Detection and Location of Microseismic Events Using The Subspace Detector

by

Ibinabo Bestmann

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Geophysics

Department of Physics
University of Alberta

Abstract

Seismic events or earthquakes occur due to stress perturbations in the sub-surface and are related to the generation of new faults or reactivation of pre-existing faults. Completeness of an earthquake catalogue is critical for gaining a better understanding of sub-surface processes and this can be achieved through detection and location of seismic events. They are especially important in applications such as monitoring reservoir changes during hydraulic fracturing treatments, as it can aid in revealing the extent of fracture growth. However, there are challenges involved in both processes. The low magnitude of most seismic events and the presence of noise may reduce the accuracy of most event detection methods. Furthermore, most event detection algorithms produce initial estimates of arrival times which sometimes contain large errors, degrading the accuracy of earthquake location procedures. The uncertainty associated with the estimated hypocenter locations can sometimes be larger than the seismic source dimensions as a result, restricting the resolution of the seismicity image.

In this work, a statistical approach known as the subspace detector is in-

vestigated and used to detect weak and variable events embedded in noise. It involves the construction of a vector space comprised of signals to be detected from seismic sources of interest. Waveforms are grouped together based on similarity to increase sensitivity of the detector to events of a particular seismic source, and waveform alignment is applied to reduce waveform variability in the subspace representation. The vector representation of the signals representing the seismic sources is shown to contain hypocenter information on the signals represented and have improved signal-to-noise ratios relative to the templates forming the subspace. Detected events are relocated relative to these vector representations, and further improvements in location accuracy can be achieved by correcting arrival time inconsistencies and reducing variation in back-azimuths via cross correlation.

The detection capabilities of the subspace detector are compared to the conventional matched filter and short-time average/long-time average (STA/LTA) detectors via tests on synthetic and real data examples. The results show that the subspace detector produces more detections than the matched filter at a reasonable false alarm rate. It can also be made general to accommodate a variety of waveforms and offers more detections relative to the STA/LTA detector with fewer false alarms. Event locations obtained by relocation relative to the vector representations are also compared to those obtained by a standard relocation technique in the form of the double difference algorithm. It is found

that the relative locations obtained using both the subspace detector and the double difference algorithm are fairly similar. There is also a significant reduction in the spatial extent of event locations after relocation with the subspace detector compared to routine absolute location techniques.

Acknowledgements

I would like to thank Dr. Mirko van der Baan for his tutelage during the course of my research. His helpful comments and guidance have improved my critical thinking ability and encouraged me to ask the right questions, which were integral to my research. I am forever grateful for the opportunity he provided me. It has been a privilege to work with my peers within the Microseismic Industry Consortium and I would like to thank them for all their helpful comments over the years. Finally I would like to thank my friends and family for encouraging and believing in me throughout the course of my research. I only came this far due to their support.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background

Seismic monitoring describes the passive recording of acoustic emissions known as earthquakes or seismic events which occur due to stress changes in a medium (Ry et al., 2017; De Meersman et al., 2009). A seismic event is normally generated at an unknown origin time and spatial coordinates known as the hypocenter, and can either be naturally occurring or artificially induced (Ry et al., 2017; Eisner et al., 2010). The elastic waves generated are usually the result of geomechanical fracture deformations, and can be digitally recorded by instruments called geophones. These devices convert ground motion into electric impulses that are proportional to its displacement, velocity or acceleration. The resulting ground motion is represented as a series of numbers which comprise what is known as a seismogram (Havskov and Alguacil, 2004). Seismic monitoring has a wide variety of applications, such as in the classification of mining explosions, screening for nuclear tests, hazard assessment of earth structure and reservoir monitoring during hydraulic fracturing treatments (De Meersman et al., 2009; Harris, 2006). The main objective in monitoring is to record the occurrence of earthquakes and obtain spatial coordinates of its hypocenter and/or its origin time in order to obtain geological information about subsurface structure and stratigraphy.

1.2 Event Detection

Usually, numerous seismic events tend to occur annually (natural and induced) and require considerable human investment if processed manually (Withers et al., 1998). To improve processing efficiency, computer algorithms are the standard means for robust detection of these events in the presence of noise (Withers et al., 1998). The earliest attempts to automatically detect seismic events used time domain methods. Some of these methods assumed that the signals to be detected were Gaussian, superimposed on Gaussian noise (Freiberger, 1963). They worked on the principle that signal differed from the noise in terms of characteristics such as energy and character. Methods such as the short time average over long time average (STA/LTA) (Vanderkulk et al., 1965; Trnkoczy, 1999), envelope detector (Allen, 1978) and the Z-detector (Swindell and Snell, 1977) were developed to this effect. However, for low magnitude signals or signals buried in noise, these methods are typically not as effective since they cannot differentiate between signal and noise. Nevertheless, these techniques remain the modern standard for event detection since they require very little in terms of parameter selection and are easy to set up. In fact, the STA/LTA detector is the primary detector of choice in most seismological monitoring networks.

Further improvements were made in event detection via the introduction of frequency domain methods (Withers et al., 1996). These techniques involved the calculation of power spectral density estimates of windowed data under the assumption that noise and signal both differed in frequency content. These methods therefore include some information on the signal to be detected to improve performance and tend to better highlight weak signals over a broad range of frequencies. Unfortunately, performance decreases if the spectral characteristics of the signals to be detected are similar to noise or for very narrow bandwidth signals. Like the time domain methods, these detectors also cannot differentiate between seismic events and some incoherent signals, and in some cases incoherent noise (spikes, bursts) (Vaezi and Van der Baan, 2015). Like the time domain methods, they also require some study of the target area to determine optimal processing parameters.

Nevertheless, the success of the frequency domain methods allowed for the use of other detection techniques which targeted signal properties to boost detection probability. For three component data, the direction of signal polarization can be obtained from the eigenvectors of its covariance matrix (Nguyen et al., 1989; Jurkevics, 1988; Montalbetti and Kanasewich, 1970). The direction of polarization can be used to define a filter to attenuate noise and boost signal information based on the direction of phase polarization. Noise may be polarized, but the direction tends to be random allowing it to be attenuated. A shortcoming of this method is its strict application to multi-component data, and the presence of noise tends to degrade the estimates of signal polarization, making this method ineffective in noisy areas (Aster et al., 1990; Montalbetti and Kanasewich, 1970).

Recently, template based detectors have been growing in popularity due to their ability to detect weak signals at fairly low detection thresholds (Gibbons and Ringdal, 2006; Harris, 2006). The matched filter technique involves the cross-correlation of a waveform template with a time series to find instances where there is a high degree of similarity to the template. Complex source mechanisms can sometimes produce varying waveforms, and this variation may not be captured by a single template. To account for this, the subspace detector expands the single template idea of the matched filter to multiple templates (Scharf and Friedlander, 1994; Harris, 2006). It utilizes a singular value decomposition on a set of templates assumed to characterize a source region to find an orthonormal representation that captures the common characteristics of the waveforms (Barrett and Beroza, 2014). This allows it to better handle waveform variation compared to the single template of the matched filter. The use of a cross-correlation approach allows for a greater probability of signal detection at a very low false alarm rate even for highly noisy conditions, but only applies to very similar sources or repeating earthquakes in the case of the matched filter. The performance of both template based methods are constrained by the availability and quality of template waveforms.

1.3 Event Location

Accurate hypocenter locations of the detected events are vital to better understand subsurface processes and reveal fracture networks and zones of mechanical instability (De Meersman et al., 2009). Arrival times of P- and/or S-waves (seismic phases) at each geophone are usually used to estimate a hypocenter location and origin time. Additional constraints such as the wavefront propagation direction obtained from three-component sensors can also be included in the location process, especially for restricted receiver distributions (Eisner et al., 2010). The accuracy of the arrival time picks have direct consequences on the accuracy of a determined hypocenter, as errors distort the image of seismicity.

Traditionally, manual picking of seismic phases has been the most reliable method of phase arrival time picking, but the ever-increasing amount of data available in recent times makes this task very difficult. Automated routines have been developed to handle this task, a majority of which are based on similar principles used in event detection. Algorithms like the short time average over long time average ratio (STA/LTA), modified Baer and Kradolfer's method and modified energy ratio (MER) utilize the difference in energy between signal and background noise to define the onset time or first break (Allen, 1978; Baer and Kradolfer, 1987; Earle and Shearer, 1994; Han, 2010). These methods tend to provide decent estimates of phase onset times but suffer from the presence of noise and the use of less than optimal parameters. The progression from energy based methods to techniques specific to signal properties was evident with the advent of cross-correlation based arrival time picking (Gibbons and Ringdal, 2006; De Meersman et al., 2009). Highly similar events tend to have onset times at the same position within the phase coda (Castellanos and van der Baan, 2013). This similarity can be used to provide high precision relative arrival time estimates.

Other approaches include autoregressive techniques which were developed on the assumption that a seismogram can be divided into segments, where the intervals before and after the onset of a signal are two different station-

ary processes (Sleeman and Van Eck, 1999; Leonard and Kennett, 1999). One of the segments is assumed to contain noise while the other has the signal of interest. An autoregressive model can be used to describe a representation of each segment, allowing for the filtering out of noise based on an autoregressive representation of the signal and noise segments, and the estimation of the phase onset position. However, when the signal-to-noise ratio (S/N) is low and arrivals are not evident, performance is poor. This method works best when a prior estimate of the signal window is known. The idea of predicting signal information via modeling was taken further, resulting in the use of neural networks (Gentili and Michelini, 2006). It is slowly becoming a means of phase onset determination via signal classification in data-driven learning schemes. Performance is constrained by the quality and quantity of data provided. In theory, it is most likely not possible to account for every signal type and noise contribution and is therefore not widely used. Nevertheless, it is currently an area of active research.

Travel time is a non-linear function of the model parameters, making it difficult to solve the event location problem easily using analytical methods. The simplest means to obtaining an optimal location and origin time for a seismic event involves a grid search. Travel times are computed from all possible hypocenters and origin times within a grid to receivers, and the point with the best agreement between the observed and predicted values is regarded as the highest probability hypocenter of the seismic event. The quality of the predicted times is dependent on the knowledge of the velocity structure of the area. Ideally, three-dimensional (3D) velocity models provide more accurate travel times but can be expensive to utilize in routine processing or not available. In most cases, a one-dimensional (1D) velocity is employed for estimating event locations. Several attempts have been made to account for the limits imposed by a 1D model. Velocity variations were partially accounted by inverting for station and/or source terms and corrections in the location procedure (Shearer, 1997; Douglas, 1967), or by simultaneously inverting the travel time data for both hypocenters and velocity structure (Crosson, 1976; Ellsworth, 1978), but these methods contributed additional computational expenses and did not al-

ways yield suitable results.

To minimize error due to an incomplete knowledge of the velocity structure, highly similar events were utilized to locate events in a relative sense. For co-located events, the ray paths between the source region and common stations are very similar. The difference in travel time can thus be attributed to the spatial offset in the region between the events, and will not be affected by unmodeled velocity structure for the parts of the propagation path common to both events (Waldhauser and Ellsworth, 2000). A common relative location method is the master event approach in which events are relocated relative to a single well located master event (Frémont and Malone, 1987). A disadvantage of this technique lies in the possible error propagation through the entire cluster and restriction on the maximum spatial extension of the cluster (Waldhauser and Ellsworth, 2000). To make the relocation procedure more robust, the joint hypocenter determination approach (JHD) was developed to jointly determine the hypocenters, origin times and station corrections for all events to be located (Douglas, 1967; Frohlich, 1979; Herrmann et al., 1981; Got et al., 1994). Waldhauser and Ellsworth (2000) further extended this idea by making use of both relative travel times (cross-correlation) and catalog travel times without the use of station corrections. Their method involves minimizing the residual between observed and calculated differential travel time for event pairs instead of a single master event. This method has been shown to improve the seismic image by collapsing locations into sharper clusters, leading to clearer identification of faults. The system of equations can become quite large for numerous events, and usually requires a large number of well distributed stations to be effective.

1.4 Motivation and contributions

Robust event detection and location plays a major role in both the completeness of a seismic event catalog, and the quality of analysis into seismicity. Template based approaches to detection have been shown to greatly improve the probability of detection at very low false alarm rates. Subspace detection techniques

can better handle waveform variation present in a data set and therefore improve detection capabilities but it is not widely used.

In this study, subspace detection will be examined and implemented on microseismic data in an attempt to characterize seismicity. It represents a set of template signals as a linear combination of orthogonal basis waveforms and spans the gap between a correlation and energy detector. It has the advantage over the standard waveform correlation of greater tolerance for source-location variations. I implement this method to characterize seismicity from a hydraulic fracturing treatment in western Canada monitored via a single vertical downhole array. The performance of the detector is compared to those of the conventional STA/LTA detector and the matched filter in terms of the number of detections, false alarms and missed detections. In addition, an approach to event location is introduced in the form of a master event relocation using the subspace detector. The singular vectors of the subspace detector tend to have higher signal-to-noise ratios than their template events and are assumed to represent the average signal properties in the design set. We therefore propose to use these waveforms to obtain more accurate relative data measurements by treating them as master events in a relocation procedure. This method will be applied to the microseismic data set and the validity and accuracy of the relocations will be compared with locations obtained using conventional methods. Finally, the performance of the method will be discussed in an attempt to establish its potential benefits and shortcomings.

Chapter 2

Subspace detection

2.1 Introduction

Microseismic monitoring has been a widely used tool to observe reservoir changes in order to optimize hydraulic fracturing stimulations (Van Der Baan et al., 2013). Typically the induced seismicity forms an elongated cloud of small earthquakes known as microseismic events, and their source locations can reveal fractures and faults, which highlight directions of increased permeability and/or porosity (De Meersman et al., 2009; Rutledge and Phillips, 2003). Detection of these microseismic events is not always straightforward. In most cases, microseismic events tend to have low magnitudes (about -2 to 3) and are usually immersed in high amplitude noise, making signal identification difficult (Van Der Baan et al., 2013). Furthermore, the low signal-to-noise ratio (SNR) limits our ability to detect low-amplitude events (Song et al., 2014).

Several algorithms have been developed to tackle this issue. Time domain techniques such as the short-time average over long-time average (STA/LTA) detector (Frémont and Malone, 1987; Vanderkulk et al., 1965) and analytic envelopes like the Hilbert transform (Earle and Shearer, 1994) are examples of commonly used techniques. They involve calculating the ratio of energy in two moving windows: a short time window which gives a measure of signal energy and a long time window which gives a measure of the background noise. These

methods suffer when phases are buried in noise due to lack of adequate signal information (Trnkoczy, 1999). Swindell and Snell (1977) developed a statistic computed by subtracting the mean from the data and then normalizing it by its standard deviation. It is advantageous because it automatically adjusts to the variance of noise, but fails to properly enhance secondary arrivals (Withers et al., 1998). Frequency domain methods such as the power spectral density technique have been shown to be superior at highlighting weak signals over a broad range of frequencies (Withers et al., 1998). Much like the time domain methods, short and long-time windows are utilized to estimate the power spectra. Vaezi and Van der Baan (2015) utilized the difference in strength of spectral content between events and background noise to find events. The average seismic background noise PSD is calculated and subtracted from the PSD estimate of the windowed data. These differences are then normalized by the standard deviations at each frequency, and segments which differ from the background noise PSD are considered events. This method breaks down when signals and noise have very similar spectra and noise is not stationary (Withers et al., 1998; Vaezi and Van der Baan, 2015).

To mitigate the effects of high amplitude noise, waveform cross-correlation, also known as the matched filter, is often employed in detection (Gibbons and Ringdal, 2006; Van Trees, 2004). A known signal is cross-correlated with successive time segments of incoming data; segments displaying a high degree of similarity to the template (master) waveform results in a high correlation coefficient. It exploits waveform similarity as a means to detect events in noisy conditions (low signal-to-noise ratio), or when there is waveform overlap (Barrett and Beroza, 2014). Occasionally repeating sources produce variable waveforms not well represented by a single template signal. This is usually due to the events being spread out over a region larger than one or two wavelengths of the dominant frequency of the repeating waveforms (Harris, 2006). This results in a poor performance of the matched filter when waveform variability is high. The subspace detector extends the matched filter concept from a single template to several templates. The template signals are modeled as a linear combination of orthogonal basis waveforms and it offers a trade-off between flexibility and

sensitivity in terms of the number of basis waveforms employed in the signal subspace; the larger the number of basis waveforms used, the closer the performance to an energy detector while a single waveform basis representation is analogous to using a single matched filter. Optimal results are usually obtained when a low-order basis representation is used to define the subspace (Jin and Friedlander, 2005; Kirsteins and Tufts, 1994; Harris, 2006).

In this chapter, several common detection techniques are introduced, and the detection problem is described in terms of a probability model for describing both noise and signal with noise, and the structure of the subspace detector is explained. The work flow for construction of the subspace detector will be outlined. Parameter selection will be discussed and analyzed and some synthetic examples will be used to demonstrate the functionality of this technique.

2.2 Event detection algorithms

2.2.1 Short time average over long time average

Most microseismic events are characterized by impulsive onsets and exponential envelopes with decreasing signal frequency over time (Lee et al., 1981). Seismic events and background noise usually differ in frequency content and character and can often be distinguished on a seismogram. This is taken advantage of by the short time average over long time average (STA/LTA) detector. The algorithm requires little to no signal information and involves the use of two windows: a short time window (STA window) which is sensitive to seismic events, and a long time window (LTA window) which gives some information on the temporal amplitudes of seismic noise (Trnkoczy, 1999). In each window, the absolute amplitude of each data sample of an incoming signal is calculated. The average of these values is calculated and the STA/LTA ratio is calculated as

$$STA(q) = \frac{1}{n_s} \sum_{i=q}^{q+ns-1} |s_i|^2, \tag{2.1a}$$

$$LTA(p) = \frac{1}{n_l} \sum_{i=p}^{p+nl-1} |s_i|^2, \tag{2.1b}$$

where $|s_i|$ represents the absolute value of the windowed data on the i^{th} time sample, n_s and n_l correspond to the number of samples within the STA and LTA time windows respectively. q and p are the starting positions of both the STA and LTA windows, with q being dependent on n_l. It can be implemented symmetrically (STA window centered within the LTA window) or asymmetrically (STA leading the LTA window with the LTA window ending right behind the start of the STA window). This determines the relative values of the STA and LTA starting values p and q respectively.

The STA time window is usually longer than a few periods of typical seismic signals, while the LTA time window is usually an integer multiple of the STA window, commonly 5-10 times the STA window length (Wong et al., 2009). A detection is declared when the ratio exceeds a threshold value across the seismic network

$$\frac{STA(q)}{LTA(p)} \geq \tau_{\text{SL}}, \tag{2.2}$$

where τ_{SL} is the threshold. Optimal parameter settings for this technique vary and do require some study by an analyst for the location of interest. In this work, a multiplexed form of the STA/LTA detector is used (Song et al., 2014).

Windowed data in N_c individual components can be analyzed independently or combined into single channel-sequential multiplexed form

$$\begin{aligned} \mathbf{x}[n] = \quad & [x_1(n)\ x_2(n)\ .....x_{N_C}(n)\ x_1(n+1)\ x_2(n+1) \\ & .....x_{N_C}(n+m-1)]^T, \end{aligned} \tag{2.3}$$

where $\mathbf{x}[n]$ is the Z $\times$ 1 time window for the n^{th} sample position from the continuous stream (Z = $N_C \times m$), $x_i[n]$ is the n^{th} time sample from the i^{th} channel of windowed data $\mathbf{x}_i$, and m represents the length of the i^{th} channel time window in samples. The STA/LTA detection statistic can then be computed on the data using the equation

$$r(n) = \frac{[\mathbf{x}_{STA}^T(n)\mathbf{x}_{STA}(n)]/n_s}{[\mathbf{x}_{LTA}^T(n)\mathbf{x}_{LTA}(n)]/n_l}, \tag{2.4}$$

where $\mathbf{x}_{STA}^T(n)$ and $\mathbf{x}_{LTA}^T(n)$ correspond to the multiplexed data in the STA and LTA windows respectively, beginning on the nth sample.

2.2.2 Matched filter

The matched filter operation can be described as an inner product between vector $w_{N,\triangle t}(t_0)$ containing N consecutive samples of a non-zero time series w(t), and vector $v_{N,\triangle t}(t_0)$, where t_0 is the time of the first sample and $\triangle$t is the temporal sampling interval (Gibbons and Ringdal, 2006)

$$\begin{aligned} \big\langle v(t_v), w(t_w) \big\rangle_{N,\triangle t} = & \ \big\langle v(t_v)_{N,\triangle t}, w(t_w)_{N,\triangle t} \big\rangle \\ & = \sum_{i=0}^{N-1} v(t_v + i\triangle t) w(t_w + i\triangle t). \end{aligned} \tag{2.5}$$

The fully normalized cross-correlation coefficient is given by

$$\begin{aligned} & CC[v(t_v), w(t_w)]_{N,\triangle t} = \\ & \\ & \frac{\big\langle v(t_v), w(t_w) \big\rangle_{N,\triangle t}}{\sqrt{\big\langle v(t_v), v(t_v) \big\rangle_{N,\triangle t} \big\langle w(t_w), w(t_w) \big\rangle_{N,\triangle t}}}. \end{aligned} \tag{2.6}$$

The template signal $w(t)$ may also be multiplexed and Eq. 2.6 applied to the multiplexed data stream to obtain a correlation coefficient at the nth sample position.

In terms of seismic data, if $w(t)$ represents recorded data and t_M represents the starting time of a data window containing a signal of interest (template/master), we can express the correlation coefficient as follows

$$CC_w(t)_{N,\triangle t} = CC[w(t), w(t_M)]_{N,\triangle t}, \tag{2.7}$$

where $w(t_M)$ corresponds to the template event with zero mean. The correla-

tion coefficient CC is a function of time that measures the similarity between a template waveform and the waveform beginning at time t. It lies in the range $[-1, 1]$ with the extreme value occurring when an exact copy of the template occurs (Gibbons and Ringdal, 2006). A detection is declared if the absolute correlation coefficient exceeds a threshold

$$|CC(t)| \geq \tau_{\mathrm{CC}}. \tag{2.8}$$

High signal-to-noise ratio templates characterizing the source to be detected are needed to apply the matched filter. These can be obtained using either an STA/LTA algorithm, visual inspection of the data stream by an analyst for suitable waveforms or from a pre-existing catalogue of events.

2.2.3 Subspace detector

The subspace detector as proposed by Scharf and Friedlander (1994) projects a sliding window of continuous data onto a vector subspace spanning a collection of source signals. The method implements a binary hypothesis test within each data window. It assumes the windowed signal contains only noise ν under the null hypothesis (H_o) or consists of a combination of signal $\mathbf{s}$ and noise ν under the alternate hypothesis (H_1). The noise ν is assumed to be zero-mean, as well as being temporally and spatially uncorrelated with unknown variance σ^2 (Harris, 2006). The signal is also assumed to be deterministic but dependent upon a vector of unknown parameters $\mathbf{a}$, and expressed as a linear combination of basis waveforms

$$\mathbf{s} = \mathbf{U}_d \mathbf{a}, \tag{2.9}$$

where $\mathbf{U}_d$ is an $m \times d$ column matrix containing the basis waveforms, with d corresponding to the number of basis waveforms, and m the number of observed time samples. The windowed data in N_c individual components can be analyzed independently or combined into a single channel-sequential multiplexed form, following the convention of eq. 2.3. The observed data when no signal is present

(H_0) is assumed to follow the probability density function $p(\mathbf{x})$ given by

$$p(\mathbf{x}|H_0) = [\frac{1}{2\pi\sigma^2}]^{m/2} \exp(-\frac{1}{2\sigma^2}\mathbf{x}^T\mathbf{x}). \tag{2.10}$$

The alternate hypothesis (H_1) follows the probability density function given by

$$\begin{aligned} p(\mathbf{x}|H_1) = [\frac{1}{2\pi\sigma^2}]^{m/2} \\ \times \exp(-\frac{1}{2\sigma^2}(\mathbf{x} - \mathbf{U}_d\mathbf{a})^T(\mathbf{x} - \mathbf{U}_d\mathbf{a})). \end{aligned} \tag{2.11}$$

In both Eqs. 2.10 & 2.11, σ^2 represents the unknown noise variance, $\mathbf{a}$ is a vector of unknown coefficients and m is the number of samples within the signal window. The presence of unknown parameters ($\mathbf{a}$, σ^2) in the probability density functions (eqs. 2.10 & 2.11) makes it convenient to utilize the maximum likelihood estimates with the available data. Hence, the Generalized Likelihood Ratio Test (GLRT) (Van Trees, 2004) is used to obtain the ratio test

$$\wedge(\mathbf{x}) = \frac{\max\limits_{\mathbf{a},\sigma} \; p(\mathbf{x}|H_1)}{\max\limits_{\sigma} \; p(\mathbf{x}|H_0)}, \tag{2.12}$$

where $\wedge$ represents the generalized likelihood ratio. Taking the natural logarithm of Eq. 2.12 eliminates the unknown variance σ^2. This leads to

$$\begin{aligned} l(\mathbf{x}) &= -\frac{m}{2}(\frac{\mathbf{x}^T\mathbf{x} - \mathbf{x}_p^T\mathbf{x}_p}{\mathbf{x}^T\mathbf{x}}) \\ &= -\frac{m}{2}(1 - c), \end{aligned} \tag{2.13}$$

where

$$\mathbf{x}_p = \mathbf{U}_d\mathbf{U}_d^T\mathbf{x}, \tag{2.14}$$

and

$$c(n) = \frac{\mathbf{x}_p^T\mathbf{x}_p}{\mathbf{x}^T\mathbf{x}}. \tag{2.15}$$

In eq. 2.13, l represents the log generalized ratio. The projected data $\mathbf{x}_p$ is a

least-squares estimate of the windowed data $\mathbf{x}$ obtained after projecting the data onto the subspace defined by matrix $\mathbf{U}_d$. c is a scalar value and corresponds to the subspace detection statistic. It is a positive quantity ranging between 0 to 1, and it gives a measure of the linear dependence between the windowed data $\mathbf{x}$ and the vectors representing the subspace. A detection is declared when c exceeds a threshold value γ, i.e. $c \geq \gamma$.

The matrix of basis waveforms $\mathbf{U}_d$ is obtained via singular value decomposition (SVD). The $m \times t$ matrix $\mathbf{S}$ containing the t template waveforms ordered in column form can be decomposed into the product of three unitary matrices

$$\mathbf{S} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T, \tag{2.16}$$

where $\mathbf{U}$ corresponds to an $m \times m$ matrix of the left singular vectors of $\mathbf{S}$, $\mathbf{\Sigma}$ is an $m \times t$ diagonal matrix of singular values, and $\mathbf{V}^T$ is a $t \times t$ matrix of right singular vectors (Wall et al., 2003). An orthonormal representation is obtained from the matrix of left singular values $\mathbf{U}$, with each column representing a basis waveform. A rank reduction can be applied to the unitary matrices by truncating the first d values of each matrix

$$\mathbf{S} = \begin{bmatrix} \mathbf{U}_d & \mathbf{U}_{t-d} \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_d & 0 \\ 0 & \mathbf{\Sigma}_{t-d} \end{bmatrix} \begin{bmatrix} \mathbf{V}_d^T \\ \mathbf{V}_{t-d}^T \end{bmatrix}. \tag{2.17}$$

In the above equation, $\mathbf{U}_d$ corresponds to the matrix of d left singular vectors which provides an orthogonal basis for the subspace representation, $\sum_{\boldsymbol{d}}$ corresponds to the truncated diagonal matrix of singular values and $\mathbf{V}_d$ represents the truncated unitary matrix of right singular vectors. $\sum_{\boldsymbol{d}} \mathbf{V}_d^T$ contains the coefficients for each of the first d largest singular values in $\sum_{\boldsymbol{d}}$ and provides an expression of the energy captured in the signal subspace $\mathbf{U}_d$ for the t events in the design set (Harris, 2006). It corresponds to matrix $\mathbf{A}_d$. The truncation accounts for the fact that a perfect representation of the signals to be detected is not possible or even desired with the chosen design set. Some noise exists and is assumed to be in the orthogonal complement to the subspace spanned by the representation $\mathbf{U}_{t-d}$. Omitting the $t - d$ smaller singular value approximates

the collection of t templates in a least-squares sense (Harris, 2006).

The fractional energy capture f_c for the i^{th} event in $\mathbf{U}_d$ can be described as

$$f_c^i = (\mathbf{a}_i^d)^T(\mathbf{a}_i^d), \tag{2.18}$$

where $\mathbf{a}_i^d$ corresponds to the i^{th} column of matrix $\mathbf{A}_d$. The fractional energy capture describes how well the individual signals are represented by the subspace representation, and ultimately the amount of linear dependence between the original data and projected data in eq 2.15. Of more interest is the average fraction of energy captured $\overline{f}_c^d$ from all t events as a function of the dimension of subspace representation d, which can thus be given as

$$\overline{f}_c^d = \frac{1}{d}\sum_{i=1}^{d} f_c^i = \frac{trace(\sum_d^T \sum_d)}{trace(\sum^T \sum)}. \tag{2.19}$$

Both the fractional energy (f_c^i) and the average fraction of energy captured ($\overline{f}_c^d$) from all t events may be plotted against the subspace dimension d, and they ranges from 0 to 1. This criterion is helpful in finding an optimal dimension of representation for the subspace. It is recommended to select a dimension of representation d that captures a sufficient amount of energy (e.g. $\geq 80\%$). The captured energy includes both signal and noise so an examination of the basis waveforms in $\mathbf{U}_d$ is recommended. Comparison of Eq. 2.16 and eq. 2.9 shows that the alternate hypothesis H_1 (eq 2.11) seeks to find the best fit between observed data $\mathbf{x}$ and possible combinations of the basis waveforms contained in the truncated left singular matrix $\mathbf{U}_d$. A high fit indicates the presence of a signal of interest is very likely which in turn produces a large similarity between the projected data $\mathbf{x}_p$ (eq. 2.14) and the observed data $\mathbf{x}$. Consequently, their energy ratio becomes close to 1 (eq. 2.15) yielding a large detection statistic c.

The detection statistic c (eq. 2.15) for the subspace detector is a ratio of sums of squares of normal random variables under the assumption of normally distributed background noise. Under this assumption, c follows a beta distribution (Harris, 2006). Normalization of both the denominator and numerator

by the noise variance σ^2 allows c to be modeled as an F distributed variable. It follows a doubly non-central F distribution when signal is present and a central F-distribution when only noise is present (Urkowitz, 1967). Therefore, theoretical detection threshold values γ can be estimated as a function of the dimension of representation d for a given false alarm rate (P_F). In a detection window of length N, the signal subspace will reside in an embedding space with degrees of freedom N under the assumption of uncorrelated noise within the data window. However, background seismic noise tends to be correlated, and the degrees of freedom in the data window decreases if data is bandpass filtered prior to detection (Wiechecki-Vergara et al., 2001). For a signal window of length N, the number of samples within the window for which noise is statistically uncorrelated across all components and the array can be estimated. This value $\hat{N}$ corresponds to the effective degrees of freedom of the signal embedding space and can be calculated by

$$\hat{N} = 1 + \frac{1}{\sigma^2_{\hat{c}}} \leq N, \tag{2.20}$$

where $\sigma^2_{\hat{c}}$ corresponds to the variance of the correlation coefficients between template events and noise (Harris, 2006; Song et al., 2014). For a given fixed false alarm rate P_F, the detection threshold γ for the detector can then be inverted for from the cumulative central F-distribution

$$P_F = 1 - F_{d,\hat{N}-d}(\frac{\gamma}{1-\gamma}\frac{\hat{N}-d}{d}), \tag{2.21}$$

where $F_{d,\hat{N}-d}$ refers to the F distribution with d and $\hat{N}-d$ degrees of freedom for the captured signal and signals assumed to have been missed by the subspace representation respectively.

This technique relies heavily on prior assumptions about the noise distribution. It is also particularly useful if short detection windows are used, since estimation variances for correlation coefficients are inversely proportional to the number of samples in a time series. Random sequences are increasingly likely to display coherent structure for a decreasing number of samples (Mendel, 1991; Oprsal and Eisner, 2014). However, noise is rarely statistically predictable. In

this study, the noise is assumed to be a white Gaussian process. Computed false alarm rates using eq. 2.21 are thus rarely honored in practice. A preferable approach to estimating the threshold value involves running the subspace detector on noise files and estimating the threshold γ from the upper limit of the histogram of detection statistics.

One dimensional subspace detector

If we consider the case where the dimensions of representation $d = 1$, Eq. 2.14 becomes

$$\mathbf{x}_p = \mathbf{u}_1 \mathbf{u}_1^T \mathbf{x}, \tag{2.22}$$

with $\mathbf{u}_1$ corresponding to a column vector with a single basis vector. The calculation involved in eq. 2.22 is essentially a dot product between the single basis vector $\mathbf{u}_1$ and the windowed data $\mathbf{x}$. Eq. 2.15 for a single subspace dimension can be written as

$$c = \frac{\mathbf{x}_p^T \mathbf{x}_p}{\mathbf{x}^T \mathbf{x}} = \frac{\mathbf{x}^T \mathbf{u}_1 \mathbf{u}_1^T \mathbf{u}_1 \mathbf{u}_1^T \mathbf{x}}{\mathbf{x}^T \mathbf{x}}. \tag{2.23}$$

since $\mathbf{u}_1$ is orthonormal, Eq. 2.23 can be expressed as

$$c = \frac{\mathbf{x}^T \mathbf{u}_1 \mathbf{u}_1^T \mathbf{x}}{\mathbf{x}^T \mathbf{x}}. \tag{2.24}$$

The fully normalized absolute cross-correlation coefficient between the template vector $\mathbf{u}_1$ and windowed data $\mathbf{x}$ is given by

$$CC[\mathbf{u}_1, \mathbf{x}] = \left| \frac{(\mathbf{u}_1^T \mathbf{x})}{\sqrt{(\mathbf{u}_1^T \mathbf{u}_1)(\mathbf{x}^T \mathbf{x})}} \right|, \tag{2.25}$$

Again, since $\mathbf{u}_1$ is orthonormal and $\mathbf{u}_1^T \mathbf{x} = \mathbf{x}^T \mathbf{u}_1$, Eq. 2.25 becomes

$$CC[\mathbf{u}, \mathbf{x}] = \left| \frac{(\mathbf{u}_1^T \mathbf{x})}{\sqrt{(\mathbf{x}^T \mathbf{x})}} \right| = \left| \sqrt{\frac{\mathbf{u}_1^T \mathbf{x} \mathbf{u}_1^T \mathbf{x}}{(\mathbf{x}^T \mathbf{x})}} \right| = \left| \sqrt{\frac{\mathbf{x}^T \mathbf{u}_1 \mathbf{u}_1^T \mathbf{x}}{(\mathbf{x}^T \mathbf{x})}} \right| \tag{2.26}$$

If we compare eq. 2.26 and eq. 2.24, we see the detection statistic c amounts to the square of the normalized cross-correlation coefficient $CC[\mathbf{u}_1, \mathbf{x}]$ (eq. 2.6). Thus the subspace detector becomes a correlation detector when a single dimension is used to define the subspace.

2.3 Design of the subspace detector

There are several steps involved in the construction and design of the subspace detector which will be elaborated on in the following subsections.

2.3.1 Obtaining template events

The subspace detector requires a suite of waveforms representing the source(s) of interest, and these could be obtained from a previously collected catalogue of detected events or by using any conventional detection technique such as the short time average over long time average (STA/LTA) detector. The subspace detector is sensitive to noise contained within the template waveforms so it is recommended to use high quality template events only.

2.3.2 Clustering template events

The advantage of subspace detection is that (1) it allows for waveform variability within the signals of interest and (2) rank reduction suppresses random noise, improving the signal-to-noise ratio of the basis waveforms contained in $\mathbf{U}_d$. Waveform variability can refer to different types of source signals (events from different locations) or diversity within a single source type (e.g. events from a similar location). The subspace detector can handle both situations but it is sometimes preferable to create expert detection systems that specialize in finding waveforms from a single source type only. These expert systems are advantageous since they allow for the use of small values of d. If the set of template events is selected automatically, without manual scrutiny of waveform similarity, it is recommended that a cluster analysis on the waveforms is performed first to group similar source signals together (Arrowsmith and Eisner, 2006; De Meersman et al., 2009). For multiple component data, the signals

can be concatenated into a single vector before clustering (eq. 2.3), or analyzed as individual traces. Given multiple channels of data, a weighted average cross-correlation can be used to group events into clusters based on similarity (Arrowsmith and Eisner, 2006). The maximum absolute amplitudes on each component can be used to weight the normalized cross-correlation functions. This helps reduce the effect of noise by down-weighting components with poor SNR.

2.3.3 Alignment of event clusters

It is strongly recommended that all events are aligned prior to applying a rank reduction. This reduces instantaneous waveform variability and leads to a smaller dimension of representation d and an overall improved performance. In this study, the iterative stacking procedure of De Meersman et al. (2009) is utilized to optimize alignment between the waveforms belonging to the cluster(s) of interest. The waveforms are first multiplexed into a single vector and stacked to get a reference waveform. This reference waveform is then correlated with the waveforms and optimal lag corrections are obtained from the maximum of the cross-correlation function. These lag corrections are applied to the template waveforms and a new reference waveform is obtained. The process is repeated till the lag correction converge to zero. The stack is the preferred choice for alignment as opposed to using a reference waveform due to the $\sqrt{n}$ increase in signal-to-noise ratio (SNR) associated with the stack compared to reference, where n represents the number of waveforms in the design set.

2.4 Information retrieval

The performances of the detectors can be described in terms of their precision, recall and accuracy. Precision is a measure of the exactness or quality of the detections made i.e. the percentage of the results which are relevant, while recall is a measure of completeness or quantity of the total number of detections made i.e. the percentage of total relevant events correctly identified by the detector (Powers, 2011). Precision and recall can be described with eqs 2.27a

and 2.27b:

$$precision = \frac{t_p}{t_p + f_p}, \tag{2.27a}$$

$$recall = \frac{t_p}{t_p + f_n}. \tag{2.27b}$$

In the above equations, t_p, f_p & f_n correspond to the number of true positives (seismic events), false positives (false alarms/type I error) and false negatives (missed events/type II error) respectively. A trade off usually exists between precision and recall since 100% recall may result in the generation of irrelevant detections (false alarms), while 100% precision may come at the risk of fewer detections. The goal in event detection is to maximum both precision and recall, and a metric that takes into account both terms is the F_1 score (Powers, 2011; Sasaki et al., 2007). This score corresponds to the harmonic mean of a detector's precision and recall and is denoted by

$$F_1 = (\frac{recall^{-1} + precision^{-1}}{2})^{-1} = 2 * (\frac{precision * recall}{precision + recall}) \tag{2.28}$$

The F_1 score is a scalar that ranges in value from 0 (worst accuracy) to 1 (perfect precision and recall) and can be used as a means to quantify overall performance.

