UNIVERSITY OF CALIFORNIA, SAN DIEGO

Improving Earthquake Source Spectrum Estimation using Multitaper Techniques

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in
Earth Sciences

by

Germán A. Prieto

Committee in charge:

Professor Frank L. Vernon, Co-Chair
Professor Peter M. Shearer Co-Chair
Professor Joel Conte
Professor Bruce Cornuelle
Professor Robert L. Parker
Professor David J. Thomson

2007

The dissertation of Germán A. Prieto is approved, and it is acceptable in quality and form for publication on microfilm:

_____

_____

_____

_____

_____
Co-Chair

_____
Co-Chair

University of California, San Diego

2007

To my beloved wife and best friend, Carolina

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to start by thanking Carolina for her constant support throughout this process. Her love and friendship have been fundamental and have helped me become a better person. Even though nobody believes it, she is the one who has to listen to me talking when we are by ourselves. Thank you for being such a patient editor. I will always love you.

I thank my two advisors, Frank Vernon and Peter Shearer. They were always supportive, guided me through my research and my life as a graduate student. Discussions included seismology and time series analysis, but also my future in academia, what good science is, tennis, etc. They gave me enough independence to develop my research which helped me in becoming a better scientist. Thank you for being there to answer all my questions.

I was very lucky to have the opportunity to interact with Robert L. Parker and David J. Thomson and to have them as part of my committee. They are among the best scientists in inverse theory and data analysis and I have learned a lot from each one of them. Given their different ways to solve problems and points of view, discussion with them gave me a very wide view of problem solving techniques, something I could never have had anywhere else. Joel Conte and Bruce Cornuelle were also very supportive of my work and provided significant comments, ideas and points of view on where the project could be improved.

I would also like to thank the large number of scientists with whom I had the honor to talk to and even get their reviews for some of my papers. Special thanks to Greg Beroza, Kevin Mayeda and Bill Walters, Rachel Abercrombie, Ralph Archuleta, Jim Brune, Hiroo Kanamori and Luis Rivera.

For two quarters, I was a teaching assistant for Gabi Laske's Natural Disasters class. Her class allowed me to experience teaching and gave me the assurance to continue in academia.

I am grateful to my family. My parents showed me love for science and education and were always supportive of my decisions. My sister Helena and my brother-in-law helped us start a life in San Diego and have been great family and friends.

Special thanks to Debi Kilb for her trust and for triggering my interest in

earthquake source physics with her willingness to discuss crustal seismology. I also thank my colleagues and classmates at IGPP for lunch hours, coffee and chat time, and our good friend, Mónica Pachón for all her energy.

Some of the chapters are reformatted version of papers that appeared in the geophysical literature. Chapter 2 (G.A. Prieto, D.J. Thomson, F.L. Vernon, P.M. Shearer and R.L. Parker., 2007, Confidence intervals of earthquake source parameters., *Geophys. J. Int.*, 168, p. 1227–1234., doi:10.1111/j.1365-246X.2006.03257.x.) is reprinted by permission of the *Geophysical Journal International* editorial staff and Blackwell Publishing. Chapters 3 (G.A. Prieto, P.M. Shearer, F.L. Vernon, and D. Kilb., 2004, Earthquake source scaling and self-similarity estimation from stacking P and S spectra., *J. Geophys. Res.*, 109, B08310, doi:10.1029/2004JB003084) and Chapter 4 (G.A. Prieto, R.L. Parker, F.L. Vernon, P.M. Shearer and D.J. Thomson., 2006, Uncertainties in earthquake source spectrum estimation using empirical Green functions., In *Earthquakes: Radiated Energy and the Physics of Faulting.* Abercrombie, McGarr, Kanamori, and di Toro eds. AGU Geophys. Monograph 170. pp 69–74) are reprinted with permission of the American Geophysical Union. In all cases I was senior author with guidance provided by the co-authors.

Financial support during my years at UCSD was provided by the NSF Grant number EAR0417983. I was also partly supported by a Lawrence Livermore Mini-Grant.

VITA

1979	Born, Jülich, Germany
2002	B.S. Geosciences, Universidad Nacional de Colombia
2004	M.S., University of California, San Diego
2005–2006	Teaching Assistant, Department of Earth Sciences University of California, San Diego
2002–2007	Research Assistant, University of California, San Diego
2007	PhD., University of California, San Diego

PUBLICATIONS

G. A. Prieto, D. J. Thomson, F. L. Vernon, P. M. Shearer and R. L. Parker. (2006) Confidence intervals of earthquake source parameters. *Geophys. J. Int.*, 168, 1227–1234, doi:10.1111/j.1365-246X.2006.03257.x

G. A. Prieto, R. L. Parker, F. L. Vernon, P. M. Shearer and D. J. Thomson. (2006) Uncertainties in earthquake source spectrum estimation using empirical Green functions. In *Earthquakes: Radiated Energy and the Physics of Faulting.*, Abercrombie, McGarr, Kanamori, and di Toro eds. AGU Geophys. Monograph 170. pp 69–74

P. M. Shearer, **G. A. Prieto**, E. Hauksson. (2006) Comprehensive Analysis of Earthquake Source Spectra in Southern California. *J. Geophys. Res.* 111, B06303, doi: 10.1029/2005JB003979.

G. A. Prieto, F. L. Vernon, T. G. Masters, and D. J. Thomson. (2005) Multitaper Wigner-Ville Spectrum for Detecting Dispersive Signals from Earthquake Records. *Proceedings of the Thirty-Ninth Asilomar Conference on Signals, Systems, and Computers*, 938–941, Pacific Grove, CA.

G. A. Prieto, P. M. Shearer, F. L. Vernon, and D. Kilb. (2004) Earthquake source scaling and self-similarity estimation from stacking P and S spectra. *J. Geophys. Res.*, 109, B08310, doi:10.1029/2004JB003084.

FIELDS OF STUDY

Major Field: Seismology

Studies in Earthquake Source Physics.
Professors Frank L. Vernon and Peter M. Shearer

Studies in Earthquake Parameter Uncertainties.
Professors Frank L. Vernon, Peter M. Shearer, Robert L. Parker and David J. Thomson

Studies in Time Series Analysis.
Professors David J. Thomson and Robert L. Parker

ABSTRACT OF THE DISSERTATION

Improving Earthquake Source Spectrum Estimation
using Multitaper Techniques

by

Germán A. Prieto
Doctor of Philosophy in Earth Sciences
University of California, San Diego, 2007
Professor Frank L. Vernon and
Professor Peter M. Shearer, Chairs

Understanding the physics of the earthquake rupture mechanism is essential, given that earthquakes are among the most harmful natural disasters. Some earthquake source parameters such as radiated seismic energy and stress drop can be used to investigate the properties and dynamics of faulting. Estimates of these parameters have large uncertainties, leading to discrepancies among different studies, particularly investigations of the scaling relations of earthquakes.

In order to understand the physics of earthquakes and their behavior as a function of magnitude, it is necessary to have an idea of the uncertainties of the estimated parameters (e.g., when comparing two earthquakes). We have developed a method to estimate the uncertainties of the source parameters as measured from the seismic wave spectra. The large uncertainties expected require improving the methodologies used to obtain the source parameters. We present two methods that take advantage of the large amounts of seismic data available.

In the first method we attempt to separate the effects of anelastic attenuation from the earthquake source spectrum characteristics. Analyzing the latter we are able to obtain source parameters with significantly reduced scatter and which indicate that the earthquake rupture is self-similar in the magnitude range 1.8 to 3.4. In the second method we perform a weighted average of spectral ratios using 160 small earthquakes as empirical Green functions to obtain estimates of the source spectrum of the 2001 M5.1 Anza earthquake. The averaging scheme significantly reduces the uncertainties

and allows us to estimate the radiated seismic energy for this earthquake with greater confidence than is otherwise possible.

Given that in the methods discussed above the seismic parameters were estimated from the spectrum of the seismic waves, we present a new multitaper algorithm that has significant bias reduction compared to standard multitaper techniques and at the same time reducing the roughness of the estimated spectrum. We show that the method has the ability to estimate both the spectrum and its slope, thus increasing the degrees of freedom if parameters are to be estimated.

1

Introduction

Every day there are an immense number of earthquakes occurring somewhere on Earth. Some of these events may be strong enough to be felt, and some, such as the recent Sumatra earthquake in December 2004 in combination with its associated tsunami, generate considerable losses in both infrastructure and human life.

Given that earthquakes are among the most harmful and costly natural disasters, it is essential to have an understanding of the physical processes that lead to their occurrence as well as a deep comprehension of the actual rupture process. Seismology – the study of the structure of the Earth and the physics of earthquakes – has thus a substantial role in mitigating the damaging effects of large earthquakes on our society.

Earthquakes can rupture along just a few meters or along hundreds of kilometers. This study focuses on understanding earthquake ruptures, and to what extent small earthquakes (such as the commonly occurring minor earthquakes in seismically active regions) and large earthquakes (such as the Sumatra earthquake) are generated by similar physical processes or if they are fundamentally different. The observable features of earthquake rupture need to be quantified for comparison between these events, and I investigate how novel time-series analysis tools and inverse problem solving methods can be brought to answer relevant questions in this field.

1.1 Earthquake physics

Earth's tectonic plates slide past each other, in some cases being accommodated by gradual sliding, in other cases by earthquake rupture that accommodates this motion by sudden slip on a fault plane. Seismologists try to understand this sudden behavior by looking at quantifiable features that can be extracted from records of the radiated elastic waves at seismic stations on or near the surface of the Earth.

An earthquake is a failure in Earth's crust. Due to plate motion there is a certain amount of potential energy (gravitational and strain energy) available within a certain region S. We may assume (for a short-term process) that the accumulated strain energy is released in the region S by the earthquake rupture. During the failure process, some energy is radiated as seismic waves (radiated energy E_S) and some energy is dissipated mechanically (fracture energy E_G) and thermally (thermal or frictional heating energy E_F). I will discuss and describe these and other terms in the following section.

1.1.1 Static and dynamic earthquake parameters

In order to understand the physics of earthquakes, it is important to quantify the behavior of some seismic parameters that describe the earthquake rupture process. I will discuss both static as well as dynamic source parameters.

Seismic Moment

Consider a point source in which a displacement offset D between the two sides of the fault occurs. It can be shown that a double-couple force can produce a displacement field equivalent to a point dislocation. The scalar seismic moment M_0 of such a double-couple source is given by (e.g., Kanamori and Anderson, 1975; Shearer, 1999):

$$M_0 = \mu D A \tag{1.1}$$

where μ is the shear modulus of the material surrounding the fault, D is the fault displacement, and A is the rupture area. The dimensions of M_0 are force $\times$ length $=$ energy, and usually the unit Nm is used.

Since the rupture in the solid Earth is irreversible, the displacement that occurs between the two sides of the fault is permanent. This displacement also occurs over some finite duration, leading to a ramp-like near-field (very close to the fault) displacement as a function of time.

The far-field displacement, on the other hand, is not permanent and is proportional to the time derivative of the near-field displacement (see Figure 1.1) . Assuming a seismic station is in the far-field (and there is no attenuation or scattering) the scalar seismic moment could be represented by the area under the displacement pulse (e.g., Madariaga, 1976; Shearer, 1999):

$$M_0 = \mu \bar{D} A \tag{1.2}$$

where $\bar{D}$ is the average displacement across the fault.

Seismic moment is believed to be the most useful and easily measured quantification of the size of an earthquake. Unlike other magnitude estimates (local magnitude or surface wave magnitude), M_0 does not saturate for large earthquakes. However, as discussed above, M_0 is a static measure of the size of an earthquake and does not provide any information on the dynamic properties of the source.

Stress Drop

As suggested above, a certain amount of stress is released by an earthquake rupture. The stress drop is defined as the average difference between the stress (we usually consider shear stress) on a fault before an earthquake and the stress after the earthquake (Kanamori and Anderson, 1975; Kanamori and Brodsky, 2004; Shearer, 1999). For a point source the stress drop is:

$$\Delta\sigma = \sigma_0 - \sigma_1 \tag{1.3}$$

where σ_0 and σ_1 represent the stress on the point before and after the earthquake, respectively. Since the stress drop can actually be highly variable in certain regions of the fault plane, we prefer the stress drop averaged over the entire fault plane (Kanamori and Anderson, 1975; Shearer, 1999):

$$\Delta\sigma = \frac{1}{A} \int_S [\sigma_0 - \sigma_1] \, dS \tag{1.4}$$

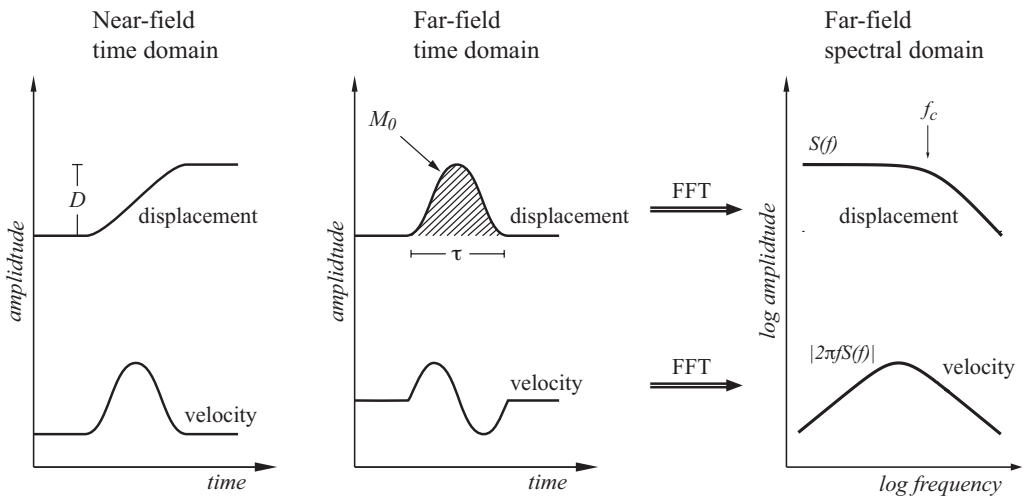

Figure 1.1: Relation between near-field and far-field displacement and velocity, and corresponding far-field amplitude spectra. The seismic moment M_0 is proportional to the shaded area under the far-field displacement curve. The two plots on the right represent the amplitude spectrum. Another source parameter seen in the near-field signal is the displacement D. The pulse duration τ and the corner frequency f_c are also represented in time and frequency domains respectively.

where the integral is performed over the surface of the fault.

The limited resolution in seismological methods does not allow determining the displacement D everywhere on the fault, and forces us to use the approximation

$$\Delta\sigma = C\mu\frac{\bar{D}}{\tilde{L}} \tag{1.5}$$

where μ is the shear modulus and comes from the relation between stress and strain ($\sigma = 2\mu\epsilon$), $\tilde{L}$ is a characteristic rupture dimension, and C is a non-dimensional constant that depends on the geometry of the fault plane.

For the particular case of a circular fault of radius r (Eshelby, 1957; Brune, 1970; Madariaga, 1976), it can be shown (plug 1.2 into 1.5) that the stress drop is related to the seismic moment M_0 by:

$$\Delta\sigma = \frac{7M_0}{16r^3} \tag{1.6}$$

Note that the stress drop is inversely proportional to the cube of the fault dimension, and thus any uncertainty in the radius of the fault will propagate into a large uncertainty in the stress drop. Once again, the stress drop is a static parameter of the seismic source.

Corner frequency and source dimension

As discussed above, the stress drop can be obtained if we know the seismic moment (e.g., from the area under the displacement pulse) and the source dimension; in this case, assuming a circular fault, the source radius. But the source radius r cannot be measured directly from seismological data and further assumptions are needed.

As suggested by Madariaga (1976), assume a circular fault with radius r with a rupture starting in the center and moving radially outside with rupture velocity $v_r = 0.9\beta$, where β is the S-wave speed. The rupture duration time is then

$$\tau = \frac{r}{v_r} \tag{1.7}$$

where τ can effectively be extracted from the pulse duration in the displacement record (Figure 1.1).

One can also study the rupture duration by looking at the spectra of the displacement records. Figure 1.1 plots the spectrum of the far-field displacement. Note that the spectrum remains constant until it reaches what is known as the corner frequency f_c and then the amplitudes decrease rapidly. The corner frequency is clearly related to the pulse duration ($\tau \propto f_c^{-1}$), and once plugged into Equation 1.7 can also be related to the radius of the fault

$$r = \frac{k\beta}{f_c} \tag{1.8}$$

where k is a nondimensional factor (0.32 for P and 0.21 for S-wave, Brune, 1970; Madariaga, 1976), depending on whether we are measuring P or S-wave corner frequencies.

Seismic Energy

When an earthquake occurs, some fraction of the total energy is radiated as seismic waves, while the rest is released as thermal and fracture energies, which together represent the energy dissipation. I now discuss the energy budget involved in the rupture process in order to provide a general understanding of the observational and the physical processes involved. I will assume that the *slip weakening model* (Ida, 1972; Palmer and Rice, 1973) is valid. An illustration of the energies related to fractures is shown in Figure 1.2.

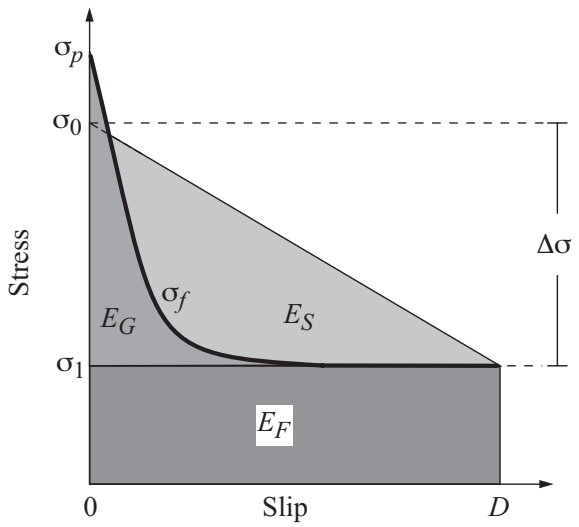

Figure 1.2: Variation of stress at a point on a fault as a function of slip based on the slip-weakening model. This model explains the partition of energy during rupture and the relation between radiated energy E_S and stress drop $\Delta\sigma$. The frictional stress $\sigma_f(s)$ is shown as the thick curve. Note that this figure represents a unit fault, the stress behavior might be different in various regions on the fault plane.

As explained by Kanamori and Rivera (2006), in an expanding crack the initial stress σ_0 increases to the peak stress σ_p (also known as yield stress) at the onset of rupture (see Figure 1.2) and then drops following the curve $\sigma_f(s)$ as a function of slip. The behavior of this curve $\sigma_f(s)$, the frictional stress, shows the particular state of stress during the rupture process. The actual variation $\sigma_f(s)$ may be very complex and seismological estimates are most likely smoothed versions of the real behavior.

The energy that is dissipated ($E_G + E_F$) is represented by the area under σ_f. This energy includes the energy used in creating new crack surface, energy breaking the surrounding rock (creating fault gauge, etc) and energy released as heat, due to friction between the two sides of the fault sliding past each other. Other types of energies (such as latent heat) may be present if, for example, phase transitions due to heating are present (Tinti *et al.*, 2005). Even though Figure 1.2 shows a clear separation between the fracture energy E_G and the thermal energy E_F, this is not always clear and is dependent on many assumptions (Abercrombie and Rice, 2005).

The seismically radiated energy is given by the difference between the entire colored area and the dissipated energy. We can only measure directly from seismograms

the energy that is radiated as seismic waves E_S.

Figure 1.2 also shows the relation between the radiated seismic energy and the stress drop. As explained before, the stress drop $\Delta\sigma$ is the difference between the initial stress on the fault σ_0 and the final stress σ_1. For example, assume that the stress drops instantaneously to σ_1, which would mean there is no fracture energy dissipation and much more energy is radiated seismically. On the other hand, if stress drops quasi-statically (e.g., creep, slow earthquakes) there is minimal or no seismic wave radiation. An intermediate case is the one shown in the example of Figure 1.2.

As suggested by many studies (Mayeda and Walter, 1996; Kanamori and Rivera, 2004; Venkataraman *et al.*, 2006, and many others) the radiated seismic energy is a dynamic measure of the size of an earthquake. Note that both seismic moment M_0 and stress drop $\Delta\sigma$ are static source parameters, while E_S is a dynamic one. The seismic moment depends on the area of the fault rupture and the average displacement, the stress drop is a function of the difference between initial and final stress states (see equation 1.2 and 1.3), while the seismic energy is a function of the behavior of the stress (the function σ_f) during earthquake rupture.

While the stress drop and seismic moment remain fixed, the radiated seismic energy E_S may be extremely variable if, for example, σ_f decreases below the final stress σ_1 and then increases back before rupture stops. Figure 1.3 shows two cases where the stress drop remains constant while the radiated seismic energy varies, depending on the rupture history and not only on the initial and final states. Similarly (Kanamori and Rivera, 2004), we can have a far-field displacement pulse from which we obtain the seismic moment and may estimate the radiated seismic energy by differentiating. One can think of a displacement pulse with a very complicated structure, but still keeping the area under the curve fixed, which would have a very large E_S. Again, the seismic energy is related to the rupture characteristics and is a dynamic parameter.

1.2 Scaling of earthquakes

These are some of the most relevant source parameters used to describe earthquake properties. Now, let us consider how they relate to each other, given that these

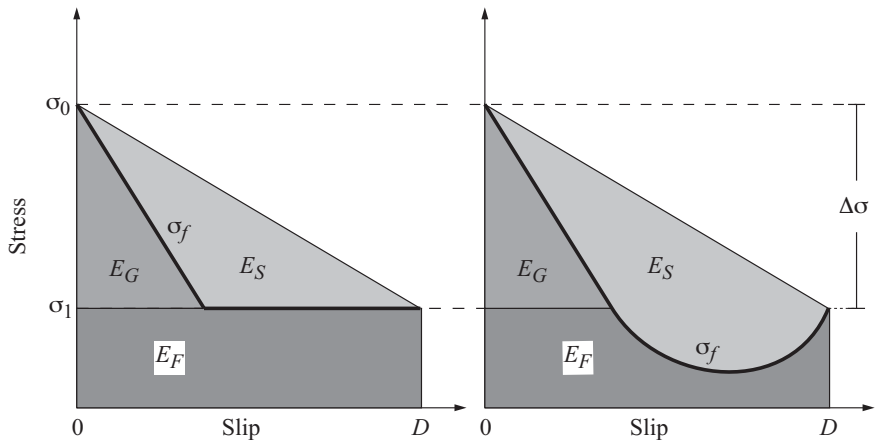

Figure 1.3: Comparison of static and dynamic source parameters. For simplicity I assume peak stress being equal to the initial stress ($\sigma_p = \sigma_0$). Note that while the stress drop only depends on initial and final stresses, the radiated seismic energy E_S is a function of the frictional stress $\sigma_f(s)$ (thick black curve) throughout the rupture process.

relations may be useful in constraining the processes involved in or around the fault.

In addition, I will discuss the behavior of these parameters as a function of the earthquake size. In other words, is the physics of the faulting mechanism associated with an M8.0 earthquake different from that of an M2.0 earthquake? The differences (or lack thereof) could potentially provide means for rapid determination of the size of an earthquake to use in early warning systems (Kanamori, 2005). Aki (1967) suggested a scale invariance of the rupture process, consistent with observations that many geological processes are similar over a wide range of scales (Abercrombie, 1995).

Seismic moment and corner frequency

The relation between the static parameters seismic moment M_0 and the length scale or characteristic rupture dimension $\tilde{L}$ (see equation 1.5) has been widely used in the literature. As seen from equation (1.8), under certain assumptions the corner frequency f_c or the source duration τ can be used as a proxy for the source dimension.

From compilations of a variety of studies (Brune, 1970; Abercrombie and Leary, 1993; Prieto *et al.*, 2004), a very common scaling relation is $M_0 \propto f_c^{-3}$. This would suggest that if this scaling holds (assuming a constant rupture velocity (see Kanamori and Rivera, 2004)), the scaling of stress drop with seismic moment would be $M_0 \propto \Delta\sigma$,

meaning that $\Delta\sigma$ is independent of earthquake size. Other studies suggest (Abercrombie, 1995; Kanamori and Rivera, 2004; Izutani and Kanamori, 2001) that in fact the scaling should be $M_0 \propto f_c^{-(3+\epsilon)}$, where $\epsilon \leqslant 1$ and $\Delta\sigma$ could be scale dependent.

Seismic moment and radiated energy

The radiated energy is of considerable interest because it has relevant information about the dynamics of rupture during an earthquake, and it can be measured with seismological methods.

A very useful dynamic parameter associated with the radiated seismic energy E_S is the apparent stress, as introduced by Wyss and Brune (1968):

$$\sigma_a = \mu \frac{E_S}{M_0} \tag{1.9}$$

which describes the dynamic properties of an earthquake. Replacing the seismic moment M_0 (Equation 1.2) we have

$$\sigma_a = \frac{E_S}{\bar{D}A} \tag{1.10}$$

and can be interpreted as the radiated seismic energy per unit area per unit displacement.

