

Virtual Receivers From Seismic Interferometry: First Steps Towards Improving Microseismic Locations

Franklin G. Horowitz and Larry D. Brown

Cornell University, Dept. of Earth and Atmospheric Sciences

Keywords

Microseismic locations, virtual seismic receivers, seismic interferometry, seismology

Introduction

A remarkable recent result from seismic interferometry allows one to create *virtual seismic receivers* (Curtis et al., 2009). By recording an original earthquake on an array of seismic receivers, one can estimate seismograms from other events as if they were recorded at the original event's location. This estimation occurs by signal processing the other events' seismic records from the array in conjunction with the original event's records.

Here, we present some *preliminary* efforts attempting to use Curtis et al.'s result to improve estimates of hypocentral locations in microseismic clouds.

Theory

Seismic Interferometry

For our purposes, the key expression from Curtis et al. (2009) is their equation (SI-16) from the Supplementary Information, which we reproduce here:

$$M_{ip}^2 M_{mq}^1 \partial_p \partial_q' G_{im}^h(\mathbf{x}_2 | \mathbf{x}_1) = iK\omega \int_S u_n(\mathbf{x}' | \mathbf{x}_2) u_n^*(\mathbf{x}' | \mathbf{x}_1) d\mathbf{x}' \quad (1)$$

This (complex valued) expression is shown in the Fourier domain – hence the multiplications are convolutions – and the explicit dependence on frequency has been suppressed in the expression. The geometry is as displayed in Figure 1 with the boundary integral due to the use of the Green/Gauss theorem – in practice the integral is approximated by a sum over discrete locations. The vectors \mathbf{x}_1 and \mathbf{x}_2 are the locations of two distinct earthquakes, while the \mathbf{x}' are the locations of the array receivers. The homogenous Green's function G_{im}^h has a source at \mathbf{x}_1 and is received at location \mathbf{x}_2 – it is a linear combination of the causal and the conjugated-a-causal Green's functions, and it is symmetric with

respect to source and receiver by reciprocity. The repeated index summation convention is used. The constant K absorbs all of the C_{ijkl} elasticity constants of the full expression, and is different for different geometries of earthquakes (Curtis et al., 2009) – we use it here as a mere constant of proportionality. Spatial gradients in the p direction are denoted ∂_p . M_{mq}^1 and M_{ip}^2 denote the moment tensors for events number 1 and 2 respectively. The $i\omega$ in the right-hand side is a temporal derivative in the Fourier Domain (transform sign convention dependent) probably being used here as an approximation of spatial gradients in this wave equation based theory. The superscripted star denotes complex conjugation, such that $u_n u_n^*$ in the RHS is a cross-correlation between the two displacement seismograms.

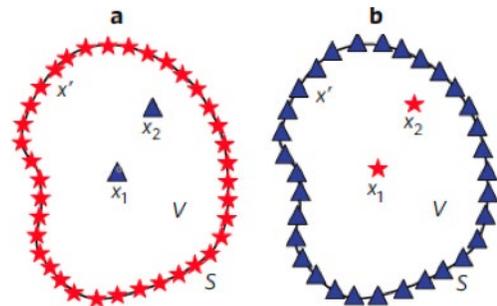


Figure 1. (From Curtis, et al. 2009). a) The standard interferometry geometry. Seismic events are depicted as red stars, receivers as blue triangles. b) The virtual receiver geometry. An array of receivers (blue triangles) receives seismograms from 2 distinct seismic events (red stars). The Green's function between the events can be estimated via Equation 1.

The full expression in Curtis et al. (2009) – for which Equation 1 is an approximation – requires unrealistic (or at the very least, rare in the real world) strain-gauge “dipole” seismometers. Equation 1 is the more practical displacement “monopole” seismometer expression, which allows arbitrary moment-tensor events to be used as either a source or a virtual receiver event. To keep the numerical experiments simple and appropriate for our first-arrival estimates, we choose to restrict our attention to explosive sources and virtual receivers with simple (spherical) moment tensor representations. Hence, after substituting Kronecker deltas for the

moment tensors appropriately (i.e. $M^l_{mq} = \delta_{mq}$ and $M^2_{ip} = \delta_{ip}$) and a little algebra using the definition of moment tensors, we find:

$$\partial_i u^h_i(\mathbf{x}_2 | \mathbf{x}_1) = iK\omega \int_S u_n(\mathbf{x}' | \mathbf{x}_2) u_n^*(\mathbf{x}' | \mathbf{x}_1) d\mathbf{x}' \quad (2)$$

In words, Equation 2 states that the homogenous volumetric strain seismogram (i.e. proportional to the seismogram recorded by a hydrophone) propagating from \mathbf{x}_1 to \mathbf{x}_2 can be estimated from observing the pressure seismograms from both earthquakes using a summation over the entire set of receiver stations as an approximation of the full surface integral.

