Continuous hypocenter and source mechanism inversion via a Green's function-based matching pursuit algorithm

Ismael Vera Rodriguez, Mauricio D. Sacchi, and Yu Jeffrey Gu, University of Alberta

The hydraulic fracturing of rock formations bearing oil and gas is a process that aims to improve well productivity. The breaking of rocks releases seismic energy that is recorded by stations usually located in nearby wells. By processing these recordings, the hypocenters of induced microseismic events are retrieved and interpreted to estimate the fractured volume. Several techniques have been developed to determine hypocenters; the most common are based in the inversion of P- and S-wave time picks from multiple stations (Pujol, 2004). There are also procedures that involve, for example, the back propagation in time of the recorded wavefields (Gajewski and Tessmer, 2005). Besides the hypocenter location of microseismic events, there is additional information embedded in the seismic recordings: the source mechanism, which is expressed via the seismic moment tensor (SMT). Not as common as the estimation of the hypocenter location, the SMT of induced microseismic events is also estimated from monitoring records (e.g., Nolen-Hoeksema and Ruff, 2001; Jechumtalova and Eisner, 2008). The SMT provides an additional source of information for understanding the fracturing process. A quasi real-time system for simultaneous event location and SMT inversion can reduce the processing and interpretation turnaround times. More notably, such a system could facilitate a realistic integration of microseismic geophysical data to significant decisions that are made during engineering operations such as fracturing.

This work discusses the development of an algorithm for a nearly real-time automatic monitoring system for microseismic events. The proposed methodology does not use rays; we prefer to use Green's functions computed by solving the elastic wave equation. The process is very demanding from the computational point of view. It offers the advantage, however, that the full multicomponent recorded wavefields are used for event location and SMT inversion. The process first focuses energy in a few nodes with potential sources via the solution of an imaging problem that resembles seismic migration. The location of the seismic events is enhanced by introducing constraints that force sparsity on a norm of the seismic moment tensor at a given node. This procedure is used to identify plausible seismic events that are later estimated one at a time via a matching procedure that uses a library of Green's functions. A quasi real-time implementation of the algorithm will require a careful investigation of techniques to compress Green's functions and fast matrix-vector multiplication schemes in compressed domains. Developments of the aforementioned directions are underway.

Representation of the problem

To start, we delimit the space where microseismic events are more likely to occur. The latter is represented via a physical grid where every node in the grid n is considered a potential source (Kawakatsu, 1998). With this assumption in mind, we represent the far-field displacement from every node source of the grid to the Green's functions for every node-receiver pair (Aki and Richards, 1980). We can express this mathematically as

$$\mathbf{u} = \mathbf{G}\mathbf{m}, \tag{1}$$

where $\mathbf{m}$ represents the six components of the SMT for all potential sources in the grid, $\mathbf{G}$ is the library of precomputed Green's functions for our area of study, and $\mathbf{u}$ is a vector containing the observed seismograms from the multicomponent receivers. The vector $\mathbf{m}$ contains $6 \times N$ elements where N is the number of grid points used to represent the subsurface. If we consider a window of analysis containing NT time samples and NG three-component receivers, then the data vector $\mathbf{u}$ contains $NT \times NG \times 3$ elements. This leads to a library of Green's functions of $(NT \times NG \times 3) \times (N \times 6)$ elements. It is clear that for standard 3D subsurface volumes, a nearly real-time inversion of Equation 1 will be required to sort out many computational challenges.

Estimation of hypocenters via focusing with a mixed-norm (l_{1-2}) regularization

We solve Equation 1 by minimizing a cost function

$$J = \|\mathbf{u} - \mathbf{G}\mathbf{m}\|_2^2 + \alpha \|\mathbf{m}\|_{1-2} \tag{2}$$

that minimizes two terms. The first term is the misfit between our observations and the model; the second term is a regularization functional (l_{1-2} norm) used to focus the solution to a few isolated grid points. During this stage of the algorithm, we are mainly concerned with hypocenter determination. We realize that the estimate of the SMT obtained by minimizing Equation 2 might not be optimal. It is important to point out, however, that the l_{1-2} norm of the SMT at a given node can be easily determined with this algorithm and used as an attribute that highlights the presence of a source. The l_{1-2} norm is a mixed-norm, which is often adopted in the context of vector-valued signal and imaging processing (Kowalski and Torresani, 2009). In our case, the mixed-norm $l_{1-2} = \|\mathbf{m}\|_{1-2}$ is defined as the l_1 norm of the l_2 norm of the SMT at each grid point. Minimization of this norm leads to solutions that promote the development of sparse amplitude maps that serve to identify the microseismic sources. The minimizer of the cost function is given by

$$\mathbf{m} = [\mathbf{G}^T\mathbf{G} + \alpha\mathbf{D}(\mathbf{m})]^{-1}\mathbf{G}^T\mathbf{u}, \tag{3}$$

where $\mathbf{D}(\mathbf{m})$ is a diagonal matrix with entries proportional to the inverse of amplitude of the SMT at a given node. The latter can be interpreted as an amplitude-dependent damp-