The behavior of apparent stress as a function of earthquake magnitude, the scaling of $\sigma_a \propto M_0$, is of key importance. Does the earthquake apparent stress change with magnitude? Does the seismic energy E_S (seismic waves that leave the source region) increase proportionally as the seismic moment M_0 (the fault area A and the slip $\bar{D}$) increases?

In a very simple way, in Figure 1.4 imagine a *unit fault* for a small and a large earthquake (the large earthquake would be composed of a large number of these unit faults) and their corresponding stress behavior as a function of slip. Clearly, the stress and slip behavior is not necessarily uniform throughout the fault plane, and we must deal with the averages during rupture for each event. The focus here is to investigate how the source parameters (frictional stress, seismic energy, stress drop, etc.) vary over a wide range of earthquake magnitudes, represented by the slip $\bar{D}$. The slip can range from millimeters for small earthquakes to meters for very large ones, and here we take the averages for individual earthquakes as representative of the relation between stress and slip.

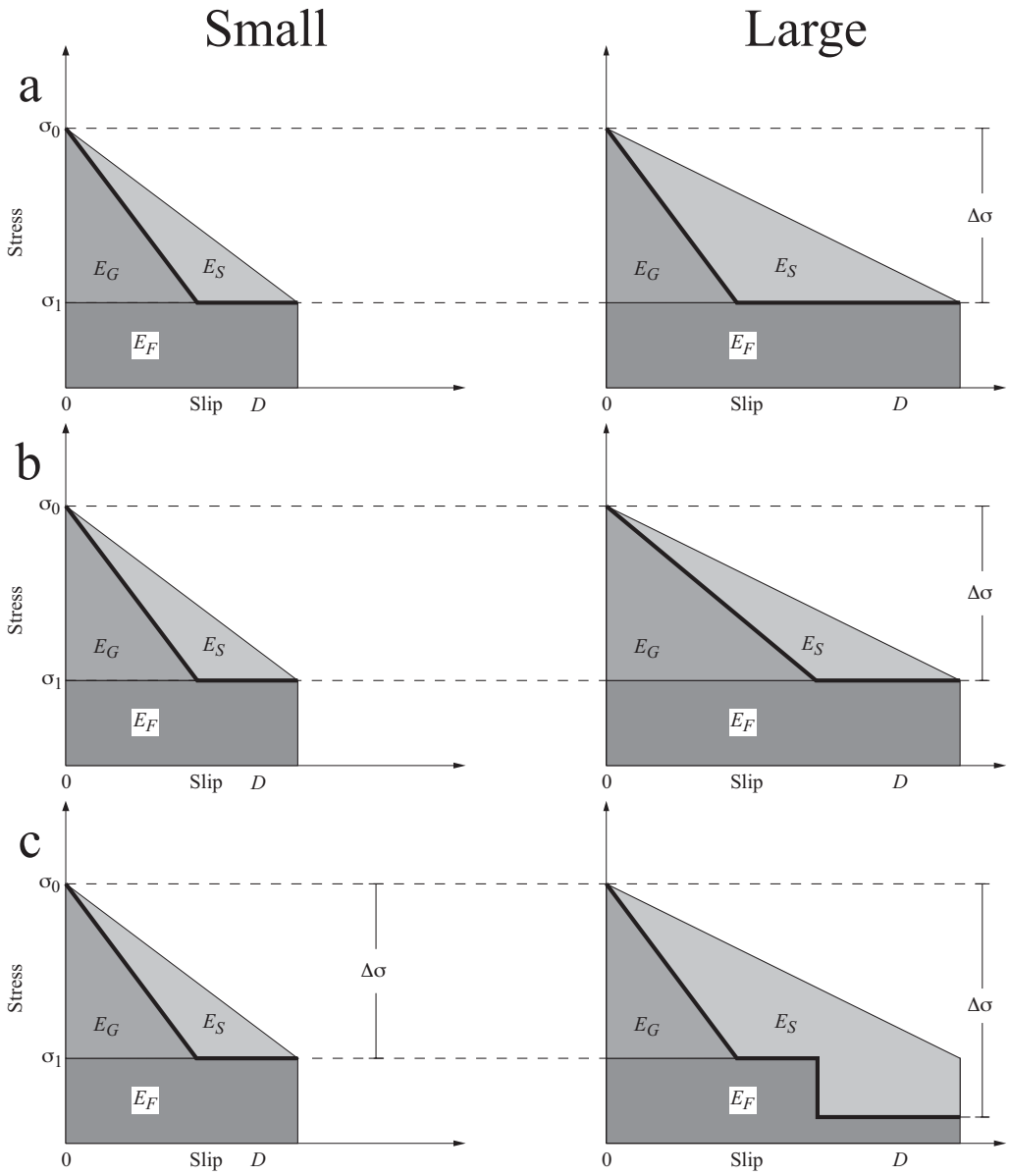

Figure 1.4: Models of earthquake rupture for large and small earthquakes. Peak stress σ_p is neglected for simplicity and the frictional stress $\sigma_f(s)$ is represented by the thick curves. Due to rupture, the initial stress state σ_0 falls to a final value σ_1 as the slip increases to the value D. a) In the self-similar model, E_S increases proportionally as a function of slip D; b) the fracture energy varies with increasing slip D, E_S does not scale proportionally as a function of D and thus apparent stress σ_a will be scale-dependent; c) the fault lubrication model of Kanamori and Heaton (2000) in which friction decreases as in the self-similar model, but after a certain slip it decreases even further, generating a larger stress drop and radiating more energy E_S.

Figure 1.4 shows a small set of models that can be used to describe the earthquake rupture process. The self-similar model (Aki, 1967; Prieto *et al.*, 2004) in Figure 1.4a assumes that the fracture energy is constant (at least for a particular region) generating an E_S that scales proportionally to the final slip D. In this case, both the stress drop $\Delta\sigma$ and the apparent stress σ_a will be constant as a function of magnitude.

If the fracture energy scales with size (Figure 1.4b) and is scale dependent, then the σ_a will not be constant. In this case, the stress drop $\Delta\sigma$ is scale independent, and the size of the earthquake is related to the fracture energy, that is, the fracture energy is a property of the fault zone and the ultimate size of the earthquake is in part governed by the fault zone properties.

A final case is shown in Figure 1.4c, which is known as the fault lubrication model (Kanamori and Heaton, 2000). In this model the rupture process may behave like the self-similar model, until it reaches a certain amount of slip, at which point an additional drop in the frictional stress occurs. This would clearly suggest very different physical processes during rupture between large and small earthquakes. The additional weakening mechanism has been explained in various ways, thermal weakening by pore fluids (Lachenbruch, 1980), elastohydrodynamic lubrication (Brodsky and Kanamori, 2001), and normal stress variations and interface separation during slip (Brune *et al.*, 1993).

If the source parameters $\Delta\sigma$ and σ_a are constant within the uncertainties over a wide magnitude range, the model in Figure 1.4a is possible, and the rupture process should be considered self-similar. If, on the other hand either or both $\Delta\sigma$ and σ_a are scale dependent, then other models must be considered. As in Figure 1.4b, the properties of the fault zone might be controlling the rupture, or as in Figure 1.4c, it might be that the rupture for very large earthquakes (large slip D) follows very different physics than the smaller earthquakes do.

1.3 How do we estimate source parameters?

So far, we have discussed some of the source parameters used for describing earthquake rupture. What is recorded at a seismic station (using mainly velocity sensor

and/or accelerometers) is the ground motion associated with the earthquake rupture and the radiated seismic energy.

Seismic waves travel inside the Earth's crust and mantle and are affected by attenuation, velocity and density variations, near surface effects and scattering until they finally arrive at the seismic station. These disturbances of the original radiated waves need to be accounted for. In this section I will discuss some of the methods used in the literature (some used in this thesis) to correct for the propagation effects in order to be able to investigate the source parameters.

There are also source effects that may be present and need to be corrected or accounted for to obtain a reliable estimate. These include the radiation pattern and the directivity effects.

The radiation pattern may alter the amplitudes of the far-field signals recorded, depending on the azimuthal direction between the earthquake source and the receiver. The directivity effect is a result of the Doppler shift or Doppler effect, where high-frequencies are expected to be radiated in the direction of rupture while lower frequencies will be present in the opposite direction. It has also been argued that the directivity effect focuses the energy radiation in the direction of rupture (Venkataraman and Kanamori, 2004). A more complete discussion about the Doppler effect and directivity in seismic sources can be found in Douglas *et al.* (1988). It is essential to have a good azimuthal coverage of stations to properly take into account these effects.

Correcting waveforms or their spectra for propagation effects is one of the challenging aspects in determining the earthquake source properties. Different methods used for correction of the propagation effects can introduce significant variability in the calculated source parameters as pointed out by Sonley and Abercrombie (2006), even when applied to the same data (e.g., Prejean and Ellsworth, 2001; Ide *et al.*, 2003).

Many researchers use the amplitude spectra of the seismic waves (Abercrombie, 1995; Ide *et al.*, 2003; Prieto *et al.*, 2004) to calculate radiated seismic energy, stress drop and other source parameters. A similar result should be obtained if working on the time-domain signals (Kanamori *et al.*, 1993; Mori *et al.*, 2003), but corrections for attenuation and deconvolutions are in many cases easier in the spectral domain. As we will discuss in the next section, in the spectral domain there are state-of-the-art methods to analyze

the frequency content of complicated signals and study their statistical reliability, which is more difficult in the time domain.

From this point on, it is assumed that the instrument response has been corrected from the recorded signals and that the effects of any incorrect instrument response correction are negligible.

Attenuation correction

To obtain earthquake source parameters, there is a need to correct for attenuation, which may vary from study to study. The basic idea is that the source spectrum is attenuated by the anelastic crust through which it travels. It is customary in earthquake physics to represent the attenuation with the inverse of the quality factor Q.

It is possible to represent the quality factor by a constant Q (Abercrombie, 1995; Prieto *et al.*, 2006) or as a frequency-dependent function $Q(f) = Q_0 f^b$, where both Q_0 and b are constants (Ide *et al.*, 2003; Sonley and Abercrombie, 2006). The calculation of the $Q(f)$ is not trivial and it is in many cases done simultaneously with some source parameters, and some trade-offs between these parameters are inevitable. Even in deep borehole stations it has been shown that a frequency dependent quality factor may be necessary (Ide *et al.*, 2003).

Empirical Green Functions

The Empirical Green Function (EGF) method (Mueller, 1985; Hartzell, 1978; Hough, 1997) takes advantage of the records of a smaller earthquake that is collocated with a larger one. We can assume that up to a certain frequency, the smaller earthquake can be approximated by a point source in time and space. This means that the ground motion recorded at a particular station for that earthquake is approximately the impulse response of the path between the source and the receiver.

This impulse response is then deconvolved from the larger earthquake, in this way accounting for attenuation, scattering, near-source and other effects. In a sense, the EGF method provides a more accurate account of path effects than the attenuation correction. Nevertheless, it also has shortcomings, namely that only events that are collocated can potentially be used, the signal-to-noise ratio for the smaller earthquakes

degrades at both low and high frequencies and that other characteristics, such as similar focal mechanisms for the two earthquakes, are needed.

In principle, this deconvolution can be performed in the time domain (e.g., Mori *et al.*, 2003) or in the frequency domain (Ide *et al.*, 2003; Abercrombie and Rice, 2005; Prieto *et al.*, 2004). As we will show in this thesis, the frequency domain methods used provide an advantage in terms of uncertainty estimation.

1.4 Earthquake source parameters

In this thesis we will focus on the study of earthquake source parameters estimated from the spectrum of the seismic signals. Assuming all the propagation effects have been accounted for, all the parameters can be estimated from the spectra of the seismic waves.

The seismic moment M_0 and the corner frequency f_c can be fitted from the displacement spectra using the Brune (1970) model:

$$S(f) = \frac{M_0}{1 + (f/f_c)^2} \tag{1.11}$$

or the Boatwright (1980) model:

$$S(f) = \frac{M_0}{\left[1 + (f/f_c)^4\right]^{0.5}} \tag{1.12}$$

As shown in Figure 1.5, the seismic moment M_0 is proportional to the zero frequency amplitude of the displacement spectrum, equivalent to the area under the displacement pulse in the time domain (see Figure 1.1). The corner frequency is given by the strong change in slope of the amplitude spectrum, which in the time domain is related to the width of the source pulse, and can be thought as a measure of the rupture duration. In general, some kind of non-linear fitting algorithm needs to be used.

The radiated seismic energy E_S can also be obtained by converting (*rotating*) the spectrum to velocity, squaring and integrating:

$$E_S = C_1 \int_0^\infty [2\pi f S(f)]^2 \, df \tag{1.13}$$

where the constant C_1 has additional parameters such as the wave speed and density of the material surrounding the fault area to obtain the correct units.

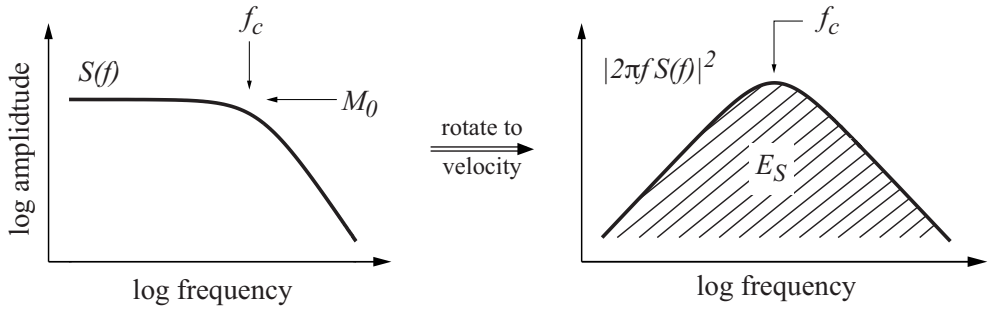

Figure 1.5: Illustration of the relation between the spectrum of the earthquake source (after correcting for all propagation and other source effects) and static and dynamic parameters. The seismic moment M_0 and corner frequency f_c can be related to the displacement source spectrum (left). The energy is related to the area under the velocity spectrum squared (left). See text for explanation.

1.5 Spectrum estimation of seismic signals

There are many applications in geophysics where relevant information contained in a given signal may be extracted from the frequency content of the spectrum. In some cases the scientist may be interested in periodic components usually immersed in some background noise (e.g., normal mode seismology (Gilbert, 1970), climate time series (Chappellaz *et al.*, 1990), etc.), in a general continuous spectrum to be estimated from a short time series (e.g., earthquake source spectra (Brune, 1970; Prieto *et al.*, 2004)) or in comparing two signals and investigating where the similarities or differences are (in seismology for example (Vernon, 1989; Hough and Field, 1996), transfer functions in electromagnetism (Constable and Constable, 2004), elastic thickness of the lithosphere (Daly *et al.*, 2004), etc.). In each of these cases, it is desirable to be able to obtain a *reasonable* spectrum with little or no bias and small uncertainties.

As suggested above, all source parameters we have discussed above can be obtained by analyzing the spectrum of the seismic signals, after other effects (attenuation, directivity, etc.) have been accounted for. But the spectrum estimation of seismic signals poses many difficulties. First, all the seismic waves we are interested in are transient and, in the case of small earthquakes, the seismic phases (e.g., body waves) are contained in a short segment within the record. Second, the signals have very high dynamic range, which might lead to severely biased estimates due to spectral leakage, where frequency

information with high amplitudes (e.g., close to the corner frequency) leaks into frequency regions with low amplitudes. A final difficulty is that the signal is non-stationary; that is, the statistical character of the data changes with position in the record.

It is standard practice to use the discrete Fourier transform (DFT) to estimate the spectrum of a particular series. However, simply using the DFT and squaring to obtain the spectrum, in other words using the periodogram, is a poor choice and should never be done (for discussion about the choice of the periodogram, see Harris, 1978; Thomson, 1982, and references therein). In general, it is much better practice to window the time series with a taper before performing the DFT and squaring to reduce spectral leakage.

Conventional tapers (in the time domain) used for spectrum estimation have a bell-shaped curve (sometimes with a flat top) and tend to zero at the edges. This approach has a major limitation, in that by applying a taper we are effectively discarding significant statistical information in a given time series. The data points at the edges of the record are down-weighted, while the data in the center is emphasized, which causes the variance of the spectral estimate to be greater than that of the periodogram.

In Thomson (1982) a different method, called the multitaper spectral analysis, was introduced. As its name suggests, the idea of the method is to use multiple orthogonal tapers to window the time series and reduce spectral leakage, and by applying the DFT and squaring, obtain almost independent estimates of the spectrum, called *eigenspectra*. As shown in Thomson (1982) and in many other studies (Vernon, 1989; Park *et al.*, 1987b; Riedel and Sidorenko, 1995) as long as only one taper is used, there will be a trade-off between the resistance to spectral leakage and the variance of a spectral estimate.

The tapers are constructed to be leakage-resistant, and sample the time series in different ways. The information that is discarded by the first taper (whose shape is very similar to the conventional tapers) is partially recovered by the second taper, and the information down-weighted by these two tapers is recovered by the third, fourth, etc.

As explained by Park *et al.* (1987b), single-taper spectral estimates have relatively large variance (increasing as a larger fraction of the data is discarded and the bias of the estimate is reduced) and are inconsistent estimates (i.e., the variance of the

estimate does not drop as one increases the number of data points). In the case of single taper methods a smooth estimate can be achieved by applying a moving-average, thus reducing the variance, while at the same time reducing the frequency resolution and increasing the bias of the estimate.

In the multitaper algorithm only a few tapers are used to construct the spectral estimate. A weighted sum of the *eigenspectra* is formed, leading to a smooth estimate of the spectrum with variance reduced due to the averaging process. By using the multiple tapers, the estimator is also consistent. It has been shown in multiple cases that the multitaper algorithm outperforms single-taper smoothed spectral estimates (Park *et al.*, 1987b; Bronez, 1992; Riedel and Sidorenko, 1995).

In terms of non-stationarity, the multitaper estimation is also a better choice. Because single taper estimates weight the data in the center of the signal more, the information present at the end of the signal is not used, which may lead to the misrepresentation of the spectrum. Multitaper estimates discard much less data and are sensitive to information from almost the entire signal.

Another very important feature of multitaper spectral estimates is the possibility of obtaining measures of uncertainties and confidence intervals from the data. Given that we have almost independent eigenspectra, it is possible to obtain error estimates of the spectrum and associated parameters (Thomson and Chave, 1991).

1.6 Objectives

The primary goal of this thesis is to develop methods to obtain better, more reliable estimates of the seismic source parameters used to investigate the physics of earthquakes. One of the issues in the physics of earthquakes is the scaling of static and dynamic parameters and what this tells us about the rupture process. As mentioned earlier, all these parameters and their scaling can be determined from the spectrum of the seismic waves, and I will focus on using and improving the state-of-the-art multitaper spectrum algorithm to obtain the most reliable parameter estimations.

The basic problem in earthquake source scaling can be observed in Figure 1.6. The figure shows a compilation of different studies on the radiated energy and apparent

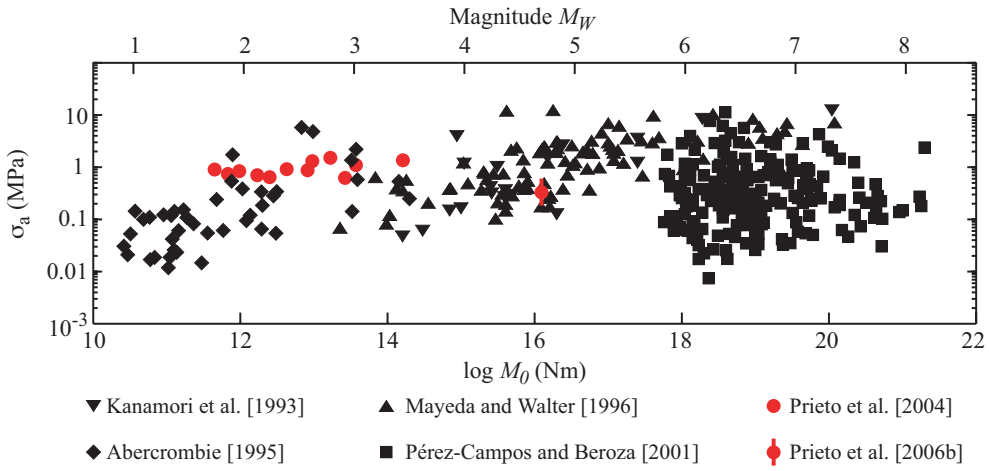

Figure 1.6: The scaling controversy on energy-seismic moment ratio and apparent stress as first compiled by Ide and Beroza (2001). Each symbol denotes a different data set as shown. In red, I added results from studies in the Anza region from Chapter 3 and the M5.1 Anza earthquake from Chapter 4 with 95% confidence bounds.

stress scaling, ranging from very small earthquakes (M0.0) to very large ones (M8.0) (Abercrombie, 1995; Mayeda and Walter, 1996; Kanamori *et al.*, 1993; Pérez-Campos and Beroza, 2001; Prieto *et al.*, 2004; Shearer *et al.*, 2006). The question is whether the dynamic parameter is constant over this wide range of magnitudes or if there is a change in behavior as the magnitude increases. Depending on the data you focus on, there could be a linear increasing trend, but with the large scatter of the data, it is also possible to have a constant scaling of the dynamics of the earthquake rupture.

I have not organized this thesis by the temporal evolution of my research, but rather by topic. The first question that may arise from looking at Figure 1.6 is what points can be believed? What are the error bars of each of these points? Is the large scatter seen in these data points real, meaning that the rupture process may have very different dynamic behavior even if the magnitude of the earthquakes is similar, or is this scatter due to large uncertainties in the individual estimates?

Data shown in Figure 1.6 come from different regions (although most of them are in California), the signal processing methods are quite different (e.g., time domain in Kanamori *et al.* (1993); multitaper spectral analysis in Abercrombie (1995); and coda wave envelopes in Mayeda and Walter (1996)) and earthquakes include strike-slip, thrust and normal faulting, making them very difficult to compare.

Many of the studies shown in the figure and many others published in the literature lack the analysis of the uncertainties associated with the methodologies used in the analysis of the seismic parameters. In order to compare two or more estimates, it is of key importance to have an understanding of the uncertainties and assumptions of the methods used.

Since this, to me, represents a major shortcoming of the present literature, in **Chapter 2** we present a method to obtain confidence intervals on earthquake source parameters using the multitaper algorithm from single-station measurements. We discuss the use of the jackknife variance applied to source parameter estimation, and show that large uncertainties are expected even in the ideal conditions of deep borehole records. An extension to multiple station measurements is also discussed. **Chapter 2** has appeared in *Geophysical Journal International* under the title "Confidence intervals for earthquake source parameters" (Prieto *et al.*, 2007a). I participated as primary author in all phases of the development of this research, including the programing of the computer algorithms.

Given that large uncertainties are expected (even in ideal conditions), we present in **Chapter 3** and **Chapter 4** two different methods to improve the estimation of the source parameters by taking advantage of the large number of earthquakes that occur in southern California using the Anza Seismic Network in Southern California. **Chapter 3** and **Chapter 4** are reformatted versions of the papers "Earthquake source scaling and self-similarity estimation from stacking P and S spectra" (in *Journal of Geophysical Research* (Prieto *et al.*, 2004)) and "Uncertainties in earthquake source spectrum estimation using empirical Green functions" (appeared in the *AGU Monograph on Radiated Energy and the Physics of Earthquake Faulting* (Prieto *et al.*, 2006)) respectively. In both cases I was senior author and under the supervision of my co-authors developed the research and computer codes that form the basis of these chapters.

In **Chapter 3**, we study the self-similarity and scaling relations of a cluster of 400 earthquakes in the range $M0.5$ to $M3.4$ by iteratively stacking spectra of P and S waves. The iterative approach is aimed at separating the propagation effects from the source spectra and, hence, be able to study the scaling properties. As I discussed above, we study the scaling by looking at source parameters (e.g., M_0, f_c, E_S, etc.) and their relation, but also introduce a test on self-similarity, which is independent of the choice

of parameterization of the earthquake spectra.

In **Chapter 4**, we investigate the problem of using Empirical Green functions (EGF) to obtain the source spectrum and associated source parameters. We show that the spectrum obtained via the EGF method has large uncertainties and introduce a method to reduce the uncertainties by using multiple EGFs instead of a single one. The method is applied to the 2001 $M5.1$ Anza earthquake in southern California.

