

FEASIBILITY OF THE INTERFEROMETRIC NEIGHBORING FRACTURE  
METHOD AND IMPROVED RELATIVE TRAVEL TIME MEASUREMENT

A Thesis

by

ILAAN E. SHTAYGRUD

Submitted to the Office of Graduate and Professional Studies of  
Texas A&M University  
in partial fulfillment of the requirements for the degree of  
MASTER OF SCIENCE

Chair of Committee,	Richard Gibson Jr.
Committee Members,	Christine Ehlig-Economides Andreas Kronenberg
Head of Department,	Rick Giardino

December 2014

Major Subject: Geophysics

Copyright 2014 Ilaan E. Shtaygrud

## ABSTRACT

The recently proposed Interferometric Neighboring Fracture (INF) localization method places unique and demanding constraints on relative travel time measurement accuracy and precision, while sampling a function of relative travel times between pairs of microseismic events as measured along a linear array. Conducting two synthetic trials, I analyze the relationship between event-receiver geometry and relative travel time measurement error and its effect on the feasibility of INF localization. The results indicate that even for typical hydraulic fracturing monitoring geometries, measurement error can exceed the feasible error limits of INF localization.

In order to mitigate this error, I propose a new relative travel time measurement technique, Modified Adaptive Steering (MAS), along with a unique preprocessing methodology, Progressive Template Extraction (PTE). Analyzing synthetic data sets with varying SNR ratios, and a field recorded microseismic data set, I compare the performance of PTE-preprocessed MAS to conventional cross-correlation (CXC). Results of both synthetic and field recorded data analysis indicate that PTE enhanced MAS outperforms CXC as a general lag measurement technique, reducing average lag error by as much as 1.25 ms at SNRs below 10. With respect to the unique constraints of the INF method, PTE-MAS produces as many as 4.2 times as many usable samplings of the relative travel time function, while reducing error in stationary position and lag by up to 15 m and 2.5 ms, respectively.

To my parents and grandparents;  
it is their work and sacrifice upon which I stand.

## ACKNOWLEDGMENTS

First and foremost, I must acknowledge the perpetual and unconditional support of my mother and father; every one of my successes is their own.

I furthermore wish to express my appreciation for the guidance of Dr. Richard Gibson over the last 2.5 years. As my committee chair, Dr. Gibson has been instrumental in giving my research direction, from initial topic selection to the careful revision of this thesis, with many words of advice given in between. I am grateful for his time and effort.

My committee members, Dr. Andreas Kronenberg and Dr. Christine Ehlig-Economides, are similarly owed my gratitude. I thank Dr. Kronenberg for his helpful and encouraging comments and always prompt responses, as well as his critical early suggestion to seek an advisor whose interests were better aligned with my own. To Dr. Economides I owe my understanding of the context of my work, and I am thankful to have been included in her own research meetings.

Finally, I extend my appreciation to the members of my research group, for bearing through my unpolished presentations, and offering helpful suggestions and assistance over the course of my studies.

# TABLE OF CONTENTS

	Page
ABSTRACT . . . . .	ii
DEDICATION . . . . .	iii
ACKNOWLEDGMENTS . . . . .	iv
TABLE OF CONTENTS . . . . .	v
LIST OF FIGURES . . . . .	vii
LIST OF TABLES . . . . .	xi
1. INTRODUCTION . . . . .	1
1.1 Microseismic Monitoring . . . . .	1
1.2 Seismic Interferometry . . . . .	2
1.3 Thesis Purpose . . . . .	4
1.4 Thesis Structure . . . . .	4
2. INTERFEROMETRIC NEIGHBORING FRACTURE LOCALIZATION . . . . .	6
2.1 Method Description . . . . .	6
2.2 Ideal Geometry and the INF Method . . . . .	11
3. MONITORING GEOMETRY AND THE FEASIBILITY OF THE INF METHOD . . . . .	15
3.1 Receiver Array Characteristics and $dt(l)$ Resolution . . . . .	15
3.1.1 Model Description . . . . .	16
3.1.2 Test Model and Procedure . . . . .	16
3.1.3 Results . . . . .	19
3.2 Array Placement and Resolvable Stationary Pair Count . . . . .	24
3.2.1 Test Model and Procedure . . . . .	24
3.2.2 Results . . . . .	27
3.3 Summary . . . . .	29
4. IMPROVED RELATIVE TRAVELTIME MEASUREMENT WITH MODIFIED ADAPTIVE STEERING . . . . .	32
4.1 Idealized Cross-Correlation Constraints and Microseismic Monitoring . . . . .	32
4.2 A Case Study of Cross-Correlation and the INF Method . . . . .	33

4.2.1	Dataset Description . . . . .	33
4.2.2	Events, Methodology, and Results . . . . .	34
4.2.3	Implications of Cross-Correlation Error . . . . .	35
4.2.4	Potential Mitigation of Misalignment Error . . . . .	37
4.2.5	Array Processing . . . . .	41
4.3	Modified Adaptive Steering and Progressive Template Extraction . .	41
4.3.1	Modified Adaptive Steering . . . . .	42
4.3.2	Progressive Template Extraction . . . . .	44
4.4	Synthetic Comparison of Modified Adaptive Steering versus Conventional Cross-Correlation . . . . .	47
4.4.1	Model Description and Comparison Framework . . . . .	47
4.4.2	General Lag Measurement Metrics and Results . . . . .	50
4.4.3	INF Focused Metrics and Results . . . . .	53
4.4.4	Indirect Performance Metrics and Results . . . . .	58
4.5	Real Data Evaluation of Modified Adaptive Steering versus Conventional Cross-Correlation . . . . .	62
4.6	Comparison of Adjacent Steering and CXC Under Idealized Conditions	65
4.6.1	Modified Adaptive Steering Advantages in Detail . . . . .	67
4.7	Summary . . . . .	70
5.	CONCLUSION . . . . .	72
	REFERENCES . . . . .	74
	APPENDIX A. SOURCES, PROPAGATION, AND WAVEFORM DISSIMILARITY . . . . .	79
	APPENDIX B. CROSS-CORRELATION ERROR IN DETAIL . . . . .	85

## LIST OF FIGURES

FIGURE		Page
2.1	Planar monitoring surface (green triangles), two hypothetical fractures with many events (red and blue points), and microseismic events $x_1$ and $x_2$ (black stars), with a stationary raypath shown by the black connecting line. . . . .	7
2.2	Linear receiver array (green triangles), reference and locatable fractures (red and blue points, respectively), and with events $x_1$ and $x_2$ (black stars). . . . .	8
2.3	For a linear array, a stationary pair of events may be rotated about the array axis without loss of stationarity, since travel times measured at the array remain constant. For a 1D velocity model, these constant-time loci are circles. . . . .	9
2.4	Relative travel time function $dt(l)$ as a function of receiver position $l$	12
2.5	Derivative of the relative travel time function $dt(l)$ as a function of receiver position $l$ . . . . .	13
3.1	Model used in the resolution trial. Two fractures, 100 events each, are separated by 400 meters. Example receiver locations are given by the blue crosses. . . . .	17
3.2	Results of trials with varying receiver spacings and orientations. Array orientations are ex, ey, and ez in the left, center, and right columns, respectively. Receiver spacings are 10 m, 20 m, and 30 m, from top to bottom row. Colorbar values are in seconds. Color scales are identical for all plots. . . . .	19
3.3	Basemap, showing locations from which histograms were taken, with reference and locatable fractures in blue and red, respectively. . . . .	22
3.4	Histograms at locations A, B, and C from the basemap in figure 3.3. $\Delta dt$ values are given by the red lines. Note how in each case, the actual distribution is somewhat skewed. . . . .	23

3.5	Stationary, resolvable pairs counted for arrays oriented in the $\hat{e}_x$ (top left), $\hat{e}_y$ (top right), and $\hat{e}_z$ (bottom) orientations. The $\hat{e}_x$ orientation maximum count is 391, while both $\hat{e}_y$ and $\hat{e}_z$ orientations reach 10000 (100%) stationary pairs. . . . .	27
3.6	Spatial distribution of nodes with at least 5000 stationary, resolvable event pairs (green). . . . .	29
4.1	Event locations given by light purple cubes. . . . .	34
4.2	Locations of events 0 and 4 (labeled stars) with respect to the receiver array (green triangles). . . . .	35
4.3	Traces for events 0 and 4. Event 4:event 0 scale ratio is approximately 10:1. . . . .	36
4.4	Measured relative arrival times for events 0 and 4. . . . .	37
4.5	Two first arrivals A and B from events 0 and 4, respectively, which could hypothetically be cross-correlated with each other to recover a relative arrival time. . . . .	38
4.6	Distribution of inter-event correlation coefficient magnitudes. . . . .	39
4.7	Receiver and event locations for this synthetic trial. Reference events are in red, locatable events are in blue. . . . .	49
4.8	$\bar{\epsilon}$ , average error of all measured lag values, as a function of SNR. CXC in blue, MAS in green, PTE enhanced MAS in red. . . . .	51
4.9	Average CXC correlation coefficients (blue), MAS inter-event correlation coefficients (green, dashed), MAS intra-event correlation coefficients (green, dotted), PTE enhanced MAS inter-event correlation coefficients (red, dashed), and PTE enhanced MAS intra-event correlation coefficients (red, dotted), as functions of SNR. . . . .	52
4.10	Distribution of lag errors calculated by conventional cross-correlation for trial with SNR=13. Bin width = 0.1625 ms. . . . .	52
4.11	Distribution of lag errors calculated by modified adaptive steering, without preprocessing, for trial with SNR=13. Bin width = 0.1625 ms. Note that the y scale has been clipped, the counts for bins 1 and 2 are 614 and 626, respectively. . . . .	53
4.12	False positive counts. CXC in blue, MAS in green, MAS with PTE in red. . . . .	55



4.13	False negative counts. CXC in blue, MAS in green, MAS with PTE in red. . . . .	56
4.14	$\bar{\epsilon}_{dt}$ , average stationary lag error for correctly identified stationary pairs, as a function of SNR. CXC in blue, MAS in green, PTE enhanced MAS in red. . . . .	57
4.15	$\bar{\epsilon}_l$ , average stationary position error for correctly identified stationary pairs, as a function of SNR. CXC in blue, MAS in green, PTE enhanced MAS in red. . . . .	58
4.16	Cross-plot of $\Delta\bar{\epsilon}$ and $\overline{\Delta\sigma}$ , as functions of SNR, with a line of best fit. The statistical correlation coefficient between the two functions is 0.92, indicating a strong correlation between standard deviation and overall error. . . . .	61
4.17	Count of $dt(l)$ measurements with consistent concavity for CXC (blue), MAS (green), and PTE enhanced MAS (red) . . . . .	62
4.18	$u$ , average difference in standard deviation of $dt(l)$ functions, as a function of SNR. MAS-CXC in green, PTEMAS-CXC in red. . . . .	63
4.19	$\langle\Delta\bar{t}\rangle$ , the average of difference in $dt(l)$ means, as a function of SNR. CXC-MAS in green, CXC-PTEMAS in red. Note that there does not appear to be any major trend with respect to SNR, although the variability in PTE enhanced MAS $\langle\Delta\bar{t}\rangle$ values is markedly smaller. . . . .	64
4.20	Traces for events 0 and 9. Event 9:event 0 scale ratio is approximately 16:1. . . . .	65
4.21	$dt(l)$ functions generated by conventional cross-correlation (black) and adaptive steering (gray). . . . .	66
4.22	Traces for event 14 which have been aligned with the adaptive steering method, along with the resulting stacked composite at the bottom. Note the large increase in relative amplitudes between the initial arrival and energy arriving after 60 ms, an example of the cancellation which MAS leverages. . . . .	69
A.1	Radiation Patterns for a double couple source. Blue lobes represent amplitudes of P-waves, red lobes show amplitudes for S-waves, and traces for receivers at varying angular offsets are shown in the top right corner. Note: lobes are not to scale. . . . .	80

A.2	Two fracture planes (blue) and two potential raypaths A and B to a hypothetical receiver array. Condition 1 is invalid, because of the angular difference from the array to the events, while condition 3 also fails to hold as the rays follow different paths. . . . .	81
B.1	Correlograms generated at each receiver for events 0 and 4 . . . . .	86

## LIST OF TABLES

TABLE	Page
3.1.1 Orientation and receiver spacing combinations used in the resolution trial. . . . .	18
3.1.2 Maximum $\overline{\Delta dt}$ values corresponding to the trials in figure 3.2. . . . .	20

# 1. INTRODUCTION

## 1.1 Microseismic Monitoring

The recent increase in tight reservoir development within the global petroleum industry has generated interest in the use of microseismic monitoring as a hydraulic stimulation monitoring tool. The study of microseismicity induced by fracturing rock is not strictly novel, as it has been a staple of geothermal field development since the 1970s (Maxwell et al., 2010), as well as a hazard detection tool within the mining industry for almost as long (Gibowicz and Kijko, 1994). Recent theoretical developments have been characterized by numerous advances, ranging from better source characterization (Sileny et al., 2009), to passive imaging and tomography (Zhang and Thurber, 2003), as well as improved hypocenter localization (Waldhauser and Ellsworth, 2000; Poliannikov et al., 2011), including the Interferometric Neighboring Fracture method which forms the focus of this thesis.

Microseismic event locations function as a direct measurement of fracture orientation, distribution, and density, qualities which enable operators to optimize hydraulic fracturing pressures, fluids, proppants, and well and stage spacings (Portis et al., 2013). Event density and distribution allow for the estimation of total stimulated reservoir volume (SRV), allowing engineers to evaluate stimulation efficacy and predict drainage (Meek et al., 2013) of stimulated reservoirs. Knowledge of fracture growth and containment, given by hypocenter distribution, may be used to delineate geological formation boundaries, supplementing well and mud logs by providing additional clues to subsurface geology in a volume surrounding the stimulation wellbore. These data additionally allow for inferences regarding reservoir connectivity, and facilitate the minimization of leakoff into and stimulation of non-reservoir forma-

tions. Waters et al. (2009) reported on the use of real time microseismic monitoring during stimulation of Barnett Shale to avoid fracture intrusion into the underlying, water bearing Ellenberger Limestone. During initial development of the Barnett Shale, microseismic analysis was used to demonstrate the existence of natural fracture networks, leading reservoir engineers to consider more complex fracture models for optimal stimulation planning (Mayerhofer et al., 2006).

## 1.2 Seismic Interferometry

Despite a number of recently proposed localization methodologies, literature reports (Havskov and Ottemöller, 2010; Schweitzer, 2001; Vesnaver et al., 2008), that contemporary seismologists still predominantly rely on Geiger and Wadati methods, and minor variations thereof, hereafter referred to as classical localization techniques. These classical methods, based on P- and S-wave arrival time differences and arrival polarization, locate events individually, and have been adapted from earthquake seismology, where meaningful analysis is possible with uncertainties on the order of hundreds or thousands of meters. The physical scales of hydraulically induced fracture networks more strongly constrain the useful limits of accuracy in event localization, such that errors of 10 s to 100 s of meters typically reported for classical localization (Hurd, 2012) restrict the quality and quantity of information which may be derived from hypocenter locations.

Recent applications of seismic interferometry to microseismic analysis have been proposed in an effort to mitigate these limitations, by generally promising to reduce the uncertainty of hypocenter inversion. Interferometric techniques compare event waveforms using cross-correlation, convolution, or some combination of the two, to supplement absolute event locations and individual event waveforms with relative event information, allowing for source-receiver virtualization (Curtis, 2009; Schuster,

2009; Poliannikov and Willis, 2011; Poliannikov et al., 2012), passive seismic imaging (Wapenaar, 2006), and event redatuming (Curtis and Halliday, 2010; Schuster and Zhou, 2006; Poliannikov and Willis, 2011). While interferometry is a powerful technique which may redatum full waveforms between pairs of events, relative event travel times alone can act as a useful set of data, particularly as input for relative localization methods, such as the Double Difference (DD) technique proposed in Waldhauser and Ellsworth (2000), as well as the Interferometric Neighboring Fracture method, and a Unified Bayesian combination of Double Difference and INF introduced in Poliannikov et al. (2013).

The Interferometric Neighboring Fracture method (INF) proposed in Poliannikov et al. (2011) is a promising relative localization methodology which is particularly useful for linear array event-receiver geometries typical of hydraulically induced microseismic events. As characteristic of interferometric techniques, INF makes use of relative event travel time information to reduce uncertainty in localized hypocenters, thereby improving the quality of information which may be inferred from microseismic event locations. The INF methodology places unique and demanding requirements on travel time measurements, such that care must be taken to ensure that event-receiver geometry and relative travel time precision are suitable for INF relocation; otherwise, as will be demonstrated in the following text, the methodology may be unable to relocate events, or it may produce nonsensical locations which may be taken as accurate without an understanding of the effect of geometrical and measurement limitations on INF localization results.

To date, only a handful of studies have succeeded in relocating microseismic events using relative event information, e.g. (Hurd, 2012; Melo et al., 2012; Phillips, 2000). Although the results of the application of Double Difference to microseismic localization in Hurd (2012) were less than ideal, the method, originally adapted

from earthquake seismology nonetheless shows promise as a microseismic relative localization tool. Based on the Geiger method in Geiger (1912), the methodology simultaneously minimizes absolute and relative traveltime residuals for all events and event pairs, reducing error induced by noise and velocity uncertainty. Compared to INF, DD is better suited to noisy data, while INF is less sensitive to velocity model error for certain geometries (Poliannikov et al., 2013). The Bayesian unification of the two methods in Poliannikov et al. (2013) leverages these situational advantages, by realizing locations under each methodology and choosing the result with the smaller uncertainty volume.

### 1.3 Thesis Purpose

Given the breadth of information that may be inferred from accurate microseismic hypocenter locations, in this thesis I propose two novel techniques, which I show to generally improve relative travel time measurement accuracy and precision, thereby reducing uncertainty in event locations generated by INF, Double Difference, and the Unified Bayesian method, as well as other localization methodologies which may rely on relative event travel times. In addition to demonstrating the circumstances under which even typical monitoring geometries may exceed INF limitations, I further establish that my methodologies, when used in tandem, reduce lag measurement error in such a way as to better satisfy the particular error constraints of the INF method, extending the feasibility of the Interferometric methodology to a wider range of acquisition geometries and data qualities, as well as reducing overall localization error.

### 1.4 Thesis Structure

In this work, I first demonstrate the relationship between source-receiver geometry and the precision and accuracy required by the Interferometric Neighboring

Fracture method, followed by an exploration of the effect of receiver array characteristics and placement on the feasibility of an INF relocation. I then demonstrate, with real data, the consequences of travel time measurements which do not satisfy the stringent requirements necessitated by the INF method, along with a discussion of the limitations of and errors caused by relative travel time measurement via conventional inter-event cross-correlation. Next, I present a Modified Adaptive Steering (MAS) technique, based on a methodology proposed in Gangi and Fairborn (1968), along with a related preprocessing technique, Progressive Template Extraction (PTE). I then use synthetic seismic data to compare the performance of my proposed methodologies with that of conventional inter-event cross-correlation. Finally, the thesis is concluded with a real data comparison, which suggests superior performance of PTE enhanced MAS with respect to conventional inter-event cross-correlation, demonstrating improvement in relative lag measurements which more closely meet the unique precision and accuracy conditions required by the INF method.



## 2. INTERFEROMETRIC NEIGHBORING FRACTURE LOCALIZATION

### 2.1 Method Description

The following description is adapted from the original method presentation in Poliannikov et al. (2011).

