

THESIS

A NEW METHOD TO TEST SHEAR WAVE SPLITTING: IMPROVING STATISTICAL  
ASSESSMENT OF SPLITTING PARAMETERS

Submitted by

Ana Corbalán Castejón

Department of Geosciences

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2016

Master's Committee:

Advisor: Derek Schutt

Jay Breidt  
Richard Aster  
Sven Egenhoff

Copyright by Ana Corbalán Castejón 2016

All Rights Reserved

## ABSTRACT

### A NEW METHOD TO TEST SHEAR WAVE SPLITTING: IMPROVING STATISTICAL ASSESSMENT OF SPLITTING PARAMETERS

Shear wave splitting has proved to be a very useful technique to probe for seismic anisotropy in the earth's interior, and measurements of seismic anisotropy are perhaps the best way to constrain the strain history of the lithosphere and asthenosphere. However, existent methods of shear wave splitting analysis do not estimate uncertainty correctly, and do not allow for careful statistical modeling of anisotropy and uncertainty in complex scenarios. Consequently, the interpretation of shear wave splitting measurements has an undesirable subjective component. This study illustrates a new method to characterize shear wave splitting and the associated uncertainty based on the cross-convolution method [*Menke and Levin, 2003*]. This new method has been tested on synthetic data and benchmarked with data from the Pasadena, California seismic station (PAS). Synthetic tests show that the method can successfully obtain the splitting parameters from observed split shear waves. PAS results are very reasonable and consistent with previous studies [*Liu et al., 1995; Özalaybey and Savage, 1995; Polet and Kanamori, 2002*]. As presented, the *Menke and Levin* [2003] method does not explicitly model the errors. Our method works on noisy data without any particular need for processing, it fully accounts for correlation structures on the noise, and it models the errors with a proper bootstrapping approach. Hence, the method presented here casts the analysis of shear wave splitting into a more formal statistical context, allowing for formal hypothesis testing and more nuanced interpretation of seismic anisotropy results.

## ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor, Derek Schutt, for his encouragement, friendship, guidance and support throughout this entire project. His motivation and enthusiasm about this challenging project made the successful completion of this project possible. I would also like to thank Richard Aster, Sven Egenhoff, and Jay Breidt for kindly serving on my committee. Special thanks to Jay Breidt for his patience as he guided me through the statistical challenges we have faced during the development of this project, for his availability and his willingness to help and to share his knowledge. I would like to thank my NR340 office colleagues, former and current, for sharing their knowledge, support and friendship along the way. Lastly, I would like to thank Javier Marquina for his transatlantic support and motivation; to Juan Corbalán, Kika Castejón and Reyes Corbalán for their encouragement; and to José Castejón, Keren Ellis and Freyja Ellis-Castejón, who made this possible in the first place. Funding for this project was provided through NSF's Division for Earth Science EAR-1252133.

## TABLE OF CONTENTS

ABSTRACT.....	ii
ACKNOWLEDGMENTS.....	iii
TABLE OF CONTENTS.....	iv
LIST OF TABLES.....	v
LIST OF FIGURES.....	vi
CHAPTER 1.....	1
1.1 Organization of Thesis.....	1
1.2 Introduction.....	1
1.2.1 Seismic Anisotropy.....	2
1.2.2 Shear Wave Splitting.....	4
1.2.3 Quantifying Shear Wave Splitting.....	6
1.3 The Transverse Component Minimization Method.....	7
1.4 The Cross-Convolution Method.....	10
1.5 Motivation.....	12
1.6 Figures.....	15
REFERENCES.....	23
CHAPTER 2.....	27
2.1 Introduction.....	27
2.2 Methodology.....	29
2.2.1 Calculation of Standard Errors.....	36
2.3 Results.....	37
2.3.1 Test on Synthetic Data Set.....	37
2.3.2 Application to Pasadena Seismic Station.....	37
2.4 Conclusions.....	38
2.5 Figures.....	40
2.6 Tables.....	46
REFERENCES.....	47
CHAPTER 3.....	50
3.1 Future Work Recommendations.....	50
3.2 Data Retrieval Procedure.....	51
3.3 Data Processing Stream.....	54
3.4 Using the R Code.....	55
3.5 Location of Code and Notes.....	56
REFERENCES.....	57
APPENDICES.....	58
APPENDIX A.....	59
APPENDIX B.....	64
APPENDIX C.....	73

## LIST OF TABLES

Table 2.1: SKS wave splitting results for the PAS station from previous studies.....	44
--	----

## LIST OF FIGURES

Figure 1.1: A natural example of an olivine lattice preferred orientation texture seen in a dunite's photomicrograph.....	15
Figure 1.2: Plot showing the different types of deformation fabrics of olivine as a function of stress and water content.....	16
Figure 1.3: Example of a split shear wave.....	17
Figure 1.4: A depiction of an SKS wave (A) and of SKS, PKS and SKKS (B).....	18
Figure 1.5: Example of a split SKS wave observed at the Flin Flon station.....	19
Figure 1.6: Plots showing the $\pi/2$ periodicity that is often observed in plots of split time versus incoming polarization, and of fast axis versus incoming polarization, when two anisotropic layers are present.....	20
Figure 1.7: Example of a shear wave splitting measurement of an SKS phase using SplitLab....	21
Figure 1.8: Example of a shear wave splitting measurement of an SKS phase using the cross-convolution method.....	22
Figure 2.1: Examples of the true signal estimation used for the bootstrapping approach to calculate the standard errors.....	40
Figure 2.2: Results for synthetic tests.....	41
Figure 2.3: Results for the Pasadena station.....	43

## CHAPTER 1

### **1.1 Organization of Thesis**

This thesis is organized into three chapters. Chapter One consists of a broad introduction to the earth process examined in this study, seismic velocity anisotropy. It also introduces the shear wave splitting technique, which is a common method used to characterize seismic anisotropy, plus a brief overview of relevant previous work on this topic. Chapter Two addresses the research undertaken in pursuit of my master's degree, consisting of writing a code in the R programming language to introduce a new method, based in the *Menke and Levin* [2003] method, to analyze and measure shear wave splitting more accurately. It also includes the results and a subsequent conclusion. This chapter has been written as a technical paper for submittal to a peer-reviewed journal. Chapter Three is targeted towards future students that will continue working on this project, either on further development of the technique or using the code as it is to get new measurements of seismic anisotropy in tectonically complex areas. Hence, this chapter puts forth recommendations for future work and explanations of all of the different procedures to run the code and to obtain results.

### **1.2 Introduction**

This work depicts the development of a new method to test and analyze shear wave splitting in the upper mantle. In order to contextualize the method and its application, this chapter summarizes some essential concepts related to anisotropy, reviews common shear wave splitting methods, and discusses the motivation behind the study.

### 1.2.1 Seismic Anisotropy

Seismic anisotropy is shorthand for seismic velocity anisotropy, the directional dependence of seismic wave velocity. Early studies on this phenomenon include *Turner* [1942] who described the preferred orientation of olivine crystals developed during deformation, and *Verma* [1960] who determined the seismic velocities of the olivine crystallographic axes. These studies suggested that deformed olivine in the mantle could lead to seismic anisotropy. The first identification of seismic anisotropy within the deeper earth was by *Hess* [1964] who observed it in the upper mantle under the ocean basins. Here, it was shown that the oceanic mantle fabric present in the Pacific Plate is perpendicular to the spreading center.

Seismic anisotropy can also be observed in other regions within the earth, such as the inner core [e.g. *Morelli et al.*, 1986], the D'' layer [e.g. *Kendall and Silver*, 1998], the lower mantle [e.g. *Kendall and Silver*, 1996], the transition zone [e.g. *Karato et al.*, 1998], and the crust [e.g. *Crampin*, 1984; *Mainprice and Nicolas*, 1989; *Crampin and Love*, 1991]. This work is motivated by an interest in upper mantle anisotropy, including the mantle lithosphere and the upper part of the asthenosphere, but the methods that are presented here are applicable to other regions where anisotropy occurs.

Several authors [*Hess*, 1964; *Nicolas and Christensen*, 1987; *Silver and Chan*, 1991; *Zhang and Karato*, 1995; *Karato et al.*, 2008] have presented evidence that upper mantle anisotropy is mainly caused by the strain-induced lattice preferred orientation (LPO) of the dominant anisotropic upper mantle minerals, olivine and orthopyroxene. LPO is also called crystalline preferred orientation (CPO) by some researchers, but it will be referred to as LPO herein. LPO textures are formed when a polycrystalline aggregate is subjected to an externally imposed deformation, and a significant number of crystals rotate and preferentially align [e.g. *Di*

*Leo et al.*, 2014] (Figure 1.1). This texture is commonly assumed to be formed by dislocation creep [*Karato et al.*, 2008; *Long and Silver*, 2009], however there are other deformation mechanisms such as: grain boundary sliding, dislocation glide, dynamic crystallization, or diffusion creep, that may play a role in the interaction between deformation and LPO anisotropy [*Di Leo et al.*, 2014]. The influence of diffusion creep on LPO texture development is controversial, and some studies have pointed out that it might even erase preexisting fabric [*Long and Silver*, 2009]. Since olivine is the primary upper mantle mineral constituent [e.g. *Frost*, 2008], it is typically assumed that olivine LPO is the primary cause of observed anisotropy [*Zhang and Karato*, 1995; *Silver*, 1996; *Menke and Levin*, 2003], although LPO of orthopyroxene, or layering of isotropic substances (called shape preferred orientation or SPO) such as partial melt, may play roles in certain circumstances [*Long and Becker*, 2010].

In the last decades, many studies [*Zhang and Karato*, 1995; *Karato et al.*, 1998, 2008; *Kaminski et al.*, 2004; *Kneller et al.*, 2005] have examined how features such as subduction zones, oceanic ridges, and plume-asthenosphere interactions affect the alignment of olivine crystallographic axes. *Karato et al.* [2008] have shown that variations in water, stress, temperature, and pressure, as well as the onset of melting, can lead to as many as five different types of olivine LPO. The influence of water and stress are the dominant factors, with the influence of temperature, pressure and partial melting playing more minor roles, particularly when the redistribution of water occurs due to changes in temperature or pressure, and/or due to partial melting [*Karato et al.*, 2008] (Figure 1.2).

Because the current study examines shear wave splitting methodology, rather than using shear wave splitting to infer the mantle's strain history, this topic will not be explored in depth here. It is sufficient to consider that patterns of seismic anisotropy found in the upper mantle are

primarily linked to deformational processes in the lithosphere and asthenosphere, as LPO is related to the strain history of the material [Zhang and Karato, 1995; Long and Becker, 2010]. Therefore, the study of mantle seismic anisotropy can shed light into the past deformation of the lithosphere and ongoing flow in the asthenosphere [Zhang and Karato, 1995; Silver, 1996; Long and Silver, 2009; Di Leo et al., 2012, 2014; Long and Wirth, 2013].

### 1.2.2 Shear Wave Splitting

One of the most frequently used techniques to characterize mantle anisotropy is shear wave splitting [Long and Becker, 2010]. This occurs when a shear wave encounters an anisotropic layer and the wave is split into two orthogonally polarized waves, one travelling faster than the other--a phenomenon analogous to optical birefringence of light, such as occurs in calcite crystals [e.g. Silver, 1996; Savage, 1999; Long and Silver, 2009; Long and Becker, 2010]. The faster travelling S-wave is polarized along what is called the “fast axis”; the slower wave is polarized orthogonal to this direction [e.g. Silver, 1996; Long and Silver, 2009; Long and Becker, 2010]. The parameters measured to quantify seismic anisotropy are the split time ( $\delta t$ )--the travel time difference between the faster and the slower travelling wave--and the fast axis orientation ( $\phi$ )--the polarization direction of the faster travelling split shear wave (Figure 1.3). The split time values depend on the geometry and thickness of the anisotropic medium that the S-wave is going through [Menke and Levin, 2003; Rasendra et al., 2014], and the fast axis values contain information about the fabric or LPO texture present in the anisotropic medium [Rasendra et al., 2014].

The seismic waves used to analyze shear wave splitting are often teleseismic core-refracted seismic waves, such as the SKS, SKKS, and PKS phases (hereafter grouped by \*KS) (Figure 1.4).

They are called teleseismic waves because they travel more than 1,000 km from the source to the receiver [*Liu and Gao, 2013*], and they are called core-refracted waves because their ray paths traverse the earth's crust, mantle and outer core. The \*KS have similar ray paths, but the SKKS bounces once internally within the outer core (Figure 1.4). Typically, the events used to analyze mantle seismic anisotropy have epicentral distances ranging from 85° to 125°, moment magnitudes ( $M_w$ ) above 5.5-6, and a signal-to-noise ratio (SNR) greater than 4 [*Liu and Gao, 2013*]. \*KS waves are particularly useful for measuring shear wave splitting for three reasons.

First, the conversion from a shear wave to a compressional wave at the core mantle boundary (CMB) removes the effects of anisotropy encountered along the source side of the ray path, allowing one to infer that any anisotropy measured must be occurring along the receiver side of the ray path [e.g. *Silver and Chan, 1991; Long and Silver, 2009; Long and Becker, 2010*].

Second, the \*KS phases are radially polarized as they emerge from the core [e.g. *Silver and Chan, 1991; Silver, 1996; Long and Silver, 2009*]. Knowledge of the polarization direction of the wave before anisotropy is encountered allows the anisotropy to be more tightly constrained (assuming the ray does not deviate from the great circle path due to velocity heterogeneities) [*Long and Silver, 2009*].

Third, \*KS phases are well suited for measuring shear wave splitting because the receiver leg of the ray path propagates through the mantle in a nearly vertical manner. This simplifies the mathematical analysis and provides good lateral resolution of anisotropy. This however makes it difficult to resolve the depth at which any measured anisotropy might be occurring [*Babuska and Cara, 1991; Long and Silver, 2009*]. Even knowing that the observed shear wave splitting measurements correspond to anisotropy in the receiver side and above the CMB, it is difficult to

determine the exact depth at which the anisotropy is found. In most seismological studies, SKS splitting is interpreted in terms of upper mantle anisotropy, ignoring the potential contribution from the lower mantle, and assuming that the effects of crustal anisotropy are much smaller than those in the upper mantle [e.g. *Long and Silver, 2009*].

### 1.2.3 Quantifying Shear Wave Splitting

In the presence of a single layer of horizontally oriented anisotropy, the effects of splitting on a vertically incident wave  $s(t)$ , as observed on the radial  $R(t)$  and transverse  $T(t)$  channels will be:

$$R(t) = s(t)\cos^2\phi + s(t - \delta t)\sin^2\phi, \quad (\text{Eqn. 1.2.2.1})$$

$$T(t) = [(s(t) - s(t - \delta t))/2]\sin 2\phi, \quad (\text{Eqn. 1.2.2.2})$$

where the splitting parameters  $\phi$  and  $\delta t$  are, respectively, the angle between the radial direction and the fast axis, and the split or delay time [e.g. *Stein and Wysession, 2012*]. The  $T(t)$  term is 0 in two cases, when there is no anisotropy and the split time is 0, or when  $\phi$  is  $0^\circ$  or  $90^\circ$ . These cases are termed “null splits”. In contrast, the observation of energy on the transverse channel is typically considered to be indicative of the presence of anisotropy, although dipping layers, ray bending off of the great circle path, or small-scale heterogeneity can also produce some T energy.

Multilayered shear wave splitting has been recognized in several regions [e.g. *Silver and Savage, 1994; Wolfe and Silver, 1998; Levin et al., 1999; Bokermann et al., 2000; Hartog and Schwartz, 2001*], and it is broadly accepted by the seismological community that a single horizontal layer model is not always enough to explain anisotropy measurements, which makes sense if the lithosphere and asthenosphere have different strain histories. A high variability of splitting parameters between nearby stations, a variation of splitting parameters for a single

station for events from different back azimuths, or a considerable number of null splits within a station or within a dataset confined to a particular area, are indications of a complex anisotropic structure [e.g. *Schutt and Humphreys*, 2001]. This complexity can be explained in terms of laterally homogeneous structure such as layered anisotropy or dipping fast axes beneath the station or in terms of lateral heterogeneities (such as an anomalous lithospheric block or a complicated symmetry system) [*Rasendra et al.*, 2014]. Generally, layered anisotropy or dipping fast axes are associated with a  $\pi/2$  back azimuthal periodicity of the splitting parameters (Figure 1.6) [*Silver and Savage*, 1994; *Savage*, 1999], while lateral heterogeneities are observed in a unique back azimuthal window [*Savage*, 1999; *Rasendra et al.*, 2014].

There are several techniques used to estimate shear wave splitting parameters; two of these are most commonly used in the seismology community. These are the *Silver and Chan* [1991] method, based on the minimization of the energy on the transverse component, and the *Menke and Levin* [2003], or cross-convolution method.

### **1.3 The Transverse Component Minimization Method**

This method was introduced by *Silver and Chan* [1991], and it is probably the most commonly used means for studying upper mantle shear wave splitting. This technique (hereafter referred to as SC) seeks to identify the pair of splitting parameters ( $\phi$ ,  $\delta t$ ) that best minimizes the amount of energy present on the transverse component after correcting for an assumed single layer of anisotropy [*Silver and Chan*, 1991; *Long and Silver*, 2009]. To do so, a grid search is performed over all plausible splitting parameters. In each step of the grid search, the observed data is rotated and time-shifted based on the test fast axis angle and the split time, respectively, and the residual energy on the transverse component is calculated. The splitting parameter pair that is associated with the minimum value of corrected transverse energy is considered to be the

best estimate of anisotropy. This residual energy provides a measure of the variance of the noise process, and it is used as the basis for calculating a confidence region for the splitting parameters [Silver and Chan, 1991]. The pair that best minimizes the amount of energy on the transverse component also best linearizes the corrected particle motion (Figure 1.7).

Silver and Chan [1991] base their error estimation on the minimization of the second eigenvalue,  $\lambda_2^{min}$  of the covariance matrix of the seismogram particle motion (the radial seismogram plotted against the transverse seismogram) for the pairs  $(\phi, \delta t)$  tested via the grid search. The pair  $(\hat{\phi}, \hat{\delta t})$  is the one that best minimizes the amount of energy on the transverse component, and hence the second eigenvalue. The value of  $\lambda_2^{min}$  is supposed to represent a measurement of the variance of the noise process, and it is the foundation for calculating the confidence region. Silver and Chan [1991] suggest that for an n-point discrete time series,  $\lambda_2^{min}$  is the sum of squares of the integrated noise in the signal, which they assumed to be a filtered Gaussian process, and therefore approximately  $\chi^2$  (chi-square) distributed. Based on this assumption, an  $F$ -test is used to calculate confidence intervals. The degrees of freedom ( $\nu$ ) are calculated by:

$$\nu = 2 \left( \frac{2E^2}{E_4} - 1 \right)$$

$$E_4 = \sum_{n=1}^{N-1} |f_n|^4 |g_n|^4 + \frac{1}{3} (|f_0|^4 |g_0|^4 + |f_N|^4 |g_N|^4),$$

$$E \simeq \sum_{n=1}^{N-1} |f_n|^2 |g_n|^2 + \frac{1}{2} (|f_0|^2 |g_0|^2 + |f_N|^2 |g_N|^2), \quad (\text{Eqn. 1.3.1})$$

where  $g(t)$  is a Gaussian noise process uncorrelated to the noise present in the data,  $f(t)$  is a filter,  $E$  is the confidence region determined from the sum of squares function, and  $E_4$  is the confidence region in the frequency domain. Note that SC assumes the noise process present in the data to be the convolution of  $g(t)$  and  $f(t)$ . Based on their observations, SC concludes that

the degrees of freedom are about one degree of freedom for each second of record, independent of the sampling interval and the network. They assess the standard errors, the  $1\sigma$  uncertainties, by calculating a quarter of the width and length of the confidence region.

Most often, this method is applied to individual seismograms, and for multiple observations at a given station, best estimates of the splitting parameters are derived from weighted averaged of the individual splitting parameters. Some authors stack the grids of transverse energy, based on the assumption that the transverse energy behaves as a chi-squared variable. With this assumption, one can obtain more reliable average values and more fully consider the covariance between  $\phi$  and  $\delta t$  [Wolfe and Silver, 1998; Schutt and Humphreys, 2001]. This stacking also assumes a single layer of horizontal anisotropy.

Several studies have shown that the SC method underestimates the uncertainty in splitting parameters [Sandvol and Hearn, 1994; Levin et al., 1999; Savage, 1999; Menke and Levin, 2003; Walsh et al., 2013]. Sandvol and Hearn [1994] suggest that the main misstep in the SC error estimation method is the assumption that the sum of squares amplitude for a n-point time series behaves as a  $\chi^2$  variable with the number of degrees of freedom equal to approximately one per second. Walsh et al. [2013] present a thorough statistical analysis of the SC method in which they describe the main inaccuracies in the SC error estimation method as follows: (i) it is not always possible to assume that the background noise in the seismogram is a filtered Gaussian process, and therefore it does not always follow a normal distribution; (ii) the estimation of the number of degrees of freedom is positive biased, so SC overestimates the degrees of freedom by a factor of 4/3; and (iii) there is a negative bias in the estimation of the error for the split time which results in a typical underestimation of error.

Some authors have attempted to produce more statistically reliable errors, such as a bootstrapping method to estimate the splitting errors [Sandvol and Hearn, 1994], or as the stacking of the resulting SC misfit surfaces mentioned above. Walsh et al. [2013] argue that the bootstrapping error estimation method is slightly better than the SC method since it avoids making assumptions about the error distribution, and it does not need to calculate the degrees of freedom. However, it assumes that the observations are all independent and identically distributed, which they contend is highly unlikely. The misfit stacking method incorrectly assumes the noise is  $\chi^2$  distributed.

Conversely, several shear wave splitting reviews [Silver, 1996; Savage, 1999; Long and Silver, 2009; Long and Becker, 2010] consider that the SC method obtains the most reliable results for simple anisotropy models. The SC method seems to yield the most robust results for noisy data, and it does not necessarily need a rich back azimuth coverage, although it seems to be the more susceptible to getting unreasonable splitting parameters in the presence of complex anisotropy [Long and Silver, 2009]. A better approach for considering complex anisotropy is the cross-convolution method of Menke and Levin [2003].

#### 1.4 The Cross-Convolution Method

Menke and Levin [2003] introduced this approach with the objective of creating a method that can be used on a range of anisotropic systems. The method consists of the minimization of the error associated with the mismatch between two time series that are both a function of the observed seismograms and the predicted seismograms (Figure 1.8) [Long and Silver, 2009]. To do this, the method needs two constructed time series,  $x(t)$  and  $y(t)$ , as follows:

$$x(t) = \tau_m * R, \quad (\text{Eqn. 1.4.1})$$

$$y(t) = \rho_m * T, \quad (\text{Eqn. 1.4.2})$$

where the symbol  $*$  denotes convolution,  $R$  and  $T$  represent the observed radial and tangential components, respectively, and  $\rho_m$  and  $\tau_m$  are the predicted (synthetic) radial and tangential impulse responses, respectively, for a hypothetical Earth model,  $m$ . The best-fitting Earth model is considered one that minimizes the normalized squared difference between  $x(t)$  and  $y(t)$ . This approach has several advantages over the more traditional techniques: it can be generalized to multilayered anisotropic media, it is not necessary to assume a vertically incident ray, and it can also discern whether a data set is consistent with simple or complex anisotropic models by comparing the one layer model fit to the two layer model fit [Menke and Levin, 2003; Yuan et al., 2008; Long and Silver, 2009]. In general, this method can work on any anisotropic system as long as  $\rho_m$  and  $\tau_m$  have similar paths.