The multitaper algorithm presented by Thomson (1982) reduces spectral leakage effectively, but suffers from *local bias*. By performing the weighted averaging of the different eigenspectra as proposed by Thomson (1982), it is assumed that the spectrum is white within a certain frequency band. As shown by Riedel and Sidorenko (1995), there is considerable bias present due to the curvature of the spectrum. In **Chapter 5**, we present an extension to the multitaper, which reduces the curvature bias and provides additional information about the derivative or slope of the spectrum of a particular signal. We show that this Quadratic multitaper method provides improved estimates of the spectrum, especially where strong structure is present, for example around the corner frequency f_c. **Chapter 5** is under consideration for publication in *Geophysical Journal International* as "Quadratic Multitaper Spectrum" (Prieto *et al.*, 2007b).

Chapter 6 summarizes the main conclusions of the thesis.

Acknowledgments

I would like to thank Greg Beroza and Luis Rivera and Luciana Astiz for comments on early versions of this chapter. Funding for this research was provided by NSF Grant number EAR0417983.

2

Confidence intervals for earthquake source parameters

We develop a method to obtain confidence intervals of earthquake source parameters, such as stress drop, seismic moment and corner frequency, from single station measurements. We use the idea of jackknife variance combined with a multitaper spectrum estimation to obtain the confidence regions. The approximately independent spectral estimates provide an ideal case to perform jackknife analysis. Given the particular properties of the problem to solve for source parameters, including high dynamic range, non-negativity, nonlinearity, etc., a log transformation is necessary before performing the jackknife analysis. We use a Student-t distribution after transformation to obtain accurate confidence intervals. Even without the distribution assumption, we can generate typical standard deviation confidence regions. We apply this approach to four earthquakes recorded at 1.5 km and 2.9 km depth at Cajon Pass, California. It is necessary to propagate the errors from all unknowns to obtain reliable confidence regions. From the example, it is shown that a 50% error in stress drop is not unrealistic, and even higher errors are expected if velocity structure and location errors are present. An extension to multiple station measurement is discussed.

2.1 Introduction

There is a long-standing controversy on whether stress drop increases with earthquake magnitude or remains constant over a wide range of earthquake sizes (Aki, 1967; Archuleta *et al.*, 1982; Kanamori *et al.*, 1993; Abercrombie, 1995; Mayeda and Walter, 1996; Ide and Beroza, 2001). The behavior of source parameters including stress drop, corner frequency, radiated seismic energy and apparent stress are of key importance in understanding the physics of earthquakes. However, it is difficult to estimate stress drop reliably from seismograms since it is dependent on the cube of the corner frequency f_c and in turn, f_c is dependent on an accurate account of seismic attenuation, path effects, etc. These factors lead to considerable uncertainty in estimates of stress drop and other source parameters.

Abercrombie (1995) used records from a 2.5 km deep borehole in Cajon Pass and showed that the data supported a constant stress drop, but also an increasing apparent stress with earthquake magnitude. Also from deep borehole data, Prejean and Ellsworth (2001) reported a similar result. A magnitude dependency has been supported by some studies (e.g., Kanamori *et al.*, 1993; Mayeda and Walter, 1996; Mori *et al.*, 2003). while other studies have suggested scale independence (e.g., McGarr, 1999; Ide and Beroza, 2001; Ide *et al.*, 2003), finding no evidence of increasing stress drop or apparent stress with magnitude.

More recently Abercrombie and Rice (2005) revisited some of the Cajon Pass seismograms and using both spectral fitting and Empirical Green Functions (EGF) concluded that both apparent stress and stress drop may increase with increasing earthquake size, but noted that the uncertainties were still large and scale independence could not be entirely discarded.

So, what are the uncertainties of the estimated source parameters? Error analysis for source parameters has been attempted before (e.g., Archuleta *et al.*, 1982; Fletcher *et al.*, 1984) but seems to have been neglected more recently. Recently Prieto *et al.* (2006) developed an approach to obtain uncertainties in earthquake source spectrum using EGF and applied it to obtain confidence intervals of radiated seismic energy. As pointed out by Tukey (1960):

"Probably the greatest ultimate importance, among all types of statistical

> procedures we now know, belongs to *confidence procedures* which, by making interval estimates, attempt to reach as strong conclusions as are reasonable by pointing out, not single likely values, but rather whole classes (intervals, regions, etc.) of *possible* values, so chosen that there can be high confidence that the 'true' value is *somewhere among them.* Such procedures are clearly quantitative conclusion procedures. ... "

In this paper we use the idea of the jackknife variance (Tukey, 1958) and follow a similar recipe to the one applied for spectra (Vernon, 1989; Thomson and Chave, 1991) to construct confidence intervals for earthquake source parameters. This is applied to single station seismograms but can easily be extended to multiple station and spectral ratios and EGF techniques. The confidence intervals are of paramount importance to obtain meaningful scaling relations when different studies, regions, etc. are compared.

We present an example from data recorded at the Cajon Pass Borehole Experiment Phase II with some records also used in Abercrombie (1997) and show the resultant confidence intervals for stress drop and other source parameters for four small earthquakes (Table 2.1) that were also recorded by the Southern California Seismic Network (SCSN). Figure (2.1) shows a map with the relocated earthquakes and the borehole location.

2.2 The Jackknife Method

The *jackknife* was first introduced by Quenouille (1949) and then named and extended by Tukey (1958) to estimate variances. It is one of many resampling methods used for statistical inference. One of the great advantages of the jackknife is that one does not need to know the statistical distribution of the parameter in question and that it works on complicated processes reliably (a detailed proof is given in Reeds, 1978). In this paper we will use the so-called delete-one jackknife, which we will refer to simply as the jackknife. A good review can be found in Miller (1974) and Efron (1982).

Assume $X_1, X_2, \ldots, X_K$ are K independent random observations taken from an unknown probability distribution characterized by a parameter θ which is to be estimated. The estimate of θ using all observations is:

$$\hat{\theta} = \hat{\theta}[X_1, X_2, \ldots, X_K] \tag{2.1}$$

Table 2.1: Hypocentral parameters for the earthquakes recorded at Cajon Pass used in this study. The relocations by Shearer *et al.* (2005), model SHLK 1.02, are given for SCSN earthquakes.

ID	Year	Month	Day	Hour	Min	Sec	Lat (°)	Lon (°)	Depth (km)	M_L	Dist (km)	CuspID
01	1994	01	01	14	56	42.087	34.2139	−117.4194	14.98	1.1	12.37	3138796
02	1994	01	01	17	47	31.257	34.3863	−117.0185	11.04	3.5	42.86	3138805
03	1994	08	22	21	27	51.366	34.3522	−117.6339	11.03	2.3	14.98	3181641
04	1994	09	21	04	16	51.762	34.2316	−117.4623	13.63	2.3	9.32	3185485

M_L is the local magnitude given by the SHLK catalog, Dist is the hypocentral distance from the borehole station, CuspID is the SCSN ID for the earthquakes.

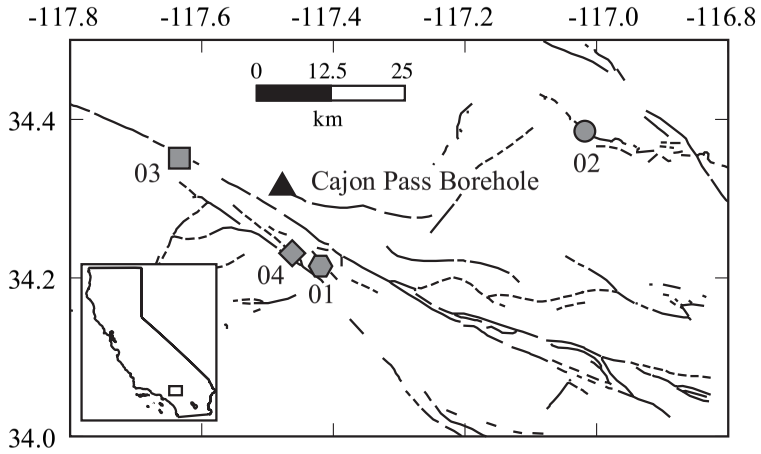

Figure 2.1: Map showing the Cajon Pass Borehole location and relocated earthquakes considered in this study. Relocations from Shearer, Hauksson and Lin (2005) model SHLK 1.02

Let

$$\hat{\theta}_{\widehat{i}} = \hat{\theta}[X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_K] \tag{2.2}$$

be the delete-one estimate of θ, where the ith observation X_i is not used to estimate $\hat{\theta}_{\widehat{i}}$. The data are thus subdivided in K groups of size $(K-1)$ by deleting each entry in turn.

An important application of the jackknife was suggested by Tukey (1958), and is the jackknife estimate of the variance of $\hat{\theta}$

$$var\{\hat{\theta}\} = \frac{K-1}{K} \sum_{i=1}^{K} \left[\hat{\theta}_{\widehat{i}} - \hat{\theta}_{\widehat{\cdot}}\right]^2 \tag{2.3}$$

where

$$\hat{\theta}_{\widehat{\cdot}} = \frac{1}{K} \sum_{i=1}^{K} \hat{\theta}_{\widehat{i}} \tag{2.4}$$

is the mean of the delete-one estimates (2.2). Although it has been proposed (Wu, 1986) that deleting an *arbitrary* number of observations might have better convergence properties, we use throughout the paper the delete-one jackknife because of its simplicity, efficiency and independence of an arbitrary chosen subdivision for the groups.

As suggested by Miller (1974) and applied in spectrum estimation (Vernon, 1989; Thomson and Chave, 1991) it is sometimes necessary to use a transformation that stabilizes the variance, especially when the statistic being investigated is bounded or its distribution is strongly non-gaussian. This can be important when estimating errors in stress drop $\Delta\tau$, seismic moment M_0, and corner frequency f_c, all with a range $[0, \infty)$.

2.2.1 Jackknife in Regression Problems

Consider the regression problem for a basic model

$$\mathbf{Y} = \mathbf{A}\beta + \mathbf{e} \tag{2.5}$$

where $\mathbf{Y}$, $\mathbf{e}$ are m sized vectors of the data and the errors, $\mathbf{A}$ is a $m \times p$ matrix from the model, and β is a p size vector of the parameters we wish to find.

Miller (1974) examined the traditional jackknife approach by deleting rows of both $\mathbf{Y}$ and $\mathbf{A}$ simultaneously and showed the asymptotic normality of the jackknife

solution vector and its variance under general conditions. The delete-one estimate is given by solving

$$\mathbf{Y}_{\widehat{i}} = \mathbf{A}_{\widehat{i}}\, \hat{\beta}_{\widehat{i}} \tag{2.6}$$

where $\mathbf{Y}_{\widehat{i}}$ and $\mathbf{A}_{\widehat{i}}$ have the ith row removed. As will be clear in the subsequent sections, the problem to solve for source parameters is non-linear and the model (2.5) is not appropriate. Instead we have

$$y_i = g_i(\beta) + e_i \tag{2.7}$$

where g_i is a nonlinear smooth function of the parameters in β (Fox *et al.*, 1980; Wu, 1986). In an analogous way, we want to obtain the delete-one estimates $\hat{\beta}_{\widehat{i}}$ that satisfy (2.7) by means of one of many non-linear parameter estimation techniques (non-linear least-squares, grid search, etc).

2.3 Multitaper Spectrum estimates

The multitaper spectrum algorithm was introduced by Thomson (1982) and has been widely used in the geophysical community (e.g., Park *et al.*, 1987b; Vernon, 1989; Chappellaz *et al.*, 1990; Lees and Park, 1995; Abercrombie, 1995). The method takes advantage of a family of orthogonal tapers which are resistant to spectral leakage.

Given a time series $x(t)$ with N contiguous data samples and assuming unit sampling, we multiply the time series by a sequence $a(t)$ called a taper and apply a DFT

$$Y(f) = \sum_{t=0}^{N-1} x(t)a(t)e^{-2\pi i f t} \qquad \text{with} \qquad \sum_{t=0}^{N-1} |a(t)|^2 = 1 \tag{2.8}$$

to obtain a direct estimate of the true spectrum $S(f)$ of the signal

$$\hat{S}(f) = |Y(f)|^2 \tag{2.9}$$

The question is then what taper to use? Is there a reason to prefer one taper over the other?

Spectral leakage is the bias introduced by energy *leaking* from frequencies different from the frequency f for $S(f)$. Now the question becomes: what taper $a(t)$ has

the greatest concentration of energy in its Fourier transform? Spectral properties of the taper can be studied from its DFT

$$A(f) = \sum_{t=0}^{N-1} a(t) e^{-2\pi i f t} \tag{2.10}$$

The function $|A(f)|$ for conventional tapers has a broad main lobe and a succession of smaller sidelobes. The larger the sidelobes, the more spectral leakage is biasing $\hat{S}(f)$.

We can express the estimate in equation (2.9) as a convolution of the taper transform (2.10) and the true spectrum $S(f)$ (see Thomson, 1982; Park *et al.*, 1987b, for derivation):

$$\hat{S}(f) = \int_{-1/2}^{1/2} |A(f - f')|^2 S(f')\, df' \tag{2.11}$$

The interpretation of this equation is as a convolution describing the *smearing* of the true spectrum as a consequence of the discrete sampling. A good taper will have a spectral window with low amplitudes whenever $|f - f'|$ gets large and large amplitudes whenever $|f - f'|$ is small.

Slepian (1978) suggested choosing a frequency W, where $0 < |W| \leq 1/2$ (unit sampling) and maximizing the fraction of energy of A at frequencies from $(-W, W)$. In mathematical form this is equivalent to:

$$\lambda(N, W) = \frac{\int_{-W}^{W} |A(f)|^2 df}{\int_{-1/2}^{1/2} |A(f)|^2 df} \tag{2.12}$$

Since no finite time series can be completely band-limited, $\lambda < 1$. The spectral leakage comes from the sidelobes of $A(f)$ convolved with the spectrum outside the band $(f - W, f + W)$. One can think of $\lambda(N, W)$ as the amount of spectral energy at $\hat{S}(f)$ that comes from $(f - W, f + W)$ and $1 - \lambda$ as the amount that comes from outside the band or as the *bias* from outside the band.

We wish to maximize the value of λ by choosing $A(f)$ appropriately. Substitute (2.10) into (2.12) and represent $a(t)$ by an N-vector of coefficients $\mathbf{a}$; taking the gradient

of λ with respect to $\mathbf{a}$ and setting to zero leads to the matrix eigenvalue problem:

$$\mathbf{D} \cdot \mathbf{a} - \lambda \mathbf{a} = 0 \tag{2.13}$$

where $\mathbf{D}$ is a symmetric matrix

$$D(t, t') = \frac{sin 2\pi W(t - t')}{\pi(t - t')} \tag{2.14}$$

with eigenvalues $1 > \lambda_0 > \lambda_1 > \cdots > \lambda_{N-1} > 0$ and associated eigenvectors $v_k(t; N, W)$ called the Slepian sequences (Slepian, 1978). From now on we will drop the explicit dependence on N and W.

The eigenvector with the largest eigenvalue is the best possible taper for the suppression of spectral leakage, and in practice we find λ_0 is usually extraordinarily close to one. But in fact it can be proved that the first $2NW - 1$ eigenvalues are also very close to one, leading to a whole family of excellent tapers. The multitaper method exploits this fact by using all of these tapers rather than merely the first one. Because the eigentapers are orthogonal (both in time and frequency domains), the estimates based on them are statistically independent of each other and can therefore be combined together to yield a more reliable overall estimate as we will explain.

In practice we choose a bandwidth W over which the spectrum is to be smoothed, thus fixing NW, which is called the time-bandwidth product of the system under study. For practical problems we always choose $NW > 1$, because we cannot expect to obtain good concentration into a frequency band narrower than $f_R = 1/N$, the Rayleigh resolution.

Figure (2.2) shows the Slepian sequences and their Fourier transforms with corresponding eigenvalues for a time series with $N = 100$ samples, $NW = 4$, and $W = 0.04$ for unit sampling. The horizontal axis is shown in Rayleigh units, basically equivalent to the frequency sampling. Figure (2.2) suggests that $NW = 4$ is equivalent to saying that the smoothing will take place over NW Rayleigh bins around the frequency of interest. Note that here we have assumed unit sampling – if that is not the case, then the time-bandwidth product is actually $\Delta t NW = 4$, where Δt is the sampling rate, in order to maintain the proper units of W.

Turning back to the spectral estimation problem, given a particular bandwidth

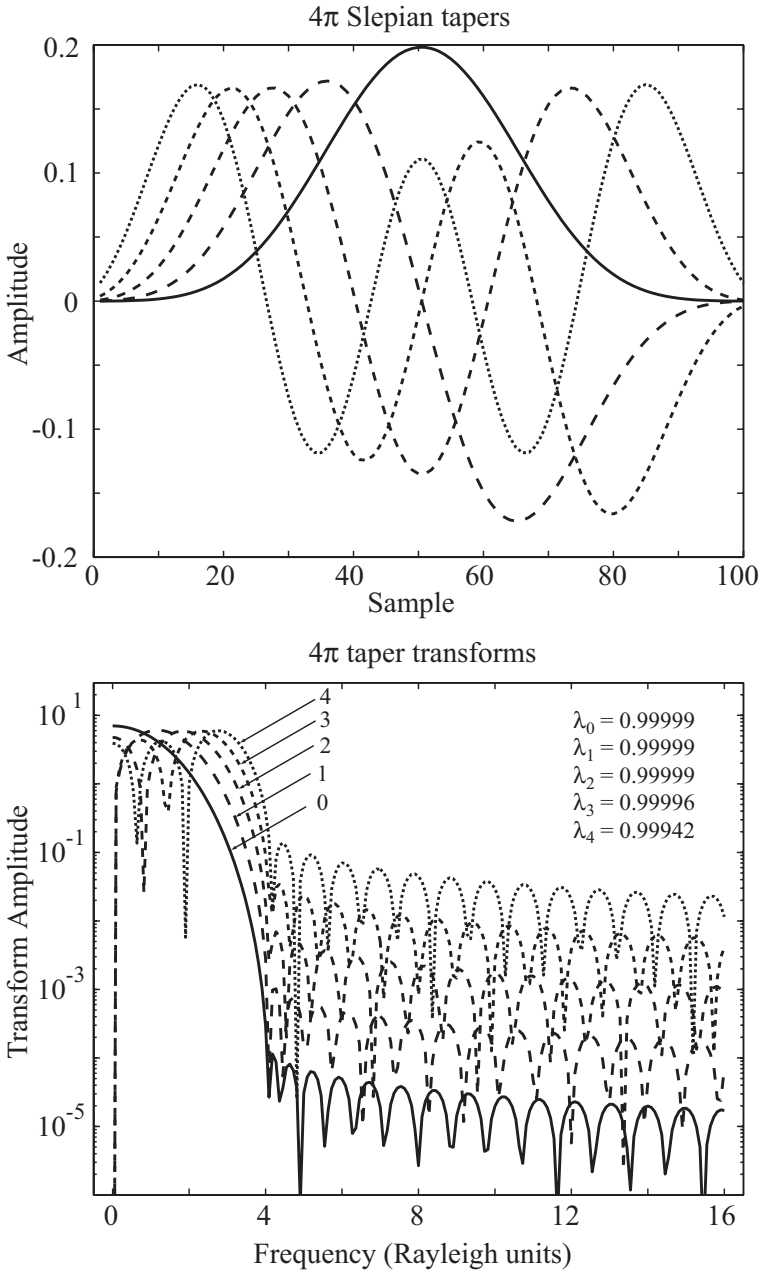

Figure 2.2: The five lowest order 4π Slepian tapers (top panel) and corresponding Fourier transform amplitudes (lower panel). Solid lines correspond to the zero order sequence $v_0(t)$, higher order tapers are plotted with dashed lines. We have used $N = 100$ and $NW = 4$. Estimated eigenvalues are also provided.

W, we compute DFTs of the tapered data $Y_k(f)$, called the eigencomponents,

$$Y_k(f) = \sum_{t=0}^{N-1} x(t) v_k(t) e^{-2\pi i f t} \tag{2.15}$$

We generally use $k = 1, \ldots, K$, where $K = 2NW - 1$. As expressed above, the corresponding eigenvalues are $\lambda_k \approx 1$ with good leakage properties.

As suggested by Thomson (1982) we use the adaptive weighting procedure

$$\mathcal{Y}_k(f) = d_k(f) Y_k(f) \tag{2.16}$$

and the corresponding adaptive spectral estimate

$$\hat{S}(f) = \frac{\sum_{k=0}^{K-1} |\mathcal{Y}_k(f)|^2}{\sum_{k=0}^{K-1} |d_k(f)|^2} \tag{2.17}$$

where the weights $d_k(f)$ are chosen to reduce bias from spectral leakage. The frequency dependent weights are useful in the analysis of high dynamic range spectral processes. The weights work as follows. At frequencies where the spectrum is reasonably flat, the weights $d_k(f) \approx 1$, thus reducing the variance of the spectral estimate by averaging over all the eigencomponents Y_k. At frequencies where the spectrum has a large dynamic range the higher order eigencomponents might be biased and the weights reduce the contributions from these components.

The optimal weights $d_k(f)$ can be found by minimizing the misfit between the estimated spectrum and the true spectrum $S(f)$. The approximate optimum weights are

$$d_k(f) \approx \frac{\sqrt{\lambda_k} S(f)}{\lambda_k S(f) + (1 - \lambda_k)\sigma^2} \tag{2.18}$$

where σ^2 represents the variance of the time series. The term $(1 - \lambda_k)\sigma^2$ represents an approximation to the bias from spectral leakage. Since we do not know the true spectrum, we replace $S(f)$ by an estimate $\hat{S}(f)$.

We find the weights and estimated spectrum $\hat{S}(f)$ by iteration. As an initial estimate of $S(f)$ we take the arithmetic average of the first two squared eigencomponents $|Y_0(f)|^2$ and $|Y_1(f)|^2$ and substitute in (2.18) to obtain estimates of $d_k(f)$. The weights are then used in (2.16 - 2.17) to obtain a new spectral estimate $\hat{S}(f)$ and this process is

repeated. Convergence is rapid and only a few cycles are necessary. Note that both the tapers and weights are normalized in order to keep the spectrum in physical units.

The kth eigenspectrum is

$$\hat{S}_k(f) = \left|\hat{\mathcal{Y}}_k(f)\right|^2 \tag{2.19}$$

For the jackknife approach, we will use the $\hat{S}_k$ as the K independent estimates of the spectrum. At each frequency f the multitaper estimate of the log spectrum is given by

$$\ln \hat{S} = \ln \left[\frac{1}{K} \sum_{k=1}^{K} \hat{S}_k \right] \tag{2.20}$$

and we also define the delete-one spectrum

$$\ln \hat{S}_{\widehat{i}} = \ln \left[\frac{1}{K-1} \sum_{k=1,k\neq i}^{K} \hat{S}_k \right] \tag{2.21}$$

The logarithmic transformation of the spectrum is suggested in Thomson and Chave (1991), providing a more symmetric distribution than the standard χ^2 for spectral estimates.

2.4 Source parameter jackknife

A general source model of the displacement spectra of both P and S waves (e.g., Abercrombie, 1995) is:

$$u(f) = \frac{\Omega_0 \, e^{-(\pi f t/Q)}}{\left[1 + (f/f_c)^{n\gamma}\right]^{1/\gamma}} \tag{2.22}$$

where Ω_0 is the long period amplitude, f is the frequency, f_c is the corner frequency, n the high frequency fall-off rate, γ is a constant, t is the travel time, and Q a frequency independent quality factor. Modified versions of spectral shapes proposed by Brune (1970) and Boatwright (1980) can be obtained by changing γ. Based on previous studies of data from the Cajon Pass Borehole (Abercrombie, 1995, 1997) a value $\gamma = 2$ and a variable fall-off n fits the spectra reasonably well. In this paper we will use

$$u(f) = \frac{\Omega_0 \, e^{-(\pi f t/Q)}}{\left[1 + (f/f_c)^{2n}\right]^{0.5}} \tag{2.23}$$

Following Ide *et al.* (2003) take the logarithm

$$\ln u(f) = g(f;\beta) \tag{2.24}$$

$$= \ln \Omega_0 - 0.5 \ln(1 + (f/f_c)^{2n}) - \frac{\pi f t}{Q} \tag{2.25}$$

where β is a vector of three components given by the parameters we are searching for, namely Ω_0, f_c, Q. The function $g(f, \beta)$ is clearly non-linear and some kind of non-linear inversion is necessary. It is not the aim of this paper to discuss the difficulties encountered in solving this problem, and we suggest reading Bard (1974) on nonlinear parameter estimation, as applied to source physics (see Abercrombie, 1995; Ide *et al.*, 2003, and references therein).

The estimate $\hat{\beta}$ of β using all observations (as in equation 2.1) is given by the solution of (2.24) using the multitaper spectrum estimate (2.20). The delete-one $\hat{\beta}_{\widehat{i}}$ parameter instead uses the delete-one spectrum (2.21). The result is K delete-one estimates of the long period amplitude, corner frequency and quality factor, denoted respectively $\Omega_{0,\widehat{i}}, f_{c,\widehat{i}}, Q_{\widehat{i}}$.

