}(\partial_i\hat{G}(\boldsymbol{x}_B,\boldsymbol{x},\omega)\hat{G}^*(\boldsymbol{x}_A,\boldsymbol{x},\omega) - \hat{G}(\boldsymbol{x}_B,\boldsymbol{x},\omega)\partial_i\hat{G}^*(\boldsymbol{x}_A,\boldsymbol{x},\omega))n_i dS\ , \quad (1)$$

where $\hat{G}(\boldsymbol{x}_A,\boldsymbol{x}_B,\omega)$ is the frequency domain GF for a receiver at $\boldsymbol{x}_A$ and a source at $\boldsymbol{x}_B$ corresponding the to causal GF in the time domain, $\hat{G}^*(\boldsymbol{x}_A,\boldsymbol{x}_B,\omega)$ is the complex conjugate GF corresponding to the anti-causal GF in the time domain, $\rho(\boldsymbol{x})$ is the density, ω is the angular frequency, $i = \sqrt{(-1)}$, and S is the closed surface of sources surrounding the receivers. We recall that the physical interpretation of $\partial_i\hat{G}(\boldsymbol{x}_A,\boldsymbol{x},\omega)$ is as the GF from a dipole source at $\boldsymbol{x}$ recorded at $\boldsymbol{x}_A$ and $\hat{G}(\boldsymbol{x}_A,\boldsymbol{x},\omega)$ is the GF from a monopole source at $\boldsymbol{x}$ recorded at $\boldsymbol{x}_A$. Wapenaar and Fokkema (2006) simplify equation 1 by making the following assumptions: 1) all sources lie in the far-field (i.e., the distance from the source to the receivers and scatterers is large compared to the wavelength); 2) rays take off approximately normal from the integration surface S; 3) the medium outside the integration surface S is homogeneous, such that no energy going outward from the surface is scattered back into the system; 4) the medium around the source is locally smooth (the high frequency approximation). With these assumptions, the spatial derivative is approximated as $n_i \partial_i \hat{G} \approx (i\omega \hat{G})/c(\boldsymbol{x})$ in equation 1, which simplifies to

$$\hat{G}(\boldsymbol{x}_A, \boldsymbol{x}_B, \omega) + \hat{G}^*(\boldsymbol{x}_A, \boldsymbol{x}_B, \omega) \approx \oint_S \frac{-2i\omega \hat{G}(\boldsymbol{x}_B, \boldsymbol{x}, \omega)\hat{G}^*(\boldsymbol{x}_A, \boldsymbol{x}, \omega)}{\rho(\boldsymbol{x})c(\boldsymbol{x})} dS \ . \quad (2)$$

We now have an equation for crosscorrelation SI that requires only monopoles sources. In general, we only have monopole seismic sources; therefore, equation 2 is used in the SI method and the amplitudes of the GF are lost.

We now explore decomposing the crosscorrelogram using SVD. Consider a matrix $\boldsymbol{A}$ such that each column of $\boldsymbol{A}$ corresponds to a crosscorrelated trace for one source. Then $\boldsymbol{A} = \boldsymbol{A}(t, s)$ is the crosscorrelogram. Suppose n_s is the number of sources, n_t is the number of time samples, t is time, s is the source, $\boldsymbol{e} = (1, ..., 1)$ is the $n_s \times 1$ vector with of ones, $\mathbf{b} = \mathbf{b}(t)$ is the GF (i.e., the stack of the crosscorrelogram $\boldsymbol{A}$ along the source dimension).

Using this notation, the stack of the crosscorrelogram is

$$\boldsymbol{A}\boldsymbol{e} = \mathbf{b} \ . \quad (3)$$

Consider the SVD decomposition of the crosscorrelogram matrix $\boldsymbol{A}$

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^t \ , \quad (4)$$

where $\boldsymbol{U}$ is the orthonormal $n_t \times n_t$ matrix of left singular vectors (a basis for the columns of A), Σ is the $n_t \times n_s$ diagonal matrix whose elements are the singular values of $\boldsymbol{A}$, and $\boldsymbol{V}$ is the orthonormal $n_s \times n_s$ right singular vector matrix (a basis for the rows of $\boldsymbol{A}$).

We can use the SVD decomposition to express the GF. Stacking $\boldsymbol{A}$ over the source dimension we get

$$\boldsymbol{A}\boldsymbol{e} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^t\boldsymbol{e} = \mathbf{b} \ . \quad (5)$$

We call Σ' the reduced Σ matrix, containing only the largest singular values. Then we construct a lower-rank approximation of the crosscorrelogram

$$\boldsymbol{A}'\boldsymbol{e} = \boldsymbol{U}\boldsymbol{\Sigma}'\boldsymbol{V}^t\boldsymbol{e} = \mathbf{b}', \quad (6)$$

where $\mathbf{b}'$ is the approximate GF.