Menke and Levin [2003] define the impulse response of a normally incident S-wave in one anisotropic layer as follows:

$$\begin{aligned}
 \rho_m &= \rho_1 \delta(t) + \rho_2 \delta(t - dt) \\
 \tau_m &= \tau_1 \delta(t) + \tau_2 \delta(t - dt) \\
 \rho_1 &= \cos^2(\phi - \theta) \\
 \rho_2 &= \sin^2(\phi - \theta) \\
 \tau_1 = \tau_2 &= \cos(\phi - \theta) \sin(\phi - \theta), \quad (\text{Eqn. 1.4.3})
 \end{aligned}$$

where  $dt$  is the split or delay time,  $\phi$  is the fast axis angle, and  $\theta$  is the back azimuth angle. The formulas used in (Eqn. 1.4.3) to calculate the impulse response for a radially polarized phase (e.g. SKS) assuming azimuthal anisotropy are the formulas used in the method presented in this document. Menke and Levin [2003] note that the application of these formulas to an SKS phase is only an approximation, since their propagation path is not exactly vertical. However, a nearly vertical path to an SKS phase is often assumed in the studies of shear wave splitting in order to

simplify the computation of the splitting parameters. They tested their models on synthetic data, and they obtained solutions accounting for about 99% of the error, thus validating the use of the one layer impulse response estimator (Eqn. 1.4.3), while the residual 1% is mainly associated with crustal reverberations that are not modeled [Menke and Levin, 2003]. Note that tests of the Cross-Convolution Method on synthetic and real data yields similar results to traditional methods in the presence of simple anisotropy [Menke and Levin, 2003; Long and Silver, 2009].

### **1.5 Motivation**

The pursuance of this project is based on the idea that a new procedure to test shear wave splitting is needed in order to obtain more accurate and reliable measurements of upper mantle seismic anisotropy. The results obtained with the code developed as a part of this thesis will yield a better statistical characterization of shear wave splitting measurements, and therefore, a better understanding and characterization of past lithospheric deformation and ongoing asthenospheric flow. Three examples where the robustness and accuracy of shear wave splitting measurements will be very valuable are as follows:

(i) Upper mantle seismic anisotropy interpretation in tectonically complex areas like the High Lava Plains (HLP) in eastern Oregon. Using data from a dense temporary broadband array a previous study detected a significant variability in the observed anisotropy among nearby stations located in northern Oregon, close to the Wallowa Mountain region [Long *et al.*, 2009]. With this new approach we can potentially test and discern if this variability is statistically significant, or an outcome of mischaracterized uncertainty. A better characterization of uncertainty would also allow for formal hypothesis tests such as whether additional anisotropic complexity, such as multiple or dipping layers of anisotropy, is needed to explain observations at a given seismometer.

(ii) Another relevant instance for the development of the new method comes from the debate about a common shear wave splitting pattern observed in subduction zones. Here, shear wave splitting observations often show trench-parallel fast axes near the trench, and trench-perpendicular fast axes further away. The pattern is conjectured to imply trench-parallel flow close to the trench and trench-normal flow away from the trench [*Russo and Silver, 1994; Smith et al., 2001; Nakajima and Hasegawa, 2004; Christensen and Abers, 2010*], assuming that the anisotropic fabric existent in the mantle is the A-type olivine. Note that the A-type olivine is often expected to be present in the upper mantle [*Savage, 1999; Becker et al., 2003*]. Nevertheless, other authors [*Kneller et al., 2005; Karato et al., 2008*] suggest that the trench-parallel shear wave splitting detected is due to an unusual olivine fabric, the B-type fabric. Several authors have expressed their concerns about assumptions regarding the type of Olivine present and have pointed out that caution needs to be exercised when inferring mantle flow direction from shear wave splitting in complex geodynamic zones [*Ribe, 1989; Dawson and Wenk, 2000; Jadamec and Billen, 2010, 2012; Di Leo et al., 2012, 2014*].

It may also be possible that the observed complexity can be produced by the presence of two different anisotropic layers in the medium. Recently, several regions where two anisotropic layers exist have been identified [*Silver and Savage, 1994; Wolfe and Silver, 1998; Levin et al., 1999; Bokermann et al., 2000; Hartog and Schwartz, 2001*]. The method presented in this thesis is meant to lay the groundwork to develop robust tests for multilayered seismic anisotropy, so it will potentially shed light on this complicated topic, for instance, allowing for formal tests of one versus two layer anisotropy.

(iii) Lastly, another issue frequently debated is the determination of the contribution of lithospheric deformation or of asthenosphere convection to the upper mantle seismic anisotropy

observed beneath the continents [*Vinnik et al.*, 1992; *Babuška et al.*, 1993; *Long et al.*, 2009]. There are two diverging positions in this debate--whether the observed anisotropy in the upper mantle is due to simple asthenospheric flow [*Vinnik et al.*, 1992], or whether the observed anisotropy is due to vertically coherent deformation of the crust and subcontinental mantle [*Silver*, 1996]. In the multilayered seismic anisotropy case, the top layer is generally interpreted to be due to vertically coherent deformation of the lithosphere, while the lower layer anisotropy may be due to simple asthenospheric flow in either the asthenosphere or the lower lithosphere [*Walker et al.*, 2001]. The method presented in this document will be a first step towards gaining a better understanding of the contribution of the lithosphere and the asthenosphere to observed shear wave splitting measurements.

## 1.6 Figures

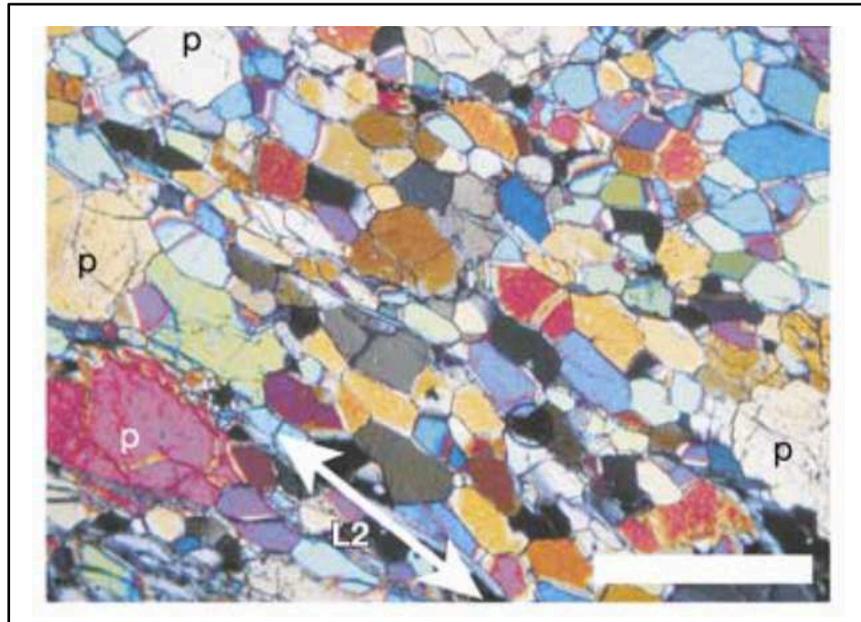


Figure 1.1: A natural example of an olivine lattice preferred orientation texture seen in a photomicrograph of a dunite from the Higashi-akaishi ultramafic body in the Sanbagawa metamorphic belt, southwest Japan. Coarse porphyroclasts (p) and neoblasts of olivine can be observed. The grain shape and orientation of the olivine crystals define the L2 lineation. The length of the white scale bar is 1.0 mm. After *Mizukami et al.* [2004].

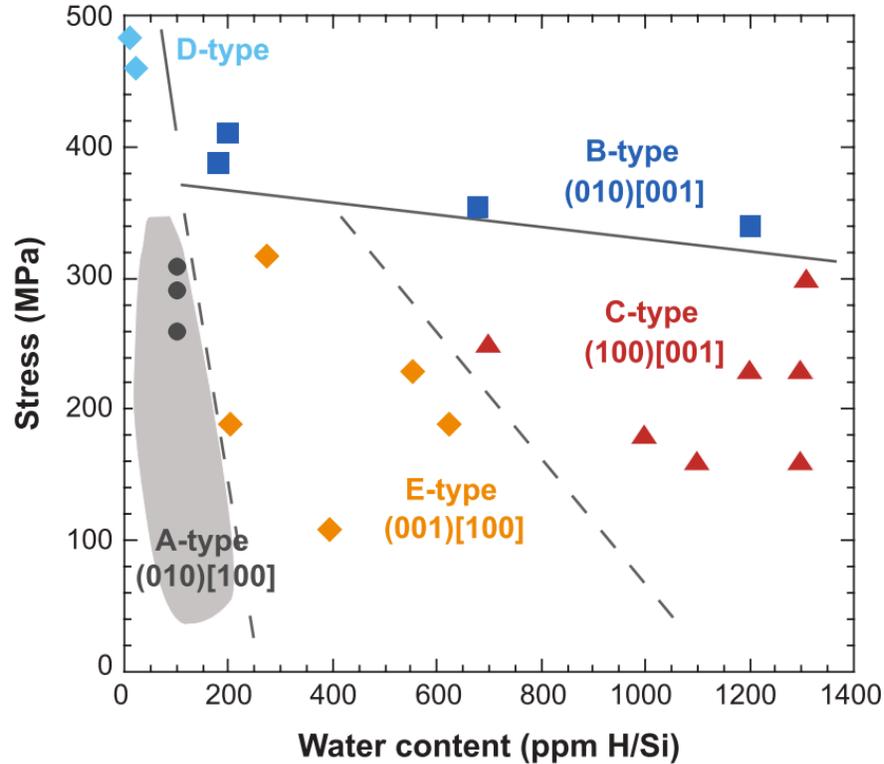


Figure 1.2: Plot from *Karato et al.* [2008] that shows the different types of deformation fabrics of olivine at high temperatures ( $T \sim 1470\text{-}1570\text{ K}$ ), as a function of stress and water content. A-type olivine lattice preferred orientation is found in water-poor and low stress conditions, while at higher stresses D-type olivine fabric is found. As the water content increases, E-type, C-type and B-type fabrics occur. High stress and water-rich conditions favor the B-type olivine fabric, as has been suggested in several studies for subduction settings [e.g. *Conder and Wiens, 2007; Collings et al., 2013; Long and Wirth, 2013*]. The different types of deformation lead to differing ways the olivine crystallographic axes align in response to strain.

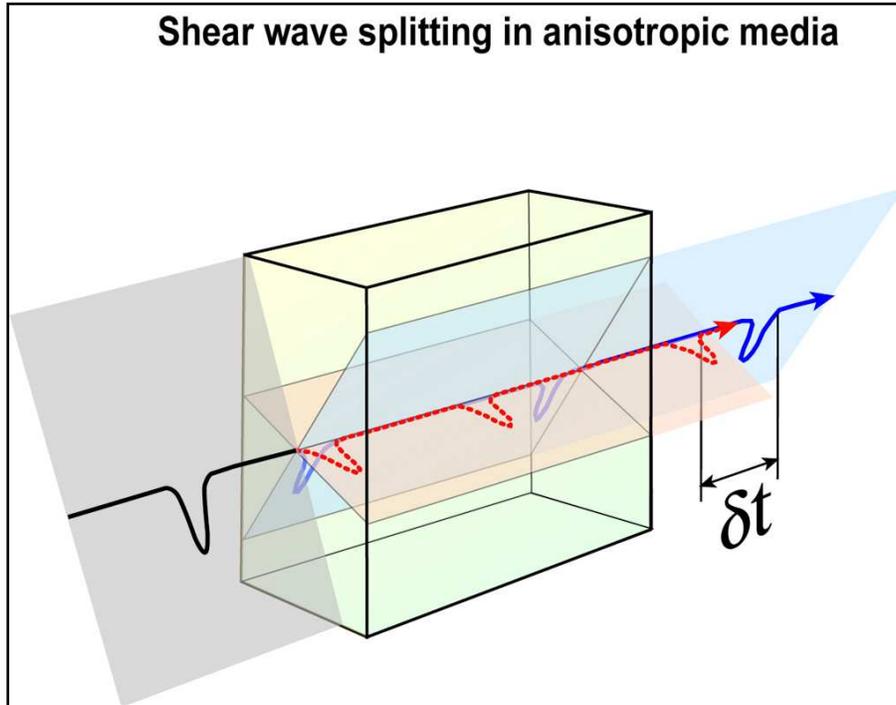


Figure 1.3: Example of a split shear wave. The black line shows a source S-wavelet, which splits into two polarized waves upon encountering the anisotropic medium (the yellow prism). Note that the blue wave travels faster than the red wave. Split shear waves are often described in terms of two parameters, the split or delay time ( $\delta t$ ) and the fast axis orientation ( $\phi$ ), which is the polarization direction of the faster travelling split shear wave, in this figure the blue wave. After *Garnero* [Unpublished, 2012] ([http://garnero.asu.edu/research\\_images/images\\_anisotropy.html](http://garnero.asu.edu/research_images/images_anisotropy.html)).

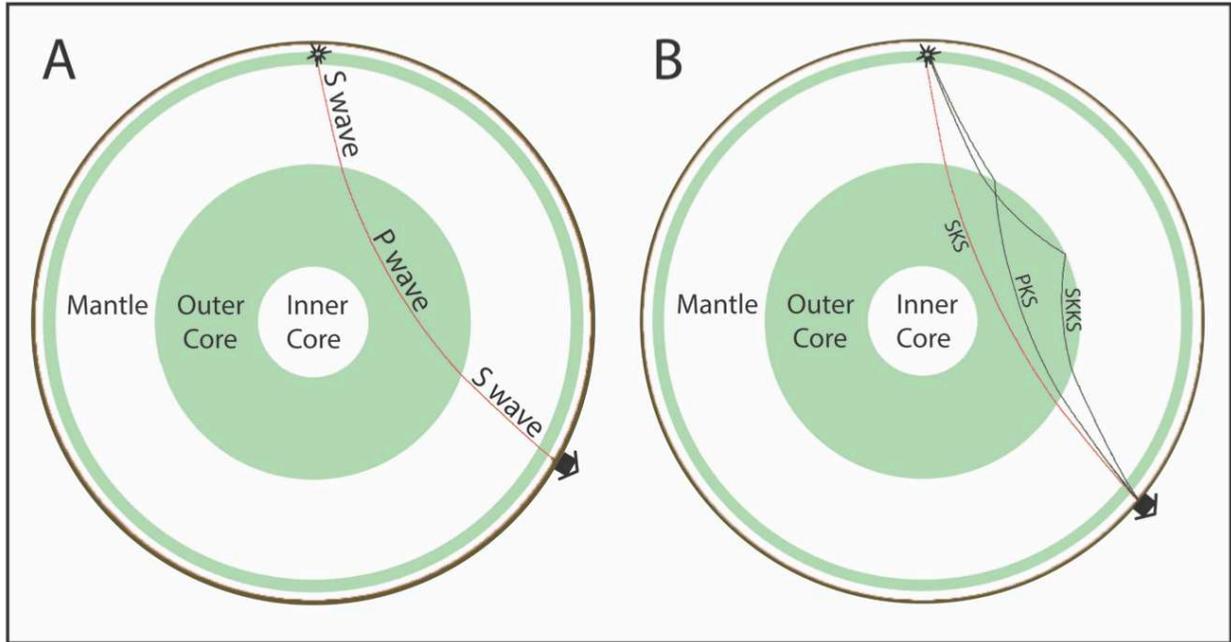


Figure 1.4: A. Representation of an SKS ray, which is characterized by shear waves (S) travelling through the crust and mantle, and by a compressional wave (P) travelling through the earth's outer core. Note that an S wave cannot travel through the outer core due to its liquid state. The conversion from a P wave to an S wave in the core mantle boundary has two major implications to the study of seismic anisotropy in the upper mantle: (i) the observed anisotropy is localized in the receiver side of the ray path; (ii) the S wave emerges from this boundary radially polarized. B. Representation of the SKS, PKS, and SKKS ray paths, termed \*KS phases in this document. After *Crotwell et al.* [1999].

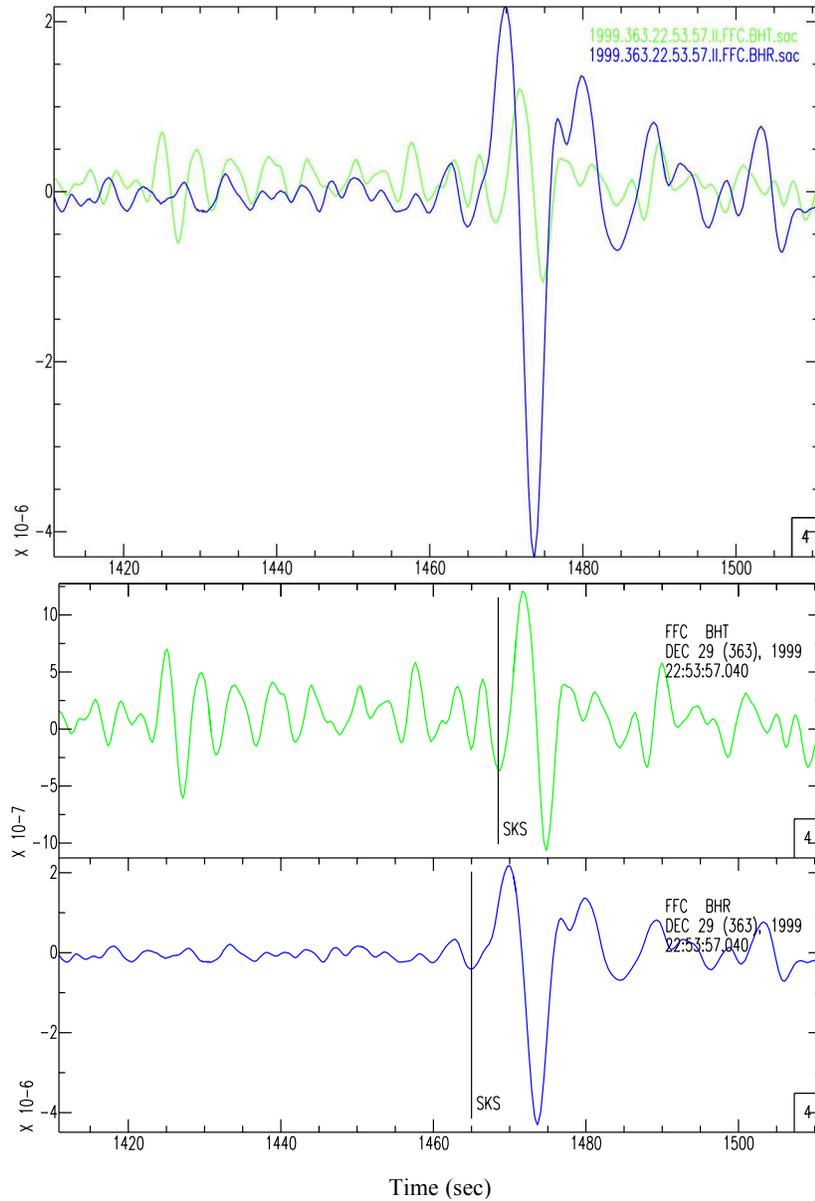


Figure 1.5: Example of a split SKS wave observed at the FFC station, Flin Flon, in Canada, part of the Global Seismographic Network (GSN). The upper figure shows the R component seismogram (blue line) and the T component seismogram (green line) for the seismic event on December 29 1999 at 22:53:57 GMT. The lower figure shows these seismograms in independent windows, with a vertical line indicating the approximate SKS arrival. The presence of a clear signal in the transverse channel (BHT in green) is a strong indicator of anisotropy. Figure obtained with Seismic Analysis Code (SAC) [Goldstein *et al.*, 1993; Goldstein and Snoke, 2005].

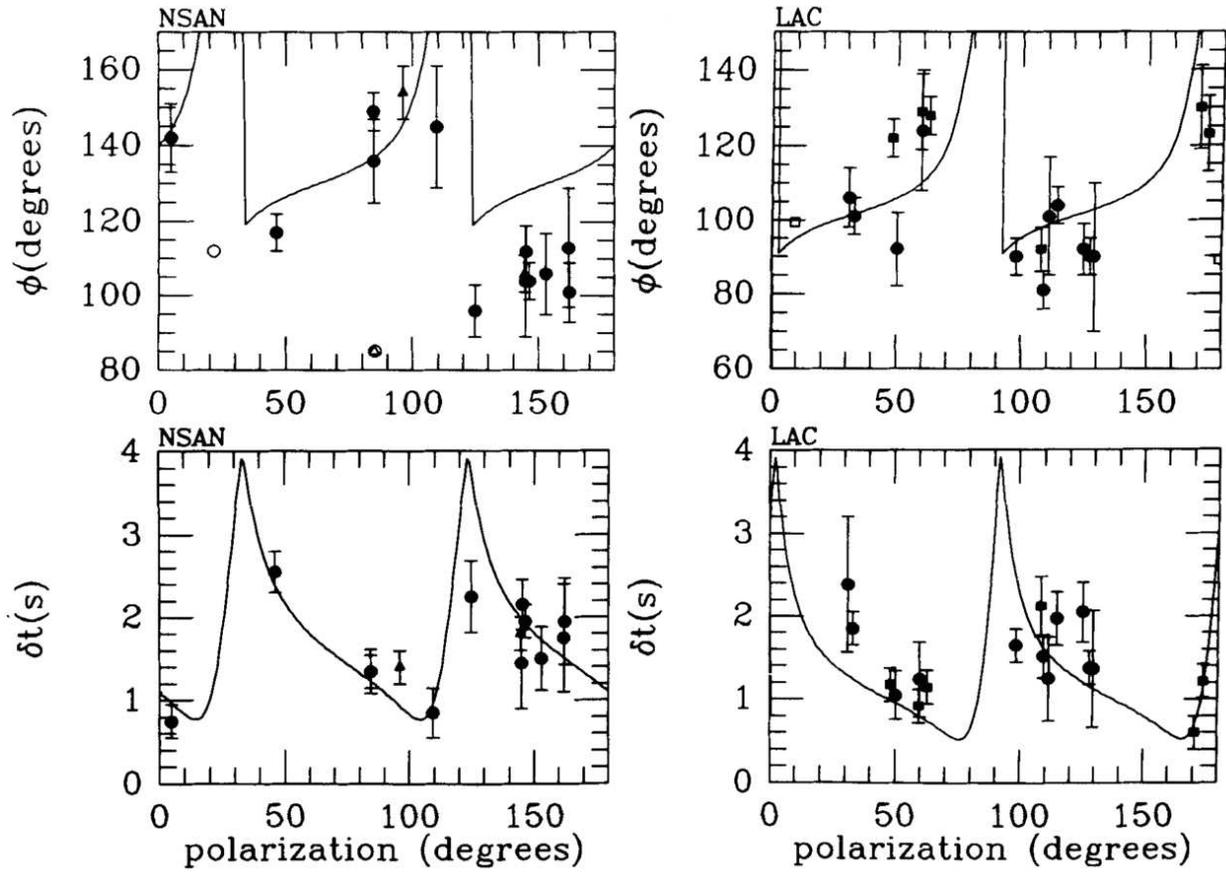


Figure 1.6: Plots with individual measurements of split time ( $\delta t$ ) and fast axis orientation ( $\phi$ ) with their error bars, as a function of the incoming polarization from the NSAN and LAC stations, both located near the San Andreas Fault in California. Circles represent SKS, triangles SKKS, and squares S. Solid symbols represent well-constrained measurements, while open symbols represent null measurements. Note that the four plots exhibit a  $\pi/2$  periodicity as a function of the incoming polarization. This behavior suggests that two anisotropic layers may be present, rather than just one. The black lines in plots are the curves calculated for the double-layer anisotropic parameters in [Silver and Savage, 1994] study. After Silver and Savage [1994].