2.4.1 Transformations

The use of transformations before performing the jackknife is in some cases necessary. In terms of source parameters a logarithmic transformation should provide more stable estimates of variance. Some reasons for this are as follows:

- The three seismic parameters are non-negative. If a simple Gaussian distribution is assumed, the tails give a nonzero probability that a parameter is negative, which is not physical. The jackknife does not constrain variances to be positive.
- As suggested by (Archuleta *et al.*, 1982), if no transformation is performed, the arithmetical average will be biased to the larger values, while taking a transformed parameter gives equal weight to all independent estimates.
- A closer to normal distribution of the errors is achieved by such a transformation.

In this respect we will perform the jackknife on $\theta = \ln \Omega_0$ rather than Ω_0.

2.4.2 Confidence intervals

As a result of section 4.1 we obtain an estimate $\hat{\theta} = \ln \hat{\beta}$ and the variance of the transformed variable

$$\tilde{\sigma}^2 = var\{\ln \hat{\beta}\} \tag{2.26}$$

$$= \frac{K-1}{K} \sum_{i=1}^{K} \left[\ln \hat{\beta}_{\widehat{i}} - \ln \hat{\beta}_{\widehat{\cdot}}\right]^2 \tag{2.27}$$

where $\hat{\beta}_{\widehat{i}}$ is any one of $\hat{\Omega}_{0,\widehat{i}}$, $\hat{f}_{c,\widehat{i}}$, and $\hat{Q}_{\widehat{i}}$.

Tukey (1958) suggested that $(\ln \hat{\beta}_{\widehat{i}} - \ln \hat{\beta}_{\widehat{\cdot}})/\tilde{\sigma}$ is nearly distributed as Student's t with $K-1$ degrees of freedom for small samples. Hinkley (1977) on the other hand stated that if the data have strongly nonnormal distributions, the Student t approximation can lead to substantial errors. However, if the transformation performed leads to more nearly normal distributions, the approximation is reasonably accurate (Davidson and Hinkley, 1997). Note that this distribution is very close to the Gaussian distribution and for 30 or more degrees of freedom they are almost indistinguishable. With this in mind, the double-sided $1-\alpha$ confidence interval of the long period amplitude is

$$\hat{\Omega}_0 \, e^{-t_{K-1}(1-\alpha/2)\tilde{\sigma}} < \hat{\Omega}_0 \leq \hat{\Omega}_0 \, e^{t_{K-1}(1-\alpha/2)\tilde{\sigma}} \tag{2.28}$$

and similar for $\hat{f}_c$ and $\hat{Q}$. If the Student t approximation is not appropriate for the particular data, one can always simply plot the $\pm\tilde{\sigma}$ bounds by adjusting (2.28). Note that because of the transformation, the lower limit is never negative.

2.4.3 Seismic Moment, Source Radius and Stress Drop

Other important source parameters estimated from the spectrum are the seismic moment (M_0), the source radius (r) and the stress drop ($\Delta\tau$) and are often calculated assuming a circular fault (Brune, 1970; Madariaga, 1976), in which case

$$M_0 = \frac{4\pi\rho c^3 R \hat{\Omega}_0}{U_{\theta\phi}} \tag{2.29}$$

$$r = \frac{k\beta}{\hat{f}_c} \tag{2.30}$$

where a constant rupture velocity is assumed. From the mean estimates of the previous two equations,

$$\Delta\tau = \frac{7M_0}{16r^3} \tag{2.31}$$

where ρ, c, R, $U_{\theta\phi}$, β are density, wave velocity, hypocentral distance, the mean radiation pattern (0.52 and 0.63 for P and S waves) and the shear wave velocity at the source. k is 0.32 and 0.21 for P waves and S waves respectively, assuming the rupture velocity is 0.9β (Madariaga, 1976).

We will assume that the parameters not associated with the source (shear wave speed, density, etc.) are known exactly, that is, do not contribute to the uncertainties of seismic moment, source radius and stress drop. We will use the idea of propagation of errors (Taylor, 1997) to obtain confidence limits of these parameters.

We perform the propagation of errors in the log domain, since it is where we have variance estimates of $\ln \Omega_0$ and $\ln f_c$, denoted respectively $\sigma^2_{\Omega_0}$ and $\sigma^2_{f_c}$. The idea is to obtain the variance of seismic moment $\sigma^2_{M_0} = var\{\ln M_0\}$, source radius $\sigma^2_r = var\{\ln r\}$, and stress drop $\sigma^2_{\Delta\tau} = var\{\ln \Delta\tau\}$. After this, equation (2.28) can be used to obtain confidence intervals. Some rules of propagation of errors are shown in the appendix.

The relation of errors between the source and spectral parameters are

$$\sigma^2_{M_0} = \sigma^2_{\Omega_0} \tag{2.32}$$

$$\sigma^2_r = \sigma^2_{f_c} \tag{2.33}$$

and a more complicated relation is obtained for the stress drop, since it depends on two variables

$$\begin{aligned} \sigma^2_{\Delta\tau} &= \sigma^2_{M_0} + 9\,\sigma^2_r \\ &= \sigma^2_{\Omega_0} + 9\,\sigma^2_{f_c} \end{aligned} \tag{2.34}$$

where it is assumed that the covariance of Ω_0 and f_c is negligible. This relation was used by Fletcher *et al.* (1984) to estimate uncertainties of stress drop using a multiplicative error. Again, the bounds (either $\pm\sigma$ or confidence intervals using the Student t approximation) can be transformed back to the linear domain using (2.28).

2.5 Application to Cajon Pass Data

Because attenuation can also cause fall-off at high-frequencies it is important to correct observed spectra for Q effects. Full consideration of these effects is beyond our

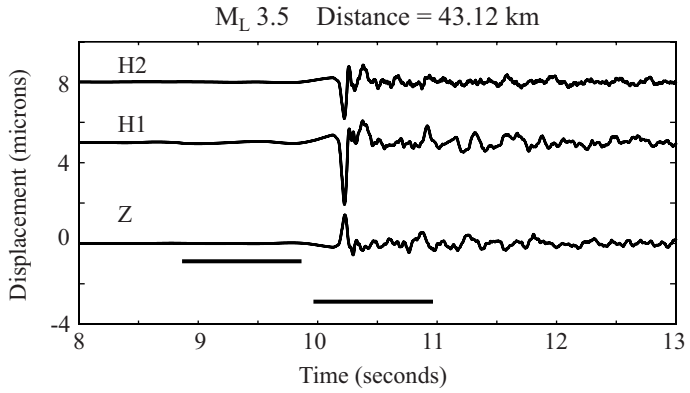

Figure 2.3: Example seismogram of the largest event used in this study, $M_L3.5$ recorded at the deepest borehole sensor 2.9 km. Seismograms have been corrected for instrument response and are flat to displacement between 2 and 300 Hz. Horizontal bars show the choice of noise and P wave window.

focus here, therefore as a demonstration of the jackknife procedure to obtain variance and confidence intervals we choose seismograms recorded at two different depths (1.5 and 2.9 km) at the Cajon Pass Borehole, where attenuation effects are relatively small and have been previously modeled by Abercrombie (1995). The seismometers that recorded this data set are 10-Hz L-15LA high temperature geophones, with sample rates of 1000 samples/sec.

Figure 2.3 shows displacement seismograms recorded at the 2.9 km depth sensor for a $M_L3.5$ earthquake 43 km away (ID 02 in Table 2.1). The spectrum is computed for a 1 second window, starting 0.15 seconds before the P pick at the station, similar to windows used in previous work (e.g., Abercrombie, 1995; Prieto *et al.*, 2004; Abercrombie and Rice, 2005) for small earthquakes.

For each of the three components (Z, H1, H2) we estimate the amplitude spectrum using a time-bandwidth product $NW = 4$ and work with $K = 7$ tapers. This means we also compute for each component 7 delete-one spectra as in equation (2.21). The final amplitude spectrum is then computed by vector summation of the three component spectra, and similarly for the delete-one spectra.

Figure 2.4 shows the complete set of delete-one spectra for the $M_L3.5$ earthquake recorded at the 2.9 km sensor. The spectra have been shifted for comparison. In general the spectral shapes are very similar and only slight differences at very low

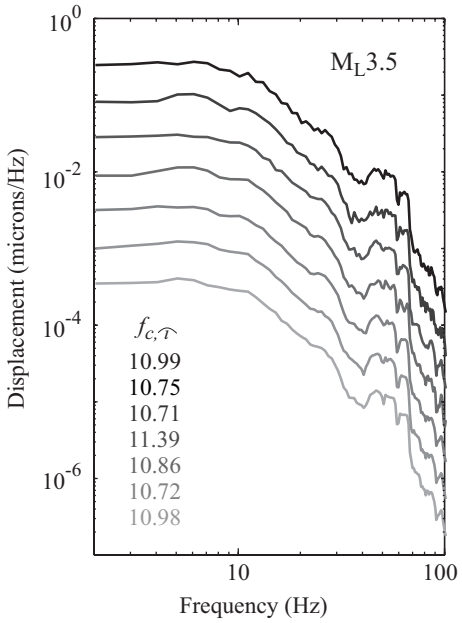

Figure 2.4: The delete-one spectrum for the $M_L3.5$ earthquake, from signal in figure 2.3 at the 2.9 km sensor. The spectra have been shifted for comparison purposes. All spectra show a similar behavior and slight differences are seen. For each spectrum, the delete-one corner frequency estimates are listed.

frequencies and roughness at higher frequencies are visible. Note that in Figure 2.4 the delete-one spectra are plotted, which are not independent estimates. Only $\hat{S}_k(f)$ are treated as independent, given the orthogonality properties of the Slepian tapers.

Here we have a good example of the properties of the multitaper algorithm. The data from the Cajon Pass have a very strong 60 Hz signal. In this work, we have $N = 1000$ samples, $dt = 0.001$, we chose $NW = 4$, the band $W = 4$ Hz and we use $K = 7$ tapers. This means that the 60 Hz peak will be smoothed over the band between $56 - 64$ Hz. In Figure 2.4 we can see a very sharp discontinuity at 64 Hz due to the fact that outside the band, very little energy is leaked to frequencies $f > 64$ Hz.

Following Abercrombie (1995) and Abercrombie and Rice (2005) we use $Q = 1000$ and correct the spectra before performing spectral fitting. We vary the fitting bandwidth to obtain optimal fits, but the same bandwidth is used for all delete-one spectra and the average spectra. An example fit to the model (2.25) is shown in Figure 2.5 for the largest and smallest earthquake in this study. Note however that the 1.5 km sensor data are used for the small earthquake, due to complicated resonances present at

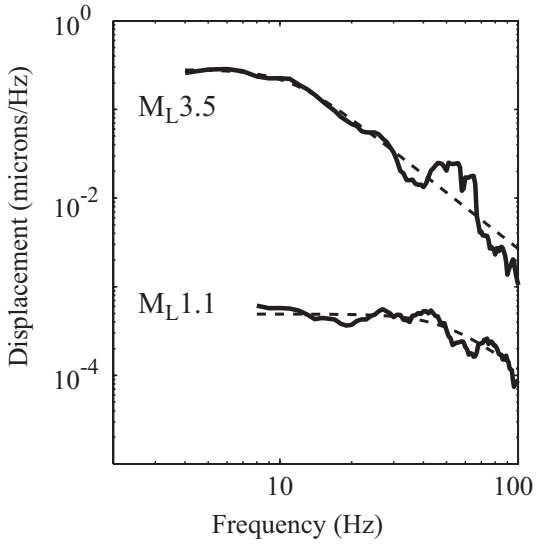

Figure 2.5: Spectral modeling of P waves for two different sized earthquakes. The bandwidth use to fit the source model varies depending on event size. Attenuation correction was previously performed, using $Q = 1000$ as suggested by Abercrombie (1995).

the 2.9 km sensor spectra that may have affected the results.

As in Abercrombie and Rice (2005) we use density $\rho = 2700$ kg/m^3, $\alpha = 6000$ m/s and $\beta = \alpha/\sqrt{3}$ and using equations (2.29 to 2.31) we estimate the source parameters $M_0, f_c, r, \Delta\tau$ from the average spectra and the jackknife parameters (e.g., $M_{0,\widehat{i}}$) from the delete-one spectra to get the jackknife variance and confidence intervals. Table 2.2 shows the source parameters and 5-95% confidence limits. Figure 2.6 shows plots of seismic moment and corner frequencies and seismic moment and stress drop for the data used in this paper.

It is important to note the assumptions and unknowns in the calculations. For example, we have assumed that the wave speeds (α, β) are known exactly. If there are errors (and certainly there are) associated with the wave speed, errors will propagate to M_0, r and subsequently to the stress drop. Assuming a 5% error in the S wave speed, thus affecting the radius uncertainties (and rupture speed), the confidence region for the stress drop for the M_L3.5 earthquake recorded at 2.9 km sensor would be (47, 108), a change of about 10%. Other sources of errors for this example include the attenuation correction, the constant Q assumption used, earthquake location errors, radiation pattern

and directivity, etc. Perhaps most importantly, we assume the validity of the source model; our method provides an estimate of the errors in $\Delta\tau$ with respect to random fluctuations in the data but is not a test of the validity of the model itself.

2.5.1 Extension to multiple stations

A generalization of the jackknife to multiple stations is desirable. A major source of uncertainty would be directivity, because as expected from directivity, the pulse width of the source time functions is narrower in the direction of rupture and broader in the opposite direction, also changing the corner frequency (e.g., McGuire, 2004).

One approach is to treat the different station estimates of corner frequencies f_c and seismic moment M_0 as independent, and, after suitable transformations, compute confidence intervals as explained in sections 2 and 4. If the earthquake source spectrum is simple (no directivity effects, radiation pattern correctly accounted for, etc.) all stations would return a similar estimate, thus having small uncertainties, while if there is strong directivity, certain stations will have considerably different corner frequencies, increasing uncertainties. A recipe for a multiple station jackknife of f_c is as follows:

(1). Compute f_c for every station that recorded the earthquake.

(2). Use the log transformation.

(3). Compute a mean $\ln \hat{f}_c$ (eq. 2.1).

(4). Compute delete-one $\ln \hat{f}_{c,\widehat{i}}$ (eq. 2.2).

(5). Compute variance $var\{\ln f_c\} = \sigma^2_{f_c}$ (eq. 2.3).

(6). Obtain confidence intervals (eq. 2.28).

A similar approach could be used for other parameters such as long period amplitude Ω_0, Q, etc. Propagation of errors (eqs 2.32 to 2.34) is necessary to obtain confidence intervals on seismic moment, source radii and stress drop.

Table 2.2: Source parameters and confidence intervals obtained by spectral fitting and jackknife analysis.

ID	M_L	Sensor	M_0 (Nm)	M_0 conf	f_c (Hz)	f_c conf	n	r(m)	$\Delta\tau$ (MPa)	$\Delta\tau$ conf
01	1.1	1.5	0.96e+11	(0.78 1.17)e+11	43.80	(33.29 57.63)	1.72	25	2.60	(1.11 6.07)
02	3.5	1.5	1.73e+14	(1.11 2.68)e+14	6.92	(4.77 10.03)	1.72	160	18.34	(5.75 58.47)
02	3.5	2.9	1.72e+14	(1.35 2.19)e+14	10.91	(9.91 12.00)	2.09	101	71.55	(52.78 96.98)
03	2.3	2.9	1.53e+12	(1.31 1.79)e+12	20.74	(17.46 24.63)	1.77	53	4.55	(2.72 7.64)
04	2.3	2.9	1.36e+12	(1.00 1.85)e+12	21.66	(18.74 25.02)	2.03	51	4.66	(2.90 7.50)

M_0 conf, f_c conf and $\Delta\tau$ conf are the 5-95% confidence intervals for the seismic moment, corner frequency and stress drop respectively. n is the high-frequency fall-off for the P waves. $Q = 1000$ is assumed. The column named Sensor represents the depth in km. of the sensor recording the earthquake.

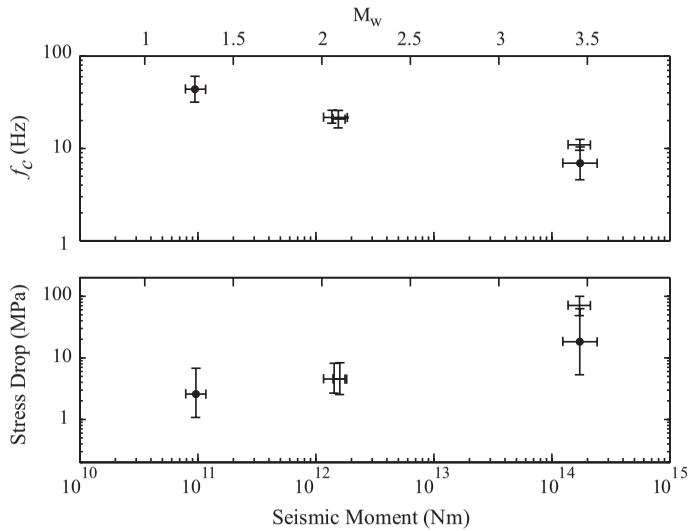

Figure 2.6: Source parameters and 5-95% confidence intervals using a Student t approximation. Symbols with dot at the center have parameters determined at the 1.5 km deep sensor. Note that the confidence regions may vary from event to event.

2.6 Conclusions

In estimating source parameters from the seismic spectrum, it is important not only to obtain a measure of the source parameters but also to obtain a measure of the uncertainties, by means of confidence intervals. The jackknife is a reliable way of estimating the variance of source parameters, and, given suitable transformations, confidence intervals. It should be easy to extend this approach to multiple station studies, where other sources of error include the radiation pattern and directivity effects which might generate different corner frequencies and radiated energy.

We calculate source parameters and confidence intervals for four small earthquakes as an example of the use of the jackknife approach. The error analysis is necessary if the data are to be used to constrain rupture models (Abercrombie and Rice, 2005), examine scaling relations and the size dependence of earthquake parameters, in order to conclude, within a reasonable reliability, something about the physics of earthquakes.

From Figure (2.6) there appears to be a slight increase of stress drop with earthquake magnitude, which, unless the errors are kept small, would pass unnoticed. Note also that the larger uncertainties are associated with the 1.5 km sensor, compared to at 2.9 km. If some of the assumptions such as radiation pattern and earthquake location contribute to the errors, the stress drop scaling would be less apparent, suggesting the need to find ways of reducing uncertainties.

The M3.5 earthquake (ID02) was recorded at two different depths and the corner frequency confidence regions barely overlap. This variation is likely explained by other sources of error such as near site effects at the shallower station. This also shows that even stations close to each other may have very different estimates of source parameters and uncertainties are needed to address the significance of these estimates.

Acknowledgments

We thank R. E. Abercrombie for help on using the Cajon Pass data and two anonymous reviewers for their constructive comments. Funding for this research was provided by NSF Grant number EAR0417983. This research was also supported by the Southern California Earthquake Center. SCEC is funded by NSF Cooperative Agree-

ment EAR-0106924 and USGS Cooperative Agreement 02HQAG0008. The SCEC contribution number for this paper is 979. We also thank IRIS and the staff at the Data Management Center for access to the Cajon Pass data.

3

Earthquake source scaling and self-similarity estimation from stacking *P* and *S* spectra

We study the scaling relationships of source parameters and the self-similarity of earthquake spectra by analyzing a cluster of over 400 small earthquakes (M_L = 0.5 to 3.4) recorded by the ANZA seismic network in southern California. We compute P, S and pre-event noise spectra from each seismogram using a multitaper technique and approximate source and receiver terms by iteratively stacking the spectra. To estimate scaling relationships, we average the spectra in size bins based on their relative moment. We correct for attenuation by using the smallest moment bin as an empirical Green's function (EGF) for the stacked spectra in the larger moment bins. The shapes of the log spectra agree within their estimated uncertainties after shifting along the ω^{-3} line expected for self-similarity of the source spectra. We also estimate corner frequencies and radiated energy from the relative source spectra using a simple source model. The ratio between radiated seismic energy and seismic moment (proportional to apparent stress) is nearly constant with increasing moment over the magnitude range of our EGF corrected data (M_L = 1.8 to 3.4). Corner frequencies vary inversely as the cube root of moment, as expected from the observed self-similarity in the spectra. The ratio between P and S corner frequencies is observed to be 1.6 ± 0.2. We obtain values for absolute moment and energy by calibrating our results to local magnitudes for these earthquakes. This yields a S to P energy ratio of 9 ± 1.5 and a value of apparent stress of about 1 MPa.

3.1 Introduction

A major question in seismology is whether the faulting mechanism of large and small earthquakes involves different physics. That is, is a $M = 8$ earthquake just a $M = 2$ earthquake scaled upward by a large factor or is something fundamentally different occurring? Aki (1967) proposed scale invariance of the rupture process, consistent with observations that many geological processes are similar over a wide range of scales (Abercrombie, 1995). There is currently a debate regarding whether earthquakes are truly self-similar over their entire size range or if systematic departures from self-similarity are observed (see, for example Abercrombie, 1995; Ide and Beroza, 2001). Thus, although many mechanisms have been proposed for differences in the physics of larger earthquakes, including shear melting (Jeffreys, 1942; Kanamori and Heaton, 2000), acoustic fluidization (Melosh, 1979), rough fault sliding-induced normal stress reduction (Brune *et al.*, 1993), fluid pressurization (Sibson, 1973), and elastohydrodynamic lubrication (Brodsky and Kanamori, 2001), the need for different mechanisms is not yet firmly established.

Studies of earthquake scaling generally involve comparisons between static measures of size (e.g., moment) and dynamic measures of size (e.g., energy). Both measures are typically derived from spectra of seismograms recorded at some distance from the earthquakes. Because moment is obtained from the low frequency part of the spectra, it is usually measured much more reliably than energy or corner frequency measurements, which require the high-frequency part of the spectra where correcting for attenuation and other path effects can be difficult. Current estimates of seismic moment made independently from local, regional and teleseismic data usually agree within about a factor of two. In contrast, estimates of seismically radiated energy by different investigators for the same earthquake often differ by more than an order of magnitude (e.g., Singh and Ordaz, 1994; Mayeda and Walter, 1996).

This uncertainty in seismic energy leads to different interpretations of the energy density of earthquakes, as measured by the Energy/Moment ratio, which is often scaled by rigidity to represent the apparent stress. Several authors find evidence that apparent stress increases with magnitude (Kanamori *et al.*, 1993; Abercrombie, 1995;

Mayeda and Walter, 1996; Izutani and Kanamori, 2001; Mori *et al.*, 2003) while others argue that apparent stress is approximately constant (Choy and Boatwright, 1995; McGarr, 1999; Ide and Beroza, 2001; Ide *et al.*, 2003). Constant apparent stress implies similar physics for small and large earthquakes, while increasing apparent stress with magnitude implies that large earthquakes are more efficient radiators of seismic energy than small ones.

Our approach here is to improve the reliability and stability of source spectra by stacking and averaging thousands of records from the Anza seismic network in southern California. We use a simple method (Warren and Shearer, 2000, 2002) to isolate the relative source spectra from the path and site effects by stacking the computed log spectra after subtracting the appropriate path-site terms. This approach is also similar to that used by Andrews (1986) to analyze spectra of the 1980 Mammoth Lakes California earthquake sequence. Rather than obtaining an absolute measure of individual source spectra, we obtain relative shapes of spectra with respect to other earthquakes. We then stack the spectra in bins of similar moment to obtain average spectra (and estimated uncertainties) as a function of earthquake size and apply attenuation corrections using the smallest earthquakes as empirical Green's functions (EGF) (e.g., Mueller, 1985; Hough, 1997). The resulting spectra are sufficiently smooth that direct tests of the self-similarity hypothesis are possible, as well as measurements of corner frequency and energy. All of our results indicate self-similarity is closely obeyed over the $M_L = 1.8$ to 3.4 size range of our EGF corrected data.