2.5 Synthetic Examples

2.5.1 Waveform variation

To assess the effects of signal variation on template-based detection, subspace detection and matched filtering are applied to waveform data. Six three-component template signals with zero mean from a microseismic data set (Eaton et al., 2014) are selected for analysis and shown in figure 2.1. Their average signal length is about 0.4s, with an average correlation coefficient among the events of 0.50, highlighting the presence of some modest signal variation. The waveforms are first multiplexed and then aligned in time using an iterative stacking procedure. Finally, a 0.375s data window is extracted which encompasses both the P- and S-waves as shown in figure 2.2.

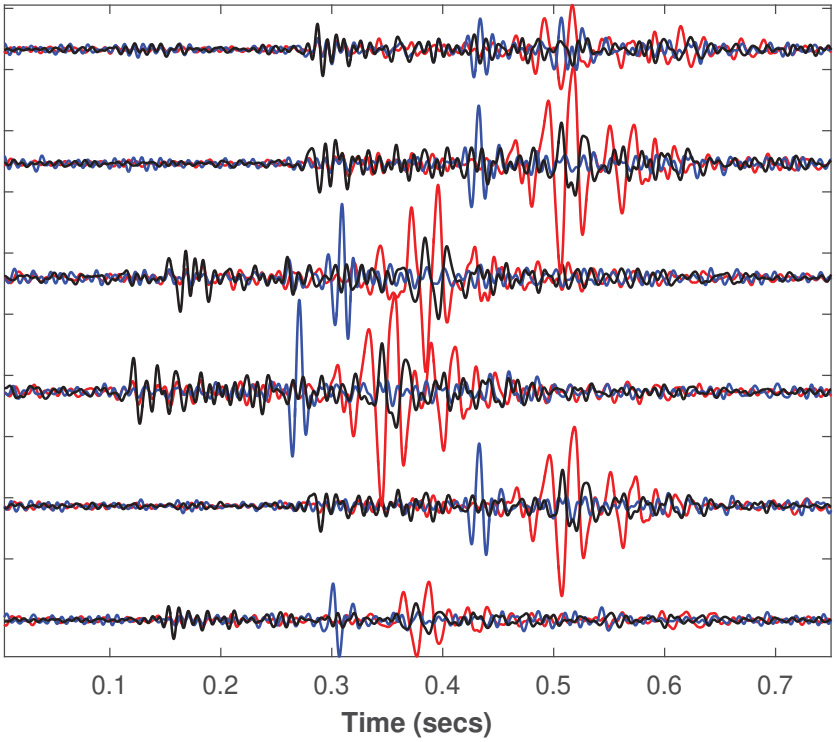
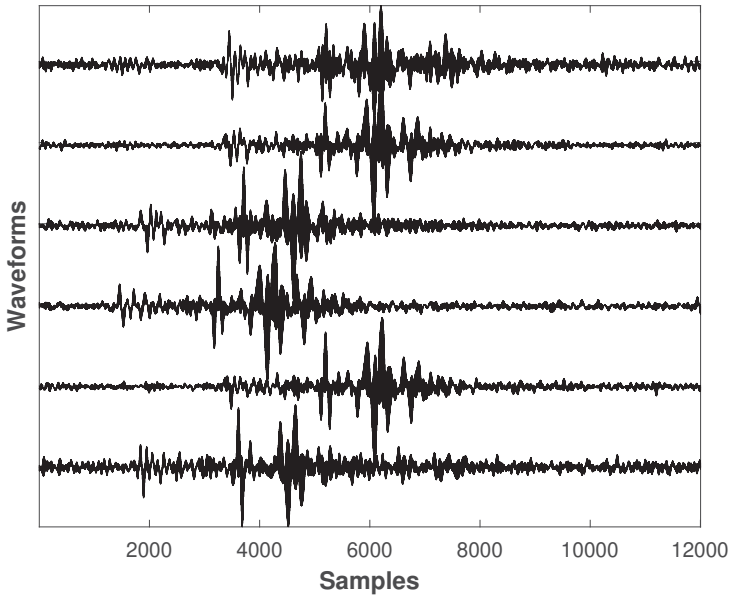
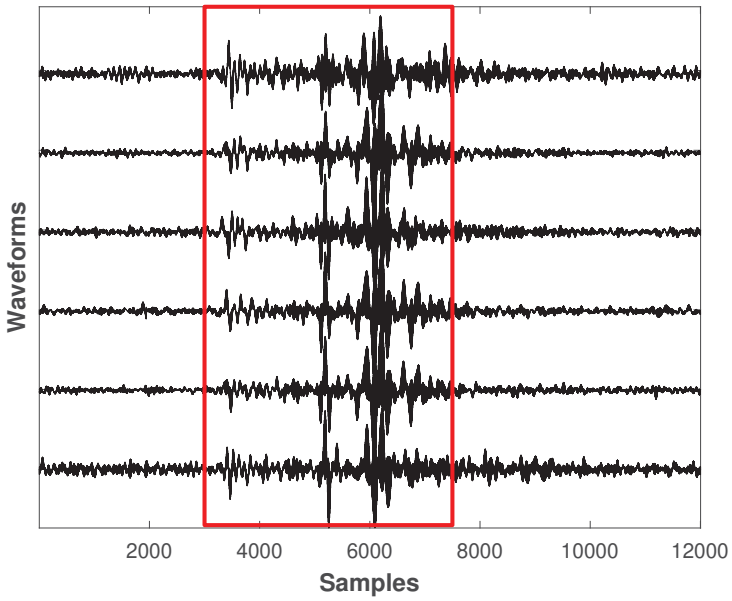


Figure 2.1: Three-component (3C) template waveforms from a microseismic data set. Signals 1 to 6 from the bottom. Vertical, north and east components shown by red, blue and black colors respectively.



(a)



(b)

Figure 2.2: a) Multiplexed template waveforms before time alignment. b) Multiplexed template waveforms after time alignment. Red box indicates selected data window.

For the matched filter test, each template is selected as a master waveform and eq 2.6 is applied to every other template in the design set to obtain a correlation coefficient value CC. For the subspace detector, a singular value decomposition (SVD) is applied to the aligned waveforms. The dimension of the subspace representation is then varied and the detection statistic c is computed for every dimension using eqs. 2.15. The results are shown in table 2.1 & 2.2.

Template number	1	2	3	4	5	6
1	1.00	0.83	0.78	0.79	0.81	0.51
2	0.83	1.00	0.95	0.80	0.97	0.65
3	0.78	0.95	1.00	0.77	0.95	0.61
4	0.79	0.80	0.77	1.00	0.77	0.48
5	0.81	0.97	0.95	0.77	1.00	0.64
6	0.51	0.65	0.61	0.48	0.64	1.00

Table 2.1: Correlation coefficient values between template signals

Dimension	Template number					
	1	2	3	4	5	6
1	0.71	0.96	0.95	0.73	0.95	0.42
2	0.74	0.96	0.96	0.99	0.97	0.56
3	0.74	0.97	0.98	1.00	0.97	0.99
4	0.82	0.97	1.00	1.00	0.98	1.00
5	0.99	0.97	1.00	1.00	0.99	1.00
6	1.00	1.00	1.00	1.00	1.00	1.00

Table 2.2: Detection statistics c as a function of subspace dimension of representation

We found the detection statistics c for the subspace detector are higher than the individual matched filter correlation coefficients CC between the template signals. The correlation coefficients were reasonably high for most of the tem-

plates except for template 6, which was only adequately captured by itself. Although signal 2 had the highest correlation coefficients with the other signals, its values were smaller than the detection statistics obtained using a subspace representation of 4. Overall, the subspace detector appears more sensitive to all templates present after a higher number of dimensions is used in the detection procedure.

2.5.2 Improvement in signal-to-noise ratio

Stacking is a robust operation commonly used in improving signal-to-noise ratio in seismic data processing (Yilmaz, 2001). It amplifies coherent signal and suppresses incoherent noise. Barrett and Beroza (2014) found the 1^{st} left singular vector closely approximates the stack of the template events to be represented. We investigate this relationship by comparing the similarity of the 1^{st} singular vector with the stack of the selected template waveforms. 50 iterations of Gaussian noise scaled to represent selected signal-to-noise ratios (SNR) are added to the aligned templates in figure 2.2. At each iteration, both the stack and 1^{st} left singular vector of the templates are obtained and the correlation coefficient between both calculated. In this test, the signal-to-noise ratio is defined as the ratio of the standard deviation between signal and noise. The standard deviations of the correlation coefficients and their mean value are then calculated after all 50 iterations at each SNR. To prevent errors implying correlation coefficients in excess of ± 1, a Fisher transform is applied to the correlation values. The Fisher transform is a variance-stabilizing technique which changes the distribution of the correlation coefficients to approximate a Gaussian distribution (Ehlers, 2002). The results are shown in figure 2.3. For the range of SNRs used, the average correlation values between the stack and 1^{st} left singular value were consistently high. In fact, it is almost 1 over the range of SNRs used signifying marginal differences between both waveforms. This phenomenon serves as evidence of the apparent noise suppression in the left singular vectors derived via the SVD.

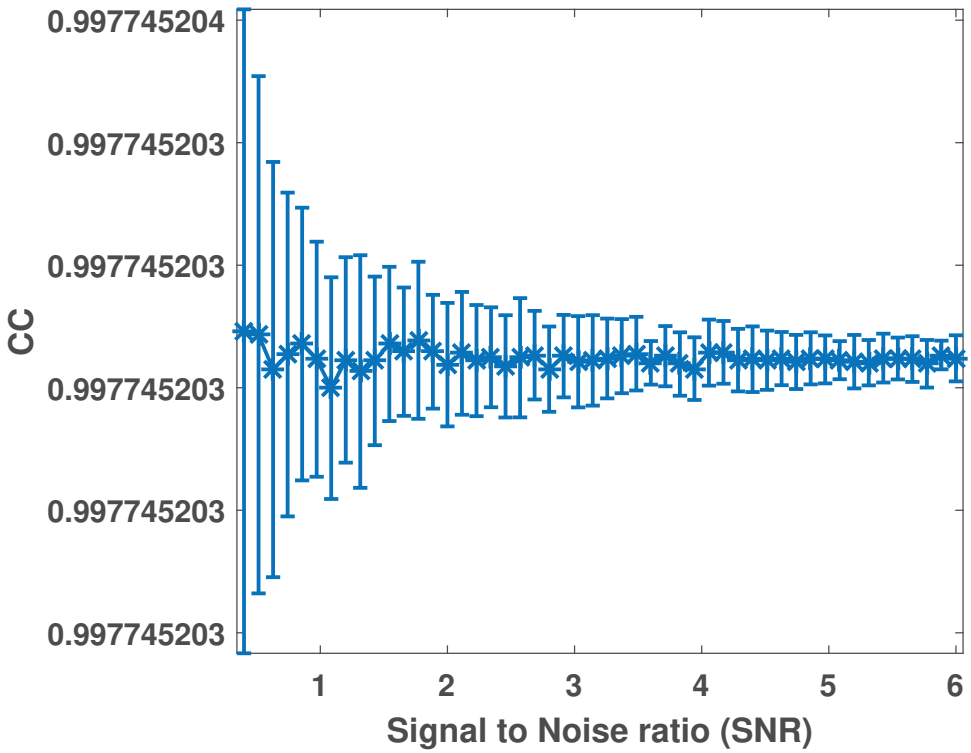


Figure 2.3: Correlation coefficient (CC) of stack and 1^{st} left singular vector as a function of SNR. The number of significant figures on the correlation coefficients has been modified to highlight the negligible differences across the different SNRs.

2.5.3 Detection threshold

Selecting an optimal detection threshold is an important processing step in event detection as it allows for the capture of as many events as possible at reasonable false alarm rates. Under the assumption of uncorrelated white noise, theoretical detection threshold values for template based detectors can be determined using Eq. 2.21. Due to the often unpredictable nature of noise, detection threshold may be directly estimated from noise data. This can be achieved by cross-correlating microseismic events with noise data, and observing the distribution of the correlation coefficients. This allows for the determination of an upper bound which would minimize false detections.

These methods of estimating the detection thresholds were applied to a mi-

croseismic data set, obtained during a hydraulic fracturing treatment (Eaton et al., 2014). Ten minutes of noise from different stages of treatment were selected in order to get fairly robust estimates of the noise levels present in the data, and values of 10^{-15} and 10^{-6} were selected as trial probabilities of false alarms P_F. Ten three-component microseismic events were selected from the data as templates, multiplexed and correlated with the noise files using 0.441s data windows N . The average variance $\sigma^2_{\hat{c}}$ of the resulting noise correlation coefficients was estimated to be 0.0014, and the embedding dimension $\hat{N}$ was calculated to be 0.064s using Eq. 2.20. The detection threshold value γ for the subspace detector was then inverted for from Eq. 2.21 for $d = 1$. To obtain a theoretical detection threshold for the matched filter, the subspace detector with a single dimension has coefficients which are the square of the matched filter (Eqs 2.25 & 2.26). The square root of the calculated subspace threshold value corresponded to the matched filter threshold. The theoretical thresholds are shown in table 2.3.

Detector	Predicted Threshold ($P_F = 10^{-6}$)	Predicted Threshold ($P_F = 10^{-15}$)
Subspace detector	0.0324	0.0841
Matched filter	0.18	0.29

Table 2.3: Theoretical detection thresholds for the subspace detector and matched filter for $d = 1$.

Direct estimates for the detection thresholds can also be made by observing the distribution of the noise correlation coefficients. Since the noise is assumed to be uncorrelated white noise, a Gaussian fit was made to the correlation coefficient population and shown in figure 2.4.

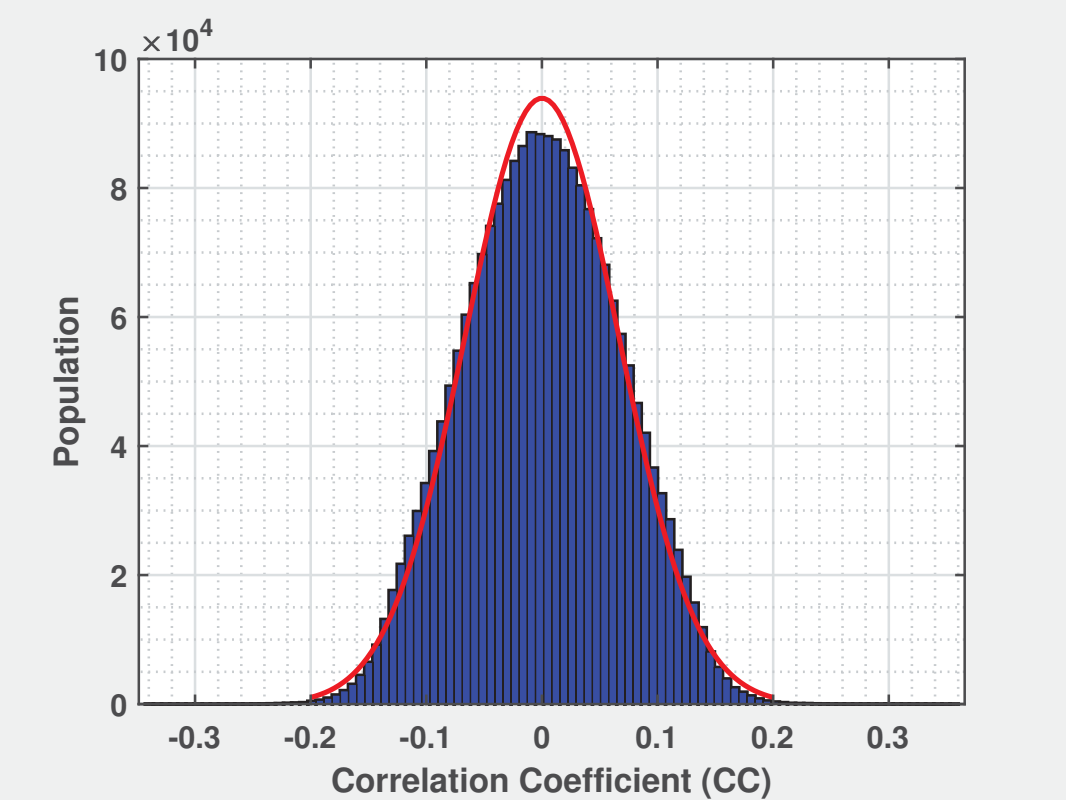


Figure 2.4: Distribution of the noise correlation coefficients for a matched filter template. A Gaussian fit is applied to the noise population, shown as the red curve.

The statistical fit to the noise population does a reasonable job in describing the data, and the threshold estimates in 2.3 roughly agree with the upper bound around the tail end of the fit. Although the fits do not completely describe the noise correlation coefficient population, they do serve as a decent guide for setting reasonable detection thresholds. We considered the case where $d = 1$, but higher dimensions may be used in subspace detection. Taking the waveforms in 2.2, a subspace of dimension 3 was constructed, and Eqs 2.14 & 2.15 were directly applied to the noise data. Since the detection statistic c was assumed to follow a beta distribution, a beta fit was applied to the population of noise detection statistics and shown in figure 2.5. We also estimated the theoretical threshold for $d = 3$, and these are shown in table 2.4.

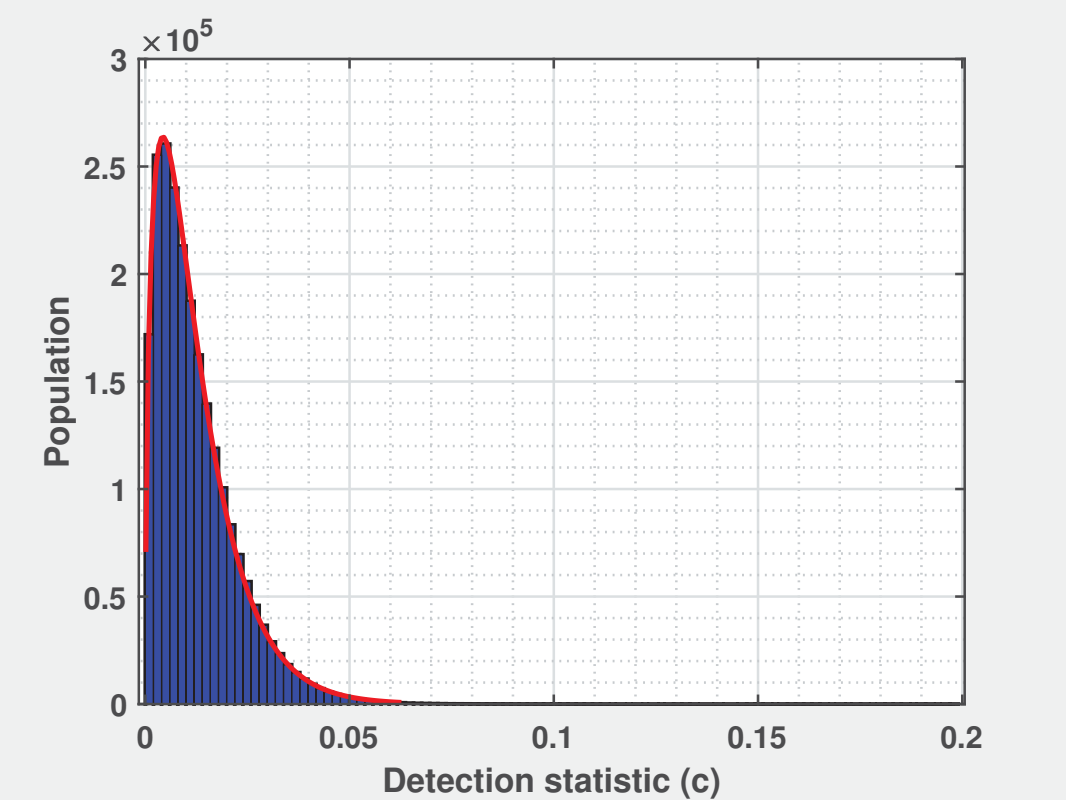


Figure 2.5: Distribution of the detection statistics on noise data for a subspace of dimension 3. A beta fit is applied to the noise population, shown as the red curve.

Detector	Predicted Threshold ($P_F = 10^{-6}$)	Predicted Threshold ($P_F = 10^{-15}$)
Subspace detector	0.0421	0.0973

Table 2.4: Theoretical detection thresholds for the subspace detector with $d = 3$.

Based on this plot, the theoretical predication again roughly agree with directly estimated values. This phenomena suggests it is possible to estimate a wide range of detection thresholds for different dimensions of representations by performing correlation of noise data with the events. This can increase efficiency since individual subspaces do not have to be constructed each time to get estimates for detection thresholds. Based on this plot, a threshold value $\gamma = 0.2$ would be ideal for minimizing false detections. The noise files are

only a sample of the total noise population and may not account for sporadic bursts of high amplitude noise or different types of noise that may occur within the data. Therefore it is imperative that the estimated threshold be high enough to account for those instances. The square root relationship between the coefficient of a matched filter and a single dimension subspace detector was used to set a comparable detection threshold for the matched filter, and $\tau_{\mathrm{CC}} = 0.45$ was set for the matched filter.

2.6 Summary

Subspace detection has been shown to be a generalization of standard template detection to multiple dimensions. The underlying mathematical principles behind the detector have been described and its design outlined in a series of steps. Through synthetic examples, the subspace detector is shown to adequately capture variations present within a set of waveforms as well as having improved signal-to-noise ratios especially in the 1^{st} left singular vector. Finally, theoretically estimating detection thresholds were shown to have reasonable agreement with estimates determined directly from the data, and may allow for efficient determination of detection thresholds for a range of subspace dimensions.

Chapter 3

Location methods in microseismic monitoring

3.1 Introduction

Microseismic events can occur anywhere in a reservoir or surrounding rock and are sometimes indicative of the re-rupturing of pre-existing faults (Arrowsmith and Eisner, 2006). Anthropogenic activities such as fluid injection and wastewater disposal can increase pore pressure in the subsurface, allowing faults to slide under pre-existing shear stresses (Van der Elst et al., 2013). The resulting stress perturbations can reactivate faults and open or close new fracture sets, which in turn generates microseismic events (De Meersman et al., 2009). Location of these microseismic events is of paramount importance in monitoring reservoir changes as precise hypocenter estimates better outline fracture geometry and the progress of fluid fronts during production, providing important constraints on fracture distribution and development (Oye and Roth, 2003; Chen et al., 2017). Earthquake locations, also known as hypocenters generally fall into two categories; absolute and relative locations. Absolute locations describe earthquake hypocenters that are specified within a fixed geographic system and time frame, while relative earthquake locations are determined with respect to another spatio-temporal object (e.g earthquake or seismic source) which may have an unknown or uncertain absolute location (Husen and Hardebeck, 2011).

Earthquake locations can be obtained using a variety of methods, and they range from linearized techniques to direct searches of a solution space. Some of these methods involve finding earthquake locations independently of other earthquakes, while other methods simultaneously determine the locations of several earthquakes at once. Each technique relies on certain assumptions which determine its relative strengths and weakness in a location procedure. Seismic data tend to have uncertainties such as random errors due to noise, and systematic errors due to un-modeled earth structure. Having a good knowledge of these limitations and strengths can be beneficial in terms of selecting the best location strategy for a given dataset or acquisition geometry (Husen and Hardebeck, 2011).

In this chapter, the location problem is described in terms of the relationship between earthquake arrival times at a given station and its hypocentral characteristics. The mathematical principles behind different location techniques are introduced and their implementation described. The benefits and disadvantages of each method are also be outlined.

3.2 Theory

Most earthquakes occur beneath the earth's surface and their source locations must be determined using observations from recording stations. The primary observations are usually the time it takes for the seismic energy to reach the receiver, referred to as arrival times. The arrival time T for an earthquake i to a seismic station k can be expressed as a path integral along the ray

$$T_k^i = \tau^i + \int_i^k u ds, \tag{3.1}$$

where τ represents the origin time of event i, u is the slowness field and ds is an element of path length. The travel time is a function of the known station's spatial coordinates (x_k, y_k and z_k), an assumed hypocenter location of the earthquake (x_i, y_i, z_i) including its origin time τ, and the velocity model

used to model the sub-surface. It is a function of the known station's spatial coordinates and velocity model and can be described as

$$t_k^i = \frac{\sqrt{(x^i - x^k)^2 + (y^i - y^k)^2 + (z^i - z^k)^2}}{v}, \tag{3.2}$$

with v representing the velocity of the medium, while the spatial coordinates of the event and receivers are given by x^i, y^i, z^i and x^k, y^k, z^k respectively. The equation shows travel time to be a nonlinear function of the model parameters. Eq. 3.2 can be applied to a 1D layered velocity model or 3D velocity model, where the velocity v becomes a function of either depth or spatial coordinates. In most cases the velocity model is kept fixed during the computation of the earthquake location, but can also be simultaneously determined using a set of well-constrained earthquakes (Husen and Hardebeck, 2011). With four unknowns (spatial coordinates and origin time), earthquake location can be described as an inverse problem. Usually four arrival time observations from at least three stations are needed to determine an earthquake's hypocenter and origin time (Waldhauser and Ellsworth, 2000; Havskov and Ottemöller, 2010). A solution is usually sought that minimizes the sum of the difference between the observed and calculated earthquake travel times i.e minimizes the residuals. The residual for station k are expressed as

$$r_k^i = (t_i^{obs} - t_i^{theo})_k, \tag{3.3}$$

where t_i^{obs} and t_i^{theo} are the observed and calculated travel times at station k for earthquake i.

For a single vertical receiver array, source locations can only be determined if the P-wave polarizations are included in the location procedure to constrain the source back azimuth (Phillips et al., 1998). The inclination of the ray from a source (x^i, y^i, z^i), at a distance d to a receiver k (x^k, y^k, z^k) may also be included as an additional constraint. The azimuths θ_k^i and inclinations γ_k^i (Oye and Roth, 2003) are given by

$$\theta_k^i = \arctan(\frac{x^i - x^k}{y^i - y^k}), \tag{3.4}$$

$$\gamma_k^i = \arccos(-\frac{z^i - z^k}{d}). \tag{3.5}$$

3.2.1 Phase arrival time picking

The standard method for obtaining arrival time picks requires visual inspection of the traces to accurately determine onset times. Due to the availability of large data volumes and the subjectivity involved, this method is fairly inefficient especially in real time monitoring. Accurate determination of the onset times of phase arrivals is necessary for precise event location and source mechanism analysis. To deal with the large data volumes, automatic procedures are necessary which provide similar to equivalent accuracy relative to manual processed arrival times. Numerous algorithms exist for picking purposes, such as the short time average over long time average (STA/LTA) detector (Allen, 1978; Trnkoczy, 1999), autoregressive modeling (Sleeman and Van Eck, 1999; Leonard and Kennett, 1999) and correlation techniques (De Meersman et al., 2009; Molyneux and Schmitt, 1999). In this thesis, cross correlation and autoregressive methods are employed for phase arrival time picking.

Cross-correlation

Cross correlation is an especially useful tool for correcting time picking inconsistencies between seismic multiplets since it provides very high precision arrival time measurements. If two events are highly similar, their time picks must be at the same position (Geller and Mueller, 1980). The time lag corresponding to the maximum peak of the cross correlation function between both events can be used to correct inconsistencies in arrival time picks. A time window around the phase of interest is selected, with the window starting around the initial arrival time pick of the phase. The windowed phase data is then correlated with the time series (Eq. 2.6) and the time at which the correlation function is maximum is considered the phase onset time. For multi-component data, the component with the maximum correlation is selected for time picking or the data may be multiplexed and a single estimate obtained. To avoid instances of noise correlation, a strict correlation threshold is usually required before an

arrival time pick is declared.

Autoregressive modeling & Akaike Information Criteria (AIC)

Seismic data are often modeled using an autoregressive (AR) process, where the value at a particular time depends on a linear combination of past values Leonard and Kennett (1999). For a time series containing a seismic signal ($x_n = x_1 ....., x_N$) where N is the length of the time series, it is assumed the intervals before and after the onset time are two different stationary time series. A model of order M can be fit to the data in both intervals as

$$x_t = \sum_{m=1}^{M} a_m^i x_{t-m} + e_t^i, \tag{3.6}$$

where $t = 1, ...M$ in the interval before the current time sample, and $t = N - M + 1, ...M$ for the interval after the current time sample ($i = 1, 2$ corresponds to the intervals), a_m^i are the autoregressive coefficients for order M ($m = 1...M$) and e_t^i represents the noise or non-deterministic part of the time-series (Sleeman and Van Eck, 1999). At each division point, the maximum likelihood function for both models as a function of the order of representation is derived and used to define the Akaike information criteria (AIC) (Sleeman and Van Eck, 1999). The AIC is used to determine the optimal order for the AR process since it gives a measure of the unreliability of the model fit (Zhang et al., 2003). For a fixed model order, the AIC for both models as a function of the merging point K of both intervals is expressed as

$$AIC(K) = (K - M)log(\sigma_{1,max}^2) + (N - M - K)log(\sigma_{2,max}^2) + C_2, \tag{3.7}$$

where σ corresponds to the maximum variance on the specified intervals at merging point K, M is the order of the model, N corresponds to the length of the time series and C_2 is an arbitrary constant. The point at which the AIC function is a minimum corresponds to the optimal separation of the two stationary time series, and is denoted the onset time (Sleeman and Van Eck, 1999). Usually the order of the autoregressive process is obtained by trial and

error before estimation of the AR coefficients and AIC calculations. However, the AIC function can also be calculated without the need for computing AR coefficients. For a time-series x of length N, the AIC is a function of the variance var in the two data segments and is given as

$$AIC(K) = Klog(var(x_1, ...x_K)) + (N - K - 1)log(var(x_k + 1, ...x_N)). \quad (3.8)$$

It is important to ensure that the data window on which the AIC is computed contains the onset of the phase in question. This is due to the fact the AIC defines the onset as the global minimum, and thus will always provide a minimum even if the onset of the signal is not contained within the analysis window. In situations where there are large amounts of noise or multiple phases, the uncertainty in the onset determination increases. Therefore it is important to fine-tune the picker in terms of the appropriate position of the time window used for onset time determination (Zhang et al., 2003).

3.2.2 P-wave polarization

P-wave polarizations generally provide estimates of the back-azimuths between a seismic event and a station, and are an especially useful parameter for constraining earthquake locations (De Meersman et al., 2006). Hodogram analysis and eigenvalue decomposition (EVD) are examples of techniques used to estimate polarization properties of a signal (Han, 2010; De Meersman et al., 2006; Vidale, 1986).

For three-component data $\mathbf{u}^j$, where j corresponds to the component (from 1 to 3) in a Cartesian frame of reference arranged as [$\mathbf{e}$(t),$\mathbf{n}$(t),$\mathbf{z}$(t)], the covariance matrix $\mathbf{S}$ in a time window around the signal is expressed as

$$\mathbf{S}_{jk} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{u}_i^j \mathbf{u}_i^k, \quad (3.9)$$

where N represents the number of samples in the time window (Bokelmann, 1995). Time-domain analysis tends to be restricted to linearly polarized phases

and suffers from stability problems in short time windows around zero crossings (De Meersman et al., 2006). The time domain signal can be converted to an analytic signal by introducing a complex component to it. The use of analytic signals is useful if signals with elliptical polarizations are analyzed (Vidale, 1986). It also tends to be less sensitive to the length of the signal window used (Chen et al., 2005). The analytic signal of the windowed is obtained by adding their Hilbert transform as the complex component. An eigenvalue decomposition of the resulting Hermitian covariance matrix S yields

$$\mathbf{S} = \mathbf{V}^* \wedge \mathbf{V}^{*T} = \sum_{i=1}^{3} \mathbf{v}_i \lambda_i \mathbf{v}_i^T, \tag{3.10}$$

where $\mathbf{V}^*$ and $\wedge$ correspond to our matrices of complex eigenvectors and real eigenvalues respectively. The eigenvector $\mathbf{a}_1$ associated with the largest eigenvalue points in the direction of the largest amount of polarization. It corresponds to a 3×1 orthonormal vector ($\mathbf{a}_1 = [\mathbf{a}_1(1), \mathbf{a}_1(2), \mathbf{a}_1(3)]^T$). From this eigenvector both the incidence angle and back-azimuth of the signal from the source to a receiver can be estimated. The incidence angle ϕ can be expressed as

$$\phi = \frac{\sqrt{Re(\mathbf{a}_1^2(1)) + Re(\mathbf{a}_1^2(2))}}{Re(\mathbf{a}_1(3))}, \tag{3.11}$$

where we denote the real part of the vector as *Re*. The back-azimuth θ is derived by

$$\theta = tan^{-1}(\frac{Re(\mathbf{a}_1(1))}{Re(\mathbf{a}_1(2))}). \tag{3.12}$$

The back-azimuths and incidence angles obtained using eqs 3.12 & 3.11 range from -90^o to 90^o.

3.3 Absolute location

3.3.1 Grid search

A simple approach to hypocenter location involves calculating the travel times and origin times in a given model space from all possible locations to each

station based on a given velocity model (Lomax et al., 2007). The point in the model producing the best agreement between observed and calculated arrival times is selected as the preferred hypocentral location and origin time. This location is obtained from the minimum of the sum of the squared residuals e from the N observations

$$e = \sqrt{\frac{\sum_{j=1}^{N} (r_j)^2}{N}}, \tag{3.13}$$

where e is the root mean square (RMS) error of the travel time residuals. The RMS error value is usually calculated at all grid points in the model, and the position with the lowest RMS error is selected as the solution. Sometimes it is possible within the model to have several locations with similar RMS error values. A measure of the uncertainty of the solution can then be obtained in the form of a contour map of rms values around the vicinity of the selected hypocenter location (Havskov and Ottemöller, 2010). Similarly we can calculate probability density functions of the derived hypocenters by assuming the azimuths and arrival times are normally distributed around the true locations (Eisner et al., 2010). The probability density of a hypocenter due to a single receiver and uncertain measurement is

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp^{-(x-x_m-x_o)^2/2\sigma^2}, \tag{3.14}$$

where x is the measured variable (P-, S-phase arrival times, azimuths or incidence angles), x_m is the predicted value obtained with a velocity model, x_o represents the origin time for the phase arrivals ($x_o = 0$ for azimuthal and inclination measurements), and σ is the standard deviation of the measured variable x. Under the assumption that each receiver and measured quantity is mutually independent, the probability density functions (PDFs) can be combined by multiplying the individual probabilities. For N receivers this can be expressed as

$$p(t_P, t_S, A) = Re^{-\sum_N (t_P - t_{mP} - t_o)^2/2\sigma_P^2} e^{-\sum_N (t_s - t_{mS} - t_o)^2/2\sigma_S^2} e^{-\sum_N (\theta - \theta_m)^2/2\sigma_\theta^2}. \tag{3.15}$$

In Eq. 3.15, t_P and t_S are the measured P- and S-wave arrival times, while θ denotes the measured azimuths. t_{mP}, t_{mS} and θ_m are the calculated P-, S-wave travel times and back-azimuths from a given model. t_o is the origin time and σ_P, σ_S and σ_θ are the standard deviations of the observed P- and S-wave arrival times and azimuths. The normalization constant R ensures the integral of eq. 3.15 over all possible locations is equal to one. The standard deviations here give a measure of the data quality, with smaller values highlighting more focused event locations. They are usually derived from the data. Eq. 3.15 may also use data from a single seismic phase and may not require the azimuthal values if the acquisition geometry has several boreholes and provide good azimuthal coverage of the events. However, using the arrival times from a single phase leads to a trade-off between the origin time and distance for a vertical receiver array since its poorly constrained (Eisner et al., 2010). By utilizing both P- and S- phase arrival times, this trade-off is minimized. Using eq. 3.15, hypocenter locations correspond to positions in the grid where the joint PDF has the maximum value.

Eq. 3.15 is based on least squares and produces good results if the misfits correspond to a Gaussian distribution. However, large outliers have fairly large impacts on the misfit since everything is squared. Minimizing the sum of absolute residuals could minimize the effects of outliers in this case and is considered more robust (Menke, 1999). The RMS error will thus be represented as

$$e = \sqrt{\frac{\sum_{j=1}^{N} |r_j|}{N}}. \tag{3.16}$$

This is known as the L1 norm solution. In this case, the probability density functions are given as

$$p(t_P, t_S, A) = Re^{-\sum_N |(t_P - t_{mP} - t_o)|/2\sigma_P^2} e^{-\sum_N |(t_s - t_{mS} - t_o)|/2\sigma_S^2} e^{-\sum_N |(\theta) - (\theta_m)|/2\sigma_\theta^2}. \tag{3.17}$$

This technique is not widely used since residuals with large outliers can be

down-weighted using the L2 norm approach (Waldhauser and Ellsworth, 2000).

The L2 and L1 approaches described provide the maximum likelihood hypocenter and origin time for an earthquake. Usually, a priori information may exist about the spatial location of an earthquake, but it is generally impossible to have a priori information about the origin time that is independent from the data (Tarantola & Valette, 1982). If the observed and predicted times are uncorrelated, a probabilitiy density function can be derived which involves minimizing the difference of arrival and travel times (Lomax et al., 2007)

$$p(t_P, t_S, A) = Re^{-\sum_N((t_P - t_{P_{avg}}) - (t_{mP} - t_{mP_{avg}}))^2/2\sigma_P^2} e^{-\sum_N((t_S - t_{S_{avg}}) - (t_{mS} - t_{mS_{avg}}))^2/2\sigma_S^2} e^{-\sum_N(\theta - \theta_m)^2/2\sigma_\theta^2}. \tag{3.18}$$

In the above equation, $t_{P_{avg}}$ and $t_{S_{avg}}$ refer to the average P- and S-wave arrival times recorded across the receivers, while $t_{mP_{avg}}$ and $t_{mS_{avg}}$ are the average calculated P- and S-wave travel times. Subtracting the average values from both time measurements standardizes them, and eliminates the origin time information from the arrival times.