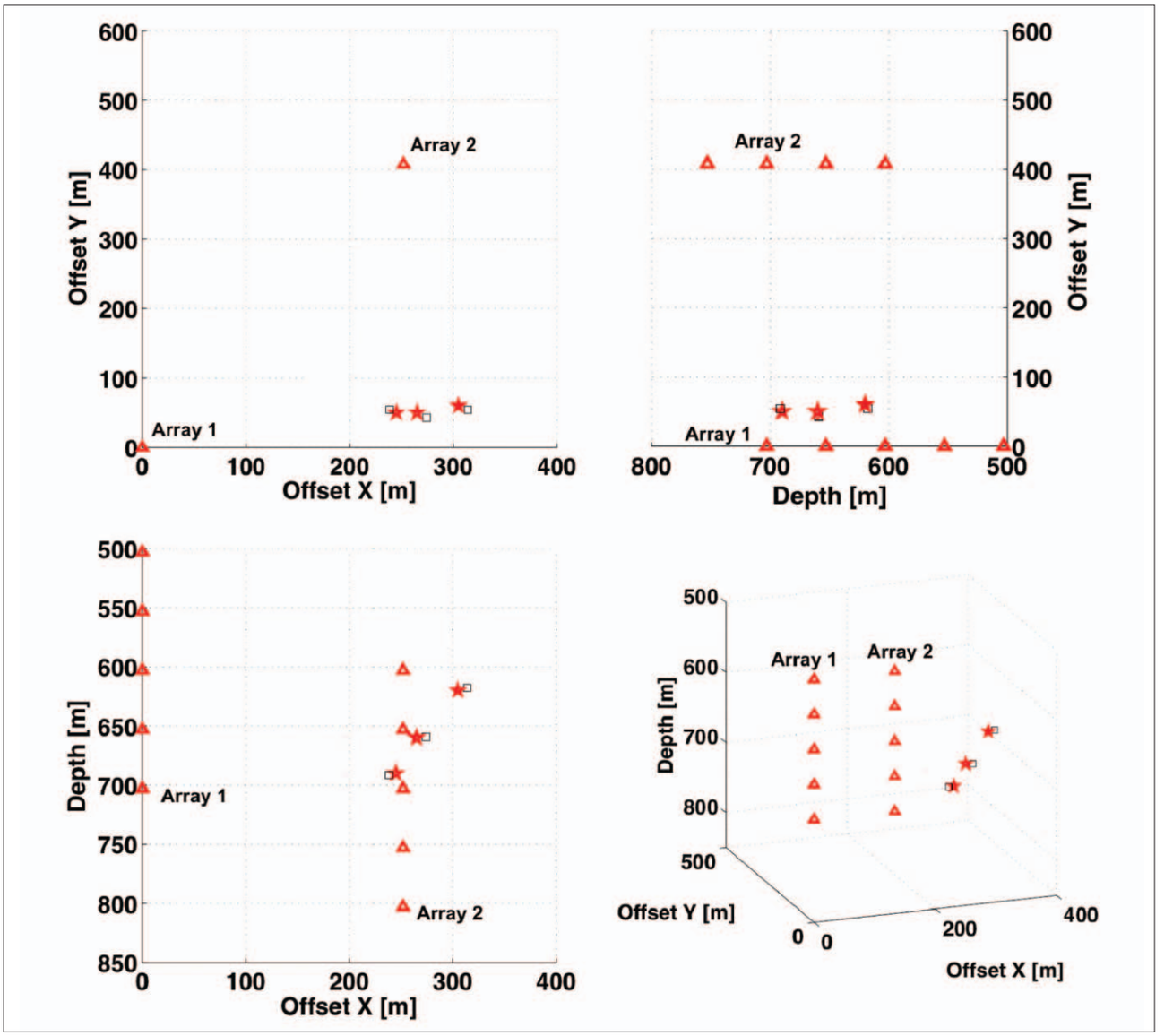

Figure 1. *Inversion results for three seismic events with closely spaced origin times, arbitrary hypocenter locations, and seismic moment tensors. Red triangles denote receiver locations. Red stars are the inverted source locations and black squares are the true source locations.*