3.2 Data Processing

We used records from the Anza seismic network (Berger *et al.*, 1984; Vernon, 1989) [*Berger et al.*, 1984; *Vernon*, 1989], 9 high-quality, three-component stations located on hard rock sites near an active part of the Clark Lake segment of the San Jacinto fault in southern California (Figure 3.1). We began this study by selecting about 800 earthquakes located in a tightly clustered volume (4.5 km sided area, with most of the events between 5 and 12 km depth) near the Toro Peak station (TRO) and 50 km from the most distant station (RDM). In this region, the database is complete to about $M_L \geq 0.5$

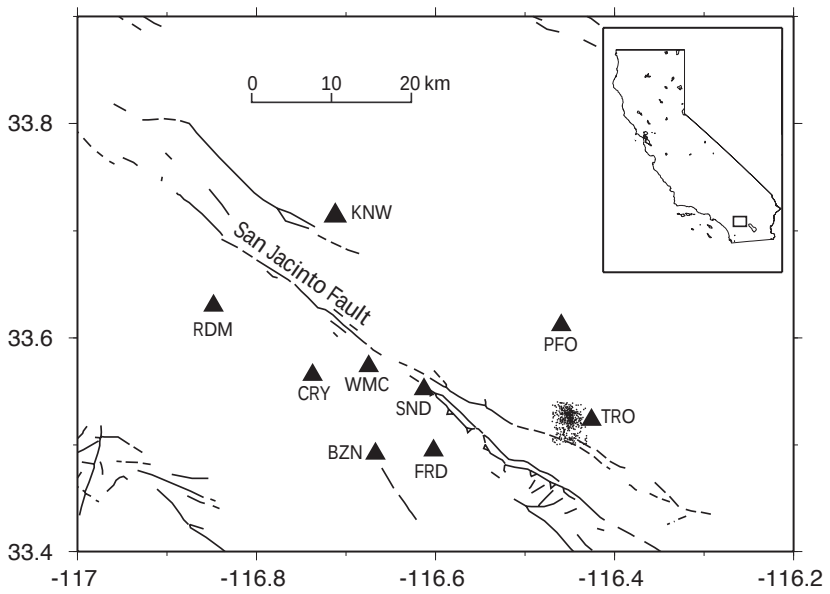

Figure 3.1: Map showing the cluster of over 400 earthquakes (small black dots) and ANZA stations (solid triangles) used in this study. The inset shows the location of the study area (rectangle) in the state of California.

with generally good signal-to-noise ratio records. The earthquakes occurred from 1983 to 1993, at which time the network recorded at 250 samples per second with Geospace HS–10 2–Hz seismometers. We selected a relatively compact group of earthquakes so that the path to each station would be similar between different earthquakes, permitting the use of simple corrections for attenuation and other path effects.

We use both P and S waves and select time windows for P on the vertical component and time windows for S on all three components. Both windows start 0.5 seconds before the analyst pick of the arrival, with a total window length of 1.28 seconds. We also select a noise window of the same length, with the last data point just before the P-wave window. The velocity spectrum is estimated using the multi-taper algorithm (Park *et al.*, 1987b) and then corrected for the instrument response function. The S wave spectrum is calculated as the vector summation of spectra from all three components. Figure 3.2 shows an example of this process for a vertical-component record from station FRD.

We apply a signal-to-noise ratio cutoff, where we use spectra only when the mean ratio is greater than 5.0 in the 0–80 Hz frequency band and the ratio is greater

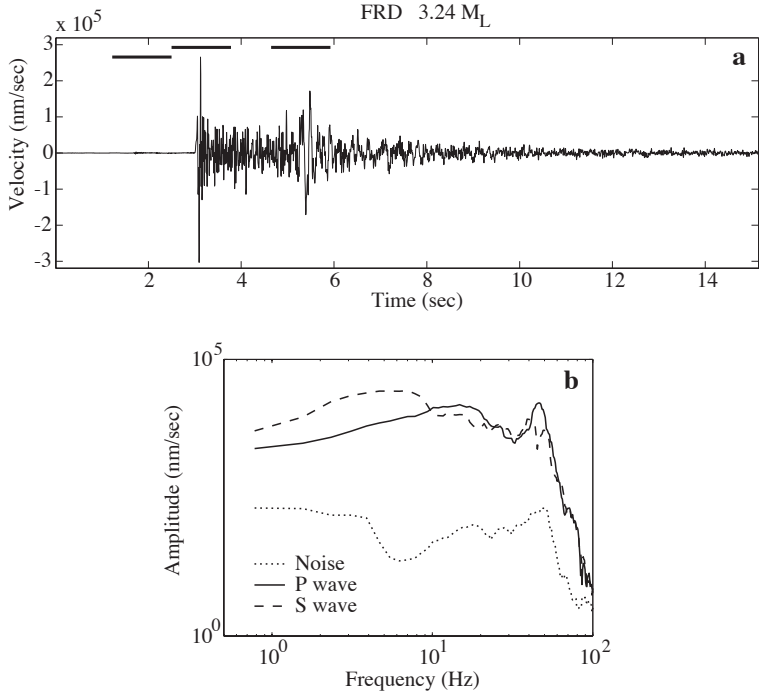

Figure 3.2: Example of computed spectra from the largest magnitude earthquake in the study area ($M_L = 3.4$) recorded at station FRD, vertical component. (a) The time series, with horizontal bars showing the noise, P, and S windows used to compute the spectra; in this case the S-wave shows up more clearly on the horizontal components. (b) Spectra for the windows shown in (a), computed using a multitaper method. Note the rapid decrease in signal to noise ratio at the higher frequencies.

than 3.0 at 80 Hz. At higher frequencies, the signal to noise ratio decreases very rapidly (see Figure 3.2), so we limit our analysis to frequencies below 80 Hz. After applying the cutoff, we have 2735 records (including both P and S waves) from 470 earthquakes. Because of their larger amplitudes, the S waves have generally higher signal-to-noise ratios than the P waves; thus our signal-to-noise cutoff excludes P waves from the smallest earthquakes in our data set, which are represented only by S wave spectra.

One possible concern is that P-wave coda may be contaminating the S-wave window. This potential source of bias is likely to have its largest effect on the closest stations where the $S - P$ time is the smallest. To test what effect this may be having on our results, we repeated our analyses using subsets of the data where we removed the closest, the two closest and the four closest stations from the source region. Although there was some increase in the variability of the stacked spectra as we reduced the number

of data in the stacks, there were no systematic changes in the S spectral shapes. Thus it does not appear that P contamination of S is a significant factor in our analyses.

Since multiple stations record every earthquake and many earthquakes are recorded at each station, we can isolate the source and receiver contributions to the spectra. Because our source region is relatively compact, the receiver contributions will also include most of the path effects. Following the method described by Warren and Shearer (2002) it is possible to isolate the relative source spectrum (Figure 3.3) if we assume that the observed spectrum $D_{ij}(f)$ from each source and receiver (denoted S_i for the ith earthquake and R_j for the jth station) is a product of source effects and path-site effects. We iteratively stack all log spectra from each earthquake, after removing the appropriate station term, to obtain the earthquake term:

$$\log(S_i) = \frac{1}{n}\sum_{j=1}^{n}[\log(D_{ij}) - \log(R_j)] \tag{3.1}$$

and we also stack all log spectra from each station, after removing the earthquake term, to obtain the path-station term:

$$\log(R_j) = \frac{1}{m}\sum_{i=0}^{m}[\log(D_{ij}) - \log(S_i)] \tag{3.2}$$

where D_{ij} is the computed spectrum, S_i the earthquake term for the ith earthquake and R_j the path-station term for the jth station. Since the earthquake term and the path-station term are dependent upon each other, we solve the set of equations iteratively until we reach a stable result, where the fractional change in either the source or path-station terms is less than 10^{-4}. We normalize the average log source spectra for all our earthquakes to unity, as a starting point for the iteration process. In practice we are mapping the deviations of the source spectra from this reference flat spectrum.

After source and path-station terms are separated we obtain 470 relative source spectra and 9 path-station spectra (separately for P and S waves). Figure 3.4 shows the P and S path-site spectra for nine different stations. Because we have not yet assumed a source model (e.g., ω^{-2}, etc.), the shape of each of these spectra will include both source and attenuation contributions. The information is contained in the differences between these curves, which are significant because all of the stations recorded the same set of earthquakes. Variations in attenuation among the stations can be seen in the position at

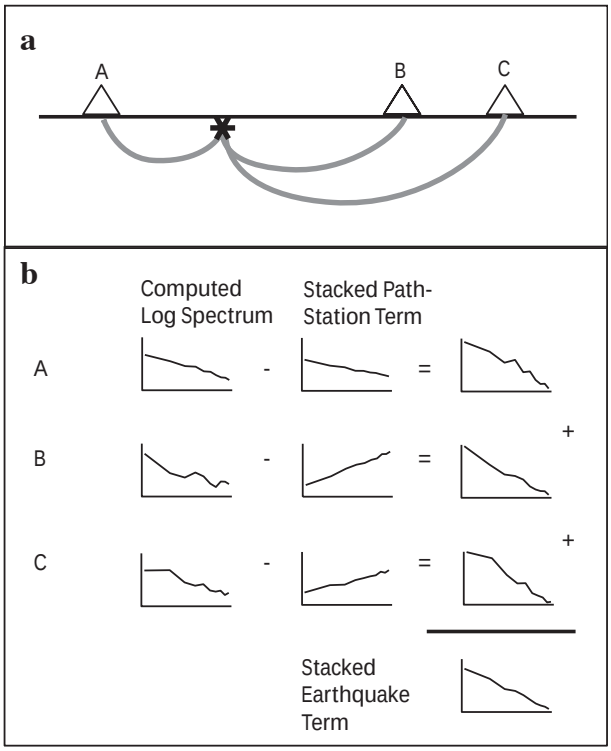

Figure 3.3: Cartoon explaining how spectral stacking is used to obtain the earthquake term, as in Warren and Shearer (2002). If a given earthquake (star) is recorded at stations A, B, and C (Figure 3a), the earthquake term is computed by stacking the log spectrum from earthquake 1 computed for stations A, B and C after removing the path-station terms for these stations (Figure 3b). An analogous procedure is used to compute the station terms.

which the spectra begin to falloff at high frequencies. For example, it is clear that station SND, located within 100 meters of the surface trace of the San Jacinto fault, records a more attenuating path than station FRD, despite the fact that FRD is located slightly closer to the earthquake cluster. In general, there is no clear distance dependence to the observed path-station spectra, suggesting that local site effects beneath each station are dominating the spectral differences among the stations.

Each of the 470 relative source spectra represents the average log spectra of all stations recording the earthquake, after correcting for differences among the path-station terms. To study how these source spectra vary as a function of earthquake size, we divide our data into 20 bins in relative moment, which is estimated from the low-frequency

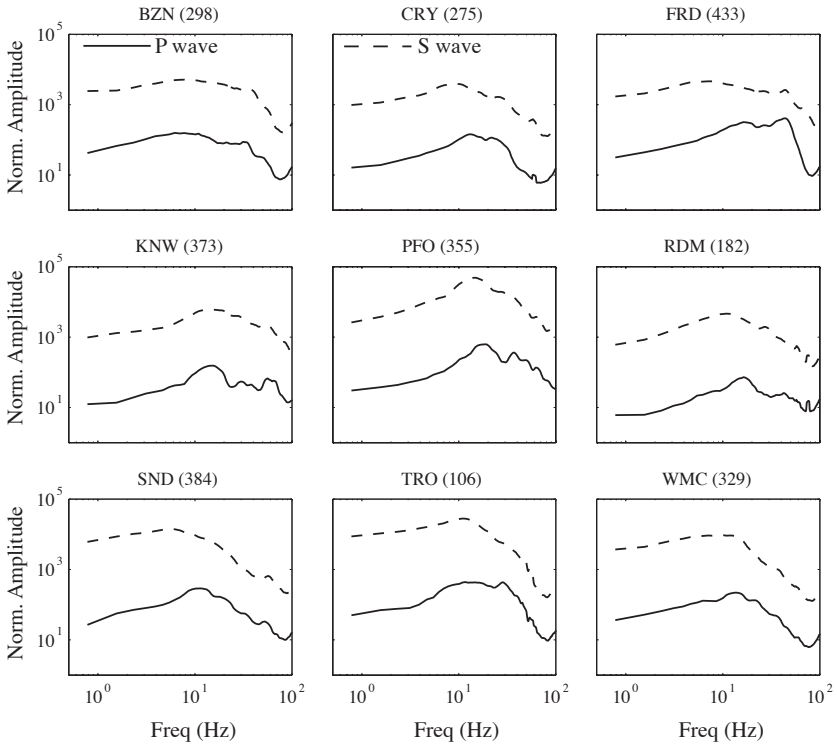

Figure 3.4: Examples of path-station terms for P (solid line) and S (dashed line) waves. In parenthesis is the number of earthquakes recorded at each particular station. Note that the path-station term for the S-waves is always larger, reflecting the higher amplitude of S compared to P.

spectral amplitude. Because our S-wave data span a larger total moment range than the P-wave data, the moment range within each S-wave bin is larger than in the P-wave bins. At this stage in our analysis, we do not compare P and S amplitudes directly; rather we process the P and S spectra separately and obtain independent results for each phase (later comparisons between P and S corner frequencies will jointly consider the data).

Selected S wave results are plotted in Figure 3.5a. Each binned source spectra is the result of averaging between 1 and 86 earthquake spectra (each of which is itself a stack of spectra from different stations recording the earthquake). The resulting binned source spectra are much smoother than the individual spectra that go into the stacks. There are generally many more earthquakes in the bins at smaller moments because of the much larger number of smaller earthquakes in the data set. The relative moments among the bins can be seen in the low frequency limit of the spectra (i.e., at about

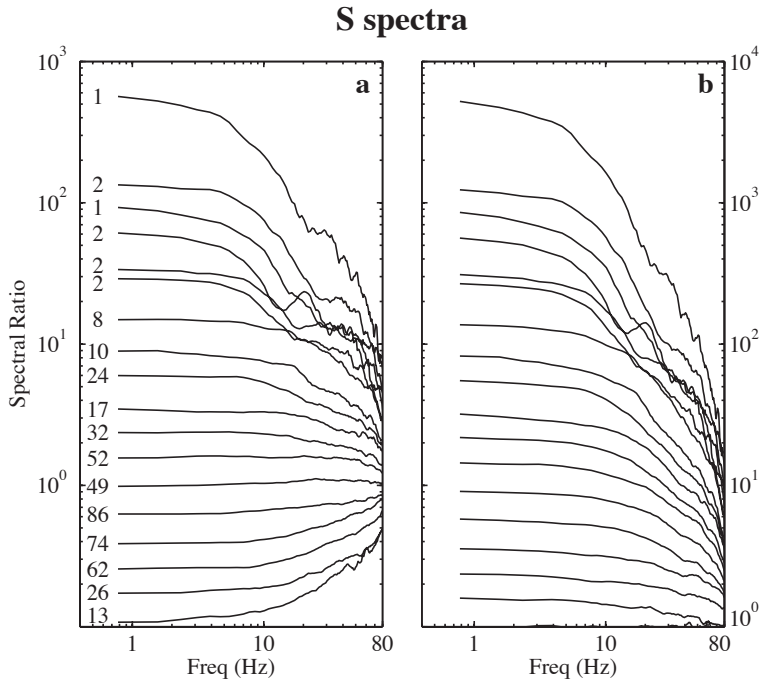

Figure 3.5: Relative source spectral shapes for some selected bins. In Figure 3.5a the shapes are relative to the average spectra, which was forced to have a constant log amplitude of 1; numbers to the left of each spectra indicate the number of earthquakes in that particular bin. Figure 3.5b shows the corrected source shapes, after applying the smallest magnitude bin as an empirical Green's function (EGF). To avoid passing unwanted line components present in the smallest bin, we smoothed the reference spectrum with a 20 point moving average before subtracting it from the data.

1 Hz). These moments are not evenly spaced in Figure 3.5 because the moments of the earthquakes within each bin are not always evenly distributed. We use a bootstrap technique that randomly resamples the earthquakes within each source spectral bin in order to estimate uncertainties on the binned spectra, and later to estimate uncertainties on properties, such as corner frequency and energy, that we compute from these spectra.

As in the case of the path-site terms, the absolute shape of the spectra plotted in Figure 3.5a is unconstrained (owing to the intrinsic tradeoff between the average source spectrum and the average path-site response function). We resolve this tradeoff in our iterative method by forcing the average source spectrum to unity. This is why the spectra for the small earthquake bins curve upward at high-frequency. This indicates that, as expected, these earthquakes have a shallower falloff at high frequencies compared

to larger earthquakes. To obtain an estimate of the true spectral shapes of the sources, we use the smallest moment bin as an empirical Green's function (EGF) for all the other bins.

Figure 3.5b shows the results of subtracting the log spectra of the smallest bin from the others. These EGF corrected spectra have the features expected for source spectra—a flat response out to a corner frequency that increases with decreasing earthquake size, and a rapid falloff beyond the corner frequency. As we will discuss later this falloff closely agrees with the Brune ω^{-2} source model (Brune, 1970). We plot all of the EGF corrected spectra in Figure 3.5, but our later analyses will focus only on those bins at least one order of magnitude larger in moment than the EGF reference bin. For comparison, Mori *et al.* (2003) used a $M \sim 1.5$ smaller EGF and Frankel *et al.* (1986) used earthquakes with $M \leq 2.1$ as EGF of $M \sim 3$ earthquakes. As previously noted, due to signal-to-noise limitations, we do not use P-wave data from the smallest earthquakes so the smallest P-wave EGF bin represents the same bin as the third smallest S-wave EGF bin.

3.3 Implications of Self-Similarity

Before further analysis, it is instructive to consider the predicted effects of earthquake self-similarity on recorded spectra (e.g., Aki, 1967). Figure 3.6 illustrates the expected change in the pulse shape and spectrum for an earthquake rupture that is increased in size by a factor b. Assuming the dimensions of the larger rupture are scaled proportionally, then the fault area, A, will increase by a factor b^2, the displacement, D, will increase by b, and the moment, $M_0 = \mu DA$, will increase by a factor of b^3. Figure 3.6b shows the resulting change in a displacement pulse, $u(t)$, recorded in the far field, assuming identical source and receiver locations and no attenuation. The exact form of the shape of this pulse depends upon details of the source, but, assuming simple scaling between the two earthquakes, the pulse shape will change in a predictable way. In particular, assuming the rupture speed is constant between the earthquakes (as simple self-similarity predicts), the pulse length will increase by a factor of b and the pulse height will increase by a factor of b^2. This is necessary in order for the moment, which

is proportional to the integrated area under the pulse, to increase by b^3.

It follows that the displacement pulse, u^*, recorded by the second earthquake can be expressed as

$$u^*(t) = b^2 u(t/b) \tag{3.3}$$

where $u(t)$ is the recorded displacement pulse of the first earthquake. The seismic energy, E_s, in the recorded pulse will be proportional to $\int \dot{u}^2(t)\, dt$ (the integrated square of the slope of the pulse), so the second pulse will contain a factor b^3 more energy than the first pulse. Thus the energy density (E_s/M_0) remains constant.

Using the similarity theorem for the Fourier transform, it follows that the spectrum of the second earthquake is given by

$$u^*(\omega) = b^3 u(b\omega) \tag{3.4}$$

where $u(\omega)$ is the spectrum of the first earthquake. This relationship predicts that the shape of all spectra on a log-log plot will be identical, but offset along a line of ω^{-3} (Figure 3.6c).

This provides a possible test of self-similarity that does not depend upon any assumptions regarding which source model is most appropriate (ω^{-2}, ω^{-3}, etc.). We perform this test (Figure 3.7) by shifting the EGF corrected spectra along an ω^{-3} line and find that the shapes are in agreement within their estimated uncertainties. Furthermore, there is no systematic dependence with moment exhibited in the alignment of the binned spectra (Figures 3.7c and 3.7d). The P-wave spectra do not align as closely as the S-wave spectra at low frequencies (≤ 1 Hz) because the individual P-wave stacks are not flat at low frequencies (Figure 3.7a). Although we do not fully understand the reason for this behavior, the shapes of the P spectra are nonetheless similar within their uncertainties. It is likely that this anomaly in the P-wave spectra is related to decreasing signal-to-noise ratios at low frequencies, which could bias the EGF reference stack because it is derived from the smallest earthquakes.

The S-wave spectra are noticeably smoother and provide our most reliable constraints on the similarity of the spectra as a function of moment. This is the most fundamental result in our study and suggests the self-similarity hypothesis is valid for our data set. The great advantage of this analysis is that we can check if spectral shapes

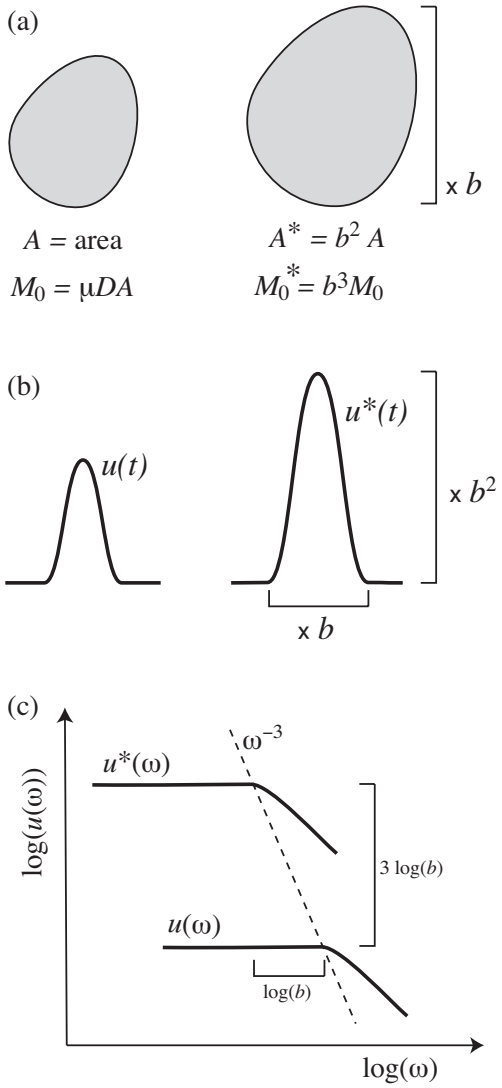

Figure 3.6: An illustration of the effects of self-similarity when an earthquake is increased in size by a factor of b. (a) The rupture area increases by b^2, the displacement by b and the moment and energy by b^3. (b) A recorded far-field displacement pulse will increase in length by b and in height by b^2. (c) Log-log plots of the spectra will have identical shapes, but shifted along an ω^{-3} line.

are self-similar or if there are systematic differences in the shapes as magnitude increases, without assuming a particular model of corner frequency and high-frequency falloff. In contrast, conventional methods for making inferences about source scaling are heavily focused on parametric data derived from the spectra rather than the spectra themselves. These parameters do, however, provide further insight regarding source properties.

Implicit in our spectral comparisons is that the focal mechanisms and rupture directions do not vary systematically between smaller and larger earthquakes because this could bias the results obtained at particular stations. We have not examined the focal mechanisms for our earthquakes but have no evidence that this is the case. Such problems are likely to be minimized in our analysis because we are averaging results from many stations at different azimuths and distances from the earthquakes. Furthermore it seems unlikely that these possible biases would have the effect of producing apparent self-similarity in our measured spectra without self-similarity being present in the earthquakes themselves.

3.4 Source Parameter Modeling

The source parameters seismic moment (M_0), corner frequency (f_c) and radiated energy (E_s) can be estimated from the source spectra and are important in the understanding of the physics of the earthquake source, as well as for computing apparent stress (σ_a), defined as $\mu E_s/M_0$, where μ is the rigidity. If self-similarity holds, as tested in the previous section, this ratio should remain constant over the same range of magnitudes.