The interferometric neighboring fracture method (INF) is a derivative of more general interferometric redatuming, which theoretically enables full Green's function recovery under idealized receiver coverage (e.g., Wapenaar, 2004). This technique requires only estimates of relative travel times between event pairs. Consider the following problem setup: a flat, vertical surface of geophones and two sources in three dimensional space,  $x_1$  and  $x_2$ , illustrated in Figure 2.1. In the case that the location of  $x_1$  is known and that of  $x_2$  is unknown, we may estimate the coordinates of  $x_2$  by recovering the traveltimes difference between the two sources along the ray which originates at source  $x_2$  and passes through source  $x_1$  before it is recorded at a receiver. We define a function,

$$dt(p) = t_{x_2} - t_{x_1} \quad (2.1)$$

to represent the measured relative difference in travel times  $t_{x_2}$  and  $t_{x_1}$  from events  $x_2$  and  $x_1$  to point  $p$  on the receiver array surface. Provided that such a ray, from point  $x_2$  through point  $x_1$  to a point on the receiver array, exists, the function  $dt(p)$  will contain a maximum lag time  $dt(p_{stat})$  equivalent to the traveltimes difference we seek to derive, where  $p_{stat}$  is the position on the 2D monitoring surface corresponding to this maximum lag  $\tau_{max}$ . The presence of a stationary lag time implies the existence of the common raypath, shared by events  $x_1$  and  $x_2$ , from point  $x_1$  to the receiver array

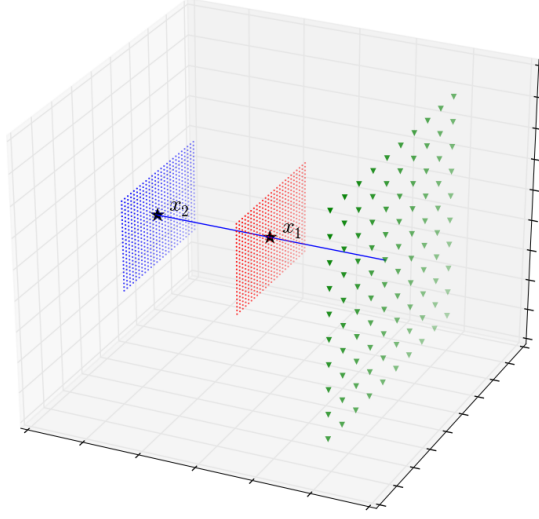


Figure 2.1: Planar monitoring surface (green triangles), two hypothetical fractures with many events (red and blue points), and microseismic events  $x_1$  and  $x_2$  (black stars), with a stationary raypath shown by the black connecting line.

surface, and combining the recovered lag time  $\tau_{max}$  with conventional ray tracing allows for inversion of the position of  $x_2$ .

Consider instead the case shown by Figure 2.2, in which the monitoring surface is replaced with a single vertical monitoring array, and in addition to the two sources  $x_1$  and  $x_2$ , two rectangular fracture planes each represent potentially hundreds of events. Assume furthermore that the plane closest to the array, containing point  $x_1$ , is a reference fracture where event locations have been determined. We now redefine the function  $dt$  as a function of  $l$ ,

$$dt(l) = t_{x_2} - t_{x_1} \quad (2.2)$$

representing the difference in travel times  $t_{x_2}$  and  $t_{x_1}$  to a parameterized point  $l$  on the linear array. In the case that events are monitored with a linear array, rotating

the positions  $x_1$  and  $x_2$  around the array does not change the measured absolute or relative travel times, such that a full spatial stationary analysis is no longer possible. However, a measured maximum in the function  $dt(l)$  implies unique coordinates in two dimensions, defining the angle of the ray from the stationary receiver position  $l$ , through the constant travel-time locus formed by rotating point  $x_1$ , to the constant travel-time locus formed by rotating point  $x_2$ , as shown in Figure 2.3.

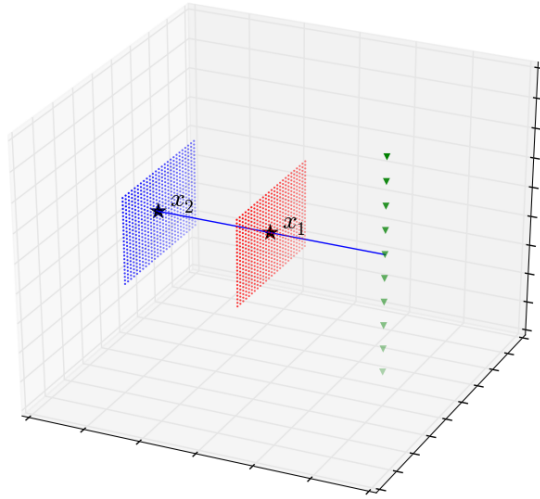


Figure 2.2: Linear receiver array (green triangles), reference and locatable fractures (red and blue points, respectively), and with events  $x_1$  and  $x_2$  (black stars).

In the example of a vertical array, tracing a ray from the receiver through point  $x_1$  using the measured stationary lag allows for the recovery of depth and offset of point  $x_2$  relative to point  $x_1$ , with a resulting azimuthal ambiguity, which may be recovered by classically locating point  $x_2$ . In this manner, hypocenter  $x_2$  is partially redatumed to hypocenter  $x_1$ , which functions as a virtualized receiver.

There are at least three potential benefits offered by the INF method over classical localization. Firstly, localization uncertainty increases with event-receiver offset,

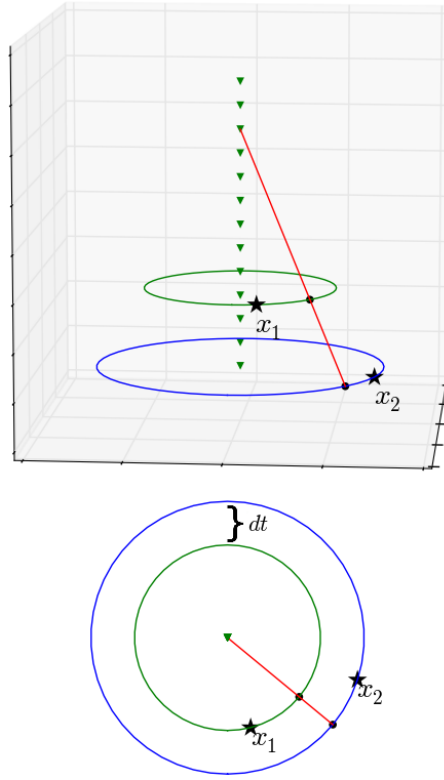


Figure 2.3: For a linear array, a stationary pair of events may be rotated about the array axis without loss of stationarity, since travel times measured at the array remain constant. For a 1D velocity model, these constant-time loci are circles.

such that it is possible to reduce distant hypocenter uncertainty by using information from more precisely localized nearby fracture events. Secondly, given the typical geometry of a hydraulically induced events, numerous reference-locatable event pairs will follow similar raypaths to the receiver array, allowing for velocity model error reduction along the common portion of the raypath, a frequently cited attribute of interferometric redatuming (e.g., Schuster and Zhou, 2006; Schuster, 2009; Waldhauser and Ellsworth, 2000; Poliannikov and Willis, 2011). Thirdly, and most significantly, the large number of events available in the reference fracture can produce multiple estimates of localized hypocenter coordinates, which may be stacked to reduce

localization error compared to classical localization methods.

Poliannikov et al. (2011) present a full uncertainty analysis. To summarize the major points and form part of the foundation for section 3, the uncertainty in coordinates recovered with classical methods is inversely proportional to the square root of the number of independent receiver realizations, which may be stacked

$$\sigma_c \propto 1/\sqrt{N_{rec}}. \quad (2.3)$$

In contrast, INF method uncertainty is inversely proportional to the square root of the number of reference events over which a localized hypocenter may be stacked

$$\sigma_{INF} \propto 1/\sqrt{N_s}. \quad (2.4)$$

Because the number of reference events  $N_s$  is expected to be much greater than the number of receivers  $N_{rec}$ , it follows that location estimates recovered by the interferometric neighboring fracture method should be more precise than those realized classically. A universal threshold of reference events required to surpass classical uncertainty is difficult to establish, since classical and INF uncertainties are complicated functions of many variables, e.g. measurement error, event-receiver offset, and stationary receiver uncertainty, which vary among data sets. The important point, however, is that an idealized application of the INF method is predicated upon the maximization of the number of reference events per locatable event. Conversely, if monitoring and/or data conditions are such that few redundant estimates of non-reference hypocenters are available, an INF based relocation of microseismic events is simply not feasible, as it cannot be expected to outperform classical localization.

## 2.2 Ideal Geometry and the INF Method

Poliannikov et al. (2011) present the INF methodology under idealized circumstances. In practice, a number of factors involving receiver placement and arrival time measurement can affect not only the accuracy but the feasibility of an INF analysis of microseismic data. Compared to other interferometric methodologies, INF places unique and demanding constraints onto relative arrival time precision and accuracy. The stationary analysis central to neighboring fracture interferometry relies upon relative arrival times as recorded along a receiver array to discretely sample a continuous relative traveltime function  $dt(l)$  in the vicinity of a stationary point. Expanding upon the definition of  $dt(l)$  given by equation 2.2 in section 2.1, and assuming a homogeneous medium, this function is given by

$$dt(l) = \frac{||p(l) - x_2|| - ||p(l) - x_1||}{v} \quad (2.5)$$

where  $p(l)$  is the parameterized spatial position  $p$  along a linear monitoring array,  $x_2$  is the distant event position to be relocated, and  $x_1$  is the reference event position. For illustrative purposes, we assume a simple vertical linear array 4000 m in length, centered on the point  $p(0) = (0, 0, 0)$ , with event positions at  $x_2 = (600, 0, 0)$  and  $x_1 = (300, 0, 0)$ , and a P-velocity of 4000 m/s. With these substitutions, the equation is plotted in Figure 2.4. Note that the function is smooth, asymptotically approaches a constant as  $l$  approaches  $\pm\infty$  and, as expected for this particular geometry, the stationary point occurs at  $l = 0$ . More importantly, observe that for a pair of receivers on the array at positions  $l = 0$  and  $l = 100$ , there is a difference in  $dt(l)$  of just 2 ms. The derivative of  $dt(l)$  is plotted in Figure 2.5.

The derivative of  $dt(l)$  ranges from .065 ms/m to -.065 ms/m at 300 m from the stationary point where it predictably goes to zero. In order for relocation of

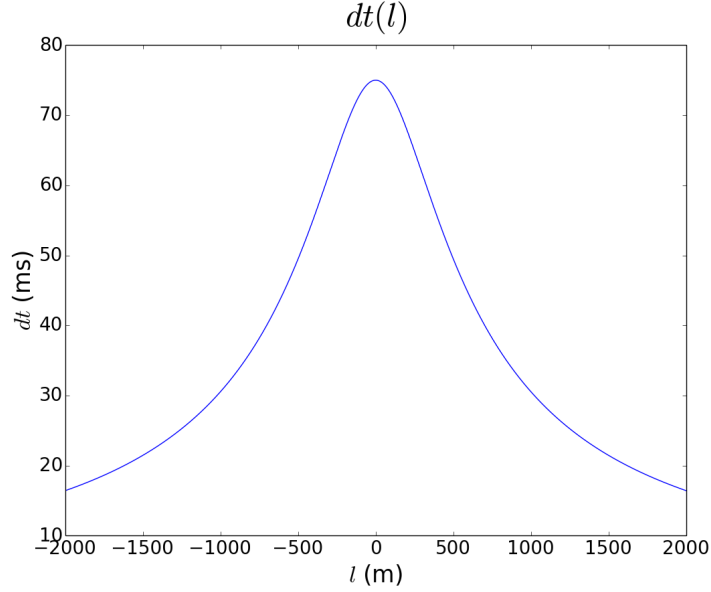


Figure 2.4: Relative travel time function  $dt(l)$  as a function of receiver position  $l$

an event to take place, the stationary point in  $dt(l)$  must be identified and, ideally, enough data points should be available for accurate interpolation of the true value of  $l$  for which  $\frac{\partial}{\partial l}dt(l) = 0$ . Because both  $dt(l)$  and its derivative depend upon event-receiver geometry, there is a limited range of options for receiver placement relative to microseism distribution in this particular case, for example, if a 7 receiver array with a receiver spacing of 15 m were centered at  $(0,0,0)$ , there would be no more than 1 ms of travel time difference between any two receivers, and a much smaller difference between adjacent receivers. This implies that, for certain geometries, the function  $dt(l)$  cannot be sampled precisely enough to resolve stationary event pairs; Waldhauser and Ellsworth (2000) report that under the best field circumstances, a relative arrival time precision of 1 ms is optimistic.

It follows that care must be taken in evaluating the potential feasibility of an INF analysis with respect to receiver array placement and array length, since even rela-

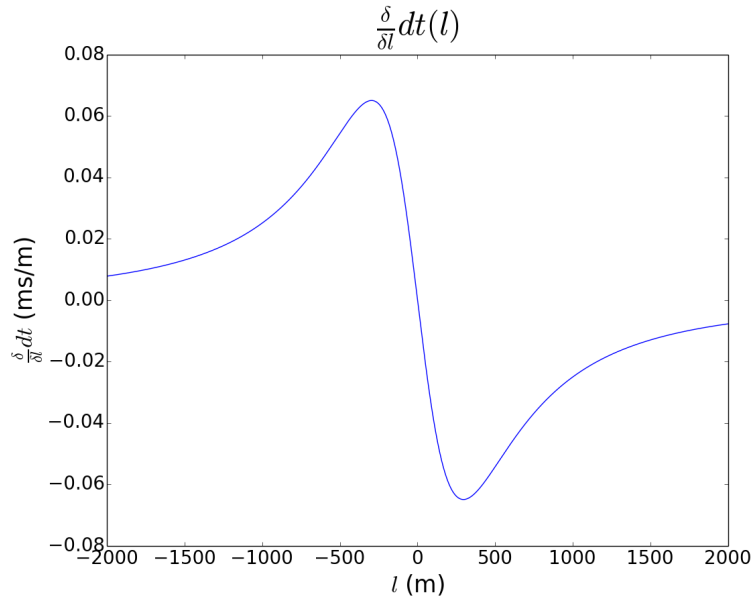


Figure 2.5: Derivative of the relative travel time function  $dt(l)$  as a function of receiver position  $l$

tively typical monitoring arrangements, such as the example presented above, may easily surpass the limits of relative travel time measurement precision. If an interferometric neighboring fracture analysis is attempted on a highly non-ideal data set, few stationary pairs may be available for relocation, or worse, noisy samplings of  $dt(l)$  may falsely identify event pairs as stationary and result in nonsensical relocations.

This behavior is the source of the unusual demand which the INF stationary analysis places on arrival time measurement. Whereas other methodologies, e.g. Double Difference, may minimize localization uncertainty by simultaneously operating on all events, so long as errors are normally distributed about true values, this condition is not sufficient for INF relocation. Because INF evaluates data on a pair by pair basis, with the number of samples equal to the number of receivers, large, random measurement error will obscure stationary  $dt(l)$  values, regardless of error distribution. Although some degree of error is unavoidable, INF requires that the error is not



necessarily small, but consistent enough to resolve the curvature of the  $dt(l)$  curve, to correctly identify stationary pairs.

### 3. MONITORING GEOMETRY AND THE FEASIBILITY OF THE INF METHOD

Prior to using the interferometric neighboring fracture method as a relocation tool, the effects of monitoring geometry must be considered in order to avoid analysis of a data set which surpasses the strict limitations of interferometric neighboring fracture analysis, as described in section 2.2. Specifically, it should be ensured that  $dt(l)$  may be resolved for the majority of event pairs in the presence of measurement error, and that a large number of stationary reference events are available per locatable event to effectively exploit the INF feature of common-event stacking. In this portion of the text, therefore, the effects of event-receiver geometry on the resolution of  $dt(l)$  and stationary pair count are explored.

#### 3.1 Receiver Array Characteristics and $dt(l)$ Resolution

As demonstrated in section 3, the feasibility of the INF method is limited by the ability of monitoring receivers to resolve the  $dt(l)$  function for a given pair of events. To summarize, if the magnitude difference in relative arrival times between receivers along an array is comparable to the magnitude of measurement error, stationary pairs may be missed or misidentified, and a correctly identified stationary pair may not be suitable for localization because of the confounding effect of measurement error on stationary point interpolation. The robustness of stationary analysis results will therefore be maximized when receivers are located in such a manner as to maximize the total change in  $dt(l)$  along the receiver array. To simplify the problem, we consider an array with just two receivers, and seek to generally maximize the

following function for all pairs of microseisms

$$\Delta dt = |dt_1 - dt_2| \quad (3.1)$$

where  $dt_1$  and  $dt_2$  are the relative lags for a given pair of events as measured at receivers 1 and 2, respectively. In this manner,  $\Delta dt$  represents the difference of two discrete samples of the  $dt(l)$  function as given in equation 2.5, at different locations on a receiver array.

Because  $\Delta dt$  is a function of 12 variables (spatial coordinates for each receiver and each event), it is difficult to provide a concise, analytic demonstration of the relationship between receiver placement and  $dt(l)$  resolution. Instead, conclusions are derived from the following numerical trial.

### *3.1.1 Model Description*

To evaluate the relationship between receiver placement and travel time measurement within the context of a hydraulically stimulated well, 200 total events are randomly positioned within two planar fracture stages, separated by 400 m. Both fractures are 100 meters wide, 75 meters tall, and 10 meters deep, each containing 100 events. A simple, homogeneous, isotropic velocity model is assumed throughout the medium, with  $v = 4000$  m/s. The model is graphically displayed in figure 3.1, with the reference fracture in red, and the locatable fracture in blue.

### *3.1.2 Test Model and Procedure*

A single trial consists of calculating the  $\Delta dt$  value averaged over all reference-locatable pairs of events at equally spaced grid nodes on the xy plane,  $z = 0$ . Between trials, receiver spacing and orientation are varied, for a total of 9 trials, with properties summarized in table 3.1.2

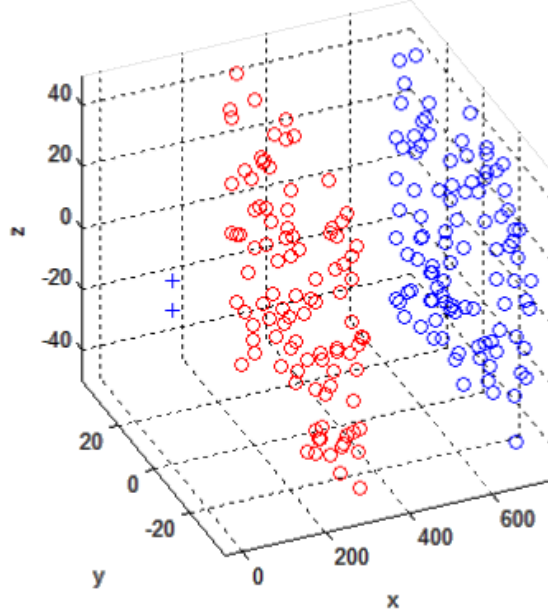


Figure 3.1: Model used in the resolution trial. Two fractures, 100 events each, are separated by 400 meters. Example receiver locations are given by the blue crosses.

In this manner, I explore the effects of array orientation, spacing between receivers, and receiver position on the difference in relative travel times as measured between two receivers. For each node in the model grid space, a value  $\overline{\Delta dt}$  is calculated as

$$\overline{\Delta dt} = \frac{1}{N} \sum_{i=0}^N |dt_1 - dt_2| \quad (3.2)$$

where the summation is over all  $N$  (reference, locatable) event pairs.  $\overline{\Delta dt}$  therefore is the average difference between time lags as measured at receivers 1 and 2 over all pairs of events. In order to avoid artifacts resulting from unrealistic equidistant event spacing, event positions are generated randomly once and kept constant for all trials.

Trial Number	Receiver Spacing	Array Orientation
1	10 m	$\hat{e}_x$
2	10 m	$\hat{e}_y$
3	10 m	$\hat{e}_z$
4	20 m	$\hat{e}_x$
5	20 m	$\hat{e}_y$
6	20 m	$\hat{e}_z$
7	30 m	$\hat{e}_x$
8	30 m	$\hat{e}_y$
9	30 m	$\hat{e}_z$

Table 3.1.1: Orientation and receiver spacing combinations used in the resolution trial.