Grid search methods take non-linearity of travel times into account by not utilizing partial derivatives. The searches utilized with these techniques may be exhaustive or directed (Lomax et al., 2007; Husen and Hardebeck, 2011), but usually results in complete location PDFs containing global and local maximums as they often sample a wide area. However, exhaustive grid searches within a volume can be expensive, especially if the search region is large, and can be up to 1000 times slower than location by iterative means (Havskov and Ottemöller, 2010). This could be mediated by running the computation once and saving the resulting grid into memory via a lookup table, avoiding the need for repeated calculations. However, depending on the size of the area being investigated this may not always be a feasible option. This technique is

mainly applicable to single events at a time since each earthquake location is done independently of other earthquakes.

3.3.2 Iterative methods

Iterative methods do not involve large-scale searchs and are predominantly computationally faster than direct search methods. The location problem can be linearized via a truncated Taylor series expansion of eq 3.1 around an initial guess of a hypocenter and origin time (x_0, y_0, z_0, t_0) (Geiger, 1912). It can be assumed that the resulting residuals are due to the error in the initial guess, and the guess was close enough to the true hypocenter location that the corrections needed to the initial estimates are small. In this case, the travel time residuals r at station k for earthquake i can be expressed as

$$r_k^i = \frac{\partial t_k^i}{\partial m} \triangle \mathbf{m}^i. \tag{3.19}$$

In eq 3.19, $\triangle \mathbf{m}^i = (\triangle x_0^i, \triangle y_0^i, \triangle z_0^i, \triangle t_0^i)$ are the perturbations to the four hypocentral parameters, while $\frac{\partial t_k^i}{\partial m}$ is the partial derivative of the travel time at the k^{th} station with respect to the four hypocentral parameters evaluated at the initial guess. Eq 3.19 can be written in matrix notation as:

$$\mathbf{r} = \mathbf{G} \triangle \mathbf{m}, \tag{3.20}$$

where $\mathbf{G}$ is the N $\times$ 4 matrix of partial derivatives at N stations, $\mathbf{r}$ is the N $\times$ 1 vector of travel time residuals and $\triangle \mathbf{m}$ is the 4 $\times$ 1 vector with the adjustments to the hypocentral parameters. The solution to the linear system of equations is obtained iteratively using standard least squares techniques. After each iteration, $\triangle \mathbf{m}$ is used to update the location and origin time, and the process continues until convergence is reached or a stopping criterion is met (Jones et al., 2008; Havskov and Ottemöller, 2010; Waldhauser and Ellsworth, 2000). P-wave polarization estimates may also be included in the system of equations if azimuthal coverage is poor. The back-azumuth residuals can also be expressed in terms of perturbation to the hypocentral parameters

$$r^i_{\theta k} = \frac{\partial \theta^i_k}{\partial m} \triangle \mathbf{m}^i. \tag{3.21}$$

Eq3.21 can be expressed in the form of eq 3.20 and added to the system of equations to be solved

$$\mathbf{r}_\theta = \mathbf{G}_\theta \triangle \mathbf{m}, \tag{3.22}$$

Convergence is usually rapid using this method but breaks down if the initial guess is far away from the optimal solution. The solutions can also become unstable when the complete solution is irregular or multiple solutions exist due to insufficient data or the presence of outliers (Husen and Hardebeck, 2011; Lomax et al., 2007). The solutions obtained this way may be poor representations of the complete solution since it could converge to a local minimum which may not be known unless the residuals are very bad. Usually a test with a grid search program could highlight if the solution is a local minimum or not or tests could be made with several starting locations (Havskov and Ottemöller, 2010).

3.4 Relative location

Relative location techniques are commonly used techniques to precisely relocate earthquake clusters containing highly similar waveforms. Repeating source mechanisms tend to produce highly similar waveforms and the precisions of their locations can be improved especially if waveform cross-correlation is utilized to improve arrival time estimates and determine high-precision relative arrival times. There are several techniques which utilize this principle, two of which are most common and examined in this chapter.

3.4.1 Master event relocation

The master event technique has been a widely used tool in the relative location of seismic events (Stoddard and Woods, 1990; Bouchaala et al., 2013; Frémont and Malone, 1987). In this method, events are relocated relative to a single well

located master event under the assumption that the events are highly similar, and thus co-located spatially. If the master event and child event are in close proximity, the ray paths for both events to a common receiver can be assumed to be similar, and the travel time differences between them can be attributed to the velocity heterogeneity in the source region between both events (Jones et al., 2008; Waldhauser and Ellsworth, 2000). These assumptions hold provided the separation distance between the events is small compared to their distance to the receiver.

The master event relocation method is illustrated by considering a seismic event i recorded at station k. The arrival time T_k^i can be written as

$$T_k^i = t_0^i + t_k^i + \Delta t_k^i + \epsilon_k^i, \tag{3.23}$$

where t_0^i is the origin time, t_k^i is the predicted travel time obtained from a reference velocity model, Δt_k^i is the unknown travel time anomaly caused by differences between the true unknown velocity model and the reference velocity model, and ϵ_k^i is the arrival time reading error. The arrival time for the master event ME is given by

$$T_k^{ME} = t_0^{ME} + t_k^{ME} + \Delta t_k^{ME} + \epsilon_k^{ME}. \tag{3.24}$$

We can rewrite Eq.3.23 in terms of the master event

$$T_k^i = t_0^{ME} + \Delta t_0 + t_k^{ME} + \frac{\partial t_k^{ME}}{\partial x_0}\Delta x_0 + \frac{\partial t_k^{ME}}{\partial y_0}\Delta y_0 + \frac{\partial t_k^{ME}}{\partial z_0}\Delta z_0 + \Delta t_k^i + \epsilon_k^i, \tag{3.25}$$

where Δt_0, Δx_0, Δy_0, Δy_0 and Δz_0 are perturbations to the hypocentral parameters and origin time of the master event to obtain arrival time of the seismic event. Since the events are assumed to be in close proximity with similar ray paths, then the errors due to velocity are roughly equal i.e. $\Delta t_k^i \approxeq \Delta t_k^{ME}$. Subtracting eq 3.24 from eq 3.25 then gives

$$T_k^i - T_k^{ME} = \Delta t_0 + \frac{\partial T_k^{ME}}{\partial x_0}\Delta x_0 + \frac{\partial T_k^{ME}}{\partial y_0}\Delta y_0 + \frac{\partial T_k^{ME}}{\partial z_0}\Delta z_0. \tag{3.26}$$

The error in the arrival time difference can further be significantly reduced via the use of cross-correlation to correct arrival picking inconsistencies. Thus, for a seismic phase at station k, the relative hypocentral parameters between a master event ME and a child event i can be expressed as

$$r_k^{iME} = \frac{\partial t_k^{iME}}{\partial m}\triangle \mathbf{m}^{iME}, \tag{3.27}$$

where $r_k^{iME} = (T^i - T^{ME})_k$ corresponds to the difference in arrival times between the child event and the master event at the k^{th} station, and $\triangle \mathbf{m}^{iME}$ $= (\triangle dt_o^{iME}, \triangle dx_o^{iME}, \triangle dy_o^{iME}, \triangle dz_o^{iME})$ corresponds to the relative changes in hypocentral parameters between the child event and master event (Jones et al., 2008). The observables may also correspond to either absolute travel time differences, or cross-correlation relative travel time differences. Eq 3.27 may be written in matrix notation and is given by Eq 3.28:

$$\Delta \mathbf{r}_m = \mathbf{A}_m \Delta \mathbf{m}, \tag{3.28}$$

where $\Delta \mathbf{r}_m$ is the N $\times$ 1 vector of the arrival time differences at N stations, and $\triangle \mathbf{m}$ is the 4 $\times$ 1 vector of unknown deviations from the master event hypocentral coordinates and origin time. The observables in Eq 3.27 can also be travel time differences between the master event and child event if origin time estimates are available. The elements of the $N \times 4$ partial derivatives matrix $\mathbf{A}_m$ are computed at the master event location. Similarly, the system of equations can be extended to include azimuth information, especially in the case of restricted acquisition geometries. Noise as well as small deviations of the borehole could result in large errors in determining the azimuthal distribution of the fractures (Bulant et al., 2007). If the ray paths are similar between the master event and the child event, then it is reasonable to expect the azimuthal values to contain common errors. Therefore, the differential azimuth values should be more accurate than the absolute values. The resulting set of

equations is given by

$$\theta_k^i - \theta_k^{ME} = \frac{\partial \theta_k^{ME}}{\partial x_0} \Delta x_0 + \frac{\partial \theta_k^{ME}}{\partial y_0} \Delta y_0 \tag{3.29}$$

$$r_{\theta k}^{iME} = \frac{\partial \theta_k^{iME}}{\partial m} \triangle \mathbf{m}^{iME}, \tag{3.30}$$

In matrix form, it corresponds to

$$\Delta \mathbf{r}_{\theta m} = \mathbf{A}_{\theta m} \Delta \mathbf{m}, \tag{3.31}$$

where $\mathbf{A}_\theta$ is the $N \times 4$ matrix containing the partial derivatives of the azimuths at the master event location and $\Delta\theta$ is the N $\times$ 1 vector of the azimuthal differences between the master and child events.

The combined system of equations is given by

$$\begin{pmatrix} \mathbf{W}_t \mathbf{A}_m \\ \mathbf{W}_\theta \mathbf{A}_{\theta m} \end{pmatrix} \Delta \mathbf{m} = \begin{pmatrix} \mathbf{W} \Delta \mathbf{r}_m \\ \mathbf{W} \Delta \mathbf{r}_{\theta m} \end{pmatrix} \tag{3.32}$$

$\mathbf{W}_t$ and $\mathbf{W}_\theta$ are diagonal matrices to weight each equation with *a priori* weights. No iteration is required in solving eq 3.28 and 3.32 since the location of the master event is already known.

Using this technique, systematic errors due to an incorrect velocity model are the same for all events since the ray paths are similar and this has no effect on the accuracy of the relative event locations (Deichmann and Garcia-Fernandez, 1992). The absolute location of the cluster is thus dependent on the errors in the arrival time estimates of the master. A disadvantage of this technique lies in the possible error propagation through the entire event cluster due to correlation of noise that may be present in the master. The maximum spatial extension of the cluster that can be relocated using this method is restricted since all events must correlate with the master event, and location error increases with distance away from the master (Waldhauser and Ellsworth, 2000). Its application is therefore limited to groups of highly similar events.

3.4.2 Double difference relocation

Like the master event technique, the double difference method of Waldhauser and Ellsworth (2000) takes advantage of the close proximity of seismic events in its relocation procedure due to the assumed similarity of the ray paths. The difference between observed and predicted travel time differences for event pairs are minimized as opposed to minimizing the residual for single events. For a pair of seismic events i and j recorded at station k, the difference in eq 3.19 for both events is taken and given by

$$dr_k^{ij} = \frac{\partial t_k^i}{\partial m}\Delta \mathbf{m}^i - \frac{\partial t_k^j}{\partial m}\Delta \mathbf{m}^j, \tag{3.33}$$

where dr_k^{ij} is the residual between the observed and calculated differential travel times between the two events. dr_k^{ij} is defined as the double difference and corresponds to

$$dr_k^{ij} = (t_k^i - t_k^j)^{obs} - (t_k^i - t_k^j)^{pred} \tag{3.34}$$

The observables in eq 3.33 could be absolute travel times t or cross-correlation relative travel-time differences. Eq 3.33 can be written out in full as

$$dr_k^{ij} = \frac{\partial t_k^i}{\partial x}\Delta x^i + \frac{\partial t_k^i}{\partial y}\Delta y^i + \frac{\partial t_k^i}{\partial x}\Delta z^i + \Delta t_0^i - \frac{\partial t_i^j}{\partial x}\Delta x^j - \frac{\partial t_i^j}{\partial y}\Delta y^j - \frac{\partial t_i^j}{\partial x}\Delta z^j - \Delta t_0^j \tag{3.35}$$

The partial derivatives of the travel times at the k^{th} stations for events i and j in eq 3.33 are evaluated at the event hypocenters. $\Delta \mathbf{m}^i$ and $\Delta \mathbf{m}^j$ correspond to the changes in the hypocentral parameters for both events to make the model better fit the data. For a set of N events, eq 3.35 can be combined for all event pairs over all stations to form a system of linear equations of the form

$$\mathbf{Em} = \mathbf{dr}, \tag{3.36}$$

where $\mathbf{E}$ is the $M \times 4N$ matrix of partial derivatives with respect to travel time, $\mathbf{dr}$ is the $M \times 1$ data vector containing the double differences and $\mathbf{m}$ is the $4N \times 1$ vector containing the changes in hypocentral parameters. In this system

of equations, M represents the number of double-difference observations. Eq. 3.36 may be extended by 4 equations to ensure the mean shift of all earthquakes during relocation is zero. This is represented by

$$\sum_{i=1}^{N} \Delta \mathbf{m}_i = 0 \tag{3.37}$$

for each coordinate and travel time (Waldhauser and Ellsworth, 2000). An initial solution to eq 3.36 is obtained from the starting hypocenter location and *a priori* weights. The locations are updated and a new partial derivative matrix $\mathbf{E}$ is calculated based on the updated locations. This process continues until a stable solution is obtained. Matrix $\mathbf{E}$ can be obtained by applying a differencing operator $\mathbf{Q}_{DD}$ to matrix $\mathbf{G}$ from eq 3.20 i.e. $\mathbf{E} = \mathbf{Q}_{DD}\mathbf{G}$. Azimuths can also be included in the procedure

$$dr_{\theta k}^{ij} = \frac{\partial \theta_k^i}{\partial m} \Delta \mathbf{m}^i - \frac{\partial \theta_k^j}{\partial m} \Delta \mathbf{m}^j, \tag{3.38}$$

written in full this become

$$dr_{\theta k}^{ij} = \frac{\partial \theta_k^i}{\partial x} \Delta x^i + \frac{\partial \theta_k^i}{\partial y} \Delta y^i - \frac{\partial \theta_i^j}{\partial x} \Delta x^j - \frac{\partial \theta_i^j}{\partial y} \Delta y^j, \tag{3.39}$$

where $dr_{\theta k}^{ij}$ is defined as

$$dr_{\theta k}^{ij} = (\theta_k^i - \theta_k^j)^{obs} - (\theta_k^i - \theta_k^j)^{pred}. \tag{3.40}$$

Eq. 3.39 can also be expressed in matrix form

$$\mathbf{E}_\theta \mathbf{m} = \mathbf{d}_\theta, \tag{3.41}$$

where $\mathbf{E}_\theta$ is the $M \times 4N$ matrix of back-azimuth partial derivatives, $\mathbf{d}_\theta$ is the $M \times 1$ data vector containing the back-azimuth double differences and $\mathbf{m}$ is the $4N \times 1$ vector containing the changes in hypocentral parameters. In this case $\mathbf{E}_\theta = \mathbf{Q}_{DD}\mathbf{G}_\theta$.

The resulting system of equations using both arrival times and back-azimuths is expressed as

$$\begin{pmatrix} \mathbf{W}_t\mathbf{E} \\ w\mathbf{G} \\ \mathbf{W}_\theta\mathbf{E}_\theta \\ u\mathbf{G}_\theta \end{pmatrix} \mathbf{m} = \begin{pmatrix} \mathbf{W}_t\mathbf{Q}_{DD} \\ w\mathbf{I} \\ \mathbf{W}_\theta\mathbf{Q}_{DD} \\ u\mathbf{I} \end{pmatrix} \mathbf{d}, \tag{3.42}$$

where w and u are scalar values which give the relative weighting between the absolute and differential arrival times and azimuths, $\mathbf{I}$ is the identity matrix and $\mathbf{d}$ is the data vector containing both travel time and azimuth observations ($\mathbf{d} = [\mathbf{dr}\ \Delta\mathbf{r}\ \mathbf{d}_\theta\ \Delta\mathbf{r}_\theta]^T$).

In this technique, interevent distances between correlated events are determined to the accuracy of the cross-correlation data while simultaneously determining the relative locations of other multiplets and uncorrelated events to the accuracy of the catalogue data. This method has been shown to improve the image of seismicity by collapsing locations into sharper clusters, leading to clearer identification of faults. Like the master event technique, similar ray paths for highly similar events help reduce systematic errors due to erroneous velocity models. A drawback to using cross-correlation techniques is the fact that they usually provide relative locations if only correlation data are used. Furthermore, the system of equations can become quite large for numerous events and screening of the data is necessary to optimize the linkage between the events and minimize redundancy in the data set (Waldhauser and Ellsworth, 2000). The system of equations may also become ill-conditioned especially if one event is poorly linked to other events, poor azimuthal coverage and errors in data measurements. This is usually reduced by damping which helps stabilize the system of equationss. Further stabilization can be achieved by applying preconditioning to the system via normalization of the columns of the derivative matrices, usually by its L2 norm.

3.5 Summary

We have described earthquake location in terms of a minimization between observed and predicted variables. Different earthquake location methods were introduced and the underlying principles governing each method described. Finally, their relative strengths and weaknesses were outlined in terms of their efficiency, accuracy and applicability.

Chapter 4

Relative location using the subspace detector

4.1 Introduction

Automated detection and location of seismic events is an important tool in the operation of global and local seismic networks (Lee et al., 1981; Oye and Roth, 2003). Microseismic events tend to be more abundant than larger magnitude earthquakes, and require precise detection and location to provide a clear picture of actively deforming regions over short observation periods (Hansen and Schmandt, 2015). With an increasing number of recording stations and higher sampling rates, robust and efficient techniques are required to keep up with the amount of active seismicity occurring and provide real-time images of a reservoir. This is especially valuable in hydraulic fracturing which is associated with high microseismic activity (Oye and Roth, 2003).

Most microseismic events are low magnitude ($M \leq 2$) and contaminated by high amplitude noise, which poses challenges in terms of event detection and accurate hypocenter localization due to difficulties in picking precise arrival time readings and inaccurate back-azimuth estimates (Hansen and Schmandt, 2015; Song et al., 2014). Furthermore, other factors such as restricted network geometries and ambiguous arrival time readings negatively impact the ability to

resolve the fine structure of seismicity (Waldhauser and Ellsworth, 2000). Even when the earthquake data have good signal-to-noise ratios, errors in the velocity model introduce systematic biases in the location estimates (Li et al., 2013).

Several techniques are used in tackling the detection problem, one of which is subspace detection (Harris, 2006; Song et al., 2014; Barrett and Beroza, 2014). It involves the construction of a library of waveform templates which are used to detect other weaker events using waveform similarity. The library of waveforms can be constructed using a family of similar waveforms or a diverse group of waveforms, making it fairly robust in handling waveform variation. The algorithm assumes that the undetected events can be represented as a linear combination of the largest singular vectors, formed from the library of template events (Barrett and Beroza, 2014). The first few singular vectors tend to have improved signal-to-noise ratios relative to the templates since they predominantly represent the signal energy of the design set. This definition implies that the singular vectors may contain some information on the source mechanism(s) within the design set, and may have similar characteristics to the template events in terms of their frequency content and waveform character. Two events are considered highly similar if their maximum event separation is not greater than a quarter wavelength of the dominant frequency; i.e they are co-located (Geller and Mueller, 1980). This similarity suggests the left singular vectors may contain hypocentral characteristics similar to the template events and could be used for location in a relative sense. If the singular vectors are representations of the source mechanism(s) contained in the design set, they could potentially be used as master events.

This chapter aims to introduce the subspace detector as a means of performing relative location of microseismic events. Unlike other relative location routines, the subspace detector offers the potential for an efficient means of simultaneous detection and location of seismic events. The higher signal-to-noise ratio (SNR) of the 1^{st} singular vector compared to the raw templates may improve the accuracy and precision of earthquake locations. In the first part of this chapter, the processing steps involved in obtaining the 1^{st} singular vec-

tor representation at each recording station will be outlined, and a procedure for obtaining its hypocenter location will be described and explained. Then the relative location methodology using the subspace detector will be introduced. In the second part of this chapter, the properties of the singular vectors will be examined in terms of their waveform characteristics in relation to the templates forming its subspace, and their hypocenter location. Finally, the relative location capabilities of the subspace detector will be compared to that of the conventional double difference and master event methods via synthetic examples.

4.2 Absolute location of the first singular vector

Locating an earthquake usually involves the use of a velocity model and phase arrival times at different receivers. In situations where a restricted acquisition geometry is employed, P-wave polarizations are needed to constrain the source back-azimuth. The first singular vector for a given subspace describes the main characteristics of the design set waveforms, and this may include hypocenter information. Obtaining an absolute location for this vector requires the same parameters necessary for locating routine earthquakes, and these are dependent on the processing steps involved in constructing the subspace.

4.2.1 Methodology

Singular vector representation

The processing steps involved in constructing the signal subspace are explained in more detail in section 2.3. Template events are obtained either using conventional detection methods, or from a previously collected catalog. These are then clustered together into groups based on their similarity using cross-correlation. A group of interest is selected, the waveforms are multiplexed and aligned in time at each receiver using cross-correlation. After alignment, a data window is defined around the phase(s) of interest. The window should be

wide enough to include the onset of the phase, and its position is kept fixed across all receivers. Fixing the window position is important as it allows for the capture of the relative time moveout of the singular vector arrival times across the receivers. An SVD is applied to the aligned waveforms within the chosen temporal window, and the matrix of left singular vectors is obtained.

P- and S-wave arrival times

After application of the SVD, the first singular vector is demultiplexed back into 3C form, and its arrival times are picked at each receiver. Since the singular vectors are not physical events, the picked arrival times are arbitrary with unknown origin times. However, the relative arrival time differences across the receivers allow for the estimation of the most likely hypocenter for this vector.

P-wave polarizations

If a single vertical borehole acquisition is used, P-wave polarization estimates are required to constrain the source back-azimuth. A time window encompassing the P-wave on the first singular vector is defined, and its covariance matrix computed. An eigenvalue decomposition is applied to the matrix, and polarization estimates at each receiver are obtained using eq 3.11.

Hypocenter location

With the arrival times and polarization estimates, a hypocenter location can be obtained for the first singular vector using a grid search approach. Due to the arbitrary nature of the arrival times, the issue of the origin times can be bypassed by utilizing eq 3.18. The average arrival time and predicted travel time are subtracted from the observed arrival times and predicted travel times respectively before the minimization procedure. Once a location is determined, an origin time may be obtained by subtracting the predicted travel times from the observed arrival times. This origin time is then subtracted from the arrival times to obtain a set of observed travel time measurements.

4.3 Relative location

The relative locations between events can often be obtained with greater accuracy than the absolute locations since they minimize the influence of travel time errors and velocity model uncertainties on the accuracy of earthquake locations. Several techniques such as the double difference algorithm and the master event method are used to obtain these locations. In this section, the main relative location methods considered in this study are introduced, and their methodology outlined.

4.3.1 Double difference relocation

The double difference relocation method utilizes all the child events in a pairwise sense. The processing steps involved in this process are as follows:

1. For each detected event, obtain arrival time and P-wave polarization estimates at each receiver. With these values, obtain initial hypocenter locations for the events.

2. Using cross correlation, either obtain relative travel time differences between event pairs, or absolute travel time differences. In the case of the single vertical borehole, the relative back azimuths may be computed by differencing the back azimuths, or analytically obtained using the method described in Appendix B.

4. With the relative travel times and back-azimuth measurements, eq 3.42 is solved to obtain the vector of relative changes in hypocentral parameters between the event pairs.

4.3.2 Master event relocation

In this method, an event is chosen from a set of detected events and used as a master. This event is usually closest to all the other events and has the best signal-to-noise ratio. The processing steps involved in this process are as follows

1. Select the master event from the set of detected events, obtain ar-

rival time and back-azimuth estimates across the receiver array and estimate a hypocenter location for this event.

2. For each detected event, get arrival time and P-wave polarization estimates at each receiver.

3. Obtain cross correlation relative travel time differences between the master and detected events, or absolute travel time differences. In the case of the single vertical borehole, the relative back azimuths may be computed by differencing the absolute back-azimuths, or obtained analytically using the method described in Appendix B. The 180^0 ambiguity associated with these estimates is normally solved by constraining the events to the quadrant containing the treatment well or source.

4. With the relative travel times and back-azimuth measurements, eq 3.32 is solved to obtain the vector of relative changes in hypocentral parameters between the detected events and master event.

4.3.3 Subspace relocation

Assuming the location of the first singular vector is in close proximity to the events forming the subspace, this vector can be used as a master event in a relocation procedure. The higher signal-to-noise ratios of this singular vector relative to the detected events may offer the possibility of improved absolute locations due to more accurate arrival time readings, which in turn will lead to more accurate relocations. The processing steps involved in this process are as follows:

1. Obtain the first singular vector representation across the receiver array, and get its arrival time and P-wave polarization estimates. With those parameters, obtain its hypocenter location.

2. For each event detected by the subspace detector, obtain arrival time and P-wave polarization estimates at each receiver.

3. Obtain cross correlation relative travel time differences between the singular vectors and detected events, or absolute travel time differences. In

the case of the single vertical borehole, the relative back-azimuths may be computed by differencing the absolute back-azimuths, or obtained analytically using the method described in Appendix B.

4. With the relative travel times and back-azimuth measurements, eq 3.32 is solved to obtain the vector of relative changes in hypocentral parameters between the child event and singular vector.

4.4 Tests and applications

4.4.1 Effects of waveform variation

In most seismic operations, seismic events may be due to different source mechanisms, or could be the result of highly similar sources. Depending on the application, all detected events may be used at once in the construction of the signal subspace, or detailed analysis of waveform multiplet groups may be undertaken. The behavior of the left singular vectors might change depending on the chosen application so an understanding of how waveform character influences the left singular vector representation is important if a relative location procedure is to be employed. In this test, the influence of waveform similarity on the first singular vector will be examined.

Methodology

Four synthetic microseismic events were generated in a homogeneous velocity medium, with P-wave and S-wave velocities of $V_p = 4000$m/s and $V_s = 2300$m/s respectively. The sampling rate (dt) was set to 0.001s, and the events were recorded at a single receiver station with no attenuation in the medium. The origin time of the events t_0 was set to 0. No noise was added to the system. In this experiment, two cases were considered; the case where the events were due to the same source mechanism, and the case where they were from different sources. In the case where the events are similar, the source mechanism was assumed to be double couple (DC), and a Ricker wavelet was used as the source wavelet. Two of the events were assumed to be the result of slight variations

in the source mechanism. Figure 4.1 shows the location of the events relative to the receiver.

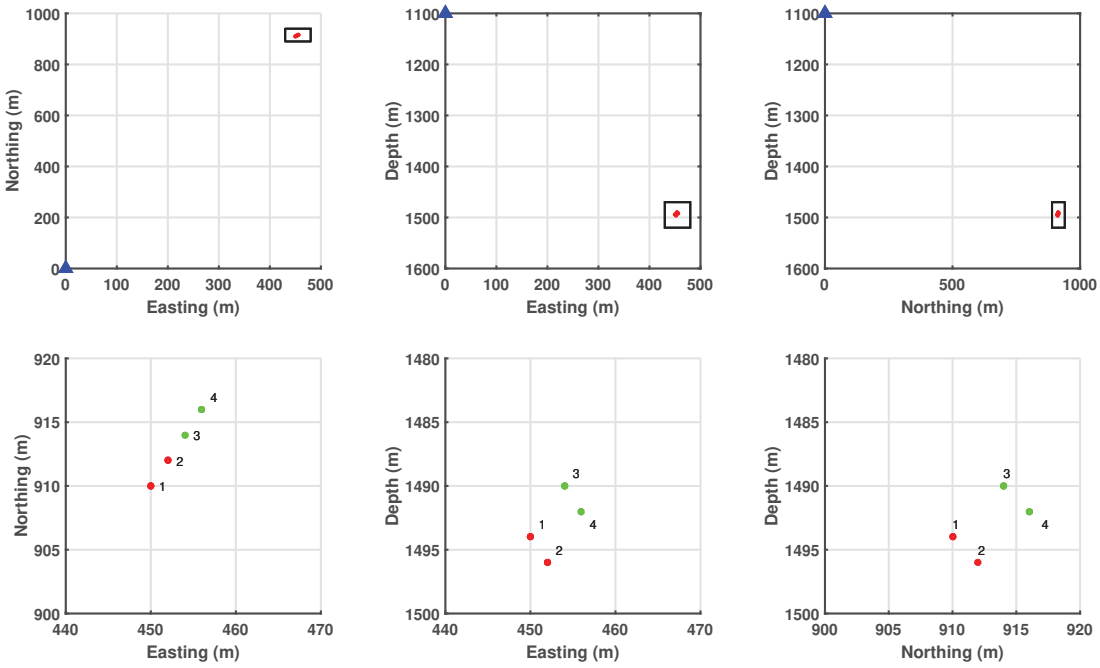


Figure 4.1: Location of synthetic seismic events relative to single receiver showing in 2D plots. Event cluster shown as red dot and receiver shown as blue triangle in the *Top row*. A close up view of the events is shown in the *Bottom row* and corresponds to the region highlighted by the black box. Events are labeled. Red dots indicate the original DC source, while green dots correspond to slight variations within the DC source.

In the next case, two of the events were assumed to be due to a DC source represented by a Ricker wavelet, while the other two originated from a compensated linear vector dipole (CLVD) source represented by a Meyr wavelet. The locations of the CLVD source events were also moved slightly further away from the DC source. The source locations are shown in figure 4.2.

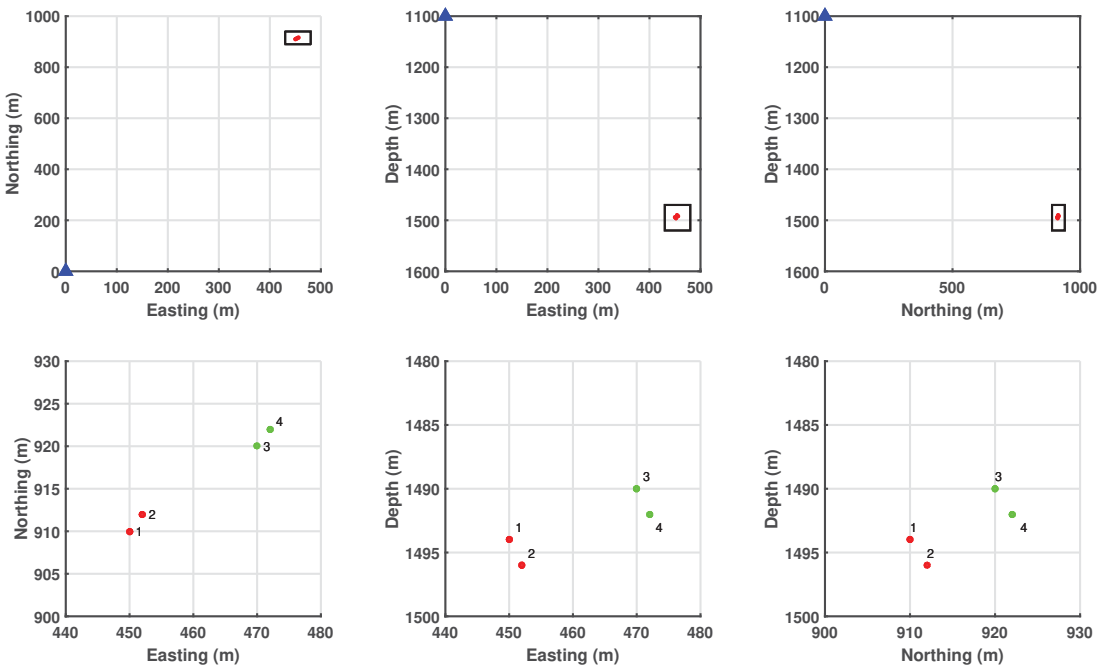
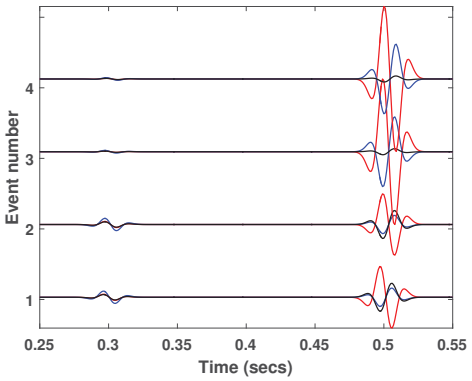
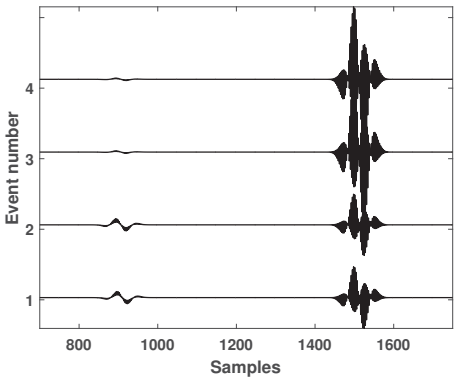


Figure 4.2: Location of synthetic seismic events relative to single receiver showing in 2D plots. Event cluster shown as red dot and receiver shown as blue triangle in the *Top row*. A close up view of the events is shown in the *Bottom row* and corresponds to the region highlighted by the black box. Events are labeled. Red dots indicate the original DC source, while green dots correspond to the CLVD source.

The 3C waveforms were multiplexed and aligned in time using an iterative stacking procedure (De Meersman et al., 2006), and are shown in figures 4.3 and 4.4.

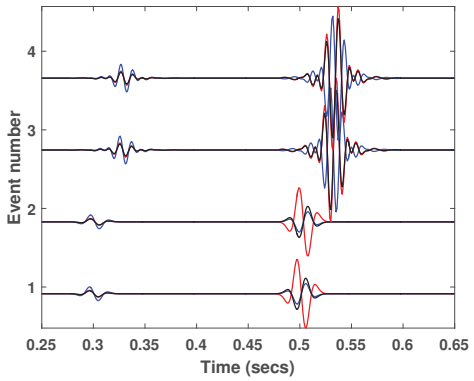


(a) Three component synthetic events

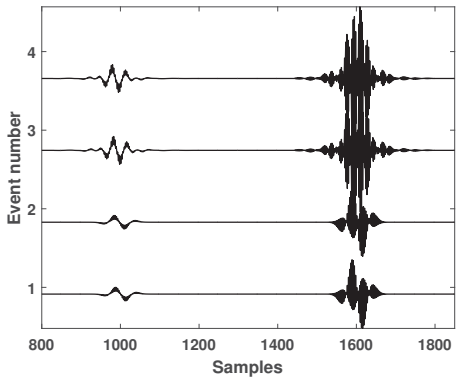


(b) Aligned waveforms

Figure 4.3: Three component synthetic signals and their time aligned multiplexed forms. East, north and vertical components shown in red, blue and black colors respectively. Events are due to a DC source mechanism.



(a) Three component synthetic events



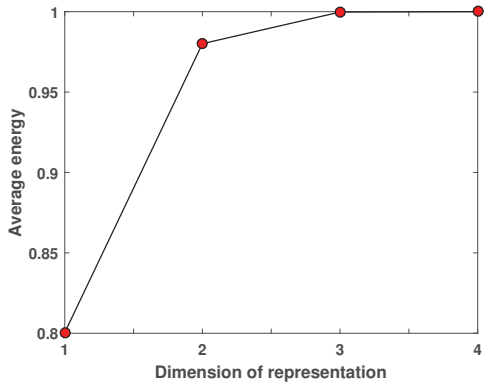
(b) Aligned waveforms

Figure 4.4: Three component synthetic signals and their time aligned multiplexed forms. East, north and vertical components are shown in red, blue and black colors respectively. Events are due to DC and CLVD source mechanisms.

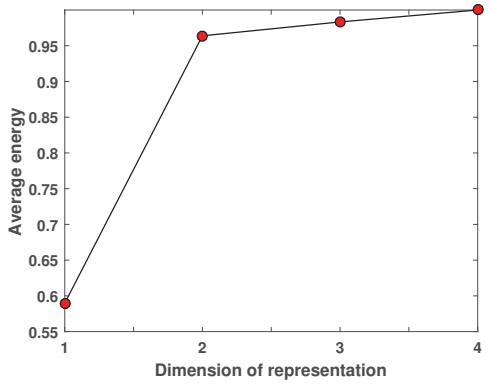
A singular value decomposition was then applied to the aligned waveforms to obtain the left singular vectors. To quantify the similarity between the singular vectors and template events, correlation coefficients between the synthetic waveforms and singular vectors were calculated on the P- and S-waves separately. The P-wave polarizations were also used as an additional constraint to define similarity. Eq 3.12 was applied on a 60msec window centered around the P-wave of the singular vector to obtain polarization estimates, and these were compared to the true back-azimuths of the templates.

Results

The average energy capture as a function of the subspace dimension of representation was calculated for both cases and shown in figure 4.5a.



(a) Similar events

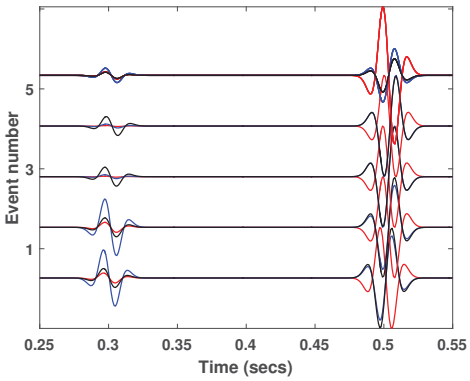


(b) Different events

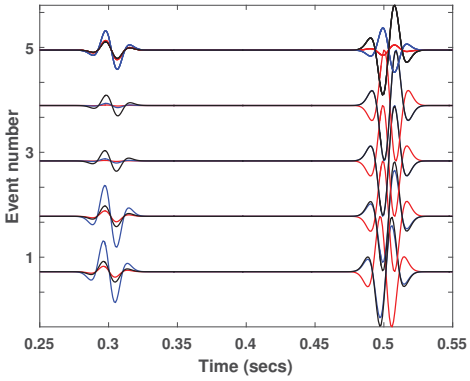
Figure 4.5: Average energy capture as a function of the dimension of representation.

The majority of the signal energy was described by the first two singular vectors in both cases. When the events were very similar, the first singular vector described a significant amount of the total signal energy ($\approx 80\%$), with a minor contribution from the second singular vector ($\approx 18\%$ of the total energy). The contribution of the second singular vector in the energy capture increased when the waveforms differed from each other, jumping to $\approx 35\%$ of the total signal energy, while the influence of the first singular vector diminished, as it only managed to describe $\approx 60\%$ of the signal energy in that case. The implications of the different energy capture trends on the singular vectors may

be understood by observing the singular vectors themselves in relation to their templates. Figures 4.6 and 4.7 show the 3C templates and their first two singular vectors.