We initially fit our P and S displacement spectral stacks with a general source model (e.g., Abercrombie, 1995):

$$u(f) = \frac{\Omega_0}{\left[1 + \left(\frac{f}{f_c}\right)^{\gamma n}\right]^{1/\gamma}} \tag{3.5}$$

where Ω_0 is the long-period amplitude (relative seismic moment), f is the frequency, f_c is the corner frequency, n is the high-frequency falloff rate and γ is a constant. We allowed the values of n to vary while using γ=1 as well as γ=2, that is we experimented with both the Brune (1970) and Boatwright (1980) models, allowing the falloff term to

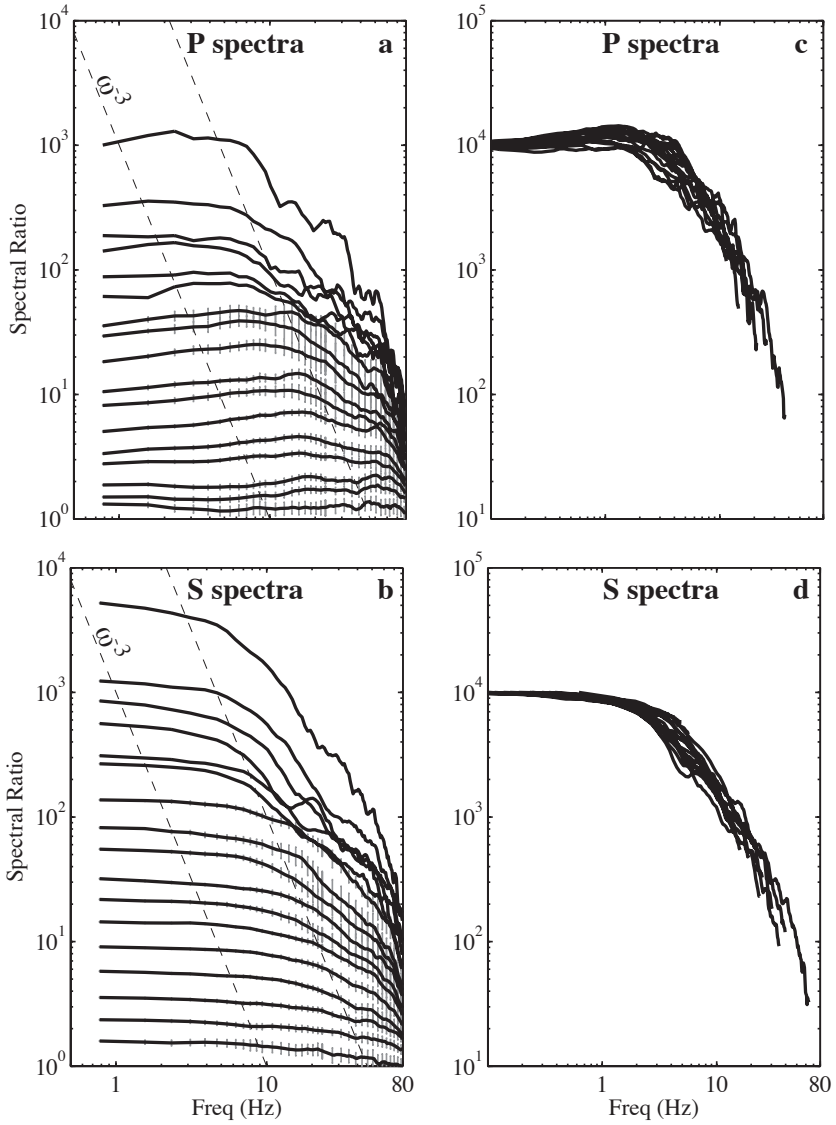

Figure 3.7: EGF corrected stacked spectra for bins of different source moment, showing the self-similarity of the spectra when shifted along an ω^{-3} line. (a) P-wave spectra and (b) S-wave spectra, with 1-σ error bars estimated using a bootstrap resampling method. The spectra shifted along an ω^{-3} line (dashed lines at left) for (c) P-waves and (d) S-waves. The spectra agree in shape within their estimated errors, consistent with the earthquake self-similarity hypothesis.

vary as well as corner frequency and relative seismic moment. We used a grid search technique to find the best-fitting set parameters (Ω_0, f_c, n, and γ). We restricted this procedure to those size bins that have relative moments ten times larger than that of the EGF bin (see Figure 3.5).

In general we found that a simple ω^{-n} model (i.e., $\gamma = 1$) worked reasonably well with values of n ranging from 1.8 to 2.2 (i.e., very close to the Brune ω^{-2} model), and that allowing additional free parameters did not significantly improve the fit. Predictions obtained using $\gamma = 2$ yielded spectra with sharper corners than are seen in our stacked spectra. It is possible that individual events have spectra with these sharp corners, but, given some variability in the positions of the corners, the corner is smoothed and widened in the stacks over many events so that the Brune model gives the best fit. We therefore used the model

$$u(f) = \frac{\Omega_0}{1 + (f/f_c)^n} \tag{3.6}$$

and solved for the best-fitting Ω_0, f_c and n for the results presented here (see Figure 3.8 for examples of the resulting fits to the stacked spectra).

The radiated seismic energy is proportional to $\int \dot{u}^2(t)\, dt$, the integrated square of the measured velocity. We perform this calculation in the frequency domain by converting the displacement spectra to velocity, squaring and integrating (e.g., following Ide and Beroza, 2001), being careful to extrapolate to very high frequencies using the model falloff rate. In this study we compute energy from the best fitting model rather than directly from the data, i.e., we use

$$I = \int_0^\infty \left[\frac{2\pi f \Omega_0}{1 + (f/f_c)^n} \right]^2 df \tag{3.7}$$

where I is the relative seismic energy. Because we are integrating the model predicted spectrum rather than the data, we can extend the upper integration limit to a sufficiently high frequency to avoid any underestimation of the energy.

3.5 Calibration to absolute moment and energy

Our results described so far involve only relative estimates of moment and seismic energy. To obtain absolute measures of these parameters directly from our data,

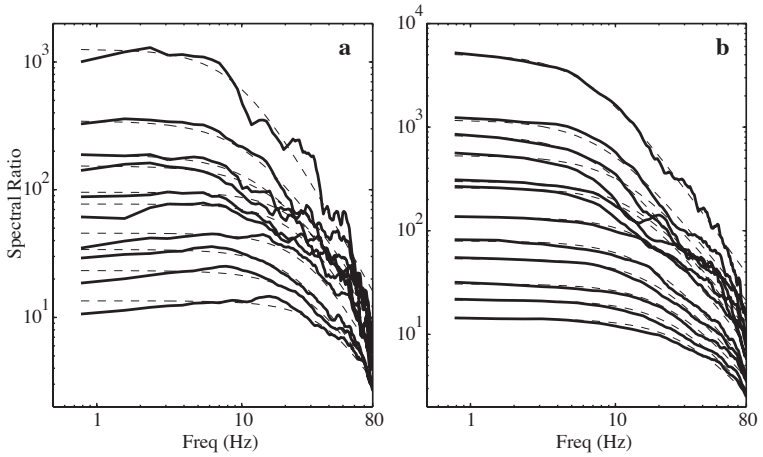

Figure 3.8: EGF corrected stacked spectra and best-fitting source models for (a) P-waves and (b) S-waves. For clarity only some of the moment bins are plotted.

we would need to apply corrections for geometrical spreading, radiation pattern, free-surface, and source-receiver impedance contrast effects. Because these corrections are often difficult to estimate precisely, this will introduce considerable uncertainty into our results. However, because our earthquakes are in a single compact region, these correction factors are likely to be highly correlated, implying that relative measures of moment and energy among our earthquakes are determined more accurately than their absolute level. Thus, our most precise results involve relative measurements among our earthquakes. However, for comparisons to other studies it is useful to have some measure of absolute moment and energy. Our approach to this problem is to exploit the fact that these earthquakes were also recorded by the Southern California Seismic Network (SCSN), which provides well-calibrated local magnitude estimates (moment is not routinely computed for earthquakes this small).

Assuming that $M_L \approx M_W$, we can estimate moment M_0 using the Kanamori (1977) relation

$$M_W = (2/3) \log_{10} M_0 - 10.7 \tag{3.8}$$

In this way, we can compute a scaling factor to relate our relative moment estimates Ω_0 to local magnitude and to true moment M_0. A comparison between SCSN mean catalog M_L versus our estimated M_L (Figure 3.9) shows a linear relationship with a slope close to unity, as expected if the 2/3 factor in (3.8) is accurate. Previous studies have shown

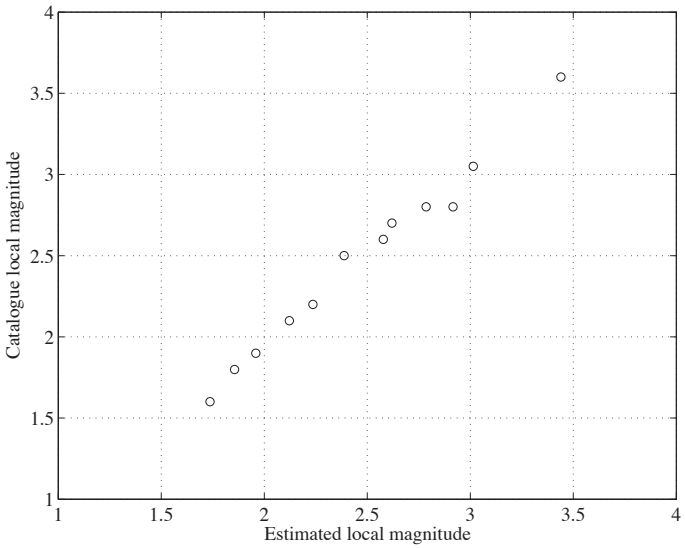

Figure 3.9: A comparison between M_L as measured by the southern California Seismic Network (SCSN) and M_L as estimated from our relative moment measures using an empirical scaling factor.

that in the magnitude range of our data set, a linear relation between $\log(M_0)$ and M_W fits the data in Southern California, although with some variations on the 2/3 factor (see Hanks and Boore, 1984; Abercrombie, 1996, for more detailed discussion). A change in the scaling factor would change the absolute moment magnitude after calibration, but the relative moment between the different earthquake bins should remain constant.

Now consider the theoretical relationships for M_0 and E_s for a double-couple source in the far field in a uniform wholespace. The standard formula (e.g., Aki and Richards, 1980; Kanamori and Rivera, 2004) for the moment in this case is

$$M_0 = 4\pi\rho c^3 r \, {}_cU_{\phi\theta}^{-1} \, \Omega_0 \tag{3.9}$$

where ρ is the density, c is the seismic velocity (either α for P wave or β for S wave), r is the source-receiver distance, ${}_cU_{\phi\theta}$ is the radiation pattern, and Ω_0 is the observed long-period amplitude. Now assume that we know M_0, ρ and c independently. We can then rewrite (3.9) as

$$\frac{{}_cU_{\phi\theta}}{r} = 4\pi\rho c^3 \frac{\Omega_0}{M_0} \tag{3.10}$$

Note that $1/r$ represents a geometrical spreading term that could be generalized to a more complicated model.

For the same whole-space double-couple model, the radiated seismic energy may be expressed as (e.g., Boatwright and Fletcher, 1984)

$$E_s^c = 4\pi\rho c r^2 \frac{\langle {}_cU_{\phi\theta}{}^2 \rangle}{{}_cU_{\phi\theta}{}^2} I \tag{3.11}$$

where $\langle {}_cU_{\phi\theta}{}^2 \rangle$ is the mean over the focal sphere of $({}_cU_{\phi\theta})^2$ (= 4/15 for P waves and 2/5 for S waves) and I is the measured relative energy (i.e., the integrated velocity squared). Because this equation involves the ratio of ${}_cU_{\phi\theta}$ and r we can use (3.10) to obtain

$$E_s^c = \frac{\langle {}_cU_{\theta\phi}{}^2 \rangle}{4\pi\rho c^5} M_0^2 \frac{I}{\Omega_0^2} \tag{3.12}$$

which is independent of the geometrical spreading and radiation pattern terms. This equation remains accurate if free-surface corrections are applied or if the instrument gain is incorrectly known, provided M_0 is determined independently. In the case where ρ and c vary between source and receiver, carrying through the impedance correction terms shows that (3.12) is still valid provided ρ and c are taken at the source.

Because the estimated energy varies inversely as c^5, the results are very sensitive to errors in velocity at the source. A 15% error in c will produce about a factor of two error in E_s. In this study we use $\alpha = 6.0$ km/s, $\beta = \alpha/\sqrt{3}$, and $\rho = 2.7$ kg/m^3, which leads to the value $\mu = 3.24 \times 10^{10}$ Pa. The values of velocity are very close to those from a 3-D seismic velocity inversion (Scott *et al.*, 1994) for the source region. We estimate the uncertainty in our source velocity estimates to be less than 5%.

The total radiated seismic energy is obtained by adding the energy for P and S waves

$$E_s = E_s^P + E_s^S \tag{3.13}$$

Finally it is important to recognize that absolute energy estimates are also very sensitive to attenuation corrections. We assume here that the EGF approach has correctly removed attenuation effects, but this remains another possible source of uncertainty in our results.

3.6 Results for corner frequency and apparent stress

To compare P and S corner frequencies, we performed a separate analysis in which the relative moment of each earthquake was estimated from both the P and S

spectra so that the same earthquakes would be contained in each moment bin. As discussed in the previous section, this relative moment will later be calibrated with estimates of moment from local magnitude determinations. We find that the P wave corner frequencies determined here are systematically higher than those estimated for S waves from the same earthquakes (Figure 3.10). The ratio $f_c(P)/f_c(S)$ is about 1.6 (individual measurements range from 1.3 to 2.0), consistent with the model of Madariaga (1976) and very close to values determined using borehole recordings at 2.5-km depth in the Cajon Pass, California by Abercrombie (1995). This ratio is likely to correspond principally to source effects since attenuation and other path-site effects have been removed.

As analyzed by Abercrombie (1995) the ratio of S to P wave energy (known as q) is also very important. From Boatwright and Fletcher (1984) we have

$$q = \frac{3}{2}\left(\frac{\alpha}{\beta}\right)^5 \left(\frac{f_c(S)}{f_c(P)}\right)^3 \tag{3.14}$$

where $f_c(S)$ and $f_c(P)$ are the corner frequencies for S and P (which are assumed to have the same falloff rate at high frequencies). Note that $q = 23.4$ for a Poisson solid if the corner frequencies are identical. In our study, $f_c(P)$ is about 1.6 times larger than $f_c(S)$, reducing the predicted value of q to about 6.

Our estimated P and S energies (Figure 3.11), calculated using equation (3.12) for the different moment bins, which yield $q = 9 \pm 1.5$, the difference from the predicted value ($q = 6$) resulting from the fact that our models permit the falloff exponent to vary slightly between P and S waves. Previous studies have found considerable variation in q estimates, as they are highly dependent upon corner frequency shifts, but our results are in reasonable agreement with, for example, Boatwright and Fletcher (1984) ($q = 13.7 \pm 7.3$) and Abercrombie (1995) ($q = 14.31$ with values from 4.43 to 46.26). We did not directly obtain P-wave energy for the two smallest spectral bins because their relative moment was not ten times larger than the P-wave EGF. To obtain P energy for these bins, we divided the S-wave energy by the $q = 9$ scaling parameter estimated from the other bins.

Another commonly applied test of self-similarity (e.g., Abercrombie, 1995; Ide *et al.*, 2003; Kanamori and Rivera, 2004) is to plot corner frequency versus seismic moment. As previously discussed, self-similarity predicts that $M_0 \propto f_c^{-3}$. We determined

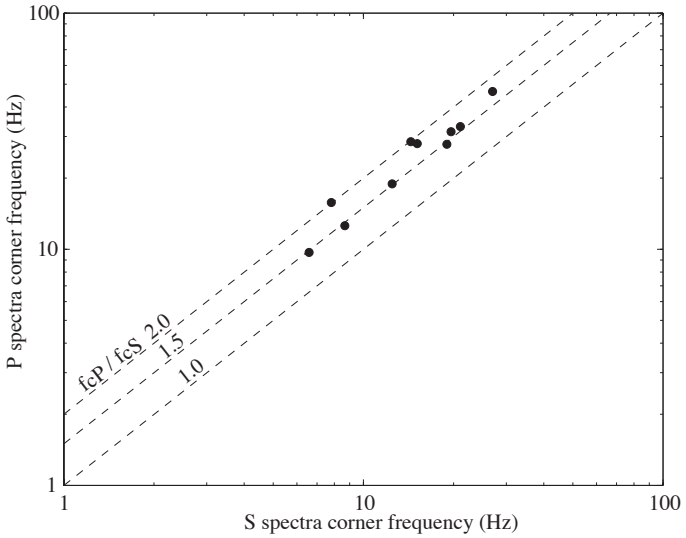

Figure 3.10: A comparison between P and S-wave corner frequencies as measured for the different moment bins. Dashed lines represent different scaling factors. The data suggest $f_c(P) \approx 1.6 f_c(S)$.

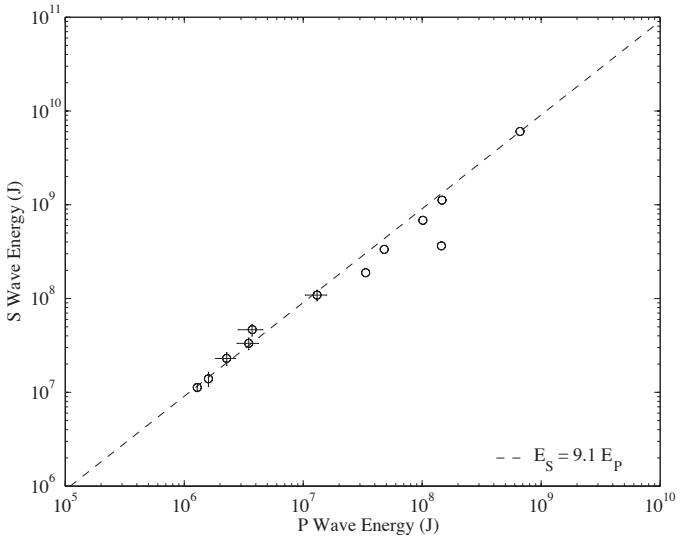

Figure 3.11: Radiated S-wave energy E_s^S versus P-wave energy E_s^P estimated from the different moment bins. The best fitting line is for $E_s^S = 9E_s^P$.

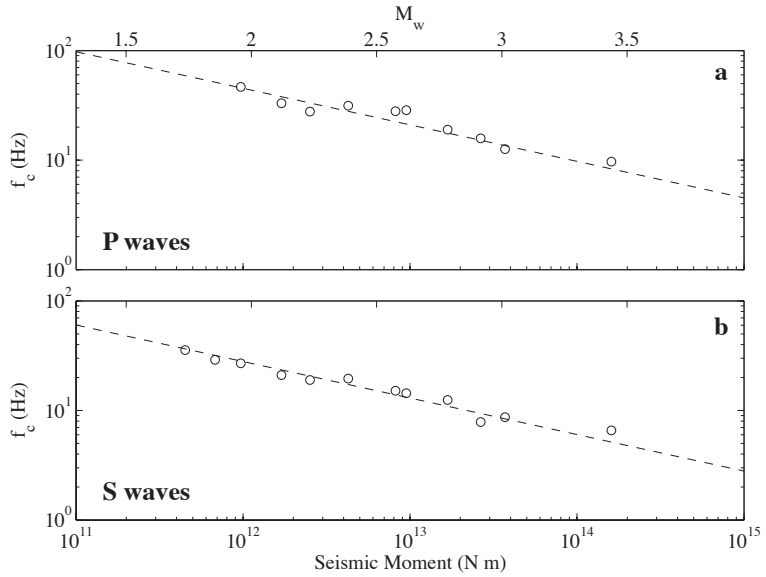

Figure 3.12: P and S corner frequencies versus moment, derived from the stacked spectra for the different moment bins. The results indicate $M_0 \propto f_c^{-3}$ as shown by the dashed lines.

the relative seismic moment and corner frequencies for P and S waves independently (see Figure 3.12). The relative moment is scaled to obtain an approximation of the absolute seismic moment (see section 3.5). Corner frequencies follow the cube root scaling expected from self-similarity, as previously observed by Ide *et al.* (2003). Of course this is not surprising, given that the spectra themselves obey self-similarity scaling (see Figure 3.7 and prior discussion). Due to the corner frequency shift for P and S waves, we plot this relationship independently. It is possible that the 40 Hz and higher corner frequencies for the smaller moment bins are constrained less accurately than the corner frequencies for the larger earthquakes because our analysis extends only to 80 Hz. However, there are more earthquakes in the smaller moment bins, resulting in smoother stacked spectra (see Figure 3.7 and 3.8), which likely improves the reliability of the corner frequency estimates even when less of the spectrum is available. It is clear from Figures 3.7 and 3.12 that the available part of the spectra are consistent with the self-similarity hypothesis.

The relationship between seismic moment and radiated seismic energy is also important and has been a focus of many previous studies. This relationship is commonly expressed in terms of apparent stress, defined as $\sigma_a = \mu E_s/M_0$. Figure 3.13 shows apparent stress plotted as a function of moment for our spectral stacks. The E_s/M_0

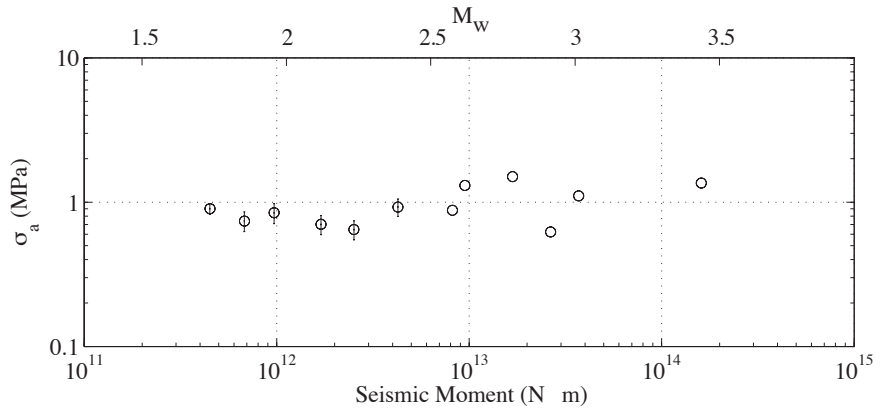

Figure 3.13: Apparent stress σ_a versus moment for the different moment bins. Apparent stress is nearly constant over $M_W = 1.8$ to 3.4, with an average value of about 1 MPa.

ratio is approximately constant as moment increases, as predicted if self-similarity is obeyed and apparent stress is constant as a function of earthquake size. Given the scatter in our data, a small degree of scaling is possible. A weighted least squares fit to the points in Figure 3.13 results in $E_s/M_0 \propto M_0^{0.08\pm0.10}$, providing relatively tight error bounds that include the zero exponent result expected from self-similarity.

The average apparent stress is about 1 MPa, but for the reasons discussed above, this value is less well constrained than the relative σ_a between our different earthquake bins. The two largest sources of error in our absolute σ_a estimate are likely to be: (1) our calibration factor between relative moment and M_W, and (2) the assumed S velocity at the source. Any calibration factor error will scale directly as M_0. From (3.8) we see that if, for example, our M_W estimates (assumed equal to the SCSN M_L values) are 0.2 units too large, this will result in σ_a estimates that are about two times too large. A 5% error in our assumed source S velocity will yield about 30% uncertainty in σ_a. Given these uncertainties and the scatter shown in Figure 3.13, a reasonable range for the possible values of the average apparent stress is 0.3 to 3.0 MPa.

3.7 Discussion

Our study indicates that self-similarity of the earthquake source is consistent with data from over 400 small earthquakes in our study region, as shown by the scaling of

source parameters such as corner frequencies and apparent stress as well as the similarity in the shapes of the source spectra themselves, independent of any particular source model. This conclusion is based on stacks of earthquake spectra in bins of similar seismic moment, a process that averages the properties of earthquakes in these bins. Spectra of individual earthquakes may also be obtained using our technique; these show much greater variability in corner frequency and apparent stress but their average properties are consistent with the results presented here. Although we do not take into account possible biasing effects, such as systematic changes in focal mechanism or rupture directivity, it is likely that these effects are minimized by averaging over stations at different distances and azimuths from the source region.

Our study supports models in which the average apparent stress is constant as a function of earthquake size, as suggested by Ide *et al.* (2003) and others. Our results are limited by the small magnitude range spanned by our earthquakes ($1.8 \leq M \leq 3.4$ for the EGF corrected data), but have sufficiently low scatter that fairly tight constraints can be placed on any possible moment dependence of apparent stress. Mayeda and Walter (1996) proposed that E_s/M_0 is proportional to $M_0^{1/4}$ over the magnitude range $3.3 \leq M \leq 7.3$, consistent with the suggestion of Abercrombie (1995) that apparent stress appears to increase gradually with moment over a magnitude range from 0 to 7. Such a strong dependence on M_0 is not supported by our results over the limited size range of our data (our best fitting scaling is $M_0^{0.08\pm0.10}$). Comparisons with other studies can extend the applicability of our results. Our estimated average apparent stress of 1 MPa is above most of the estimates of Abercrombie (1995) for similar size earthquakes (i.e., $M_W = 1.8$ to 3.4) and is consistent with the suggestion of Ide and Beroza (2001) that apparent stress has a nearly constant value of 1 MPa over the entire observed range of earthquake sizes.

A large number of studies have suggested that the source spectra might have more complex behavior than that estimated from simple corner frequency models (e.g., Singh and Ordaz, 1994; Mayeda and Walter, 1996) and should include intermediate falloffs. Differences in the results obtained in different studies might be due to model assumptions that depend upon parametric data derived from the spectra rather than the spectra themselves. An advantage of our approach is that we can directly use the

shapes of the spectra to test for self-similarity without any source model assumptions. A source of concern for our parametric analysis is whether the maximum frequency of 80 Hz that we use in our study is affecting our results, especially for estimates of the corner frequency for the smallest earthquakes. This does not appear to be a problem because we observe no saturation of the corner frequencies for the small events (see Figure 3.12).

The values of apparent stress that we obtain have much less scatter than those seen in most previous studies, probably because of the averaging that we perform within each moment bin. Thus, although our study spans a quite limited magnitude range, our nearly constant values of apparent stress place fairly tight constraints on the amount of any scaling with moment that could be present within our data. Recently Mayeda *et al.* (2004) have argued that a potential problem exists in comparing apparent stress for events over a broad region because the regional scatter of the estimates could make resolving scaling variations problematic. Also, some of the trends of previous studies might be masking (or exposing) the true trend, because of the large range of apparent stress uncertainties. Our study has the advantage of being restricted to a specific source region and of averaging over a large number of earthquakes, reducing the scatter and likely biases in our apparent stress estimates.