### 3.1.3 Results

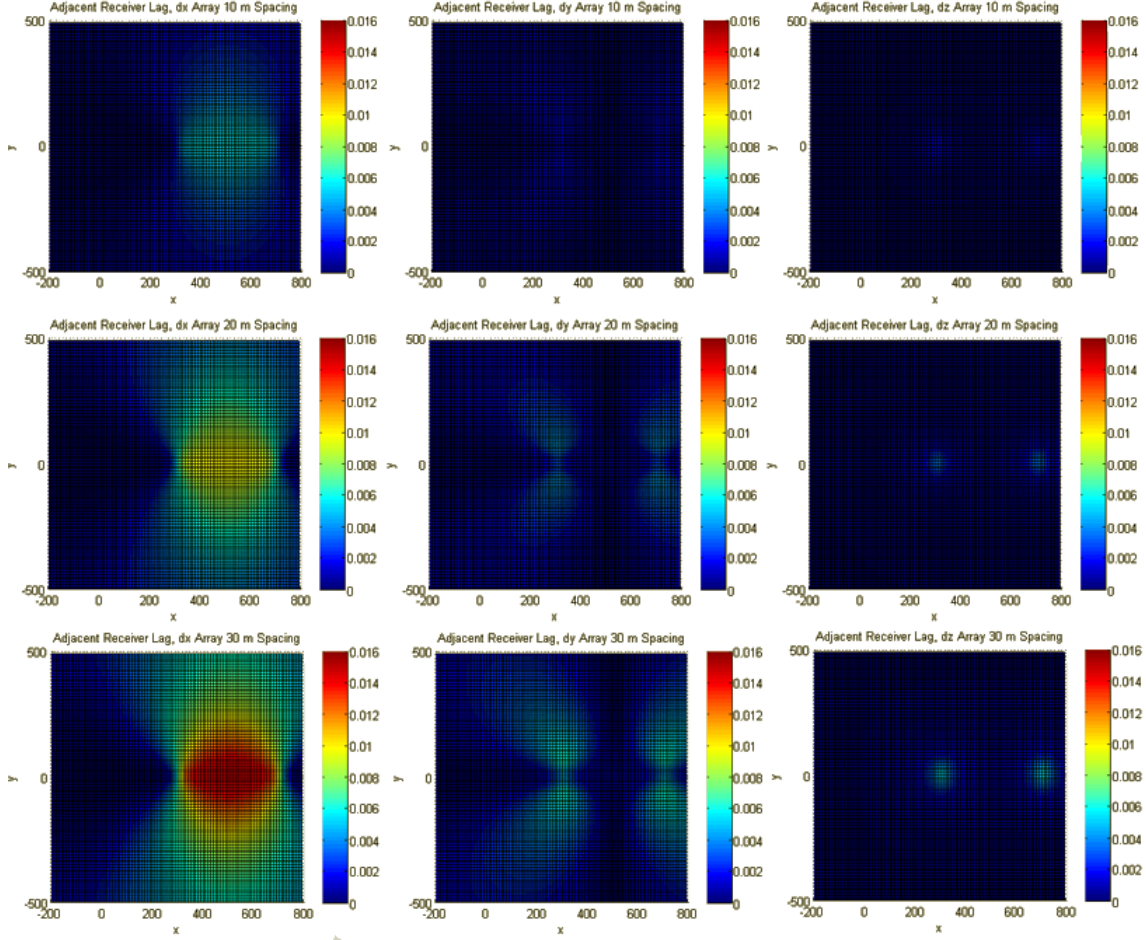


Figure 3.2: Results of trials with varying receiver spacings and orientations. Array orientations are  $ex$ ,  $ey$ , and  $ez$  in the left, center, and right columns, respectively. Receiver spacings are 10 m, 20 m, and 30 m, from top to bottom row. Colorbar values are in seconds. Color scales are identical for all plots.

	$\hat{e}_x$	$\hat{e}_y$	$\hat{e}_z$
10 m	0.0049 s	0.0019 s	0.0018 s
20 m	0.0099 s	0.0038 s	0.0034 s
30 m	0.0138 s	0.0057 s	0.0050 s

Table 3.1.2: Maximum  $\overline{\Delta dt}$  values corresponding to the trials in figure 3.2.

Two observations are immediately apparent. First, plot symmetries are only dependent upon array orientation, not receiver spacing. Second, calculated inter-receiver time lag differences  $\overline{\Delta dt}$  increase with receiver spacing for a given orientation.  $\overline{\Delta dt}$  is largest when receiver spacing is large, and the receiver array is oriented in the  $\hat{e}_x$  direction, midway between the reference and locatable fractures.

If a conservative constraint of 0.01 s is placed on  $\overline{\Delta dt}$ , simulating a minimum precision tolerance which would be acceptable in the presence of measurement error, receivers must be spaced at least 20 m apart, and in the case that the spacing is less than 30 m, only an  $\hat{e}_x$  oriented array will satisfy the minimum  $\overline{\Delta dt}$  requirement, and even then there is a limited range of locations within which  $\overline{\Delta dt}$  is greater than 0.01 s, bounded by the circular yellow region in the bottom left hand corner of figure 3.1.3.

Observe also that  $\overline{\Delta dt}$  goes to 0 as the receiver array is moved away from the fracture planes. This has an important consequence, in that the location of a monitoring array within the vertical section of a well pad, even for typical monitoring geometries, may produce data which cannot be reliably relocated with the INF method. As an example, even the idealized geometry presented in Poliannikov et al. (2011), with receivers, reference, and locatable fractures centered at  $x = 0, 100$ , and 200 meters, respectively, produces a  $\overline{\Delta dt}$  value of 0.0022 s for a velocity of 4000 m/s for an adjacent receiver separation of 30 m, far less than a practical 0.01 s limit. If

two vertical receivers are placed with a centroid at  $(0, 0, dz)$ , the minimum spacing required to satisfy the 0.01 s constraint is approximately 260 m for the Poliannikov model and 230 m for the model used in the trial above.

A surface monitoring array presents a similar problem, being geometrically similar to the  $\hat{e}_y$  trials above, with more extreme y offsets than the 500 m test limits. Even at approximately 500 m of distance between fracture centroids and the monitoring array,  $\overline{\Delta dt}$  goes to 0 s for a 30 m receiver spacing, and distances from the surface to stimulated wellbores generally surpass 1 km, where a receiver spacing in excess of 100 m is required to meet the 0.01 s resolution constraint.

Both cases allude to the physical relationship between travel-times and receiver-fracture geometry: as microseismic hypocenters move away from a monitoring array towards infinity, the angles from receivers to events approach 0, such that the difference in raypath lengths measured by  $\Delta dt$  also rapidly approaches 0.

The average value  $\overline{\Delta dt}$  may be somewhat misleading, as it does not guarantee that all event pairs will meet the minimum resolution value for a given receiver placement. Unfortunately, the distribution of  $\Delta dt$  also varies with receiver placement and orientation; nonetheless, the distribution follows a clear trend with respect to array position. To explore the relationship between  $\Delta dt$  and the actual proportion of event pairs which can be expected to be resolved, histograms were generated for three select locations from the 30 m  $\hat{e}_z$  trial, shown in figure 3.4 with a location basemap in figure 3.3.



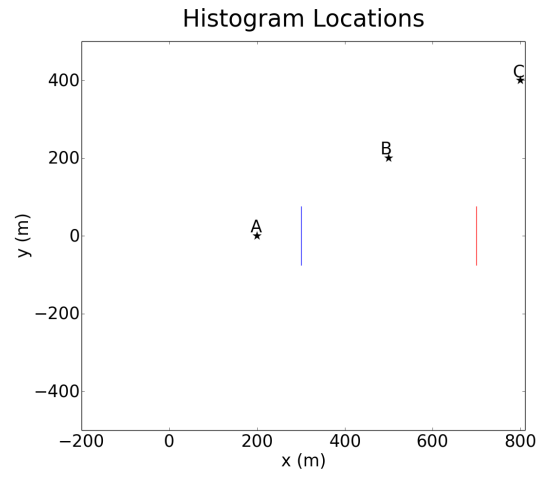


Figure 3.3: Basemap, showing locations from which histograms were taken, with reference and locatable fractures in blue and red, respectively.

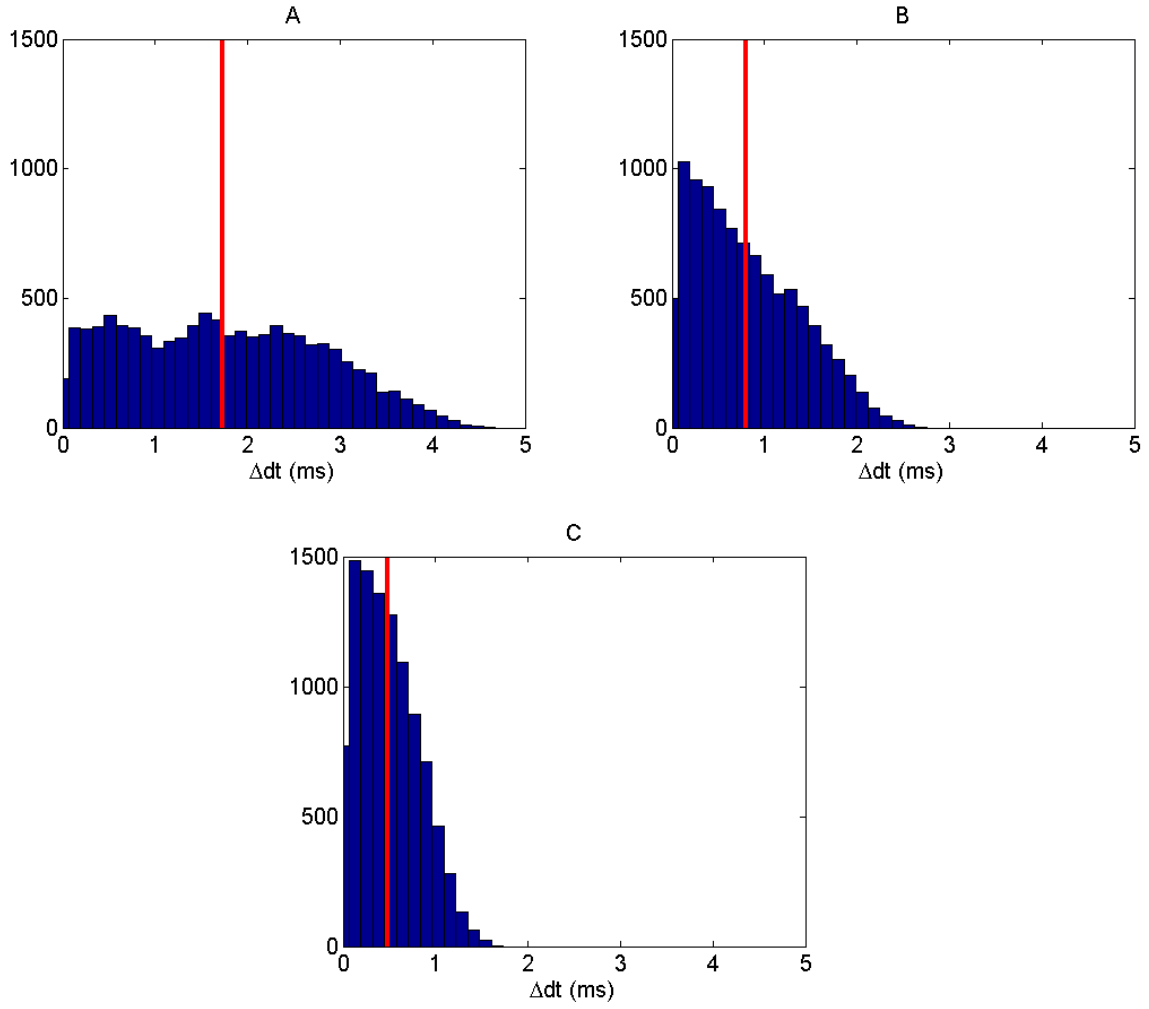


Figure 3.4: Histograms at locations A, B, and C from the basemap in figure 3.3.  $\overline{\Delta dt}$  values are given by the red lines. Note how in each case, the actual distribution is somewhat skewed.

As the array is moved away from the borehole axis at  $y = 0$ , the distribution tends to skew more and more toward lower values. The  $\overline{\Delta dt}$  values, represented by the red circles in figure 3.4, are 1.7 ms, 0.80 ms, and 0.49 ms, while the proportions

of event pairs with  $\Delta dt$  greater than  $\overline{\Delta dt}$  are .48, .44, and .45, for locations A, B, and C, respectively. This suggests that if  $\overline{\Delta dt}$  at a particular receiver location is equal to the minimum resolvable time, only about 45%-48% of event pairs will produce resolvable  $dt(l)$  values.

### 3.2 Array Placement and Resolvable Stationary Pair Count

While the analysis above may provide a useful exploration of the relationship between  $dt(l)$  resolution and receiver placement, it ignores the influence of event-receiver geometry on the number of stationary pairs potentially available for relocation. Given that a synthetically demonstrated strength of the INF methodology lies in the stacking of multiple stationary pairs over a common relocated event, optimal event-receiver geometry can only be realized through the additional consideration of stationary pair count maximization.

The previous trial is therefore expanded upon with a second numerical exploration, this time simultaneously considering resolvability as well as stationary pair count with respect to receiver position.

#### 3.2.1 Test Model and Procedure

This numerical test uses the same velocity/hypocenter model as that in section 3.1 above, although rather than evaluating travel times at two receivers, a receiver array is represented by a continuous line. The test simply counts the number of event pairs for a given array spacing, position, and orientation, which satisfy the following two constraints:

1. A stationary point must exist along the array, and not at the first or last receivers
2. The stationary point must be resolvable

To test for condition 1 above, for a simple 1D model, since both events in a

stationary pair may be rotated around the receiver array axis without a loss of stationarity, the fracture planes are first flattened to a single azimuth. The stationary test then simply tests for geometrical co-linearity of each reference-locatable pair with some point on the array.

Condition 2 depends on receiver spacing, array length, measurement precision, and the minimum number of resolvable receivers desired. Because the stationary point will fall between receivers in practice, its value must be interpolated, and in the presence of error, multiple  $dt(l)$  measurements are necessary for accurate interpolation. The condition can be reformulated: what is the minimum distance along the array from the stationary point at which the next measurement must be taken so that the difference  $\Delta dt$  in two measurements is greater than the desired precision? If we assume a constant receiver spacing, and that  $\Delta dt$  increases as we move away from the stationary point, then this distance  $d$  relative to array length  $l$  may be used to estimate the number of receivers which will be able to resolve  $\Delta dt$  values:

$$N_r = N_t(1 - 2\frac{d}{l}) \quad (3.3)$$

where  $N_r$  is the number of receivers which can measure  $\Delta dt$  and  $N_t$  is the total number of receivers. The factor of 2 arises from the fact that receivers on either side of the stationary point within distance  $d$  will be unable to resolve  $\Delta dt$ . This equation is valid under the assumption that there is at least  $d$  distance from the stationary point to either end of the array.

Assuming 10 receivers, I choose a cutoff of  $d/l = .25$ , in which case a minimum of approximately 5 receivers should be able to measure  $\Delta dt$  for interpolation, assuming an optimistic minimum resolvable  $\Delta dt$  of 1 ms. In other words, if  $d/l > .25$ , the distance from the stationary point to the first array position at which  $\Delta dt$  exceeds

1 ms is greater than  $1/4$  of the total array length, and condition 2 is not satisfied.

As the previous trial has demonstrated that increased inter-receiver spacing increases resolvability, I now seek to demonstrate the array position which maximizes the number of resolvable stationary points. To simplify the analysis, I assume the equivalent of a 10 receiver array with 40 m inter-receiver spacing, for a total array length of 400 m, and a minimum 1 ms resolution cutoff. For a given fracture geometry, point pairs which satisfy conditions 1 and 2 can only increase with increased receiver spacing and array length, so the emphasis here is on the effects of location and orientation of the receiver array relative to fracture geometry. In the same style as the previous trial, the centroid of the array is shifted through all points of a grid, and at each point the number of pairs which satisfy conditions 1 and 2 above are counted. This procedure is repeated for three possible array orientations  $\hat{e}_x$ ,  $\hat{e}_y$ , and  $\hat{e}_z$ . For the  $\hat{e}_z$  trial, the array center is placed 200 meters above the center of the fractures, to represent monitoring from within the vertical portion of a horizontal well.  $\hat{e}_x$  and  $\hat{e}_y$  array centers are kept level with fracture centers, to mimic horizontal monitoring from within adjacent horizontal wellbores.

### 3.2.2 Results

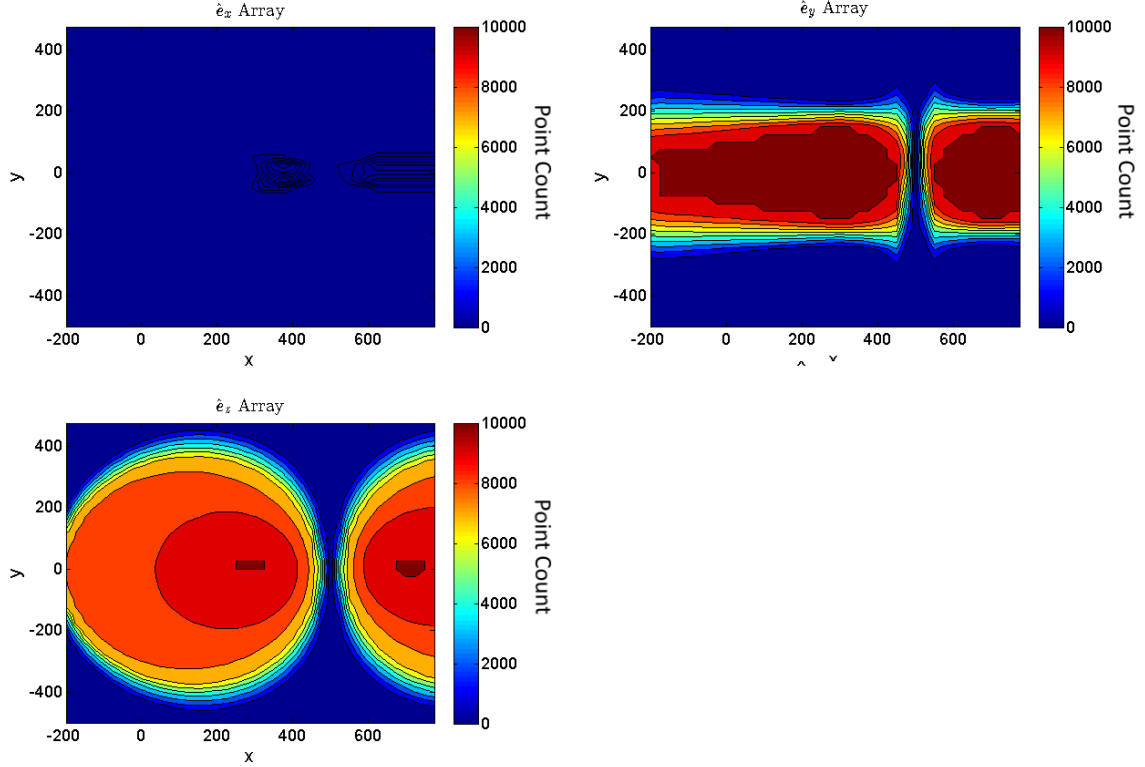


Figure 3.5: Stationary, resolvable pairs counted for arrays oriented in the  $\hat{e}_x$  (top left),  $\hat{e}_y$  (top right), and  $\hat{e}_z$  (bottom) orientations. The  $\hat{e}_x$  orientation maximum count is 391, while both  $\hat{e}_y$  and  $\hat{e}_z$  orientations reach 10000 (100%) stationary pairs.

At a glance, the results in Figure 3.5 seem to contradict those of section 3.1.3, as the  $\hat{e}_x$  orientation was previously shown to offer the best resolution for a given receiver spacing, whereas in this case the  $\hat{e}_x$  oriented array produces just 391 resolvable stationary pairs in the optimal location, or just 3.9% of all possible event pairs,

compared to 10,000 (100%) for the  $\hat{e}_y$  and  $\hat{e}_z$  oriented arrays. This discrepancy highlights the importance of the geometrical relationship between fractures, monitoring receivers, and the suitability of an INF analysis, since any reasonably long array oriented perpendicular to fracture planes ( $\hat{e}_x$  in this case) will not produce enough stationary pairs for relocation. In this case the analysis is not limited by resolution, but by the simple lack of event pairs which are stationary, i.e. are co-linear, after rotation, with any point on the array.

The  $\hat{e}_y$  and  $\hat{e}_z$  orientations on the other hand produce more favorable results. To simplify, we assume a cutoff value of 5000 total resolvable stationary pairs, below which an array location does not resolve enough stationary pairs for an INF analysis. This translates to 50 potential relocations per non-reference event, although depending on array placement, locatable events will be stationary with more or less than 50 reference events. The  $\hat{e}_z$  trial satisfies this condition at 59% of tested locations, compared to 39% for  $\hat{e}_y$ . It is worth noting, however, that the average counts of all locations which satisfy the 5000 pair minimum are 9134 pairs for  $\hat{e}_y$  and 8410 for  $\hat{e}_z$ , such that neither orientation is obviously better suited for INF analysis. Plots showing nodes which satisfy the minimum 5000 stationary pair condition are shown below. The asymmetrical nature of the plots is a result of the random hypocenter distribution detailed previously.