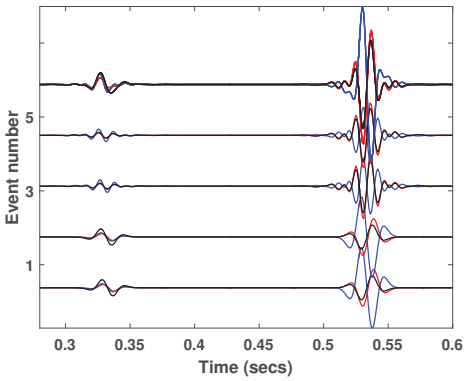


(a) Template events and their 1^{st} left singular vector

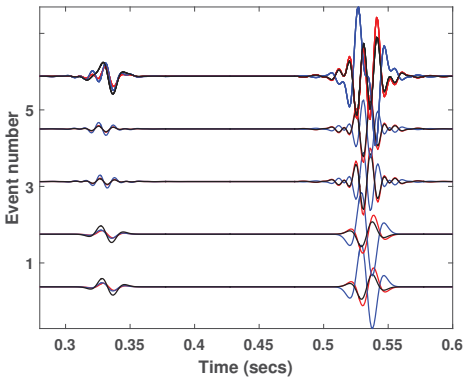


(b) Template events and their 2^{nd} left singular vector

Figure 4.6: Template events and their 1^{st} and 2^{nd} left singular vectors in three-component form. The singular vectors correspond to event number 5 (top waveform) in both figures. East, north and vertical components are shown in red, blue and black colors respectively.



(a) Template events and their 1^{st} left singular vector



(b) Template events and their 2^{nd} left singular vector

Figure 4.7: Template events and their 1^{st} and 2^{nd} left singular vectors in three-component form. The singular vectors correspond to event number 5 (top waveform) in both figures. East, north and vertical components are shown in red, blue and black colors respectively.

Both singular vectors were very similar to their templates when the source mechanism was similar, with the main differences seen in the waveform amplitudes. When waveform variability became more pronounced, as was the case with the two different source mechanisms, the singular vectors appeared to contain characteristics of the waveforms in the design set to varying degrees, with none of the singular vectors describing any of the waveforms uniquely. These differences between the singular vectors and their templates were quantified by examining their similarity. Tables 4.1 and 4.2 show the correlations and back-azimuth values.

Waveform	Back-azimuth (θ)	CC1 (P)	CC1 (S)	CC2 (P)	CC2 (S)
1 (Red)	26.31	0.99	0.98	0.99	0.92
2 (Red)	26.36	0.99	0.98	0.99	0.92
3 (Green)	26.41	0.99	0.98	0.99	0.79
4 (Green)	26.46	0.99	0.98	0.99	0.79
Vector 1	26.35	-	-	-	-
Vector 2	26.33	-	-	-	-

Table 4.1: Back-azimuths and correlation coefficients between the P-waves (P) and S-waves (S) of the events and their first two left singular vector. CC1 and CC2 correspond to correlations with the 1^{st} and 2^{nd} left singular vectors respectively. The events are all due to a DC source mechanism.

Waveform	Back-azimuth (θ)	CC1 (P)	CC1 (S)	CC2 (P)	CC2 (S)
1 (Red)	26.31	0.87	0.85	0.84	0.57
2 (Red)	26.36	0.87	0.85	0.84	0.57
3 (Green)	27.06	0.67	0.85	0.68	0.58
4 (Green)	27.11	0.67	0.85	0.68	0.58
Vector 1	33.10	-	-	-	-
Vector 2	38.18	-	-	-	-

Table 4.2: Back-azimuths and correlation coefficients between the P-waves (P) and S-waves (S) of the events and their first two left singular vector. CC1 and CC2 correspond to correlations with the 1^{st} and 2^{nd} left singular vectors respectively. The events are due to DC and CLVD source mechanisms.

For the very similar events, the overall correlations with their singular vectors were high, with back-azimuth estimates that were fairly similar between the templates and singular vectors. Both singular vectors in this case had back-azimuth estimates that were fairly close to the average value of the templates ($\theta_{avg} = 26.39^0$). The S-waves correlations were where the main differences appeared. Their correlations with the first singular vector were consistently high

across all the templates, whereas there was a dip in similarity between the second singular vector and the events representing the variation of the DC source. Nevertheless, the overall correlations were still good.

The correlations between the singular vectors and templates displayed moderate variability when the waveforms differed from each other. The first singular vector still maintained fairly high correlations with the templates overall, the exception being on the P-wave of the CLVD events. The second singular vector displayed poor correlations with all the templates, most notably on the S-waves. The back-azimuth estimates also differed between the singular vectors and their events, with estimates for both singular vectors lying outside the range of the template values. The results suggest the first singular vector adequately captures the dominant characteristics of the design set, as long as the templates are very similar to each other. Variability of the templates results in waveform distortion of the singular vectors, which leads to errors in the back-azimuth estimates.

4.4.2 Proximity of singular vectors to their templates

The success of any relocation technique is dependent on the relative distances of the events being relocated to each other or the master event. If the ray paths between the pair of events to a common receiver are assumed to be similar, then differences in travel time can be attributed to velocity variation in the source region between the events. If subspace relocation is to be successful, the hypocenter location of the singular vectors must be fairly close to the events being relocated to achieve accurate results. In this test, the location of the singular vector relative to the templates in the design set for different acquisition geometries is examined.

Methodology

Ten synthetic microseismic events were generated in a homogeneous velocity medium and used as templates in the subspace construction. The acquisition geometries utilized in these tests were a down-hole array consisting of five

boreholes, with each borehole containing four receivers, and a vertical twelve receiver array. The sampling rate (dt) was set to 1ms and the source mechanism of the events was a double couple (DC) with no attenuation. A Ricker wavelet was used as the source wavelet. We assumed P-wave and S-wave velocities of $V_p = 4000m/s$ and $V_s = 2300m/s$ respectively. The origin time t_0 was set to 0, the source locations, arrival times and back-azimuths of the templates to the receiver stations are perfectly known. No noise was added to the system. The events were assumed to be highly similar to each other, with some events occurring at the same location. The array configurations are shown in figure 4.8.

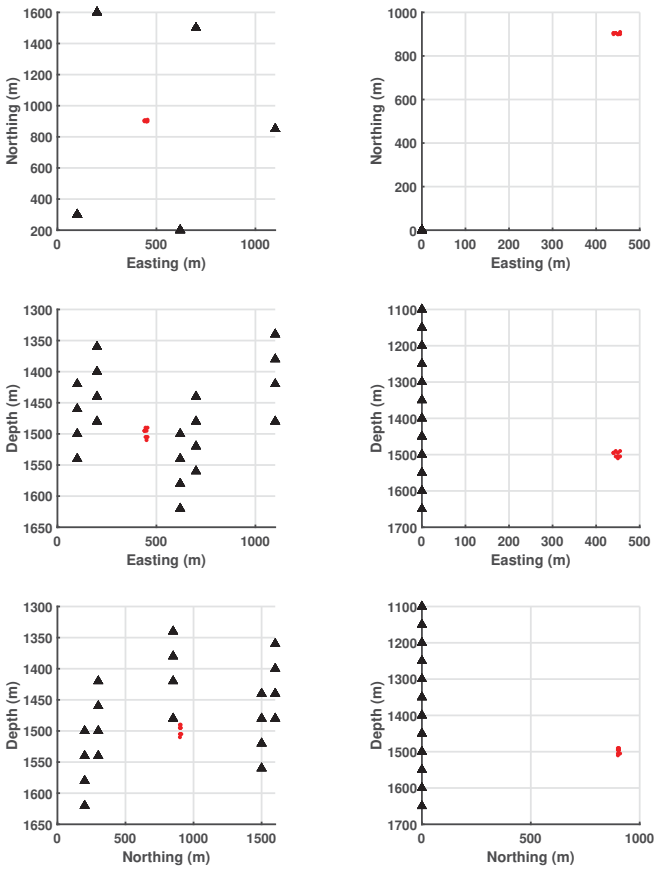
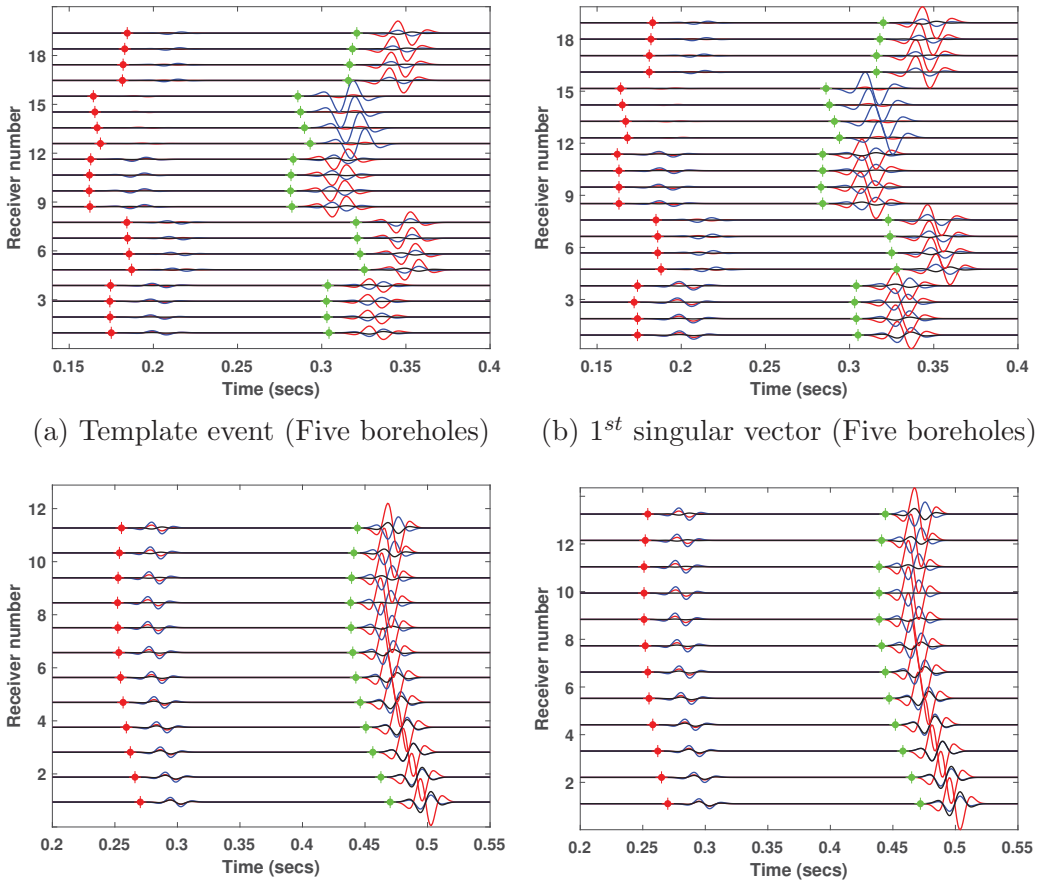


Figure 4.8: Location of synthetic microseismic events relative to receiver arrays. *Left column*: Array configuration for five boreholes setup. *Right column*: Array configuration for single vertical borehole. Receivers shown as black triangles and events shown as red dots. The source locations are the same in both cases.

For each acquisition geometry, the left singular vector representations were

obtained at each receiver using the procedure outlined in sections 2.3 and 4.2, and the first left singular vector was selected as a master event. Arrival times were then picked on the singular vector at each receiver, as shown in figure 4.9.



(a) Template event (Five boreholes)

(b) 1^{st} singular vector (Five boreholes)

(c) Template event (Single borehole)

(d) 1^{st} singular vector (Single borehole)

Figure 4.9: A template event and the corresponding 1^{st} left singular vector representation. The travel time picks for the P and S-waves are indicated by red and green dots. The true travel time estimates are shown for the template event while the picked estimates are shown for the singular vectors. East, north and vertical components are shown in red, blue and black colors respectively

For the five borehole array, only the arrival times were used in the location procedure, while the P-wave polarizations were included to constrain the horizontal location for the single vertical array. The polarization estimates for the singular vector were obtained by extracting 60ms windows centered around the P-waves, starting from the picked arrival time estimates at each receiver. The absolute locations of the singular vectors were then estimated using a grid

search approach. The grid volume was defined as 100m × 100m × 100m cube in x, y and z respectively, with a 1m spacing between grid nodes. The maximum likelihood hypocenter for the 1^{st} singular vector was then obtained via eq 3.18.

Results

Figure 4.10 shows a close up view of the microseismic source locations of the templates in relation to their 1^{st} left singular vector.

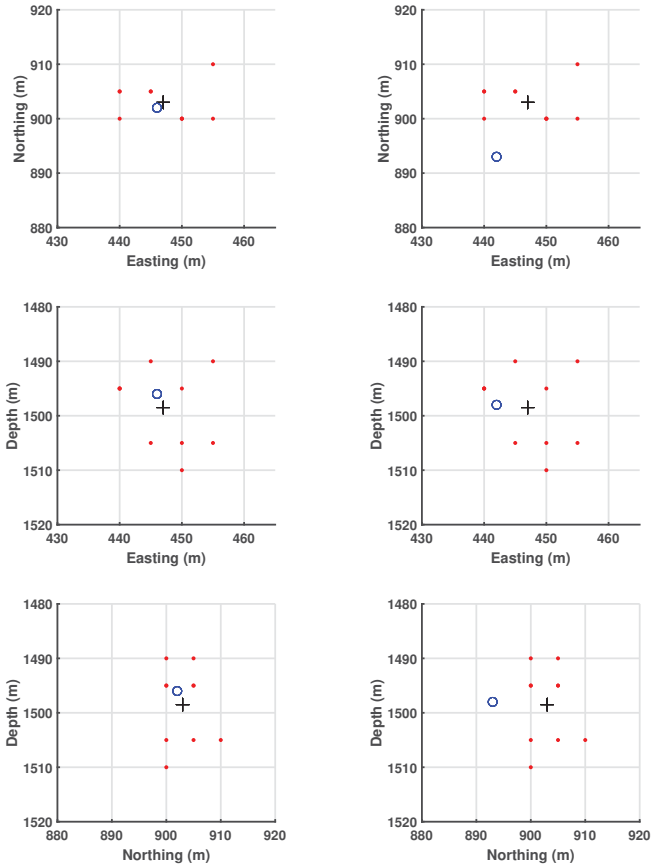
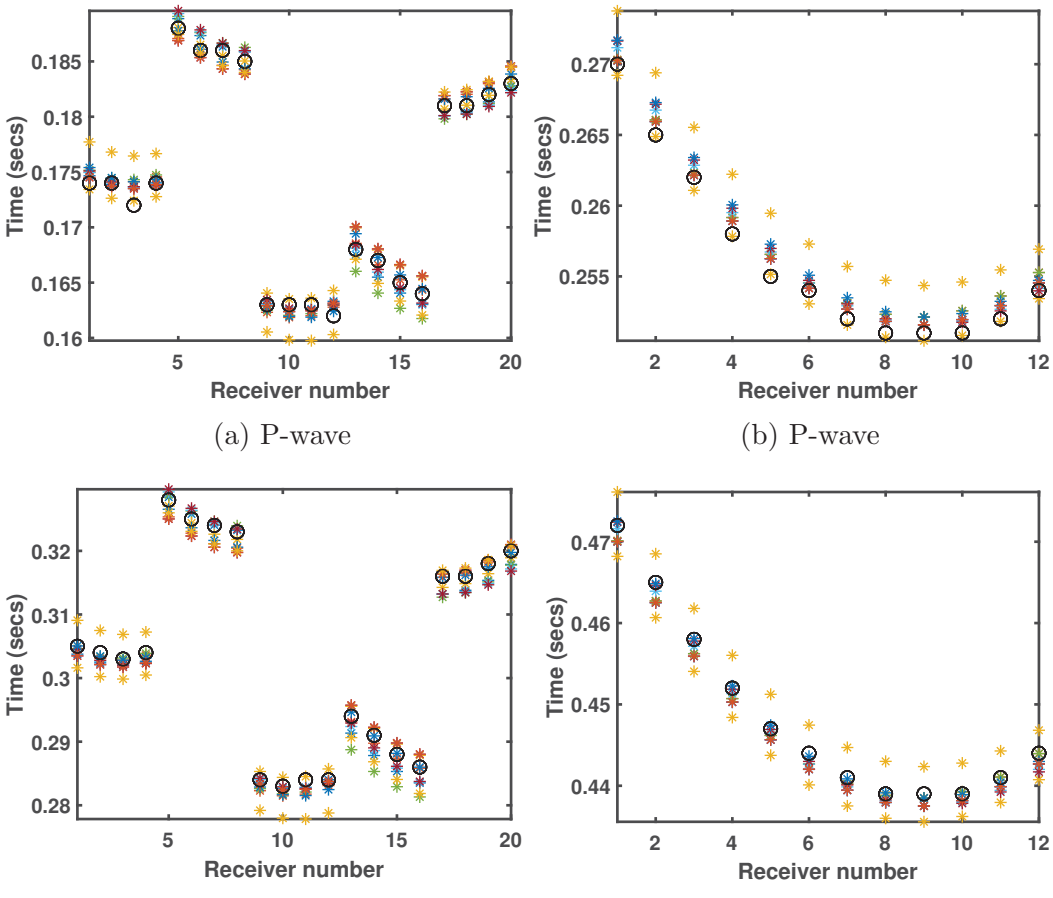


Figure 4.10: Location of synthetic microseismic events relative to their 1^{st} left singular vector for two acquisition geometries. *Left column*: Five boreholes. *Right column*: Single vertical borehole. Source locations shown as red dots, while the singular vector location is denoted by a blue circle. The cluster centroid is highlighted by a black cross.

In both cases the singular vectors were proximal to the template events, with a maximum observed separation distance between the events and singular

vector of 25m across both setups. Its location was situated close to the cluster centroid in the five borehole steps, with a slightly larger offset in the case of the single vertical borehole. The depth estimates from both acquisition styles were both fairly close to the cluster centroid, with the main deviations seen in the epicentral estimates. These observations may be understood by observing the distribution of the travel times and back-azimuth estimates, shown in figures 4.11 and 4.12.



(a) P-wave (b) P-wave

(c) S-wave (d) S-wave

Figure 4.11: Distribution of observed travel times for template events and their 1^{st} singular vector for two acquisition geometries. *Left column*: Five boreholes. *Right column*: Single vertical borehole. The travel times of the template events are shown as stars while black circles denote the singular vector travel times.

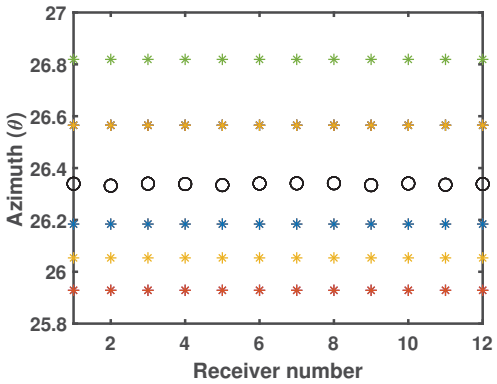


Figure 4.12: Distribution of the back-azimuths for template events and their 1^{st} singular vector. The back-azimuths of the template events are shown as stars while black circles denote the singular vector back-azimuths.

The travel time estimates for the singular vectors were very similar to the templates in both acquisition setups. The back-azimuth estimates of the singular vector were also within the range of the template values, and very closely approximated the average azimuth for all the templates. The high similarity of the travel time and polarization estimates imply the hypocenter location of the events should lie close to the cluster, which was the case in this example. Based on these results, the differences in the singular vector location between both acquisition could be attributed to the parameters used in location. Only travel times were used when the azimuthal coverage was decent, but polarization estimates had to be included when constraints were needed on the horizontal position of the vector for the single vertical borehole.

4.4.3 Comparison to other relative location techniques

The relative relocation capabilities of the subspace detector may be analyzed by comparing it to established relative location routines. In this test, the subspace relocation method was compared to both the conventional double difference and master event relocation techniques.

Methodology

The ten synthetic microseismic events generated in section 4.4.2 were used as templates in this experiment, with the same source locations and seismic velocities used in both acquisition geometries. The events were also assumed to occur at varying times i.e have different origin times. The left singular vector representations were obtained at each receiver using the procedure outlined in sections 2.3 and 4.2, and the 1^{st} singular vector was selected as a master event for subspace relocation. For the master event relocation technique, a template event closest to all the other events was selected as a master event. Next, random Gaussian noise with zero mean and standard deviations of 6 ms and 5^0 respectively were added to the synthetic event arrival times and back-azimuths for the events. No noise was added to the singular vector absolute arrival times and back-azimuths since it was assumed these values were more accurate relative to the events. The same noise was added to the absolute data across all three methods to ensure the initial source locations were the same.

Initial source locations for the events and singular vector were obtained using a grid search, with the same grid spacing outlined in section 4.4.2. With these source locations, initial origin time estimates were obtained by subtracting the predicted travel times from the observed arrival times and averaging the vector of origin times. The relative travel times were also contaminated with random zero mean Gaussian noise. To simulate a real world scenario where these relative times were obtained from waveform cross-correlation which are usually an order of magnitude more accurate than absolute data, the standard deviation of the noise added was set to 1 ms. The relative back-azimuth data are usually more accurate than absolute back-azimuths as common path errors and errors due to deviation of the borehole are minimized. Therefore, the standard deviation of the added random zero mean Gaussian noise was set to 1^0.

The initial source locations from the grid search, absolute data and differential data were used as inputs to the double difference algorithm. Eq. 3.42 was solved iteratively using least squares, and the solution vector was used to update the initial source locations. The absolute data was given a relative

weight of 0.1 in the relocation procedure for the double difference relocation. For the master event relocation technique, the initial source location of the master event was obtained from the grid search, and was used as an input to the algorithm in addition to the differential data. The same process was applied to subspace relocation, using the singular vector location and differential data as inputs to the algorithm. The location of the singular vector was used to obtained predicted travel times, which were subtracted from the observed arrival times to obtain an estimate for its origin time. For both the master event and subspace relocation techniques, eq 3.32 was solved, and the solution vector of hypocentral deviations was added to the location of the master event and singular vector respectively.

Results

Initial source locations for the synthetic events recorded by the five borehole receiver array are shown in figure 4.13.

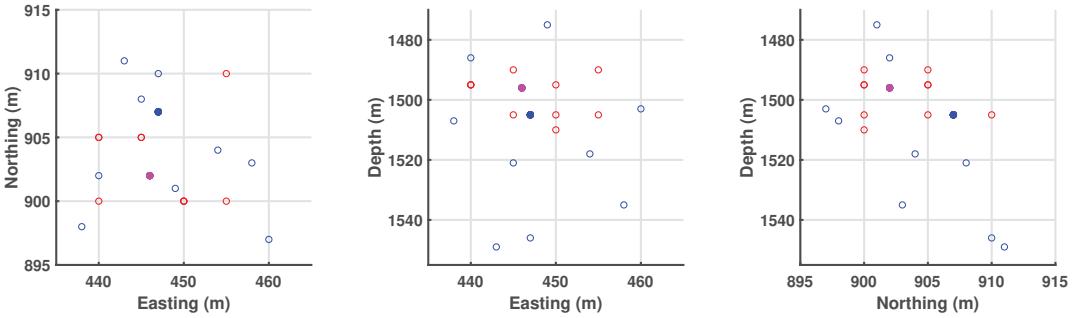


Figure 4.13: Source locations of the synthetic events recorded by a five borehole array. Red and blue open circles indicate true and initial source locations from grid search respectively. Blue and magenta closed circles highlight the location of the selected master event and singular vector respectively.

The origin time estimates obtained for these events are also shown in table 4.3

Event	True (s)	Initial (s)
1	10	9.2
2	5	4.4
3	2	0.8
4	20	20.3
5	30	29.1
6	40	38.6
7	6	6.6
8	100	98.9
9	3	2
10	4	5.1

Table 4.3: Table of origin times for the synthetic microseismic events before relocation for the five boreholes setup.

The estimated origin time for the singular vector in this case was ≈ 0 so no constant was added to the singular vector arrival times. The initial source locations were fairly spread out, with a major distortion in the shape of the cluster apparent. In particular, the errors appear to be more pronounced in depth. Relocation was performed using all three relative location techniques, and the results are shown in figure 4.14.

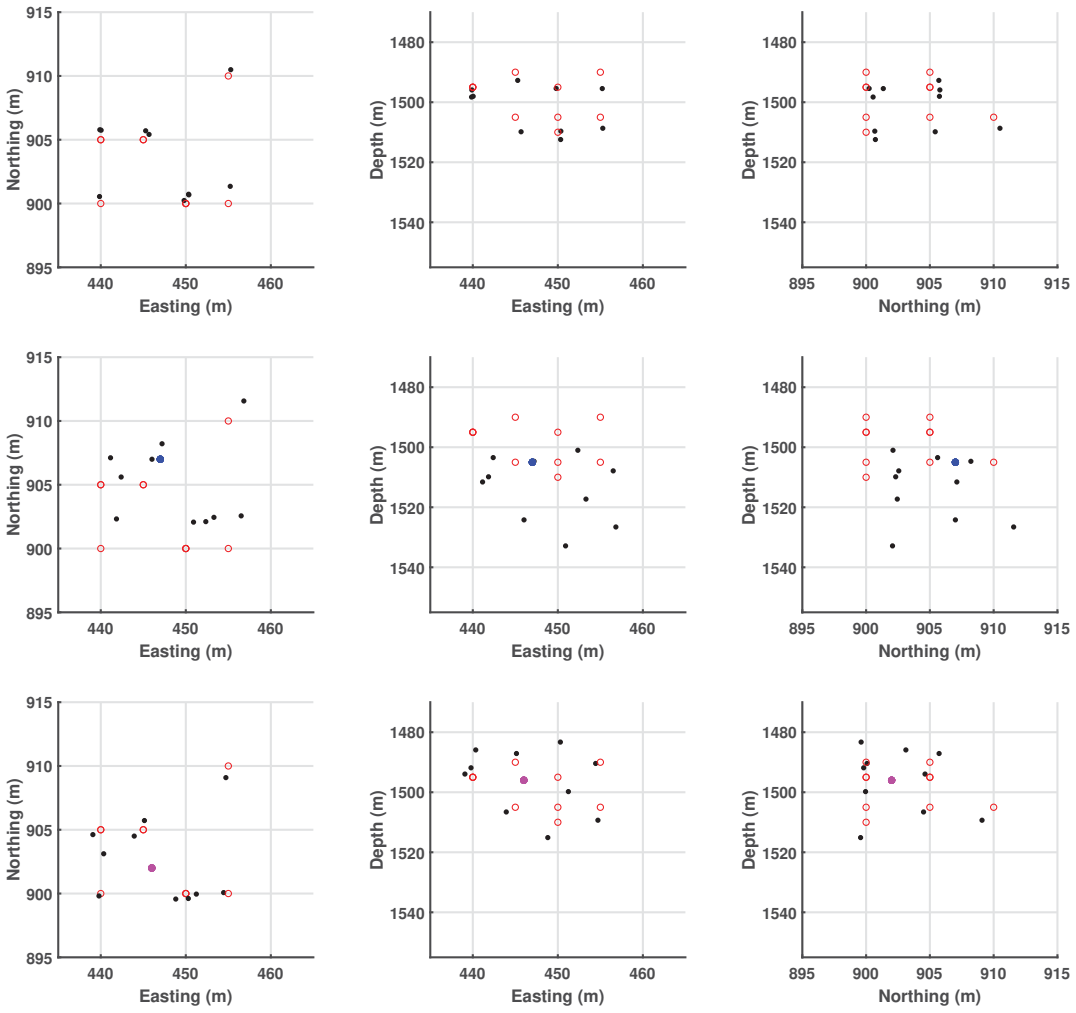


Figure 4.14: Source locations of the synthetic events recorded by a five borehole array. *Top Row:* Double difference relocations. *Middle Row:* Master event relocations. *Bottom Row:* Subspace relocations. Red open circles indicate true source locations, while black closed circles represent the relocations. Blue and magenta closed circles highlight the location of the master event and singular vector respectively.

Overall, the relocations were a lot closer to their true positions compared to the initial grid search locations, with the relative distances between the events somewhat similar across the methods. The double difference relocations were the closest in absolute location to the true source locations, and had smaller relative distances between the events relative to the other techniques. In both the master event and subspace relocation techniques, the events were pulled closer to their respective masters. The relocations were also very similar between the

master event technique and subspace relocation, with the main differences seen in the absolute location of the clusters. Overall, the cluster shapes were better approximated after relocation compared to the grid search locations. The true cluster shape is decently captured by the relocations, but some stretching and rotation is apparent in the relocations, leading to a distortion of the seismic image. The origin time estimates after relocation are shown in table 4.4.

Event	True (s)	Initial (s)	DD (s)	MER (s)	SR (s)
1	10	9.2	10	9.4	10
2	5	4.4	5	4.4	5
3	2	0.8	2	1.4	2
4	20	20.3	20	19.4	20
5	30	29.1	30	29.4	30
6	40	38.6	40	39.4	40
7	6	6.6	6	5.4	6
8	100	98.9	100	99.4	100
9	3	2	3	2.4	3
10	4	5.1	4	3.4	4

Table 4.4: Table of origin times for the synthetic microseismic events after relocation for the five borehole setup. All initial origin times for the events were used in the double difference technique. The initial origin time of event 2 was selected as the master event origin time for the MER technique while the origin time estimated for the singular vector (≈ 0) was used in the SR technique.

The true origin time estimates were recovered using the double difference technique, further highlighting this method's robustness. The origin times obtained using the master event relocation technique were all systematically shifted from their true values by 0.6s, the same shift observed between the true and initial estimates for the master event (event 2). Subspace relocation captured the true origin times, suggesting the use of accurate absolute data minimizes the error in the origin time estimates after relocation in this method.

Application of the relocation techniques to the single vertical borehole array also displayed similar results. Figure 4.15 shows the initial source locations for the synthetic events for this acquisition geometry.

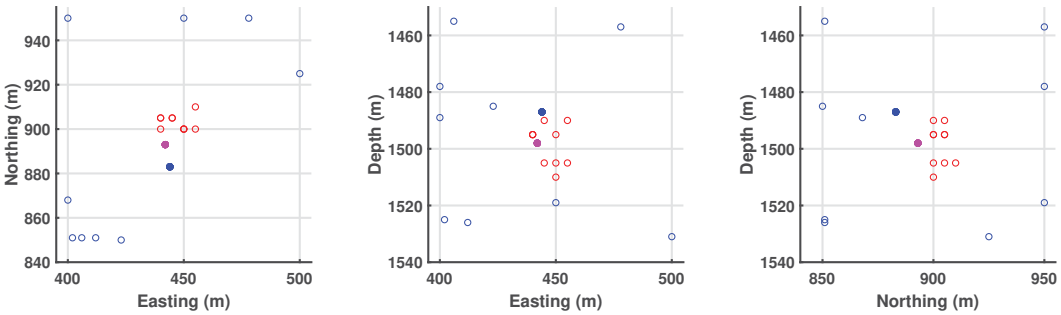


Figure 4.15: Source locations of the synthetic events recorded by a single borehole array. Red and blue open circles indicate true and initial source locations from grid search respectively. Blue and magenta closed circles highlight the location of the selected master event and singular vector respectively.

Event	True (s)	Initial (s)
1	10	9.4
2	5	4.6
3	2	1.1
4	20	19.7
5	30	29.7
6	40	39.9
7	6	5.8
8	100	99.5
9	3	2.2
10	4	4.6

Table 4.5: Table of origin times for the synthetic microseismic events before relocation for the single borehole setup.

The inclusion of azimuthal data increased the uncertainty in the system, leading to an increase in the location error. The relocation results for this

scenario are shown in figure 4.16.

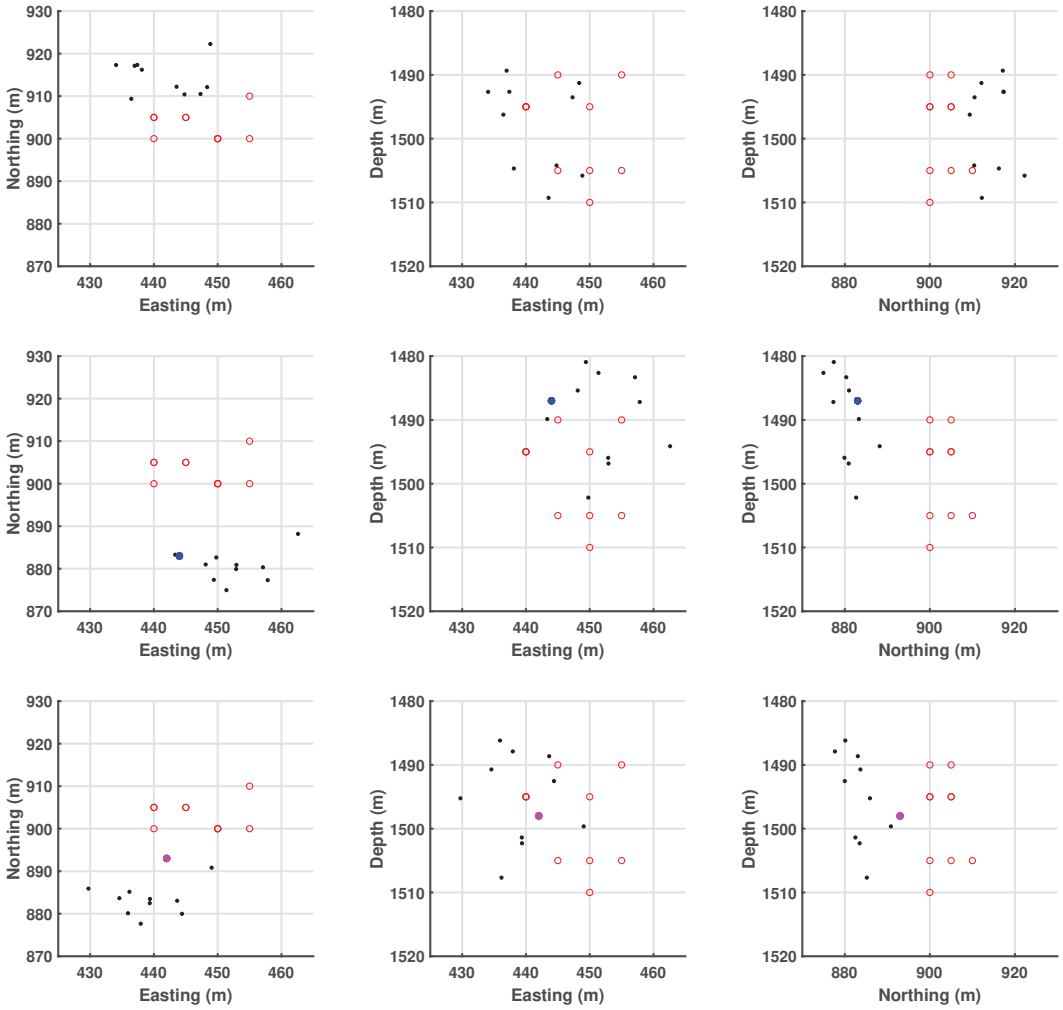


Figure 4.16: Source locations of the synthetic events recorded by a single borehole array. *Top Row:* Double difference relocations. *Middle Row:* Master event relocations. *Bottom Row:* Subspace relocations. Red open circles indicate true source locations, while black closed circles represent the relocations. Blue and magenta closed circles highlight the location of the master event and singular vector respectively.

Again, the absolute locations of the relocated cluster were closer to the true locations compared to the initial source locations across the different techniques. The distortion of the cluster shape is more pronounced for this acquisition geometry compared to the five bore hole setup. The double difference relocations were the closest approximations to the cluster shape compared to

the techniques, but there wasn't a significant difference in the relative distances between the events after relocation. The master event and subspace relocation results were more rotated and stretched compared to the double difference, but still retained elements of the original cluster orientation. The origin time estimates after relocation for this acquisition setup displayed the same trends observed in the five borehole setup, and are shown in table 4.6.

Event	True (s)	Initial (s)	DD (s)	MER (s)	SR (s)
1	10	9.4	10	9.7	10
2	5	4.6	5	4.7	5
3	2	1.1	2	1.7	2
4	20	19.7	20	19.7	20
5	30	29.7	30	29.7	30
6	40	39.9	40	39.7	40
7	6	5.8	6	5.7	6
8	100	99.5	100	99.7	100
9	3	2.2	3	2.7	3
10	4	4.6	4	3.7	4

Table 4.6: Table of origin times for the synthetic microseismic events after relocation for the single borehole setup. All initial origin times for the events were used in the double difference technique. The initial origin time of event 4 was selected as the master event origin time for the MER technique while the origin time estimated for the singular vector (≈ 0) was used in the SR technique.

Again the double difference and subspace relocation techniques accurately recovered the true origin times, while the master event results showed a systematic shift equal to the shift between the true value and utilized value of the master event.

4.5 Discussion

The dominant characteristics of a design set are represented adequately by a small number of singular vectors, but the relative importance of each vector depend on the extent of waveform similarity. In the situation where waveform similarity was quite high, the 1^{st} singular vector represented the majority of the signal energy, with minor contributions from the subsequent vectors. It also exhibited large similarity to its templates in waveform character, which was reflected in travel time and polarization information. A look at figure 4.6 shows no discernible lags between the singular vector and its template, suggesting proximal locations, and the similarity of the back-azimuth estimates further support this argument. The 2^{nd} singular vector in this experiment also exhibited the same traits as the 1^{st} singular vector, but its relatively minor contribution to the description of the signal energy implies its more likely to contain noise energy, which will lead to less than optimal results. The 1^{st} singular vector still described the dominant characteristics of the design set when waveform variability becomes more significant, but the approximations do not describe the waveform types uniquely. These lead to waveform characteristics of the singular vectors that did not appear to have any physical meaning since they represented a blend of different waveform types. This phenomenon translates to errors in travel times, a fact that is illustrated in figure 4.17, where only the P-wave components of one of the DC and CLVD events are shown, as well as their 1^{st} singular vector.