To illustrate this idea, consider the source-receiver geometry in a constant velocity/density medium, figure 1. The sources directly to the left and right of the receivers are in the stationary region, and the sources above and below are in non-stationary regions (Snieder, 2004). The gaps in sources break the requirement of a closed surface S in equation 2, i.e. that the receivers are completely surrounded by sources. In our numerical modeling we use 13 sources in each of the stationary zones and nine sources in each of the non-stationay zones. We used a Ricker wavelet as our source function and all receiver data are modeled analytically. Figure 2(a) shows the standard crosscorrelogram as well as the standard stack, figure 2(c), estimated by stacking the crosscorrelogram over source azimuth. The standard crosscorrelogram contains energy from stationary sources that contribute to the real GF and residual energy from non-stationary sources that causes artifacts due to incomplete cancellation. The rank-2 crosscorrelogram, figure 2(b), and corresponding GF, figure 2(d), are free from the non-stationary energy and fluctuations. We thus conclude that in this example, using SVD leads to a more accurate estimate of the GF. Next we show a more realistic 2D acquisition and discuss the application of SVD to a situation with multiple wave modes.

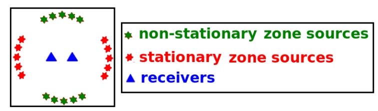

Figure 1: Source-receiver geometry for homogeneous model.

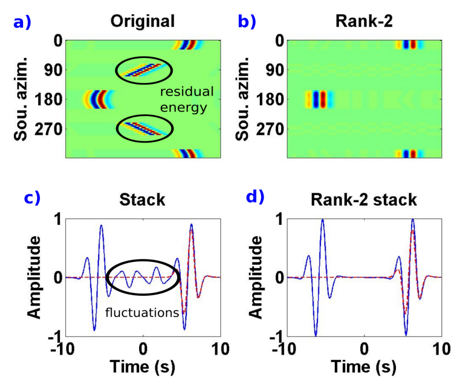

Figure 2: Standard (a) and rank-2 crosscorrelograms (b) and corresponding GFs (c, d) for the source-receiver geometry shown in figure 1. The blue lines are the recovered GFs and the red lines are the causal part of exact GFs for comparison.

EXAMPLES

We now apply the SVD technique discussed above to the same synthetic data set used by Mikesell et al. (2009). Consider the 2-layer acoustic model shown in figure 3. The top layer has velocity $V_0 = 1250m/s$, the bottom layer has velocity $V_1 = 1750m/s$, and density is constant throughout the model. We attempt to create a virtual shot gather as if there were a source at receiver r_1 using SI. Figure 4 shows the real shot gather for comparison. Note the direct, reflected, and refracted waves present in the 'real' shot gather.

A 2D array of 110 sources is placed to the left of the receiver line (figure 3), resembling a standard 2D off-end seismic survey. The wavefield generated by each source is recorded at

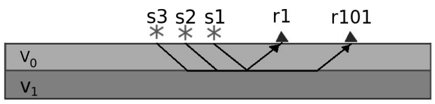

Figure 3: Virtual refraction source-receiver geometry.

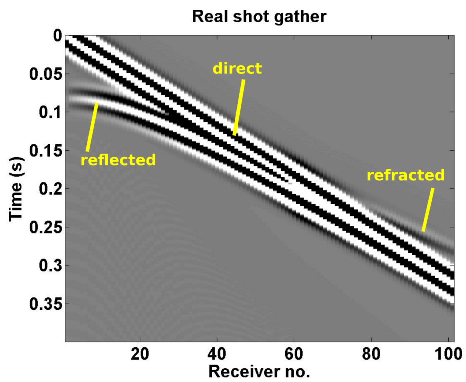

Figure 4: Modeled shot gather with a source placed at the location of receiver r_1. We have convolved the real shot gather with the 40 Hz Ricker wavelet used as a source.

each receiver. The GF between r_1 and each of the other receivers is obtained using crosscorrelation SI.

Figure 5(a)-(b) shows the standard and the rank-1 crosscorrelograms for receivers r_1 and r_{18}. Figure 5(c)-(d) shows the causal part of the corresponding GF estimates in blue and the exact GF in red. The phase difference between the estimate and the exact GF's is caused by the approximation to the closed surface integral over S. The amplitude of the reflected arrival, relative to that of the direct arrival is more accurately recovered using the rank-1 approximate crosscorrelogram.