Our results are limited to the cluster of earthquakes in our study region but the spectral stacking method should readily be applicable to other data sets. In particular, it would be useful to study clusters or aftershock sequences that contain larger earthquakes to extend the magnitude range. There are a number of possible candidates in southern California for such an analysis, including the Northridge and Landers aftershock sequences. In addition, studies of large numbers of distributed earthquakes, as recorded by local and regional seismic networks, might reveal spatial patterns in source properties. In this case, corrections for attenuation effects will be more complicated than when the earthquakes are restricted to a single cluster, but in principle attenuation and source effects can be still be separated using a spectral stacking approach.

Acknowledgments

We thank the personnel of IGPP and the Anza Seismic group at UCSD who recorded, picked and archived the seismograms. Rachel Abercrombie, Greg Beroza, Jack

Boatwright and Arthur McGarr provided useful comments and reviews. Funding for this research was provided by NEHRP/USGS grants 03HQPA0001 and 01HQAG0021. This research was also supported by the Southern California Earthquake Center. SCEC is funded by NSF Cooperative Agreement EAR-0106924 and USGS Cooperative Agreement 02HQAG0008. The SCEC contribution number for this paper is 766.

4

Uncertainties in earthquake source spectrum estimation using empirical Green functions

We analyze the problem of reliably estimating uncertainties of the earthquake source spectrum and related source parameters using Empirical Green Functions (EGF). We take advantage of the large dataset available from 10 seismic stations at hypocentral distances (10 km $< d <$ 50 km) to average spectral ratios of the 2001 M5.1 Anza earthquake and 160 nearby aftershocks. We estimate the uncertainty of the average source spectrum of the M5.1 target earthquake by performing propagation of errors, which, due to the large number of EGFs used, is significantly smaller than that obtained using a single EGF. Our approach provides estimates of both the earthquake source spectrum and its uncertainties, plus confidence intervals on related source parameters such as radiated seismic energy or apparent stress, allowing the assessment of statistical significance. This is of paramount importance when comparing different sized earthquakes and analyzing source scaling of the earthquake rupture process. Our best estimate of radiated energy for the target earthquake is 1.24×10^{11} Joules with 95% confidence intervals $(0.73 \times 10^{11}, 2.28 \times 10^{11})$. The estimated apparent stress of 0.33 (0.19, 0.59) MPa is relatively low compared to previous estimates from smaller earthquakes (1MPa) in the same region.

4.1 Introduction

A fundamental problem in seismology is accurate estimation of the radiated seismic energy of an earthquake. When an earthquake occurs, some fraction of the total energy is radiated as seismic waves, providing important information about the earthquake rupture process. Determining the amount of radiated seismic energy can be difficult and large differences in these estimates have been found among different techniques and groups (Pérez-Campos *et al.*, 2003).

Partly the difficulty lies in the need to estimate radiated seismic energy over a large dynamic range and a wide frequency band, from the low frequencies needed to define the seismic moment to well beyond the corner frequency. Another difficulty comes from the need for path and site corrections; these corrections have much larger uncertainties at higher frequencies. Finally, the earthquake source might have more complicated features than first expected, including directivity effects that need to be taken into account in order to avoid biasing the estimation.

Over the last 15 years there have been many studies of the radiated seismic energy of different sized earthquakes in different regions of the Earth (e.g., Kanamori *et al.*, 1993; Abercrombie, 1995; Choy and Boatwright, 1995; Mayeda and Walter, 1996; Ide and Beroza, 2001; Venkataraman *et al.*, 2002; Mori *et al.*, 2003; Prieto *et al.*, 2004, and many others). What is often missing in these studies is a measure of the uncertainty of each individual estimate. The question of the uncertainty of estimates is of key importance for describing the significance of one measurement compared to another. In particular, how do the errors in the assumed attenuation model *propagate* into the uncertainty of the source spectrum?

The major unknown in the system is the transfer function between source and receiver. We will focus in this paper on developing a technique to use Empirical Green Functions and estimating and reducing the uncertainties by averaging over a set of suitable aftershocks. By using many EGFs we effectively are randomly sampling the errors in the path effects and averaging over the propagation space. As an example we will present results for the October 31 2001 M5.1 Anza earthquake, using data from local stations.

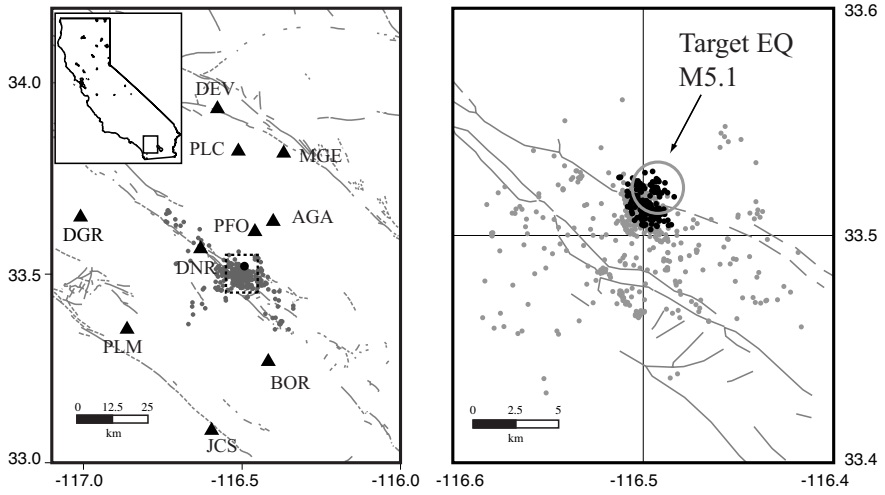

Figure 4.1: Maps of the study area. The left map covers the entire study region, showing the seismic stations (black triangles), the $M5.1$ target Anza earthquake (black circle) and aftershocks M2.9 and lower (gray circles). The inset shows the location of the study area and the state of California. A close-up region (dashed box) is shown in the right hand map. The target earthquake is shown as an open gray circle of radius proportional to a 1MPa stress drop event. Aftershocks used in this study (black circles) and general aftershock seismicity of the region (gray circles) are shown.

4.2 Data Processing

A M_L 5.1 earthquake occurred on 31 October 2001 in the Anza region in southern California (hypocenter 33.5081°N, 116.5143°W, depth 15.2km, M_W4.7) The earthquake exhibited thrust motion on a vertical fault striking N35°E (Hauksson *et al.*, 2002). The data comes from Anza broadband velocity sensors and IDA strong motion sensors at station PFO and both broadband and strong motion sensors from TriNet.

All stations are within 10 to 50 km from the source region. The choice of using local stations to obtain the source spectrum is motivated by the need to obtain reliable estimates of the spectrum at high frequencies and also the need to have good signal-to-noise ratios for the smaller aftershocks. The epicentral distance between the mainshock and all the aftershocks used is less than 2 km (see map in Figure 4.1)

Since S-P times for the closer stations are small, we take a 12 second window that includes both P and S waves, starting 1 second before the P wave pick (e.g., Venkataraman *et al.*, 2002). We also select a noise window of the same length, immedi-

ately preceding the signal window. A signal-to-noise ratio (SNR) is obtained by the ratio of the signal and noise spectra. The spectra of the waveforms are estimated using the multitaper technique (Thomson, 1982), which not only allows a variance reduction of the spectra within a certain bandwidth, but also provides an estimate of the uncertainties at each frequency bin using a jackknife approach (Vernon, 1989; Thomson and Chave, 1991). For each waveform at each station we obtain the spectrum and the $5-95\%$ confidence interval.

Since we are interested in the source spectrum of the $M5.1$ earthquake, we perform spectral division between the mainshock and all the aftershocks recorded at each station. Following the theory of propagation of errors (Taylor, 1997), dividing two random variables results in a new random variable $R_i(f)$ (the spectral ratio), where the i term represents the ith aftershock used for deconvolution, with uncertainties being a function of the uncertainties of the spectra of the mainshock and aftershock. As shown in Figure 4.2, the relative uncertainties in $R(f)$ are larger than the individual components in $X(f)$ and $Y(f)$, due to the instability of the deconvolution process, or simply because we are dividing two noisy spectra. The gray shaded area is larger in Figure 4.2 for the spectral ratio.

As noted by Aki and Richards (1980) (Volume II, Chapter 11.5.5) simple spectral density estimates have only two degrees of freedom and are distributed as χ^2_2; consequently, their ratio is distributed statistically as $F_{2,2}$, a distribution so broad it has infinite variance. We improve on this situation by making use of multitapers: each spectral estimate is made by weighted average over 7 tapers (and time-bandwidth product 4). Then the number of degrees of freedom increases to about 10 (less than the expected 14 because of the weighting).

We take the log of the spectral ratios and stack our data in 15 bins divided by the magnitude of the EGF used ($\Delta M_L = 0.2$). Given that variations in spectra are generally observed to be log-normally distributed, throughout the paper we will average and stack spectra in the log domain, which is equivalent to using a geometric average rather than the arithmetic mean. This also ensures that low amplitude spectra contribute equally to average spectral shapes compared to high amplitude spectra. Figure 4.3 shows a set of typical spectral ratios $R_i(f)$ for EGFs with a range of earthquake magnitudes. Note the

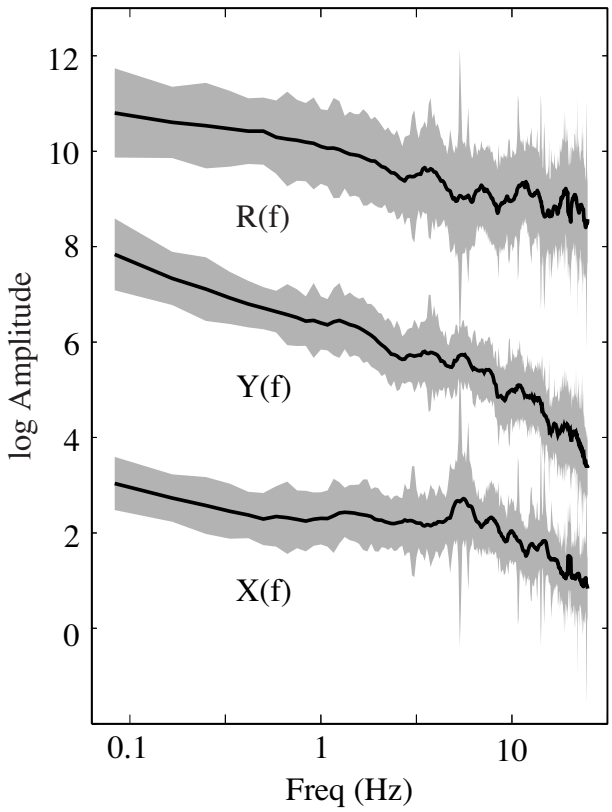

Figure 4.2: Estimates of the spectrum for ground motion associated with the M5.1 earthquake $Y(f)$ at station AGA and the spectrum of the largest aftershock $M2.9$ $X(f)$ at the same station and corresponding spectral ratio $R(f) = Y(f)/X(f)$ (offset for comparison purposes). The gray area represents the 95% confidence interval estimated using the jackknife approach and propagation of errors. Note how uncertainties grow on $R(f)$ after spectral division.

effect of the corner frequency of the larger aftershocks and the effect of the low SNR at low frequencies for the smaller aftershocks. The uncertainties (not shown in Figure 4.3) for each particular bin are again estimated using the method of propagation of errors (Taylor, 1997). For each station a scaling factor for the spectral ratios is determined over a narrow band where the shapes are consistent (gray area in Figure 4.3).

4.3 The Combined Empirical Green Function

In order to take advantage of the large wealth of data available from the local stations, including very small aftershocks (M 0–0.5), we will average together the dif-

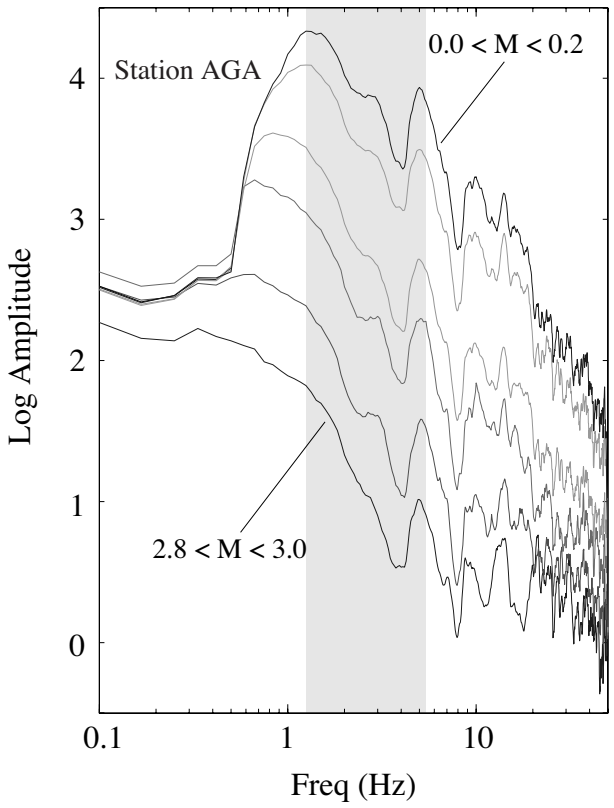

Figure 4.3: Selected bins of spectral ratios ($\Delta M_L = 0.2$) at station AGA. The effect of the corner frequency of the $2.8 < M < 3.0$ earthquakes (lower line) flattens the spectral ratio at high frequencies, while the low SNR affects lower frequencies of the smaller EGFs, but gives reasonable spectral fall-off at high frequencies. The gray shaded area shows the band used to estimate a scaling factor for combining the spectral ratios.

ferent estimates of spectral ratios. This will enable us to bring down the variance or uncertainties of the source spectrum.

The spectral ratio $R(f)$ contains two sources of error, (a) the variance due to the intrinsic spectral estimation process, which we take from the multitaper algorithm and (b) systematic errors (which we will call bias) due to the effect of the corner frequency of the EGF. As described by Ide *et al.* (2003), the spectral ratio $R(f)$ of two earthquakes located close to each other, assuming the same focal mechanisms and path effects, can be expressed as

$$R(f) = \frac{Y(f)}{X(f)} = \frac{S(f)}{G(f)} \tag{4.1}$$

where $Y(f)$, $X(f)$ represent the spectrum of the ground motions of the mainshock and aftershock and $S(f)$, $G(f)$ represent the source spectrum of the mainshock and the

aftershock respectively.

An approximate form of the source spectrum (Brune, 1970) is

$$G(f) = \frac{M_0}{1 + (f/f_c)^2} \tag{4.2}$$

where M_0 is the seismic moment and f_c is the corner frequency of the EGF. An ideal case for an EGF would be to use an earthquake whose corner frequency f_c was very large, giving then in the log domain

$$\log R(f) = \log S(f) - \log M_0 \tag{4.3}$$

resulting in a scaled version of the source spectrum, without changing its shape.

The bias of the log spectrum that is created by a finite corner frequency is then

$$b(f) = \log(1 + (f/f_c)^2) \tag{4.4}$$

which clearly shows that as the frequency f grows and approaches the corner frequency of the EGF, the bias of the source spectrum increases.

We do not know the true corner frequency of the EGF and the relatively low sampling rate of the stations in the network (100 sps) is not enough to estimate it from spectral ratios (as in Hough, 1997; Ide *et al.*, 2003). Instead, we use a simple scaling relation to get an approximate corner frequency. We assume that f_c can be approximated (Venkataraman *et al.*, 2002) from $f_c = 0.49\beta(\Delta\tau/M_0)^{1/3}$ where $\Delta\tau$ is stress drop, and we use a value of 1 MPa. This choice of stress drop is rather arbitrary but certainly within the average in southern California and the study region (Vernon, 1989; Shearer *et al.*, 2006). We chose a rather low stress drop, as a conservative value to obtain small f_c's for the EGFs and limit the potential bias at high frequencies. We estimate the seismic moment following the procedure from Prieto *et al.* (2004), that is, we assume local magnitude $M_L = M_w$ for the small earthquakes and use the Kanamori (1977) relation to obtain M_0. One could argue that this approximation is not accurate, but as it turns out, even allowing the corner frequency of the EGF to change by $20 - 30\%$ does not substantially affect the results, changing the radiated energy estimate of the mainshock by less than 3% in our example.

A common technique for dealing with estimates that contain variance and bias as sources of error is the *mean-square error* (MSE) (Rice, 1995). At a given station we

construct the source spectrum of the target earthquake $S(f)$ from a linear combination of the spectral ratios

$$\log S(f) = \sum_{i=1}^{N} w_i(f)\ \log R_i(f) \tag{4.5}$$

where the index i in the sum runs over the events, w_i are the weights for each spectral ratio, and N is the number of EGF available at a particular station. As explained earlier, the idea is to create a weighted average of the spectral ratios.

The MSE is the sum of the variance and the bias squared of the estimate. Applying this idea to our linear combination of spectra

$$mse^2 = \sum_{i=1}^{N} w_i^2 \sigma_i^2 + \left(\sum_{i=1}^{N} w_i b_i \right)^2 - \lambda \sum_{i=1}^{N} w_i \tag{4.6}$$

where the first term represents the variance (σ_i^2), the second term is the bias squared and we added a normalization constraint as a Lagrange multiplier λ.

Taking the derivative of (4.6) with respect to the unknowns w_i and λ and minimizing, we obtain two sets of linear equations

$$\sum_{j=1}^{N} \left[\sigma_i^2 \delta_{ij} + b_i b_j \right] w_j \ - \lambda = 0 \tag{4.7}$$

$$\sum_{j=1}^{N} w_j = 1 \tag{4.8}$$

We solve the linear set of equations for each individual frequency f with a non-negative least squares approach (see Lawson and Hanson, 1974), to obtain only positive weights. We find that requiring a SNR of 5 or larger for a particular $R_i(f)$ leads to better results. This approach will discard most of the spectra below 0.8 Hz due to low SNR for the smaller magnitude EGFs, and will only use the set of larger EGFs. At higher frequencies more spectral ratios have good SNR, but the larger earthquakes have either larger bias terms or the variance is much larger than for the smaller EGF (since the network records many small aftershocks, thus the variance of a particular spectral ratio bin is decreased), so that the weights prefer the smaller EGF. This means we keep the part of each spectrum that has good SNR and discard only the frequency points with low SNR.

Figure 4.3 shows $R_i(f)$ for a range of EGF earthquake magnitudes at station AGA. We scale the spectral ratios before combining the estimates in (4.5), since the absolute amplitudes depend on the seismic moment of the EGF used. The scaling factor is obtained from a narrow band (shaded area in Figure 4.3) where the spectral ratios are consistent. We tested the effect of the scaling factor by changing the band width used for shifting the spectra. Because the resultant energy estimate does not vary significantly, we choose to ignore this effect.

Solving for the weights in (4.6) and applying the result in (4.5) we obtain a relative source spectrum for each station in the network. Following Prieto *et al.* (2004) we calibrate the relative source spectrum at each station to the seismic moment of the target earthquake. We express the mean source spectrum by averaging over the different seismic stations.

$$\log \overline{S}(f) = \frac{1}{K} \sum_{k=1}^{K} \log S_k(f) \tag{4.9}$$

where K are the number of stations and S_k is the outcome of (4.5) for station k. The processing and averaging is done one frequency at a time, at individual stations, and the mean source spectrum $\overline{S}$ represents a station average.

Figure 4.4 shows the source spectrum $\overline{S}(f)$ for our target event, with $5-95\%$ confidence intervals at each frequency point. Note how for the lower frequencies the uncertainties are larger, since the weights are non-zero only for the larger aftershocks, while at higher frequencies many events are available, an average over many independent estimates, increasing the number of degrees of freedom and decreasing the variance.

4.4 Results for radiated seismic energy

From the mean source spectrum $\overline{S}(f)$ and the uncertainties, it is possible now to obtain an estimate of the radiated seismic energy (Vossiliou and Kanamori, 1982) and its uncertainties for our target earthquake:

$$E_R = \frac{4\pi}{5\rho} \left[\frac{1}{3\alpha^5} + \frac{1}{2\beta^5} \right] \int_0^\infty f^2 |\overline{S}(f)|^2 df \tag{4.10}$$

where ρ is the density, and α and β are the P and S wave velocity at the source. We set $\rho = 2700$ kg/m^3, and the velocities are taken from Scott *et al.* (1994), with $\alpha = 6000$

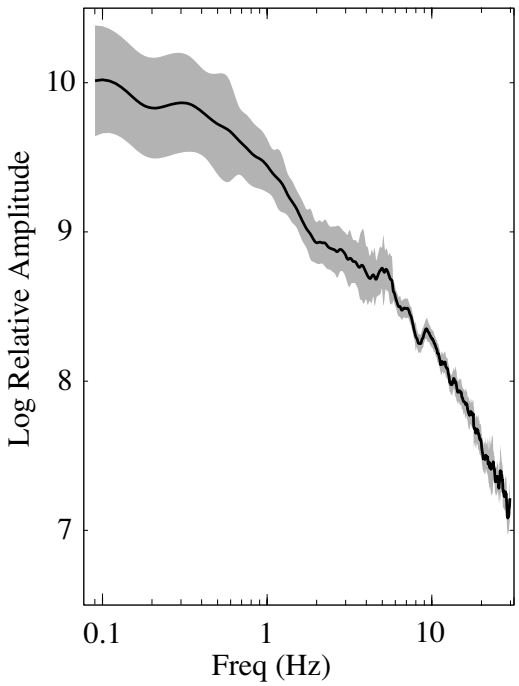

Figure 4.4: Mean source spectrum over 8 stations with 95% confidence intervals (gray area) for the M5.1 target earthquake. Compared to Figure 2, the uncertainties have been decreased due to averaging over different stations and by the combined EGF, especially at high frequencies where many small aftershocks can be used to perform spectral division.

m/s and $\beta = 3465$ m/s. The apparent stress is

$$\tau_a = \mu E_R / M_0 \tag{4.11}$$

where we use $\mu = 3.24 \times 10^4$ MPa.

Since the S wave contains more than 90% of the energy we only use the second term of (4.10). The radiated energy is calculated by integrating (numerically) the mean source spectrum up to the highest frequency possible (30 Hz). We also extrapolate at the higher frequencies assuming a fall-off rate of ω^{-1} in velocity, which contains less than 5% of the total energy. The 95% bounds are integrated to obtain energy uncertainties. The value of radiated energy obtained is 1.24×10^{11} $(0.73 \times 10^{11}, 2.28 \times 10^{11})$ Joules, with the 95% confidence interval in parenthesis. We also compute a value of the apparent stress (proportional to E_R/M_0) and obtain 0.33 (0.19, 0.59) MPa, with the 95% confidence interval.

The apparent stress obtained is within the values of previous results for similar

sized earthquakes, and a little lower than smaller events in the same region (Prieto *et al.*, 2004). But note in this case we have not only obtained the measure of radiated energy, but also the uncertainties, which because we have used over 100 EGFs, have been reduced significantly compared to what one would obtain using just one EGF (i.e., compare Figures 4.2 and 4.4). Our analysis focuses on random variations in our input data (spectral ratios) rather than uncertainties in model parameters. Therefore we do not attempt to propagate errors associated with parameters such as wave speed or density, even though these errors might be important. From equation (4.10) it would be straight forward to propagate such errors, if known, following Taylor (1997).

4.5 Discussion

As explained before, using spectral ratios will increase the variance and a larger number of source models (corner frequencies, fall-off rates) are to be allowed within the uncertainties. Since stress drop varies as f_c^3, the uncertainties of the stress drop as estimated from f_c will grow considerably, making the inference of scaling features between different earthquakes more difficult. Our goal is to show a consistent way of estimating the uncertainties of the earthquake source spectrum and some source parameters using the method of propagation of errors and the variance of the spectrum estimation procedure. A source spectrum is then constructed from a weighted set of spectral ratios in order to reduce the variance.

As pointed out by Sonley and Abercrombie (2006) it is good practice to check whether a realistic source pulse is obtained after deconvolution, by inverse FFT of the multitaper eigencomponents. We compute source pulses for our target M5.1 earthquake and obtain results similar to those of McGuire [pers. com., 2005] for the same event. However it is not possible to check all our EGFs since SNR limitations at low frequencies affect the very small EGFs and it is not always possible to recover source pulses.

In our data set we find that there is no single EGF with good SNR and appropriate bias reduction on the entire frequency band of interest (about 0.1 - 30 Hz) and it is necessary to use multiple EGFs. Even if an ideal EGF is found, the uncertainties of the radiated energy estimate would be considerably larger than presented here and

should be taken into account when comparing different results or looking for scaling of energy with earthquake magnitude.