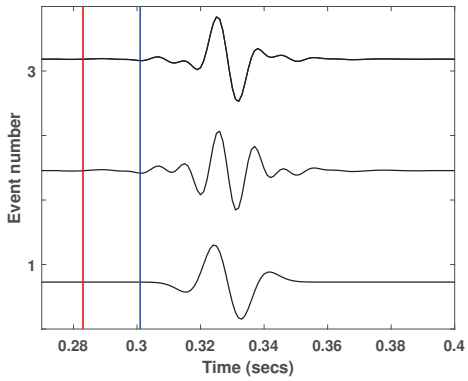
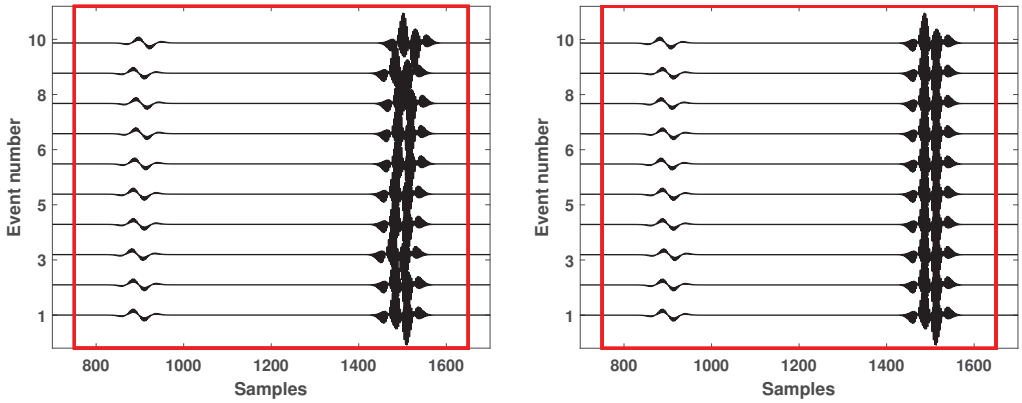


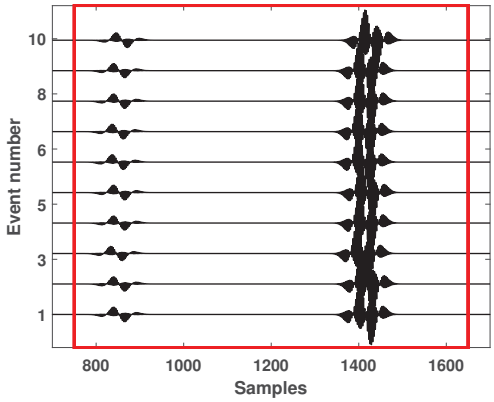
Figure 4.17: DC source (event 1), CLVD source (event 2) and their first left singular vector (event 3) after alignment. The arrival time of the DC source is shown as a blue line while the arrival time of the CLVD source is shown as a red line. Only the vertical component is shown.

The CLVD event had a slightly longer duration than the DC event and after alignment, the earliest arrival time was seen on the CLVD event in figures 4.4 and 4.4. The singular vector had an identical arrival time to the CLVD event, suggesting the singular vectors are sensitive to the earliest arriving waveform. The back-azimuth estimates of the singular vectors were also erroneous, lying outside the range of the template events. The similarity of the relative moveout times across the receiver array between the singular vectors and templates in both acquisition geometries suggests the methodology employed to get the singular vector representation is accurate. Using a fixed window during the alignment of the waveforms allows for consistency across the receiver array. This is illustrated in figure 4.18 which shows multiplexed waveforms before and after alignment for two different receivers.

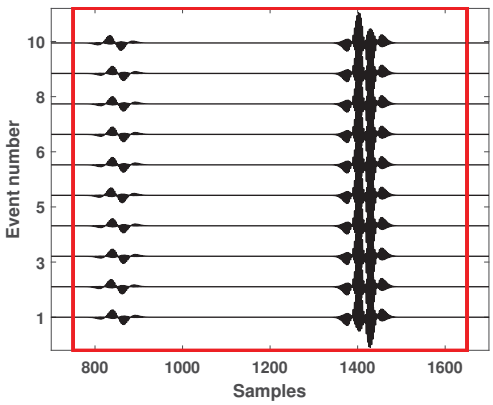


(a) Receiver 1 (Unaligned)

(b) Receiver 1 (Aligned)



(c) Receiver 12 (Unaligned)



(d) Receiver 12 (Aligned)

Figure 4.18: Multiplexed synthetic waveforms before and after alignment from two different receivers. Red box indicates fixed temporal window used for obtaining singular vector representation.

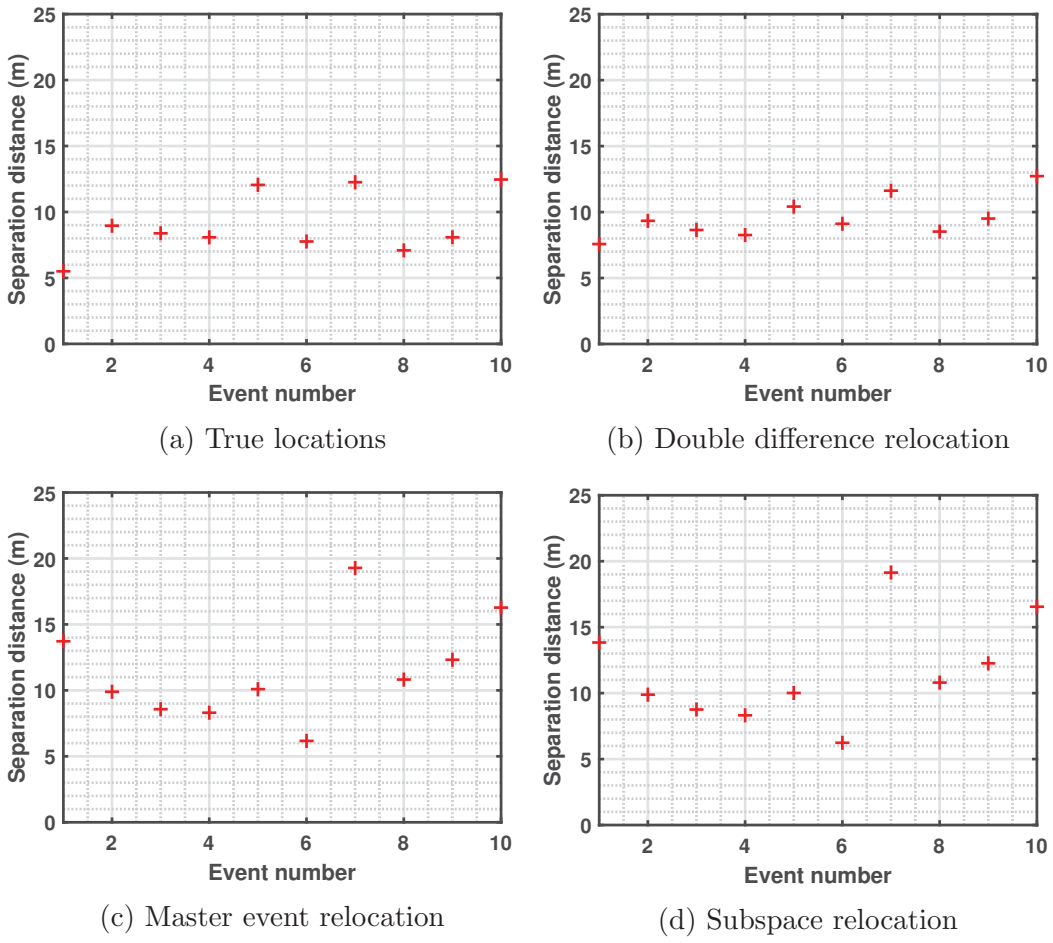
Before alignment there are relative time differences observed between the receivers since the travel times to the receivers differ between events. This relative time difference was roughly maintained after alignment. The waveforms are shifted predominantly around the average position of the events, and the use of a combined window including P- and S-waves has a negligible effect on the alignments due to the very similar S-P differential times. This is further supported by figure 4.11 which shows the observed travel times of the singular vectors roughly follow those of the templates, and lie within the template values. The location of the singular vectors were fairly similar, especially in the depth estimates. Azimuthal information was included for the single vertical

borehole which is the main constraint on the horizontal position of an event, and that explains why the main differences are seen in the epicenter estimates between both setups. Nevertheless, the differences were fairly minor and the estimated hypocenters were still within the general vicinity of the templates.

Application of the relocation techniques to uncertain arrival time measurements showed an improvement in absolute locations of the relocated cluster relative to the initial grid search locations. In all three cases, the cluster shape was fairly well estimated, but some stretching and rotation was evident, leading to a distortion of the cluster shape. The double difference technique produced the most robust results, as their relocations were not only closer to the true source locations, but were subject to less distortion compared to both the master event relocation technique and subspace relocation. In both the master event and subspace relocation techniques, the relocated events were pulled closer to the position of the master event and singular vector respectively. These two techniques displayed very similar relative distances amongst the event, with the main difference seen in the absolute location of the clusters. The singular vector was assumed to be more accurately located since it had higher signal-to-noise ratios relative to the other events. Subsequently, its location was much closer to the true source locations. The origin time information were overall fairly close to the true value, but the master event method showed some systematic bias due to error in the estimate for the master. The double difference technique was robust enough to accurately obtain the true values.

Computational costs are an important factor to consider since they determine the relative practicality of the relocation techniques. For N events recorded at k receivers, the differential data needed by the double difference algorithm will be greater than both the master event and subspace relocation methods by a factor of $0.5k(N-1)$. The inclusion of absolute data further increase the number of data measurements by Nk. Coupled with the iterative method through which a solution is obtained, and the need for data damping due to ill conditioned systems, the double difference approach method may incur some significantly processing costs due to the large system of observations

to be solved. Therefore, it is important to quantify just how successful the relocation results are relative to each other. The accuracy of the relative event locations may be quantified by comparing the distances of the events from their centroid. Figures 4.19 and 4.20 show the observed centroid-event separation distances for both acquisition setups.



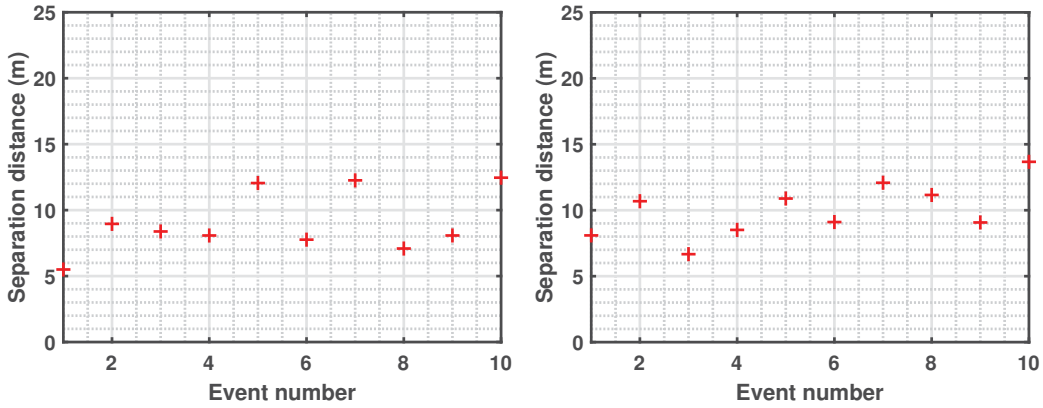
(a) True locations

(b) Double difference relocation

(c) Master event relocation

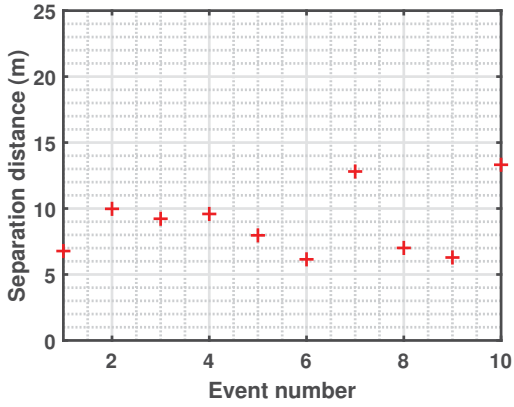
(d) Subspace relocation

Figure 4.19: Separation distance between the events and their cluster centroid for the five borehole acquisition geometry.

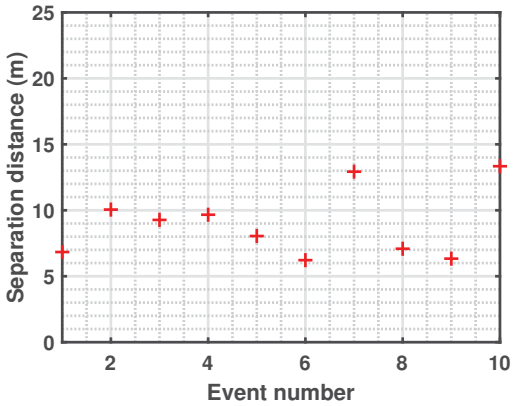


(a) True locations

(b) Double difference relocation



(c) Master event relocation



(d) Subspace relocation

Figure 4.20: Separation distance between the events and their cluster centroid for the single vertical borehole acquisition geometry.

The overall cluster centroid separation distances between the events were similar for both the true locations and the relocations, with the double difference centroid distances displaying more consistency across both setups. However, these differences are fairly small, which imply similar accuracy in relocations across the relocation techniques.

4.6 Summary

The subspace detector has been examined in terms of the hypocentral characteristics of the left singular vectors in relation to the templates constituting the subspace. The results suggest the 1^{st} left singular vector sufficiently rep-

resents the hypocentral characteristics of the design set in terms of waveform character, polarization and arrival time. This representation is highly dependent on high similarity amongst the templates, and becomes erroneous when the source mechanisms become more diverse. The methodology for performing subspace relocation was described and validated by comparing the similarity of arrival time and back-azimuth measurements between singular vectors and their templates. The hypocenters of the singular vectors were fairly consistent between both acquisition geometries, and were fairly close to their respective templates positions. The main differences were due to the inclusion of polarization information in the single vertical borehole case.

Application of the relocation techniques to noisy data showed the double-difference approach provided the most robust results. The subspace relocation method had very similar results to the master event technique, with the main difference occurring in the absolute location of the cluster. Unlike the standard master event relocation procedure which utilizes the template with the best signal-to-noise ratio (SNR), the 1^{st} left singular vectors represent a $\sqrt{n}$ increase in SNR (n = number of templates) since it closely approximates the stack of the templates. This implies an increase in location accuracy of the cluster due to both more accurate hypocenter locations for the singular vector, and higher correlations with the templates. The overall differences in the relative event distances were not too significant, suggesting subspace relocation may be a practical means to determine accurate relative locations at reasonable processing costs.

Chapter 5

Microseismic monitoring of a hydraulic fracture treatment

5.1 Introduction

The matched filter is a useful technique for detecting highly similar events, and offers increased sensitivity to waveforms at low false alarm rates (Gibbons and Ringdal, 2006). This method of event detection could often lead to increased processing times especially if a large number of template events are used, and may result in redundancy due to similar templates providing identical results. Energy detectors on the other hand, require no signal information and strictly detect events based on thresholds related to the ratio of signal to noise energy. This lack of signal sensitivity often leads to the omission of low magnitude events, and an increased false alarm rate when aggressive detection thresholds are set (Trnkoczy, 1999).

The redundancy in using similar matched filter templates may be alleviated by using low rank approximations of waveform groups in the detection procedure. The subspace detector provides an orthonormal representation of a set of waveforms using singular value decomposition, which highlight the common characteristics present amongst a set of waveforms (Harris, 2006; Barrett and Beroza, 2014; Freiberger, 1963). This representation offers increased sensitiv-

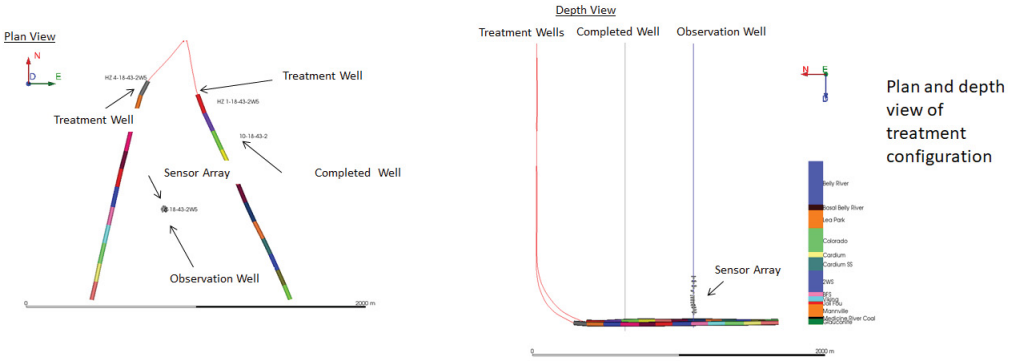
ity to waveform variation compared to the use of a single template and may increase detection capability at a lower processing cost. This enhanced sensitivity to waveform could also increase detection capability if template coverage is sufficient, leading to performances comparable to energy detectors at much lower false alarm rates.

The main aim of this chapter is to demonstrate the functionality of the subspace method, using data from a microseismic monitoring operation during a hydraulic fracturing treatment. First, the subspace detector is compared to the matched filter in terms of its efficiency and sensitivity to waveform. Next, the effects of the subspace design on its performance are investigated, and finally, its ability to increase catalog information compared to the conventional STA/LTA detector at low false alarm rates is analyzed and discussed.

5.2 Geological setting and field layout

The Hoadley field was discovered in 1977, and is predominantly a gas condensate field hosted by the Glauconite member of the Lower Cretaceous Mannville Group (Eaton et al., 2014; Newbert et al., 1987). The Mannville group comprise the oldest Cretaceous rocks for most of the Western Canada Sedimentary Basin (WCSB) and are indicative of a major episode of subsidence and sedimentation, which preceded a long period of uplift, exposure and erosion (Eaton et al., 2014; Hayes et al., 1994). Gas is produced from the lower Cretaceous Glauconite sand, which represents a southwest to northeast trending marine barrier bar system known as the Hoadley barrier complex (Newbert et al., 1987; Hayes et al., 1994; Eaton et al., 2014). It marks the northern limit of a continental to marginal-marine depositional environment (Hayes et al., 1994; Eaton et al., 2014). The lithology of the barrier bar varies from porous sand facies to interbar lagoon facies and tidal channel facies and the system contains progradational marine sandstone bodies up to 32km in length, each hosting several distinct reservoirs (Newbert et al., 1987). The hydraulic fracturing experiment occurred 58.6km northwest of Red deer in the Hoadley gas field of central Alberta, Canada (Eaton et al., 2014). The treatment zone consists

of interbedded tight sandstone reservoirs deposited in a wide range of near-shore environments (Reynolds et al., 2012). The site location and treatment configuration are shown in figure 5.1.



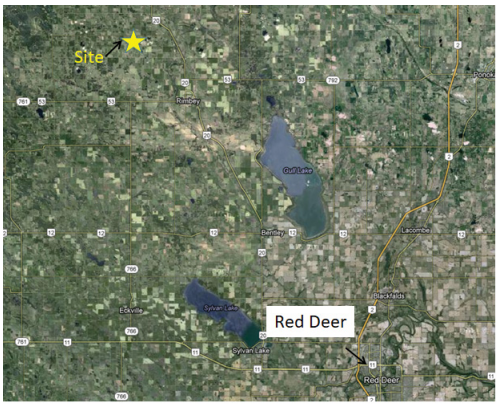


Figure 5.1: *Top*: Plan and depth view of treatment configuration. Treatment well and sensor array denoted by arrows. *Bottom*: Site of treatment: Treatment wells located approximately 58.6km northwest of Red Deer, AB and highlighted by yellow star.

5.3 Methodology

5.3.1 Pre-processing

Processing of data commenced with an analysis of the frequency content of the microseismic signals to suppress noise and improve the signal-to-noise ratio. Figure 5.2 displays a section of data containing a raw microseismic signal, its

spectrogram, and the resulting filtered signal. The spectrogram was obtained by computing discrete Fourier transforms on 0.0625s time windows of the microseismic signal, with 80% overlap (0.05s) between them. Most of the signal energy for the event is present within the 40-130Hz frequency band. Based on this information, a trapezoidal bandpass filter with corner frequencies [20 40 145 160] Hz was applied to the signal to suppress as much noise as possible. A notch filter was also applied to remove 60Hz frequencies corresponding to possible power lines in the area.

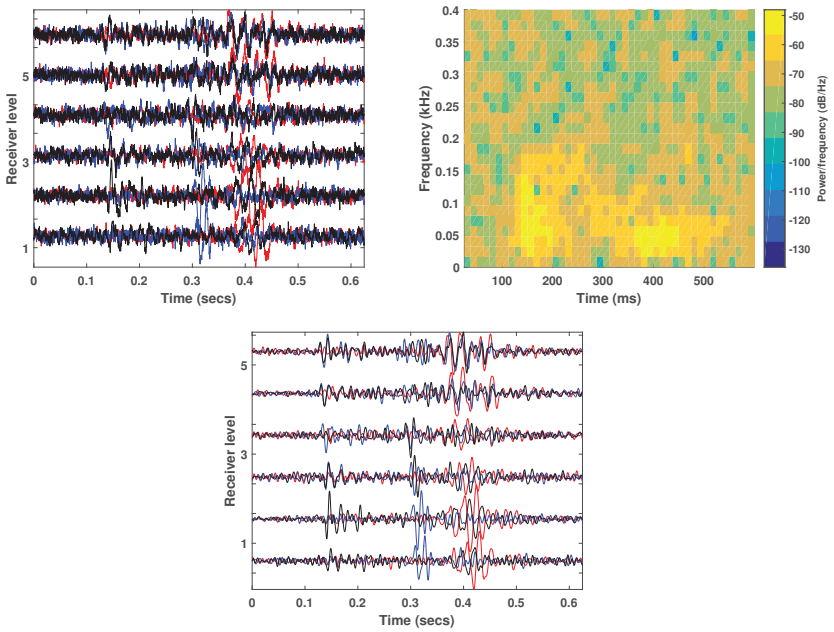


Figure 5.2: Time-frequency analysis of microseismic event. Signal energy is predominantly within the 40-130Hz frequency range. *Top left to bottom*: 1) Unfiltered microseismic signal. 2) Time-frequency representation (spectrogram). A blocky color scale has been used to emphasize the frequency band corresponding to signal energy. 3) Filtered microseismic event. Vertical, north and east components shown by red, blue and black colors respectively.

5.3.2 Construction of template library

Template waveforms used for both correlation based detectors were obtained with an STA/LTA detector. A conservative threshold value τ_{SL} of 10 was used to declare a detection at a receiver, and 8 out of 12 receivers were required to trigger before a global detection was declared to ensure good quality detec-

tions. STA and LTA windows were selected to be 4 (0.06s) and 9 times (0.15s) the dominant period of the microseismic signals respectively. Those window sizes were deemed large enough to adequately capture both signal and noise information, and short enough to avoid averaging signal energy. The detector was then run over the entire first day of data which corresponded to 7 hours of data, leading to a total of 517 detections made. The detection results were then inspected visually to determine good quality microseismic events. In this study, a microseismic event is defined as having visible and coherent seismic phases (P- and/or S-waves) and reasonably small S-P differential times. After inspection, incoherent signals and non-microseismic events were removed, leaving 506 templates for analysis.

5.3.3 Multiplet analysis

A multiplet is defined as a group of n events (n $\geq$ 2) where each event is very similar with at least one other event in the group (Arrowsmith and Eisner, 2006). The idea is events of similar origin will be grouped together as they tend to be very similar up to a distance of a quarter of their dominant wavelength, according to Geller and Mueller (1980). In this study, the multiplets are defined as groups for which the average cross-correlation coefficients is $\geq$ 0.8 i.e. every member of a group must have a correlation coefficient greater or equal to this value with at least one other member of the group. Due to the large number of template events obtained with the STA/LTA detector, this value is deemed sufficient as it is high enough to ensure high similarity between events in multiplet groups, but low enough to allow for some signal variability within multiplet groups, allowing for differentiation between multiplet groups. This similarity is measured by array-averaged time domain cross correlation functions between the events.

Starting 50ms before the P-wave arrival time estimates, a 500ms window encompassing both P- and S-waves was used to extract waveform data before calculating the peak correlation coefficient value between event pairs. Both phases were utilized in a combined window since similar waveforms have sim-

ilar P- and S-waves, making it more efficient to process both phases at once. Events were then linked together in a chainlike fashion as a function of the array-averaged weighted correlation coefficient known as single link clustering (Arrowsmith and Eisner, 2006). The weighting was done by calculating the correlation coefficient on each component of a geophone, then weighting it by the maximum data amplitude for the receiver before taking the average to emphasize components with higher signal-to-noise ratios. This procedure was applied to every event pair, and the correlation coefficient values were stored in a symmetric matrix (figure 5.3).

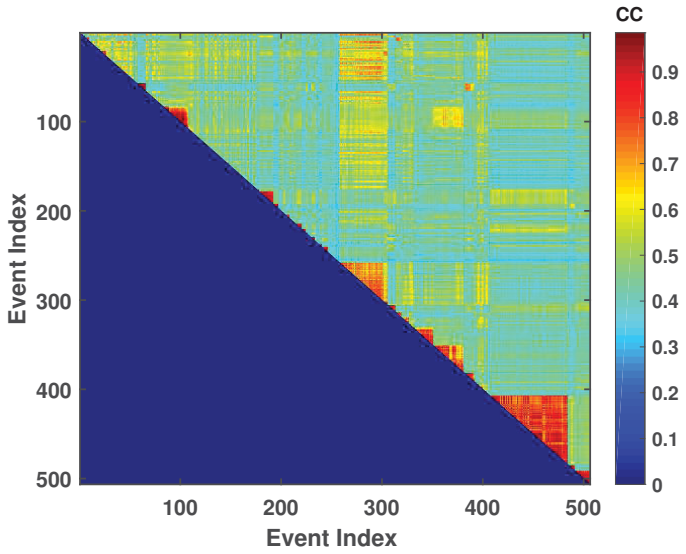


Figure 5.3: Matrix of correlation coefficient (CC) between 506 template microseismic events sorted into their respective multiplet groups. Every cell corresponds to the maximum correlation coefficient event between event pairs.

Separation of events into multiplet groups resulted in the formation of 201 groups, 35 of which contained multiple events. The multiplet groups comprised 340 events, with the remaining 166 groups forming single event clusters. In this study, 4 multiplet groups containing events from the target data section (one hour) were selected for analysis. For each multiplet group, a subspace detector was constructed. A summary of the selected multiplet groups is shown table 5.1. The templates contained in these multiplet groups predominantly spanned the mid to late stages of the chosen data segment.

Multiplet group	# of events in group
1	14
2	15
3	2
4	8

Table 5.1: Table of multiplet groups.

5.3.4 Subspace construction

After sorting the events into multiplet groups, waveforms in each of the selected multiplet groups were multiplexed and aligned in time using the iterative stacking procedure of (De Meersman et al., 2009). Starting 50ms before the P-wave arrival time estimates at each receiver, 330ms data windows encompassing the aligned multiplexed P- and S-waves were extracted for use in construction of the signal subspace. The data windows were kept as short as possible to minimize the amount of noise present in the subspace representation. The matrix of aligned events was then be decomposed into three unitary matrices via a singular value decomposition (SVD) (eq 2.16), and an orthonormal representation of the aligned events $\mathbf{U}$ was obtained from the matrix of left singular values.

Dimension of representation

Determining the optimal dimension of representation for the subspace is crucial in the detection procedure. A sufficient amount of signal energy must be captured by the subspace to adequately represent the signals in a group. A measure of the average energy capture in each multiplet group as a function of the subspace dimension of representation d was obtained by applying eq 2.19 to the aligned waveforms. Ideally, the dimension of representation selected should capture a reasonably high percentage of the average signal energy in order to be effective. The average energy capture in this case represents both signal and noise energy, so an inspection of the basis vectors was necessary to ensure

limited noise capture in the subspace representation.

5.3.5 Matched filter analysis

Individual templates may also be selected for analysis in a matched filter routine. The three-component templates were first multiplexed and a time window extracted around the phases is extracted, starting 50ms before the P-wave arrival. Eq 2.6 was then applied to the data. The length of the data windows used for the templates was 330ms.

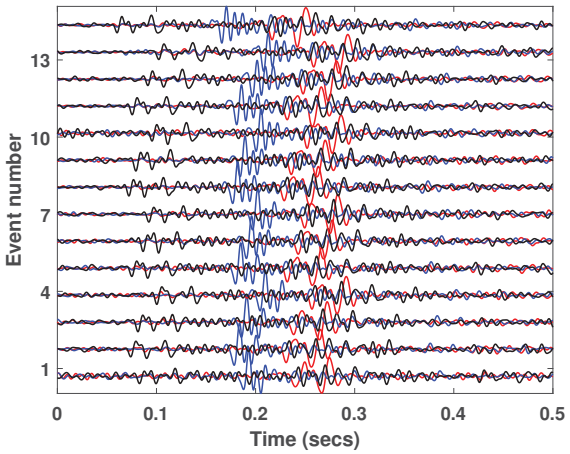
5.4 Tests and application

5.4.1 Processing efficiency of template based detectors

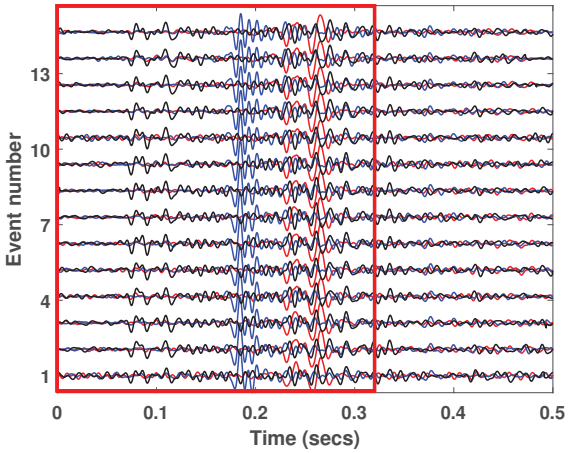
Using all multiplet group templates in a detection routine provides the most complete detection picture for the group. However, redundancy becomes a factor with an increasing number of templates, especially if most of the templates are fairly similar to each other. The rank reduction offered by the subspace detector may provide processing gains in terms of efficiency since it represents the multiplet group with a smaller number of templates. Furthermore, the use of multiple dimensions in the subspace representation could increase sensitivity to waveform variation compared to the use of a single template, and may be similar or better in performance to using all templates. We investigated this hypothesis by comparing the detections made by the subspace detector to those made using all templates in a multiplet group. The performances were further evaluated by comparing their precision, recall and F1-score, which were described in detail in section 2.4. In this test, multiplet group 1 was selected for analysis.

Methodology

The processing steps outlined in the construction of the signal subspace (section 2.3) and matched filter were followed in this test. Data windows encompassing the aligned waveforms were extracted, as shown in figure 5.4.



(a) Before alignment



(b) After alignment

Figure 5.4: Three component waveforms from multiplet group 1. a) Before time alignment. b) After time alignment. Red box indicates selected data window used in construction of signal subspace. Vertical, north and east components shown by red, blue and black colors respectively.

After extraction of the data window, a subspace representation was constructed using the methodology described in section 2.3.4. The average energy capture as a function of subspace dimension was computed in order to determine the optimal value of d, after which subspace detection was applied to the data. The energy capture curve for multiplet group 1 is shown in figure 5.5.

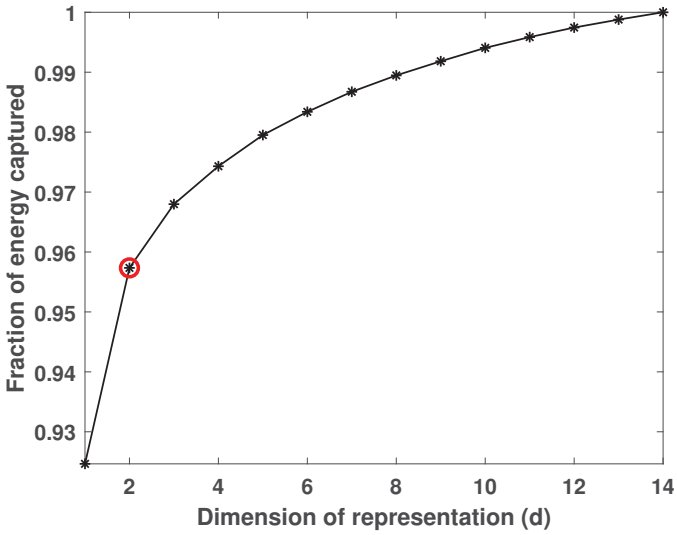


Figure 5.5: Average energy capture for multiplet group #1 as a function of the dimension of representation d. The chosen optimal dimension of representation is highlighted by a red circle.

The dimension d was selected to be 2 since after that dimension, the increases in the energy capture were marginal. An analysis was also carried out to see how the 1D subspace detector performed relative to the higher dimension detector, and the individual multiplet templates. Therefore, subspace detection was also applied to the data using a dimension $d = 1$. Each template in the multiplet group was also selected for use in a matched filter operation. The detection threshold for the subspace detector γ was set to 0.2 based on results from section 1.4.3, while an equivalent detection threshold τ_{CC} for the matched filter of 0.45 was set. A minimum of 8 receivers across the 12 receiver array were required to exceed the detection threshold before a detection was declared. These were then inspected visually to differentiate microseismic events from false alarms.

Results

The alignment procedure resulted in no discernible lag differences between the events on both the P- and S-waves as seen in figure 5.4. Removal of the inter-event lags is important as it allows for a low rank approximation of the waveforms in the subspace representation. The detection results from the matched

filters and subspace detectors were tallied and shown in table 5.2

Detector type	Events	False alarms	Missed events
Subspace detector ($d = 1$)	170	2	50
Subspace detector ($d = 2$)	208	7	12
Matched filter (all templates)	184	7	36

Table 5.2: Table showing results after application of subspace detection and matched filtering to a segment of data using a multiplet group. A total of 220 events were detected by both detectors.

Based on these results, the detector performances were evaluated in terms of their precision, recall and F1 score, and the results are shown in table 5.3.

Detector type	Precision	Recall	F1-score
Subspace detector ($d = 1$)	0.99	0.77	0.87
Subspace detector ($d = 2$)	0.97	0.95	0.96
Matched filter (all templates)	0.96	0.84	0.90

Table 5.3: Precision, recall and accuracy (F1-score) of the subspace detectors and matched filters constructed from multiplet group 1

Across both detectors, precision was very high which is an indication of high quality detections. The recall of the detectors was where the major differences lay. Using the 2D subspace detector resulted in the most number of detections relative to the other detectors, with an identical number of false

alarms to the results obtained using all 14 templates in matched filtering. This was reflected in its recall, which was the highest of all the detectors. The 1D subspace detector captured less events relative to the combined matched filter. This distinction was not very significant, as seen in the fairly small difference in recall between both results. It did have the highest precision, recording only 2 false alarms compared to the 7 detected by both the matched filter templates and 2D subspace detector.

The times taken to apply each detector to a single receiver are shown in table 5.4.

Detector type	Run time (s)
Subspace detector ($d = 1$)	$\approx$ 826
Subspace detector ($d = 2$)	$\approx$ 1004
Matched filter (1 template)	$\approx$ 826
Matched filter (14 template)	$\approx$ 11,564

Table 5.4: Processing times for both detectors after application to one hour of data on a single receiver.

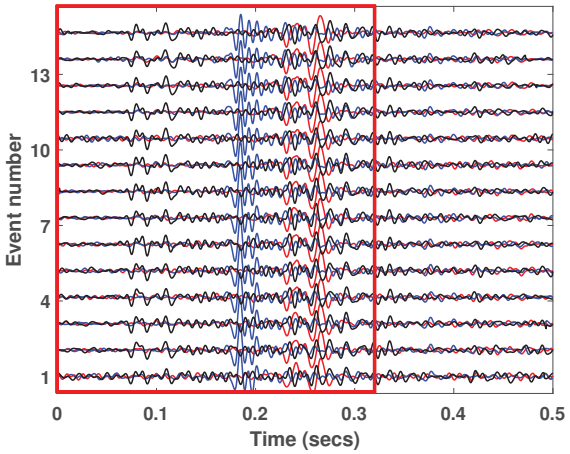
The single matched filter template and 1D subspace detector both had similar run times when applied to the data, with the 2D detector taking a bit more time. In relative terms, the 2D detector represent a processing gain of about 12x the speed of running all 14 templates, and a further 14x gain if the 1D detector was employed.

5.4.2 Effects of subspace design on waveform sensitivity

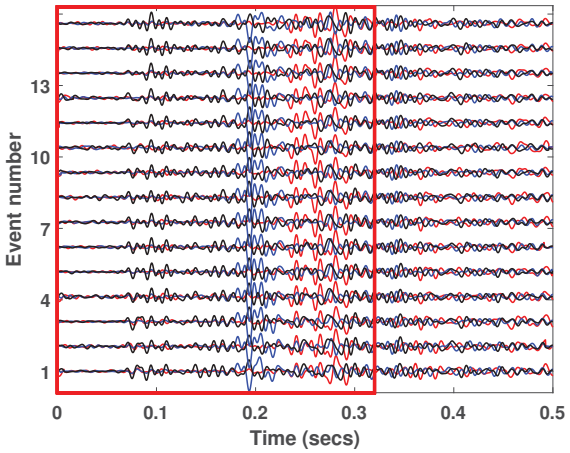
The correlation level at which the waveforms are separated into multiplet groups is dependent on the goals of the study. A high threshold may be more useful if diversity within a specific source is of interest, and a low threshold groups different source signals together. Applying subspace detection to individual multiplet groups is a means of increasing processing efficiency due to the rank reduction offered by the SVD at the heart of the detector. However, utilizing all the templates without performing multiplet analysis may be an even more efficient means of performing detection since it bypasses the processing step of grouping events. Examining the advantages and disadvantages of both methods of subspace construction is important in determining an optimal detection routine. In this test, multiplet groups 1 and 2 were selected for analysis.

Methodology

A subspace representation was constructed for each group using 330ms data windows centered around the aligned waveforms, encompassing both phases. Next, the waveforms from both groups were merged into a single group and also aligned in time. The waveform alignments are shown in figures 5.6 and 5.7.



(a) Multiplet group 1



(b) Multiplet group 2

Figure 5.6: Three component waveforms after alignment a) Multiplet group 1 b) Multiplet group 2. Red box indicates selected data window used in construction of signal subspace. Vertical, north and east components shown by red, blue and black colors respectively.

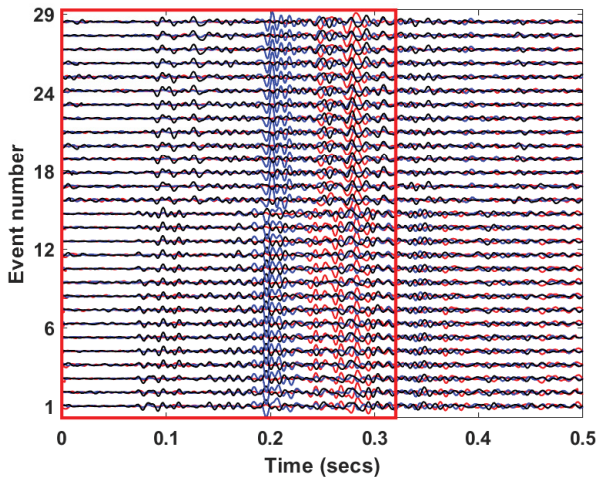


Figure 5.7: Three component waveforms from both multiplet groups 1 and 2 after alignment. Red box indicates selected data window used in construction of signal subspace. Vertical, north and east components shown by red, blue and black colors respectively.