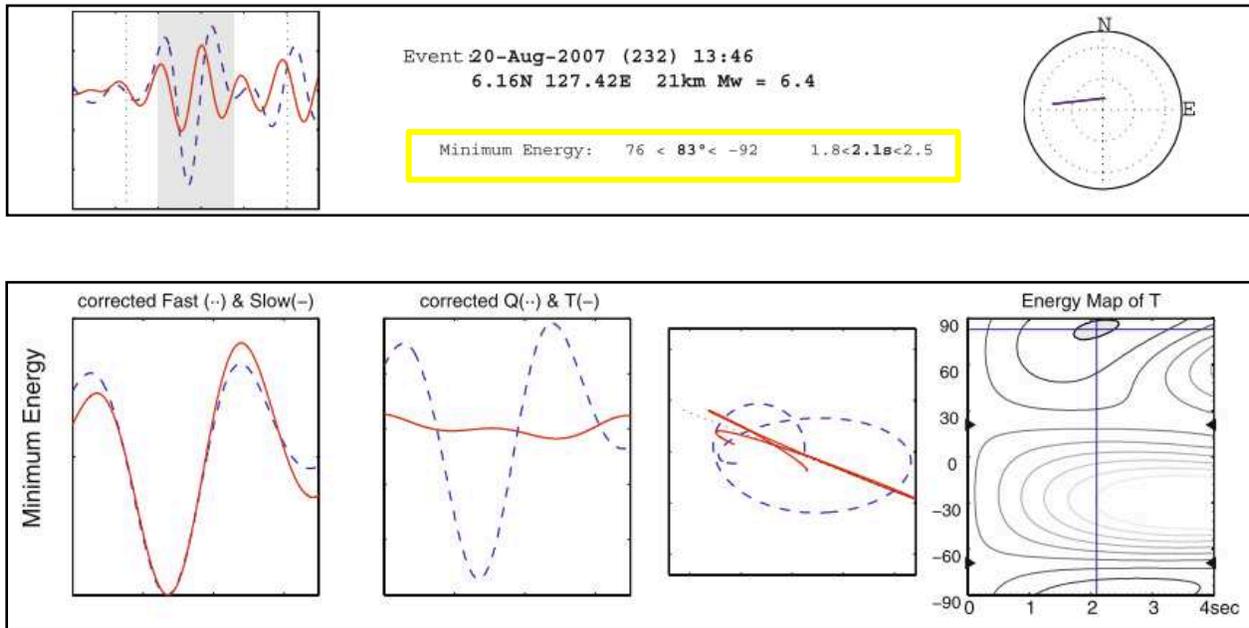


Figure 1.7: Example of a shear wave splitting measurement of an SKS phase using the SplitLab software [Wüstefeld *et al.*, 2008] that implements the transverse component minimization method [Silver and Chan, 1991]. The top panel shows the information about the event used to measure the splitting, and the numeric results (yellow square). The dashed blue line is the radial component, and the solid red line is the transverse component. In the bottom panel, the far left graph shows the corrected fast (dashed blue line) and slow (solid red line) components. The center left graph shows the corrected radial (dashed blue line) and transverse (solid red line) components. The center right graph shows the initial particle motion (dashed blue line) and the corrected particle motion (solid red line). It can be observed that once the effect of splitting is corrected there is no energy on the corrected transverse component, and that the corrected particle motion is linear. The far right graph is the contour plot of the correlation coefficient of split time ( $x$  axis) and fast axis orientation ( $y$  axis). The best-fitting splitting parameters ( $\delta t$ ,  $\phi$ ) are shown where the blue lines intersect, and the 95% confidence region (shaded in gray) is also shown. After Long and Silver [2009].

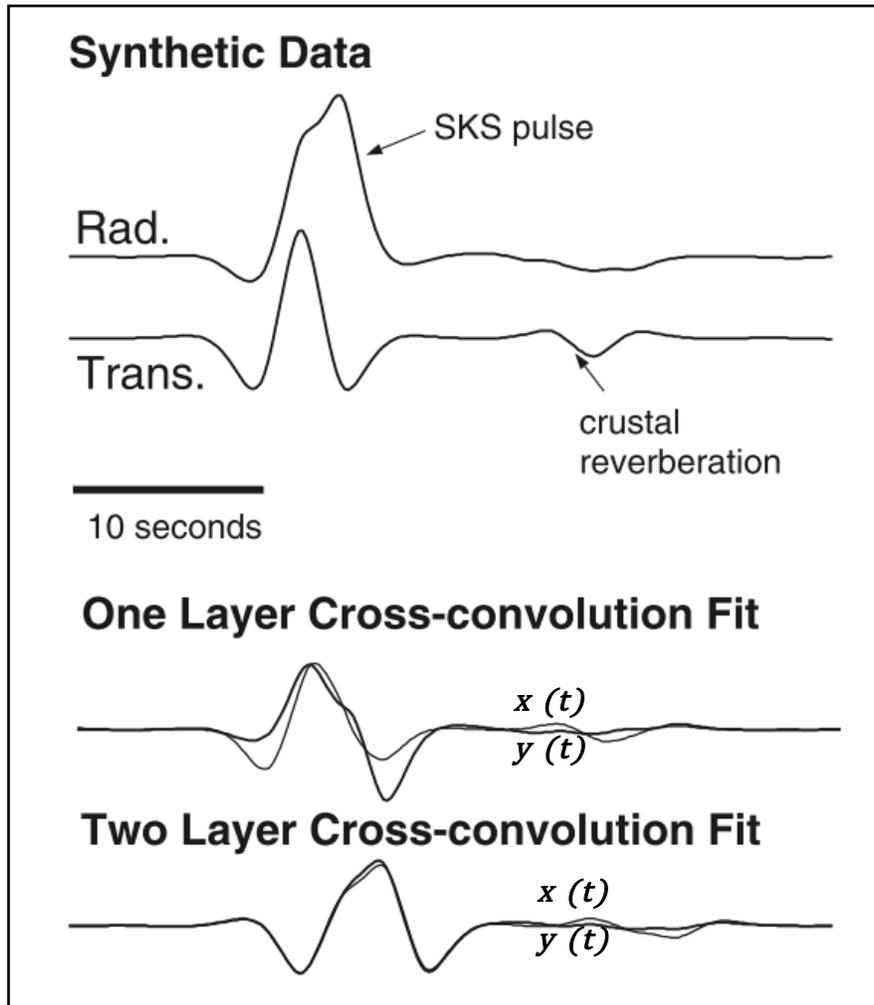


Figure 1.8: Example of shear wave splitting measurement of an SKS phase using the cross-convolution method [Menke and Levin, 2003]. The top traces are the radial (Rad.) and transverse (Trans.) components used to measure the splitting. The bottom traces are the fit obtained by performing the one layer and the two layer cross-convolution method. The  $x(t)$  and  $y(t)$  functions are the radial and transverse traces cross-convolved with the predicted horizontal and vertical impulse response functions for an hypothetical Earth model, respectively (Eqn. 1.4.1 and Eqn. 1.4.2). In this example, the two layer fit is significantly better than the one layer fit, therefore a two anisotropic layer model should be taken into account. After Long and Silver [2009].

## REFERENCES

- Babuška, V., J. Plomerová, and J. Šílený (1993), Models of seismic anisotropy in the deep continental lithosphere, *Phys. Earth Planet. Inter.*, 78(3-4), 167–191, doi:10.1016/0031-9201(93)90154-2, doi:10.1016/0031-9201(93)90154-2.
- Becker, T. W., J. B. Kellogg, G. Ekström, and R. J. O’Connell (2003), Comparison of azimuthal seismic anisotropy from surface waves and finite strain from global mantle-circulation models, *Geophys. J. Int.*, 155(2), 696–714, doi:10.1046/j.1365-246X.2003.02085.x.
- Bokelmann, G. H. R., P. G. Silver, H. R. Bokelmann, and P. G. Silver (2000), Mantle variation within the Canadian Shield: Travel times from the portable broadband Archean-Proterozoic Transect 1989, *J. Geophys. Res. Earth*, 105(B1), 579–605, doi:10.1029/1999jb900387.
- Christensen, D. H., and G. a. Abers (2010), Seismic anisotropy under central Alaska from SKS splitting observations, *J. Geophys. Res.*, 115(B4), 1–12, doi:10.1029/2009JB006712.
- Collings, R., A. Rietbrock, D. Lange, F. Tilmann, S. Nippres, and D. Natawidjaja (2013), Seismic anisotropy in the Sumatra subduction zone, *J. Geophys. Res. Solid Earth*, 118(10), 5372–5390, doi:10.1002/jgrb.50157.
- Conder, J. A., and D. A. Wiens (2007), Rapid mantle flow beneath the Tonga volcanic arc, *Earth Planet. Sci. Lett.*, 264(1-2), 299–307, doi:10.1016/j.epsl.2007.10.014.
- Crampin, S. (1984), Effective anisotropic elastic constants for wave propagation through cracked solids, *Geophys. J. R. Astron. Soc.*, 76, 135–145, doi:10.1111/j.1365-246X.1984.tb05029.x.
- Crampin, S., and J. H. Love (1991), A decade of shear-wave splitting in the Earth’s crust: what does it mean? what we can make of it? and what should we do next?, *Geophys. J. Int.*, 387–407, doi: 10.1111/j.1365-246X.1991.tb01401.x.
- Crotwell, H. P., T. J. Owens, and J. Ritsema (1999), The TauP Toolkit : Flexible Seismic Travel-time and Ray-path Utilities, *Seismol. Res. Lett.*, 70(2), 154–160.
- Dawson, P. R., and H.-R. R. Wenk (2000), Texturing of the upper mantle during convection, *Philos. Mag. A*, 80(3), 573–598, doi:10.1080/014186100250723.
- Frost, D. J. (2008), The Upper Mantle and Transition Zone, *Elements*, 4(3), 171–176, doi:10.2113/GSELEMENTS.4.3.171.
- Goldstein, P., and J. A. Snoke (2005), SAC Availability for the IRIS Community, *Inc. Res. Institutions Seismol. Newsl.*, 7, 6.
- Goldstein, P., D. Dodge, M. Firpo, L. Minner, and J. A. Snoke (1993), SAC2000: Signal Processing and Analysis Tools for Seismologists and Engineers, *Data Manag. Cent. Electron. Newsl.*, 8(1), 1–7.
- Hartog, R., and S. Y. Schwartz (2001), Depth-dependent mantle anisotropy below the San Andreas fault system: Apparent splitting parameters and waveforms, *J. Geophys. Res.*, 106(B3), 4155, doi:10.1029/2000JB900382.
- Hess, H. H. (1964), Seismic Anisotropy of the Uppermost Mantle under Oceans, *Nature*, 203, 629–631, doi:10.1038/203629a0.
- Jadamec, M. a, and M. I. Billen (2010), Reconciling surface plate motions with rapid three-dimensional mantle flow around a slab edge., *Nature*, 465(7296), 338–41, doi:10.1038/nature09053.

- Jadamec, M. a., and M. I. Billen (2012), The role of rheology and slab shape on rapid mantle flow: Three-dimensional numerical models of the Alaska slab edge, *J. Geophys. Res. Solid Earth*, *117*(2), 1–20, doi:10.1029/2011JB008563.
- Kaminski, É., N. M. Ribe, and J. T. Browaey (2004), D-Rex, a program for calculation of seismic anisotropy due to crystal lattice preferred orientation in the convective upper mantle, *Geophys. J. Int.*, *158*, 744–752, doi:10.1111/j.1365-246X.2004.02308.x.
- Karato, S., H. Jung, I. Katayama, and P. Skemer (2008), Geodynamic Significance of Seismic Anisotropy of the Upper Mantle: New Insights from Laboratory Studies, *Annu. Rev. Earth Planet. Sci.*, *36*(1), 59–95, doi:10.1146/annurev.earth.36.031207.124120.
- Karato, S.-I., C. Dupas-Bruzek, and D. C. Rubie (1998), Plastic deformation of silicate spinel under the transition-zone conditions of the Earth's mantle, *Nature*, *395*(September), 266–269, doi:10.1038/26206.
- Kendall, J.-M., and P. G. Silver (1996), Constraints from seismic anisotropy on the nature of the lowermost mantle, *Nature*, *381*(6581), 409–412, doi:10.1038/381409a0.
- Kneller, E. A., P. E. van Keken, S. Karato, and J. Park (2005), B-type olivine fabric in the mantle wedge: Insights from high-resolution non-Newtonian subduction zone models, *Earth Planet. Sci. Lett.*, *237*(3-4), 781–797, doi:10.1016/j.epsl.2005.06.049.
- Di Leo, J. F., J. Wookey, J. O. S. Hammond, J. M. Kendall, S. Kaneshima, H. Inoue, T. Yamashina, and P. Harjadi (2012), Mantle flow in regions of complex tectonics: Insights from Indonesia, *Geochemistry, Geophys. Geosystems*, *13*(12), 1–20, doi:10.1029/2012GC004417.
- Di Leo, J. F., a. M. Walker, Z. H. Li, J. Wookey, N. M. Ribe, J. M. Kendall, and a. Tommasi (2014), Development of texture and seismic anisotropy during the onset of subduction, *Geochemistry, Geophys. Geosystems*, *15*(1), 192–212, doi:10.1002/2013GC005032.
- Levin, V., W. Menke, and J. Park (1999), Shear wave splitting in the Appalachians and the Urals: A case for multilayered anisotropy, *J. Geophys. Res.*, *104*(B8), 17975–17993, doi:10.1029/1999JB900168.
- Liu, H., P. M. Davis, and S. Gao (1995), SKS splitting beneath southern California, *Geophys. Res. Lett.*, *22*(7), 767, doi:10.1029/95GL00487.
- Liu, K. H., and S. S. Gao (2013), Making reliable shear-wave splitting measurements, *Bull. Seismol. Soc. Am.*, *103*(5), 2680–2693, doi:10.1785/0120120355.
- Long, M. D., and T. W. Becker (2010), Mantle dynamics and seismic anisotropy, *Earth Planet. Sci. Lett.*, *297*(3-4), 341–354, doi:10.1016/j.epsl.2010.06.036.
- Long, M. D., and P. G. Silver (2009), Shear Wave Splitting and Mantle Anisotropy: Measurements, Interpretations, and New Directions, *Surv. Geophys.*, *30*(4-5), 407–461, doi:10.1007/s10712-009-9075-1.
- Long, M. D., and E. A. Wirth (2013), Mantle flow in subduction systems: The mantle wedge flow field and implications for wedge processes, *J. Geophys. Res. Solid Earth*, *118*(2), 583–606, doi:10.1002/jgrb.50063.
- Long, M. D., H. Gao, A. Klaus, L. S. Wagner, M. J. Fouch, D. E. James, and E. Humphreys (2009), Shear wave splitting and the pattern of mantle flow beneath eastern Oregon, *Earth Planet. Sci. Lett.*, *288*(3-4), 359–369, doi:10.1016/j.epsl.2009.09.039.
- Mainprice, D., and A. Nicolas (1989), Development of shape and lattice preferred orientations: application to the seismic anisotropy of the lower crust, *J. Struct. Geol.*, *11*(1), 175–189, doi:10.1016/0191-8141(89)90042-4.

- Menke, W., and V. Levin (2003), The cross-convolution method for interpreting SKS splitting observations, with application to one and two-layer anisotropic earth models, *Geophys. J. Int.*, *154*(2), 379–392, doi:10.1046/j.1365-246X.2003.01937.x.
- Mizukami, T., S. R. Wallis, and J. Yamamoto (2004), Natural examples of olivine lattice preferred orientation patterns with a flow-normal a-axis maximum., *Nature*, *427*(6973), 432–436, doi:10.1038/nature02179.
- Morelli, A., A. M. Dziewonski, and J. H. Woodhouse (1986), Anisotropy of the inner core inferred from PKIKP travel times, *Geophys. Res. Lett.*, *13*(13), 1545–1548, doi:10.1029/GL013i013p01545.
- Nakajima, J., and A. Hasegawa (2004), Shear-wave polarization anisotropy and subduction-induced flow in the mantle wedge of northeastern Japan, *Earth Planet. Sci. Lett.*, *225*(3-4), 365–377, doi:10.1016/j.epsl.2004.06.011.
- Nicolas, A., and N. I. Christensen (1987), Formation of anisotropy in upper mantle peridotites - a review, *Am. Geophys. Union*, *16*, 111–123, doi:10.1029/GD016p0111.
- Özalaybey, S., and M. K. Savage (1995), Shear-wave splitting beneath western United States in relation to plate tectonics, *J. Geophys. Res.*, *100*(B9), 18135, doi:10.1029/95JB00715.
- Polet, J., and H. Kanamori (2002), Anisotropy beneath California: shear wave splitting measurements using a dense broadband array, *Geophys. J. Int.*, *149*(2), 313–327, doi:10.1046/j.1365-246X.2002.01630.x.
- Rasendra, N., M. Bonnin, S. Mazzotti, and C. Tiberi (2014), Crustal and upper-mantle anisotropy related to fossilized transpression fabric along the denali fault, Northern Canadian Cordillera, *Bull. Seismol. Soc. Am.*, *104*(4), 1964–1975, doi:10.1785/0120130233.
- Ribe, N. (1989), Seismic anisotropy and mantle flow, *J. Geophys. Res.*, *94*(B4), 4213–4223.
- Russo, R. M., and P. G. Silver (1994), Trench-parallel flow beneath the nazca plate from seismic anisotropy., *Science*, *263*(5150), 1105–1111, doi:10.1126/science.263.5150.1105.
- Sandvol, E., and T. Hearn (1994), Bootstrapping shear-wave splitting errors, *Bull. Seismol. Soc. Am.*, *84*(6), 1971–1977.
- Savage, M. S. (1999), Seismic anisotropy and mantle deformation: What have we learned from shear wave splitting?, *Rev. Geophys.*, *37*(1), 65–106, doi:10.1029/98RG02075.
- Schutt, D. L., and E. D. Humphreys (2001), Evidence for a deep asthenosphere beneath North America from western United States SKS splits, *Geology*, *29*(4), 291–294, doi:10.1130/0091-7613(2001)029<0291:EFADAB>2.0.CO;2.
- Silver, P. G. (1996), SEISMIC ANISOTROPY BENEATH THE CONTINENTS: Probing the Depths of Geology, *Annu. Rev. Earth Planet. Sci.*, *24*(1), 385–432, doi:10.1146/annurev.earth.24.1.385.
- Silver, P. G., and W. W. Chan (1991), Shear Wave Splitting and Sub continental Mantle Deformation, *J. Geophys. Res.*, *96*, 429–454, doi:10.1029/91JB00899.
- Silver, P. G., and M. K. Savage (1994), The interpretation of shear-wave splitting parameters in the presence of two anisotropic layers, *Geophys. J. Int.*, *119*(3), 949–963, doi:10.1111/j.1365-246X.1994.tb04027.x.
- Smith, G. P., D. a Wiens, K. M. Fischer, L. M. Dorman, S. C. Webb, and J. a Hildebrand (2001), A complex pattern of mantle flow in the Lau backarc., *Science*, *292*(5517), 713–6, doi:10.1126/science.1058763.
- Vinnik, L. P., L. I. Makeyeva, A. Milev, and A. Y. Usenko (1992), Global patterns of azimuthal anisotropy and deformations in the continental mantle, *Geophys. J. Int.*, *111*(3), 433–447, doi:10.1111/j.1365-246X.1992.tb02102.x.

- Walker, K. T., G. H. R. Bokelmann, and S. L. Klemperer (2001), Shear-wave splitting to test mantle deformation models around Hawaii, *Geophys. Res. Lett.*, *28*(22), 4319–4322, doi:10.1029/2001GL013299.
- Walsh, E., R. Arnold, and M. K. Savage (2013), Silver and Chan revisited, *J. Geophys. Res. Solid Earth*, *118*(September), 5500–5515, doi:10.1002/jgrb.50386.
- Wolfe, C. J., and P. G. Silver (1998), Seismic anisotropy of oceanic upper mantle: Shear wave splitting methodologies and observations, *J. Geophys. Res.*, *103*(B1), 749, doi:10.1029/97JB02023.
- Wüstefeld, A., G. Bokelmann, C. Zaroli, and G. Barruol (2008), SplitLab: A shear-wave splitting environment in Matlab, *Comput. Geosci.*, *34*(5), 515–528, doi:10.1016/j.cageo.2007.08.002.
- Yuan, H., K. Dueker, and D. L. Schutt (2008), Testing five of the simplest upper mantle anisotropic velocity parameterizations using teleseismic S and SKS data from the Billings, Montana PASSCAL array, *J. Geophys. Res. Earth*, *113*(B3), B03304, doi:B03304 Artn b03304.
- Zhang, S., and S. Karato (1995), Lattice preferred orientation of olivine aggregates deformed in simple shear, *Nature*, *375*, 774–777, doi:10.1038/375774a0.

## CHAPTER 2

### 2.1 Introduction

Shear wave splitting has proved to be a very useful technique to probe for seismic anisotropy in the earth's interior. Upper mantle anisotropy is probably the anisotropy most commonly studied, since it constrains past (“fossilized”) and present mantle deformation [e.g. *Silver, 1996; Long and Silver, 2009; Long and Wirth, 2013; Di Leo et al., 2014*]. At upper mantle depths the primary cause of anisotropy is strain-induced lattice preferred orientation of olivine [*Nicolas and Christensen, 1987; Silver and Chan, 1991; Zhang and Karato, 1995; Karato et al., 2008*], the predominant anisotropic upper mantle mineral. Because of the relationship between strain and anisotropy, it can be linked to deformational processes mantle [*Zhang and Karato, 1995; Long and Becker, 2010*]. However, several authors [e.g. *Ribe, 1989; Dawson and Wenk, 2000; Jadamec and Billen, 2010, 2012; Di Leo et al., 2012, 2014*] have pointed out that caution needs to be exercised when inferring mantle flow from shear wave splitting measurements, especially in complex geodynamics zones. Most seismic anisotropy studies evaluate the anisotropy with simplified Earth models, such as one horizontal anisotropic layer, and the results may not be an accurate representation of the anisotropic structure.