In essence, we have created a virtual hydrophone receiver at position \mathbf{x}_2 !

An interesting feature of Equation 2 (or indeed, Equation 1), is that the math places no restrictions on whether $u_n(\mathbf{x}' | \mathbf{x}_2)$ occurs before or after $u_n(\mathbf{x}' | \mathbf{x}_1)$ in time. As long as we have both sets of seismograms recorded, we can employ the equation. For our present purposes, this means that the choice of the best located (or “nicest” moment tensor) virtual receiver(s) event(s) can be made or adjusted at any time after the activity of a microseismic swarm. There are other (related) important consequences of this order-blind feature that we hope to flesh out with future work.

Improving the Microseismic Locations

We focus on the Double Difference (DD) location algorithm of Waldhauser and Ellsworth (2000), and its variants (e.g. Zhang and Thurber, 2003; or the HypoCC code of Foulger and Julian, 2013). The basic technique minimizes traveltime residuals of the form:

$$dr_k^{ij} = (t_k^i - t_k^j)^{obs} - (t_k^i - t_k^j)^{cal} \quad (3)$$

In this expression, dr is a DD residual, the t 's are travel times, the superscripted indices refer to the earthquake events, the subscripted index refers to the recording station, and the superscripted “*obs*” and “*cal*” refer to the measured *observation* and the velocity model derived *calculated* values respectively. The actual residual being minimized by an inversion procedure is the sum over all stations of Equation 3.

Notice that the DD residual is composed of two conceptually different parts. The “*obs*” component relates to the difference in travel time between two events observed from the same instrument. This is the primary observation of the method, and contains all of the information about the relative hypocentral locations, source times, and the velocity field encountered by the actual source-receiver ray-paths. The “*cal*” component is composed of differences in travel times due to the assumed or estimated velocity models.

Virtual receivers can help improve inversions based on such residuals in several ways. By using Equations 1 or 2 to increase the effective number of arrival time observations, the statistics in the DD inversion are improved. By placing virtual receivers closer to the swarm of microseismicity than possible with real receivers, the effects of inaccurate velocity models can be ameliorated. The virtual receivers’ close-in geometry also allows for better ray-path coverage for tomographic estimates –improving velocity model estimates for the rockmass surrounding the swarm, hence allowing the possibility of iterative improvement in the absolute locations. There are probably other reasons that will become apparent as this new technique matures.

Simulations

We simulate a microseismic cluster (Figure 2) as being located on a planar fault with attitude N45W 45NE, passing through the point (6000,5000,1000) meters. The (right handed) coordinate system is shown and has its x component increasing due north, its y component increasing due east, and its z component increasing downwards. The receiver array is arranged at 30-degree increments around a circle on the surface with radius 2500 meters, and its center located at (5000,5000,0). Note that the cluster is offset with respect to the center of the receiver circle. There are 25 (blue) microseismic cluster events distributed on the fault plane, with the center event (red) designated as the “Virtual Receiver” (VR) – which plays the role of \mathbf{x}_2 in equations 1 or 2. The numerical model extends 10x10x2 km as shown. There is a free surface at $z=0$, and all other boundary surfaces use the non-reflecting boundary conditions implemented in the modeling code.

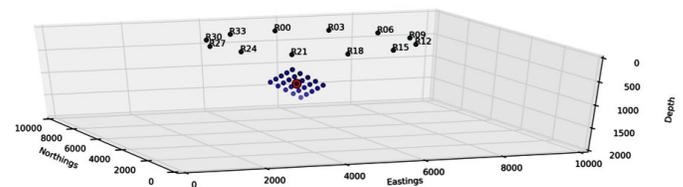


Figure 2. Model Geometry. The event cloud is dipping away from you towards the NE, and the entire model domain is displayed.