In each case, the subspace dimension d was set to 2, the detection threshold γ was set to 0.2, and a minimum of 8 receivers across the 12 receiver array were required to exceed this threshold before a detection was detected. The detections were then visually inspected to differentiate microseismic events from false alarms.

Results

The detection results using both the individual and combined multiplet groups were tallied, and the results are shown in table 5.5.

Detector type	Events	False alarms	Missed events
Subspace detector (Groups 1 + 2)	302	12	0
Subspace detector (Merged)	286	5	16

Table 5.5: Detection results from subspace detection applied to both individual multiplets and combined multiplet groups. A total of 302 events were detected across both detectors.

The detections using the individual groups exceeded those of the subspace detector constructed from the combined group, displaying greater sensitivity to waveform. This came at an increased number of false alarms relative to the merged group. However, the difference in the number of events detected exceeded the differences in false alarms.

5.4.3 Detection optimization

Detecting as many events as possible at low false alarm rates is the main objective of any event detection routine. An increase in false alarms puts a burden on system resources, as well as processing time due to the increased screening required. The subspace detector displays sensitivity to events corresponding to the particular multiplet group of interest, but its performance may be constrained by the availability of templates. The STA/LTA detector on the other hand requires no signal information in its detection routine, making it a fairy straight forward and cost-effective technique. This lack of signal information can also be a draw back especially in cases where signal and noise information don't differ by a great deal, or overlapping events are present. This test aims to demonstrate how the subspace detector compares to the conventional STA/LTA detector in terms of detection performance, and how the number of templates used in subspace detection influences the results.

Methodology

The STA and LTA windows used in obtaining templates for the waveform library were retained (0.06s and 0.15s respectively), and the detection threshold τ_{SL} was lowered to a value of 3. This value was deemed equivalent to the subspace detector threshold γ of 0.2 since it minimized the number of false alarms. For each multiplet group, the waveforms were aligned and 330ms data windows extracted. Average energy capture curves were then used to obtain optimal dimension of representation d for each group, as shown in table 5.6.

Multiplet group	# of events in group	Dimension of representation
1	14	2
2	15	2
3	2	1
4	8	2

Table 5.6: Multiplet group and the dimension of representation for the subspace representation

For each multiplet group, eq 2.19 was applied to the aligned waveforms to determine the average energy capture. In this study, the minimum energy capture required for each group was 80%. The dimensions were further selected after physical examination of the left singular vectors. For both detectors, a minimum of 8 receivers across the 12 receiver array were required to exceed the specified detection threshold before a detection was declared. The detections were then visually inspected to differentiate microseismic events from false alarms.

Results

All four multiplet groups were used in subspace detection, while the STA/LTA detection algorithm was applied to the data segment. The results are shown in table 5.7.

Detector type	Events	False alarms	Missed events
STA/LTA	268	10	125
Subspace detector (Total)	351	15	42

Table 5.7: Detection results from subspace detection and STA/LTA applied to one hour of data. A total of 393 events were detected across both the detectors.

The performance of the detectors were further evaluated by looking at their precision, recall and accuracy via the F1 score. Table 5.8 shows these metrics.

Detector type	Precision	Recall	F1-score
STA/LTA	0.96	0.68	0.80
Subspace detector (Total)	0.96	0.89	0.92

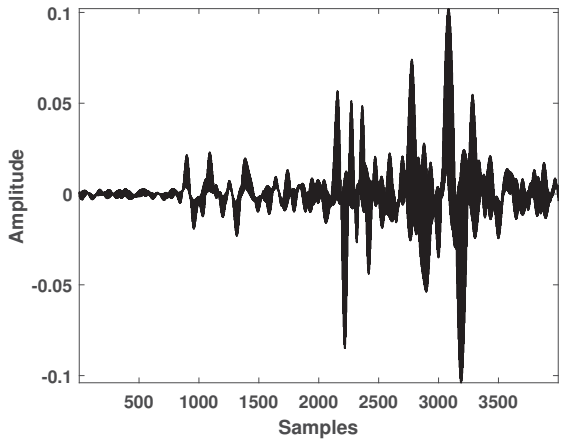
Table 5.8: Precision, recall and harmonic mean of the subspace detector and STA/LTA detector after application to microseismic data.

The individual subspace detectors had identical precision rates to the STA/LTA detector, validating the chosen detection thresholds. The difference in the number of events detector by the subspace detector was substantial, with a recall of 0.89 compared to the 0.68 of the STA/LTA. Both detectors displayed low alarm rates, with slightly more false alarms detected by the subspace detector. However, the differences were very marginal between both methods. Overall, the subspace detector displayed better efficiency, showing great balance between precision and recall with an F1 score of 0.92.

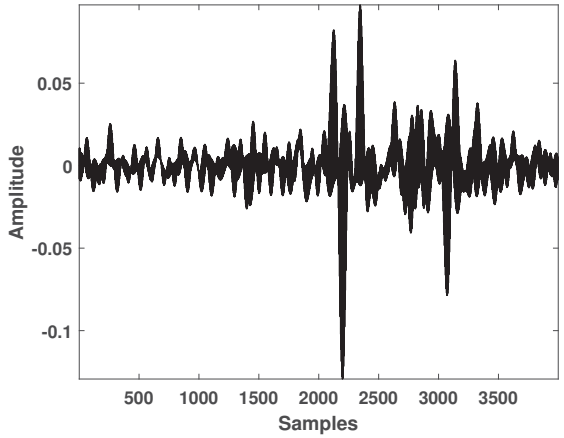
5.5 Discussion

Comparison of the subspace method to the matched filter revealed the dimension of the subspace detector (i.e the number of singular vectors) plays

an important role in its waveform sensitivity. Figure 5.8 shows the first two singular vectors of multiplet group 1 in multiplexed form.



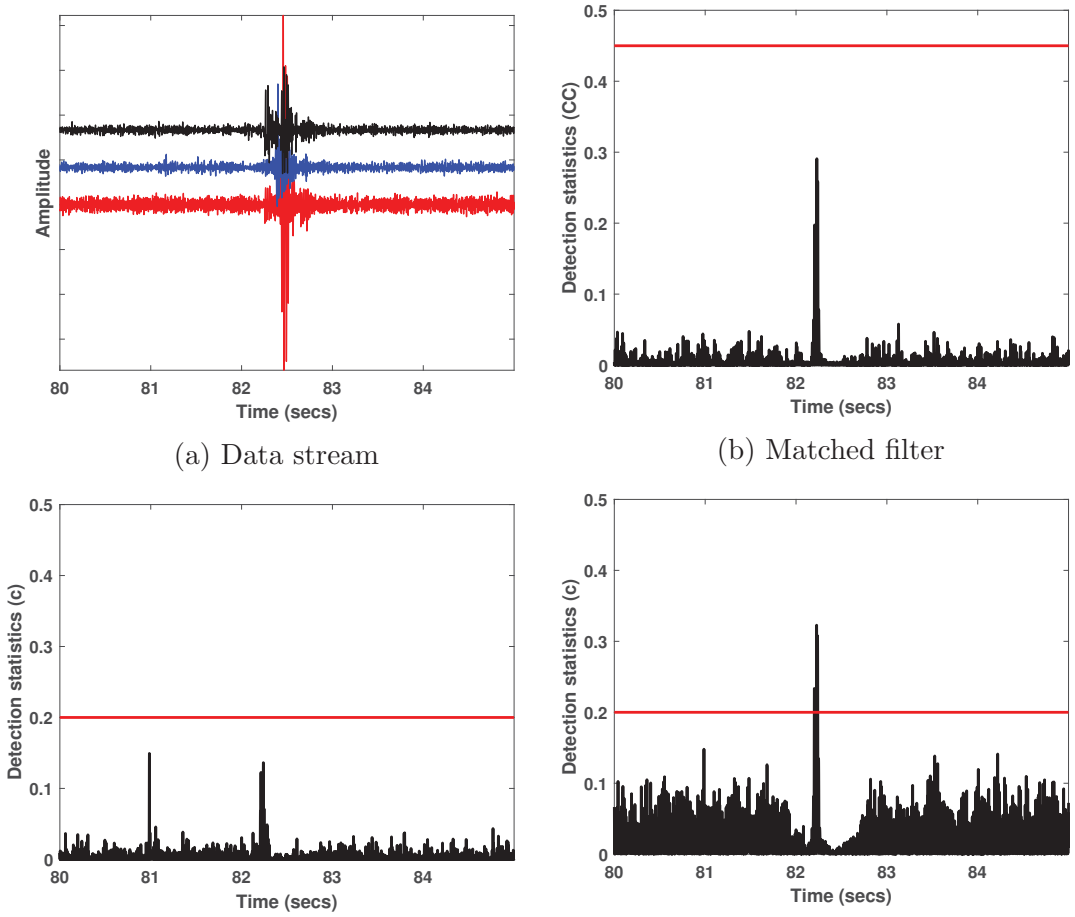
(a) 1^{st} singular vector



(b) 2^{nd} singular vector

Figure 5.8: First two singular vectors of multiplet group 1

The 1^{st} singular vector exhibited good signal-to-noise (SNR) ratio, with very visible P- and S-waves. The 2^{nd} singular vector displayed a lower SNR, with only the S-waves visible above the noise. Nevertheless, the use of both singular vectors supplied extra waveform information, which increases sensitivity to waveform variation. This is substantiated in figure 5.9 which shows an event missed by both the 1D subspace detector and matched filter templates but captured by the 2D subspace detector.



(a) Data stream

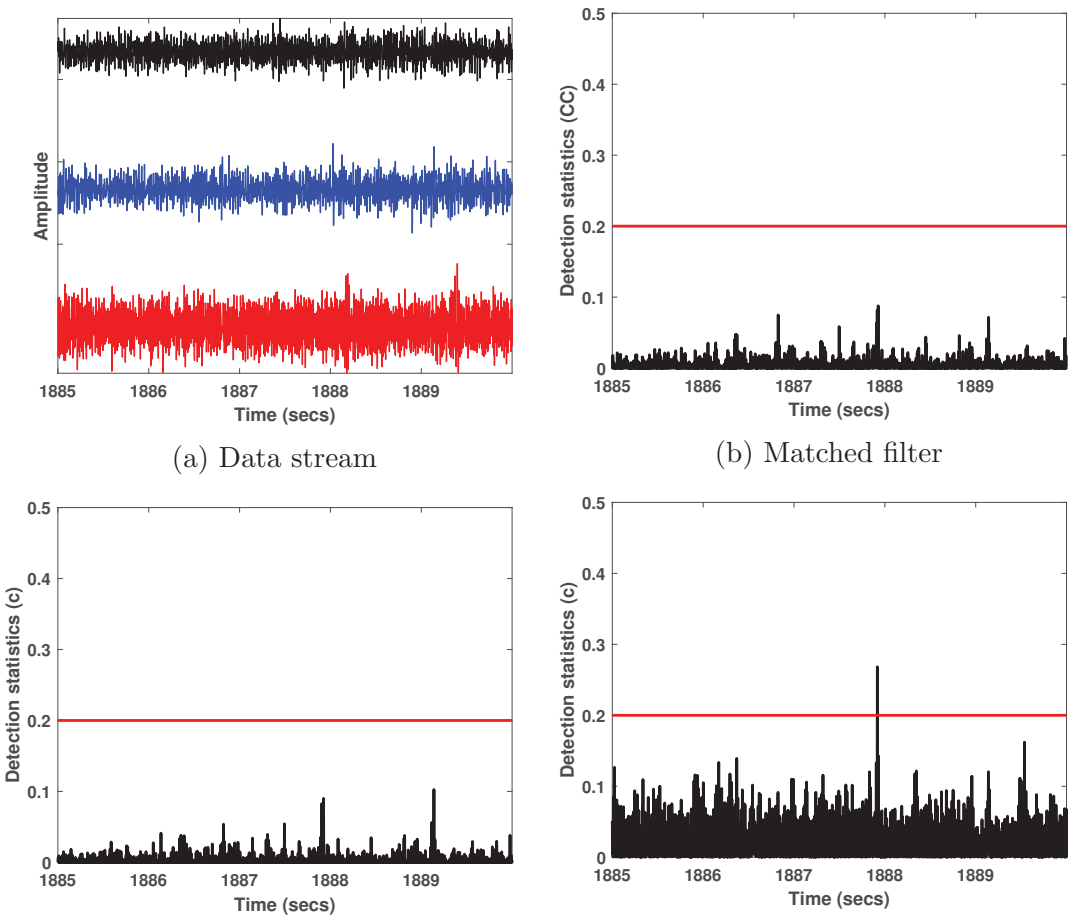
(b) Matched filter

(c) 1D subspace detector

(d) 2D subspace detector

Figure 5.9: Detection results applied on data stream. Vertical, north and east components shown in red, blue and black colors respectively. (a) Data interval contains an event of interest around 82s (b) Matched filter template has a peak lower than the detection threshold. (c) 1D subspace detector has a peak smaller than the detection threshold (d) 2D subspace detector shows the same peak as the 1D subspace detector, but with a larger amplitude which exceeds the threshold.

Here, the detection statistic is higher on the 2D detector relative to the 1D detector for the event. This increased waveform sensitivity comes at the cost of a higher probability of false alarms, since the use of higher dimensions increases the noise content in the subspace representation. This explains the detection made by the subspace detector around the 1888s mark in figure 5.10, where there is no apparent waveform information present.



(a) Data stream

(b) Matched filter

(c) 1D subspace detector

(d) 2D subspace detector

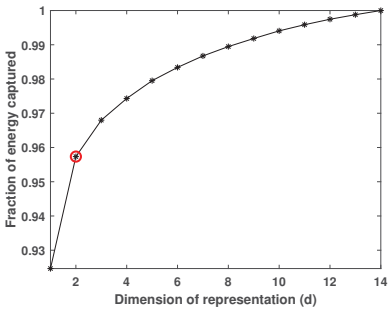
Figure 5.10: Detection results applied on data stream. Vertical, north and east components shown in red, blue and black colors respectively. (a) Data interval contains no discernible signal (b) Matched filter registers no detections. (c) 1D subspace detector registers no detection (d) 2D subspace detector has a peak which exceeds the threshold at about 1888s.

Computationally, the use of a subspace detector offers benefits both in terms of waveform sensitivity and speed. The 2D subspace detector not only detects more events relative to using all the templates in a matched filter operation, but it runs approximately twelve times as fast, which implies significant boosts in processing. The instances of multiple detection of the same event are also minimized or eliminated for multiplet groups since the detector captures the performance of all the templates in a single run. The 1D subspace detector detected less compared to the combined matched filter results with a lower false alarm rate. It would be expected that the use of a single dimension has a

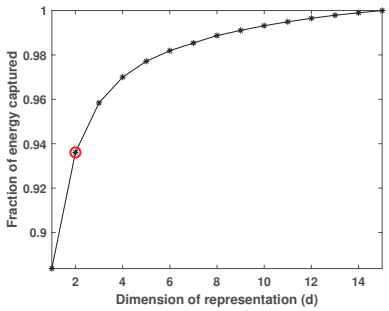
lower probability of producing false alarms relative to the individual templates since the 1^{st} singular vector is very similar to the stack of the waveforms, and thus would have an increased SNR. However, there is a loss in sensitivity to waveform variation since only the average waveform is used. Nevertheless, the number of events recovered were still fairly similar, with a recall of 0.77 compared to the 0.84 obtained using all the templates. Depending on the goals of the study, if only highly similar events to the waveform group are required for analysis, then the potential processing efficiency may be increased even further if a 1D detector is used. In this example, the speedup is roughly 14 times that of using all the templates. Overall, the template based detectors had high precision results, with the highest precision seen on the single templates and 1D subspace detector. The trade off between precision and recall dictates that the less precise detectors had better recall, which was the case with all 14 templates and 2D subspace detector. In terms of overall performance on both metrics, the 2D detector was superior with an F1 score of 0.97.

The design of the subspace detector also appears to play a major role in its sensitivity to waveform. A look at figures 5.6 and 5.7 showed optimal alignment when the waveforms came from the same multiplet group, and a relative shift in the alignment when multiple groups were combined into the same subspace representation. The lack of proper alignment increases the rank of the subspace representation since the lags become a characteristic of the design set. This is reflected in figure 5.11 which shows at the same dimension of representation, the energy capture for the merged group is less than those of the individual groups. At $d = 1$, the energy capture of the merged group is below 80%, compared to the individual multiplet groups are both above 80%. This suggests a reduction in waveform sensitivity of the singular vectors for the merged group as they become more representative of the averaged waveforms between the multiplet groups. The consequences of these observations can be seen after application to the microseismic data, in which the merged multiplet group detected less events than the individual subspaces. If the variability increases in the subspace representation due to the inclusion of waveforms from different sources or with varying waveform character, the distortion in the singular vectors would be

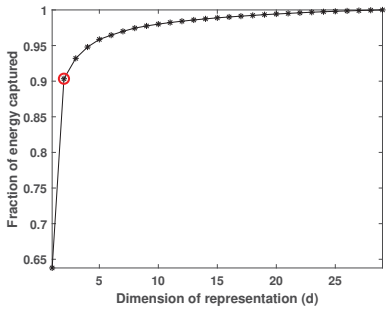
expected to increase further, minimizing waveform sensitivity.



(a) Multiplet group 1



(b) Multiplet group 2

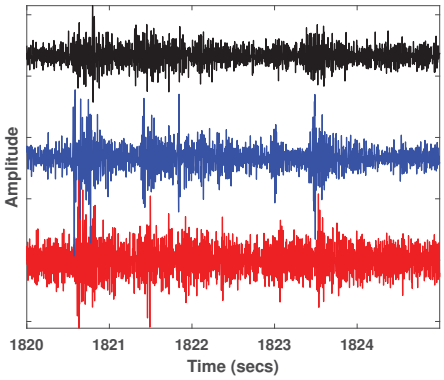


(c) Merged group

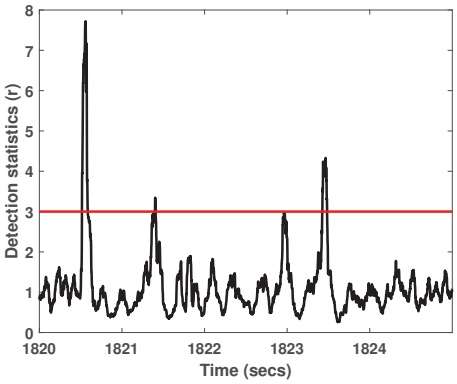
Figure 5.11: Average energy capture for multiplet groups 1 and 2, as well the merged group as a function of the dimension of representation d. The selected dimension of representation is highlighted by a red circle and corresponds to 2.

The overall detections across the subspace detectors exceeded those of the STA/LTA detector, with an overall F1 score greater than the energy detector. The performance of the subspace method is directly related to availability and diversity of the templates from a given data section. Having a large number of templates from different parts of the data increases the probability of detecting more events compared to having a lot of templates from the same period since there's less redundancy. The processing costs were also very reasonable, with a total of 7 dimensions needed to detect more than the energy detector at a comparable false alarm. The caveat to the increased sensitivity of the subspace detector to waveform variation lies in the fact that it strictly applies to waveforms that are either identical, or very similar to its subspace representation. This is in stark contrast to the STA/LTA detector which does not rely on waveform information. This explain why some waveforms present in the data

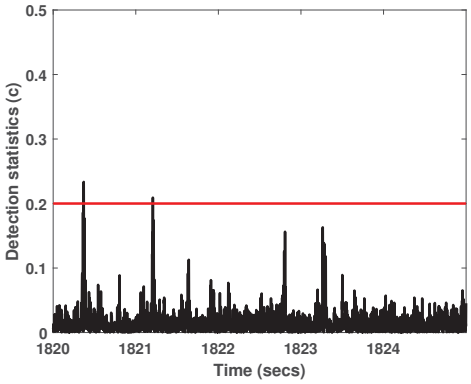
segment shown in figure 5.12 were missed by the subspace detector, but picked up by the STA/LTA detector.



(a) Data stream



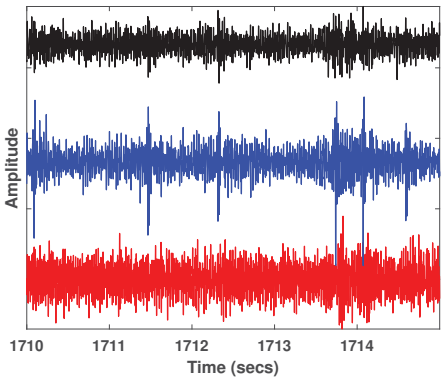
(b) STA/LTA detector



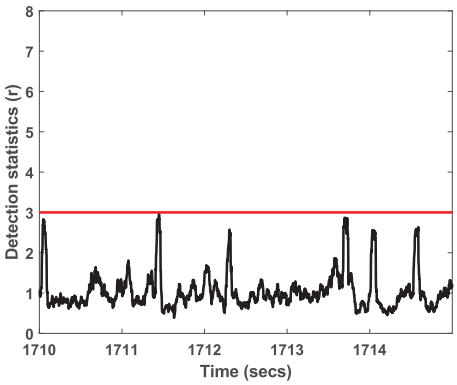
(c) 2D subspace detector (Group 2)

Figure 5.12: Detection results applied on data stream. Vertical, north and east components shown in red, blue and black colors respectively. (a) Data interval contains several events (b) STA/LTA detector registers four detections (c) Subspace detector only picks up two events.

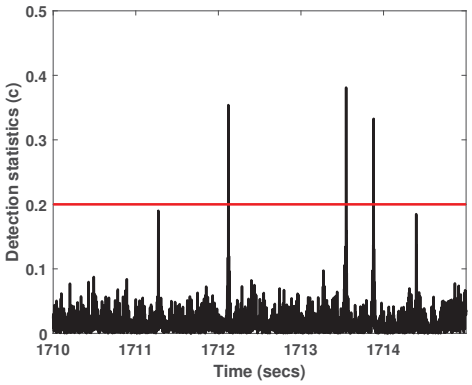
This insensitivity of this detector to waveform also implies it cannot differentiate signal from noise, leading to greater probability of missed events if they are low magnitude, buried in noise or overlapping events. The subspace detector on the other hand displays high sensitivity to events represented by its subspace, especially with higher dimensions. This is substantiated in figures 5.13 and 5.14 which shows a section of data containing waveforms that is missed by the STA/LTA detector but picked up by the subspace detector, and a data section highlighting the STA/LTA's lack of waveform sensitivity.



(a) Data stream

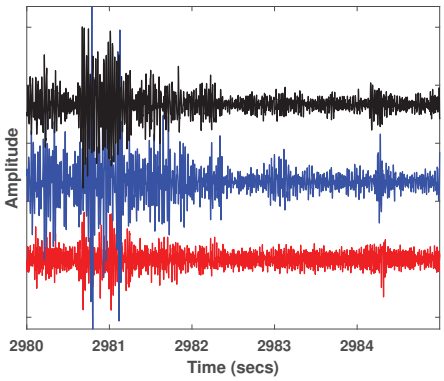


(b) STA/LTA detector

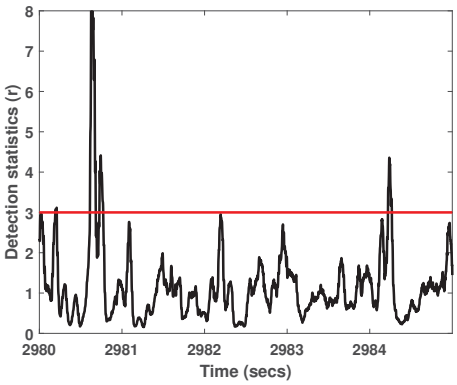


(c) 2D subspace detector (Group 2)

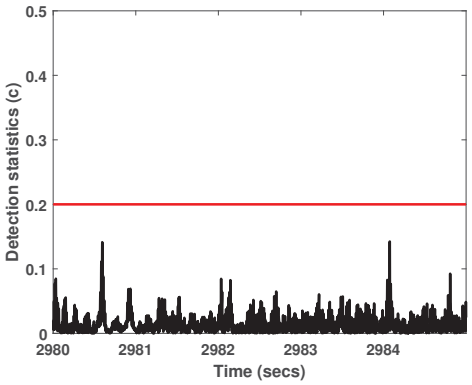
Figure 5.13: Detection results applied on data stream. Vertical, north and east components shown in red, blue and black colors respectively. (a) Data interval contains several events (b) STA/LTA detector registers no detections (c) Subspace detector only picks up three events.



(a) Data stream



(b) STA/LTA detector



(c) 2D subspace detector (Group 2)

Figure 5.14: Detection results applied on data stream. Vertical, north and east components shown in red, blue and black colors respectively. (a) Data interval contains a regional high amplitude seismic event and a low magnitude microseismic signal (b) STA/LTA detector detects both the seismic events and a false alarm around 2982s (c) Subspace detector registers no detections

5.6 Summary

The subspace detector is a robust technique that can handle waveform variability even in adverse noise conditions. Matched filters are very sensitive to events generated by specific sources, with very low false alarm rates and high detection probability. However, they cannot capture waveform variations and a degradation in performance occurs with increasing change in waveform character, resulting in fewer detections. The use of multiplet groups in subspace detection increases processing efficiency and speed relative to the standard matched

filter, while increasing sensitivity to waveform compared to utilizing different waveform groups at once. The STA/LTA trigger's insensitivity to waveform allows for more capturing of signals, but at the cost of an increased number of false alarms if a relaxed detection threshold is used. Unlike the template based detectors, its detection ability is heavily influenced by the noise levels present in the data. The use of several subspace detectors has been shown to increase event detection capabilities without incurring heavy processing costs at reasonable false alarm rates.

Chapter 6

Microseismic imaging of fracture complexity in a tight sand reservoir

6.1 Introduction

The double difference algorithm is an efficient technique for collapsing diffuse catalog locations into sharp images of seismicity and outlining the narrow region of faults (Waldhauser and Ellsworth, 2000). The use of catalog and cross-correlation data in a combined system of equations allows for simultaneous relocations between multiplet groups and uncorrelated events in a robust manner. However, the system of equations to be solved can become substantially large for numerous events which increases processing times, and may become difficult to solve if the system is ill-conditioned. The master event technique on the other hand offers the advantage of faster processing times, but the signal-to-noise ratio of the selected master event and its location relative to the events to be relocated place constraints on its accuracy (Stoddard and Woods, 1990; Bouchaala et al., 2013; Jones et al., 2008). A drawback to this technique lies in the use of a single event in the relocation technique, which is less robust than using all event pairs.

The subspace detector offers the potential for precise detection of events of interest at low false alarm rates, and can increase the number of events related to specific sources via multiplet groups. Additionally, the 1^{st} singular vector obtained from application of singular value decomposition (SVD) to a group of very similar events is assumed to represent their general hypocenter characteristics, and its high signal-to-noise ratio offers the potential for more accurate relative event locations.

In this work, the ability of the subspace detector to simultaneously improve event detection and location capabilities is demonstrated on microseismic data from treatment of an unconventional tight sand reservoir, which occurred over a two day period in the Hoadley field, Alberta. A total of 506 events were detected and separated into multiplet groups based on waveform similarity. Six multiplet groups from different periods, spanning the entire treatment were selected for analysis. First, the methodology for detection and location of the microseismic events is outlined. Next, absolute and relative locations are obtained for these events, and their accuracy compared. Finally, the seismic images obtained will be analyzed and discussed.

6.2 Methodology

6.2.1 Pre-processing

The dataset analyzed comes from the Hoadley field, which is described in more detail in chapter 5 (figure 5.1). The processing steps described in section 5.3.1 were applied to the data. A trapezoidal bandpass filter with corner frequencies [20 40 145 160] Hz was applied to the data to suppress as much noise as possible. A notch filter was also applied to remove 60Hz frequencies corresponding to possible power lines in the area. Next, an STA/LTA detector was applied to all 7 hours of data to obtain templates for use in the subspace detector. STA and LTA windows were selected to be 0.06s and 0.15s respectively, and a conservative threshold value τ_{SL} of 10 was used to declare a detection at a receiver. 8 out of 12 receivers were also required to trigger before a global

detection was declared to ensure good quality detections. After application of the detection routine and data screening, a total of 506 templates were selected for analysis.

6.2.2 Multiplet analysis

In this study, multiplet groups were defined as groups for which the average cross-correlation coefficients is ≥ 0.8 i.e. every member of a group must have a correlation coefficient greater or equal to this value with at least one other member of the group. Starting 50ms before the P-wave arrival time estimates, a 500ms window encompassing both P- and S-waves was used to extract waveform data before calculating the peak correlation coefficient value between event pairs. This process is described in more detail in section 5.3.3. A total of 201 multiplet groups were formed, 35 of which contained multiple events. In this study, 6 multiplet groups were selected for analysis. The events contained in these groups spanned the entire treatment period and are shown in table 6.1.

Multiplet group	Number of events
1	6
2	7
3	30
4	77
5	6
6	15

Table 6.1: Number of events contained within the multiplet groups.

6.2.3 Subspace detection

With the selected multiplet groups, the events were multiplexed and aligned in time using the iterative stacking procedure of (De Meersman et al., 2009). Starting 50ms before the P-wave arrival time estimates at each receiver, 330ms

data windows encompassing the aligned multiplexed P- and S-waves were extracted for use in construction of the signal subspace for each group. The matrix of aligned events was then decomposed into three unitary matrices via a singular value decomposition (SVD) (eq 2.16), and an orthonormal representation of the aligned events **U** was obtained from the matrix of left singular values. The subspace dimension of representation d was then determined by applying eq 2.19 to the aligned waveforms. Figure 6.1 shows an energy capture curve for a multiplet group at a receiver.

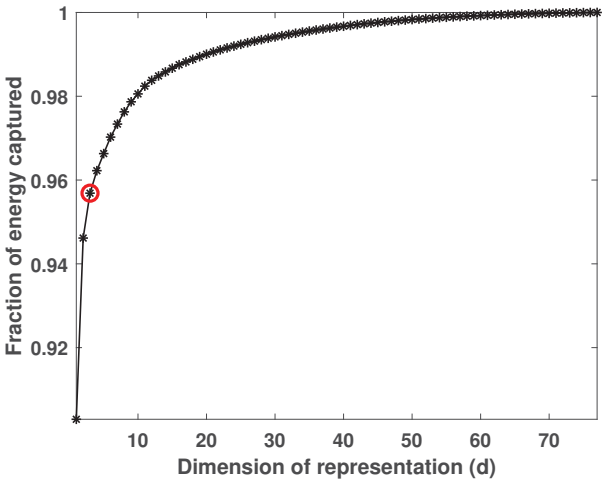


Figure 6.1: Average energy capture as a function of subspace dimension of representation d for multiplet group 4. The selected dimension of representation is highlighted by red circle.

Based on the curves and visual inspection of the singular vectors, dimensions of representation were chosen for each multiplet group. Finally, the detection threshold for the subspace detector γ was set to 0.4 to limit bad quality events. In this study, usable events were defined as events that had P- and S-waves which were distinct from noise.

6.2.4 Absolute data

In order to obtain hypocenter estimates for the microseismic events, arrival time and polarization estimates were required. However, noisy conditions make arrival time picking more difficult, and can also have an adverse effect on po-

larization estimates. The left singular vectors are assumed to represent higher signal-to-noise ratio generalized descriptions of the detected waveforms. In this study, the 1^{st} singular vector was used to obtain initial arrival time estimates for the child events. The process involved can be outlined in a series of steps:

1. The singular vector representation was obtained at each receiver using the methodology described in sections 4.2.1.

2. Arrival time estimates were then obtained for this vector at each receiver. Since this vector is our assumed master event, the estimated arrival time picks at each receiver must be as accurate as possible to minimize errors in the absolute location.

3. 75ms data windows around the initial arrival time estimates for both the P and S-wave on the singular vectors were then extracted. For each event, a sample-by-sample correlation (Eq. 2.25) was applied to the data stream using the P- and S-waves from the singular vector. This operation is applied starting around the initial time pick estimate for the child event at each receiver. The position with the maximum correlation coefficient is selected as the arrival time pick for that receiver. Due to poor data quality, a minimum correlation coefficient of 0.7 was required to register an arrival time pick at a receiver.

4. Receivers which did not meet this minimum correlation requirement were not assigned any time picks. After the initial time picks were obtained, the median of all the nonzero time picks was calculated and the AIC function (eq 3.8) was computed using a 12.5ms data window on either side of the median to obtain arrival time picks on receivers with no initial time picks (Zhang et al., 2003).

5. After obtaining arrival time estimates at all receivers, a new median time pick was computed, and the AIC process was repeated again on receivers with arrival time estimates greater than 2.5 times the standard deviation of the new picks. The picks were then manually adjusted if necessary.

6. P-wave polarizations were obtained using the final arrival time picks. Eq. 3.9 and 3.12 were applied to a 50ms data window beginning 12.5ms before the P-wave arrival time picks to estimate their back-azimuth.

6.2.5 Differential data

The presence of arrival time inconsistencies in the data can have negative effects on the quality of the event locations since random errors in arrival time estimates or noisy back-azimuth estimates translate to location uncertainties. The high degree of similarity between similar events offers the potential for obtaining high precision relative times using cross-correlation. The relative times can help minimize common errors from both the velocity model and travel times since the events are assumed to be closely spaced and path errors cancel out. The small separation distances also imply common errors in the back-azimuth errors which could be minimized by using the relative polarization estimates. A procedure for estimating both differential travel times and back-azimuths between the 1^{st} singular vector and child events is introduced which does not involve computing covariance matrices. The steps are as follows:

1. Extract a time window around both the singular vector and child events for both phases. We utilized a 50ms window around both P- and S-waves. For both phases, the window began 10ms before the estimated arrival time estimates.
2. Next, the windowed waveforms are multiplexed following the convention of eq 2.3 and cross-correlation is performed between the multiplexed child events and singular vectors.
3. The maximum correlation value is obtained and its corresponding lag is applied to align the phases. This lag is divided by three to account for multiplexing.
4. After alignment, the relative azimuth between the child events and singular vector is analytically estimated using the horizontal components of both signals. This procedure is described in more detail in Appendix B.

A master event was also selected from each cluster, and the above steps were applied to obtain differential data for master event relocation.

6.2.6 Event location

A calibrated 1D velocity model was provided for this dataset and is shown in figure 6.2. In this work, a grid search approach was used to determine initial absolute event locations. The grid volume was defined as a 2000m $\times$ 2000m $\times$ 280m cube in x, y and z respectively, with a 10m spacing between grid nodes, and it covered the area of interest. Hypocenter locations for the child events and 1^{st} singular vector were then obtained by minimizing the squared difference between observed and predicted arrival times and azimuths using Eq. 3.18.

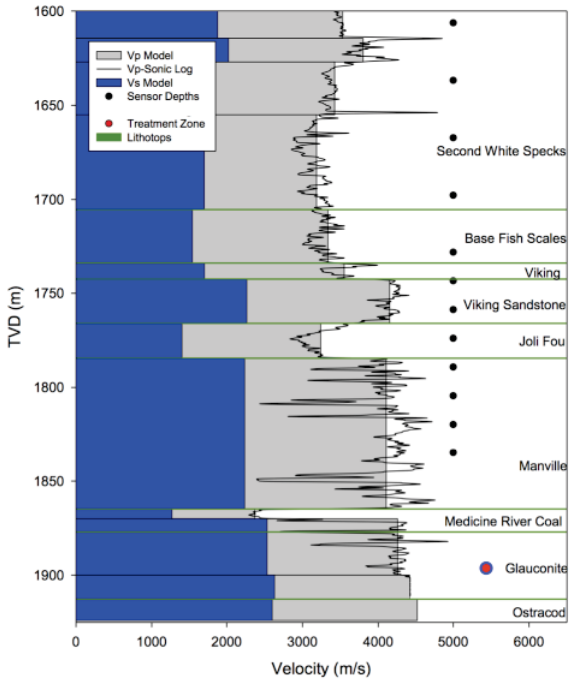


Figure 6.2: Schematic of the velocity model used to compute hypocenter locations. Geophones at their respective depths are indicated by black dots. Formation boundaries are also shown. Treatment occurred within the Glauconite formation around a depth of 1890m. Taken from (Eaton et al., 2014).

With the hypocenter locations, travel times were predicted to the receiver stations and subtracted from the arrival times to get origin time estimates. These were then removed from the arrival times to obtain the observed travel times. The child events were then relocated relative to the singular vector hypocenter. Only the differential data were used as observables in the relocation scheme i.e the cross-correlation derived lags and relative azimuths. An event

from each cluster was chosen as a master event and used in a master event relocation procedure, and the results were compared to those from subspace relocation.

6.3 Results

6.3.1 Detections

A total of 365 events were obtained by the 6 subspace detectors. The events spanned all 7 hours of treatment, having the greatest concentration of events in the early stages. Due to the poor horizontal constraints of the acquisition geometry and high noise levels, only events with reasonably clear P-wave arrivals were selected for further processing. Events contaminated by high amplitude noise were also excluded from the study. Screening of the data resulted in 296 events left for analysis. The dimensions of representation, multiplet size and number of detected events made by the groups are shown in table 6.2

Multiplet group	Number of events	Dimension of representation	Detected events in cluster
1	6	2	30
2	7	2	12
3	30	3	76
4	77	3	134
5	6	2	10
6	15	2	34

Table 6.2: Dimension of representation and number of detected events in each cluster.

6.3.2 Hypocenter locations

With the arrival time and polarization estimates, the absolute locations of the events and singular vector were resolved by minimizing the squared difference between the observed and predicted arrival times and azimuths using eq 3.18,

based on the 1D velocity model. Next, using the differential data, the events were relocated using subspace relocation, using the absolute locations of the singular vector obtained from the grid search. These locations are shown in figure 6.3.