Probably the most frequently used phases in shear wave splitting studies are the SKS, SKKS, and PKS [*Wüstefeld et al., 2008; Long and Silver, 2009; Long and Becker, 2010; Liu and Gao, 2013*]. These phases are particularly useful because they limit the location of the observed anisotropy to the receiver side of the ray path [*Silver and Chan, 1991; Long and Silver, 2009; Long and Becker, 2010*]. They also have a known polarization (radial) because of the P-to-S conversion at the core mantle boundary [*Silver and Chan, 1991; Silver and Savage, 1994; Silver,*

1996; *Long and Silver, 2009*]. Additionally, their near-vertical incidence simplifies the mathematical analysis and provides an excellent lateral resolution. However, the vertical resolution is poor [*Alsina and Snieder, 1995; Silver, 1996; Savage, 1999; Becker et al., 2003; Wang et al., 2008; Wüstefeld et al., 2009; Yuan and Romanowicz, 2010*].

The transverse component minimization method [*Silver and Chan, 1991*] is probably the method most commonly used to study shear wave splitting. However, several studies have noted that this method underestimates the uncertainty in the splitting parameters [*Sandvol and Hearn, 1994; Levin et al., 1999; Savage, 1999; Menke and Levin, 2003; Walsh et al., 2013*], and it makes many assumptions about the anisotropic nature of the mantle. Some authors have attempted to obtain more statistically reliable errors, such as a stacking of the misfits for multiple observations [*Schutt et al., 1998; Wolfe and Silver, 1998*], or bootstrapping to estimate uncertainties [*Sandvol and Hearn, 1994*]. *Menke and Levin [2003]* introduced the cross-convolution method, which has more flexibility than the *Silver and Chan [1991]* method, in that it can constrain more complex anisotropic structures, and can integrate other observables such as receiver functions [e.g. *Bodin et al., 2013*].

With an exponential increase in teleseismic data, it becomes useful to put the analysis of shear wave splitting into a more careful statistical context, allowing for formal hypothesis testing and more nuanced interpretation of results. For example, a goal could be to test whether the fast axis orientation variation between two stations is significant, or just a consequence of uncertainty. A particular concern is that the *Silver and Chan [1991]* uncertainty estimation is incorrect due to the assumption that the minimized squared tangential energy can be approximated by a  $\chi^2$  distribution, with degrees of freedom approximated via a moment matching argument. This is an attempt to correct for the strong autocorrelation in the

microseismic noise. [Sandvol and Hearn, 1994; Levin et al., 1999; Savage, 1999; Menke and Levin, 2003; Walsh et al., 2013] suggest the autocorrelation may still be biasing the statistics. The other widely used method of [Menke and Levin, 2003] has no explicit error modeling. It further relies on an ad hoc estimate of degrees of freedom for an approximate  $F$ -test, without any rigorous justification.

Hence, we have developed a new method of shear wave splitting measurements based on the cross-convolution method [Menke and Levin, 2003], with which we aim to better characterize uncertainty, and which will also form the kernel of future codes that could formally stack data from different events, use Bayesian constraints from nearby stations, and consider more complex anisotropy (within the range that the Menke and Levin [2003] method is valid).

## 2.2 Methodology

The foundation of the new method is the cross-convolution approach of Menke and Levin [2003], who noted that if the radial and transverse waves travel similar ray paths, then for the correct anisotropic model ( $m_0$ ), the convolution of the radial impulse response of this model  $\rho_{m_0}$  with the observed transverse waveform  $T$ , should equal the convolution of the transverse impulse response  $\tau_{m_0}$  with the observed radial waveform  $R$ . Hence, to estimate  $m_0$ , search over geologically plausible values of anisotropy ( $m$ ) to find the minimizer of the objective function,

$$\|\rho_m * T - \tau_m * R\|^2, \quad (\text{Eqn. 2.2.1})$$

and use the minimizer  $\hat{m}$  as the estimator of  $m_0$ .

The goal of this study is to quantify the effect of microseismic noise on the estimation of  $m_0$ , which is characterized in terms of the fast axis azimuth,  $\phi$ , and split time,  $\delta t$ . For simplicity, this initial work assumes vertically oriented ray paths, propagating through a single layer of

anisotropy in which the fast axis is vertical and the anisotropic symmetry system is hexagonal. However, this method is readily adapted to more complex scenarios.

Consider the horizontal radial and transverse seismograms for any given station, on which the SKS signal and its pre-event noise have been windowed to isolate them from other interfering phases. Each time series (radial and transverse) can be cut into two subsets: the pre-event noise, from the beginning time of the windowed seismogram to immediately before the SKS arrival; and the event itself, from the end of the pre-event noise to the end of the observed SKS energy. Let  $r_k^*$  and  $t_k^*$  denote the pre-event noise in the radial and transverse series, where  $k$  is the integer-valued time index. Let  $B$  denote the backshift operator  $B^h X_k = X_{k-h}$ . Similar to [Menke and Levin, 2003], the two event series can be presented as convolutions of a common source signal  $S_k$  with impulse response functions, but we consider the inclusion of additive noise,

$$\begin{aligned} R_k &= \rho_{m_0}(B)S_k + r_k, \\ T_k &= \tau_{m_0}(B)S_k + t_k, \end{aligned} \quad (\text{Eqn. 2.2.2})$$

where  $r_k$  is an independent noise sequence with the same autocorrelation structure as  $r_k^*$ ,  $t_k$  is an independent noise sequence with the same autocorrelation structure as  $t_k^*$ , and  $\rho_{m_0}(B)$  and  $\tau_{m_0}(B)$  are polynomials in the backshift operator, evaluated at the corrected anisotropic model ( $m_0$ ). These polynomials involve study of the transformed time series

$$\begin{aligned} \rho_m(B)T_k - \tau_m(B)R_k &= \rho_m(B)\{\tau_{m_0}(B)S_k + t_k\} - \tau_m(B)\{\rho_{m_0}(B)S_k + r_k\} \\ &= \{\rho_m(B)\tau_{m_0}(B) - \tau_m(B)\rho_{m_0}(B)\}S_k \\ &\quad + \rho_m(B)t_k - \tau_m(B)r_k, \end{aligned} \quad (\text{Eqn. 2.2.3})$$

for each candidate model  $m$ . If  $m = m_0$ , then the cross-convolved signal term  $\{\rho_m(B)\tau_{m_0}(B) - \tau_m(B)\rho_{m_0}(B)\}S_k$  is exactly zero, and only the cross-convolved noise  $\rho_m(B)t_k - \tau_m(B)r_k$

remains. Understanding the structure of the noise is therefore essential to understanding the properties of any estimator built on cross-convolution. Fortunately, for any candidate model  $m$ , the properties of the unobservable series  $\rho_m(B)t_k - \tau_m(B)r_k$  are identical to the observable series  $\rho_m(B)t_k^* - \tau_m(B)r_k^*$ . Our method relies on modeling the characteristics of the latter series and using that model to develop a formal inferential procedure for the cross-convolved series.

It is particularly convenient to conduct the analysis in the frequency domain, via Fourier transforms of the time-domain sequences, since convolutions have simple forms in the frequency domain. The autocorrelation structure of the noise sequences in the time domain can equivalently be represented as spectral densities in the frequency domain, and the effect of the convolutions with the filters  $\tau_m(B)$  and  $\rho_m(B)$  is easily computed using power spectrum theory [e.g. *Welch*, 1967].

Prior to the frequency domain analysis, we pre-process the data so that the transformed pre-event noise sequences can be plausibly modeled as realizations of weakly stationary time series (models with time-invariant means and autocovariance functions). Our pre-processing includes differencing, centering and scaling. For  $n + 1$  observations  $k = 0, 1, \dots, n$  of  $r_k^*$  and  $t_k^*$ , we compute

$$\begin{aligned} r_k^{*(d)} &= r_k^* - r_{k-1}^* = (1 - B)r_k^* \\ t_k^{*(d)} &= t_k^* - t_{k-1}^* = (1 - B)t_k^* \end{aligned} \tag{Eqn. 2.2.4}$$

$$\begin{aligned} r_k^{*(c)} &= r_k^{*(d)} - n^{-1} \sum_{k=1}^n r_k^{*(d)} \\ t_k^{*(c)} &= t_k^{*(d)} - n^{-1} \sum_{k=1}^n t_k^{*(d)} \end{aligned} \tag{Eqn. 2.2.5}$$

$$\sigma = \frac{1}{2} \sqrt{\frac{\sum_{k=1}^n (r_k^{*(c)})^2}{n-1}} + \frac{1}{2} \sqrt{\frac{\sum_{k=1}^n (t_k^{*(c)})^2}{n-1}}$$

$$r_k^{*(p)} = \frac{r_k^{*(c)}}{\sigma}, \quad t_k^{*(p)} = \frac{t_k^{*(c)}}{\sigma}; \quad (k = 1, \dots, n), \quad (\text{Eqn. 2.2.6})$$

where  $B$  denotes the backshift operator as presented in (Eqn. 2.2.2),  $r_k^{*(d)}$  and  $t_k^{*(d)}$  stand for the differenced radial and transverse pre-event noise, respectively,  $r_k^{*(c)}$  and  $t_k^{*(c)}$  stand for the centered radial and transverse pre-event noise, respectively, and  $r_k^{*(p)}$  and  $t_k^{*(p)}$  stand for the pre-processed radial and transverse pre-event noise, respectively.

Our analysis will be conducted not with the original, raw series but with the preprocessed series. Next, we transform the pre-processed, pre-event noise time series from the time domain to the frequency domain, to estimate the autocorrelation structures of the pre-processed noise sequences with the corresponding spectral densities in the frequency domain. Denote the  $n - 1$  non-zero Fourier frequencies by

$$\omega_j = \frac{2\pi j}{n} \text{ for } j \in \mathcal{F}_n = \left\{ -\left\lfloor \frac{n-1}{2} \right\rfloor, \dots, \left\lfloor \frac{n}{2} \right\rfloor \right\} \setminus \{0\}, \quad (\text{Eqn. 2.2.7})$$

where  $\lfloor \cdot \rfloor$  denotes the integer part or floor function. Let  $\{I_R(\omega_j)\}_{j \in \mathcal{F}_n}$  and  $\{I_T(\omega_j)\}_{j \in \mathcal{F}_n}$  denote the normalized periodograms [Schuster, 1898] for the  $r_k^{*(p)}$  and  $t_k^{*(p)}$  data, respectively; that is, the squared modulus of the discrete Fourier transform divided by  $2\pi n$ .

The periodogram requires smoothing to yield a consistent estimator of the spectral density. In our smoothing approach, we use a flexible, semiparametric regression methodology to estimate the log-spectral density of each data set, regressing the log-periodogram on the non-zero Fourier frequencies using penalized spline regression. Let  $\text{ln}f_R(\omega)$  and  $\text{ln}f_T(\omega)$  denote the resulting estimated functions. The spectral densities describe the dependence structure in the pre-processed, pre-event noise sequences, and are used in the cross-convolution method for the pre-processed event sequences. Because the sequences can be very long, we treat  $\sigma$ ,  $\text{ln}f_R(\omega)$  and  $\text{ln}f_T(\omega)$  as known in what follows.

We turn now to the event series. We pre-process the event series in the same way as the pre-event noise series, but use the same constant  $\sigma$  for rescaling. Let  $M$  denote the maximum integer-valued lag to be considered. Accordingly, for  $N + M$  observations  $k = 0, 1, \dots, N + M$  of  $R_k$  and  $T_k$ , we compute

$$\begin{aligned} R_k^{(d)} &= R_k - R_{k-1} = (1 - B)R_k \\ T_k^{(d)} &= T_k - T_{k-1} = (1 - B)T_k \end{aligned} \quad (\text{Eqn. 2.2.8})$$

$$\begin{aligned} R_k^{(c)} &= R_k^{(d)} - n^{-1} \sum_{k=1}^{N+M} R_k^{(d)} \\ T_k^{(c)} &= T_k^{(d)} - n^{-1} \sum_{k=1}^{N+M} T_k^{(d)} \end{aligned} \quad (\text{Eqn. 2.2.9})$$

$$R_k^{(p)} = \frac{R_k^{(c)}}{\sigma}, \quad T_k^{(p)} = \frac{T_k^{(c)}}{\sigma}; \quad (k = 1, \dots, N + M), \quad (\text{Eqn. 2.2.10})$$

and plug in to the cross-convolution formula to yield

$$\begin{aligned} \rho_m(B)T_k^{(p)} - \tau_m(B)R_k^{(p)} &= \rho_m(B) \left\{ \tau_{m_0}(B)S_k^{(p)} + t_k^{(p)} \right\} - \tau_m(B) \left\{ \rho_{m_0}(B)S_k^{(p)} + r_k^{(p)} \right\} \\ &= \left\{ \rho_m(B) \tau_{m_0}(B) - \tau_m(B) \rho_{m_0}(B) \right\} S_k^{(p)} + \rho_m(B)t_k^{(p)} - \tau_m(B)r_k^{(p)}, \end{aligned} \quad (\text{Eqn. 2.2.11})$$

by the same argument leading to (Eqn. 2.2.3). If  $m = m_0$ , then the cross-convolved pre-processed signal term  $\left\{ \rho_m(B) \tau_{m_0}(B) - \tau_m(B) \rho_{m_0}(B) \right\} S_k^{(p)}$  is exactly zero, and only the cross-convolved pre-processed noise  $\rho_m(B)t_k^{(p)} - \tau_m(B)r_k^{(p)}$  remains. The autocorrelation structure and spectral density of the unobserved series  $\rho_m(B)t_k^{(p)} - \tau_m(B)r_k^{(p)}$  are the same as the autocorrelation structure of the observed series  $\rho_m(B)t_k^{*(p)} - \tau_m(B)r_k^{*(p)}$ , for any model  $m$ . This common spectral density is

$$f_x(\omega) = \left| \rho_m(e^{-i\omega}) \right|^2 f_T(\omega) + \left| \tau_m(e^{-i\omega}) \right|^2 f_R(\omega), \quad (\text{Eqn. 2.2.12})$$

where  $f_T(\omega)$  and  $f_R(\omega)$  have previously been computed, and the power transfer functions

$\left| \rho_m(e^{-i\omega}) \right|^2$  and  $\left| \tau_m(e^{-i\omega}) \right|^2$  depend on the form of the impulse response functions.

Let  $\phi$  denote the unknown fast axis, and let  $\theta$  denote the known back azimuth. To calculate the predicted impulse response functions we follow the *Menke and Levin* [2003] approximation of one anisotropic layer to a normally incident shear wave from below,

$$\begin{aligned}\rho_m(B) &= \rho_1 + \rho_2 B^L, \quad \rho_1 = \cos^2(\psi), \quad \rho_2 = \sin^2(\psi), \\ \tau_m(B) &= \tau_1 + \tau_2 B^L, \quad \tau_1 = \cos(\psi) \sin(\psi), \quad \tau_2 = -\tau_1,\end{aligned}\quad (\text{Eqn. 2.2.13})$$

where  $m = (\phi, \delta t)$  and  $\psi = \phi - \theta$ . The lag  $L$  is integer-valued, and it is a function of the split time  $\delta t$  (in seconds) and of the sampling interval  $\text{SI}$  (in seconds). With this choice of impulse response functions, we have

$$\begin{aligned}f_x(\omega) &= |\rho_m(e^{-i\omega})|^2 f_T(\omega) + |\tau_m(e^{-i\omega})|^2 f_R(\omega) \\ f_x(\omega) &= |\rho_1 + \rho_2(e^{-i\omega})^L|^2 f_T(\omega) + |\tau_1 + \tau_2(e^{-i\omega})^L|^2 f_R(\omega) \\ f_x(\omega) &= [(\rho_1 + \rho_2 \cos L\omega)^2 + (-\rho_2 \sin L\omega)^2] f_T(\omega) \\ &\quad + [(\tau_1 + \tau_2 \cos L\omega)^2 + (-\tau_2 \sin L\omega)^2] f_R(\omega) \\ f_x(\omega) &= \{\rho_1^2 + \rho_2^2 + 2\rho_1\rho_2 \cos(L\omega)\} f_T(\omega) \\ &\quad + \{\tau_1^2 + \tau_2^2 + 2\tau_1\tau_2 \cos(L\omega)\} f_R(\omega).\end{aligned}\quad (\text{Eqn. 2.2.14})$$

That is, if the candidate model  $m$  is the true model  $m_0$ , then the theoretical spectral density for the cross-convolved, pre-processed event series is given by (Eqn. 2.2.14). Our proposed approach is therefore to find  $\phi$  and  $L$  (and hence  $\delta t$ ) such that the cross-convolved series have normalized periodogram (empirical spectral densities) as consistent as possible, in some sense, with the theoretical spectral density in (Eqn. 2.2.14). Such consistency is achieved by maximizing the Whittle likelihood [*Whittle*, 1953; *Brockwell and Davis*, 1991] in a grid search over all plausible values of fast axis orientation,  $\phi$ , and split time,  $\delta t$ .

Computing the periodogram of the pre-processed, event signal sequences, we window the data so that models with different lags use the same number of observations,  $N$ . Let  $M$  denote the maximum integer-valued lag to be considered. Then, for a given candidate model  $m = (\phi, \delta t)$ , the data set will consist of  $N$  cross-convolved observations,

$$\{X_t\}_{k=M+1}^{M+N} = \left\{ \rho_m(B)T_k^{(p)} - \tau_m(B)R_k^{(p)} \right\}_{k=M+1}^N. \quad (\text{Eqn. 2.2.15})$$

Denote the  $N - 1$  non-zero Fourier frequencies by

$$\Omega_j = \frac{2\pi j}{N} \text{ for } j \in \mathcal{F}_N = \left\{ -\left\lfloor \frac{N-1}{2} \right\rfloor, \dots, \left\lfloor \frac{N}{2} \right\rfloor \right\} \setminus \{0\}. \quad (\text{Eqn. 2.2.16})$$

If  $\phi$  and  $L$  are the data-generating parameters, then by (Eqn. 2.2.14) the normalized periodogram ordinates  $\{I_X(\Omega_j)\}_{j \in \mathcal{F}_N}$  are approximately independent exponential random variables with expected value  $E[I_X(\Omega_j)] = f_X(\Omega_j)$ . Accordingly, the Whittle approximation to the log-likelihood of  $\{X_k\}_{k=M+1}^{M+N}$  at lag  $L$  is

$$\mathcal{L}(L, \phi) = -\sum_{j \in \mathcal{F}_N} \ln f_X(\Omega_j) - \sum_{j \in \mathcal{F}_N} I_X(\Omega_j) / f_X(\Omega_j). \quad (\text{Eqn. 2.2.17})$$

Now (Eqn. 2.2.17) is a function of lag  $L$ , which in turn depends on  $\delta t$ , the unknown split time (in seconds) and on the SI, the known sampling interval (in seconds). Since  $\delta/\text{SI}$  is not necessarily integer-valued, we compute the log-likelihood for both the lower integer-valued lag  $L_1 = \lfloor \delta/\text{SI} \rfloor \leq \delta/\text{SI}$  and the upper integer-valued lag  $L_2 = L_1 + 1 > \delta/\text{SI}$ . We then linearly interpolate results between these two values, yielding the interpolated log-likelihood for split time and fast axis,

$$\widetilde{\mathcal{L}}(\delta, \phi) = (L_2 - \delta/\text{SI})\mathcal{L}(L_1, \phi) + (\delta/\text{SI} - L_1)\mathcal{L}(L_2, \phi), \quad (\text{Eqn. 2.2.18})$$

which equals  $\mathcal{L}(L_1, \phi)$  if  $L_1 = \delta/\text{SI}$ . We maximize  $\widetilde{\mathcal{L}}(\delta, \phi)$  via a grid search over  $\delta t$  and  $\phi$  to obtain the estimates,  $\widehat{\delta t}$  and  $\widehat{\phi}$ .

### 2.2.1 Calculation of Standard Errors

To calculate the standard errors we have developed a bootstrapping procedure to work together with our code.

Since we can present our time series as a filtered source signal  $S_k$  plus noise, we first estimate the source signal  $S_k$  with a least-square regression model fit,

$$\begin{bmatrix} R_{L+1} \\ \vdots \\ R_N \\ T_{L+1} \\ \vdots \\ T_N \end{bmatrix} = \begin{bmatrix} \rho_2 & 0 & \dots & 0 & \rho_1 \\ & \ddots & & & \ddots \\ & & \rho_2 & 0 & \dots & 0 & \rho_1 \\ & & \tau_2 & 0 & \dots & 0 & \tau_1 \\ & & & \ddots & & & \ddots \\ & & & & \tau_2 & 0 & \dots & 0 & \tau_1 \end{bmatrix} \begin{bmatrix} S_1 \\ \vdots \\ S_N \end{bmatrix} + \begin{bmatrix} r_{L+1} \\ \vdots \\ r_N \\ t_{L+1} \\ \vdots \\ t_N \end{bmatrix} \quad (\text{Eqn.2.2.1.1})$$

plugging in estimated values of the filter coefficients  $\rho_1$ ,  $\rho_2$ ,  $\tau_1$  and  $\tau_2$ , which are functions of the estimates,  $\widehat{\delta t}$  and  $\widehat{\phi}$ . We test the estimation on synthetic data with different noise levels, and for a very low noise level the estimated source signal  $\widehat{S}_k$  reproduces the source signal  $S_k$  used to create the synthetic data, a Ricker wavelet (Figure 2.1).

Next, we produce noise replicates surrogated from the pre-processed radial and transverse pre-event noise  $r_k^{*(p)}$  and  $t_k^{*(p)}$ , respectively. The noise replicates are uncorrelated with the pre-event noise or with each other, but they all share the same spectrogram characteristics, i.e. the replicates have the same autocorrelation function but different randomized phase in each realization. We produce as many replicates as pairs of bootstrap shear wave splitting results as we need to obtain the standard errors and their distribution. We consider that 100 bootstrap replicates are enough to describe the error distribution and obtain reliable standard errors [Efron and Tibshirani, 1991; Sandvol and Hearn, 1994]. We add the noise replicates to the estimated source signal  $\widehat{S}_k$ , so we create a set of N replicates of  $\{r_k^*, R_k\}_b$  and  $\{t_k^*, T_k\}_b$ . We resample N times the estimated fast axis,  $\widehat{\phi}$ , and the estimated split time,  $\widehat{\delta t}$ . The standard error is the

empirical standard deviation of the  $N$  bootstrap replicates distribution for each of the estimates,  $\{\hat{\delta t}\}_b$  and  $\{\hat{\phi}\}_b$ .