We use the parallel fully elastic 3D finite difference Wave Propagation Program from Lawrence Livermore (WPP; e.g. Petersson and Sjögreen, 2011) for our wavefield simulator. The grid spacing is 20 meters in all three directions. We use a P speed of 2500 m/s, an S speed of 1500 m/s, and a mass density of 2600 kg/m³ throughout the entire volume. After trial and error, we use their predefined source-time-function called “Smoothwave” – a 7th order polynomial in time with reasonable spectral behavior – as producing solutions meeting the stability criteria of their simulator. (Triangular source-time-functions have too broad a spectrum for stable solutions.) WPP allows one to record seismograms from any position in the model including in the interior, so we can compare the seismograms estimated via equation 1 with those directly computed by WPP.

Model run-times for the 3-second seismograms we generated were on the order of 5 minutes per simulation on a 24 core (hyperthreaded) machine, and there were 25 simulations required (one per aftershock) to generate the full suite of seismograms.

Results

Earth Surface Array Only

Shown in Figure 3 are estimated seismograms formed via Equation 2 at location VR compared with “true” seismograms at VR (available from the WPP code). The only \mathbf{x}' contributions to the Equation 2 approximation come from the model “physical” array (the green “R” locations in Figure 2). All seismograms are displayed for 1 second’s duration even though three to four seconds were computed.

From experience using other source-time-functions, we infer that the high frequencies evident in some of the red seismograms are an indicator of incipient numerical instability in the WPP solution.

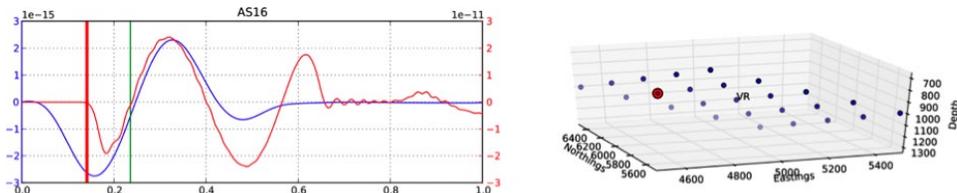


Figure 3. An example of Equation 2 in action. Plotted on the right is the location (in red) of an example aftershock being analysed as if being recorded at the location VR. Plotted on the left are two seismograms, on a horizontal axis in units of seconds, and a vertical axis in (nondimensional) volumetric strain. The seismogram in red is the actual “recorded” seismogram found by retrieving the results at that location from the simulation program WPP. The seismogram in blue is estimated for location via equation 1 from the surface seismograms recorded at the green “R” locations of Figure 2. Also shown is a red bar at the theoretical time of the P arrival (known from the simple velocity structure) as well as a green bar at the S arrival time. The amplitude mismatch between the two seismograms is attributable to the incorrect value of $K=1$ in Equation 2.

By examining the full suite of seismogram estimates, we infer that the first minimum in the Equation 1 estimated seismogram is a good “pick” for first arrival times – in a fashion not unlike zero-phase arrival times.

Full Surface Integral Approximation

The integral appearing in Equations 2 or 3 is meant to be a fully continuous surface integral over the entire bounding surface of the volume of interest. However, in real-world applications of the theory we are only able to approximate that integral via a discrete sum over our physical receivers. Such receiver geometries, because of obvious operational constraints for seismic surveys, are unlikely to be reasonable geometric approximations of the full surface integral.

To investigate how much of an improvement is available from a better surface integral approximation, we recorded seismograms via WPP from locations arrayed on a 1km grid over the bounding surface (Figure 4). The resulting Equation 2 estimate of the seismogram corresponding to that displayed in Figure 3 is shown in Figure 5. Encouragingly, the quality of the virtual receiver estimates formed via the Figure 2 array is quite comparable to the quality from those from the Figure 4 array. Once again we see that the (zero-phase style) deepest minimum of the estimate is a

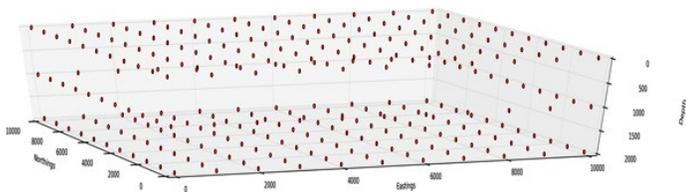


Figure 4. Summing over these locations forms a better approximation to the Equation 2 surface integral. Each receiver is on a 1km grid on all of the boundaries of the original volume displayed in Figure 2.