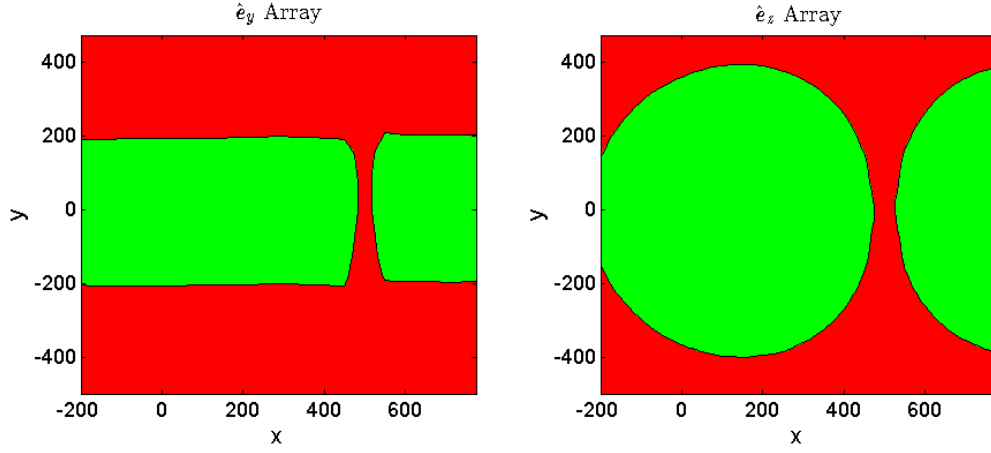


Figure 3.6: Spatial distribution of nodes with at least 5000 stationary, resolvable event pairs (green).

On a final note, the primary focus of this section is the general shape of the plots above, as the total area of suitable receiver array locations can be increased by increasing array length or reducing the minimum desired  $\Delta dt$  value.

### 3.3 Summary

Array length and positioning in the field are constrained by economic considerations, generally limiting monitoring locations to either downhole horizontal arrays lying parallel to the stimulation well, and perpendicular to idealized fracture geometry ( $\hat{e}_x$  cases above), or within the vertical section of a horizontal pad ( $\hat{e}_z$  cases above). Because stimulation wells are typically drilled in the direction of minimum horizontal stress, and arbitrarily oriented monitoring wells are not drilled, event-receiver geometries resembling the tested  $\hat{e}_y$  cases do not occur in practice. As the resolvable-stationary pair trial of the previous section has shown, microseismic recordings from a downhole array parallel to the stimulated borehole cannot be expected to produce



enough stationary pairs for an INF analysis. The only practical monitoring arrangement which may therefore produce data suitable for an Interferometric Neighboring Fracture analysis is a vertical downhole array.

The results in section 3.1 demonstrate that as the offset between an array and the monitored fracture planes is increased, the difference in measured relative arrival times between adjacent receivers,  $\Delta dt$ , decreases. Combining this observation with the measurement requirements of  $dt(1)$  detailed in section 2.2, it is clear that care must be taken to ensure that the expected  $\overline{\Delta dt}$  value for a given monitoring geometry is greater than the expected relative arrival time measurement precision prior to performing an INF analysis, in order to avoid the aforementioned erroneous relocations that may otherwise result.

Given the range of choices for array length and receiver count, as well as the variability among hypocenter distributions in practice, the suitability of a data set for INF relocation should be evaluated on a case-by-case basis. If preliminary hypocenters are available, the analyses above are simple enough to be repeated with true receiver positions and available classical hypocenter locations prior to investing time in a full INF analysis. Additionally, such a preliminary step can avoid erroneous locations returned by the INF analysis of an infeasible data set.

Finally, results in 3.2.2 may be used to establish a rule of thumb regarding maximum permissible offset between a vertical downhole array and monitored fractures. Ignoring the effects of noise, with a generous array length of 400 m and a conservative minimum resolvable  $\Delta dt$  of 1 ms, figure 3.2 shows that the stationary-resolvable pair count rapidly drops below 5000 toward 0 at array offsets exceeding 500 m to the first fracture. For shorter arrays or larger minimum  $\Delta dt$ , this maximum allowable offset decreases. Therefore, monitoring geometries approaching 500 m of array-fracture offset must be carefully examined to determine the feasibility of associated hypocenter

relocation with the INF method.

#### 4. IMPROVED RELATIVE TRAVELTIME MEASUREMENT WITH MODIFIED ADAPTIVE STEERING

The INF method does not place an explicit constraint on the method used to recover relative travel times between events, although the present standard for interferometric methods is cross-correlation (Molyneux and Schmitt, 1999; Waldhauser and Ellsworth, 2000; Schuster, 2009). However, under the circumstances typical of hydraulic monitoring, a number of assumptions regarding event geometry and source characteristics break down, resulting in poor correlation coefficients and relative time measurement error when times are measured with inter-event cross-correlation. The following sections briefly demonstrate these non-ideal behaviors in the context of a real set of data, providing empirical justification for the use of a newly modified adaptive steering based relative travel time estimation methodology presented in section 4.3, proposed to mitigate demonstrated travel time errors associated with cross-correlation or subtraction of picked arrival times.

##### 4.1 Idealized Cross-Correlation Constraints and Microseismic Monitoring

Cross-correlation as a relative traveltime estimation tool has been adapted for microseismic analysis from general Earthquake seismology (Schuster, 2009), and it is accurate to the degree that the two waveforms being cross-correlated are similar. For correlation on global scales and geometries typical of earthquake seismology, some combination of the following assumptions is expected to ensure generally similar waveforms (Arrowsmith and Eisner, 2006; Poupinet et al., 1984; Waldhauser and Ellsworth, 2000; Geller and Mueller, 1980; Phillips, 2000):

1. Seismic sources are generated by the same source mechanism
2. Seismic sources are separated by roughly  $1/4$  of the dominant wavelength or

less

### 3. Correlated events follow a relatively common raypath

In fact, the technique of multiplet analysis is predicated upon these conditions. These assumptions are typically not valid during microseismic monitoring, resulting in largely dissimilar microseismic waveforms and poor normalized correlation coefficients during cross-correlation. The breakdown of these conditions manifests as differences in first arrival and coda expression between events. For a detailed explanation of the invalidity of these conditions and their effects on recorded waveforms, the reader is referred to Appendix A.

As demonstrated below, when cross-correlation is applied to dissimilar waveforms, the accuracy of relative arrival time recovery is compromised, and within the context of an INF analysis,  $dt(l)$  functions cannot be reliably measured. The ultimate result is that event pairs must be discarded without relocation, and if the number of remaining valid pairs is too small, INF cannot be expected to outperform classical localization, if any relocations can be performed at all.

## 4.2 A Case Study of Cross-Correlation and the INF Method

### 4.2.1 Dataset Description

The microseismic data examined below (and in later) sections was recorded during the stimulation of a shale reservoir. A 7 receiver downhole monitoring array was placed in a horizontal well adjacent to the stimulated borehole, approximately 500 m from the stimulation well. 6 of the receivers were separated by 15 m, while the easternmost receiver was offset by 30 m from the rest of the array. Because of the proximity of receivers to the stimulated volume, S/N ratios are typically greater than 10. 212 events were identified and located by a third party processing company, and their origin times were used to link provided locations to recorded arrivals for all 212

events. A plot of event locations and the monitoring array is shown in Figure 4.1.

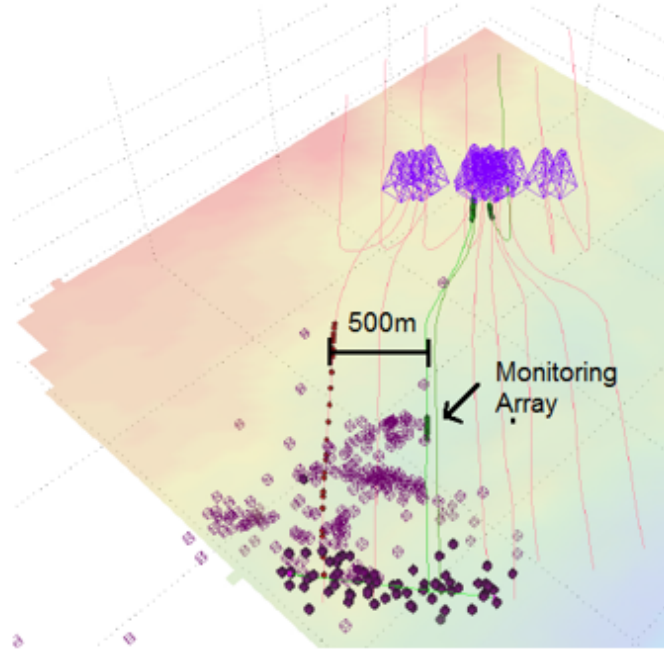


Figure 4.1: Event locations given by light purple cubes.

#### 4.2.2 Events, Methodology, and Results

In order to exemplify the errors introduced by cross-correlation of dissimilar waveforms, as well as the consequences that such errors may have on INF relocation, P-arrivals for two events, numbers 0 and 4, are examined, with their locations and traces given below. To simplify the analysis, 8 ms windows of only the Z-component of receiver recordings are considered, in which direction both events are predominantly polarized. The relative amplitudes are approximately 10:1, event 4 to event 0.

For both events, the exact first arrival time is difficult to estimate visually. It

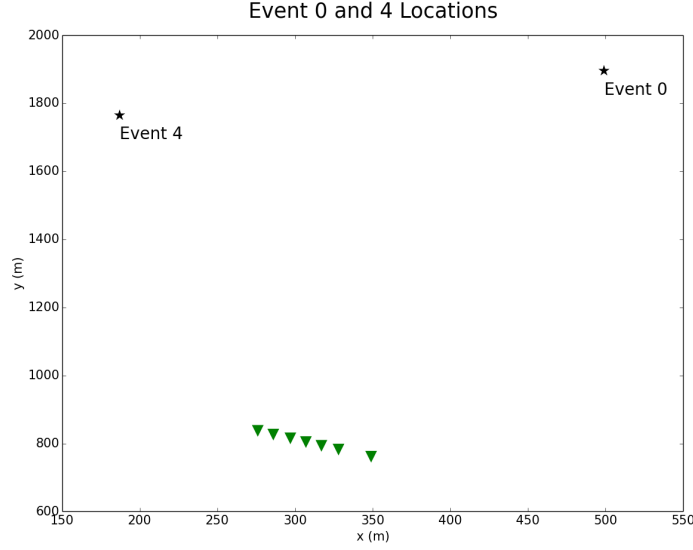


Figure 4.2: Locations of events 0 and 4 (labeled stars) with respect to the receiver array (green triangles).

should be observed additionally that these two events are markedly different in both first arrival and post first arrival expression. Furthermore, note that the post first arrivals for both events appears to surpass the amplitude of first arrivals by factors larger than 10 on some traces. Finally, observing each series of event traces individually, it is clear that waveforms of the same event recorded by different receivers are inconsistent. In typical interferometric fashion, and following the INF methodology described previously, I cross-correlate event waveforms by receiver, and attempt to recover the relative traveltime function  $dt(l)$  over the receiver array, with the results shown in Figure 4.4.

#### 4.2.3 Implications of Cross-Correlation Error

It is apparent that the resulting measurements of  $dt(l)$  as shown in Figure 4.4 cannot be accurate, for at least two reasons. First, as described in section 2.2, the

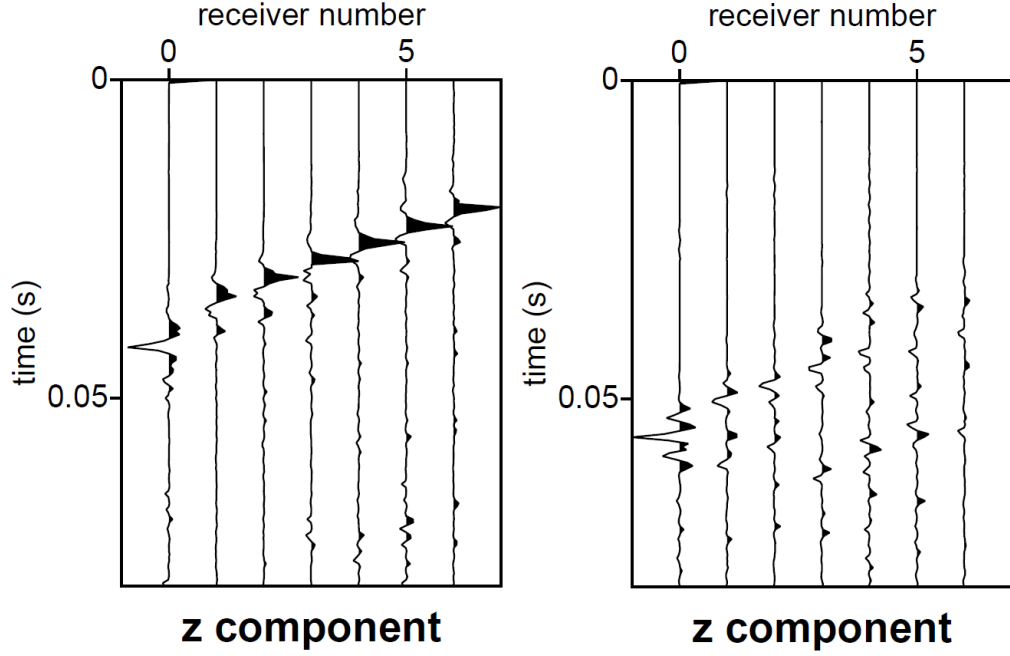


Figure 4.3: Traces for events 0 and 4. Event 4:event 0 scale ratio is approximately 10:1.

true function  $dt(l)$  is smooth, particularly on the scale of 100 m, in contrast to the samples shown above. Second, a visual examination of moveouts for events 0 and 4 indicates that the measured values of  $dt(l)$  should be relatively consistent in magnitude and curvature, whereas the plot above ranges in lags from 15 ms to 34 ms, or 18% to 42.5% of the total signal window width.

If the reported hypocenters for events 0 and 4 are correct, the  $dt(l)$  function in Figure 4.4 should not contain a maximum, since the pairs are not spatially collinear with any point on the receiver array, even after accounting for azimuthal ambiguity. It follows that a naive application of the stationary analysis, without consideration for the accuracy and precision of cross-correlation derived relative arrival times, could erroneously identify event pairs as stationary, as indicated by the apparent maximum in  $dt(l)$  above. Furthermore, for results like this, it would be impossible to accurately

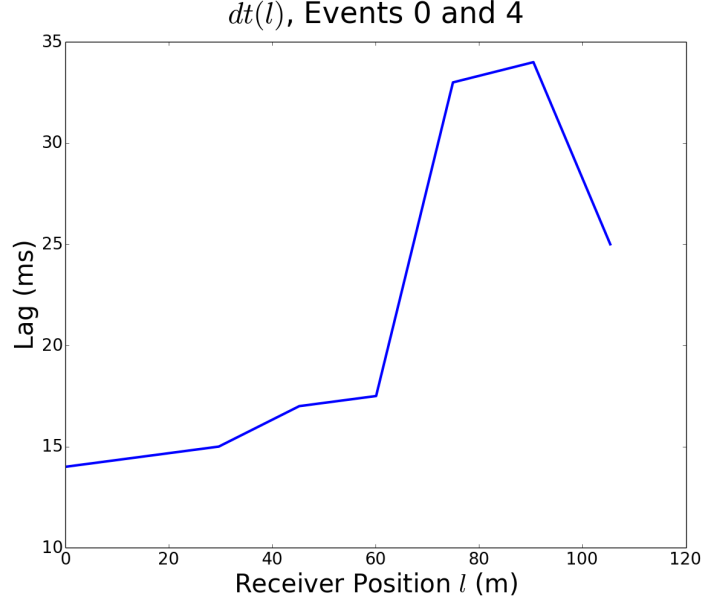


Figure 4.4: Measured relative arrival times for events 0 and 4.

interpolate for the true maximum lag, as the curvature of lag function samples is not only incorrect but also inconsistent.

The major source of error in this case, and for the majority of event pairs that are compared with conventional cross-correlation, is a misalignment of first arrivals, as discussed in detail in Appendix B. Related to the concept of cycle skipping (Schimmel and Assumpcao, 2003; Schaff and Richards, 2004), the problem arises when dissimilar, ringy waveforms produce multiply peaked correlograms on cross-correlation. Occasionally, correlogram maxima do not correspond to true relative arrival lags, resulting in inconsistent relative arrival time measurements for event pairs.

#### 4.2.4 Potential Mitigation of Misalignment Error

A number of options for reducing the potential for misalignment and limiting the associated error are apparent. We could seemingly accomplish both tasks by



minimizing the time width of the signals being cross-correlated; however, consider the two first arrivals in Figure 4.5, taken from events 0 and 4 and windowed to 12 ms.

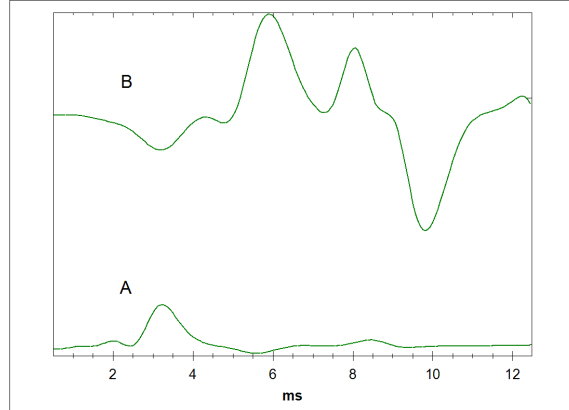


Figure 4.5: Two first arrivals A and B from events 0 and 4, respectively, which could hypothetically be cross-correlated with each other to recover a relative arrival time.

The maximum value of the cross-correlogram (not shown) appears at the lag corresponding to the alignment of the large peak in trace A with the final peak in trace B, suggesting that even for small time window widths which themselves approach the limits of cross-correlation, there is still the potential for misalignment. Furthermore, as window widths are made arbitrarily small, they begin to defeat the purpose of cross-correlation if we assume that they contain true first arrivals, and as we approach the limits of first arrival picking precision, we run the risk of windowing out the true first arrival. Given the nature of the data set, misalignment of cross-correlated traces may be the norm, since mathematically cross-correlation tends to align functions by their largest peaks, which for this data set frequently do not correspond to first arrivals. For both window widths in this example, note that

apparent error is much greater than the required precision suggested by the idealized analysis in section 2.2, indicating that conventional cross-correlation cannot resolve the values of  $dt(l)$  with the precision and accuracy necessary for stationary analyses.

Because the accuracy of cross-correlation is predicated upon signal similarity, there is the seemingly plausible option of limiting stationary analyses to event pairs with some acceptable minimum coefficient value. Indeed, it would be expected that geometrically stationary event pairs would satisfy at least one of the three previously discussed conditions for signal similarity, namely that the event signals would follow a common raypath, and would therefore tend to present more similar waveforms. To examine this possibility, all pairs of 8 ms p-arrival windows of events within the data set were cross-correlated, with magnitude correlation-coefficients summarized in Figure 4.6.

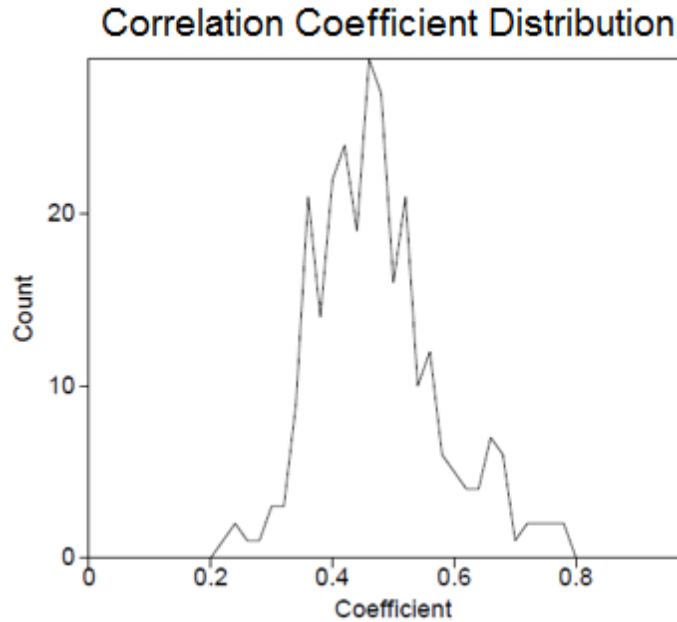


Figure 4.6: Distribution of inter-event correlation coefficient magnitudes.