Note that the standard errors calculated for the results presented in this document are counter-intuitively small—particularly for the splitting time. While the development and presentation of the new method resulting in the calculation of the likelihood surface has been a significant project in itself, having proper uncertainty bounds is also essential. Thus, the reader should be aware that we continue to examine the properties of the likelihood surface to determine the best approach by which to assess uncertainty.

The likelihood surfaces we obtain from typical SKS data are not easy to properly characterize, as they usually present different ridges and plateaus, and not the expected nicely-resolved smooth peaks. Generally, with well-resolved likelihood surfaces, the standard errors can be calculated with a simple quadratic fit of the peak at which the maximum likelihood estimator is located. However, the geometry of the retrieved likelihood surfaces is such that another method is needed, so we have chosen to use a bootstrapping technique to estimate the standard errors.

## **2.3 Results**

### *2.3.1 Test on Synthetic Data Set*

A synthetic data set has been produced to test the reliability of the method presented here. First, we calculate the impulse responses of a split SKS for a range of varying back azimuths from  $0^\circ$  to  $180^\circ$ . The impulse responses are calculated according to equation (Eqn. 2.2.2) for a split time of 1.2 seconds and a fast axis of 20 degrees. Second, we convolve the impulse responses with a Ricker wavelet, and we add phase-randomize seismic noise to the convolved waveform to simulate real seismic data. We scale the variance of the noise so that the signal-to-

noise ratio of the Ricker wavelet with the noise is 1. Because the variance of the impulse responses vary by back azimuth, the signal-to-noise ratio of the synthetics varies from about 0 to 1 depending on the variance of the impulse responses for each back azimuth event. The method successfully reproduces the input splitting parameters ( $\hat{\phi}$ ,  $\hat{\delta t}$ ) in all the synthetic tests (Figure 2.2). In the case of null splits (Figure 2.2 A) it is worth remarking that the logarithmic likelihood surfaces correctly constrain the two possible values of the fast axis polarization ( $\phi = 20^\circ$  and  $\phi + 90 = 110^\circ$ ), although the splitting estimates are not relevant.

### 2.3.2 Application to Pasadena Seismic Station

Station PAS was selected as a test location to compare the new method to previous studies results, and because it is well-studied and seems to have a single-layer of well-resolved anisotropy [*Liu et al.*, 1995a; *Özalaybey and Savage*, 1995; *Polet and Kanamori*, 2002], although some small lateral variations in the anisotropic medium may be present [*Polet and Kanamori*, 2002]. The PAS station is located in southern California, in Pasadena, about 60 kilometers from the San Andreas Fault. It is part of the TERRAScope Broadband Network. This station is used to benchmark our method, so we do not attempt to explain the implications of the measured splitting. Splitting measurements around the PAS station in southern California, and west of the San Andreas Fault, are generally consistent with each other and indicate an average E-W fast axis direction, and a delay time ranging from 1 s to 1.5 s [*Özalaybey and Savage*, 1995; *Polet and Kanamori*, 2002]. Table 2.1 shows the splitting results for three different studies undertaken in southern California [*Liu et al.*, 1995a; *Özalaybey and Savage*, 1995; *Polet and Kanamori*, 2002]. They all use the transverse component minimization method to analyze the SKS splitting.

The PAS data set was retrieved by means of the Standing Order for Data (SOD) software [*Owens et al.*, 2004]. Data from April the 5<sup>th</sup> 1998 to December the 31<sup>st</sup> 2015, within a

magnitude range of 5.8 – 10, including all types of magnitudes, with distance ranges within 85° and 125°, and with signal-to-noise ratio greater than 2.5 was requested. A total number of 122 events were retrieved. After a visual inspection, 12 events were used. The resulting estimates,  $\hat{\delta t}$  and  $\hat{\phi}$  are consistent with results from previous studies (Table 2.1), so we conclude that our method can correctly resolve the splitting parameter of a split shear wave.

## 2.4 Conclusions

We have presented a new method to estimate shear wave splitting parameters based on the cross-convolution approach [Menke and Levin, 2003]. The application of this new method will yield a better statistical characterization of shear wave splitting, and therefore a better understanding of past lithospheric deformation and ongoing asthenospheric flow. Synthetic tests show that the method can successfully obtain the split time and the fast axis orientation from observed split shear waves. We have also tested our method on SKS events retrieved from station PAS (Pasadena, California), and compared our results with results from other studies that use the transverse component minimization method [Silver and Chan, 1991]. Said comparison shows that our results are very reasonable. The method presented here is the groundwork for further development of this procedure, which could include more complex anisotropic models such as multiple layers, dipping layer(s), and non-vertical incidence angle. Therefore, it is a first step towards testing complex seismic anisotropy, and gaining a better understanding of the geodynamics in complex settings.

## 2.5 Figures

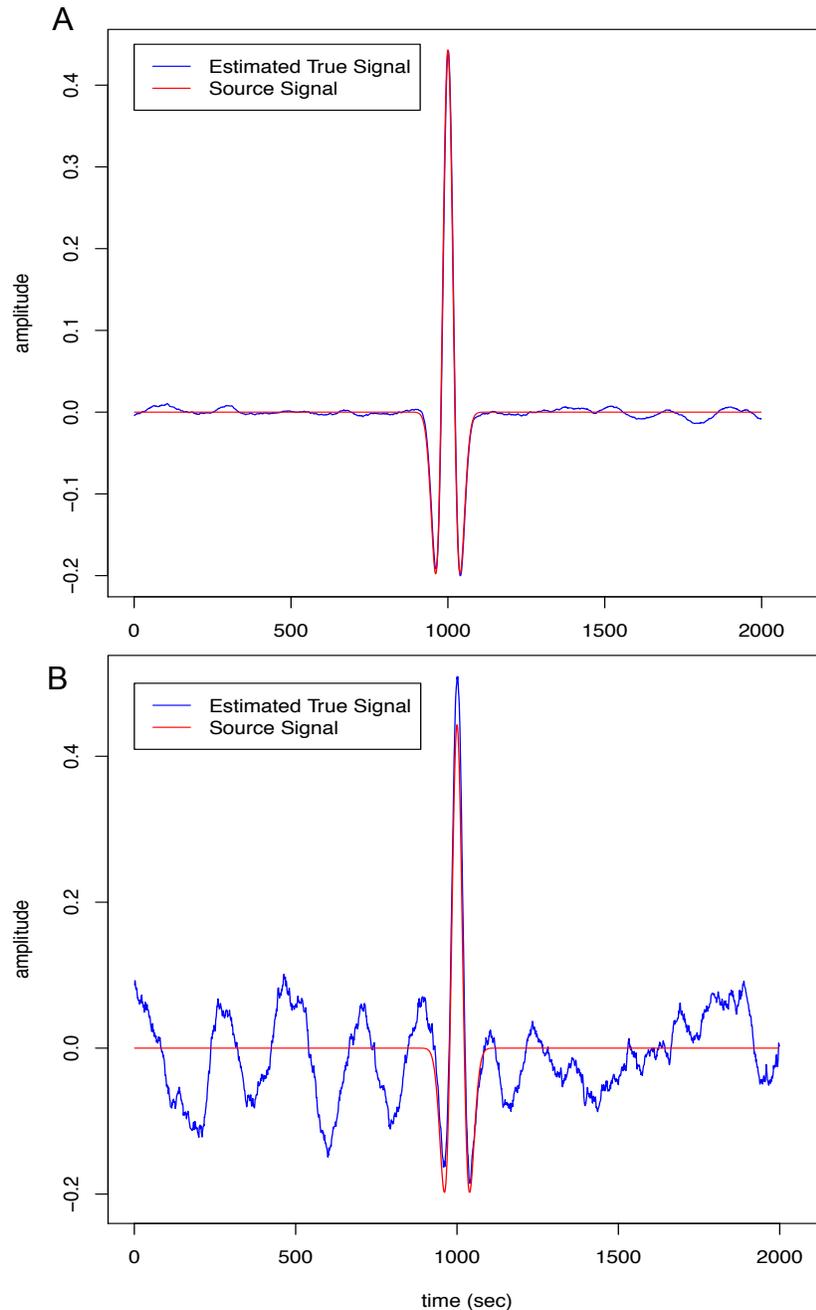
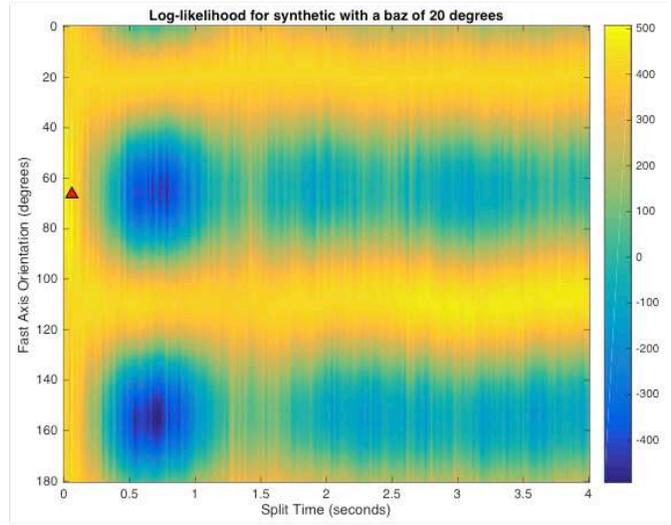
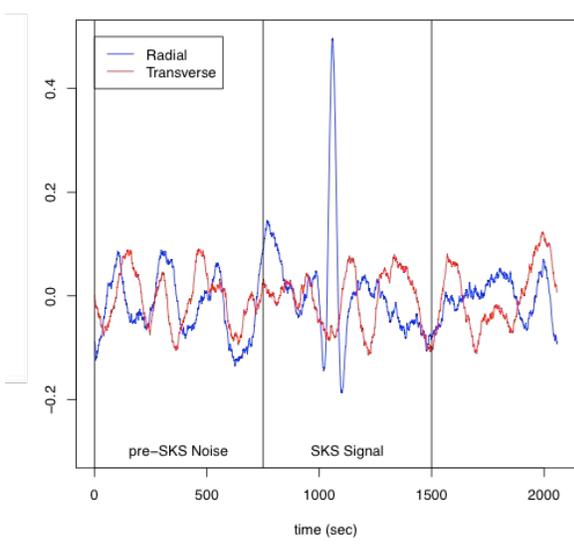
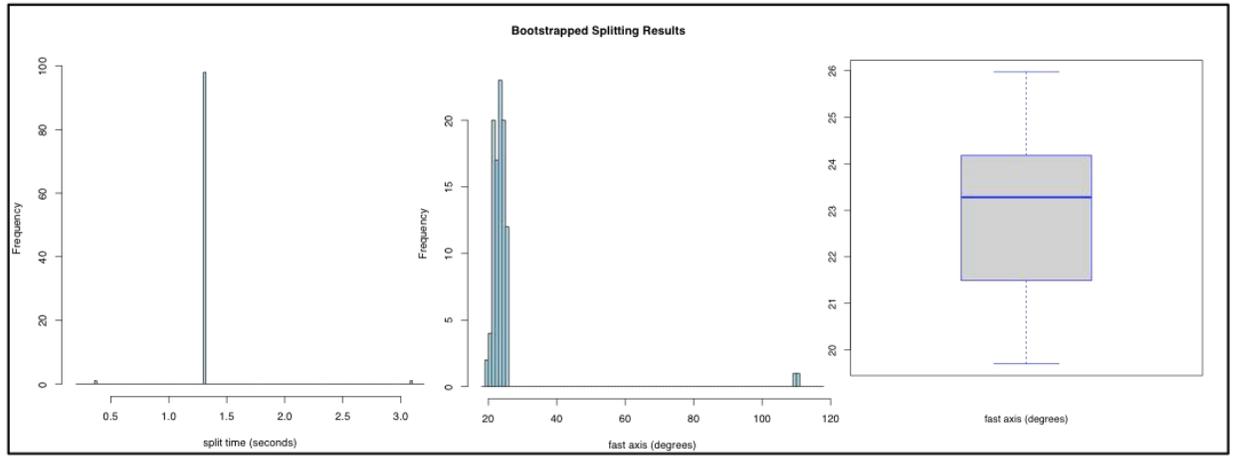
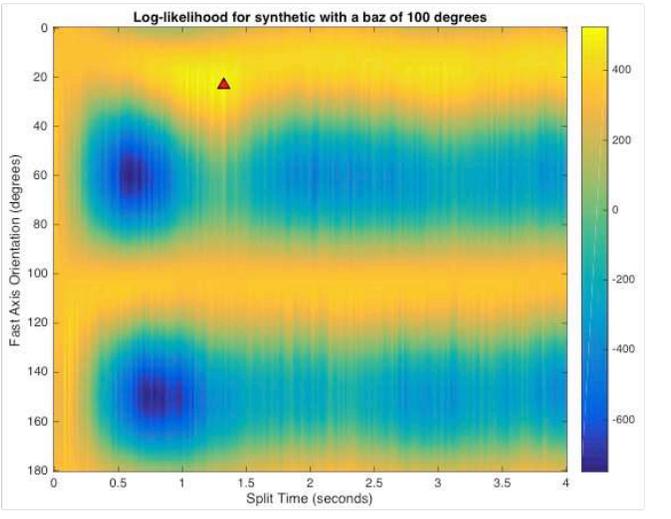
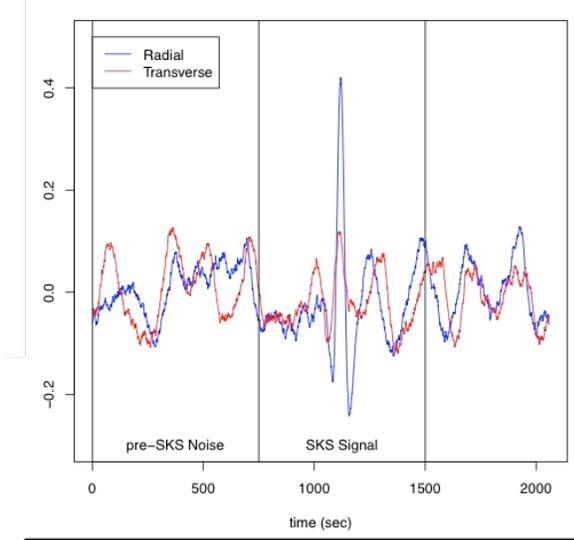


Figure 2.1: Examples of the true signal estimation on synthetic data set used for the bootstrapping approach to calculate the standard errors. A. Comparison of the estimated true signal (blue line) with the source signal (red line). For very low noise levels in the synthetic data, the true signal estimation represents accurately the source signal. B. Comparison of the estimated true signal (blue line) with the source signal (red line). For typical noisy data, the true signal estimation is well resolved but the noise is not removed.

**A Synthetic event for a back azimuth of 20 degrees**



**B Synthetic event for a back azimuth of 100 degrees**



C

Synthetic event for a back azimuth of 140 degrees

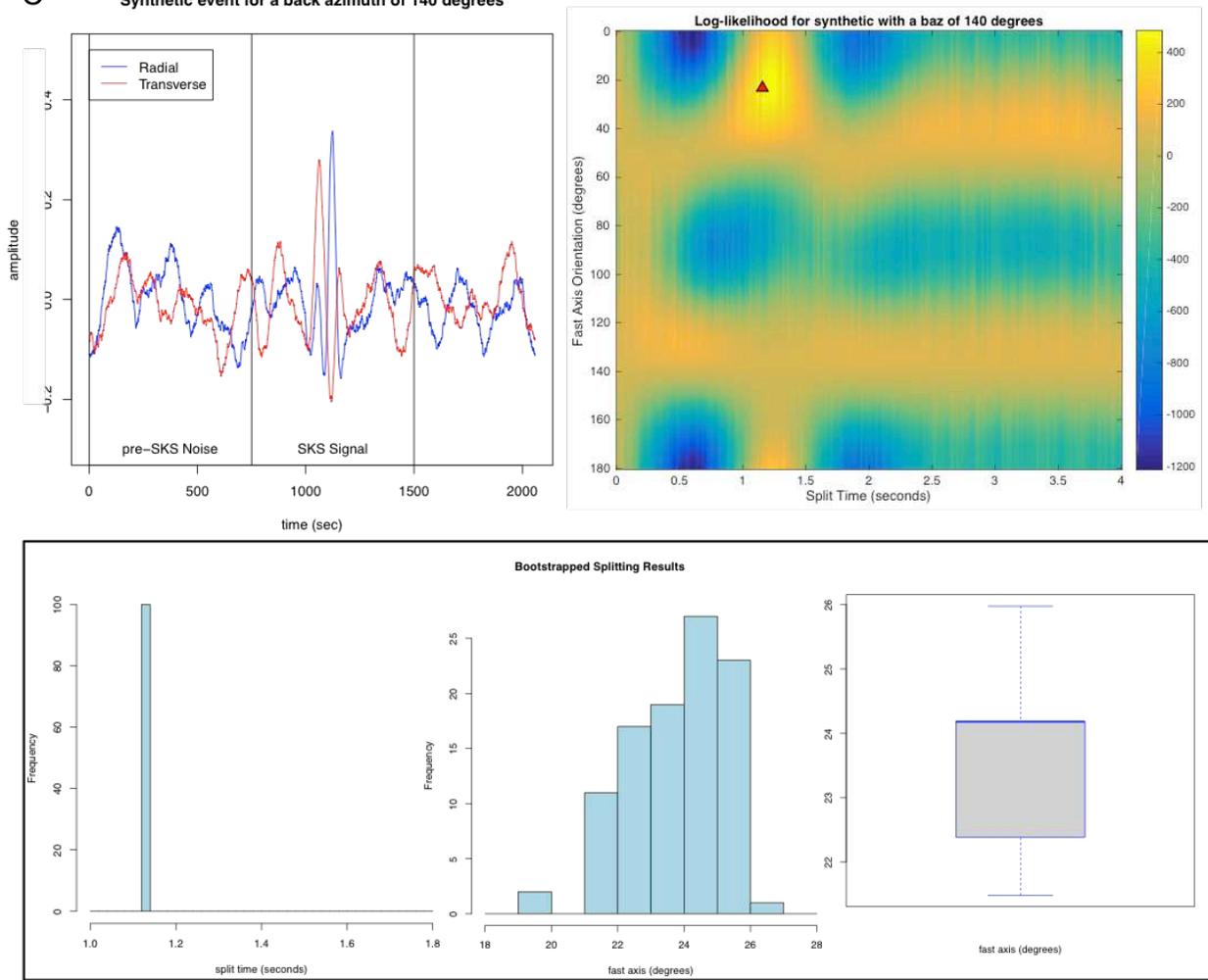


Figure 2.2: Results from the three synthetic tests using an input of  $\widehat{\delta t} = 1.2$  s and  $\widehat{\phi} + 20^\circ$ . The top left panel of each of the three subfigures shows the radial (blue line) and transverse (red line) waveforms, and the logarithmic likelihood surface in the right. The sections of the seismogram user-identified as pre-event noise and as signal for processing is indicated. Two magenta lines are superimposed on the logarithmic likelihood surface to indicate the synthetic values of split time and fast axis orientation, and their intersection shows the real value of the  $\widehat{\delta t}$  and  $\widehat{\phi}$ . This can be compared to the location of the maximum likelihood, indicated by the red triangle. The bottom panel shows the histogram of the bootstrapped split time results in the left, the histogram of the bootstrapped fast axis results in the center, and the bootstrapped fast axis results (without outliers, if they exist) in a box plot. The distributions of the histograms are the distributions of our errors, and the standard errors are calculated as the empirical standard deviation of those distributions. A. Synthetic test for a back azimuth of 20 degrees, which is a null split. B. Synthetic test for a back azimuth of 100, a near-null split. The estimates are  $\widehat{\delta t} = 1.32 \pm 0.20$  s and  $\widehat{\phi} = 23.28^\circ \pm 12.27^\circ$ . C. Synthetic test for a back azimuth of 140°, maximum split. The estimates are  $\widehat{\delta t} = 1.16 \pm 0.00$  s and  $\widehat{\phi} = 23.28^\circ \pm 1.42^\circ$ .