ing term that enhances focusing of the most energetic nodes, which are clearly associated to microseismic sources. The resulting algorithm (Equation 3) resembles the solver utilized by the high-resolution Radon transform (Sacchi and Ulrych, 1995). It is important to stress, however, that rather than promoting the sparsity of a scalar quantity at each grid point, the current implementation is promoting the sparsity of the amplitude of the SMT at a given node. In other words, the sparsity of the six elements of the seismic moment tensor as a group is promoted and not the sparsity of each individual element of the SMT. A similar approach to focusing was proposed by Wang and Sacchi (2007), where they showed how a hybrid norm could be used to impose sparsity and continuity in images estimated via least-squares, wave-equation migration.

In order to circumvent the computational problem associated to the inversion of large matrices, the kernel $\mathbf{G}^T\mathbf{G}$ is replaced by its diagonal. With this approximation, we have an iterative algorithm to obtain the map of potential hypocenters. The term $\mathbf{G}^T\mathbf{u}$ is the well-known back-projection operator that, in this case, is implemented by means of the cross-correlation of the library of Green's functions and the data (Kawakatsu and Mongtaner, 2008; Leaney, 2008). Our algorithm gives a sharp image of the location and origin time of the sources in the grid. One needs to realize, however, that the estimates of the SMT given by our solution are not accurate. The diagonal operator can be computed in a preprocessing stage and saved in memory. Similarly, the operator $\mathbf{G}^T$ can also be saved and applied to new data streams entering the algorithm (Tsuruoka et al., 2009). An advantage of our formulation is that the full displacement field (seismograms) is used for the inversion. In addition, one can show

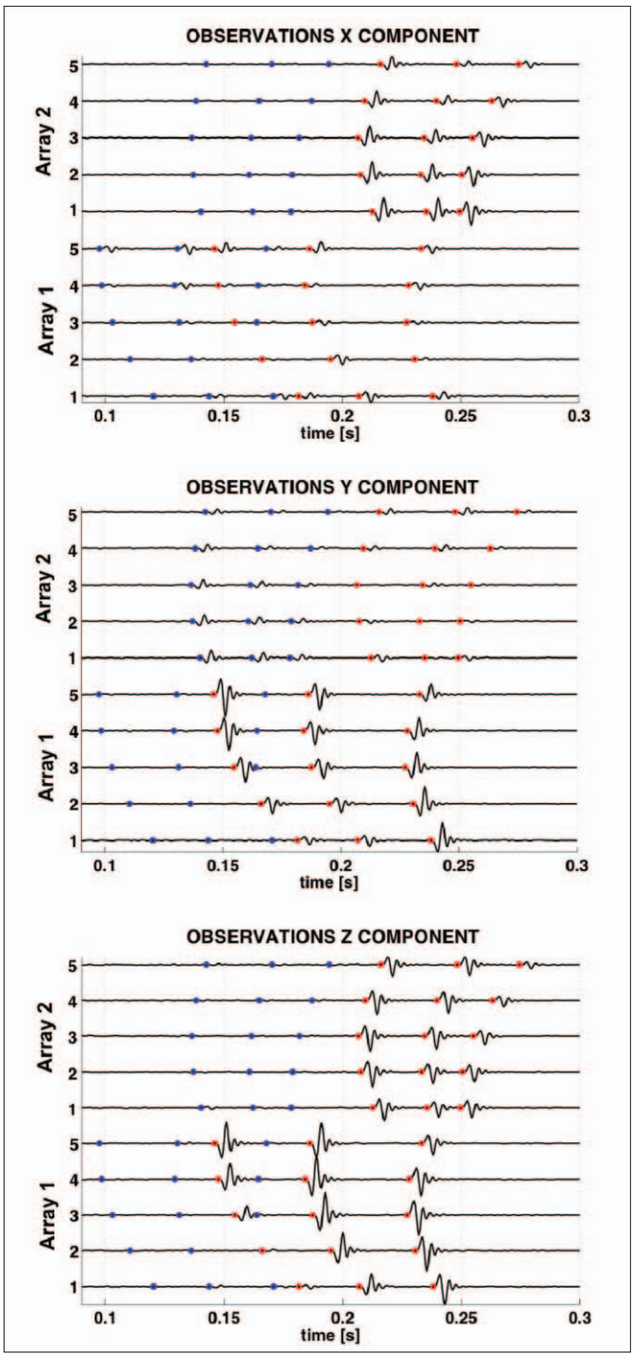

Figure 2. *Synthetic observations for the three sources with arbitrary seismic moment tensors and hypocenter locations. The signal-to-noise ratio (SNR) is 70%, computed with respect to the highest amplitude present in all traces. The additive random noise is Gaussian and uncorrelated. First break traveltimes for P- (blue) and S- (red) waves are shown as a reference.*