The vast majority of event-event correlations produce coefficients of less than 0.6, which is generally not taken to be large enough to expect accurate relative arrival time measurement via cross-correlation. It should be noted that applying a Butterworth 5 Hz-200 Hz bandpass filter, following the work of Hurd (2012), did not appreciably improve the spread of the coefficients, and that Hurd reported similarly low correlation coefficients. Even if a minimum coefficient of 0.6 is taken as a minimum acceptable value, only 2042 event pairs, or 12%, are potentially available for relocation, with questionable accuracy.

Unfortunately, the event-receiver geometry of the data set is not suitable for a full trial of the INF method. Using provided classical locations,  $dt(l)$  values were predicted at the given receiver positions, and 90% of event pairs are beyond resolution, as  $dt(l)$  values among receivers do not differ by more than 1 ms over the entire array. The result was that upon cross-correlation, relative arrival times were drowned out by misalignment, and  $dt(l)$  samples for the majority of event pairs could not be used for relocation, even when they did contain a measured maximum, since the data points could not be used for interpolation, as in the example for events 4 and 0 above. Moreover, the inconsistency of the measured  $dt(l)$  functions calls the apparent maxima and their associated lags into question.

The problem of accurate relative arrival time measurement therefore remains. For this data set, conventional event-event cross-correlation appears to fail to return accurate time lags, while the non-ideal characteristics of the measured waveforms create an excellent arena in which an accuracy improvement may be demonstrated, if a suitable alternative measurement methodology can be developed.

#### 4.2.5 *Array Processing*

The process of determining relative arrival times at receivers along an array is closely related to the general problem of seismological array processing, which has been studied since the 1950s (Rost and Thomas, 2009). These techniques are generally designed to improve signal to noise ratios through various forms of stacking, via cancellation of incoherent noise. One historic method in particular, the process of beam forming, alternatively referred to as array steering or migration, is still commonly used today in both global and petroleum seismology, and involves delaying and stacking seismic arrivals, with varying methods of normalization and delay calculation (Ram and Mereu, 1975; Rawlinson and Kennett, 2004; Gibbons and Ringdal, 2006).

While these techniques are designed for improving estimates of absolute arrival times, if these absolute times are sufficiently accurate, they may simply be subtracted to calculate the relative arrival times frequently measured by direct cross-correlation. Because direct event-event cross-correlation has been shown to fail for this data set, and microseismic characteristics suggest that waveform dissimilarity is to be generally expected for microseismic data (Appendix A), I propose a number of modifications to an adaptive array steering method proposed in Gangi and Fairborn (1968) and referenced more recently, with slight modification, in Rawlinson and Kennett (2004).

### 4.3 Modified Adaptive Steering and Progressive Template Extraction

The modified adaptive steering (MAS) methodology proposed herein improves lag accuracy and precision over conventional cross-correlation by reducing the number of inter-event, dissimilar waveform cross-correlations, and predominantly relying instead on absolute arrival lags derived from cross-correlation of stacked, intra-event (similar) waveforms. I additionally propose a new pre-processing technique, dubbed

progressive template extraction (PTE), to extend the utility of this modified array steering methodology to signal to noise ratios approaching 1.0, where modified adaptive steering tends to fail alone. As I demonstrate in section 4.5, the modified adaptive steering methodology, in tandem with progressive template extraction, improves the accuracy of  $dt(l)$  measurements for both synthetic and field microseismic data, minimizing measurement inconsistency resulting from event-event cross-correlation misalignment, thereby increasing the number of event pairs available for INF based relocation.

Rawlinson and Kennett (2004) note that the original adaptive steering methodology as presented in Gangi and Fairborn (1968) is well suited for dissimilar waveforms. The additions discussed below are designed to further reduce lag measurement error for low SNR ratios, as well as adapt the methodology for measurement of relative event lags.

#### *4.3.1 Modified Adaptive Steering*

##### *4.3.1.1 Original Method Description*

As described in Gangi and Fairborn (1968), Modified Adaptive Steering begins with the picking of time windows containing the arrivals of interest on each receiver. From these absolute picks, initial relative arrival times  $t_i^c$  at each receiver  $i$  are determined and used to perform an initial linear stack of arrival traces  $u_i(t)$  over all  $N$  receivers, to derive an initial composite trace,

$$u_c(t) = \frac{1}{N} \sum_{i=1}^N u_i(t - t_i^c). \quad (4.1)$$

Next, each receiver trace  $u_i$  is cross-correlated with the initial composite trace  $u_c(t)$ , to derive the corrective lag  $\tau_{ic}$  at which the normalized correlation coefficient  $c_i$  is

maximized

$$c_i = \max[f_x(u_c(t), u_i(t - \tau))] \quad (4.2)$$

where  $f_x$  denotes the normalized cross-correlation function. Once these corrective lags have been found, the individual traces are shifted again by  $\tau_{ic}$  to reform the composite trace  $u_c(t)$ ,

$$u_c(t) = \frac{1}{N} \sum_{i=1}^N u_i(t - t_i^c - \tau_{ic}) \quad (4.3)$$

and the cross-correlation procedure is repeated, with the goal of maximizing the average correlation coefficient,

$$\bar{c}_i = \frac{1}{N} \sum_{i=1}^N \max[f_x(u_c(t), u_i(t - \tau))] \quad (4.4)$$

The process is repeated until the average correlation coefficient  $\bar{c}$  converges. The result is a maximally aligned composite trace  $u_m(t)$ , and a series of raw relative arrival lags  $\tau_{ri}$  measured with respect to the composite trace, for a single event.

#### 4.3.1.2 Proposed Modifications

During initial synthetic testing with 7 receivers, it was observed that for signal to white noise ratios below approximately 5, the adaptive steering methodology, as described, failed to make any corrections to the starting arrival times used for initial stacking. Because the initial arrival times do not align traces for optimal signal stacking, corrective lags are frequently 0 for large noise levels, as cross-correlation between stacked traces and individual traces preferentially aligns trace waveforms with their original positions within the stacked trace.

In order to encourage corrective lag recovery under noisy conditions, the original methodology is modified, such that  $N$  unique composite stacks are generated, one for each trace  $i$ . The  $j^{th}$  composite trace  $u_{c,j}(t)$  is instead formed by stacking over all

traces  $u_i$  such that  $i \neq j$ . Equations 4.1 and 4.3 are therefore respectively modified as follows:

$$u_{c,j}(t) = \frac{1}{N-1} \sum_{i=1}^N u_i(t - t_i^c), \quad i \neq j \quad (4.5)$$

$$u_{c,j}(t) = \frac{1}{N-1} \sum_{i=1}^N u_i(t - t_i^c - \tau_{ic}), \quad i \neq j. \quad (4.6)$$

Corrective lags and correlation coefficients are then derived via cross-correlation between individual traces and the corresponding composite stacks generated without them, with the same goal of maximizing the average of the individual correlation coefficients, given by

$$c_j = \max[f_x(u_{c,j}(t), u_i(t - \tau))], \quad i = j. \quad (4.7)$$

Once either correlation coefficients converge, or corrections go to zero, the traces are shifted by their corrective lags, and a single, final composite trace is generated using all  $N$  traces, as given by equation 4.3. For two events  $j$  and  $k$ , the process therefore produce two composite traces  $u_{m,j}(t)$  and  $u_{m,k}(t)$ , and two sets of raw lags at each receiver  $i$ ,  $\tau_{ri,j}$  and  $\tau_{ri,k}$ . The composite traces are then cross-correlated to recover a relative composite lag  $\tau_c$ , and relative event arrivals are finally calculated at each receiver as

$$a_i = \tau_c + \tau_{ri,j} - \tau_{ri,k}. \quad (4.8)$$

#### 4.3.2 Progressive Template Extraction

For SN ratios below approximately 5, even with the modifications proposed above, the adaptive steering methodology often fails to converge when initial relative arrival

times are not sufficiently accurate to coherently stack individual traces. The problem could be mitigated if some template were available to which traces could be compared for generation of the initial composite stack, and other authors have applied template matching to earthquake and even man-made seismic data, (Gibbons and Ringdal, 2006; Plenkers et al., 2013, e.g.); however, these methods generate templates, a priori, to detect subsequent events under the assumption of waveform similarity, which cannot be expected for heterogeneous microseismic first arrivals. Instead, I propose a method to gradually extract a unique template for each microseismic event recorded along a linear array, which will later be shown to improve the coherency of the initial stack generated by the modified adaptive steering technique, improving convergence toward accurate lag values when used as an initial preprocessing step.

Progressive Template Extraction (PTE) begins after first arrivals are detected and approximately windowed on all  $N$  receivers of a linear array. While phase picks will contain some degree of error, particularly in the case of noisy data, PTE only requires approximate time windows containing first arrival waveforms. Starting at one end of an  $N$  receiver array, the 1<sup>st</sup> and 2<sup>nd</sup> receiver arrival windows are cross-correlated, to calculate a lag  $\tau_1$ , used to align and stack the two traces to form the start of a progressive composite stack. This composite is then cross-correlated with the 3rd receiver arrival, which is similarly aligned and stacked with the current composite stack, at this point the summation of 3 aligned traces. Thusly, a composite stack is progressively constructed, gradually aligning and summing over all  $N$  receivers, such that the  $j^{th}$  iteration composite is given by

$$u_{c,j}(t) = \frac{1}{j} \sum_{i=1}^j u_i(t - \tau_i) \quad (4.9)$$

where  $u_i$  is the  $i^{th}$  receiver arrival window and  $\tau_i$  is the lag between the  $i^{th}$  receiver



arrival and the previous stack, which maximizes

$$f_x(u_{c,i-1}(t), u_i(t - \tau)) \quad (4.10)$$

where  $f_x$  again is the cross-correlation function.  $\tau_1$  is always 0.

When inaccurate arrival times are used to simultaneously align and stack all  $N$  traces, as in the unmodified adaptive steering methodology, there is the risk of semi-coherent interference which creates a stacked trace which does not closely resemble individual traces, and therefore results in cycle-skipping or larger misalignment error, and general divergence of subsequent corrective lags following iteration. The issue is further compounded by the gradual shift in first arrival expression along the array, brought about by interfering secondary arrivals (Appendix A.3). This progressive technique instead leverages the greater degree of first arrival similarity between adjacent receivers, and the goal is to opportunistically build a template before arrivals may be incorrectly stacked by MAS. If the first cross-correlation fails to recover the correct lag between receivers 1 and 2,  $N - 1$  additional opportunities remain to correctly stack at least one pair of arrivals to form an initial noisy template, with a boosted signal to noise ratio.

In this manner, the final composite should contain at least one pair of closely aligned traces, and, when the corresponding coherently aligned lags are used as a starting point for modified adaptive steering, the initial stack acts as an empirically extracted template which improves the accuracy of corrective lags and increases the probability of convergence.

#### 4.4 Synthetic Comparison of Modified Adaptive Steering versus Conventional Cross-Correlation

In order to compare the performance of modified adaptive steering, with and without progressive template extraction preprocessing, and conventional cross-correlation, the three methodologies are used to analyze a set of traces containing variable first arrivals to which white Gaussian noise has been added at varied signal to noise ratios. Additionally, the synthetic data is used to demonstrate the validity of two indirect relative performance measurements, before they are used in section 4.5 to evaluate relative performance of the methods as applied to a field recorded microseismic data set.

##### *4.4.1 Model Description and Comparison Framework*

200 total microseismic events were distributed equally and with random positions within two simulated fractures, centered at  $(160, 0, 0)$  and  $(470, 0, 0)$ . In order to ensure approximately equal counts of stationary and non-stationary event pairs, the vertical span of the reference fracture at  $x = 160$  m was limited to 50 m, while that of the locatable fracture at  $x = 470$  m was set to 700 m. A vertical, linear receiver array was centered on the point  $(0, 0, 0)$ , consisting of 7 geophones spaced 30 meters apart. Events and receivers are shown in Figure 4.7. Assuming a homogeneous, attenuation-free, isotropic 4000 m/s velocity model, travel times were calculated from each event to each receiver. Using these travel times, traces containing first arrivals were generated for each event.

Because adaptive steering, with and without template extraction, is explicitly proposed to outperform conventional cross-correlation when microseismic waveforms are dissimilar, synthetic first arrivals were made to represent the general first arrival heterogeneity as well as the gradual interaction of scattered subsequent arrivals

recorded along an array, as observed in the previously introduced real data set. For a given event, using calculated arrival times at each receiver, first arrivals are initially modeled as simple ricker wavelets, with an event-constant peak frequency ranging from 200 Hz-300 Hz. Following the assumptions discussed in Appendix A.2, subsequent arrivals are modeled as near-source scattered energy: the original event hypocenter is randomly perturbed by a distance of one wavelength or less, scattered arrival times are calculated, and traces containing ricker wavelets aligned to these arrival times are summed with the original first arrival trace. Each event trace randomly contained between 2 and 6 scattered arrivals.

These traces, 80 ms long, sampled at 16000 Hz, are then contaminated with white Gaussian noise at varied SNRs. Finally, each trace window is randomly offset, when read and windowed, by  $\pm 50$  ms to simulate picking error. For a given trial, the SNR was kept constant for all events, and varied in integer values from 1-20, 25, and 32, for a total of 22 trials. Event and receiver positions were kept constant between trials. Modified adaptive steering and conventional cross-correlation were used to calculate relative arrival times among all reference-locatable event pairs, and the values were compared to true relative arrival values as generated by the model. Out of 10000 total event comparisons, 4365 pairs were stationary.

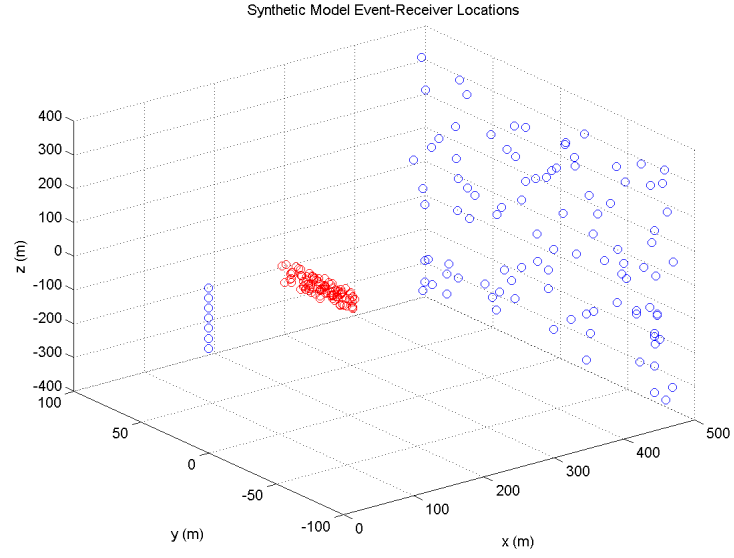


Figure 4.7: Receiver and event locations for this synthetic trial. Reference events are in red, locatable events are in blue.

Three classes of metrics were chosen to compare the performance of MAS, PTE enhanced MAS, and CXC:

1. Direct measures of general lag error
2. Direct measures of INF specific error
3. Indirect measures of relative error

Comparator descriptions and corresponding results are segregated and presented by class in the following subsections.

#### 4.4.2 General Lag Measurement Metrics and Results

In order to evaluate the techniques as general relative arrival time measurement tools, the overall lag error was calculated by averaging over all lag measurements

$$\bar{\epsilon} = \frac{1}{N_p * N_r} \sum_i^{N_r} \sum_j^{N_p} |\tau_{ij,m} - \tau_{ij,t}| \quad (4.11)$$

where  $N_p$  and  $N_r$  are the number of event pairs and receivers, respectively, and  $\tau_{ij,m}$  and  $\tau_{ij,t}$  are the measured and true lag values measured at the  $j^{th}$  receiver for the  $i^{th}$  pair. Additionally, three average correlation coefficients are generated: the average of all conventional inter-event cross-correlations, the average of all intra-event adaptive steering correlations, and the average of all inter-event adaptive steering stack correlations.

Without progressive template extraction, lag error associated with the modified adaptive steering technique was generally larger (Figure 4.8) than that of cross-correlation over all signal to noise ratios. This is partially attributable to the simulated 50 ms initial pick error, which caused iterative corrections to converge to incorrect lags. MAS with PTE, on the other hand, performed on par with conventional cross-correlation, until approximately  $\text{SNR} = 10$ , where the average lag error was lower by as much as 1.25 ms.

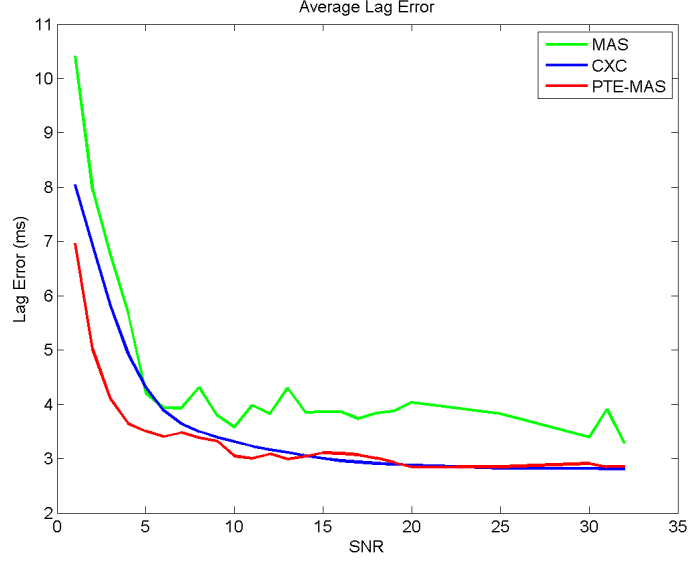


Figure 4.8:  $\bar{\epsilon}$ , average error of all measured lag values, as a function of SNR. CXC in blue, MAS in green, PTE enhanced MAS in red.

The large differences in inter-event, intra-event, and conventional correlation coefficients in Figure 4.9 seem to contradict the relative performance implied by Figure 4.8, as the larger correlation coefficients suggest much better overall lag accuracy. However, as shown by Figures 4.10 and 4.11, modified adaptive steering is prone to generating large outliers. The result is that, although the bulk of MAS errors are closer to 0 than those of CXC, the overall mean error is greater for MAS. The outlier effect is reduced by PTE preprocessing, however it appears that MAS may be able to outperform conventional cross-correlation without preprocessing if outliers can be accurately identified and discarded, and similarly, MAS performance with PTE may be further refined.

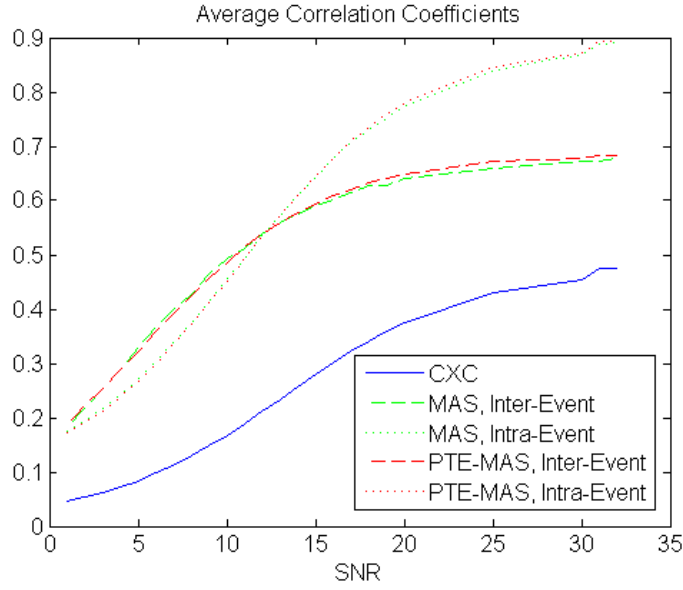


Figure 4.9: Average CXC correlation coefficients (blue), MAS inter-event correlation coefficients (green, dashed), MAS intra-event correlation coefficients (green, dotted), PTE enhanced MAS inter-event correlation coefficients (red, dashed), and PTE enhanced MAS intra-event correlation coefficients (red, dotted), as functions of SNR.