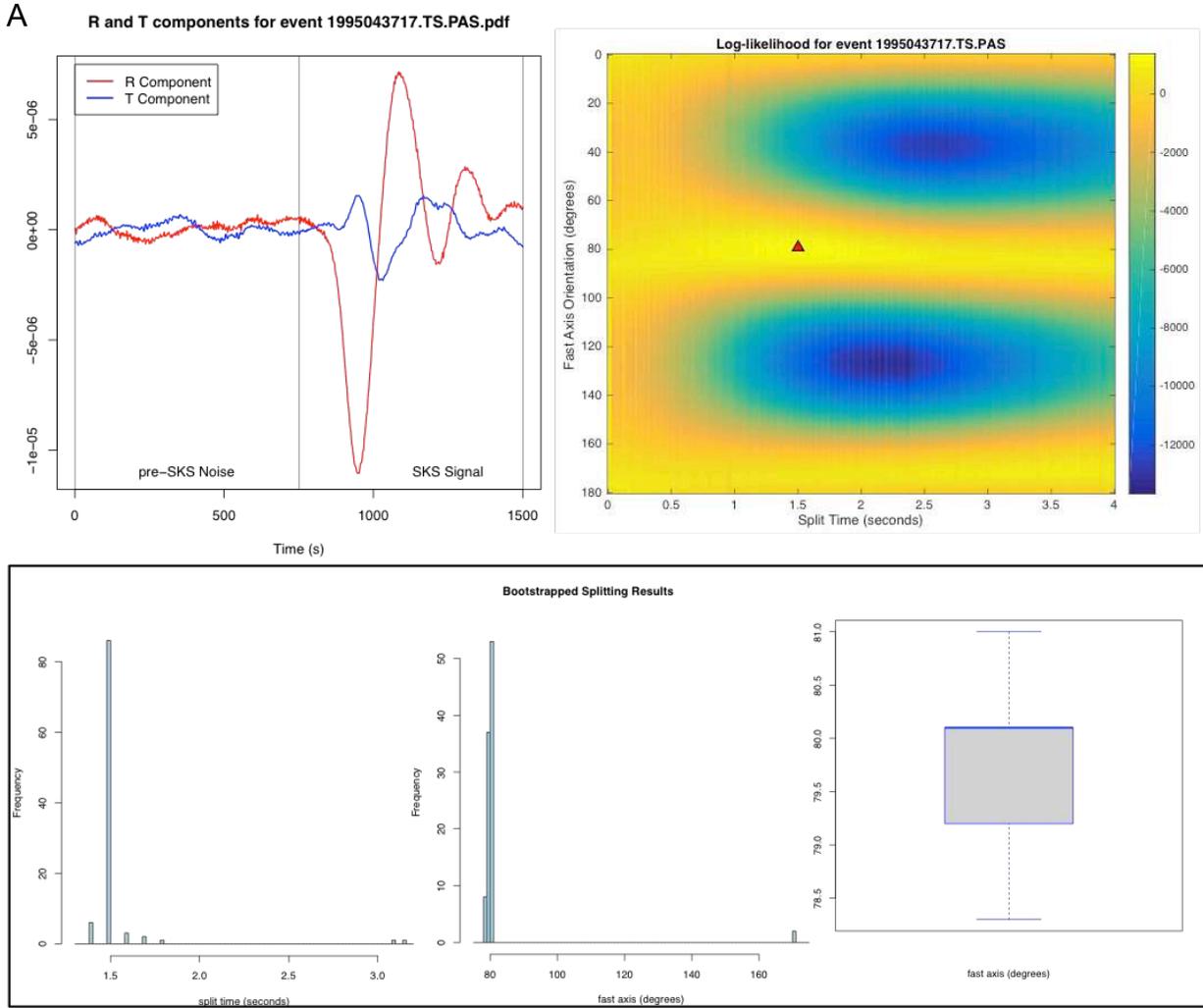
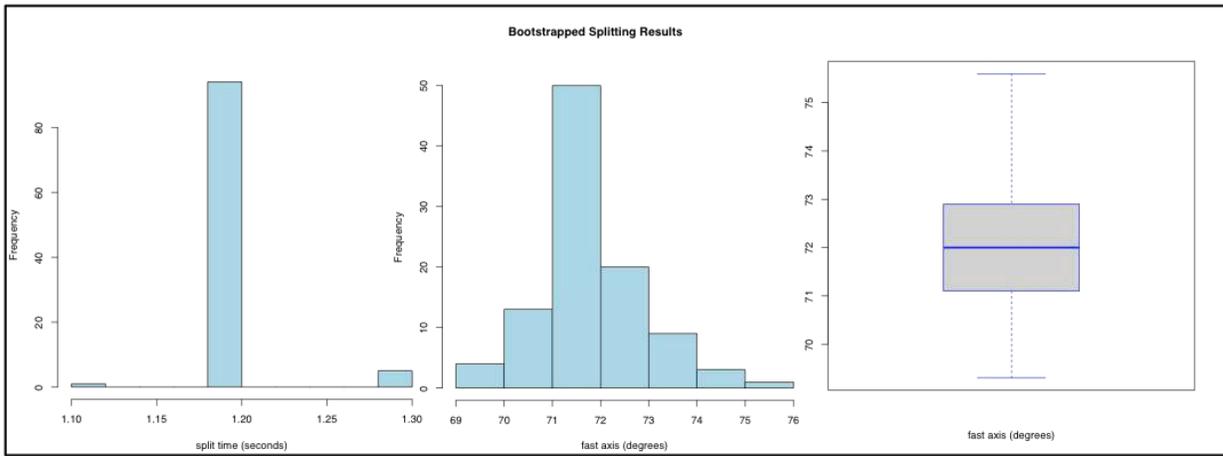
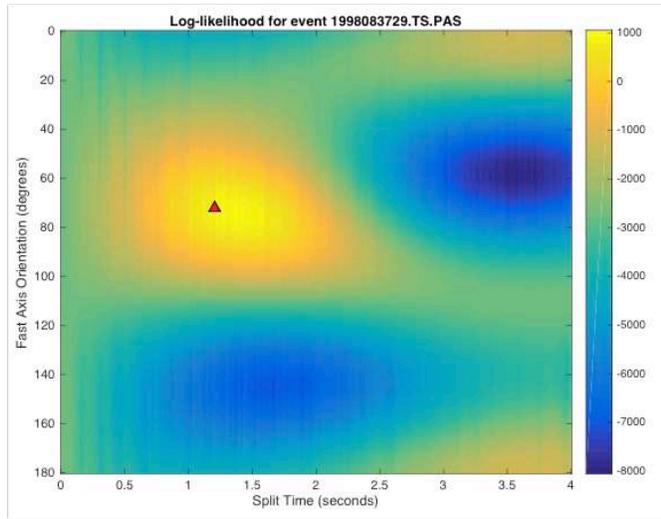
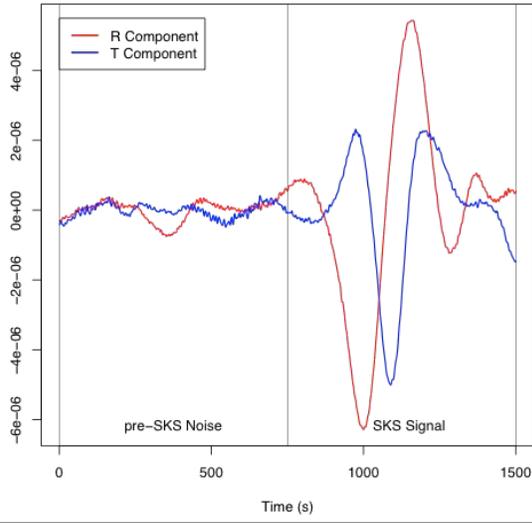


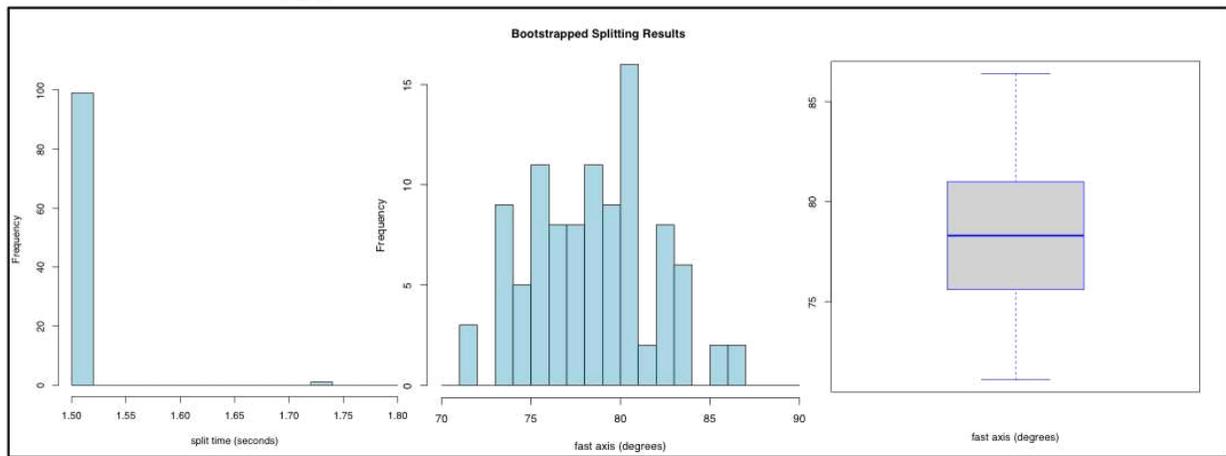
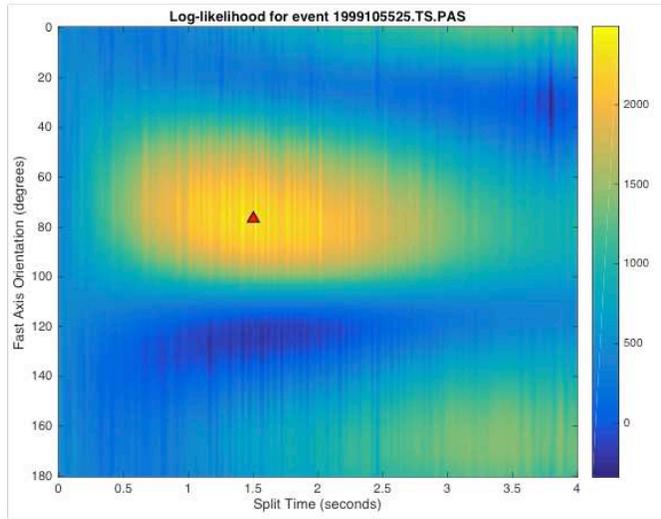
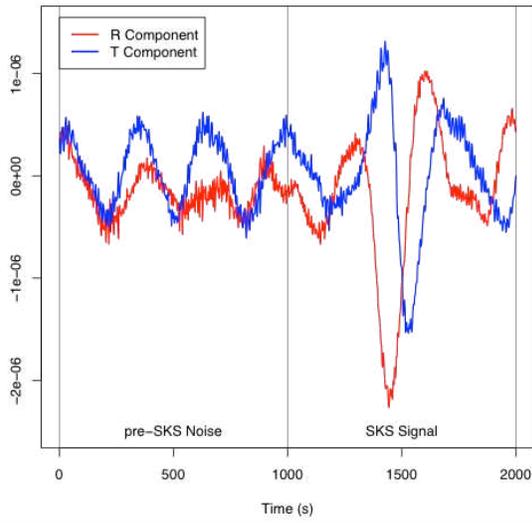
Figure 2.3: Results from the Pasadena (PAS) seismic station. Each subfigure is as described in Figure 2.2. A. Results from an event that occurred in 1995 at 04:37:17 GMT;  $\hat{\delta t} = 1.5 \pm 0.23$  s and  $\hat{\phi} = 79.2^\circ \pm 12.86^\circ$ . B. Results from an event that occurred in 1998 at 08:37:29 GMT;  $\hat{\delta t} = 1.2 \pm 0.02$  s and  $\hat{\phi} = 72^\circ \pm 1.31^\circ$ . C. Results from an event that occurred in 1999 at 10:55:25 GMT;  $\hat{\delta t} = 1.5 \pm 0.02$  s and  $\hat{\phi} = 76.5^\circ \pm 3.45^\circ$ .

**B** R and T components for event 1998083729.TS.PAS.pdf



C

R and T components for event 1999105525.TS.PAS.pdf



## 2.6 Tables

Author of study	Averaged SKS wave splitting results for PAS station	
	$\delta t$ (seconds)	$\phi$ (degree)
[ <i>Liu et al.</i> , 1995b]	$1.2 \pm 0.1$	$72 \pm 4$
[ <i>Özalaybey and Savage</i> , 1995]	$1.2 \pm 0.2$	$82 \pm 5$
[ <i>Polet and Kanamori</i> , 2002]	$0.95 \pm 0.25$	$77 \pm 12$

Table 2.1: Table with SKS wave splitting results for the PAS station from previous studies. We use these values to benchmark our method.

## REFERENCES

- Alsina, D., and R. Snieder (1995), Small-scale sublithospheric continental mantle deformation : constraints from SKS splitting observations, *Geophysical Journal International*, 431–448, doi:10.1111/j.1365-246X.1995.tb06864.x.
- Becker, T. W., J. B. Kellogg, G. Ekström, and R. J. O’Connell (2003), Comparison of azimuthal seismic anisotropy from surface waves and finite strain from global mantle-circulation models, *Geophys. J. Int.*, 155(2), 696–714, doi:10.1046/j.1365-246X.2003.02085.x.
- Bodin, T., H. Yuan, and B. Romanowicz (2013), Inversion of receiver functions without deconvolution--application to the Indian craton, *Geophys. J. Int.*, 196, 1025–1033, doi:10.1093/gji/ggt431.
- Brockwell, P. J., and R. A. Davis (1991), *Time Series: Theory and Methods*.
- Dawson, P. R., and H.-R. R. Wenk (2000), Texturing of the upper mantle during convection, *Philos. Mag. A*, 80(3), 573–598, doi:10.1080/014186100250723.
- Efron, B., and R. Tibshirani (1991), Statistical Data Analysis in the Computer Age, *Sci. New Ser.*, 253(5018), 390–395, doi: 10.1126/science.253.5018.390.
- Jadamec, M. a., and M. I. Billen (2010), Reconciling surface plate motions with rapid three-dimensional mantle flow around a slab edge., *Nature*, 465(7296), 338–41, doi:10.1038/nature09053.
- Jadamec, M. a., and M. I. Billen (2012), The role of rheology and slab shape on rapid mantle flow: Three-dimensional numerical models of the Alaska slab edge, *J. Geophys. Res. Solid Earth*, 117(2), 1–20, doi:10.1029/2011JB008563.
- Karato, S., H. Jung, I. Katayama, and P. Skemer (2008), Geodynamic Significance of Seismic Anisotropy of the Upper Mantle: New Insights from Laboratory Studies, *Annu. Rev. Earth Planet. Sci.*, 36(1), 59–95, doi:10.1146/annurev.earth.36.031207.124120.
- Di Leo, J. F., J. Wookey, J. O. S. Hammond, J. M. Kendall, S. Kaneshima, H. Inoue, T. Yamashina, and P. Harjadi (2012), Mantle flow in regions of complex tectonics: Insights from Indonesia, *Geochemistry, Geophys. Geosystems*, 13(12), 1–20, doi:10.1029/2012GC004417.
- Di Leo, J. F., a. M. Walker, Z. H. Li, J. Wookey, N. M. Ribe, J. M. Kendall, and a. Tommasi (2014), Development of texture and seismic anisotropy during the onset of subduction, *Geochemistry, Geophys. Geosystems*, 15(1), 192–212, doi:10.1002/2013GC005032.
- Levin, V., W. Menke, and J. Park (1999), Shear wave splitting in the Appalachians and the Urals: A case for multilayered anisotropy, *J. Geophys. Res.*, 104(B8), 17975–17993, doi:10.1029/1999JB900168.
- Liu, H., P. M. Davis, and S. Gao (1995a), SKS splitting beneath southern California, *Geophys. Res. Lett.*, 22(7), 767, doi:10.1029/95GL00487.
- Liu, H., P. M. Davis, and S. Gao (1995b), SKS splitting beneath southern California, *Geophys. Res. Lett.*, 22(7), 767, doi:10.1029/95GL00487.
- Liu, K. H., and S. S. Gao (2013), Making reliable shear-wave splitting measurements, *Bull. Seismol. Soc. Am.*, 103(5), 2680–2693, doi:10.1785/0120120355.

- Long, M. D., and T. W. Becker (2010), Mantle dynamics and seismic anisotropy, *Earth Planet. Sci. Lett.*, 297(3-4), 341–354, doi:10.1016/j.epsl.2010.06.036.
- Long, M. D., and P. G. Silver (2009), Shear Wave Splitting and Mantle Anisotropy: Measurements, Interpretations, and New Directions, *Surv. Geophys.*, 30(4-5), 407–461, doi:10.1007/s10712-009-9075-1.
- Long, M. D., and E. A. Wirth (2013), Mantle flow in subduction systems: The mantle wedge flow field and implications for wedge processes, *J. Geophys. Res. Solid Earth*, 118(2), 583–606, doi:10.1002/jgrb.50063.
- Menke, W., and V. Levin (2003), The cross-convolution method for interpreting SKS splitting observations, with application to one and two-layer anisotropic earth models, *Geophys. J. Int.*, 154(2), 379–392, doi:10.1046/j.1365-246X.2003.01937.x.
- Nicolas, A., and N. I. Christensen (1987), Formation of anisotropy in upper mantle peridotites - a review, *Am. Geophys. Union*, 16, 111–123, doi:10.1029/GD016p0111.
- Owens, T. J., H. P. Crotwell, C. Groves, and P. Oliver-Paul (2004), Sod: Standing Order for Data, *Seismol. Res. Lett.*, 75(4), 515–520, doi:10.1785/gssrl.75.4.515-a.
- Özalaybey, S., and M. K. Savage (1995), Shear-wave splitting beneath western United States in relation to plate tectonics, *J. Geophys. Res.*, 100(B9), 18135, doi:10.1029/95JB00715.
- Polet, J., and H. Kanamori (2002), Anisotropy beneath California: shear wave splitting measurements using a dense broadband array, *Geophys. J. Int.*, 149(2), 313–327, doi:10.1046/j.1365-246X.2002.01630.x.
- Ribe, N. (1989), Seismic anisotropy and mantle flow, *J. Geophys. Res.*, 94(B4), 4213–4223.
- Sandvol, E., and T. Hearn (1994), Bootstrapping shear-wave splitting errors, *Bull. Seismol. Soc. Am.*, 84(6), 1971–1977.
- Savage, M. S. (1999), Seismic anisotropy and mantle deformation: What have we learned from shear wave splitting?, *Rev. Geophys.*, 37(1), 65–106, doi:10.1029/98RG02075.
- Schuster, A. (1898), On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena, *Terr. Magn.*, 3(1), 13–41, doi:10.1029/TM003i001p00013.
- Schutt, D., E. D. Humphreys, and K. Dueker (1998), Anisotropy of the Yellowstone Hot Spot Wake, Eastern Snake River Plain, Idaho, *Pure Appl. Geophys.*, 151, 443–462, doi:10.1007/s000240050122.
- Silver, P. G. (1996), SEISMIC ANISOTROPY BENEATH THE CONTINENTS: Probing the Depths of Geology, *Annu. Rev. Earth Planet. Sci.*, 24(1), 385–432, doi:10.1146/annurev.earth.24.1.385.
- Silver, P. G., and W. W. Chan (1991), Shear Wave Splitting and Sub continental Mantle Deformation, *J. Geophys. Res.*, 96, 429–454, doi:10.1029/91JB00899.
- Silver, P. G., and M. K. Savage (1994), The interpretation of shear-wave splitting parameters in the presence of two anisotropic layers, *Geophys. J. Int.*, 119(3), 949–963, doi:10.1111/j.1365-246X.1994.tb04027.x.
- Walsh, E., R. Arnold, and M. K. Savage (2013), Silver and Chan revisited, *J. Geophys. Res. Solid Earth*, 118(September), 5500–5515, doi:10.1002/jgrb.50386.
- Wang, X., J. F. Ni, R. Aster, E. Sandvol, D. Wilson, C. Sine, S. P. Grand, and W. S. Baldrige (2008), Shear-Wave Splitting and Mantle Flow beneath the Colorado Plateau and Its Boundary with the Great Basin, *Bull. Seismol. Soc. Am.*, 98(5), 2526–2532, doi:10.1785/0120080107.

- Welch, P. D. (1967), The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms, *IEEE Trans. Audio Electroacoust.*, 15(2), 70–73, doi:10.1109/TAU.1967.1161901.
- Wolfe, C. J., and P. G. Silver (1998), Seismic anisotropy of oceanic upper mantle: Shear wave splitting methodologies and observations, *J. Geophys. Res.*, 103(B1), 749, doi:10.1029/97JB02023.
- Wüstefeld, A., G. Bokelmann, C. Zaroli, and G. Barruol (2008), SplitLab: A shear-wave splitting environment in Matlab, *Comput. Geosci.*, 34(5), 515–528, doi:10.1016/j.cageo.2007.08.002.
- Wüstefeld, A., G. Bokelmann, G. Barruol, and J. P. Montagner (2009), Identifying global seismic anisotropy patterns by correlating shear-wave splitting and surface-wave data, *Phys. Earth Planet. Inter.*, 176(3-4), 198–212, doi:10.1016/j.pepi.2009.05.006.
- Yuan, H., and B. Romanowicz (2010), Depth dependent azimuthal anisotropy in the western US upper mantle, *Earth Planet. Sci. Lett.*, 300(3-4), 385–394, doi:10.1016/j.epsl.2010.10.020.
- Zhang, S., and S. Karato (1995), Lattice preferred orientation of olivine aggregates deformed in simple shear, *Nature*, 375, 774–777, doi:10.1038/375774a0.

## CHAPTER 3

### 3.1 Future Work Recommendations

This chapter is oriented towards helping future graduate students, and includes information on downloading data from Canadian stations, using the new shear wave splitting code, and future research and development ideas.

During the initial stages of my research, I spent a fair amount of time on seismic data acquisition and culling bad events. The area of interest encompasses southwest Alaska and northwest Canada. The data downloaded will not be analyzed as part of this thesis, but will be safely stored to be used by students in the future. Thus, the first recommendation of future work regarding this project will be to update the data downloaded and to obtain shear wave splitting results with the method presented in this document. The new splitting results, with reliable uncertainties, should shed light on the seismic anisotropy pattern in the targeted area, and help constrain the mechanisms leading to past and present orogenic deformation. Moreover, there has not been a recent regional scale analysis of shear wave splitting in the area, and the new method will help in building a better understanding of lateral or depth variations in anisotropy.

The second recommendation for future work related to this project is to modify the code to accommodate off-vertical ray paths. Although most shear wave splitting studies assume a nearly vertical SKS ray-path, augmenting the method by considering the angle of incidence of the rays would remove a source of systematic error. It would also be a substantial step forwards towards considering dipping anisotropy and phases like S-waves from deep events that come in at angles significantly removed from vertical. This addition could be done with the use of a

functional fit to impulse responses for rays having a variety of ray parameters, and some preliminary work has been done along these lines.

Another line of work that can be followed to improve the code is to translate it to different programming languages. The code is written in R, which is a free platform-independent language that is frequently used in statistical studies, although less often in seismological studies. Hence, to encourage wide adoption of the method, it would be useful to have the code available also in Matlab and Python, two languages commonly used by the seismological community. An ideal complement would be to create an application or Matlab package to be used by researchers without any coding knowledge, similar to the Splitlab Software [*Wüstefeld et al.*, 2008].

The last recommendation, and most challenging but also most exciting, is to modify the code to account for complex anisotropy, i.e. two anisotropic layers, two layers with dipping axis, or one layer with a dipping axis. This implies some computational and mathematical challenges, but the groundwork is done and presented in this thesis.

Note that the standard errors that we are calculating with the bootstrapping approach seem to be very small (Figure 2.2 and Figure 2.3), and intuitively we believe an error might be present in the bootstrapping code. We are still working on this standard error calculation, and we aim to polish the bootstrapping approach before submitting Chapter Two as a paper to a peer-reviewed journal.

### **3.2 Data Retrieval Procedure**

As mentioned above, an initial goal of my research was to examine the mantle strain fabric in southwest Alaska and northwest Canada. However, the development of the new method and writing of code turned out to be a more involved process than initially envisioned.

Consequently my research was refocused on developing and implementing our new shear wave splitting method.

The data that was previously downloaded has been safely stored in our research group's server in order to be analyzed by graduate students down the line with the method presented in this document.

The location of the raw data is stored in the server path `/data/seismo/Raw`. There are three folders that correspond to data mentioned here: `/SOD_Alaska_Ana_SKS_May_2014`; `/SOD_Canada_Ana_SKS_May_2014`; and `/CNSN_Canada_Ana_SKS_March_2015`. The Alaskan data set was retrieved by means of the Standing Order for Data (SOD) software [Owens *et al.*, 2004]. Data from January the 1<sup>st</sup> 1999 to May the 26<sup>th</sup> 2014, within a magnitude range of 5.7 – 10, including all types of magnitudes, and with distance ranges within 85° and 105° was requested (see “Sumer2014AlaskaRecipe.xml”). A total number of 202 events were retrieved. The Canadian data set was more challenging to retrieve. Some events of the data set were retrieved also through SOD since there is some Canadian stations data available from the Incorporated Research Institutions for Seismology Data Management Center (IRIS DMC). Data from January the 1<sup>st</sup> 1999 to May the 26<sup>th</sup> 2014, within a magnitude range of 5.8 – 10, including all types of magnitudes, and with distance ranges within 85° and 105° was requested (see “Fall2014CanadaRecipe.xml”). A total number of 276 events were retrieved through this mechanism. However, data from the majority of seismic stations across Canada are not available from IRIS DMC. Therefore, data from those stations has to be retrieved directly from the Canadian National Data Centre (CNDC) by an automated data retrieval system based on electronic mail and on File Transfer Protocol (FTP) named the CNDC AutoDRM

(<http://earthquakecanada.nrcan.gc.ca/stndon/AutoDRM>). The AutoDRM system is based on the analogous system of the Swiss Seismological Service (<http://www.seismo.ethz.ch/prod/autodrm>).