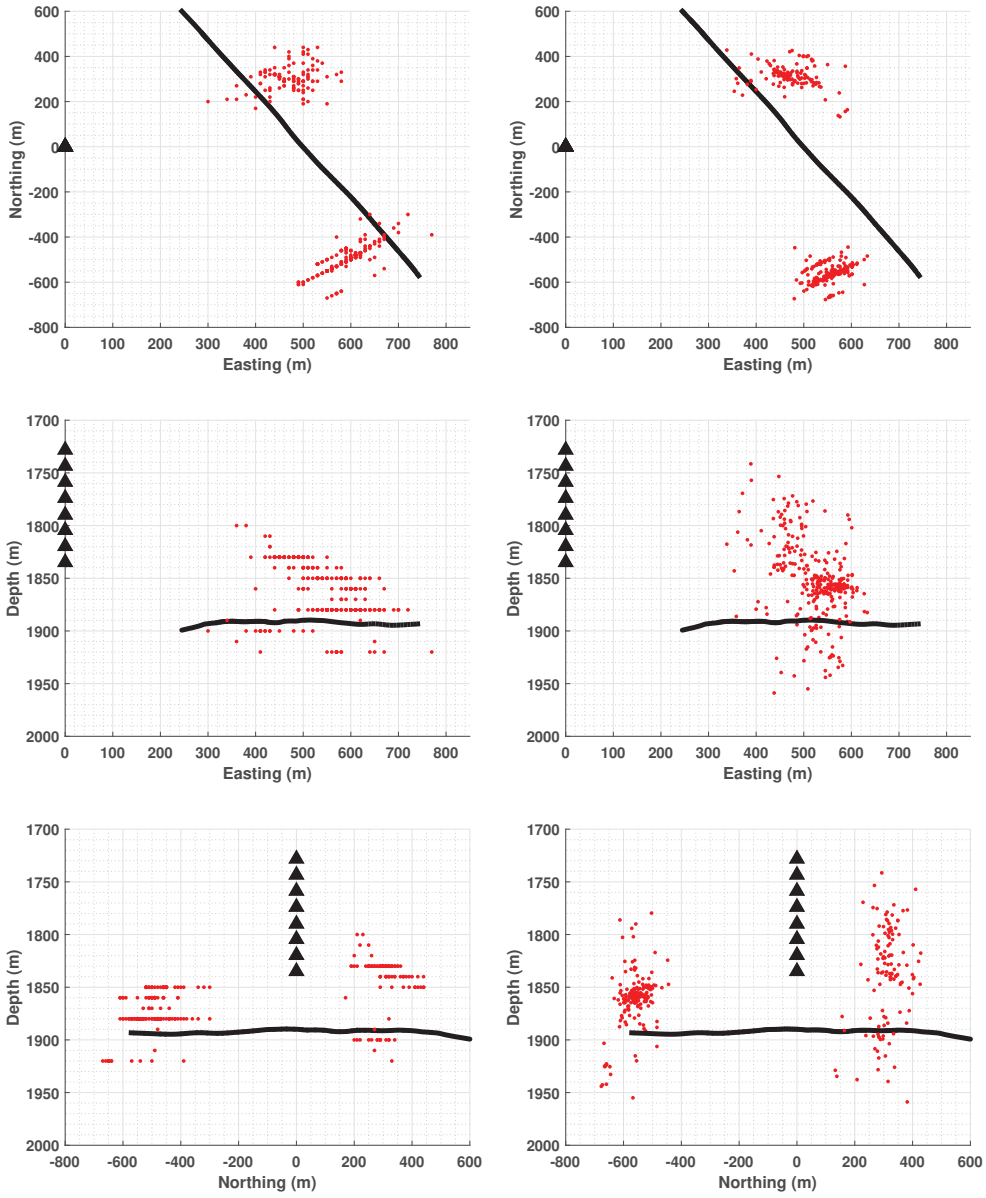
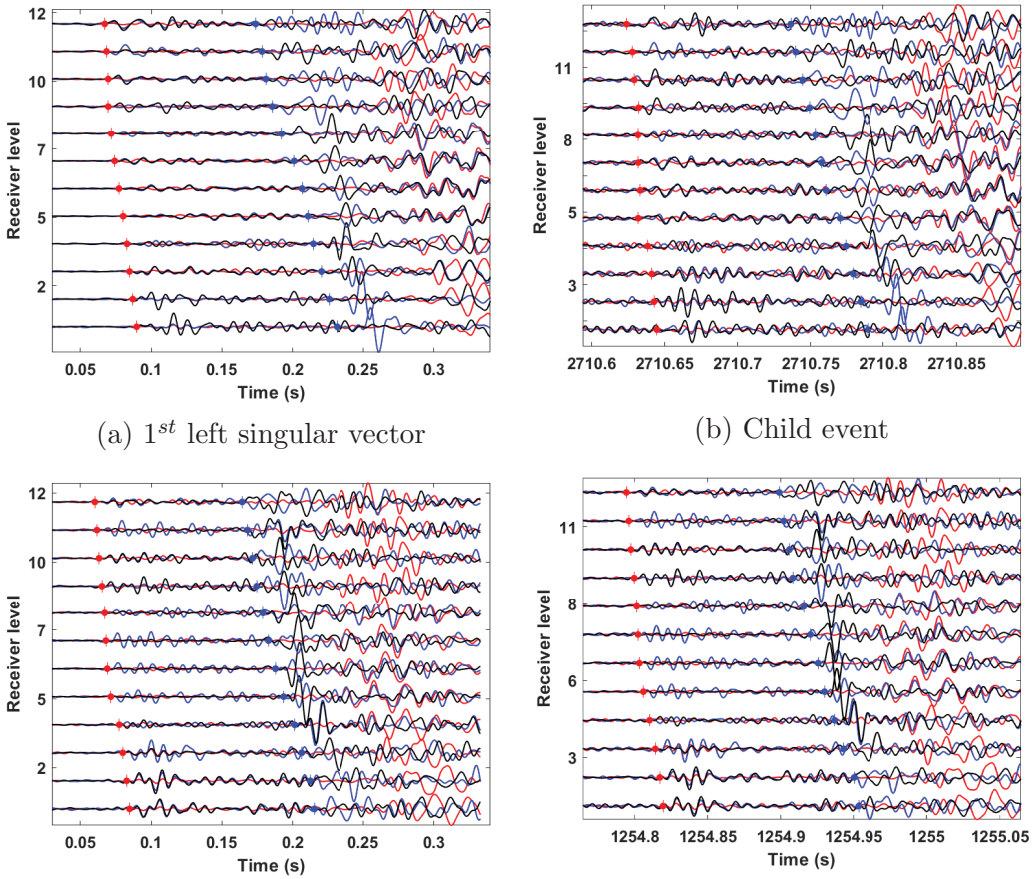


Figure 6.3: Event locations shown as 2D illustrations. Receiver array shown as black triangles while treatment well shown as solid black line. The events are highlighted as red blocks. *Column 1*: Absolute locations. *Column 2*: Subspace relocations.

The seismic image obtained from the grid search was fairly blocky, as most events appeared in narrow bands, especially in depth. This is a consequence of the grid spacing used to define the grid. If the average model P-wave velocity of 3643m/s is considered, then a minimum time difference δt of 2.7ms between events will result in a distance change equal to the selected grid spacing of 10m. Time differences lower than this will all locate on the closest grid point. Using a grid spacing of 5m reduces δt to 1.37ms which would allow for better resolution of location differences, but increases the grid size by a factor of $\approx$ 8, which is computationally expensive in terms of speed and processing costs. After relocation, the majority of the events have been pulled closer to each other, with more defined strike and dip directions observed in the seismic image. Additionally, the events have been separated into more distinct bands in the horizontal plane, suggesting the existence of multiple sub-structures within the larger structure. However, some events were pushed further away and constituted outliers in the relocation procedure. Overall, more detail was evident in the seismic image.

6.3.3 Quality of the absolute data

Figure 6.4 shows arrival time picks on some singular vectors and events from their corresponding cluster.



(a) 1^{st} left singular vector

(b) Child event

(c) 1^{st} left singular vector

(d) Child event

Figure 6.4: 1^{st} left singular vector representation and child events from two clusters. (a): 1^{st} left singular vector representation for cluster 4. (b): Child event from cluster 4. (c): 1^{st} left singular vector representation for cluster 6. (d): Child event from cluster 6. P and S-phase arrival time picks are shown as red and blue dots respectively.

Visually, the singular vectors display high similarity to their respective child events in terms of the waveform character and time moveout across the receiver array. The P-wave and S-waves are also very evident relative to the noise on this signal. The use of the singular vector to estimate the arrival times served as a guide in terms of identifying the phase arrivals in the data, especially in

cases where the noise levels made it difficult to identify the signal onset. The success of the relative location procedure is very dependent on the accuracy of the absolute location of the chosen master event for both subspace and master event relocation. The more accurate the master event location, the closer the events are pulled to their true positions for highly similar events. Figure 6.5 shows the positions of the singular events and the selected master events from each cluster.

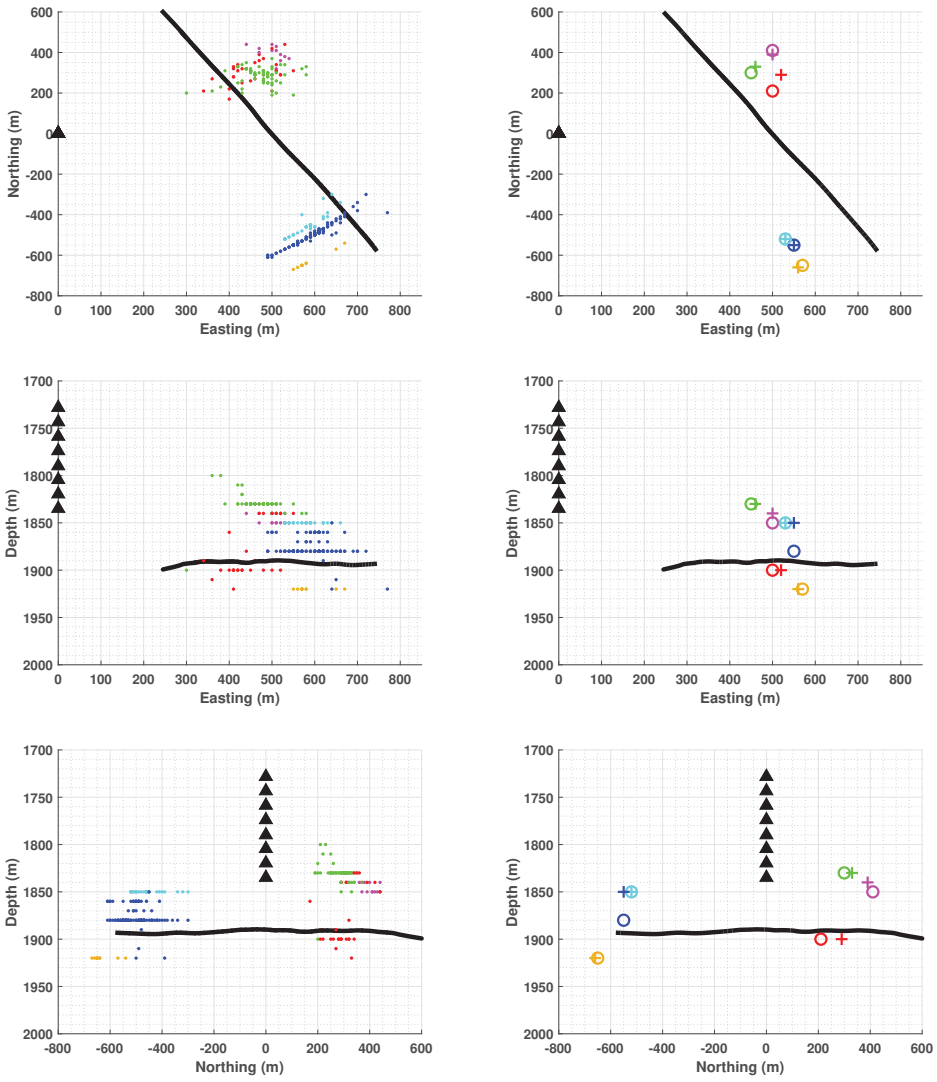


Figure 6.5: Absolute locations from grid search shown as 2D illustrations. Receiver array shown as black triangles while treatment well shown as solid black line. The colours correspond to the different clusters and are denoted by: Red (1), Magenta (2), Green (3), Blue (4), Orange (5) and Cyan (6). *Column 1*: Locations of the detected events. *Column 2*: Locations of the singular vectors (crosses) and master events (circles).

Both the singular vectors and master events located within the vicinity of their respective clusters, which implies realistic relative locations. The validity of the absolute locations were further examined by observing the event locations relative to the stages during which they occurred. Figure 6.6 shows the absolute locations of the events groups based on the stage at which they occurred.

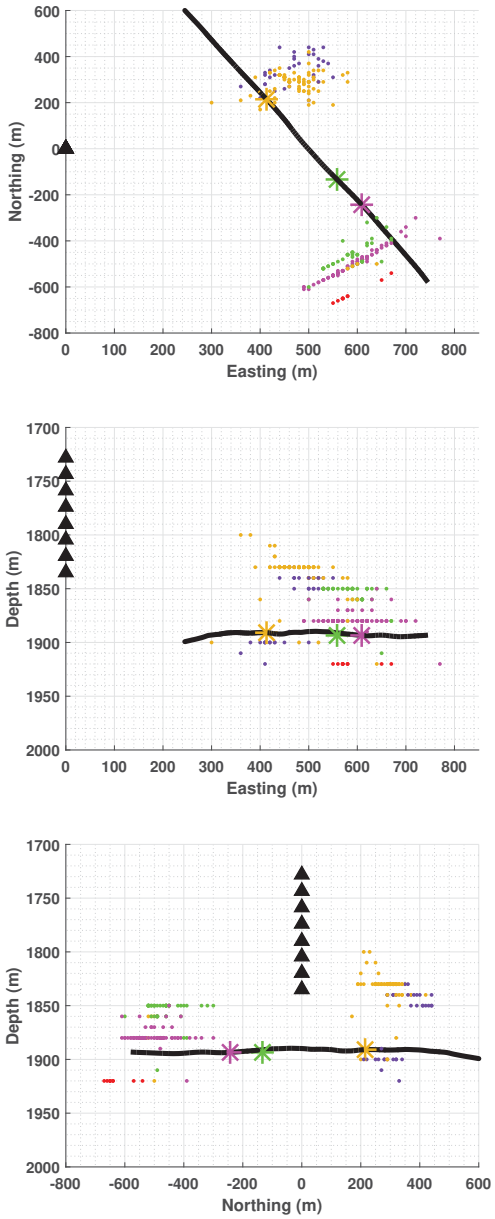
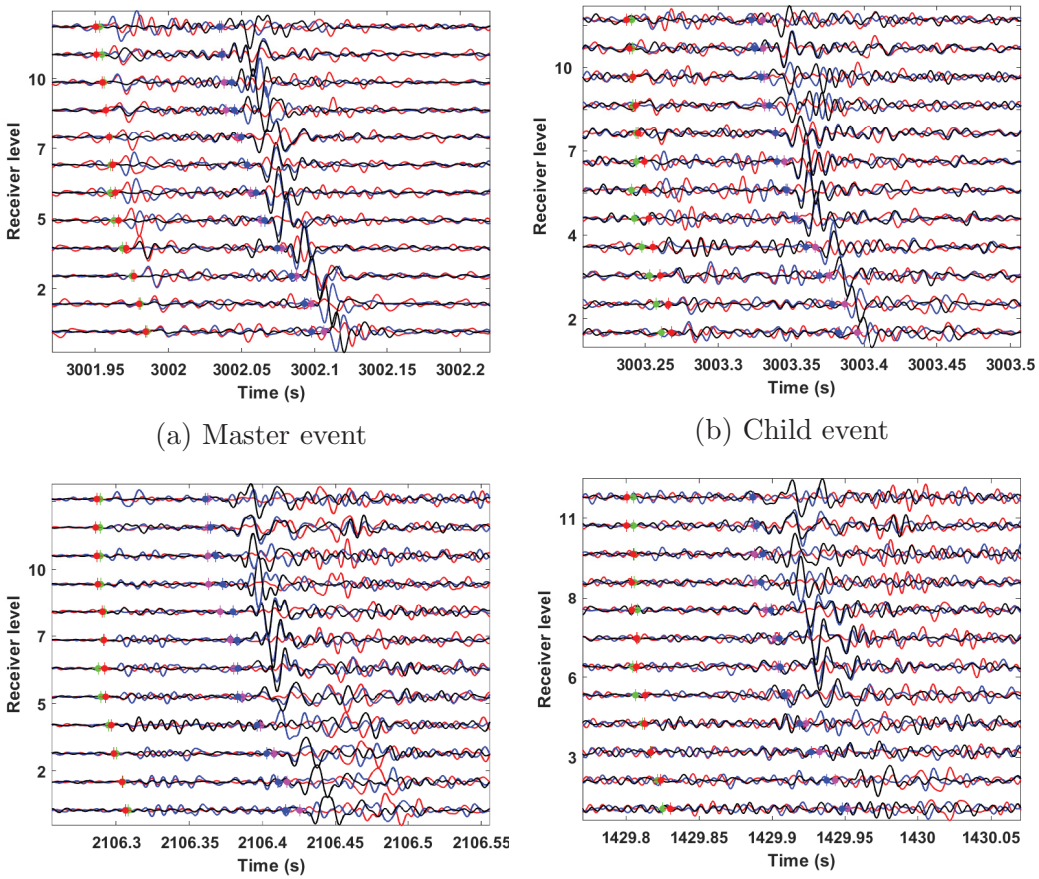


Figure 6.6: Absolute locations of the detected events grouped by the treatment stage during which they occurred. Receiver array shown as black triangles while treatment well shown as solid black line. The injection points are shown as stars. The colours correspond to the different stages and are denoted by: Red (1), Magenta (4), Green (5), Orange (8) and Purple (post pumping). Only stages containing more than 5 events are shown

At each stage, the events appear to locate reasonably close to the stage specific injection point which suggests the absolute locations are realistic. However,

the most important determinant of absolute location accuracy is how reliable the predicted travel times are in terms of explaining the waveform. Most of the master events were close to the respective singular vector for their clusters. It is therefore reasonable to use the results of the master event relocation to validate the subspace relocation results. The master events were picked based on a good signal to noise ratio, as well as the reliability of their predicted travel times. Figure 6.7 shows master events and child events from clusters 1 and 3 with their observed and predicted travel times.



(a) Master event

(b) Child event

(c) Master event

(d) Child event

Figure 6.7: Master event and child events from two clusters, with their observed and predicted arrival times. (a): Master event (cluster 1). (b) Child event (cluster 1). (c): Master event (cluster 3). (d): Child event (cluster 3). The red and blue dots correspond to the picked observed P-wave and S-wave arrival times, while the green and magenta dots denote the predicted P-wave and S-wave arrival times.

The signal-to-noise ratio of the master events were higher relative to the child events and this allowed for more precise picking of the arrival times, with the lower noise content having less of an influence on the polarization estimates relative to the child events. The high signal-to-noise ratio is especially important since cross-correlation lags are utilized in the relocation procedure, and higher correlations correspond to higher quality lag estimates. Furthermore, the predicted arrival times were close to the expected onset times on the master, which is a good indicator of accuracy since it means the predicted location explains the event reasonably well. This is also suggests that the grid spacing used in the location procedure was reasonable.

6.3.4 Quality of the differential data

The inconsistencies in the arrival time picks may be minimized by using cross-correlation. Figure 6.8 shows waveform data for the 1^{st} singular vector from cluster 4 and a child event at a common receiver.

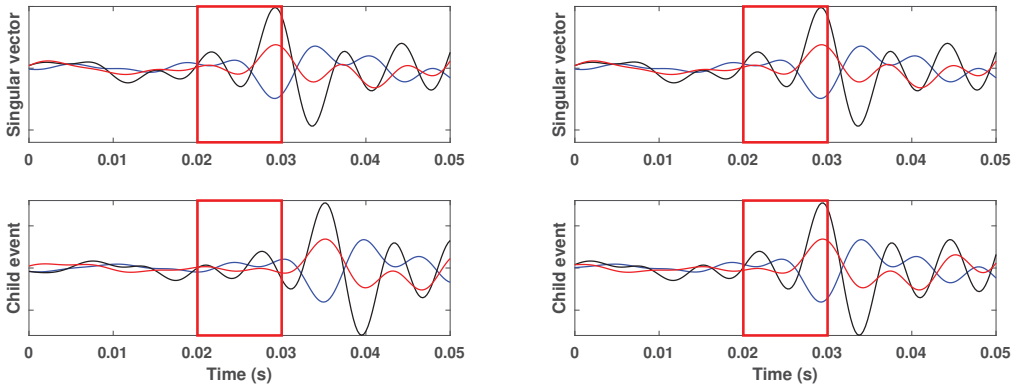


Figure 6.8: Windowed P-wave from a microseismic event and 1^{st} left singular vector from cluster 4 recorded at a common receiver. Inaccuracy in arrival time picking resulted in a misalignment between both waveforms. *Left column*: P-waves before alignment. *Right column*: P-waves after alignment. The 1^{st} left singular is shown as the top figure in both columns while the child event is the bottom figure. The red box highlights a section of the waveforms to illustrate the lag correction.

The high similarity of these two events suggests their corresponding arrival time picks should be in the same position, a fact that is addressed by cross-

correlation. The relative arrival time inconsistencies between both waveforms is corrected using the peak time lag from the cross-correlation function. This operation helps minimize the effects of random errors in the arrival time estimates. Figure 6.9 shows the resulting time lags after cross-correlation correction using the singular vectors.

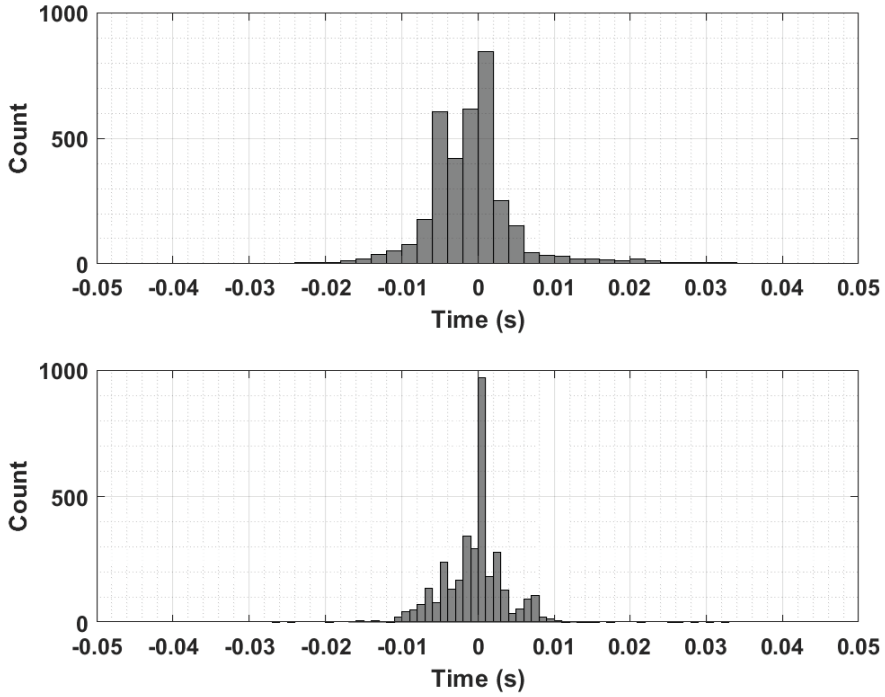
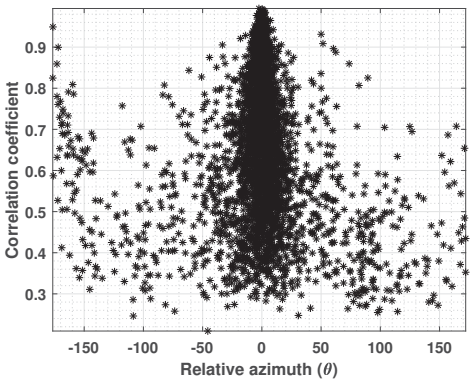


Figure 6.9: Histogram of lag corrections between the events and their first singular vector. *Top panel*: P-wave correlations. *Bottom panel*: S-wave correlations.

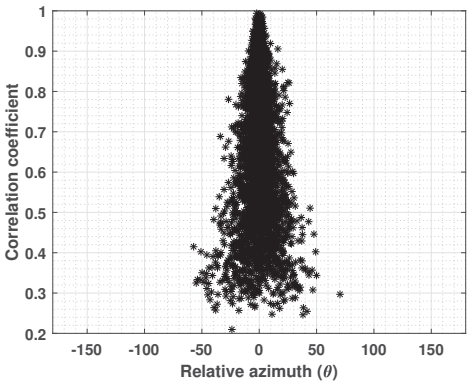
The distribution suggests most of the possible picking errors fall within 10 msecs, with some extending as far as 40 msecs as seen on the P-wave lags. If the average P-wave velocity of 3643m/s is considered, this translates to location errors as large $\approx$ 150m from the travel time data alone. With the inclusion of noisy back-azimuth estimates, this error estimate increase.

The horizontal locations of the microseismic events are mainly constrained by back-azimuth estimates for a single vertical borehole, and these parameters are highly sensitive to noise. Although taking the difference of the back-

azimuths derived using eqs. 3.10 and 3.12 can minimize common error due to similar ray paths, a large degree of uncertainty may remain since the absolute back-azimuths are estimated independently, and noise effects on the estimates may not be similar. Additionally, the computation of covariance matrices for individual events at each receiver can increase processing time. Obtaining relative estimates between the singular vectors and child events analytically may be advantageous in terms of speed and consistency. We compared the relative azimuths obtained using the difference between the absolute back-azimuths to those analytically determined using cross correlation. Highly similar waveforms are assumed to have small separation distances suggesting their angular differences should also be relatively small. We therefore used that assumption as a quality control measure by observing how the relative azimuths varied with the correlation coefficients between the events and singular vector. The results are shown in figure 6.10.



(a) Difference of absolute azimuths



(b) Cross-correlation derived relative azimuths

Figure 6.10: Relative back-azimuths between the events and their first singular vector as a function of the correlation coefficient across all receivers. a) Estimates obtained by differencing absolute back-azimuths. b) Estimates obtained analytically through cross-correlation.

There is greater uncertainty in the relative azimuths obtained from the difference of two independently estimated absolute azimuths compared to analytically determined estimates. In fact, for correlation coefficients of 0.8 and greater, the analytically determined relative azimuths do not exceed 30^0 as shown in figure 6.11, compared to the differencing approach which had differences as large as 180^0 at those correlation levels.

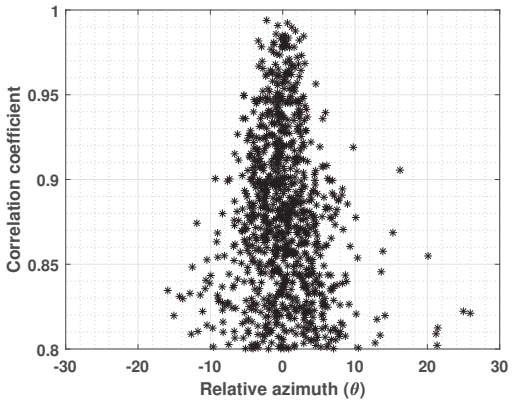


Figure 6.11: Relative azimuths obtained analytically for events with correlations of 0.8 and greater across all receivers. The variance in the relative azimuth estimates decreases with increasing values of the correlation coefficients. Most of the differential azimuths lie around 10^0.

By aligning the waveforms using cross-correlation, the signal-to-noise ratio is maximized resulting in more accurately determined relative back-azimuths. With the differential data, both subspace relocation and master event relocation were applied to the events. The resulting locations from subspace relocation and master event relocation are shown in figure 6.12.

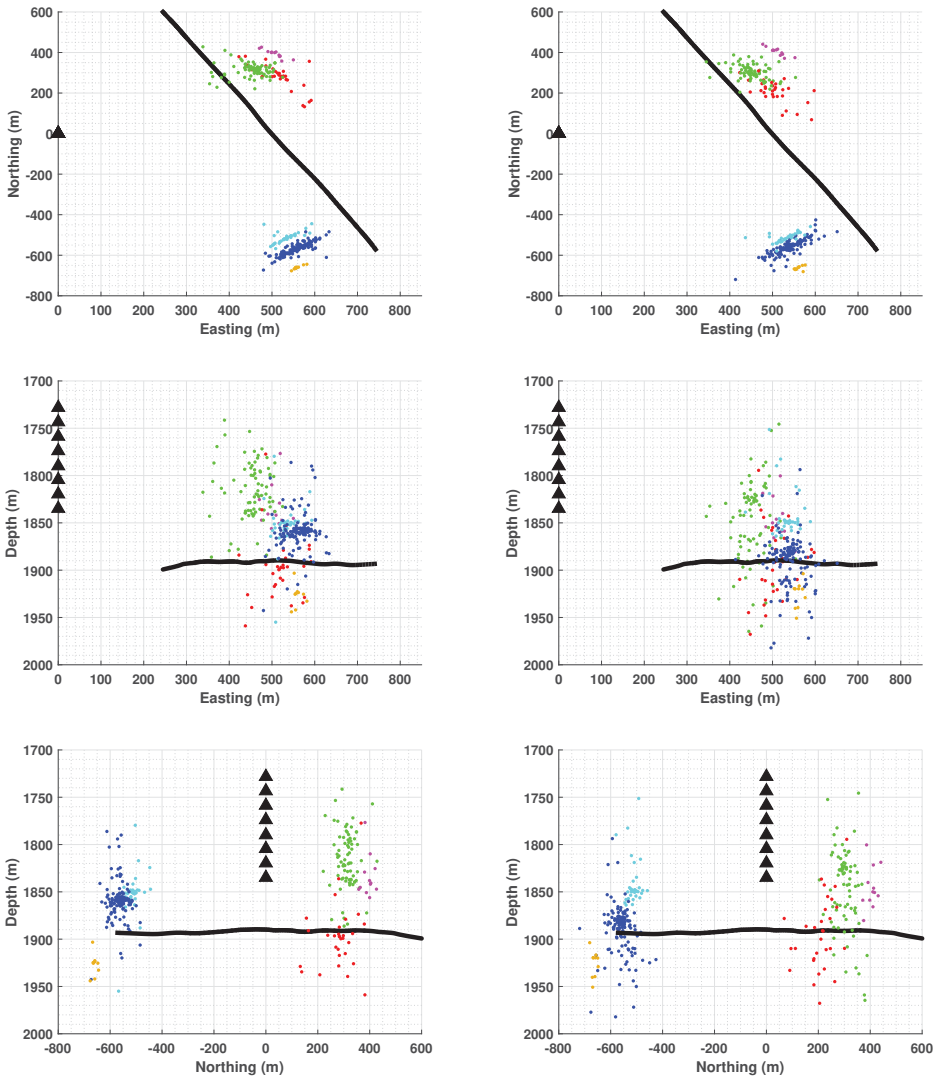


Figure 6.12: Relocated events shown as 2D illustrations. Receiver array shown as black triangles while treatment well shown as solid black line. The colours correspond to the different event clusters and are denoted by: Red (1), Magenta (2), Green (3), Blue (4), Orange (5) and Cyan (6). *Column 1*: Subspace relocation *Column 2*: Master event relocation.

In both relocation procedures the resulting seismic images are very similar, with the absolute locations of the clusters locating close to their respective master event positions. However, tighter clustering of the events is evident in the subspace relocation results, which led to more defined orientation of fault structures. The strike directions especially are better defined in subspace relocation compared to the master event results. The events were also plotted

by stage after relocation and are shown in figure 6.13.

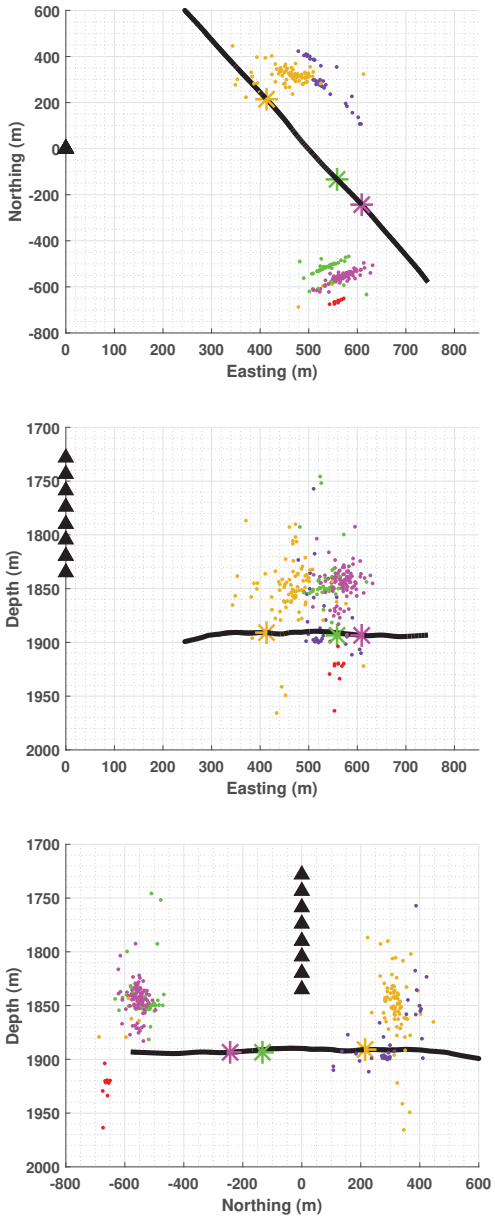


Figure 6.13: Locations of the detected events grouped by the treatment stage during which they occurred after subspace relocation. Receiver array shown as black triangles while treatment well shown as solid black line. The injection points are shown as stars. The colours correspond to the different stages and are denoted by: Red (1), Magenta (4), Green (5), Orange (8) and Purple (post pumping). Only stages containing more than 5 events are shown

The majority of the different clusters appear to be constrained to the stages

during which they occurred. However, some events from the same stages occur at different locations. This is especially evident in stage 5 (green) which has a few events occurring around the same location as stage 4 (magenta). A small number events from stage 8 (orange) also occur around the location of the stage 4 (magenta) events.

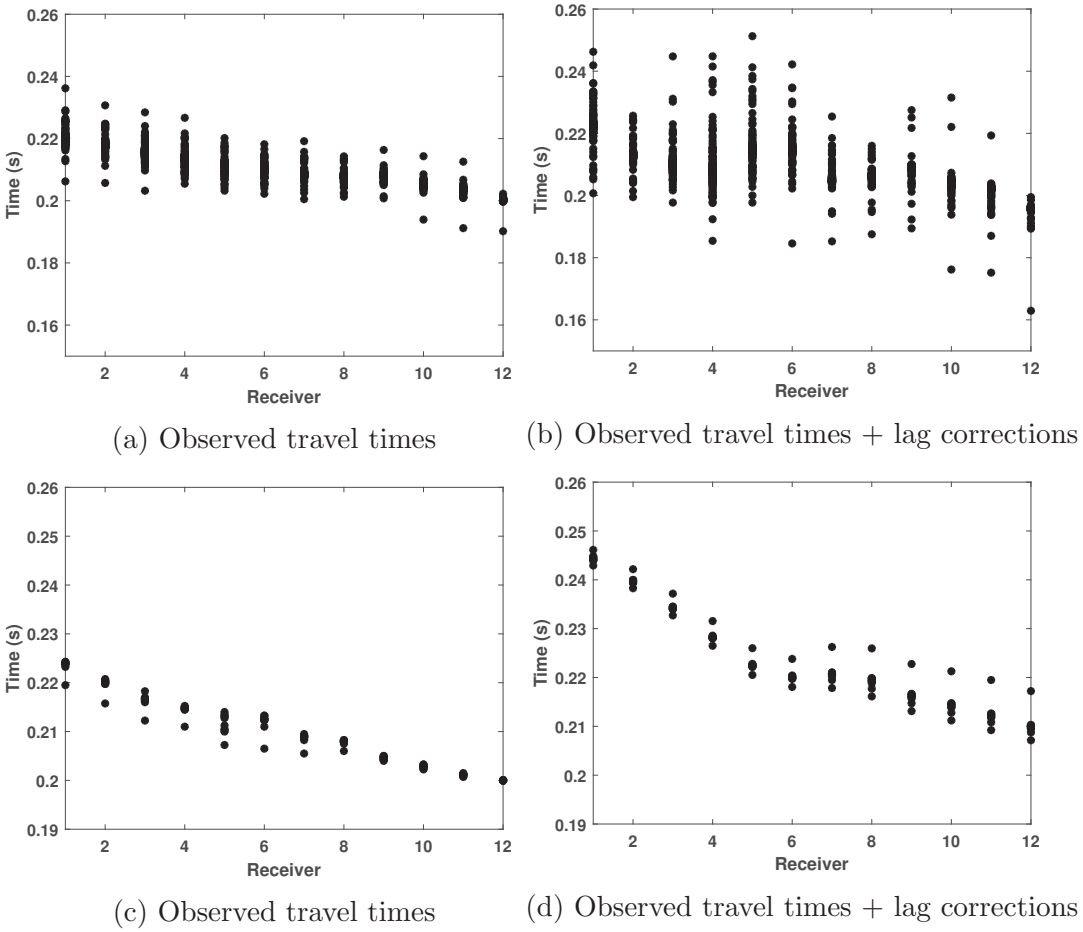
6.4 Discussion

The use of the 6 multiplet groups containing a total of 141 events resulted in a three-fold increase in the number of events. However, poor azimuthal coverage of the acquisition geometry, as well as high amplitude noise restricted the number of events available for processing, especially waveforms with no evident P-wave energy. Nevertheless, the events left for processing amounted to a two-fold increase in the event catalogue, which allows for increased resolution of the seismic image. The distribution of the detected microseismic events from the absolute locations suggest the presence of two main fault structures at different locations across the treatment well. At each treatment stage, the microseismic events migrated up the well bore as seen in figure 6.6. However, the resolution of the seismic image is poor from the grid search locations. Only one of the structures had a well defined strike, and very poor resolution on the dip of the structures. The use of wider grid spacing means events which are closer than the limits of the grid locate on the same position in space. This did not appear to have a significant negative impact on the absolute locations as the predicted travel times were fairly close to the observed times. Additionally, the locations of the events were consistent with the location of the stage specific injection points.