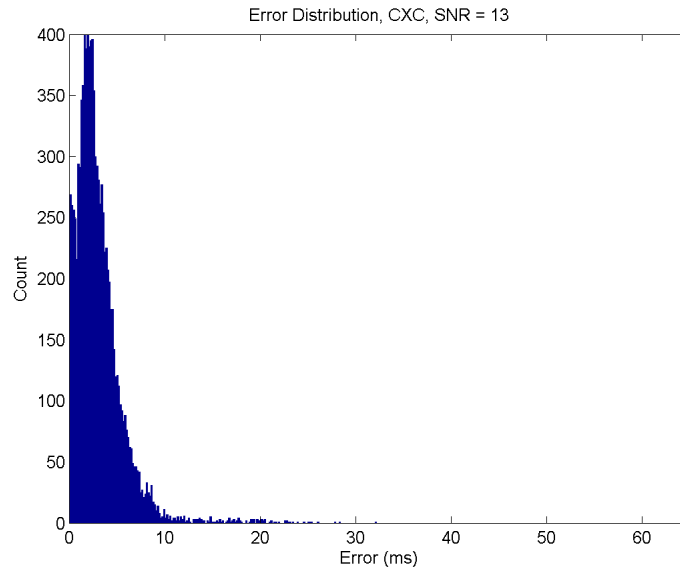


Figure 4.10: Distribution of lag errors calculated by conventional cross-correlation for trial with SNR=13. Bin width = 0.1625 ms.

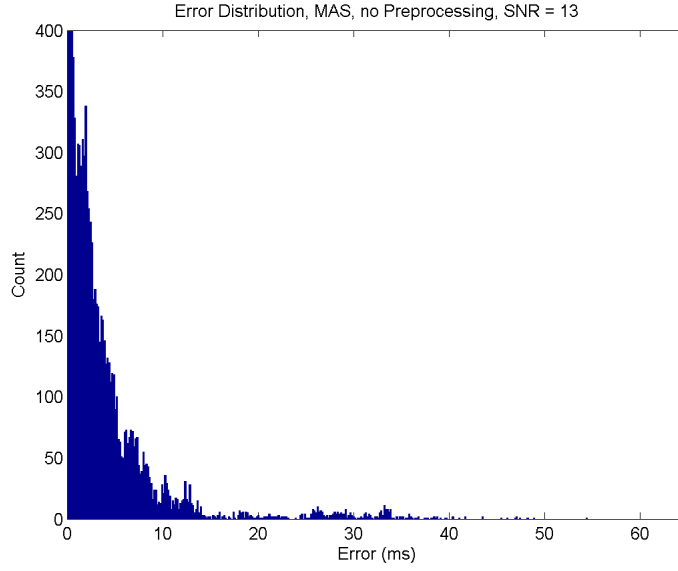


Figure 4.11: Distribution of lag errors calculated by modified adaptive steering, without preprocessing, for trial with SNR=13. Bin width = 0.1625 ms. Note that the y scale has been clipped, the counts for bins 1 and 2 are 614 and 626, respectively.

#### 4.4.3 INF Focused Metrics and Results

The second group of measurements examines the minimization of error within the context of an INF analysis. Counts are conducted of falsely identified stationary pairs (false positives) and incorrectly identified non-stationary pairs (false negatives), both of which may result from cross-correlation error as in the example of events 0 and 4. For each correctly identified stationary pair, the error in the measured  $dt(l)$  maximum is evaluated by calculating the stationary position error

$$\epsilon_l = |l_{stat,m} - l_{stat,t}| \quad (4.12)$$



and the stationary lag error

$$\epsilon_{dt} = |dt_{stat,m} - dt_{stat,t}| \quad (4.13)$$

where subscripts  $m$  and  $t$  indicate measured and true values, respectively. For a given SNR trial, the values are averaged over all correctly identified stationary pairs, such that

$$\bar{\epsilon}_l = \frac{1}{N_s} \sum_i^{N_s} \epsilon_{l,i} \quad (4.14)$$

and

$$\bar{\epsilon}_{dt} = \frac{1}{N_s} \sum_i^{N_s} \epsilon_{dt,i} \quad (4.15)$$

where  $N_s$  is the total number of correctly identified stationary pairs. The stationary position error  $\epsilon_l$  is a measure of the error in the origin of the ray traced from the stationary receiver, through the reference location, onto the location of the distant event. The contribution of  $\epsilon_l$  to final hypocenter error is not straightforward, since it is a function of event-receiver geometry, however it has the effect of modifying the takeoff angle in the ray. The effect of stationary lag error  $\epsilon_{dt}$  is more obvious, with associated hypocenter error  $\epsilon_{x,dt}$  given by

$$\epsilon_{x,dt} = \epsilon_{dt} * V \quad (4.16)$$

where  $V$  is the velocity near the termination of the traced ray, assuming a homogeneous model.

False positive and false negative counts (Figures 4.12 and 4.13) for modified adaptive steering, with and without preprocessing, were both consistently smaller than those of conventional cross-correlation, over almost all tested SNR ratios. For

both counts, PTE-enhanced MAS once again outperformed non-preprocessed MAS, although the margin of improvement was much smaller than that over CXC.

False positives occur when measured  $dt(l)$  functions erroneously contain maxima, while false negatives occur when theoretically stationary  $dt(l)$  functions do not contain detectable maxima. The false positive count is particularly sensitive to cycle-skipping or misalignment, which tends to occur inconsistently for a given pair of events. It is therefore no surprise that, because modified adaptive steering minimizes the propensity for mis-alignment by increasing waveform similarity, this method outperforms CXC in both false positive and false negative counts. At SNR = 1, MAS with PTE produced 17% fewer false negatives and 24% fewer false positives than conventional cross-correlation, showing a reduced propensity for erroneous relocation during INF analysis.

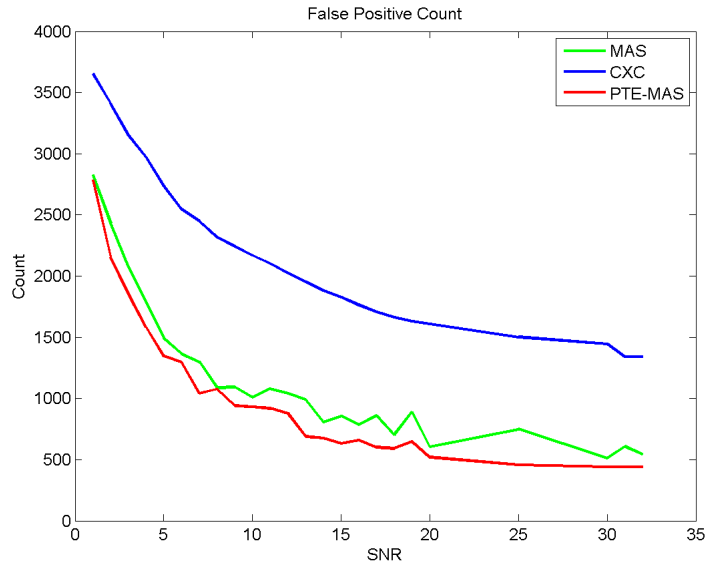


Figure 4.12: False positive counts. CXC in blue, MAS in green, MAS with PTE in red.

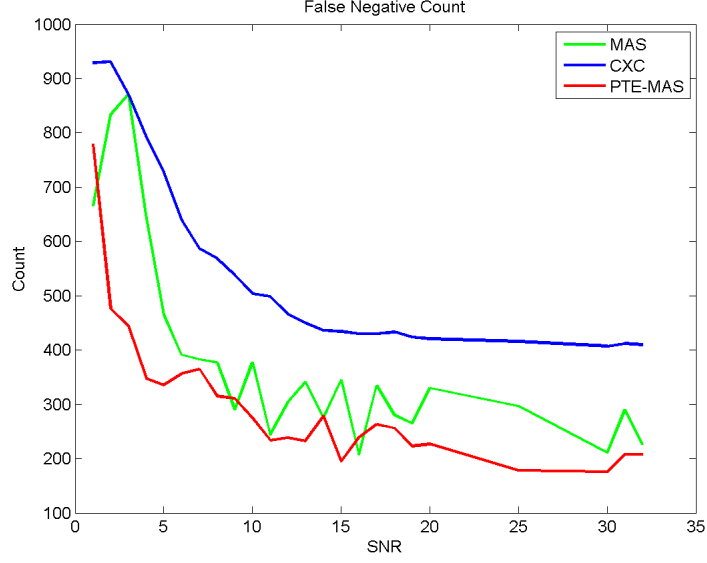


Figure 4.13: False negative counts. CXC in blue, MAS in green, MAS with PTE in red.

Of the correctly identified stationary pairs, the values for CXC and adaptive steering average lag error  $\bar{\epsilon}_l$  (Figure 4.14) resemble those of overall lag error  $\epsilon$ , both in magnitude and in trend. Non-preprocessed MAS produced erratic results, failing to outperform conventional cross-correlation for most signal to noise ratios, likely a result of the outlier effect mentioned in section 4.4.2. With PTE, adaptive steering produced a smaller average error for all SNR values, particularly at SNRs below approximately 6, with an average stationary lag error reduction of 24%, or 2.7ms, for  $\text{SNR} = 1$ . This translates to an average relocation error reduction of 10.8 meters for a velocity of 4000 m/s.

Performance between PTE and non-PTE MAS was more consistent over average stationary position error  $\bar{\epsilon}_{dt}$  (Figure 4.15), although again preprocessed adaptive steering generally provided superior error reduction. The error was reduced con-

sistently over all noise levels, by 22%-33%, bringing identified stationary receiver positions closer to true positions by 10 m-15 m for a 180 m array.

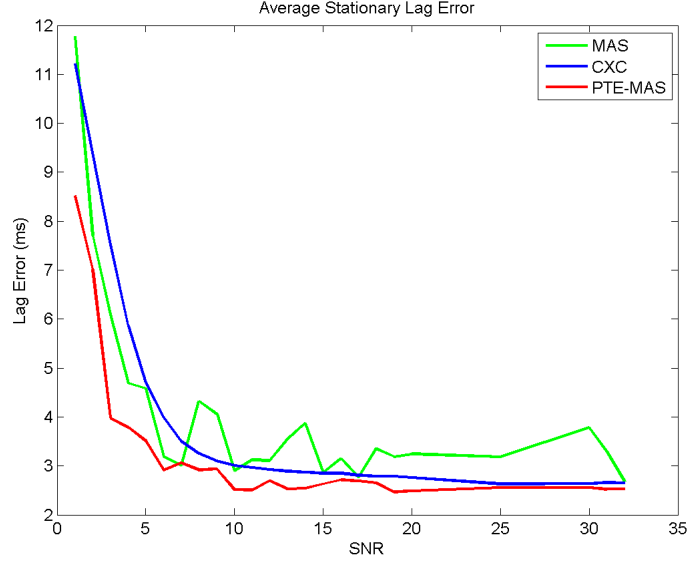


Figure 4.14:  $\bar{\epsilon}_{dt}$ , average stationary lag error for correctly identified stationary pairs, as a function of SNR. CXC in blue, MAS in green, PTE enhanced MAS in red.

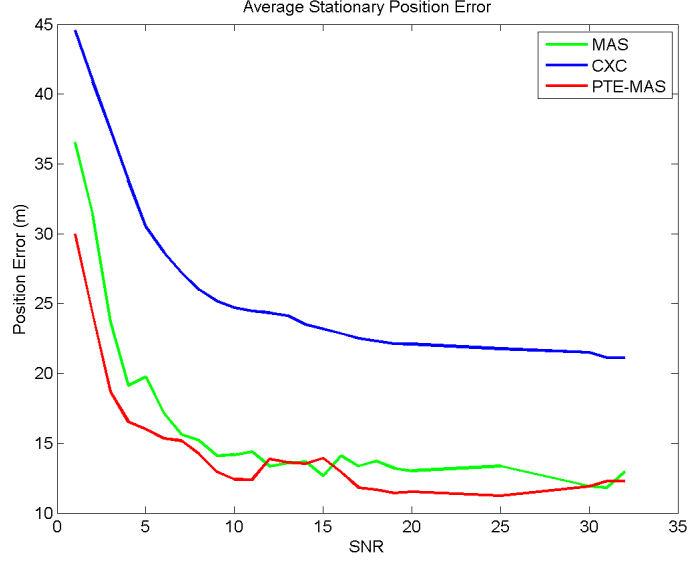


Figure 4.15:  $\bar{\epsilon}_l$ , average stationary position error for correctly identified stationary pairs, as a function of SNR. CXC in blue, MAS in green, PTE enhanced MAS in red.

#### 4.4.4 Indirect Performance Metrics and Results

The third group of measurements has been developed to measure indirectly the relative performance of adaptive steering and conventional cross-correlation in the case that true relative arrival times are not available, as would be true for a field-recorded data set. They are applied to these synthetic data to demonstrate their validity prior to their application in section 4.5.

As hinted in the event 0 and event 4 example shown previously, measured  $dt(l)$  values in cases where CXC fails to properly align dissimilar waveforms on certain receivers show an erroneously large spread. It therefore follows that any methodology with the potential to reduce misalignment should result in lower standard deviations for  $dt(l)$  values measured for any particular event pair. To compare adaptive steering

to CXC, I therefore examine the difference in standard deviations of all event pairs calculated with each methodology. Mathematically, I calculate the average

$$\overline{\Delta\sigma} = \frac{1}{N} \sum_N (\sigma_{dt(l),CXC} - \sigma_{dt(l),AS}) \quad (4.17)$$

Where  $\sigma_{dt(l),CXC}$  and  $\sigma_{dt(l),AS}$  are the standard deviations of the  $dt(l)$  functions for a given pair of events calculated using CXC and adaptive steering, respectively, and  $N$  is the total number of event pairs over which the summation is performed. Separate values are obtained for results with and without progressive template extraction preprocessing. To evaluate the use of standard deviation as a general error proxy, the statistical correlation between  $\overline{\Delta\sigma}$  and the difference  $\epsilon_{CXC} - \epsilon_{AS}$ , as functions of signal to noise ratio, will be examined.

The metric with the most immediate consequences within the framework of an INF analysis is the concept of concavity (i.e. the sign of the second derivative  $dt''(l)$ ). In the case that receiver array length is short compared to array-event distance, the second derivative of a  $dt(l)$  function corresponding to a stationary pair will be of constant sign, and a non-stationary pair may occasionally produce no more than one sign change if measured  $dt(l)$  values are accurate. More importantly, even if measured  $dt(l)$  values for a given pair are approximately accurate, the event pair cannot be used for relocation if the second derivative of the measured function has multiple sign changes, because the stationary lag cannot be reliably identified or found by interpolation. It follows that a more accurate  $dt(l)$  estimator should produce a greater number of  $dt(l)$  measurements with consistent concavity, and the number of pairs with no more than one sign change in  $dt''(l)$  are counted.

Figures 4.18 and 4.17 show the concavity counts and standard deviation differences, respectively, over all tested signal to noise ratios. In terms of usable mea-

measurements of  $dt(l)$  as defined by consistent concavity, progressive template extraction enhanced modified adaptive steering drastically outperforms conventional cross-correlation, producing 1.3 (SNR = 25) to 4.2 (SNR = 1) times as many event pairs with consistent concavity potentially available for relocation.

To evaluate the applicability of standard deviation as a proxy for overall relative error measurement, a statistical correlation coefficient was calculated between the difference in average lag error as a function of SNR,

$$\Delta\bar{\epsilon}(s) = \bar{\epsilon}_{CXC}(s) - \bar{\epsilon}_{AS}(s), \quad (4.18)$$

and the average difference in standard deviations ( $\overline{\Delta\sigma}$ , Equation 4.17) as a function of SNR for preprocessed modified adaptive steering  $dt(l)$  measurements. The cross-plot is shown in Figure 4.16.  $\rho_{\Delta\bar{\epsilon}, \overline{\Delta\sigma}} = 0.92$ , suggesting that the average difference in standard deviation increases in proportion to the relative error between measurements made with modified adaptive steering and conventional cross-correlation.

The adaptive steering methodology introduces an error associated with generating a single reference lag for each event pair. Because the subsequent relative lag measurements are aligned to this single value, the calculated  $dt(l)_{AS}$  lags may consistently under- or over-estimate true arrival lags, in the case that the relative error introduced by cross-correlation is comparatively small. While CXC has its own random error, the error is receiver-specific, such that the errors in lags for a particular  $dt(l)_{CXC}$  function should not be expected to be of equal sign. In order to loosely estimate whether adaptive steering over- or under-estimates lags compared to CXC, the mean value of  $dt(l)$  samples is calculated for each pair, and the differences between

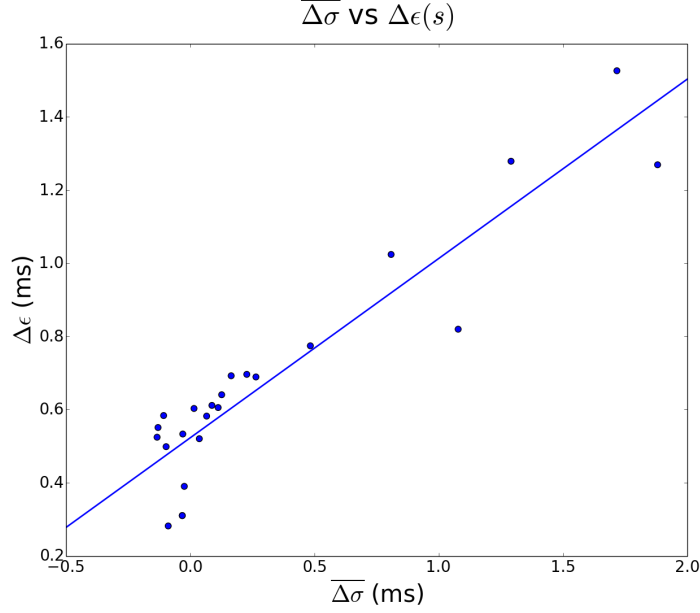


Figure 4.16: Cross-plot of  $\Delta\bar{\epsilon}$  and  $\overline{\Delta\sigma}$ , as functions of SNR, with a line of best fit. The statistical correlation coefficient between the two functions is 0.92, indicating a strong correlation between standard deviation and overall error.

CXC and adaptive steering means are averaged over all event pairs

$$\langle \Delta\bar{t} \rangle = \frac{1}{N} \sum_{i=0}^N (\overline{dt(l)_{i,CXC}} - \overline{dt(l)_{i,AS}}) \quad (4.19)$$

where  $dt(l)_{i,CXC}$  and  $dt(l)_{i,AS}$  are the lag functions returned by the two different methodologies for event pair  $i$ , and  $N$  is the total number of event pairs. Here again, two values are calculated, with and without preprocessing.

Based on the results in Figure 4.19, there does not appear to be a discernible relationship between signal to noise ratios and  $\langle \Delta\bar{t} \rangle$ . As an indirect measure of relative performance, the metric is therefore abandoned. However, Figure 4.19 does demonstrate that there is no consistent tendency for MAS to under- or over-estimate relative lags as compared to conventional cross-correlation.