Thus, there are two folders with data sets from Canada as mentioned above, one corresponding to data retrieved by SOD and other one corresponding to data retrieved from the CNDC AutoDRM.

Retrieving data from SOD is a very useful procedure that makes it easier to gather seismic data for different seismological studies. No further explanation of SOD and its use is presented here as it is well documented on the SOD website (<http://www.seis.sc.edu/sod>). However, the AutoDRM system is much more tricky to use and it does deserve some discussion. To retrieve data from AutoDRM we followed two different approaches. The first approach consisted of creating a Bash script that sends an email to [autodrm@seismo.nrcan.gc.ca](mailto:autodrm@seismo.nrcan.gc.ca) with the information concerning the event to retrieve. This script has two key components. The first is that it goes to sleep every 30 seconds in order not to clog the Canadian server, and the second is that the body text in the email has been crafted to follow the strict Canadian AutoDRM formatting requirements. Once the Canadian server receives the email, if the data exists for the requested event, they send an email back to the address provided by the user with the FTP URL to download the event file. Because having proper event information is essential, in order to search for SKS events in the area of interest, we created a Bash script (see “write\_time\_window\_table.bash”, Appendix A) that: (1) loops through a test file of the Canadian broadband stations within the region of study; (2) uses the “find\_events” command from SOD to identify events within that region that have the desired event parameters (years of occurrence, minimum and maximum magnitude, minimum and maximum distance range, and waveform phase of interest); (3) uses TauP [Crotwell *et al.*, 1999] to calculate the beginning and ending

time to make sure the phase of interest (e.g. SKS) is included in the length of the seismogram to be downloaded. However, this method takes too long to request a large data set since it was developed for requesting small data sets.

Hence, we created a new script, which also uses the previous one, to take a different approach (see “make\_CNSN\_table\_request.bash”). It consists of creating a table containing all of the events information to be requested, and sending the table directly to the peers in the Geological Survey of Canada Canadian National Seismograph Network (GSC CNSN). They can download the data set at once, pull it to the server and send the user the FTP address to retrieve it. It takes some days, potentially up to weeks, but this approach dramatically reduces the time needed to download larger datasets, and it is the approach chosen to download the Canadian data.

### **3.3 Data Processing Stream**

The standard, Silver and Chan [1991], method of shear wave splitting produces results that are dependent upon the method used for preprocessing data, which [Liu and Gao, 2013] try to minimize with a semiautomatic procedure of data selection, signal discrimination and data preprocessing. Ideally, a method would be robust with respect to changes in the seismogram that do not alter the signal, and the new method seems to be a step in that direction.

Here we suggest the following procedure for preparing data for analysis: (i) identify events that show a clear SKS signal in both the R and the T components; (ii) cut the events so the pre-SKS noise and the SKS signal have the same length (i.e. same number of time samples) - the same length condition is not required but it is a good work routine; (iii) locate the events in the folder where they are going to be analyzed with the R code. Some valuable steps to add to the preprocessing might be to check that there is not overlapping of the SKS phase with other

phases, and that there are no duplicate events. But there is no need to remove any trend on the data, to apply any tapering, or to employ any frequency filter.

Additionally, the Canadian data downloaded by means of the CNDC AutoDRM service needs to be transformed from SEED to SAC format, and then rotated into a radial-transverse reference frame.

Occasionally the R error, “`Error in readBin (...): invalid 'n' argument`”, is encountered. This is due to conditions associated with bad cut parameters using the Seismic Analysis Code (SAC) [Goldstein *et al.*, 2003; Goldstein and Snoke, 2005] tool named “`cut`”. SAC has another tool named “`cutter`” that can be used to control these types of errors, and the “`fillz`” option fills with zeros before the beginning of the file or after the ending of the file to account for the error. Thus, one should use these two SAC functions when parsing the data into signal and noise files.

Another R error occasionally encountered is “`D_Tstar [ , Lag2+1]: subscript out of bounds`”. This error is due to the frequency content and length of the data. To work around this error decrease the sampling interval (e.g. from 0.05 to 0.02). This can be done easily with another SAC command: “`interpolate`”. For reference, I usually try to use a sampling interval of 0.02 and R and T seismograms of 1500 time samples, of which the first 750 time samples correspond to the pre-SKS noise and the last 750 time samples correspond to the SKS signal. With a sampling interval of 0.02, 750 time samples equal 15 seconds long, so the total length of the R and T seismograms is 30 seconds.

### **3.4 Using the R code**

To run the R code you can do it either by using the user interface RStudio, or you can use R in the terminal typing in the command line “`R`”, assuming that the R language is installed in

your workstation. The R code consists of two R files: “Part\_1\_Setup\_Functions.R” (Appendix B) and “Part\_2\_Run\_Functions.R” (Appendix C). The first file contains all the functions needed to run the second file. Run the first file either in the terminal or in the Console in RStudio by “`source("Part1_Setup_Functions.R")`”. Note that your working directory must be the directory where you have the code (i.e. the two R files) and the data to be analyzed. Then, run the second file by “`source("Part2_Run_Functions.R")`”. Note that all the scripts are commented so you can see where to set up the variables specific to your data. The file “Part\_1\_Setup\_Functions.R” does not need to be modified, unless you want to use a different grid search size or modify some function. The file “Part\_2\_Run\_Functions.R” only needs to set up the sampling interval of the data, the back azimuth of the event, the event name, and to choose the length of the event desired to analyze the SKS splitting.

### **3.5 Location of Code and Notes**

All the scripts, the code, notes and thesis are located in the research group’s server under the path `/data/seismo`. Inside the folder `/Research/Ana` there is a folder named `/Bash_Scripts` that contains all the Bash scripts used in the development of this thesis and the code presented here. There is a folder named `/R_Scripts` that contains the R Code, and all of the R scripts used; and a folder named `/Matlab_Scripts` that contains all of the Matlab files used in this project; and a folder named `/SOD_html_recipes` that contains all of the html files used to retrieve seismic data from SOD.

All my research can be found in `/data/seismo/Ana/RESEARCH`.

This thesis is located in the path `/data/seismo/Documents/Ana`.

## REFERENCES

- Crotwell, H. P., T. J. Owens, and J. Ritsema (1999), The TauP Toolkit : Flexible Seismic Travel-time and Ray-path Utilities, *Seismol. Res. Lett.*, 70(2), 154–160, doi:10.1785/gssrl.70.2.154.
- Goldstein, P., and J. A. Snoke (2005), SAC Availability for the IRIS Community, *Inc. Res. Institutions Seismol. Newsl.*, 7, 6.
- Goldstein, P., D. Dodge, M. Firpo, and L. Minner (2003), SAC2000: Signal Processing and Analysis Tools for Seismologists and Engineers, *IASPEI Int. Handb. Earthq. Eng. Seismol.*, 81, 1613–1620.
- Liu, K. H., and S. S. Gao (2013), Making reliable shear-wave splitting measurements, *Bull. Seismol. Soc. Am.*, 103(5), 2680–2693, doi:10.1785/0120120355.
- Owens, T. J., H. P. Crotwell, C. Groves, and P. Oliver-Paul (2004), Sod: Standing Order for Data, *Seismol. Res. Lett.*, 75(4), 515–520, doi:10.1785/gssrl.75.4.515-a.
- Wüstefeld, A., G. Bokelmann, C. Zaroli, and G. Barruol (2008), SplitLab: A shear-wave splitting environment in Matlab, *Comput. Geosci.*, 34(5), 515–528, doi:10.1016/j.cageo.2007.08.002.

## APPENDICES

## APPENDIX A

```
#!/usr/bin/bash

#set -x

# write_time_window_table.bash
# This script creates a file suitable for submitting to the AutoDRM
service of the Geologic Survey of Canada's waveform database. Here are
some relevant links:

# http://www.earthquakescanada.nrcan.gc.ca/stndon/AutoDRM/index-
eng.php
# http://neic.usgs.gov/neis/AUTODRM/autodrm.help (the format to use)
# http://www.earthquakescanada.nrcan.gc.ca/stndon/AutoDRM/etiquette-
eng.php

# The goal of the script is to produce a file such as this, which will
be emailed to autodrm@seismo.nrcan.gc.ca

# BEGIN GSE2.0
# MSG_TYPE REQUEST
# MSG_ID test
# FTP address@colostate.edu
# TIME 2000/01/01 01:00:00.000 TO 2000/01/01 01:00:10.000
# STA_LIST CBB
# CHAN_LIST BH*
# WAVEFORM SEED
# OUT_FILE test_file
# STOP

# Notes
# 1) The OUTFILE command is unique to the GSC Autodrm.
# 2) The GSC suggests that power users (i.e. us) should wait 30-60
seconds between emails.

# Method
# 1) Awk the station file to get broadband stations (that have HH*
and BH* channels).
# 2) Loop through a table of station names and location. This can
be found at
http://www.earthquakescanada.nrcan.gc.ca/stndon/data\_avail-eng.php.
# 3) Use find_events with the -d option to find events that might
produce a reasonable SKS
# note: delta range for reasonable sks is about 85-125. Make a text
file containing these.
# 4) Loop through this event table and use taup to predict the
first arrival and SKS arrival times.
# 5) Use the 'date' command to convert the arrival times to
beginning and ending times that autodrm can use. Here $time_before and
$time_after can be used to adjust the time windows.
```

```

# 6) Write these into a format autoDRM can use. Also two event
information tables will be created.
# 7) email these.

# Basic example of how to email from bash:
# mail -s "hello" "example@example.com" <<EOF
# hello
# world
# EOF

# Assumptions
# 1) working_dir and table_dir below are real directories
# 2) The TauP Toolkit is working and in your path
(http://www.seis.sc.edu/downloads/TauP/)
# 3) You have the table of Canadian stations
(http://www.earthquakescanada.nrcan.gc.ca/stndon/data\_avail-eng.php)

#--- Files and Directories (make sure these exist)
working_dir="/home.local/anacorb/RESEARCH/CANADA/write_requests"
table_dir="/home.local/anacorb/RESEARCH/CANADA/write_requests" #
location of tables
drm_request_dir="/home.local/anacorb/RESEARCH/CANADA/write_requests/dr
m_requests"

station_file_root="canadian_stations"
station_file=$table_dir/canadian_stations.txt
event_file=$table_dir/events_all.txt

#--- User Defined Parameters
# define lat and long limit of stations
lon_w="-141" # This is the eastern boundary of AK
lon_e="-110"
lat_n="70"
lat_s="48" # This gets bottom of Vancouver Island

min_mag="6" # minimum magnitude to look for in find_events
begin_yr="1999" # earliest year for which to request data
time_before="60" # this is the time before the first arrival or
SKS used to indicate the beginning of the time request
time_after="120" # time after the predicted SKS

min_delta="85"
max_delta="105"

phase="SKS" # May also want to look for PKS and SKKS
#pause_time="35" # Only if sending email.
#Time in seconds to pause between sending out each email. Suggestion
is on website is between 30-60 seconds (Without the pause it will bung
up the system)

# Parameters needed for request file. Other ones will be set in the
script.
msg_id="Corbalan"

```

```

ftp_email="anacorb@lamar.colostate.edu"
autoDRM_email="autodrm@seismo.nrcan.gc.ca"

# first, let's awk the station file to get broadband stations within
our lat/lon region of interest. Will just search on the vertical
channal since I want the results to be unique by station name and
operational date. Note that it can be a little tricky to pass
variables to awk. One way is to use the -v flag before each variable.
awk -v LON_W=${lon_w} -v LON_E=${lon_e} -v LAT_N=${lat_n} -v
LAT_S=${lat_s} '{if ( ($2=="BHE") || ($2=="HHE") ) && ($4 < LON_E) &&
($4 > LON_W) && ($3 > LAT_S) && ($3 < LAT_N) ) printf "%10s %8s
%10.5f %10.5f %10.5f %15.4f \n", $1,$2,$3,$4,$5,$6}' $station_file >
$stable_dir/${station_file_root}.$phase.txt

# note that one way of doing a loop over text file, namely while read
stn_data;
# do
# find_event command
# done < find_events.txt
# doesn't work with find_events. It seems the problem is that
find_events is waiting for more command line input using the method
above

old_IFS=$IFS      # save the field separator
IFS=$'\n'        # new field separator, the end of line

# This will loop over the station file, and get data to be passed to
find_events.
for line in $(cat $stable_dir/${station_file_root}.$phase.txt); do
    echo "$line"
    sname=`echo $line | awk ' {print $1}'`
    schan=`echo $line | awk ' {print $2}'`
    slat=`echo $line | awk ' {print $3}'`
    slon=`echo $line | awk ' {print $4}'`
    selev=`echo $line | awk ' {print $5}'`
    sstartdate=`echo $line | awk ' {print $6}'`

    # get request channel, BH* or HH*.
    if [ "$schan" == "BHE" ];then
        request_chan="BH*"

    elif [ "$schan" == "HHE" ];then
        request_chan="HH*"
    fi

    echo " "
    echo "======"
    echo "finding good events for station $sname"
    echo " "
    starttime=$(date +%s)

    # get SKS events and put into a table
    echo "working on station $sname "

```

```

    echo "writing to $stable_dir/${sname}_events.txt "
    find_events -b $begin_yr -m $min_mag -d
    $slat/$slon/$min_delta/$max_delta -o '$event.getLatitude("##0.000;-
##0.000") $event.getLongitude("##0.000;-##0.000") $event.depth
$event.magnitude $event.time $event.getTime("yyyy_MM_dd_HH_mm_ss")' |
awk ' {printf " %10.5f %10.5f %8d %6s %5.2f %4s %12s %10s %5s %22s
\n", $1, $2, $3, $4, $5, $6, $7, $8, $9, $10}' > $stable_dir/${sname}_events.txt
    endtime=$(date +%s)
    time_sec=$((endtime-$starttime))
    time_min=`echo "$time_sec / 60" | bc`
    echo "event list for station $sname (find_events took $time_min
minutes)"
    num_events=`wc -l $stable_dir/${sname}_events.txt | awk '{print
$1}'`
    echo "$num_events events were found for station $sname "
    echo "$stable_dir/${station_file_root}.$phase.txt"
    sleep 10

# now loop through events table and get beginning and ending
times, then write files to email
for event_data in $(cat $stable_dir/${sname}_events.txt); do
    echo "      -----"
    echo "      getting arrival time data for event of magnitude
$emag on date $edate"
    elat=`echo $event_data | awk '{print $1}'`
    elon=`echo $event_data | awk '{print $2}'`
    edepth=`echo $event_data | awk '{print $3}'`
    emag=`echo $event_data | awk '{print $5}'`
    edate=`echo $event_data | awk '{print $7}'`
    etime=`echo $event_data | awk '{print $8}'`
    ename=`echo $event_data | awk '{print $10}'`
    taup_time -h $edepth -sta $slat $slon -evt $elat $elon >
/tmp/taup_output.txt

    # having delta might be useful
    delta=`awk 'NR==6 {print $1}' /tmp/taup_output.txt`

    #---get time range to request data.
    # get the first arrival
    first_arrival_time=`awk 'NR==6 {print $4}'
/tmp/taup_output.txt`which

    # get the SKS arrival
    SKS_arrival_time=`awk '$3 ~ /SKS/ {print $4}'
/tmp/taup_output.txt`

    # Use arrival times to get the beginning and end times of the
trace to request
    b=`echo "$first_arrival_time - $time_before" | bc`
    e=`echo "$SKS_arrival_time + $time_after" | bc`

    # Use the date command to get the actual times in a format
that autoDRM likes.  Change the date commands from "date" to "gdate"

```

```

after installing coreutils (see comments at start of script) if you
are doing this on a mac.
    begin_time=`date --date "$edate $etime $b seconds"`
    end_time=`date --date "$edate $etime $e seconds"`

    begin_time_formatted=`date --date "$begin_time" +"%Y/%m/%d
%T.000"`
    end_time_formatted=`date --date "$end_time" +"%Y/%m/%d
%T.000"`

    # TIME 2000/01/01 01:00:00.000 TO 2000/01/01 01:00:10.000
    # make text_file to send via email to the autoDRM sever
    fname="${sname}_${ename}.drmrequest.txt"
    echo "BEGIN GSE2.0" > ${drm_request_dir}/${fname}
    echo "MSG_TYPE REQUEST" >> ${drm_request_dir}/${fname}
    echo "MSG_ID $msg_id " >> ${drm_request_dir}/${fname}
    echo "FTP $ftp_email" >> ${drm_request_dir}/${fname}
    echo "TIME $begin_time_formatted to $end_time_formatted">>
${drm_request_dir}/${fname}
    echo "OUT_FILE ${sname}_${ename}.seed" >>
${drm_request_dir}/${fname}
    echo "STA_LIST $sname" >> ${drm_request_dir}/${fname}
    echo "CHAN_LIST $request_chan" >>
${drm_request_dir}/${fname}
    echo "WAVEFORM SEED" >> ${drm_request_dir}/${fname}
    echo "STOP" >> ${drm_request_dir}/${fname}
done
done

```

## APPENDIX B

```
#####  
#####  
#####  
# Set up manually the function rsac to open and read sac files in R  
rsac<- function(files, endian = .Platform$endian)  
{  
  if(length(endian) == 1 & length(files) > 1)  
    endian <- rep(endian, length(files))  
  n <- length(files)  
  data <- vector(mode = "list", length = n)  
  for(i in 1:n)  
  {  
    file <- files[i]  
    zz <- file(file, "rb")  
    h1 <- readBin(con = zz, what = numeric(), n = 70, size = 4,  
                 endian = endian[i])  
    dim(h1) <- c(5, 14)  
    h1 <- aperm(h1)  
    # NA values:  
    h1[h1 == -12345] <- NA  
    h2 <- readBin(con = zz, what = integer(), n = 35, size = 4,  
                 endian = endian[i])  
    dim(h2) <- c(5, 7)  
    h2 <- aperm(h2)  
    # NA values:  
    h2[h2 == -12345] <- NA  
    h3 <- readBin(con = zz, what = logical(), n = 5, size = 4,  
                 endian = endian[i])  
    h4 <- readBin(con = zz, what = character(), n = 1, size = 4,  
                 endian = endian[i])  
    # Define header variables:  
    dt <- h1[1, 1]  
    depmin <- h1[1, 2]  
    depmax <- h1[1, 3]  
    scale <- h1[1, 4]  
    odelta <- h1[1, 5]  
    b <- h1[2, 1]  
    e <- h1[2, 2]  
    o <- h1[2, 3]  
    a <- h1[2, 4]  
    f <- h1[5, 1]  
    stla <- h1[7, 2]  
    stlo <- h1[7, 3]  
    stel <- h1[7, 4]  
    stdp <- h1[7, 5]  
    evla <- h1[8, 1]  
    evlo <- h1[8, 2]  
    evel <- h1[8, 3]  
    evdp <- h1[8, 4]
```

```

mag <- h1[8, 5]
dist <- h1[11, 1]
az <- h1[11, 2]
baz <- h1[11, 3]
gcarc <- h1[11, 4]
cmpaz <- h1[12, 3]
cmpinc <- h1[12, 4]
nzyear <- h2[1, 1]
nzjday <- h2[1, 2]
nzhour <- h2[1, 3]
nzmin <- h2[1, 4]
nzsec <- h2[1, 5]
nzmsec <- h2[2, 1]
norid <- h2[2, 3]
nevid <- h2[2, 4]
N <- h2[2, 5]
idep <- h2[4, 2]
iztype <- h2[4, 3]
leven <- h3[1]
lpspol <- h3[2]
kstnm <- substr(h4, 1, 8)
kstnm <- sub("-12345", "      ", kstnm)
kevm <- substr(h4, 9, 24)
kevm <- sub("-12345", "      ", kevm)
khole <- substr(h4, 25, 32)
khole <- sub("-12345", "      ", khole)
ko <- substr(h4, 33, 40)
ko <- sub("-12345", "      ", ko)
ka <- substr(h4, 41, 48)
ka <- sub("-12345", "      ", ka)
kcmpnm <- substr(h4, 161, 168)
kcmpnm <- sub("-12345", "      ", kcmpnm)
knetwork <- substr(h4, 169, 176)
knetwork <- sub("-12345", "      ", knetwork)
kinst <- substr(h4, 185, 192)
kinst <- sub("-12345", "      ", kinst)
seek(con = zz, where = 632)
x <- readBin(con = zz, what = numeric(), n = N,
             size = 4, endian = endian[i])
close(zz)
data[[i]] <- list(amp = x, dt = dt, depmin = depmin, depmax =
depmax,
                 scale = scale, odelta = odelta,
                 b = b, e = e, o = o, a = a, f = f,
                 stla = stla, stlo = stlo, stel = stel, stdp =
stdp,
                 evla = evla, evlo = evlo, evel = evel, evdp =
evdp,
                 mag = mag, dist = dist, az = az, baz = baz,
                 cmpaz = cmpaz, cmpinc = cmpinc,
                 nzyear = nzyear, nzjday = nzjday, nzhour =
nzhour,

```

```

        nzmin = nzmin, nzsec = nzsec,
        nzmsec = nzmsec, norid = norid,
        nevid = nevid, N = N,
        units = idep, iztype = iztype,
        leven = leven, lspol = lspol,
        sta = kstnm, kevm = kevm, khole = khole,
        ko = ko, ka = ka,
        comp = kcmpnm, knetwork = knetwork, kinst =
kinst)
    }
    class(data) <- "rsac"
    invisible(data)
}
## Code: Thompson, E. M. and Lees, J. M. \code{Rsac}
#####
#####
#####
preprocess<-function(R,T,sigma=-9,difference=TRUE){
  # Center the data by subtracting off the mean
  if(difference){
    R<-diff(R,1,1)
    T<-diff(T,1,1)
  }# end if difference
  R<- R-mean(R)
  T<- T-mean(T)
  if(sigma<0){
    sigma_R<-sqrt(var(R))
    sigma_T<-sqrt(var(T))
    sigma<-(sigma_R+sigma_T)/2
  }# end if negative sigma
  R<-R/sigma
  T<-T/sigma
  out<-list(R=R,T=T,sigma=sigma)
  return(out)
}
#####
#####
#####
#remove all the effects of pre-processing the time series
unpreprocess<-function(R,T,Sigma=1,undifference=TRUE){
  if(undifference){
    R <- cumsum(R)
    T <- cumsum(T)
  }
  R<- R+mean(R)
  T<- T+mean(T)
  R <- Sigma*R
  T <- Sigma*T
  out<-list(R=R,T=T)
  return(out)
}
#####
#####