After relocation, tighter clustering of the events was evident, especially in the horizontal direction. In both subspace relocation and master event relocation, the events were pulled closer to their respective masters. Concentrated regions of seismicity were more evident in the relocated image and details on the fault orientations were sharper. The detected events which occurred during the early stages of treatment were aligned in the SW-NE directions, while the

microseismic events which occurred towards the end of the treatment and post pumping were aligned in the opposite direction. Eaton et al. (2014) interpreted the misaligned clusters to represent reactivation of pre-existing fractures, and the SW-NE direction to correspond to the present-day maximum horizontal stress direction. The variability in the fracture orientations was interpreted to be due to variations in reservoir facies within the barrier bar depositional environment. The occurrence of events from the same stage at different locations suggest more than one multiplet group was active in some stages. In figure 6.13, some events from stage 5 occur around the location of stage 4 events. A few stage 8 events also occurred around the vicinity of stage 4 events. This suggests that the cluster occurring during stage 4 may have been reactivated several times during the course of the treatment. This cluster also had the most events detected by the subspace detector, which may imply it was the most dominant source mechanisms present during the treatment.

To verify the accuracy of the subspace relocation procedure, travel times from the new locations obtained via subspace relocation were calculated to the receiver stations and compared to the original travel times. The cross-correlation derived lags were added to the observed travel times to see how consistent they were with the predictions. Figure 6.14 shows the observed P-wave travel times from clusters 4 and 5, as well as their corrected travel times using cross correlation lags.



(a) Observed travel times

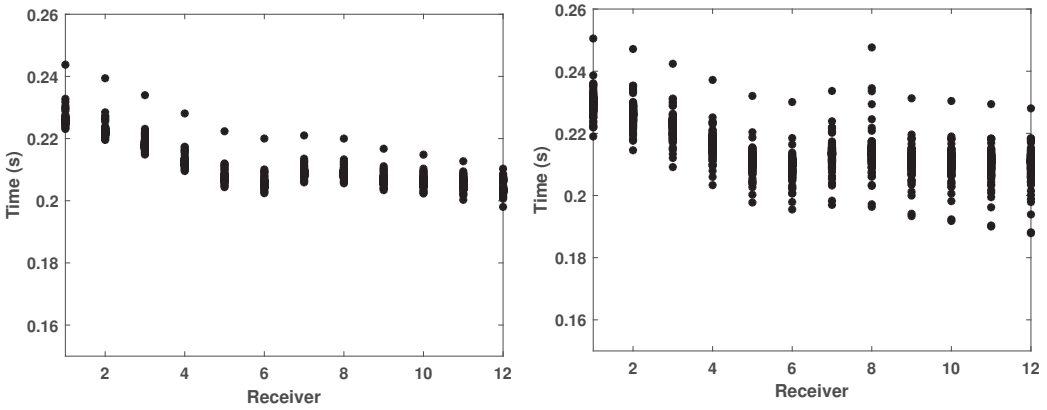
(b) Observed travel times + lag corrections

(c) Observed travel times

(d) Observed travel times + lag corrections

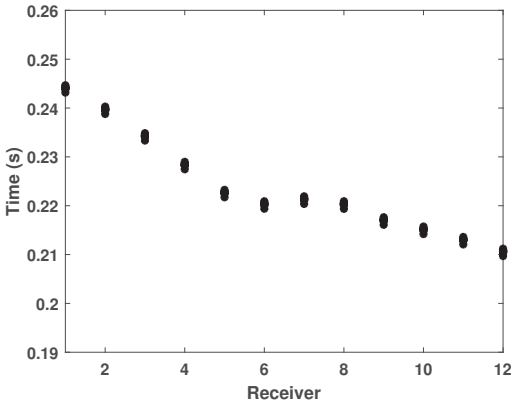
Figure 6.14: Observed P-wave travel times and the lag corrected P-wave observed times. *Top row*: Cluster 4. *Bottom row*: Cluster 5.

In both cases, the added lags have resulted in a slight upwards shifts in the observed travel times. Figure 6.15 shows the resulting predicted travel times for both the grid search locations and relocations.

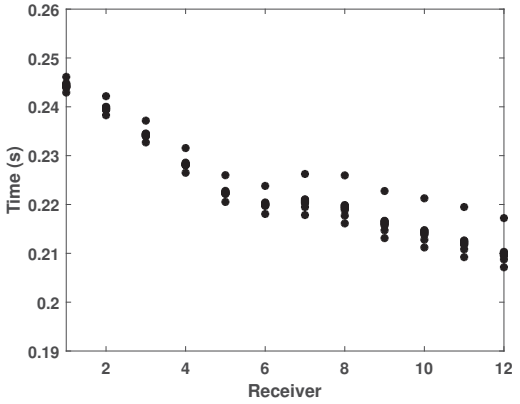


(a) Predicted travel times (grid)

(b) Predicted travel times (relocation)



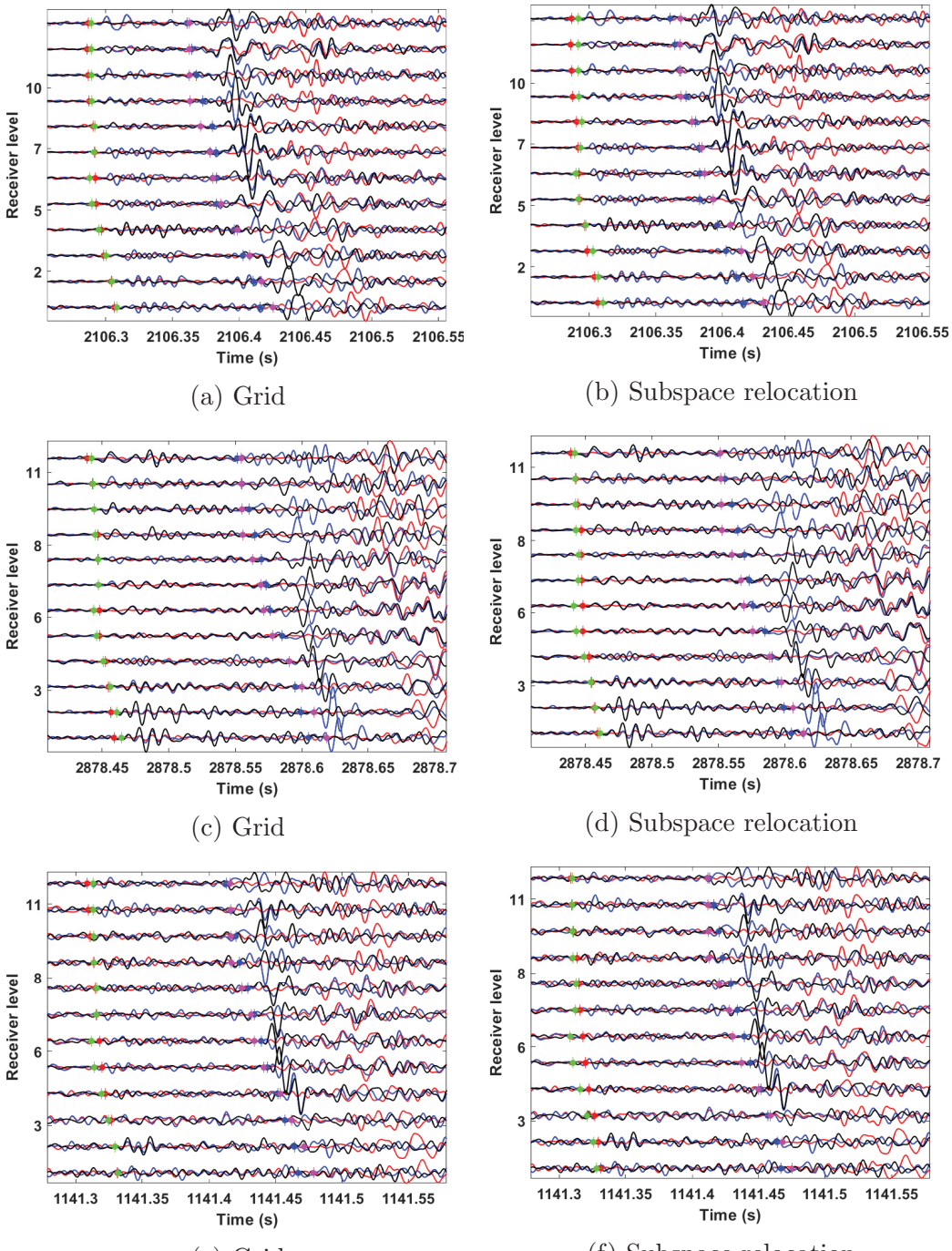
(c) Predicted travel times (grid)



(d) Predicted travel times (relocation)

Figure 6.15: Predicted travel times for clusters 4 and 5. *Top row*: Cluster 4. *Bottom row*: Cluster 5.

The new predicted times have also been shifted upwards relative to the initial predictions, which highlight the consistency between the lag corrected observed times and new locations obtained via relocation. To further validate the reliability of the relative locations, the predicted travel times were compared to the original arrival time estimates and predictions from the grid search. Figure 6.16 shows events from three clusters with the predicted travel times from the grid search and after subspace relocation.



(a) Grid

(b) Subspace relocation

(c) Grid

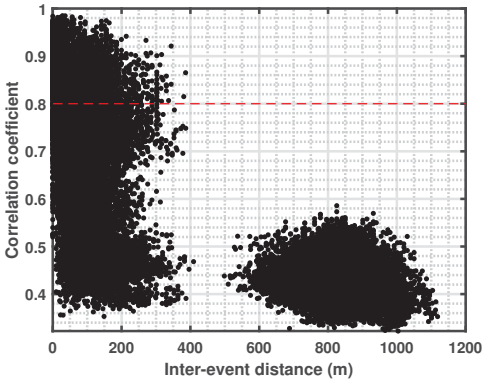
(d) Subspace relocation

(e) Grid

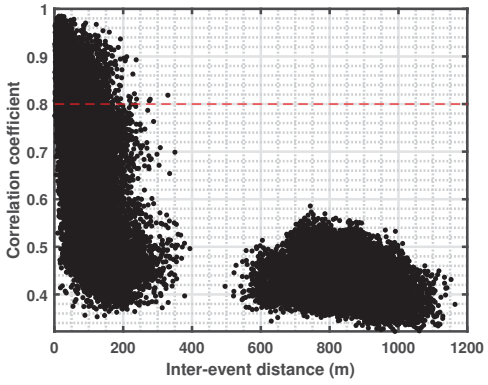
(f) Subspace relocation

Figure 6.16: Events from clusters 3 (top panel), 4 (middle panel) and 6 (bottom panel), with their observed and predicted arrival times. The red and blue dots correspond to the picked observed P-wave and S-wave arrival times, while the green and magenta dots denote the predicted P-wave and S-wave arrival times. *left column*: Grid search. *Right column*: Subspace relocation.

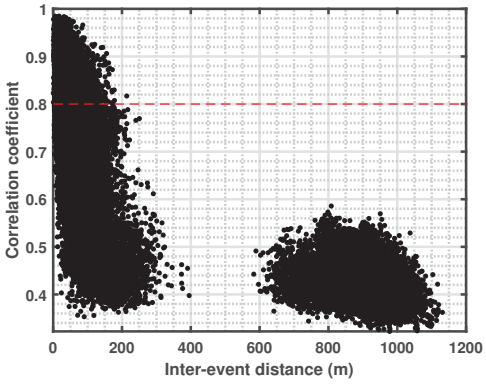
The predicted travel times from the relocated positions appear to be explaining the events pretty well, and lie reasonably close to the onset time position of the waveforms on each receiver. A measure of the accuracy of the relative locations can be obtained by looking at the separation distances between events before and after relocation. For a group of highly similar events, the expectation would be for similar events to locate close to each other. Figure 6.17 shows the inter-event distances between events from the grid search locations, subspace relocation and master event relocation as a function of their correlation coefficients, using a 330ms window which encompassed both P- and S-waves.



(a) Grid search



(b) Master event relocation

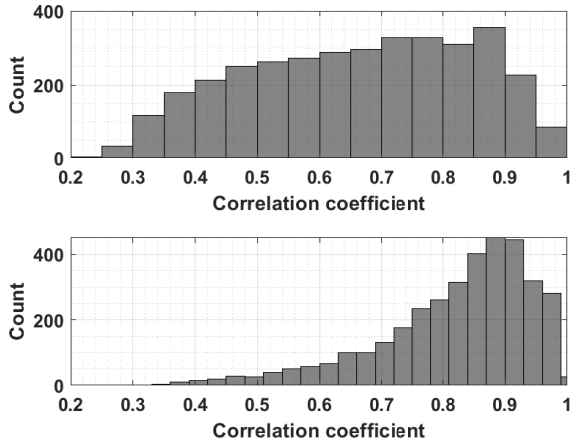


(c) Subspace relocation

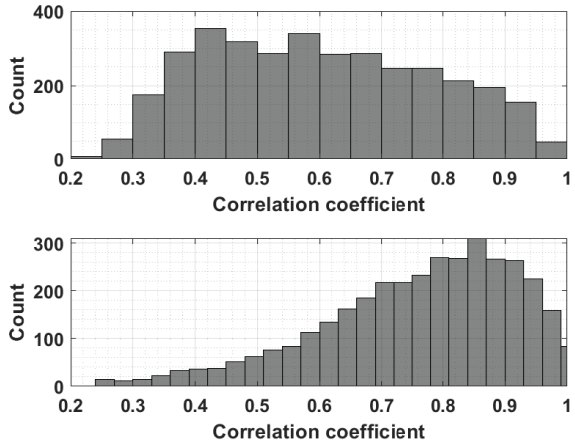
Figure 6.17: Separation distances as a function of correlation coefficient between events. The correlation level at which events are considered similar is shown as dashed red line and corresponds to 0.8

The initial grid search locations had separation distances as large as 350m

for highly similar events. After relocation, this dropped down to about 250m for master event relocation and about 150m for subspace relocation. In the relocated images, although a large majority of events from specific clusters were pulled together, some outliers were still evident. To understand these results, the distribution of correlation coefficients is examined for both the master and singular vectors. The quality of both the arrival time lag corrections and estimated relative back-azimuths are dependent on the correlation coefficients, and the distribution is shown in figures 6.18



(a) First singular vector



(b) Master event

Figure 6.18: Histogram of correlation coefficients across all the receivers between events and a: 1^{st} left singular vectors b: Master events. In both figures, the top panel shows the P-wave correlations while the bottom panel shows the S-wave correlations.

The distribution of correlation coefficients between the events and singular vectors is more right skewed compared to the master events, which suggests the singular vectors had higher signal-to-noise ratios. This in turn implies higher quality lag corrections and thus more reliable relocations. However, there is a prevalence of low correlation coefficients especially on the P-wave data in this data set, which may have introduced errors in the relocation procedure via less accurate lag estimates. Additionally, figure 6.11 reveals that even for high correlation coefficients, relative azimuths were as high as 25^o, which may have added additional error in the relative distances. In this experiment the detection window included both P- and S-waves, which may have allowed for one phase to dominant the detection procedure.

Nevertheless, the use of differential measurements allowed for more precise estimation of relative distances compared to absolute data. In this study, the lags were obtained between the multiplexed waveforms which implies they cannot be more precise than the sampling interval of the data. This limitation may be remedied by interpolation to increase the sampling rate, which allows for more precise lag corrections. The quality of the lags is also dependent on the quality of the correlation coefficients. Due to the restricted nature of this acquisition geometry, the P-wave data had to be included and this may have reduced the resolution of the relative locations. More accurate results will also be obtained with an increase in the number of recording stations as they provide additional constraints and may increase the amount of good quality data in the system.

6.5 Summary

Waveform similarity is useful for detecting similar events at low false alarm rates, as well as relocating microseismic events which originate in the same source region. The subspace detector has been shown to be a useful tool for simultaneous detection and relocation of microseismic events. The use of low-rank approximations for multiplet groups substantially increases catalog information during detection, and the improvements in the signal-to-noise ratio offered by the singular vectors was shown to minimize location uncertainty compared to regular grid search locations and the conventional master event relocation method.

Chapter 7

Conclusions

Microseismic monitoring is a useful tool in delineating fracture networks and highlighting stress fields. It has been widely applied in real-time imaging during hydraulic fracturing of unconventional hydrocarbon reservoirs, as well as in underground mining operations where it can reveal excavation damaged zones. The geometry of the fracture networks are outlined by the locations of the microseismic earthquakes and allows for interpretation of subsurface processes occurring. The level of detail possible from such an analysis is directly related to the volume of microseismic events since larger numbers reveal more information. Additionally, accurate locations provide sharper images of seismic and can allow for more precise identification of zones of interest.

Our objective in this thesis was to examine the subspace detector as a tool for enhancing detection capabilities of microseismic events and improving their hypocenter location. This objective was met with major challenges in the form of low signal-to-noise ratio of the microseismic data, restricted acquisition geometry with poor receiver coverage and large uncertainties in both the estimated arrival times of the seismic phases and back-azimuths of the events from their source to the receiver locations.

The subspace technique was successfully applied to microseismic data from a hydraulic fracturing treatment in the Hoadley gas field of central Alberta. The steps involved in constructing a subspace detector were described and we

demonstrated the ability of the detector to utilize either groups of highly similar events as templates or entire waveform libraries. Through synthetic tests, the left singular vectors were shown to have higher signal-to-noise ratios relative to the templates used in its construction. We then compared this technique to some standard detection techniques using real data and showed its ability to improve the number of detections relative to these techniques at reasonably low false alarm rates while saving on computational costs.

A novel approach to event location was also introduced which involved using the left singular vectors of the subspace detector as master events in a relocation procedure.The left singular vectors of the subspace detector describe the characteristics of the templates used in the construction of the detector and were assumed to also contain information on the hypocentral properties, a fact which we explored. We examined how the templates used in the construction of the subspace detector influenced the hypocenter location of the singular vector and discovered hypocentral information on the templates used in the subspace design were mainly contained in the 1^{st} singular vector. This phenomenon was seen to be limited to only templates with a high degree of similarity between them. We then described how this could be used in a master event relocation procedure and outlined the steps involved. This methodology was validated by performing a location of the singular vector via synthetic test, and they were foud to be close to their design set events. A comparison to the standard double difference algorithm and the master event relocation technique via synthetic examples showed fairly similar results in terms of the relative location accuracy with lower processing costs.

Application of this master event relocation technique independently for multiplet groups significantly reduced errors in the relative event locations by collapsing the clusters into more defined zones of seismicity. We outlined how accurate absolute arrival times, relative travel times and back-azimuths could be simultaneously obtained using cross-correlation as well as the waveform signal directly without the need for a covariance matrix. Additionally, we demonstrated the usefulness of weighting schemes in the inversion for our data

acquisition set up which was fairly restricted and corrupted by noise.

The travel time residuals, histograms of time-lags and correlations as well as cross-plots of event correlations versus separation distances were used as quality control measures to validate the accuracy of the event locations. The increased detail in the fault orientation after relocation allowed for some inferences regarding the temporal pattern of the fault orientation during the treatment.

7.1 Suggested future research

1. The data utilized in this work was obtained from a single vertical borehole which had poor data constraints in terms of its azimuthal coverage, limited number of receivers and high noise levels. In the case where the receivers provide good azimuthal coverage, only a single phase may be used to limit the influence of velocity uncertainty on the data with the inclusion of back-azimuths which tend to contain a lot of errors.

2. The double difference relocations have more accurate absolute locations compared to the master event since it utilizes all event pairs as opposed to a single link with the master. However, the left singular vectors have higher signal-to-noise ratios and describe the dominant hypocentral characteristics of its templates if there is high similarity of the events. It may be possible to include the left singular vectors as additional events in the double difference relocation algorithm to improve the linkage between the events and increase information in the system.

Bibliography

Allen, R. V., 1978, Automatic earthquake recognition and timing from single traces: Bulletin of the Seismological Society of America, **68**, 1521–1532.

Arrowsmith, S. J., and L. Eisner, 2006, A technique for identifying microseismic multiplets and application to the Valhall field, North Sea: Geophysics, **71**, V31–V40.

Aster, R. C., P. M. Shearer, and J. Berger, 1990, Quantitative measurements of shear wave polarizations at the Anza seismic network, southern California: Implications for shear wave splitting and earthquake prediction: Journal of Geophysical Research: Solid Earth, **95**, 12449–12473.

Baer, M., and U. Kradolfer, 1987, An automatic phase picker for local and teleseismic events: Bulletin of the Seismological Society of America, **77**, 1437–1445.

Barrett, S. A., and G. C. Beroza, 2014, An empirical approach to subspace detection: Seismological Research Letters, **85**, 594–600.

Bokelmann, G. H., 1995, P-wave array polarization analysis and effective anisotropy of the brittle crust: Geophysical Journal International, **120**, 145–162.

Bouchaala, F., V. Vavryčuk, and T. Fischer, 2013, Accuracy of the master-event and double-difference locations: synthetic tests and application to seismicity in West Bohemia, Czech Republic: Journal of seismology, **17**, 841–859.

Bulant, P., L. Eisner, I. Pšenčík, and J. Le Calvez, 2007, Importance of borehole deviation surveys for monitoring of hydraulic fracturing treatments: Geophysical Prospecting, **55**, 891–899.

Castellanos, F., and M. van der Baan, 2013, Microseismic event locations using the double-difference algorithm: CSEG Recorder, **38**, 26–37.

Chen, Y., H. Zhang, Y. Miao, Y. Zhang, and Q. Liu, 2017, Back azimuth constrained double-difference seismic location and tomography for downhole microseismic monitoring: Physics of the Earth and Planetary Interiors, **264**, 35–46.

Chen, Y., Z.-J. ZHANG, and X.-B. Tian, 2005, Complex polarization analysis based on windowed Hilbert transform and its application: Chinese Journal of Geophysics, **48**, 960–967.

Crosson, R. S., 1976, Crustal structure modeling of earthquake data: 1. Simultaneous least squares estimation of hypocenter and velocity parameters: Journal of geophysical research, **81**, 3036–3046.

De Meersman, K., J.-M. Kendall, and M. Van der Baan, 2009, The 1998 Valhall microseismic data set: An integrated study of relocated sources, seismic multiplets, and S-wave splitting: Geophysics, **74**, B183–B195.

De Meersman, K., M. Van der Baan, and J.-M. Kendall, 2006, Signal extraction and automated polarization analysis of multicomponent array data: Bulletin of the Seismological Society of America, **96**, 2415–2430.

Deichmann, N., and M. Garcia-Fernandez, 1992, Rupture geometry from high-precision relative hypocentre locations of microearthquake clusters: Geophysical Journal International, **110**, 501–517.

Douglas, A., 1967, Joint epicenter determination: Nature, **215**, 47–48.

Earle, P. S., and P. M. Shearer, 1994, Characterization of global seismograms using an automatic-picking algorithm: Bulletin of the Seismological Society of America, **84**, 366–376.

Eaton, D., E. Caffagni, M. van der Baan, and L. Matthews, 2014, Passive seismic monitoring and integrated geomechanical analysis of a tight-sand reservoir during hydraulic-fracture treatment, flowback and production: Unconventional Resources Technology Conference, Denver, Colorado, 25-27 August 2014, Society of Exploration Geophysicists, American Association of Petroleum Geologists, Society of Petroleum Engineers, 1537–1545.

Ehlers, J. F., 2002, Using the Fisher transform: TECHNICAL ANALYSIS OF STOCKS AND COMMODITIES-MAGAZINE EDITION-, **20**, 40–43.

Eisner, L., B. Hulsey, P. Duncan, D. Jurick, H. Werner, and W. Keller, 2010, Comparison of surface and borehole locations of induced seismicity: Geo-

physical Prospecting, **58**, 809–820.

Ellsworth, W. L., 1978, Three-dimensional structure of the crust and mantle beneath the island of Hawaii.: PhD thesis, Massachusetts Institute of Technology.

Freiberger, W. F., 1963, An approximate method in signal detection: Quarterly of Applied Mathematics, **20**, 373–378.

Frémont, M.-J., and S. D. Malone, 1987, High precision relative locations of earthquakes at Mount St. Helens, Washington: Journal of Geophysical Research: Solid Earth, **92**, 10223–10236.

Frohlich, C., 1979, An efficient method for joint hypocenter determination for large groups of earthquakes: Computers & Geosciences, **5**, 387–389.

Geller, R. J., and C. S. Mueller, 1980, Four similar earthquakes in central California: Geophysical Research Letters, **7**, 821–824.

Gentili, S., and A. Michelini, 2006, Automatic picking of P and S phases using a neural tree: Journal of Seismology, **10**, 39–63.

Gibbons, S. J., and F. Ringdal, 2006, The detection of low magnitude seismic events using array-based waveform correlation: Geophysical Journal International, **165**, 149–166.

Got, J.-L., J. Fréchet, and F. W. Klein, 1994, Deep fault plane geometry inferred from multiplet relative relocation beneath the south flank of Kilauea: Journal of Geophysical Research: Solid Earth, **99**, 15375–15386.

Han, L., 2010, Microseismic monitoring and hypocenter location: PhD thesis, University of Calgary.

Hansen, S. M., and B. Schmandt, 2015, Automated detection and location of microseismicity at Mount St. Helens with a large-N geophone array: Geophysical Research Letters, **42**, 7390–7397.

Harris, D. B., 2006, Subspace detectors: theory: Technical report, Lawrence Livermore National Lab.(LLNL), Livermore, CA (United States).

Havskov, J., and G. Alguacil, 2004, Instrumentation in earthquake seismology.

Havskov, J., and L. Ottemöller, 2010, *in* Routine data processing in earthquake seismology: Springer, 101–149.

Hayes, B., J. Christopher, L. Rosenthal, G. Los, B. McKercher, D. Minken, Y. Tremblay, J. Fennell, and D. Smith, 1994, Cretaceous Mannville Group of

the western Canada sedimentary basin, *in* Geological Atlas of the Western Canada sedimentary basin: Canadian Society of Petroleum Geologists and Alberta Research Council . . . , **4**, 317–334.

Herrmann, R. B., S.-K. Park, and C.-Y. Wang, 1981, The Denver earthquakes of 1967-1968: Bulletin of the Seismological Society of America, **71**, 731–745.

Husen, S., and J. Hardebeck, 2011, Understanding seismicity catalogs and their problems: Community Online Resource for Statistical Seismicity Analysis.

Jin, Y., and B. Friedlander, 2005, A CFAR adaptive subspace detector for second-order Gaussian signals: IEEE Transactions on Signal Processing, **53**, 871–884.

Jones, G., S. Nippress, A. Rietbrock, and J. Reyes-Montes, 2008, Accurate location of synthetic acoustic emissions and location sensitivity to relocation methods, velocity perturbations, and seismic anisotropy: Pure and Applied Geophysics, **165**, 235–254.

Jurkevics, A., 1988, Polarization analysis of three-component array data: Bulletin of the seismological society of America, **78**, 1725–1743.

Kirsteins, I. P., and D. W. Tufts, 1994, Adaptive detection using low rank approximation to a data matrix: IEEE Transactions on Aerospace and Electronic Systems, **30**, 55–67.

Lee, W. H. K., W. H. K. Lee, X. Lee, and S. Stewart, 1981, Principles and applications of microearthquake networks: Academic press, **2**.

Leonard, M., and B. Kennett, 1999, Multi-component autoregressive techniques for the analysis of seismograms: Physics of the Earth and Planetary Interiors, **113**, 247–263.

Li, J., H. Zhang, W. L. Rodi, and M. N. Toksoz, 2013, Joint microseismic location and anisotropic tomography using differential arrival times and differential backazimuths: Geophysical Journal International, **195**, 1917–1931.

Lomax, A., A. Michelini, A. Curtis, et al., 2007, Earthquake location, direct, global-search, *in* Encyclopedia of Complexity and System Science: Springer.

Mendel, J. M., 1991, Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications: Proceedings of the IEEE, **79**, 278–305.

Menke, W., 1999, Using waveform similarity to constrain earthquake locations:

Bulletin of the Seismological Society of America, **89**, 1143–1146.

Molyneux, J. B., and D. R. Schmitt, 1999, First-break timing: Arrival onset times by direct correlation: Geophysics, **64**, 1492–1501.

Montalbetti, J. F., and E. R. Kanasewich, 1970, Enhancement of teleseismic body phases with a polarization filter: Geophysical Journal International, **21**, 119–129.

Newbert, J., M. Trick, et al., 1987, A systematic deliverability analysis for the Hoadley gas condensate field: Presented at the SPE Annual Technical Conference and Exhibition, Society of Petroleum Engineers.

Nguyen, D., R. Brown, and D. Lawton, 1989, Polarization filter for multi-component seismic data: CREWES Research Report, **7**, 93.

Oprsal, I., and L. Eisner, 2014, Cross-correlation—An objective tool to indicate induced seismicity: Geophysical Journal International, **196**, 1536–1543.

Oye, V., and M. Roth, 2003, Automated seismic event location for hydrocarbon reservoirs: Computers & Geosciences, **29**, 851–863.

Phillips, W., T. Fairbanks, J. Rutledge, and D. Anderson, 1998, Induced microearthquake patterns and oil-producing fracture systems in the Austin chalk: Tectonophysics, **289**, 153–169.

Powers, D. M., 2011, Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation.

Reynolds, M. M., S. Thomson, D. J. Quirk, M. B. Dannish, F. Peyman, A. Hung, et al., 2012, A direct comparison of hydraulic fracture geometry and well performance between cemented liner and openhole packer completed horizontal wells in a tight gas reservoir: Presented at the SPE Hydraulic Fracturing Technology Conference, Society of Petroleum Engineers.

Rutledge, J. T., and W. S. Phillips, 2003, Hydraulic stimulation of natural fractures as revealed by induced microearthquakes, Carthage Cotton Valley gas field, east Texas Hydraulic Stimulation of Natural Fractures: Geophysics, **68**, 441–452.

Ry, R. V., T. Septyana, S. Widiyantoro, A. Nugraha, and A. Ardjuna, 2017, Improved location of microseismic events in borehole monitoring by inclusion of particle motion analysis: a case study at a CBM field in Indonesia: IOP Conference Series: Earth and Environmental Science, IOP Publishing,

012025.

Sasaki, Y., et al., 2007, The truth of the F-measure: Teach Tutor mater, **1**, 1–5.

Scharf, L. L., and B. Friedlander, 1994, Matched subspace detectors: IEEE Transactions on signal processing, **42**, 2146–2157.

Shearer, P. M., 1997, Improving local earthquake locations using the L1 norm and waveform cross correlation: Application to the Whittier Narrows, California, aftershock sequence: Journal of Geophysical Research: Solid Earth, **102**, 8269–8283.

Sleeman, R., and T. Van Eck, 1999, Robust automatic P-phase picking: an on-line implementation in the analysis of broadband seismogram recordings: Physics of the earth and planetary interiors, **113**, 265–275.

Song, F., N. R. Warpinski, M. N. Toksöz, and H. S. Kuleli, 2014, Full-waveform based microseismic event detection and signal enhancement: An application of the subspace approach: Geophysical Prospecting, **62**, 1406–1431.

Stoddard, P. R., and M. T. Woods, 1990, Master event relocation of Gorda block earthquakes: Implications for deformation: Geophysical Research Letters, **17**, 961–964.

Swindell, H., and N. Snell, 1977, Station processor automatic signal detection system, phase I: final report, station processor software development, Texas Instruments Report No: ALEX (01)-FR-77-01.

Trnkoczy, A., 1999, Topic understanding and parameter setting of STA/LTA trigger algorithm: New manual of seismological observatory practice, **2**.

Urkowitz, H., 1967, Energy detection of unknown deterministic signals: Proceedings of the IEEE, **55**, 523–531.

Vaezi, Y., and M. Van der Baan, 2015, Comparison of the STA/LTA and power spectral density methods for microseismic event detection: Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society, **203**, 1896–1908.

Van Der Baan, M., D. Eaton, M. Dusseault, et al., 2013, Microseismic monitoring developments in hydraulic fracture stimulation: Presented at the ISRM International Conference for Effective and Sustainable Hydraulic Fracturing, International Society for Rock Mechanics and Rock Engineering.

Van der Elst, N. J., H. M. Savage, K. M. Keranen, and G. A. Abers, 2013, Enhanced remote earthquake triggering at fluid-injection sites in the mid-western United States: Science, **341**, 164–167.

Van Trees, H. L., 2004, Detection, estimation, and modulation theory, part I: detection, estimation, and linear modulation theory: John Wiley & Sons.

Vanderkulk, W., F. Rosen, and S. Lorenz, 1965, Large aperture seismic array signal processing study: IBM Final Report, ARPA Contract Number SD-296.

Vidale, J. E., 1986, Complex polarization analysis of particle motion: Bulletin of the Seismological society of America, **76**, 1393–1405.

Waldhauser, F., and W. L. Ellsworth, 2000, A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, california: Bulletin of the Seismological Society of America, **90**, 1353–1368.

Wall, M. E., A. Rechtsteiner, and L. M. Rocha, 2003, Singular value decomposition and principal component analysis, *in* A practical approach to microarray data analysis: Springer, 91–109.

Wiechecki-Vergara, S., H. L. Gray, and W. A. Woodward, 2001, Statistical development in support of CTBT monitoring: Southern Methodist University.

Withers, M., R. Aster, C. Young, J. Beiriger, M. Harris, S. Moore, and J. Trujillo, 1998, A comparison of select trigger algorithms for automated global seismic phase and event detection: Bulletin of the Seismological Society of America, **88**, 95–106.

Withers, M. M., R. C. Aster, C. J. Young, and E. P. Chael, 1996, High-frequency analysis of seismic background noise as a function of wind speed and shallow depth: Bulletin of the Seismological Society of America, **86**, 1507–1515.

Wong, J., L. Han, J. Bancroft, and R. Stewart, 2009, Automatic time-picking of first arrivals on noisy microseismic data: CSEG. 0 0.2 0.4 0.6 0.8, **1**, 1–4.

Yilmaz, Ö., 2001, Seismic data analysis: Processing, inversion, and interpretation of seismic data: Society of exploration geophysicists.

Zhang, H., C. Thurber, and C. Rowe, 2003, Automatic P-wave arrival detection and picking with multiscale wavelet analysis for single-component recordings: Bulletin of the Seismological Society of America, **93**, 1904–1912.

Appendix A

Description of partial derivatives

In this section, the partial derivatives with respect to the earthquake model parameters are shown. Eq. 3.2 corresponds to the travel time equation of a ray from an earthquake i at a given source location (x^i,y^i,z^i) and origin time τ, to station k. Its partial derivatives with respect to the model parameters are given by the following equations

$$\frac{\partial t_k^i}{\partial x} = \frac{1}{v}\frac{(x^i - x^k)}{d_i} \tag{A.1}$$

$$\frac{\partial t_k^i}{\partial y} = \frac{1}{v}\frac{(y^i - y^k)}{d_i} \tag{A.2}$$

$$\frac{\partial t_k^i}{\partial z} = \frac{1}{v}\frac{(z^i - z^k)}{d_i} \tag{A.3}$$

$$\frac{\partial t_k^i}{\partial \tau} = 1 \tag{A.4}$$

In the above equations, v can become a function of depth or spatial coordinates depending on the velocity model used, while d_i corresponds to the distance between the assumed source location and receiver coordinates. The back-azimuth of a ray from a source to a receiver location is described by eq. 3.4, and its partial derivatives are given by

$$\frac{\partial \theta_k^i}{\partial x} = \frac{(y^i - y^k)}{(x^i - x^k)^2 + (y^i - y^k)^2} \tag{A.5}$$

$$\frac{\partial \theta_k^i}{\partial y} = -\frac{(x^i - x^k)}{(x^i - x^k)^2 + (y^i - y^k)^2} \tag{A.6}$$

$$\frac{\partial \theta^i}{\partial z} = 0 \tag{A.7}$$

Appendix B

Determination of relative back-azimuths

This chapter describes how the differential back-azimuths between two events are analytical calculated, based on the method of Li et al. (2013). The azimuthal angles are defined in this work as increasing clockwise from the positive y-axis, which corresponds to the true north direction. Let $\mathbf{w}_1$(t) correspond to a data vector containing a waveform of interest, beginning on the arrival time estimate of the phase. If $\mathbf{w}_1$(t) has a back-azimuth of θ_1 from its source to a reference receiver, then its north and east components can be given as

$$\mathbf{w}_{n_1}(t) = \mathbf{w}_1(t)cos(\theta_1) \tag{B.1}$$

$$\mathbf{w}_{e_1}(t) = \mathbf{w}_1(t)sin(\theta_1) \tag{B.2}$$

The same expressions can be derived for a waveform $\mathbf{w}_2$(t) with a back-azimuth of θ_2

$$\mathbf{w}_{n_2}(t) = \mathbf{w}_2(t)cos(\theta_2) \tag{B.3}$$

$$\mathbf{w}_{e_2}(t) = \mathbf{w}_2(t)sin(\theta_2) \tag{B.4}$$

We can then derive expressions for the differential azimuths in terms of cosines and sines using double angle identities

$$\mathbf{w}_{e_1}^T(t)\mathbf{w}_{n_2}(t) + \mathbf{w}_{n_1}^T(t)\mathbf{w}_{e_2}(t) = \mathbf{w}_1^T(t)\mathbf{w}_2(t)[sin(\theta_1)cos(\theta_2) - cos(\theta_1)sin(\theta_2)]$$
$$\mathbf{w}_1^T(t)\mathbf{w}_2(t)[sin(\theta_1 - \theta_2)] \tag{B.5}$$

$$\mathbf{w}_{n_1}^T(t)\mathbf{w}_{n_2}(t) + \mathbf{w}_{e_1}^T(t)\mathbf{w}_{e_2}(t) = \mathbf{w}_1^T(t)\mathbf{w}_2(t)[cos(\theta_1)cos(\theta_2) + sin(\theta_1)sin(\theta_2)]$$
$$= \mathbf{w}_1^T(t)\mathbf{w}_2(t)[cos(\theta_1 - \theta_2)] \tag{B.6}$$

The differential azimuth can then be obtained by taking the arc-tangent of the expression below

$$tan(\theta_1 - \theta_2) = \frac{\mathbf{w}_1^T(t)\mathbf{w}_2(t)[sin(\theta_1 - \theta_2)]}{\mathbf{w}_1^T(t)\mathbf{w}_2(t)[cos(\theta_1 - \theta_2)]} = \frac{\mathbf{w}_{e_1}^T(t)\mathbf{w}_{n_2}(t) + \mathbf{w}_{n_1}^T(t)\mathbf{w}_{e_2}(t)}{\mathbf{w}_{n_1}^T(t)\mathbf{w}_{n_2}(t) + \mathbf{w}_{e_1}^T(t)\mathbf{w}_{e_2}(t)} \tag{B.7}$$

The waveforms in $\mathbf{w}_1$(t) and $\mathbf{w}_2$(t) should be aligned using waveform cross-correlation to maximize the signal-to-noise ratio in the dot products and improve the accuracy of the relative azimuth estimates.