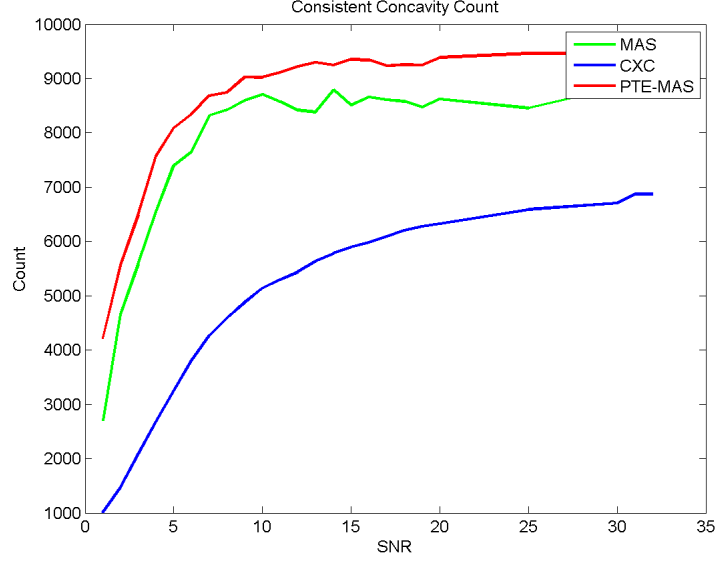


Figure 4.17: Count of  $dt(l)$  measurements with consistent concavity for CXC (blue), MAS (green), and PTE enhanced MAS (red)

#### 4.5 Real Data Evaluation of Modified Adaptive Steering versus Conventional Cross-Correlation

An analysis was conducted using the same shale data set introduced in section 4.2 to assess the relative performance of adaptive steering compared to CXC. Because true relative arrival times cannot be known for microseismic field data, two indirect metrics, standard deviation and concavity count (as introduced in section 4.4), were used for this comparison, emphasizing the measurement characteristics that are necessary for an INF analysis. For the purpose of this trial, 192 events with clear first arrivals were windowed to a time width of 0.08 s. For each pair of events, a lag function  $dt(l)$  was calculated using both CXC and progressive template extraction enhanced modified adaptive steering, and the results were examined using the assessors introduced and verified in section 4.4. First arrivals for initial adaptive

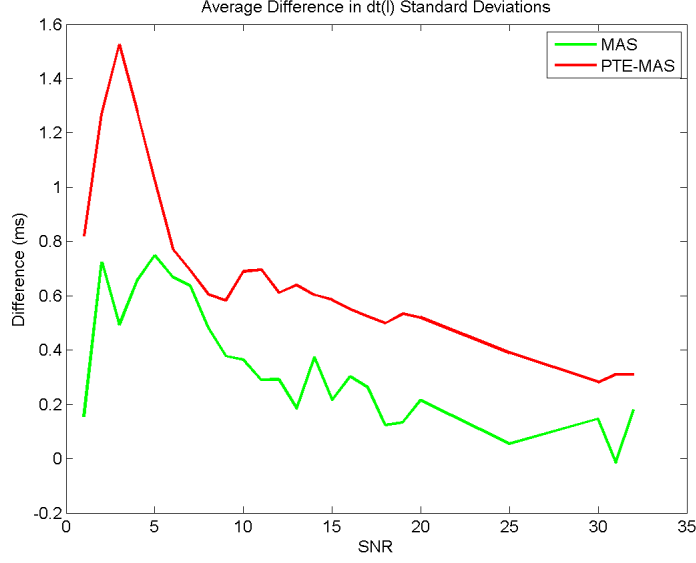


Figure 4.18:  $u$ , average difference in standard deviation of  $dt(l)$  functions, as a function of SNR. MAS-CXC in green, PTEMAS-CXC in red.

steering alignment were picked by hand.

Comparing standard deviations using equation 4.17 from section 4.4, the average difference in standard deviations of  $dt(l)$  functions,  $\overline{\Delta\sigma}$ , was 3.89 ms, suggesting that, on average, adaptive steering  $dt(l)$  functions have more tightly grouped values, and therefore are less likely to be misaligned by cross-correlation. Extrapolating the relationship between  $\overline{\Delta\sigma}$  and average lag error indicated by the synthetic comparison in section 4.4 suggests an overall average lag reduction of approximately 6.7 ms.

Out of a total of 18528 event pairs, modified adaptive steering produced 8748 consistent curves, while CXC produced just 2997, an improvement factor of 2.9 for the total number of potentially relocatable event pairs. This improvement is consistent with the reduction in standard deviations of  $dt(l)$  measurements, and suggest that for real data, the number of sporadic cross-correlation errors is generally reduced with modified adaptive steering. Note that this metric only loosely measures the

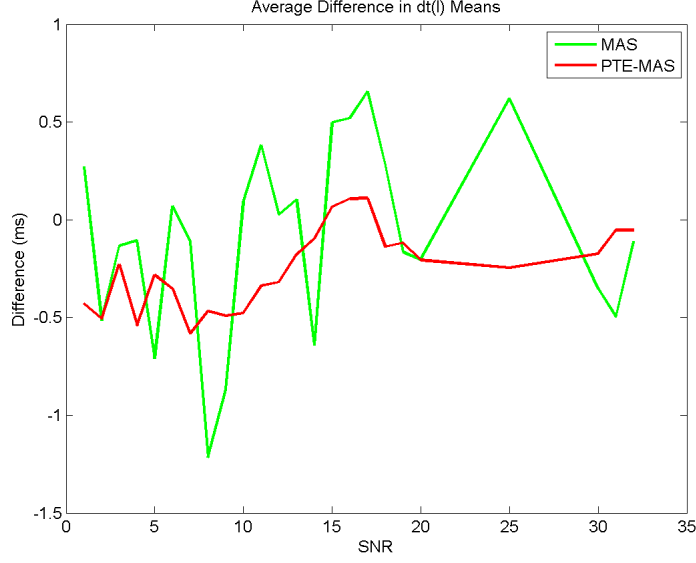


Figure 4.19:  $\langle \Delta \bar{t} \rangle$ , the average of difference in  $dt(l)$  means, as a function of SNR. CXC-MAS in green, CXC-PTEMAS in red. Note that there does not appear to be any major trend with respect to SNR, although the variability in PTE enhanced MAS  $\langle \Delta \bar{t} \rangle$  values is markedly smaller.

number of pairs available for relocation; even with 7 receivers, it should be possible to allow for an occasional outlying  $dt$  measurement and a relaxation of the concavity constraint.

Aggregate statistics, justified by the synthetic test of section 4.4, therefore suggest that the adaptive steering method is generally more accurate for recovering inter-event lags when applied to data sets with inconsistent event waveforms.  $dt(l)_{AS}$  functions have generally smaller standard deviations compared to  $dt(l)_{CXC}$  functions for the same event pairs, and  $dt(l)_{AS}$  functions are 2.9 times as likely to show consistent concavity, as would be expected for true  $dt(l)$  functions.

#### 4.6 Comparison of Adjacent Steering and CXC Under Idealized Conditions

In order to further demonstrate the validity of modified adaptive steering, its performance relative to conventional inter-event cross-correlation is examined for the using the Z-components of the single, relatively ideal pair of events 0 and 9, seismograms shown in Figure 4.20. The ratio of trace scales is approximately 16:1, event 9:event 0. Both events are predominantly polarized in the Z-component direction.

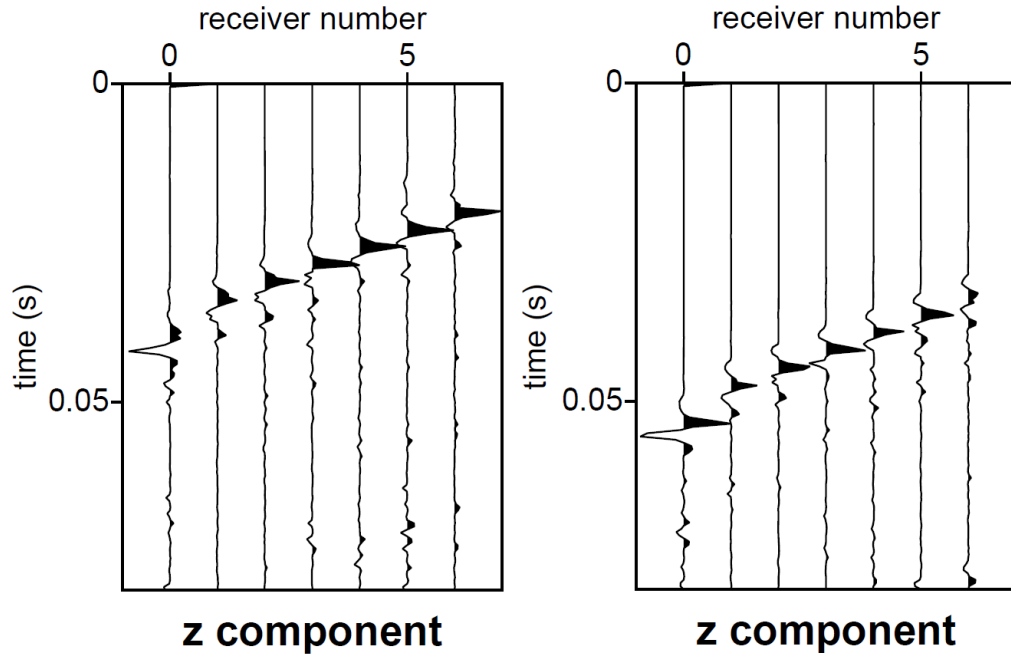


Figure 4.20: Traces for events 0 and 9. Event 9:event 0 scale ratio is approximately 16:1.

The average CXC correlation coefficient is 0.83, which indicates that inter-event waveforms are indeed similar. In contrast, even in this idealized case where both waveforms contain well defined sharp impulses, the average correlation coefficients

returned by adaptive steering after alignment are 0.88 for event 0 and 0.90 for event 9, while the coefficient for the composite trace correlation is 0.93. This difference suggests that even in the best field cases between similar waveforms, both composite and intra-event waveforms generally correlate to a greater degree than individual waveforms of different events. The  $dt(l)$  functions are shown in Figure 4.21, CXC in black, adaptive steering in gray. Note that the trends are similar, with a normalized correlation coefficient between  $dt(l)_{AS}$  and  $dt(l)_{CXC}$  of 0.84.

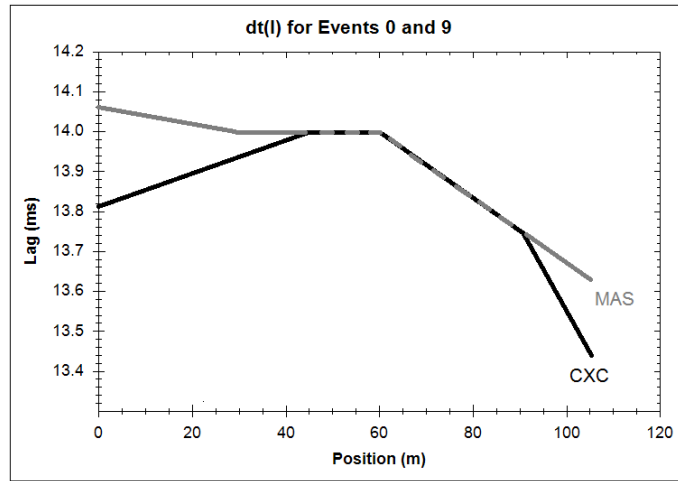


Figure 4.21:  $dt(l)$  functions generated by conventional cross-correlation (black) and adaptive steering (gray).

The standard deviations for both  $dt(l)$  curves are comparable, .18 ms and .15 ms for CXC and adaptive steering, respectively. The differences in this case are likely attributable to small trace variations in first arrival peak widths compare, for example, traces on receiver 6, where event 9 has a doubly peaked arrival.

The average time lags, 13.83 ms and 13.90 ms, for  $dt(l)_{CXC}$  and  $dt(l)_{AS}$ , respec-

tively, are comparable in magnitude. This corresponds to a negligible relocation error of approximately 0.28 meters for a P-velocity of 4000 m/s. It is not immediately obvious which method is closest to the actual time lags, as it is impossible to measure true lag times to this level of precision with any current methodology.

Both curves pass the concavity test, although only the  $dt(l)_{CXC}$  curve contains a maximum, between 40 m and 60 m. Classically derived locations provided by a third-party processing company indicate that events 0 and 9 are not stationary, suggesting that the adaptive steering methodology, for this idealized case, may be more accurate. The error in  $dt(l)_{CXC}$  is a manifestation of cross-correlation misalignment, although the error is not large enough to be attributed to fully skipped cycles. Instead, this event pair demonstrates the precision limitation of cross-correlation under relatively ideal waveform conditions the relative arrival time values differ by no more than .25 ms, which corresponds to just 4 samples at a sample frequency of 16000 Hz. At this scale, the resolution of relative arrival times with cross-correlation is compromised by small variations in peak widths, on the order of tenths of milliseconds, and cannot be improved by modified adaptive steering with progressive template extraction.

#### *4.6.1 Modified Adaptive Steering Advantages in Detail*

Compared to subtraction of picked arrival times, modified adaptive steering with progressive template extraction preprocessing produces more accurate measures of relative arrival times, because it makes use of cross-correlation for absolute arrival measurement. The methodology outperforms conventional event-event cross-correlation (CXC) for highly variable microseismic data because most cross-correlation operations are performed on self-similar waveforms from individual events, with comparisons between individual traces and their event-specific composite stacks.

In addition to general similarity, for events recorded along a linear array, first arrival polarity will generally be consistent among all traces, limiting the number of potential erroneous correlogram maxima to positive values. This is in contrast to CXC, which must consider both positive and negative maxima during lag calculation, effectively doubling the potential for error for noisy correlograms.

MAS is particularly useful within the context of an INF analysis, when considering the constraints imposed on measurement of the relative lag function  $dt(l)$  by the INF methodology. Because cross-correlation is limited to similar waveforms, the  $dt(l)$  measurements have been shown to be more consistent, with fewer mis-alignment errors may shift, occlude, or erroneously produce a stationary  $dt(l)$  point. The major error in calculating  $dt(l)$  values is instead expected to occur on cross-correlation of composite traces. While the resulting reference lag is constant at each receiver, the synthetic trial in section 4.4 demonstrated that the error is random among event pairs, such that it may be minimized during stacking over the greater number of pairs produced by PTE enhanced MAS for redundant relocation.

The SNR improvement afforded by trace stacking is especially useful in the field of microseismic processing, given that microseismic events associated with hydraulic stimulation are typically low in magnitude and can easily be drowned out by the incoherent noise which cancels during array steering. There is, however, the question of the effect that stacking has on scattered arrivals, specifically those that comprise the waveform immediately following the first arrival, since scattered energy may not be strictly incoherent, depending on the spatial of distribution of scattering bodies with respect to relatively short monitoring arrays. In the traces for event 14 shown in Figure 4.22, there is a seemingly gradual change in peak amplitudes and distributions from the 1<sup>th</sup> to the 7<sup>th</sup> receiver. This steady change with receiver position  $l$  is likely a consequence of constructive and destructive interference among

scattered arrivals with slightly different polarizations with respect to array orientation. The small variations in slowness vectors of these scattered arrivals translate to relative scattered arrival times which differ by less than one wavelength and result in a relative phase difference which varies as scattered arrivals impinge along the array. Although scattered energy is coherently recorded on multiple receivers, when traces are aligned to first arrival peaks these scattered post first arrival waveforms frequently cancel, reducing or eliminating the propensity for cycle-skipping or larger misalignment during composite trace correlation between events. Figure 4.22 shows an example of this cancellation, where event 14 traces have been aligned and stacked with the combined PTE-MAS methodology, and post first arrivals have been muted in the composite.

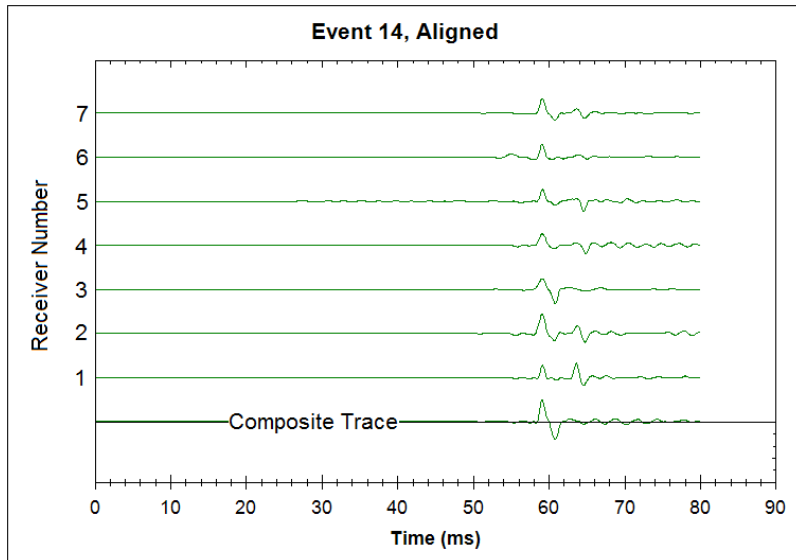


Figure 4.22: Traces for event 14 which have been aligned with the adaptive steering method, along with the resulting stacked composite at the bottom. Note the large increase in relative amplitudes between the initial arrival and energy arriving after 60 ms, an example of the cancellation which MAS leverages.



## 4.7 Summary

Ignoring the effects of scattering and propagation, even typical microseismic monitoring geometries easily approach and surpass the modern limits of relative travel time measurement techniques, including cross-correlation. Because the three assumptions that ensure cross-correlation accuracy in light of scattering and non-ideal propagation are generally invalid for microseismic datasets, there is no guarantee that event waveforms will produce large coefficients on cross-correlation, and the errors, which result from misalignment of dissimilar signals, reduce the number of stationary pairs which may be identified and used for relocation by INF, while additionally calling the validity of any identified stationary pairs into question.

While shrinking cross-correlation window widths and limiting stationary analyses to pairs with large correlation coefficients do not appear to be feasible for accuracy improvement, the proposed modified adaptive steering (MAS) technique, enhanced by progressive template extraction (PTE), outperforms conventional inter-event cross-correlation (CXC) in the synthetic case of section 4.4, comparatively minimizing overall lag, stationary position and stationary lag measurement error, reducing  $dt(l)$  spread, and improving the consistency of  $dt(l)$  concavity, returning  $dt(l)$  measurements which appear to be more consistent with the theoretical shape of the  $dt(l)$  curve, and producing more consistent, interpolatable  $dt(l)$  measurements that can be used for INF relocation. Using standard deviation and concavity counts as proxies for indirect relative performance measurement, PTE-MAS shows promise as a superior relative lag calculation methodology when applied to real microseismic data.

Because of consistency and accuracy limitations, INF relocation of microseismic events using CXC may generally not be possible; however, the applicability of INF

analysis may be greatly extended if PTE-MAS is used for relative travel-time calculation.

## 5. CONCLUSION

For typical linear array monitoring geometries associated with hydraulic stimulation, relative travel time measurements can quickly approach or surpass the precision limitations of modern measurement techniques. Although other relative localization methodologies, E.G. Double Difference, can tolerate some degree of normally distributed error by operating on large numbers of events simultaneously, the Interferometric Neighboring Fracture method is unique in that relative travel time measurements must be accurate and consistent enough for interpolation during stationary analysis of individual event pairs. Before an INF based relocation is conducted, care must be taken to ensure that array spacing, orientation, and placement relative to induced events are such that stationary analyses are possible and accurate, and that the number of expected stationary reference events per located event is greater than the number of monitoring receivers.

The high frequency content of microseismicity and the spatial distribution of hypocenters relative to monitoring arrays lead to a large variability among event waveforms, with the result that cycle-skipping and general cross-correlation misalignment error limit the accuracy of conventional inter-event cross-correlation, resulting in inconsistent  $dt(l)$  functions, which either produce erroneous stationary positions, or cannot be interpolated for relocation. As an alternative, the modified adaptive steering methodology (MAS), with progressive template extraction (PTE) preprocessing, has been shown to generate more accurate relative arrival time measurements, which meet the unique usability requirements of the INF method, by limiting errors introduced during cross-correlation of dissimilar waveforms.

If, prior to conducting an INF based relocation, an initial analysis of event-

receiver geometry suggests a large enough difference in relative lag times across a linear monitoring array to produce consistent, interpolatable relative time measurements, as well as a large number of expected stationary event pairs, then improved measurement consistency afforded by the modified adaptive steering technique will provide a larger number of well-determined stationary pairs. With the resulting increase in reference events per relocated hypocenter, the Interferometric Neighboring Fracture method can be expected to robustly relocate distant events and surpass the precision of classical, individual localization techniques.