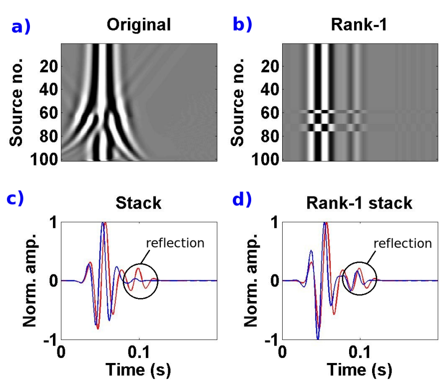

Figure 5: Standard (a) and rank-1 (b) crosscorrelograms and corresponding GFs (c, d) for receivers r_1 and r_{18}.

Repeating this procedure for all receivers, we create a standard virtual shot gather, figure 6, and a rank-1 shot gather, figure 7, for a virtual source at r_1. As is expected from the results in figure 5, the rank–1 approximation has better recovered the reflected wave than in the standard virtual shot record. In addition, the real refracted wave is clearer in the rank-1 approximation than in the standard virtual shot gather. Each trace in the shot records (real and virtual) are normalized individually such that all direct arrivals have a peak amplitude of 1. All gathers are displayed with the same clipping; they are clipped to show the weaker arrivals we are attempting to enhance. However, this means that geometrical spreading, or any other offset dependent amplitudes in general, will be lost. Also, as is standard in SI, we do not recover the exact amplitudes due to the lack of dipole sources in the approximate equation 1.

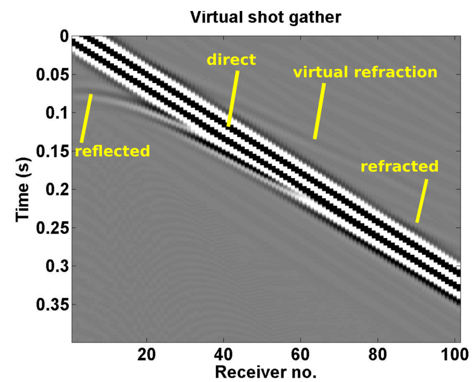

Figure 6: Interferometric virtual shot record for virtual source at r_1. See Mikesell et al. (2009) figure 4(a). We recover all the events in figure 4, plus the virtual refraction.

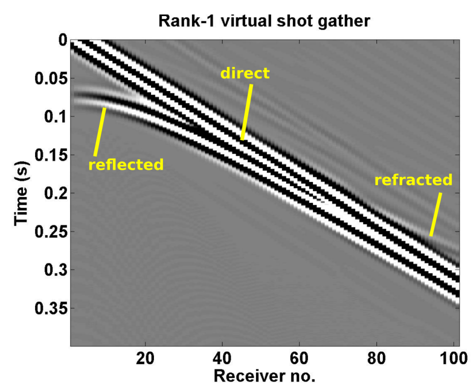

Figure 7: Rank-1 virtual shot gather. The real events are enhanced when this shot gather is compared to figure 6.

We now add pseudo-random noise to the data before crosscorrelation to test the SVD method's robustness in the presence of noise. Figure 8 shows the real shot gather plus noise. Figures 9(a)-(b) show again the standard and rank-1 crosssorrelograms for r_1 and r_{18}. Figures 9(c)-(d) show the respective GF estimates in blue and the exact GFs in red. We again see that the reflected wave GF amplitude is more accurately recovered using the rank-1 crosscorrelogram.

In the standard , figure 10, and in the rank-1 virtual shot gather, figure 11, we see that the rank-1 shot gather has recovered the reflected wave even though it is nearly obscured by noise in the standard shot record. We also see some energy from the real refracted wave in the rank-1 shot gather, whereas in the standard shot gather this arrival is completely lost in the noise.

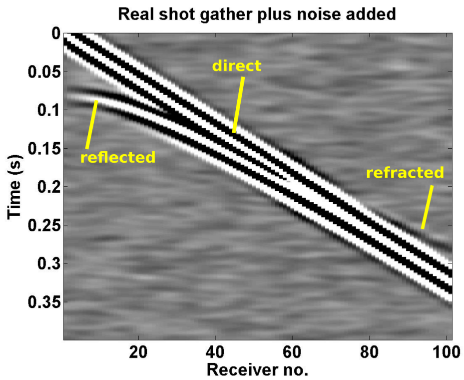

Figure 8: Modeled shot gather, similar to figure 4, but with added noise.