```

```

#####
# Function to compute the Fourier frequencies.
Fourier_Freq<-function(n){
  Ln<- floor((n-1)/2) # this means take the integer part
  Un<-floor(n/2)
  Fn<- -Ln:Un
  omega<-2*pi*Fn/n
  return(omega)
}
#####
#####
#####
# Periodogram/(2*pi) function.
Pgram<-function(y){
  n<-length(y)
  Ln<- floor((n-1)/2) # this means take the integer part
  Un<-floor(n/2)
  Fn<- -Ln:Un
  # dft in R does not divide by n; need to do that here.
  #
  dft<-fft(y)/sqrt(n*2*pi)
  #
  # fft orders these (0,2pi) instead of (-pi,pi),
  # so (pi,2pi) needs to map to (-pi,0).
  #
  tmp<-dft
  dft[1:Ln]<-tmp[(Un+2):n]
  dft[(Ln+1):(Ln+Un+1)]<-tmp[1:(Un+1)]
  #
  I_y<-Re(dft*Conj(dft))
  return(I_y)
}
#####
#####
#####
fit_p spline<-function(I,omega,K=90,eps=0.005,penalty=1){
  # Need to delete frequency zero.
  zero<-(omega==0)
  I<-I[!zero]
  omega<-omega[!zero]
  knots<-log(eps)+(1:K)*(log(pi-eps)-log(eps))/K
  knots<-c(-rev(exp(knots)),exp(knots))
  ZZ<-outer(omega,knots,"-")
  ZZ<-ZZ*(ZZ>0)
  ZZ<-cbind(rep(1,length(omega)),omega,ZZ)
  fit<-
solve(t(ZZ)%*%ZZ+diag(c(0,0,rep(penalty,2*K))))%*%t(ZZ)%*%cbind(log(I)
)
  return(fit)
}
#####
#####
#####

```

```

eval_pspline<-function(fit,omega,K=90,eps=0.005,penalty=1){
  knots<-log(eps)+(1:K)*(log(pi-eps)-log(eps))/K
  knots<-c(-rev(exp(knots)),exp(knots))
  ZZ<-outer(omega,knots,"-")
  ZZ<-ZZ*(ZZ>0)
  ZZ<-cbind(rep(1,length(omega)),omega,ZZ)
  eval<-ZZ%*%cbind(c(fit))
  return(eval)
}
#####
#####
#####
# Matrix of dft's.
DFT_Matrix<-function(y,Max_Lag){
  D<-c()
  N<-length(y)
  for(g in 0:Max_Lag){
    D<-cbind(fft(y[(g+1):(g+N-Max_Lag)]),D)
  }

  # Reorder D for consistency with periodogram.
  # fft orders these (0,2pi) instead of (-pi,pi),
  # so (pi,2pi) needs to map to (-pi,0).

  n<-N-Max_Lag
  Ln<- floor((n-1)/2) # this means take the integer part
  Un<-floor(n/2)
  tmp<-D
  D[1:Ln,]<-tmp[(Un+2):n,]
  D[(Ln+1):(Ln+Un+1),]<-tmp[1:(Un+1),]
  return(D)
}
#####
#####
#####
# Evaluate the Whittle log-likelihood
Whittle_fast<-
function(phi,DT,D_Rstar,D_Tstar,ln_f_R,ln_f_T,Omega,sample_interval=0.
02){
  Max_Lag<-floor(4/sample_interval)
  N<-dim(D_Tstar)[1]
  SS<-(1:N)[Omega>0]
  lag<-DT/sample_interval
  Lag1<-floor(lag)
  psi<-2*pi*phi/360 # convert to radians
  R1<-cos(psi)^2
  R2<-sin(psi)^2
  T1<-cos(psi)*sin(psi)
  T2<--T1
  rho2<-(R1^2+R2^2)+2*R1*R2*cos(Lag1*Omega)
  tau2<-(T1^2+T2^2)+2*T1*T2*cos(Lag1*Omega)
  spectrum_cross1<-rho2*exp(ln_f_T)+tau2*exp(ln_f_R)
}

```

```

    tmp_real<-R1*Re(D_Tstar[,1])+R2*Re(D_Tstar[,Lag1+1])-
T1*Re(D_Rstar[,1])-T2*Re(D_Rstar[,Lag1+1])
    tmp_imag<-R1*Im(D_Tstar[,1])+R2*Im(D_Tstar[,Lag1+1])-
T1*Im(D_Rstar[,1])-T2*Im(D_Rstar[,Lag1+1])
    I_cross1<-(tmp_real^2+tmp_imag^2)/(2*pi*N)
    Lag2<-Lag1+1
    I_cross2<-0
    spectrum_cross2<-1
    if(lag>Lag1){
        tmp_real<-R1*Re(D_Tstar[,1])+R2*Re(D_Tstar[,Lag2+1])-
T1*Re(D_Rstar[,1])-T2*Re(D_Rstar[,Lag2+1])
        tmp_imag<-R1*Im(D_Tstar[,1])+R2*Im(D_Tstar[,Lag2+1])-
T1*Im(D_Rstar[,1])-T2*Im(D_Rstar[,Lag2+1])
        I_cross2<-(tmp_real^2+tmp_imag^2)/(2*pi*N)
        rho2<-(R1^2+R2^2)+2*R1*R2*cos(Lag2*Omega)
        tau2<-(T1^2+T2^2)+2*T1*T2*cos(Lag2*Omega)
        spectrum_cross2<-rho2*exp(ln_f_T)+tau2*exp(ln_f_R)
    }
    like1<--sum(I_cross1/spectrum_cross1)-sum(log(spectrum_cross1))
    like2<--sum(I_cross2/spectrum_cross2)-sum(log(spectrum_cross2))
    log_like<-(Lag2-lag)*like1+(lag-Lag1)*like2
    return(log_like)
}
#####
#####
#####
# Plot a criterion surface
Surface_Plot<-function(DT,phi,criterion){
    G<-length(DT)
    h<-topo.colors(G)
    plot(range(DT),range(phi),type="n",xlab="Split Time",ylab="Fast
Axis")
    r<-rank(criterion)
    delta_DT<-max(diff(DT,1,1))
    delta_phi<-max(diff(abs(phi),1,1)) # can be messed up depending
on order of grid values
    for(i in 1:G){

        rect(DT[i],phi[i],DT[i]+delta_DT,phi[i]+delta_phi,col=h[r[i]],bor
der=NA)
    }# end loop on i
    g_max<-(1:G)[r==G]
    points(DT[g_max],phi[g_max],pch=17,col="red",cex=1.2)
}
#####
#####
#####
#Save the values of DT and phi in a list
Save_results<-function(DT,phi,criterion){
    G<-length(DT)
    r<-rank(criterion)
    g_max<-(1:G)[r==G]
    save<-list(DTm=DT[g_max],phim=phi[g_max])
}

```

```

    return(save)
}
#####
#####
#####
# Create lookup table using Menke and Levin one-layer functional
forms.
# Change manually the maximum and minimum dt & maximum and minimum
fast axis.
# Change sqrt_G for finer or coarser grid
Lookup_Table<-
function(sqrt_G=200,scale=1,max_DT=4.0,min_DT=0,max_phi=180,min_phi=0,
sample_interval=0.01,back_azimuth=baz){
  fast_axis<-rep((min_phi+((0:(sqrt_G-1))/sqrt_G)*(max_phi-
min_phi)),sqrt_G)
  delta<-fast_axis-back_azimuth
  delta<-2*pi*delta/360
  DT<-sort(rep((min_DT+((0:(sqrt_G-1))/sqrt_G)*(max_DT-
min_DT)),sqrt_G))
  R1<-scale*cos(delta)^2
  R2<-scale*sin(delta)^2
  T1<-scale*cos(delta)*sin(delta)
  T2<--T1
  Lag<-(DT/sample_interval)%/%1
  out<-
list(fast_axis=fast_axis,DT=DT,R1=R1,R2=R2,T1=T1,T2=T2,Lag=Lag)
  return(out)
}
#####
#####
#####
# Set up manually the function SurrogateData from WaveletComp package
SurrogateData <-
function(x, method = "white.noise",
        params=list(AR      = list(p=1),
                    ARIMA  = list(p=1, q=1,
                                meboot = list(trim = 0.1,
                                                force.clt = F, expand.sd = T, fiv = 5)
                    )
        ){
  if(method == "white.noise") x.sur <- rnorm(length(x))
  if(method == "shuffle")     x.sur <- sample(x, length(x))
  if(method == "Fourier.rand") x.sur <- FourierRand(x)

  if(method == "AR")
  {
    x.sur <- AR(x, params = params)
  }
}

```

```

#   if(method == "meboot")           {
#
#       trim      = params$meboot$trim
#       force.clt = params$meboot$force.clt
#       expand.sd  = params$meboot$expand.sd
#       fiv       = params$meboot$fiv
#
#       x.sur <- meboot(x, reps=2, trim = trim, force.clt = force.clt,
expand.sd = expand.sd, fiv = fiv)$sensemble[,1]
#
#   }

  if(method == "ARIMA")           {

      x.sur <- ARIMA(x, params = params)

  }

  return(invisible(x.sur))
}
## Code: Tian, H. and Cazelles, B., \code{WaveletCo}
#####
#####
#####
# Set up manually the function FourierRand from WaveletComp package
# Needed for using SurrogateData with the Fourier randomization method
FourierRand <-
function(x){
  n <- length(x)
  z <- fft(x)

  if(n%%2 == 0){
    ph <- 2*pi*runif(n/2-1)
    ph <- c(0, ph, 0, -rev(ph))}

  if(n%%2 != 0){
    ph <- 2*pi*runif((n-1)/2)
    ph <- c(0, ph, -rev(ph))}

  ph <- complex(imaginary = ph)
  z <- z * exp(ph)
  x.sur <- Re(fft(z, inverse = TRUE)/n)

  return(invisible(x.sur))
}
## Code: Tian, H. and Cazelles, B., \code{WaveletCo}
#####
#####
#####
# Inputs:
# Rstar: N-vector of radial event data. N must be even.
# Tstar: N-vector of tangential event data. N must be even.
# L: estimated lag (a single integer).

```

```

# rho: 2-vector of estimated rho[1], rho[2]
# tau: 2-vector of estimated tau[1], tau[2]
# fR=fit_R: earlier result from fit_pspline
# fT=fit_T: earlier result from fit_pspline
estimate_S_star<-
function(Rstar,Tstar,L=120,rho=c(cos(pi/9)^2,sin(pi/9)^2),tau=c(cos(pi
/9)*sin(pi/9),-cos(pi/9)*sin(pi/9)),fR=fit_R,fT=fit_T){
  N<-length(Rstar)
  OMEGA<-sort(abs(Fourier_Freq(N)))

# Compute orthonormal matrix for diagonalization.
  PP<-rbind((1/sqrt(N))*rep(1,N))
  for(j in 1:((N-2)/2)){
    tmp<-OMEGA[2*j]*(1:(N-1)) # 0 is in element 1, evens have
OMEGA_1,OMEGA_2,...
    c_j<-sqrt(2/N)*c(1,cos(tmp))
    s_j<-sqrt(2/N)*c(0,sin(tmp))
    PP<-rbind(PP,c_j,s_j)
  }
  tmp<-OMEGA[N]*(1:(N-1))
  c_j<-sqrt(1/N)*c(1,cos(tmp))
  PP<-rbind(PP,c_j) #c_j is in the even rows; s_j is in the odd rows

# Evaluate the spectral densities on the re-ordered OMEGA.
  LN_f_R<-eval_pspline(fR,OMEGA,penalty=0.5)
  LN_f_T<-eval_pspline(fT,OMEGA,penalty=0.5)

# Use diagonalization of autocovariance matrix to get generalized
least squares estimators
# of S_star, by multiplying both sides by inverse square root of
Gamma_R and Gamma_T:
  inverse_sqrt_Gamma_T<-t(PP)%*%diag(c(exp(-
0.5*LN_f_T))%*%PP*(1/sqrt(2*pi))
  inverse_sqrt_Gamma_R<-t(PP)%*%diag(c(exp(-
0.5*LN_f_R))%*%PP*(1/sqrt(2*pi))
  Z_R<-matrix(0,N,N+L)
  Z_T<-matrix(0,N,N+L)
  for(j in 1:N){
    Z_T[j,j+L]<-tau[1]
    Z_T[j,j]<-tau[2]
    Z_R[j,j+L]<-rho[1]
    Z_R[j,j]<-rho[2]
  }
  YT<-inverse_sqrt_Gamma_T%*%cbind(Tstar)
  YR<-inverse_sqrt_Gamma_R%*%cbind(Rstar)
  Y<-rbind(YR,YT)
  Z<-rbind(inverse_sqrt_Gamma_R%*%Z_R,inverse_sqrt_Gamma_T%*%Z_T)
  S_star_hat<-lm(Y~-1+Z)$coef # estimate of the unpre-processed signal
  return(S_star_hat)
}
#####
#####
#####

```

## APPENDIX C

```
##### First, run latest version of Setup_Functions.R
(Part_1_Setup_Functions_Complete.R) #####
library(foreach) # foreach package installation is needed.
library(doParallel) # doParallel package installation is needed. It is
a "parallel backend" for the foreach package using the %dopar%
function. By default, doParallel uses multicore functionality on Unix-
like systems and snow functionality on Windows.
registerDoParallel(cores=4) # to use multicore-like functionality we
need to stablish the number of cores we are going to use.

ptotaltime<-system.time({

## 1ST.- Calculate the estimated values of {DT,phi} running the code
over the data (for a single event with known BAZ)##
#####
#####
#####
#####

# Set up the variables needed:
# Need to change the sampling interval (si), the back azimuth (baz),
the event name (event_name),
# the radial and transverse files to use (Rsacfile, Tsacfile), and
their desired length (R, T)
si<-0.02
baz<-(266)
event_name<-c("1995043717.TS.PAS")
lu<-Lookup_Table(sample_interval=si,back_azimuth=baz)

Rsacfile<-rsac('1995043717.TS.PAS.BHR.sac', endian = .Platform$endian)
Tsacfile<-rsac('1995043717.TS.PAS.BHT.sac', endian = .Platform$endian)
Rsacfile<-lapply(Rsacfile,"[","amp") # Rnoi<-Rnoise[[1]]$amp
Tsacfile<-lapply(Tsacfile,"[","amp") # Tnoi<-Tnoise[[1]]$amp
Rsacfile<-unlist(Rsacfile)
Tsacfile<-unlist(Tsacfile)

R<-Rsacfile[1800:3300] #set up the length manually to get correctly
the pre-SKS event and the event below
T<-Tsacfile[1800:3300] #set up the length manually to get correctly
the pre-SKS event and the event below

#####
#####

Ri<-R[0:750] # Radial pre-SKS noise
Ti<-T[0:750] # Tangential pre-SKS noise

outi<-preprocess(Ri,Ti,difference=TRUE)
Ri_p<-outi$R
```

```

Ti_p<-outi$T
ni<-length(Ri_p)
omegai<-Fourier_Freq(ni)

I_Ri<-Pgram(Ri_p)
fit_Ri<-fit_pspline(I_Ri,omegai,penalty=0.5)
omi<-omegai[omegai>0]
ln_f_Ri<-eval_pspline(fit_Ri,omi,penalty=0.5)

I_Ti<-Pgram(Ti_p)
fit_Ti<-fit_pspline(I_Ti,omegai,penalty=0.5)
ln_f_Ti<-eval_pspline(fit_Ti,omi,penalty=0.5)

#####
#####

Rstari<-R[751:1500] # Radial SKS signal
Tstari<-T[751:1500] # Tangential SKS signal

out_stari<-preprocess(Rstari,Tstari,sigma=outi$sigma,difference=TRUE)
Rstari_p<-out_stari$R
Tstari_p<-out_stari$T
MLagi<-max(lu$Lag)
Ni<-length(Rstari_p)-MLagi
Omegai<-Fourier_Freq(Ni)

ln_f_Ri<-eval_pspline(fit_Ri,Omegai,penalty=0.5)
ln_f_Ti<-eval_pspline(fit_Ti,Omegai,penalty=0.5)

#####
#####

D_Rstari<-DFT_Matrix(Rstari_p,Max_Lag=MLagi)
D_Tstari<-DFT_Matrix(Tstari_p,Max_Lag=MLagi)

#####
#####

Gi<-length(lu$DT)
likei<-rep(0,Gi)
SSi<-(1:Ni)[Omegai>0]
for(g in 1:Gi){
  likei[g]<-Whittle_fast(lu$fast_axis[g]-
baz,lu$DT[g],D_Rstar=D_Rstari,D_Tstar=D_Tstari,ln_f_R=ln_f_Ri,ln_f_T=ln_f_Ti,Omega=Omegai,sample_interval=si)
}

pdf(paste('log_likelihood_for_ev_',event_name,'.pdf',sep=""))
Surface_Plot(lu$DT,lu$fast_axis,likei)
title(main=paste("Log-Likelihood for ",event_name))
dev.off()

save<-Save_results(lu$DT,lu$fast_axis,likei)

```

```

est_DT<-c(round(save$DTm,2)) # estimated value of DT with the code
based on the cross-convolution method
est_phi<-c(round(save$phim,2)) # estimated value of phi with the code
based on the cross-convolution method

list<-cbind(est_DT,est_phi)
write.table(list,
file=(paste('Splitting_Results_for_ev_',event_name,'.txt',sep="")),
sep="\t", row.names=FALSE, col.names=FALSE)

#####
#####
#####
#####
#####
#####

## 2ND.- Create synthetic signal "noise free" seismogram with the
estimated values of {DT,phi} calculated on the 1st step##
#####
#####
#####
#####

# Set up various parameters
split_time<-est_DT
fast_axis<-est_phi
top<-fast_axis-baz # angle between fast axis and baz in top layer
top<-2*pi*top/360

# Set up the amplitudes of impulses (from Levin and Menke)
Rho1<-cos(top)^2
Rho2<-sin(top)^2

Tau1<-cos(top)*sin(top)
Tau2<--Tau1

# Estimate the true signal that produced the data to use it as the
input wavelet
Lag<-floor(split_time/si)
Rsig<-R[-1]
Tsig<-T[-1]

true_signal<-
estimate_S_star(Rsig,Tsig,L=Lag,rho=c(Rho1,Rho2),tau=c(Tau1,Tau2),fR=f
it_Ri,fT=fit_Ti)

# Convolve the Menke & Levin coefficients with the input wavelet and
generate the synthetic seismograms
N<-length(true_signal)

St<-c(rep(0,Lag),true_signal)
RDATA<-rep(0,N)

```

```

TDATA<-RDATA

for(i in (1+Lag):(N+Lag)){
  RDATA[i]=Rho1*St[i]+Rho2*St[i-Lag]
  TDATA[i]=Tau1*St[i]+Tau2*St[i-Lag]
}

#####
#####
#####
#####
#####
#####

## 3RD.- Bootstrapping the code with the synthetic noise free
seismogram created on the 2nd step and the noise file from 1st step##
#####
#####
#####
#####

# Cut and scale the signal to the same length and same variance than
the real signal (Rstari)
Rstar_in<-RDATA[802:1550] # Be careful with this length! Make sure it
encloses the signal for doing correctly the bootstrapping.
Tstar_in<-TDATA[802:1550] # Be careful with this length! Make sure it
encloses the signal for doing correctly the bootstrapping.

#lu<-Lookup_Table(sample_interval=si,back_azimuth=back_azimuth)
B<-100 # number of times we are going to loop over the code - number
of desired simulations

#####
#####
#Randomize noise to add to the simulations of synthetic data

Rb_p<-c()
R2b_p<-c()
Tb_p<-c()
T2b_p<-c()

Rb_up<-c()
R2b_up<-c()
Tb_up<-c()
T2b_up<-c()

for(b in 1:B){

  Rb_p<-cbind(Rb_p,c(SurrogateData(Ri_p, method = "Fourier.rand")))
  R2b_p<-cbind(R2b_p,c(SurrogateData(Ri_p, method = "Fourier.rand")))

  Tb_p<-cbind(Tb_p,c(SurrogateData(Ti_p, method = "Fourier.rand")))
  T2b_p<-cbind(T2b_p,c(SurrogateData(Ti_p, method = "Fourier.rand")))
}

```

```

out<-
unpreprocess(Rb_p[,b],Tb_p[,b],Sigma=outi$sigma,undifference=TRUE)
Rb_up<-cbind(Rb_up,c(out$R))
Tb_up<-cbind(Tb_up,c(out$T))

out<-
unpreprocess(R2b_p[,b],T2b_p[,b],Sigma=outi$sigma,undifference=TRUE)
R2b_up<-cbind(R2b_up,c(out$R))
T2b_up<-cbind(T2b_up,c(out$T))

}

#####
#####
#####
#####
#LOOP STARTS HERE
# Loop over the code to get the values of DT and phi for different
repeated simulations of synthetic data foreach iterates over the
variables in parallel

ptime<-system.time({

  Results<-foreach(i=1:ncol(Rb_up), j=1:ncol(Tb_up), k=1:ncol(R2b_up),
l=1:ncol(T2b_up)) %dopar% {

    out<-preprocess(Rb_up[,i],Tb_up[,j],difference=TRUE)
    R_out<-out$R
    T_out<-out$T
    n<-length(R_out)
    omega<-Fourier_Freq(n)
    I_R<-Pgram(R_out)
    fit_R<-fit_pspline(I_R,omega,penalty=0.5)
    om<-omega[omega>0]
    ln_f_R<-eval_pspline(fit_R,om,penalty=0.5)

    I_T<-Pgram(T_out)
    fit_T<-fit_pspline(I_T,omega,penalty=0.5)
    ln_f_T<-eval_pspline(fit_T,om,penalty=0.5)

#####
#####

    Rstar<-Rstar_in + R2b_up[,k]
    Tstar<-Tstar_in + T2b_up[,l]

    out_star<-preprocess(Rstar,Tstar,sigma=out$sigma,difference=TRUE)
    Rstar_out<-out_star$R
    Tstar_out<-out_star$T
    MLag<-max(lu$Lag)
    N<-length(Rstar_out)-MLag
    Omega<-Fourier_Freq(N)

```

```

ln_f_R<-eval_pspline(fit_R,Omega,penalty=0.5)
ln_f_T<-eval_pspline(fit_T,Omega,penalty=0.5)

#####
#####

D_Rstar<-DFT_Matrix(Rstar_out,Max_Lag=MLag)
D_Tstar<-DFT_Matrix(Tstar_out,Max_Lag=MLag)

#####
#####

G<-length(lu$DT)
like<-rep(0,G)
SS<-(1:N)[Omega>0]
for(g in 1:G){
  like[g]<-Whittle_fast(lu$fast_axis[g]-
baz,lu$DT[g],D_Rstar=D_Rstar,D_Tstar=D_Tstar,ln_f_R=ln_f_R,ln_f_T=ln_f
_T,Omega=Omega,sample_interval=si)
}

#####
#####

save<-Save_results(lu$DT,lu$fast_axis,like)
results<-c(save$DTm,save$phim)

#####
#####

#####
#####

#####
#####

#####
#####

#LOOP ENDS HERE
}
Results<-t(round(matrix(unlist(Results),2,B),2))
write.table(Results,
file=(paste('Results_for_ev_',event_name,'.txt',sep="")), sep="\t",
row.names=FALSE, col.names=FALSE)
}) # ptime
}) # ptotaltime

```