that the proposed algorithm can estimate hypocenters and origin times using a single station (Fan and Wallace, 1991). In this article, we have not addressed the problem of updating the velocity model. It is clear that the latter will require a continuous updating of the Green's function library, and we believe that this is not feasible in the context of our current algorithm.

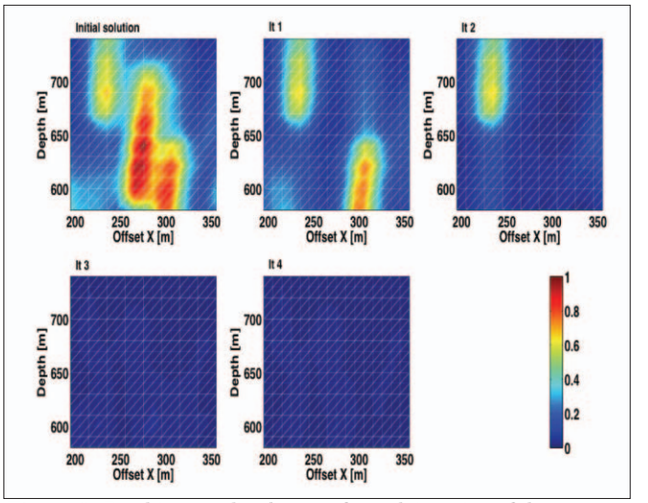

Figure 3. *Resulting amplitude maps for each iteration of the algorithm described in the article. The maps are summed over time (the 140 time samples of the processing window) and over the grid nodes in the Y direction. Notice how the amplitude anomaly in the middle of the initial solution map shows at least three apparent maxima; however, all of them are due to the same source and disappear in the first iteration output map.*

Estimation of the SMT

Once we have identified and timed the sources with the foregoing algorithm, individual inversions are carried out for each identified source. We have found that the inversion of the SMT for a given event can be accomplished in an efficient way via a matching pursuit algorithm (Tropp and Gilbert, 2007). In this case, a least-squares estimate of the SMT for the most energetic event is obtained. The resulting estimator of the SMT is used to generate synthetic data that are subtracted from the observations. The latter leads to residual observations that are used to estimate the second most energetic source and so on until exhausting the number of identified sources.

Synthetic experiments

The proposed methodology is tested using a homogenous isotropic half space and two vertical arrays of five receivers each. Three sources occurring at similar origin times, with arbitrary seismic moment tensors and at randomly chosen hypocenters are shown in Figure 1. For simplicity, we used a Hanning window as the source-time function. Random noise with Gaussian distribution has been added to the observations and a time window of 140 samples is analyzed (Figure 2). The signal-to-noise ratio is 70% (maximum amplitude of the noiseless data/standard error of the noise). The source hypocenters do not coincide with any one of the grid nodes. The frequency content of the observations was reduced to 100 Hz to decrease the computational grid. The mixed-norm regularization algorithm is used to identify potential hypocenters. This is shown in Figure 3 (first panel). The algorithm identifies one source every time it runs, and then the contribution of this source to the observations is subtracted before starting the next iteration. In this test, we allowed the algorithm to look for a maximum of five sources

SMT element	Source 1			Source 2			Source 2		
	True SMT	Estimated SMT		True SMT	Estimated SMT		True SMT	Estimated SMT	
M11	0.62	0.58		0.32	0.11		0.10	0.08	
M12	0.83	0.87		−1.22	−1.21		1.03	1.00	
M13	−1.39	−1.30		1.58	1.54		−0.93	−0.91	
M22	1.47	1.10		0.67	0.61		1.69	1.79	
M23	−1.14	−1.07		−0.63	−0.62		−1.26	−1.28	
M33	0.26	0.03		1.27	1.09		1.73	1.56	
	Source hypocenter			Source hypocenter			Source hypocenter		
	X	Y	Z	X	Y	Z	X	Y	Z
	274.13	43.00	659.28	313.88	54.35	617.62	238.39	54.46	691.47
	Orig. time error = 4.0 ms			Orig. time error = 4.0 ms			Orig. time error = 2.5 ms		
	Location error = 11.5 m			Location error = 10.8 m			Location error = 8.1 m		