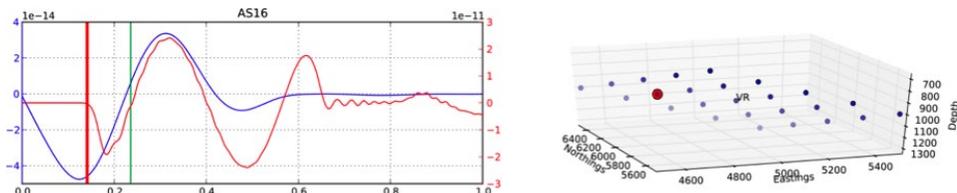


Figure 5. Everything is the same as in Figure 3 except that the blue seismogram is estimated from a better approximation of the full surface integral of Equation 2. Recording seismograms at discrete gridded locations shown in Figure 4 and summing the results form this better approximation.

reasonable “pick” for the first arrival, and indeed this is visually confirmed across all of the seismograms from all of the events on the modelled fault plane.

Stationary Phase Stacking

The surface integrals in the form of Equations 1 or 2 stack point-wise seismograms with a wide variety of phase angles (e.g. Schuster, 2009). From purely geometric considerations, those physical receivers a small angle away from the source-to-virtual-receiver ray should have phases that interfere constructively, while those physical receivers from a larger angle

away from that ray should interfere destructively. Obviously, this interference is wavelength dependent. Figure 6 shows the geometry.

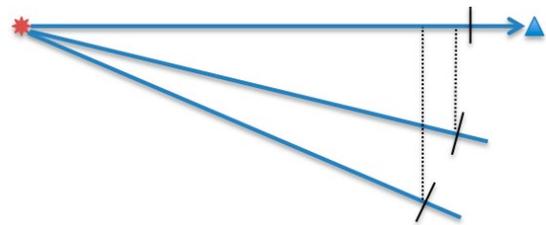


Figure 6. The red starburst is the microseismic source event. The blue triangle is the virtual receiver event connected to the source event by a blue ray-path. The thin black line is the leading edge of the propagating wave. The other blue lines shown below are ray-paths to physical receivers, with their corresponding leading edges under the action of Equations 1 or 2 shown by black lines. The dotted lines from those leading edges back to the source-VR ray-path show the projections of the physical receiver arrivals. During the stacking operation implicit in Equations 1 or 2, small angles away from the source-VR path interfere constructively, while larger angles interfere destructively. Such constructively interfering wave phenomena are known as “Stationary Phases” in asymptotics.

It is apparent from examination of the geometry in Figure 6 that the full surface integrals in Equations 1 or 2 are not necessary to construct an asymptotic approximation of the waveform. Only those ray-paths “close” in some (wavelength dependent) angular sense to the source-to-VR ray-path stack constructively. Accordingly, we now numerically investigate first-arrival-time error statistics as a function of angle from the source-to-VR ray-path.

Because we have available all of the seismograms recorded at the locations in Figure 4, we can compute Equation 2 stacks only including seismograms a specified angle from the source-VR ray-path. Figure 7 summarizes the resulting arrival time “picking error” statistics. The “picking error” is defined to be the arrival time of the deepest minimum in the Equation 2 stack minus the theoretical arrival time computed from the known distances and velocity for each of the blue microseismic events.

The y axis is arrival time error in seconds. The x axis is the angle in degrees from the source-VR rays up to which we included seismograms from the physical receiver locations of Figure 4. Angles

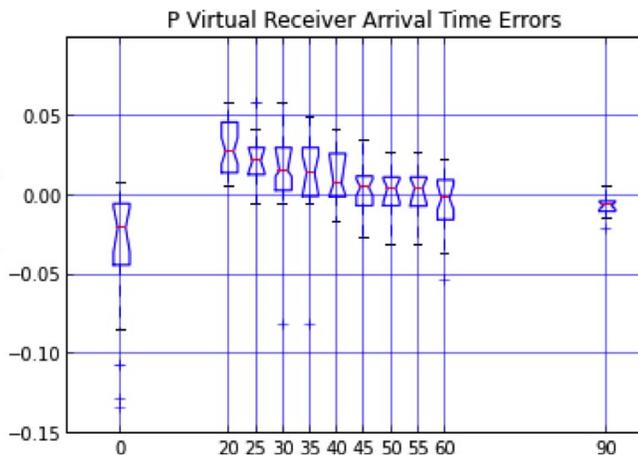


Figure 7. P arrival time error statistics for VR seismograms estimated according to Equation 2.

displayed as 0 and 90 degrees are special cases, corresponding to the “R” locations of Figure 2 and the full surface integral respectively. Box and whisker plots summarize the arrival time errors for each case. The red central line in each box is the median value of the arrival time errors collected from all non-VR sources in Figure 2. The width of the notches on each box are 95% confidence intervals for those median arrival time errors, formed via a bootstrap procedure of 10,000 samples-with-replacements. The ends of the boxes show the limits of the inner (2nd and 3rd) quartiles. The whiskers shown delimit 1.5 times the inner quartile range, while the flyers are plotted as crosses outside the whiskers.