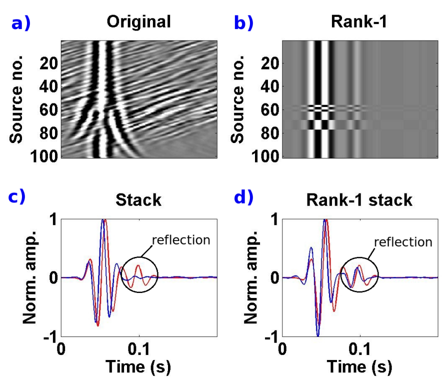

Figure 9: Standard and rank-1 crosscorrelograms and GFs, similar to figure 5. Noise has been added to the synthetic data so that the reflection and refraction are difficult to see in the standard virtual shot gather.

CONCLUSIONS

An outstanding problem in seismic interferometry remains the accurate estimation of the GF when the source coverage is not ideal. We have shown theoretically why using SVD to approximate crosscorrelograms before stacking is a promising approach to alleviate this problem. From the examples shown, we conclude that the rank-1 approximation of the correlogram highlights reflected and refracted arrivals that are not properly recovered using the standard crosscorrelogram. In addition, we find that SVD is able to enhance arrivals that would otherwise be obscured by noise. Future directions include using SVD on real data and to separate multiply-scattered signal from singly scattered signal in the crosscorrelogram as well as using the covariance of the correlogram to locate stationary points necessary, for example, to estimate the depth of a layer from the virtual refraction.

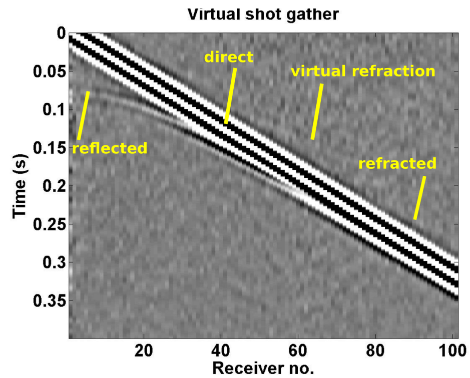

Figure 10: Standard virtual shot gather, similar to figure 6, but with added noise.

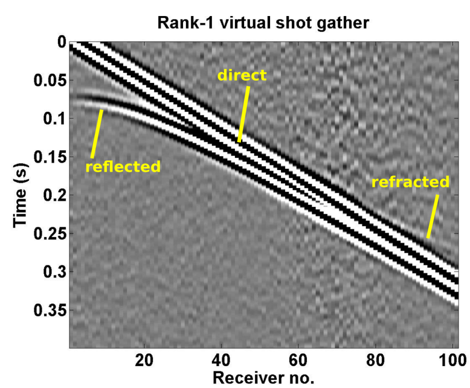

Figure 11: Rank-1 virtual shot gather with added noise. Note the enhanced reflected wave compared to the reflected wave in the standard virtual shot gather in figure 10. Some of the physical refraction energy is recovered here while it is lost in standard virtual gather.

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2010 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

Bakulin, A., and R. Calvert, 2006, The virtual source method: Theory and case study: Geophysics, **71**, no. 4, SI139–SI150, doi:10.1190/1.2216190.

Claerbout, J. F., 1968, Synthesis of a layered medium from its acoustic transmission response: Geophysics, **33**, 264–269, doi:10.1190/1.1439927.

Hansen, C., 1999, Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion: Society for Industrial Mathematics.

Hansen, P. C., M. E. Kilmer, and R. Kjeldsen, 2006, Exploiting residual information in the parameter choice for discrete illposed problems: BIT Numerical Mathematics, **46**, no. 1, 41–59, doi:10.1007/s10543-006-0042-7.

Mikesell, D., K. van Wijk, A. Calvert, and M. Haney, 2009, The virtual refraction: Useful spurious energy in seismic interferometry: Geophysics, **74**, no. 3, A13–A17, doi:10.1190/1.3095659.

Sacchi, M. D., T. J. Ulrych, and Magnuson, 1998, Eigenimage analysis of common offset sections: Signal-to-noise enhancement and pre-stack data compression.

Snieder, R., 2004, Extracting the Green's function from the correlation of coda waves: a derivation based on stationary phase: Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics, **69**, no. 4, 046610, doi:10.1103/PhysRevE.69.046610. PubMed

Ulrych, T. J., S. Freire, and P. Siston, 1988, Eigenimage processing of seismic sections: SEG Technical Program Expanded Abstracts, **7**, no. 1, 1261–1265, doi:10.1190/1.1892508.

Wapenaar, K., and J. Fokkema, 2006, Green's function representations for seismic interferometry: Geophysics, **71**, no. 4, SI33–SI46, doi:10.1190/1.2213955.