Table 1. *Inversion results for the three arbitrary sources appearing in the same processing window. The source coordinates correspond to the true hypocenters.*

at each processing window. It is easy to see that, after the third iteration, only noise residuals are present in the output (Figure 3). The average origin time uncertainty in the results is about 4 ms. This is in part due to the fact that the hypocenter locations do not coincide with the grid nodes. Nonetheless, the algorithm identified the closest nodes to the true hypocenters. The error in location is approximately 11 m. The latter is very close to the grid spacing (10 m). On the other hand, the resulting estimators of the SMT coincide quite well to the true ones. The information for this simulation is summarized in Table 1.

Conclusions

The methodology presented in this article highlights the initial steps toward the development of an automatic algorithm to estimate the origin time, hypocenter, and SMT of microseismic events. Several directions are currently under consideration to improve the efficacy of the algorithm. In particular, the problem associated with storing and applying Green's functions in a fast and efficient way is key for the development of realistic real-time schemes for microseismic monitoring. Compression methods that preserve waveform integrity and fast algorithms for matrix-time-vector multiplication in compressed spaces are also part of our ongoing research efforts.

It is important to stress that the proposed algorithm does not use ray-theory-based Green's functions. The forward model operator proposed in this article uses isotropic elastic Green's functions with the possibility of including anisotropic and anelastic effects. In other words, the structure of the algorithm for hypocenter determination and SMT inversion does not change since the model for the desired media is controlled by the Green's functions. The inclusion of the mixed-norm focusing metric leads to a resolution enhanced back-projection algorithm that was used as starting point for one source at the time inversion by means of the matching pursuit algorithm. TLE

References

Aki, K., and P. Richards, 1980, *Quantitative Seismology*, Freeman.

Fan, G., and T. Wallace, 1991, The determination of source parameters for small earthquakes from a single, very broadband seismic station *Geophysical Research Letters*, 18, 1385–1388.

Gajewski, D. and E. Tessmer, 2005, Reverse modeling for seismic event characterization, *Geophysical Journal International*, 163, 276–284.

Jechumtalova, Z. and L. Eisner, 2008, Seismic source mechanism inversion from a linear array of receivers reveals non-double-couple seismic events induced by hydraulic fracturing in sedimentary formation, *Tectonophysics*, 460, 124–133.

Kawakatsu, H., 1998, On the real time monitoring of the long-period seismic wavefield, *Bulletin of Earth Research International*, 73, 267–274.

Kawakatsu, H. and J. Montagner, 2008, Time-reversal seismic-source imaging and moment-tensor inversion, *Geophysical Journal International*, 175, 686–688.

Kowalski, M. and B. Torresani, 2009, Sparsity and persistence: mixed norms provide simple signals models with independent coefficients, *Signal, Image and Video Processing*, 3, 251–264.

Leaney, W. S., 2008, Microseismic inversion by least-squares time reversal and waveform fitting, SEG *Expanded Abstracts*, 27, 1347–1351.

Nolen-Hoeksema, R. and L. Ruff, 2001, Moment tensor inversion of microseisms from the B-sand propped hydrofracture, M-site, Colorado, *Tectonophysics*, 336, 163–181.

Pujol, J., 2004, Earthquake location tutorial: graphical approach and approximate epicentral location techniques, *Seismological Research Letters*, 75, 63–74.

Sacchi, M. D and T. J. Ulrych, 1995, Improving resolution of Radon operators using a model re-weighted least-squares procedure, *Journal of Seismic Exploration*, 4, 315–328.

Tropp, J. and A. Gilbert, 2007, Signal recovery from random measurements via Orthogonal Matching Pursuit, *IEEE Transactions on Information Theory*, 53, 4655–4666.

Tsuruoka, H., H. Kawakatsu, and T. Urabe, 2009, GRiD MT (grid-based real-time determination of moment tensors) monitoring the long-period seismic wavefield, *Physics of the Earth and Planetary Interiors*, 175, 8–16.

Wang J. and M. D. Sacchi, 2007, High-resolution wave equation AVP imaging with sparseness constraints, GEOPHYSICS, 72, 1, S11–S18.

Acknowledgments: We thank the sponsors of the Signal Analysis and Imaging Group (SAIG) at the University of Alberta for their support to develop this work. We thank Scott Leaney for sharing with us many thoughts about microseismic data processing and inversion. We also thank Shawn Maxwell for helping us to clarify the ideas presented in this article.

Corresponding author: veraodr@ualberta.ca