## REFERENCES

- 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.
- Curtis, A., 2009, Source-receiver seismic interferometry: SEG Technical Program Expanded Abstracts 2009, 3655–3659.
- Curtis, A., and D. Halliday, 2010, Source-receiver wave field interferometry: *Phys. Rev. E*, **81**, no. 4, 046601.
- Eisner, L., D. Gei, M. Hallo, I. Opral, and M. Y. Ali, 2013, The peak frequency of direct waves for microseismic events: *Geophysics*, **78**, A45–A49.
- Fehler, M., 1989, Stress control of seismicity patterns observed during hydraulic fracturing experiments at the fenton hill hot dry rock geothermal energy site, new mexico: *International Journal of Rock Mechanics and Mining Sciences and Geomechanics Abstracts*, **26**, 211 – 219.
- Gangi, A., and J. Fairborn, 1968, Accurate determination of seismic-array steering delays by an adaptive computer program: *Nuovo Cimento*, **6**, 105–115.
- Geiger, L., 1912, Probability method for the determination of earthquake epicenters from the arrival time only: *Bull.St.Louis.Univ*, **8**, 60–71.
- Geller, R. J., and C. S. Mueller, 1980, Four similar earthquakes in central california: *Geophysical Research Letters*, **7**, 821–824.
- 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.
- Gibowicz, S. J., and A. Kijko, 1994, *in* An introduction to Mining Seismology: Aca-

- demic Press Inc. Waltham, Massachusetts.
- Havskov, J., and L. Ottemöller, 2010, Routine data processing in earthquake seismology: With sample data, exercises and software, 1st edition. ed.: Springer, Boston, MA.
- Hurd, O., 2012, Geomechanical Analysis of Intraplate Earthquakes and Earthquakes Induced During Stimulation of Low Permeability Gas Reservoirs : PhD thesis, Stanford University.
- Lay, T., and T. Zhang, 1992, Near-source contributions to teleseismic p waves and p-wave coda for underground explosions: Bulletin of the Seismological Society of America, **82**, 383–405.
- Maxwell, S., J. Rutledge, R. Jones, and M. Fehler, 2010, Petroleum reservoir characterization using downhole microseismic monitoring: GEOPHYSICS, **75**, 75A129–75A137.
- Mayerhofer, M. J., E. P. Lolon, J. E. Youngblood, and J. R. Heinze, 2006, Integration of microseismic fracture mapping results with numerical fracture network production modeling in the barnett shale: Society of Petroleum Engineers.
- Meek, R., R. Hull, H. Bello, D. Portis, and B. Suliman, 2013, 237, *in* What Broke? Microseismic analysis using seismic derived rock properties and structural attributes in the Eagle Ford play: 1443–1454.
- Melo, G., A. Malcolm, and M. Fehler, 2012, 347, *in* Comparison of microearthquake locations using seismic interferometry principles: Massachusetts Institute of Technology. Earth Resources Laboratory, Cambridge, MA, 1–5.
- Molyneux, J. B., and D. R. Schmitt, 1999, First-break timing; arrival onset times by direct correlation: Geophysics, **64**, 1492–1501.
- Phillips, W. S., 2000, Precise microearthquake locations and fluid flow in the geothermal reservoir at soultz-sous-forets, france: Bulletin of the Seismological Society of

- America, **90**, 212–228.
- Plenkers, K., J. Ritter, and M. Schindler, 2013, Low signal-to-noise event detection based on waveform stacking and cross-correlation: application to a stimulation experiment: *Journal of Seismology*, **17**, 27–49.
- Poliannikov, O., A. E. Malcolm, H. Djikpesse, and M. Prange, 2011, Interferometric hydrofracture microseism localization using neighboring fracture: *GEOPHYSICS*, **76**, WC27–WC36.
- Poliannikov, O. V., M. Prange, A. Malcolm, and H. Djikpesse, 2013, A unified bayesian framework for relative microseismic location: *Geophysical Journal International*.
- Poliannikov, O. V., S. Rondenay, and L. Chen, 2012, Interferometric imaging of the underside of a subducting crust: *Geophysical Journal International*, **189**, 681–690.
- Poliannikov, O. V., and M. E. Willis, 2011, Interferometric correlogram-space analysis: *Geophysics*, **76**, SA9–SA17.
- Portis, D. H., M. Murray, H. Bello, P. Clarke, G. Barzola, and K. Canan, 2013, 177, *in* Searching for the Optimal Well Spacing in the Eagle Ford Shale: A Practical Tool-Kit: AAPG, Tulsa, OK, 237–244.
- Poupinet, G., W. L. Ellsworth, and J. Frechet, 1984, Monitoring velocity variations in the crust using earthquake doublets: An application to the calaveras fault, california: *Journal of Geophysical Research: Solid Earth*, **89**, 5719–5731.
- Ram, A., and R. F. Mereu, 1975, A comparison of adaptive processing techniques with nth root beam forming methods: *Geophysical Journal of the Royal Astronomical Society*, **42**, 653–670.
- Rawlinson, N., and B. L. N. Kennett, 2004, Rapid estimation of relative and absolute delay times across a network by adaptive stacking: *Geophysical Journal International*, **157**, 332–340.

- Rost, S., and C. Thomas, 2009, Improving seismic resolution through array processing techniques: *Surveys in Geophysics*, **30**, 271–299.
- Schaff, D. P., and P. G. Richards, 2004, Lg-wave cross correlation and double-difference location: Application to the 1999 xiuyan, china, sequence: *Bulletin of the Seismological Society of America*, **94**, 867–879.
- Schimmel, M., and M. Assumpcao, 2003, Seismic velocity anomalies beneath se brazil from p and s wave travel time inversions: *JOURNAL OF GEOPHYSICAL RESEARCH*, **108**.
- Schuster, G. T., 2009, *Seismic interferometry*: Cambridge University Press, Cambridge, **1**.
- Schuster, G. T., and M. Zhou, 2006, A theoretical overview of model-based and correlation-based redatuming methods: *Geophysics*, **71**, SI103–SI110.
- Schweitzer, J., 2001, Hyposat an enhanced routine to locate seismic events: *pure and applied geophysics*, **158**, 277–289.
- Sileny, J., D. P. Hill, L. Eisner, and F. H. Cornet, 2009, Nondouble-couple mechanisms of microearthquakes induced by hydraulic fracturing: *Journal of Geophysical Research: Solid Earth*, **114**.
- Snieder, R., A. Grt, H. Douma, and J. Scales, 2002, Coda wave interferometry for estimating nonlinear behavior in seismic velocity: *Science*, **295**, 2253–2255.
- Vesnaver, A., L. Lovisa, and G. Bohm, 2008, Full 3D relocation of microseisms for reservoir monitoring: *SEG Technical Program Expanded Abstracts*, SEG, 1441–1445.
- 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.
- Wapenaar, K., 2004, Retrieving the elastodynamic green’s function of an arbitrary



- inhomogeneous medium by cross correlation: Phys. Rev. Lett., **93**, no. 25, 254–301.
- Wapenaar, K., 2006, 46, *in* Seismic interferometry for passive and exploration data: SEGJ, Tokyo, Japan, 1–6.
- Warpinski, N. R., M. Mayerhofer, K. Agarwal, and j. Du, 2013, Hydraulic-fracture geomechanics and microseismic-source mechanisms: Society of Petroleum Engineers, **18**, no. 04, 766–780.
- Waters, G. A., H. Ramakrishnan, J. Daniels, D. Bentley, J. Belhadi, and D. Sparkman, 2009, Ss: Unlocking the unconventional oil and gas reservoirs: Utilization of real time microseismic monitoring and hydraulic fracture diversion technology in the completion of barnett shale horizontal wells.
- Zhang, H., and C. H. Thurber, 2003, Double-difference tomography: The method and its application to the hayward fault, california: Bulletin of the Seismological Society of America, **93**, 1875–1889.

## APPENDIX A

### SOURCES, PROPAGATION, AND WAVEFORM DISSIMILARITY

For correlation on global scales and geometries typical of earthquake seismology, the following assumptions are expected to ensure generally similar waveforms (e.g., Arrowsmith and Eisner, 2006; Poupinet et al., 1984; Waldhauser and Ellsworth, 2000; Geller and Mueller, 1980):

1. Seismic sources are generated by the same source mechanism
2. Seismic sources are separated by roughly of the dominant wavelength or less
3. Correlated events follow a relatively common raypath

As demonstrated below, these assumptions are typically not valid during microseismic monitoring, resulting in largely dissimilar microseismic waveforms and poor Pearson coefficients during cross-correlation.

In the context of recorded waveform character, the manifestation of non-ideal seismic propagation may be separated into two effects: a change in the initial impulse (first arrival) shape and phase, and the addition of varying post first arrival energy. When the above conditions are satisfied, it is generally valid to presume that recorded waveforms will be generated and affected similarly during propagation, and that events should therefore correlate strongly.

#### A.1 First Arrival Expression

First arrival waveshape is dependent on source mechanism and frequency dependent attenuation along the ray path. The orientation of the focal source with respect to a given receiver determines the polarity and relative amplitudes of measured compressional and shear arrivals, since all non-isotropic sources produce angularly depen-

dent radiation patterns. Consider the illustration below, which depicts the idealized compressional (blue) and shear (red) amplitudes predicted as a function of receiver angle relative to a double couple shear source:

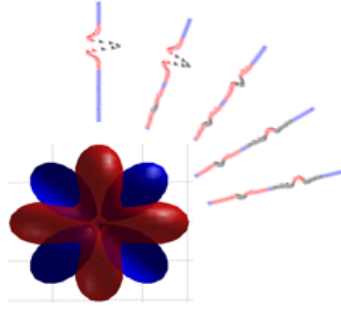


Figure A.1: Radiation Patterns for a double couple source. Blue lobes represent amplitudes of P-waves, red lobes show amplitudes for S-waves, and traces for receivers at varying angular offsets are shown in the top right corner. Note: lobes are not to scale.

Note how as the receiver is moved in the X-Y plane around the focal source, the first arrival polarizations and amplitudes vary. If conditions 1 and 3 are satisfied above, it follows that since the source mechanisms and angular offsets of two events at a given receiver are similar, the first arrivals should have the same polarity and relative phase amplitudes. Within the context of hydraulically induced microseismicity monitored by a linear array, conditions 1 and 3 are frequently not valid. Consider the following map view illustration, with two fracture planes given in blue, and two potential travel paths from events to a vertical receiver array:

Because focal mechanisms are expected to align approximately with local field stresses (e.g., Fehler, 1989), the focal mechanisms for events A and B are aligned identically with respect to each other. However, the relative angle to events A and B from the receiver array differ by more than 45, invalidating conditions 1 and 3 and

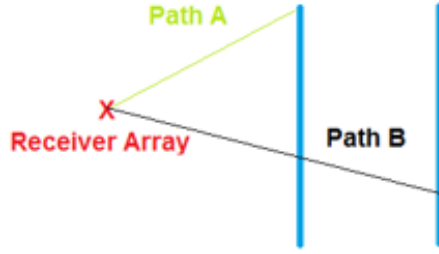


Figure A.2: Two fracture planes (blue) and two potential raypaths A and B to a hypothetical receiver array. Condition 1 is invalid, because of the angular difference from the array to the events, while condition 3 also fails to hold as the rays follow different paths.

consequently, the polarizations and amplitudes of first arrivals as measured at the receiver array cannot be expected to match.

Assumption 1 is additionally faulty when recent literature regarding source mechanisms of hydraulically induced events is considered. While conventional literature still reports that microseismic sources are strictly shear events (Warpinski et al., 2013), Sileny et al. (2009) less recently analyzed microseismic data from a Texas gas field, inverting amplitude and polarity measurements for focal mechanisms, and concluded that microseismic sources from the same stimulation are represented by event-specific linear combinations of double couple, isotropic and CLVD moment tensors. If their conclusions are correct and apply generally to microseismic events, then it is invalid to expect polarization and amplitude similarity for waveforms of cross-correlated events which have an identical angular offset with respect to receiver position, since the very radiation patterns of events may differ.

One final effect of non-ideal spatial distribution associated with condition 3 is a result of path dependent attenuation and dispersion. Frequency dependent attenuation and dispersion have the effect of widening the time signature of first arrivals by preferentially muting high frequencies, as well as shifting the phase of first arrivals.

These two phenomena both create the potential for error in relative arrival time estimation.

## A.2 Post First Arrival Expression

When the three above assumptions are not met, waveform presentation is further complicated by non ideal propagation of energy which is recorded after the first arrival. In practice there is an error associated with picking of first arrival times, compounded by multiple arrivals, signal noise, attenuation, and dispersion. Cross-correlation is typically able to mitigate the effect of these picking errors on relative arrival time measurements because it operates on time-windows which contain first arrivals. Unfortunately, seismic first arrivals are generally immediately followed by post first arrival energy, which can confound cross-correlation based relative arrival-time extraction when subsequent waveforms have amplitudes similar to or greater than first arrivals (See section 4.2 and appendix B for an empirical illustration of this error).

Post first arrival wavetrains are predominantly composed of scattered seismic energy (Snieder et al., 2002; Lay and Zhang, 1992), with additional contributions from variably dispersed/attenuated frequencies. The scattered energy may be loosely characterized by its origin, created by near-source scattering or by distant scattering, overburden, and/or free surface interactions. The main reason for this differentiation during consideration of signal similarity is the angular dependence of scattered waveforms. In the case of near-source scattering, energy scattered within 360 of a hypocenter can in principle contribute to waveforms shortly after the first arrival, since the difference between path lengths of the actual source and a scattering body is minimal. However, as scatterers move away from the hypocenter, there is a smaller angular window within which a signal may be scattered before the path length be-

comes too large to arrive in time, or at all if attenuation is considered.

Herein lies the justification for the quarter wavelength constraint placed on hypocenter separation: if two events are separated approximately by  $\frac{1}{4}$  of the dominant wavelength, it may be generally expected that the two event signals will be affected similarly by both near and far source scattering, and therefore present similar post first arrival waveforms, (e.g., Arrowsmith and Eisner, 2006; Geller and Mueller, 1980). This condition is generally met within the scope of earthquake seismology, where frequencies of interest are on the order of 1Hz or less, enabling multiplet analysis with event offsets on the order of kilometers. Microseismic frequency content however is much higher; Eisner et al. (2013) report peak frequencies of around 30Hz at a 2000m offset surface array, while signal analysis of frequency spectra of events from the Naughtwell Stage 2 data set examined later in this work show peaks in the range of 200Hz-300Hz for a downhole monitoring array with receiver-event offsets in the range of hundreds of m. Even conservatively taking the lower frequencies reported by Eisner et al., corresponding roughly to a maximum inter-event separation of 30m, it is impossible to satisfy the quarter wavelength constraint for the majority of microseismic pairs, since fracture stages alone are separated by hundreds of meters. The result, as demonstrated in section 4.2, is a large variation among post first arrival waveforms, leading to poor correlation coefficients and inaccurate relative traveltimes measurement.

### A.3 Intra-Event Waveform Similarity

A final characteristic of microseismic waveforms worth mentioning is the variation of waveforms for individual events as recorded at multiple receivers along a closely spaced array. As evidenced in the following section, individual microseismic events show startlingly large variation in both first arrival and coda representation

among signals recorded along an array. While this phenomenon has yet to be explicitly addressed in literature, it appears as though the mechanism responsible for the differences is strongly angularly dependent, since the angular difference between any two receivers to a given event is no more than a few degrees. It is therefore presumed to be a manifestation of either near-source scattering, or scattering by heterogeneities within tens of meters of array receivers. The large ratio of non-direct/direct arrival amplitudes, in many cases greater than 1, may be attributed to near source S-P conversion, in the absence of an alternative explanation.

In conclusion, for applications of the INF method to microseismic data, accurate measurement of relative arrival times by inter-event cross-correlation cannot be expected, since the three conditions which generally ensure signal similarity between two given events cannot be satisfied given the scale and geometry of hydraulic stimulation and monitoring. Source/angle variations and unacceptably large inter-event distances contribute to differences in waveforms of both first and subsequent arrivals, arising from scattering, attenuation, and differences in source mechanism and orientation relative to receivers. Section 4.2 demonstrates the specific effect that these non-ideal characteristics may have on an application of the INF method with an analysis of a real data set.

## APPENDIX B

### CROSS-CORRELATION ERROR IN DETAIL

As typical for cross-correlation based relative arrival time measurement, the extraction of a relative arrival time between two arrivals consists of generating a cross-correlogram for two signals, and identifying the lag value for which the correlation coefficient is at a maximum. Qualitatively, this lag time represents the time shift at which the signals are most similar when visually overlain and, in an ideal case in which two first arrival waveforms are perfectly impulsive, this lag would exactly recover the true relative arrival time. To understand the source of the error in the highly non-ideal case resulting in the  $dt(l)$  sampling between events 0 and 4 as presented in section 4.2, examine the cross-correlograms generated at each receiver:

Note that although in this example, correlation coefficients ranging from 0.39 to 0.75 indicate generally poor waveform similarity, the following conclusions apply even for coefficients of 0.8 and up, reportedly presumed to be sufficient coefficients for accurate traveltime measurement. Maximum coefficient values, corresponding to plotted  $dt(l)$  lag values, are circled on each correlogram.

The correlogram for receiver 0 has a single, well defined extremum, which would be expected under ideal circumstances. However, note that correlograms 2, 3, and 6 have multiple local extrema, which are within more than 85% of the global maximum coefficient value. This misalignment, occasionally referred to as cycle skipping, is a consequence of the complex coda which follow the first arrivals of event 4, since there are in effect multiple lags at which the single large arrival measured for event 0 align in time with deflections recorded for event 4, resulting in large integral values for the cross-correlation function. The result is a lag ambiguity, since we cannot discern



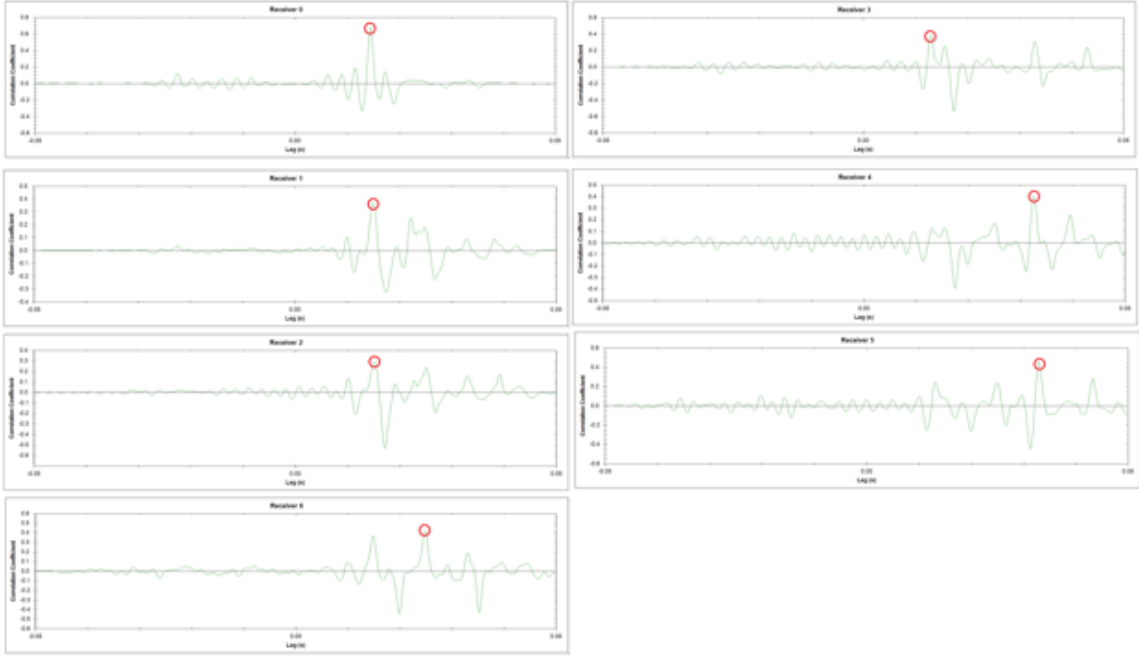


Figure B.1: Correlograms generated at each receiver for events 0 and 4

with certainty from the correlograms alone which local maxima correspond to the true relative arrival time. The ambiguity is further compounded by the consideration that the relative polarities of the arrival waveforms are unknown since, as discussed in section A, there is no guarantee that source mechanisms and their alignments will be similar for different events. This effectively doubles the number of candidates for true lags peaks, as negative minima must also be considered.

It appears therefore that the inconsistency among measured lags and the poor agreement between the expected shape of the  $dt(l)$  function and the measured lag values are both results of misalignment by the cross-correlation function. It is difficult to rigorously relate first arrival waveform characteristics to cross-correlation uncertainty beyond general qualitative statements, since the true lag values between events 0 and 4 are unknown, and to date it appears that no analytic relationship

between cross-correlation uncertainty and signal properties has been derived.