There are several things noteworthy in Figure 7: Firstly, the “R” location array arrival time errors have a median of about 20 msec., which corresponds to a median raw location error of about 50 meters from that geometry using this technique with a single virtual receiver. Obviously, incorporating multiple virtual receivers into a relocation effort could improve that raw error by averaging out the arrival time error fluctuations. Secondly, one explanation for the trend observed in the median results from 20 to 60 degrees is that more physical receivers give better counting statistics in our Equation 2 stack. That is consistent with the tight, low-median errors found for the full surface integral result (plotted at 90 degrees in the figure). Thirdly, the 20 to 60 degree median errors are all positive, implying that the stack and this “deepest minimum” picking strategy appears to produce events that arrive later than they should. We currently have no explanation for this observation. Fourthly, there are some very bad outliers evident in our method, with arrival time errors on the order of 150 msec. or so. We have no explanation for this observation either. We speculate that there are potentially some numerical instabilities still found in some of the solutions which contaminate our statistics.

Future Work

We currently expect to perform the following tasks in the near-term future.

- Integrate one of the HypoDD (Waldhauser and Ellsworth, 2000), the TomoDD (Zhang and Thurber, 2003), and/or the HypoCC (Foulger and Julian, 2013) relocation codes into our efforts. This will enable us to actually see the effects of placing virtual receivers into the relocation algorithms.
- Put in a realistic velocity structure (perhaps along with topography) and investigate its effects on the estimated waveforms.
- Because of the “clean” nature of our numerical experiments, we haven’t yet explored any of the effects of filtering, or any other signal processing techniques. This needs to be investigated.
- Investigate Equation 1 for estimating displacement seismograms rather than pressure seismograms since our target applications are land based, and we are unlikely to encounter pure explosive sources.

We hope to have accomplished the tasks higher up in the list by the time of the meeting, and will report on the new results at the meeting.

We also intend to deploy these techniques against real-world data – either in an EGS Geothermal situation, an earthquake after-shock scenario, and/or a shale gas hydraulic fracturing scenario. We are actively seeking partners for this activity. The EGS and shale gas applications of microseismic monitoring are heavily influenced by the quality of the hypocentric relocations, and are the major driving force behind our push towards deploying seismic interferometry in these kinds of problems.

Acknowledgement

FGH thanks Klaus Regenauer-Lieb of the University of Western Australia for allowing access to his parallel cluster for the WPP computations performed during this work.

References

- Curtis, A., Nicolson, H., Halliday, D., Trampert, J., & Baptie, B. (2009). Virtual seismometers in the subsurface of the Earth from seismic interferometry. *Nature Geoscience*, 2 (10), 700-704. <http://dx.doi.org/10.1038/ngeo615>.
- Foulger, G. R., & Julian, B. R. (2013). Seismological software for geothermal monitoring. In *PROCEEDINGS, Thirty-Eighth Workshop on Geothermal Reservoir Engineering*, SGP-TR-198. Stanford University. <https://pangea.stanford.edu/ERE/pdf/IGAstandard/SGW/2013/Foulger.pdf>.
- Petersson, N. A., & Sjögreen, B. (2011). User’s guide to WPP version 2.1.5. Lawrence Livermore National Laboratory. <https://computation.llnl.gov/casc/serpentine/download/WPP-2.1-UsersGuide.pdf>.
- Schuster, G. T. (2009). *Seismic Interferometry*. Cambridge University Press, first ed.
- Waldhauser, F., & Ellsworth, W. L. (2000). A double-difference earthquake location algorithm: Method and application to the Northern Hayward Fault, California. *Bulletin of the Seismological Society of America*, 90 (6), 1353-1368. <http://dx.doi.org/10.1785/0120000006>.
- Zhang, H., & Thurber, C. H. (2003). Double-difference tomography: The method and its application to the Hayward Fault, California. *Bulletin of the Seismological Society of America*, 93 (5), 1875-1889. <http://dx.doi.org/10.1785/0120020190>