A possible major source of bias is if many EGFs used in this study (especially the small ones) have consistently different focal mechanisms that are not accounted for. We believe that by using many EGFs we are sampling a wide variety of focal mechanisms. For the target event we obtain realistic source pulses, suggesting similar mechanisms. The aim of this study is to show a way of estimating and reducing uncertainties of source spectra based on the target and EGF spectra. It is not intended to completely remove all possible biases associated with the choice of events and model parameterization.

Acknowledgments

We thank Glen Offield for field operations and the Anza group at UCSD who picked and archived the seismograms. We thank M. Hellweg, V. Oye and R. Abercrombie (editor) for thoughtful comments. Funding for this research was provided by NSF Grant number EAR0417983. This research was also supported by the Southern California Earthquake Center. SCEC is funded by NSF Cooperative Agreement EAR-0106924 and USGS Cooperative Agreement 02HQAG0008. The SCEC contribution number for this paper is 933.

5

Quadratic Multitaper Spectrum

The power spectral density of geophysical signals provides relevant information about the processes that generated these particular signals. We present a new method to optimally use the multitaper spectral analysis method. The method is an extension of the algorithm by Thomson (1982) with a reduction of bias due to the curvature close to the frequency of interest. A comparison of the original and the new Quadratic multitaper with the same resolution bandwidth demonstrates the reduction of bias in areas where the signal has significant quadratic structure without the introduction of additional sidelobe leakage. In addition, the methodology provides independent estimation of the derivatives of the spectrum (e.g., the slope of the spectrum). The extra information can be implemented for parameter estimation or in comparing different signals.

5.1 Introduction

There are many applications in geophysics where relevant information contained in a given signal may be extracted from the frequency content of the spectrum. In some cases, the scientist may be interested in periodic components usually immersed in some background noise (e.g., normal mode seismology (Gilbert, 1970), climate-time series (Chappellaz *et al.*, 1990), etc.), in a general continuous spectrum to be estimated from a short time series (e.g., earthquake source spectra (Brune, 1970), bathymetry data (Goff and Jordan, 1988), etc.) or in comparing two signals and investigating where the similarities or differences are (in seismology for example (Vernon, 1989; Hough and Field, 1996); transfer functions in electromagnetism (Constable and Constable, 2004), elastic thickness of the lithosphere (Daly *et al.*, 2004), etc.). In each of these cases, it is desirable to be able to obtain a reasonable spectrum, with little or no bias and small uncertainties.

In Thomson (1982) the multitaper spectral analysis method was introduced. The original algorithm has been widely used in geophysical applications and has been shown in multiple cases to outperform the single-tapered, smoothed periodogram (Park *et al.*, 1987b; Bronez, 1992; Riedel *et al.*, 1993). In the latter, a multitaper estimate that is subsequently smoothed is preferred.

Single taper estimates have a major limitation, in the sense that by using one taper a significant portion of the signal is discarded. The data points at the extremes are down-weighted, causing the variance of the direct spectral estimate to be greater than that of the periodogram. In the multitaper algorithm, the statistical information discarded by one taper is partially recovered by the others. The tapers are constructed to optimize resistance to spectral leakage and only a few of them are computed. The multitaper spectrum is constructed by a weighted sum of these single tapered periodograms. The weighting function is defined to generate a smooth estimate with less variance than single taper methods and at the same time to have reduced bias from spectral leakage.

The weighting proposed (Thomson, 1982) is ideal for spectral leakage but suffers from local bias. What we mean by local bias is that by averaging over a number of tapered spectra, the estimate will have a broad response around the center frequency, whose width is that of the frequency resolution chosen to reduce spectral leakage.

Table 5.1: Essential notation and mathematical symbols used in this chapter. In the text, $\hat{\theta}$ means estimate of the variable θ; $E[\cdot]$ is the expected value and * represents the complex conjugate.

Symbol	Description
$x(t)$	The time series to be analyzed (assumed with time unit intervals)
N	The number of data points of the time series
t	The time variable ($t = 0, 1, 2, \ldots, N-1$)
f,f'	Continuous frequency variables
$v_k(t)$	Slepian sequences. Are a function of both N and a bandwidth W
W	Bandwidth of the windows v_k, which define the inner domain $(f-W, f+W)$
λ_k	Eigenvalues of the Slepian sequences. Also give the fraction of energy in the inner band $(-W, W)$.
K	Number of tapers or windows to be used. Number of Slepian sequences used
$dZ(f)$	Orthogonal incremental process from Cramér Spectral representation. Also known as generalized Fourier transform.
V_k	The Slepian *functions*, Fourier transform of v_k's. Functions are orthogonal in the principal domain $(-\frac{1}{2}, \frac{1}{2})$
$\mathcal{V}_k$	Orthonormal version of Slepian functions in inner domain $(-W, W)$. Defined as $\mathcal{V}_k = V_k/\lambda_k$
$Y_k(f)$	kth eigencomponent, the discrete Fourier transform of $x(t)v_k(t)$ in the domain $(-\frac{1}{2}, \frac{1}{2})$.
$\mathcal{Y}_k(f)$	Idealized kth eigencomponent in the domain $(-W, W)$. Theoretically contain information from the inner domain.
$S_k(f)$	The kth eigenspectra. A direct spectral estimate using v_k as a taper.
$C_{jk}(f)$	Covariance matrix of the idealized eigencomponents. $C_{jk} = \mathcal{Y}_j\, \mathcal{Y}_k^*$.
$S(f)$	The power spectral density (PSD) of the time series.
$T_n(f)$	The nth Chebyshev polynomial function.
α_n	The nth Chebyshev coefficient for the expansion of the spectrum $S(f)$
$d_k(f)$	Multitaper weights used to obtain estimates of the bandlimited coefficients $\mathcal{Y}_k(f)$

Riedel and Sidorenko (1995) suggested a different weighting function to minimize the local bias. They take a different set of orthogonal tapers and, assuming quadratic structure of the spectrum, reduce the bias introduced by the curvature of the spectrum. Because the chosen tapers minimize local bias, there is little spectral leakage reduction and this method does not perform as well with very high dynamic range signals.

In this paper we present an improved multitaper spectral estimate using Thomson's algorithm modified to reduce curvature bias. The second derivative of the spectrum is estimated by means of an expansion of the spectrum via the Chebyshev Polynomials. Because we use the Slepian tapers and the original weighting functions, the estimate is resistant to spectral leakage, yet reduces bias due to spectral curvature. In addition, we can estimate the derivative of the spectrum (slope), which can be used in parameter estimation or as a discriminant for comparing different signals.

This paper is organized as follows. In the next section, we outline the problem associated with the estimation of the power spectral density (PSD). After this, we provide a brief review of the multitaper spectrum estimation algorithm and discuss some of its properties. Then, we introduce the Quadratic multitaper, first explaining in section 5.4 the methodology to estimate the derivatives of the spectrum and in section 5.5 the estimation of the unbiased spectral estimate. Finally, there is a fairly extensive section in which we give a number of examples to compare Thomson's multitaper and the new algorithm. There is an application to bathymetric data and the application of the derivatives of the spectrum for comparing signals.

5.2 Spectrum Estimation

In time-series analysis it is often useful to describe the power spectral density (PSD) of the signal, given that it may have information of the background noise, periodic components, and transients. These pieces of information are fundamental in geophysics (Thomson, 1990).

To start, we need to consider a stationary stochastic process $x(t)$, a zero mean discrete time series consisting of N contiguous samples and assume that the sampling

rate is always unity, so that $t = 0, 1, 2, \ldots, N-1$. Define the Fourier transform of the observations

$$Y(f) = \sum_{t=0}^{N-1} x(t)\, e^{-2\pi i f t}, \qquad -\frac{1}{2} \leqslant f \leqslant \frac{1}{2}. \tag{5.1}$$

We assume the signal is a harmonizable process, thus has a Cramér spectral representation (Cramér, 1940)

$$x(t) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{2\pi i f t}\, dZ(f) \tag{5.2}$$

for all t, where $dZ(\cdot)$ is an orthogonal incremental process (Doob, 1952).

The random orthogonal measure $dZ(f)$ has its first moment

$$E\left[dZ(f)\right] = 0 \tag{5.3}$$

and second moment (of relevance for our purposes)

$$E\left[|dZ(f)|^2\right] = S(f)df \tag{5.4}$$

where $S(f)$ is defined as the *power spectral density function* of the process and $E[\cdot]$ is the expected value. Note here that the frequency variable f is continuous and so we are in fact trying to find a function $S(f)$ from the finite series $x(t)$.

Plugging the Cramér spectral representation (5.2) into the discrete Fourier transform (5.1), we arrive at the basic integral equation (Thomson, 1982, 1990)

$$Y(f) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{\sin N\pi(f-v)}{\sin \pi(f-v)} dZ(v) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathbb{D}(f,v) dZ(v) \tag{5.5}$$

where $\mathbb{D}(f,v)$ is the Dirichlet kernel. The basic integral equation is a convolution that can be interpreted as the *smearing* of the true $dZ(f)$ projected into $Y(f)$, due to the finite duration of the time series $x(t)$.

This paper is about obtaining an approximate solution to this equation and based on that solution (via Equation 5.4) obtaining reliable estimates of the power spectral density of the signal.

5.3 Multitaper Spectrum estimates

We give a brief review of the standard theory for multitapers. Proof of various assertions can be found in (Thomson, 1982, 1990; Park *et al.*, 1987b; Percival and Walden, 1993). Note that given the properties of (5.5) and the smoothing of the Dirichlet kernel there is no unique solution to this problem. The multitaper spectrum estimate is an approximate least-squares solution to equation (5.5) using an eigenfunction expansion. The choice of this type of solution will be explained next.

The most obvious first guess for the spectrum is the squared Fourier transform. Slightly rewriting the discrete Fourier transform (5.1) and squaring

$$\hat{S}(f) = |Y(f)|^2 = \left| \sum_{t=0}^{N-1} x(t)a(t)\, e^{-2\pi i f t} \right|^2 \tag{5.6}$$

where the sequence $a(t)$ is called a taper. In the case of (5.1) the taper $a(t) = 1$ is a boxcar function. To maintain total power correctly $a(t)$ needs to be normalized:

$$\sum_{t=0}^{N-1} |a(t)|^2 = 1 \tag{5.7}$$

In the frequency domain, the properties of the taper are deduced from its Fourier transform

$$A(f) = \sum_{t=0}^{N-1} a(t)\, e^{-2\pi i f t} \tag{5.8}$$

We call the function $A(f)$ the spectral window associated with a. For conventional tapers, $|A(f)|$ has a broad main lobe and a succession of smaller sidelobes (see Figure 5.1).

The choice of the taper can have a significant effect on the resultant spectrum estimate. One can observe this by expressing equation (5.6) as a convolution of the taper transform (5.8) and the true spectrum

$$E\left[\hat{S}(f)\right] = \int_{-\frac{1}{2}}^{\frac{1}{2}} |A(f')|^2\, S(f - f')df' \tag{5.9}$$

Here, as in (5.5), there is smearing or smoothing of the true spectrum. This means that the choice of window is important. A good taper will have a spectral window with low amplitudes whenever $|f - f'|$ is large, leading to an estimate $\hat{S}(f)$ based primarily

on information close to the frequency f of interest. The objective of the taper $a(t)$ is to prevent energy at distant frequencies from biasing the estimate at the frequency of interest. This bias is known as spectral leakage. We wish to minimize the leakage at frequency f from frequencies $f' \neq f$.

In practice, it is not sensible to be concerned about $|f' - f| \leqslant 1/N$, since this is the lowest frequency accessible from a record of length N. A bandwidth W is chosen, where $1/N < W \leq 1/2$, and the fraction of energy of A in the interval $(-W, W)$ is given by:

$$\lambda(N, W) = \frac{\int_{-W}^{W} |A(f)|^2 df}{\int_{-\frac{1}{2}}^{\frac{1}{2}} |A(f)|^2 df} \tag{5.10}$$

where λ is a measure of spectral concentration. It is clear that no choice of W can make λ greater than unity. Our task is to maximize λ.

To maximize λ substitute (5.8) in (5.10), take the gradient of λ with respect to the vector $\mathbf{a} = [a(0), a(1), \ldots, a(N-1)]$ and set to zero to obtain a matrix eigenvalue problem:

$$\mathbf{D} \cdot \mathbf{a} - \lambda \mathbf{a} = 0 \tag{5.11}$$

where the matrix $\mathbf{D}$ has components

$$D_{t,t'} = \frac{sin 2\pi W(t - t')}{\pi(t - t')}, \quad t, t' = 0, 1, \ldots, N-1 \tag{5.12}$$

and is symmetric.

The solution of (5.11) has (dropping dependence on N and W) eigenvalues $1 > \lambda_0 > \lambda_1 > \cdots > \lambda_{N-1} > 0$ and associated eigenvectors $v_k(t)$, called the Slepian sequences (Slepian, 1978). The first eigenvalue λ_0 is extraordinarily close to unity, thus making the choice $a(t) = v_0(t)$ the taper with the best possible suppression of spectral leakage for the particular choice of bandwidth W. In fact, the first $2NW - 1$ eigenvalues are also very close to one, leading to a family of very good tapers for bias reduction. The multitaper algorithm exploits the fact that a number of tapers have good spectral leakage reduction, and uses all of them rather than only one.

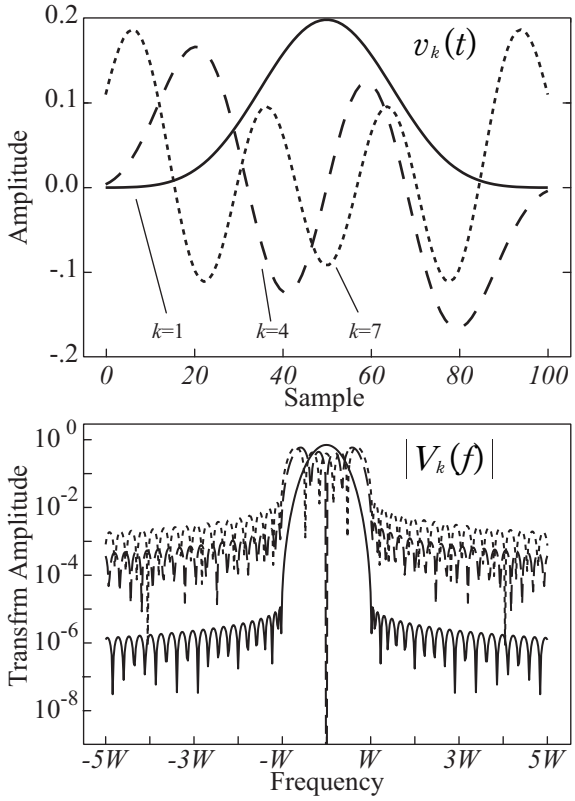

Figure 5.1: Selected Slepian sequences and corresponding Slepian functions for $N = 100$ samples and a choice of $NW = 4$. Sometimes called 4π Slepian sequences.

5.3.1 Properties of Slepian sequences and functions

The Slepian sequences are solutions of the symmetric matrix eigenvalue problem (5.11) - (5.12). The eigenvectors with associated eigenvalues λ_k are real and orthogonal as usual

$$\sum_{t=0}^{N-1} v_j(t)v_k(t) = \delta_{jk} \tag{5.13}$$

These vectors will be used as tapers in (5.6). We define the Slepian functions as the spectral windows, the Fourier transforms of the sequences

$$V_k(f) = \sum_{t=0}^{N-1} v_k(t)e^{-2\pi ift} \tag{5.14}$$

Note that the V_k's are complex functions of frequency. Figure (5.1) shows three Slepian sequences and their corresponding Slepian functions.

Orthogonality conditions also hold in the frequency domain, making the choice

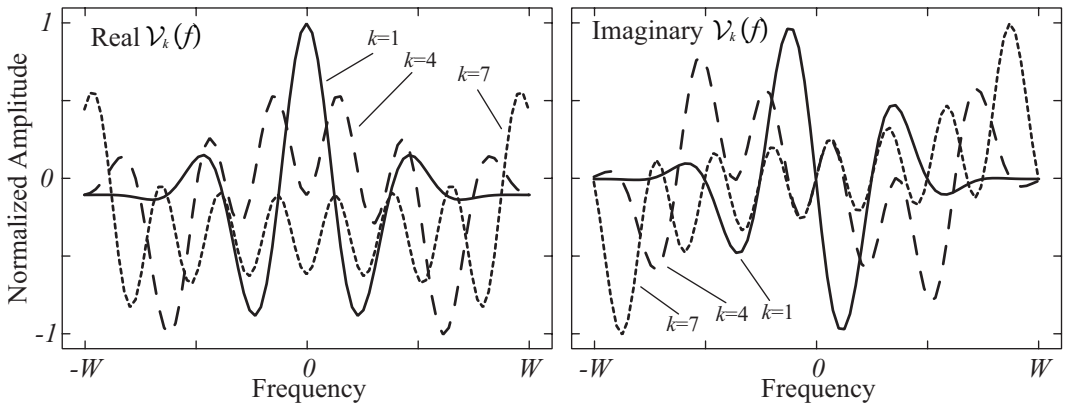

Figure 5.2: Orthonormal version of the Slepian functions in Figure 5.1, $\mathcal{V}_k(f)$ in the inner domain. Only three functions are shown in the example with their real and imaginary amplitudes normalized. Number and thin lines show the index of the corresponding Slepian function plotted. The symmetry and amplitude of the functions are of interest.

of the Slepian sequences so interesting:

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} V_j(f)V_k^*(f)df = \delta_{jk} \tag{5.15}$$

$$\int_{-W}^{W} V_j(f)V_k^*(f)df = \lambda_k\delta_{jk} \tag{5.16}$$

It is convenient to define an orthonormal version of the V_k's on the inner domain $(-W, W)$

$$\mathcal{V}_k(f) = \frac{V_k(f)}{\sqrt{\lambda_k}} \tag{5.17}$$

with the obvious property

$$\int_{-W}^{W} \mathcal{V}_j(f)\mathcal{V}_k^*(f)df = \delta_{jk} \tag{5.18}$$

This last property will be exploited in the sections to come. Figure (5.2) shows a selection of the $\mathcal{V}_k(f)$ functions in the inner interval. Note how the real part of the functions is always even and the imaginary part is odd. See also how the amplitude of the functions increases outward as the Slepian function index increases and are more sensitive to structure further from the center frequency.

5.3.2 The multitaper algorithm

The objective of this algorithm is to estimate the spectrum $S(f)$ by using K of the Slepian sequences to obtain the k *eigencomponents*:

$$Y_k(f) = \sum_{t=0}^{N-1} x(t) v_k(t) e^{-2\pi i f t} \tag{5.19}$$

and a set of K eigenspectra as in (5.6):

$$\hat{S}_k(f) = |Y_k(f)|^2 \tag{5.20}$$

from which we can form the mean spectrum

$$\bar{S}(f) = \frac{1}{K} \sum_{k=1}^{K} \hat{S}_k(f) \tag{5.21}$$

The idea of taking an average is to reduce the variance in the spectral estimate. As will be shown below, the mean spectrum is not an ideal estimate and we prefer a weighted average instead, one that minimizes some measure of discrepancy. While the spectral leakage properties of the $\hat{S}_0$ eigenspectrum are very good, since the eigenvalues are close to unity when $K < 2NW - 1$, the leakage characteristics of the succesive estimates degrade. It is clear that by using $\hat{S} = |Y_0|^2$, the least amount of spectral leakage is achieved. Nevertheless, including the other eigencomponents $(Y_1, Y_2, \ldots, Y_K)$, while increasing spectral leakage, reduces the variance of the spectral estimate and is thus preferred.

In order to estimate the discrepancy of the different eigencomponents Y_k, we combine (5.19) with (5.2)

$$\begin{aligned} Y_k(f) &= \sum_{t=0}^{N-1} x(t) v_k(t) e^{-2\pi i f t} \\ &= \sum_{t=0}^{N-1} \int_{-\frac{1}{2}}^{\frac{1}{2}} dZ(f') e^{2\pi i f' t} v_k(t) e^{-2\pi i f t} \\ &= \int_{-\frac{1}{2}}^{\frac{1}{2}} \sum_{t=0}^{N-1} v_k(t) e^{-2\pi i (f - f') t} dZ(f') \end{aligned}$$

and using the definition of the Fourier transform of the taper (5.14) we obtain:

$$Y_k(f) = \int_{-\frac{1}{2}}^{\frac{1}{2}} V_k(f - f')\, dZ(f') \tag{5.22}$$

containing information from the whole interval $(-\frac{1}{2}, \frac{1}{2})$.

If the sequence $x(t)$ were passed by a perfect bandpass filter from $f - W$ to $f + W$ before truncation to the sample size with N data points, we would obtain the *idealized eigencomponents* $\mathcal{Y}_k(f)$ that, though unobservable, would be represented by:

$$\mathcal{Y}_k(f) = \int_{-W}^{W} \frac{V_k(f')}{\sqrt{\lambda_k}} dZ(f - f') = \int_{-W}^{W} \mathcal{V}_k(f') dZ(f - f') \tag{5.23}$$

Note that here we adopt the orthonormal functions $\mathcal{V}_k$, in order to maintain the correct normalization. The $\mathcal{Y}_k$ takes only information over the inner interval $(-W, W)$.

In order to estimate $\hat{S}(f)$, we find a set of frequency dependent weights $d_k(f)$, as proposed by (Thomson, 1982, 1990):

$$d_k(f) = \frac{\sqrt{\lambda_k} S(f)}{\lambda_k S(f) + (1 - \lambda_k)\sigma^2} \tag{5.24}$$

where σ^2 is the variance of the signal $x(t)$. The multitaper spectrum is then obtained

$$\hat{S}(f) = \frac{\sum_{k=0}^{K-1} d_k^2 \left|Y_k(f)\right|^2}{\sum_{k=0}^{K-1} d_k^2} \tag{5.25}$$

Since we don't know the spectrum $S(f)$ in (5.24), we are required to assume an initial estimate of the spectrum (averaging the first two k eigenspectra $S_0 + S_1$ for example) and find the weights d_k iteratively. A complete derivation of the weights in (5.24) can be found in Thomson (1982) or Percival and Walden (1993).

5.4 Estimating the derivatives of the spectrum

More information about the spectrum can be obtained by looking at the covariance matrix of the K components:

$$C_{jk}(f) = E\left[d_j Y_j \, d_k Y_k^*\right] = E\left[\mathcal{Y}_j \, \mathcal{Y}_k^*\right] \tag{5.26}$$

with $j, k = 1, \ldots, K$. We use the orthogonal increment property (5.4) and substituting (5.23):

$$C_{jk}(f) = \int_{-W}^{W} G_{jk}(f') \, S(f - f') df' \tag{5.27}$$

where $G_{jk}(f) = \mathcal{V}_j(f)\mathcal{V}_k^*(f)$. If the spectrum does not vary in the interval $(-W, W)$ then the covariance matrix is diagonal

$$\mathbf{C}(f) = S(f)\,\mathbf{I} \tag{5.28}$$

where $\mathbf{I}$ is the $K \times K$ identity matrix. Note that the multitaper spectrum in (5.25) is equivalent to taking the trace of $\mathbf{C}(f)$ and normalizing by the weights.

When the spectrum is constant in the interval $(-W, W)$, Thomson's multitaper is unbiased and provides an appropriate estimate of the spectrum. If, however, the spectrum varies within the interval, the matrix $\mathbf{C}(f)$ will not be diagonal and (5.25) may be biased. Clearly, we rarely obtain perfectly diagonal covariance matrices and the spectrum is not perfectly resolved.

Like (5.5) equation (5.27) is a Fredholm integral of the first kind and suffers from similar non-uniqueness and smearing features, except in this case we have reduced spectral leakage and are only concerned about the interval $(-W, W)$.

Thomson (1990) suggested taking a set of orthogonal eigenfunctions to expand the spectrum to solve (5.27). Here we propose to employ the Chebyshev polynomials for estimating the derivatives of the spectrum and using these derivatives to obtain an improved solution of the spectrum $S(f)$. We prefer the Chebyshev polinomials (Mason and Handscomb, 2003) because, as can be seen from Figure (5.3), these polynomials are sensitive to structure at the edges of the interval, where the eigenfunctions proposed by Thomson (1990) have very little energy.

We write the spectrum in the inner interval as:

$$S(f - f') = \alpha_0 T_0(\frac{f'}{W}) + \alpha_1 T_1(\frac{f'}{W}) + \alpha_2 T_2(\frac{f'}{W}), \qquad -W \leqslant f' \leqslant W \tag{5.29}$$

where $T_n(x)$ is the Chebyshev polinomial of degree n (see Mason and Handscomb, 2003). We show only the first three polynomials, since these are applied throughout the study, and express the quadratic terms of the signal.

Returning to the inverse problem (5.27) in spectrum estimation and inserting (5.29):

$$C_{jk}(f) = \alpha_0 H_{jk}^{(0)} + \alpha_1 H_{jk}^{(1)} + \alpha_2 H_{jk}^{(2)} + O(f - f')^3 \tag{5.30}$$

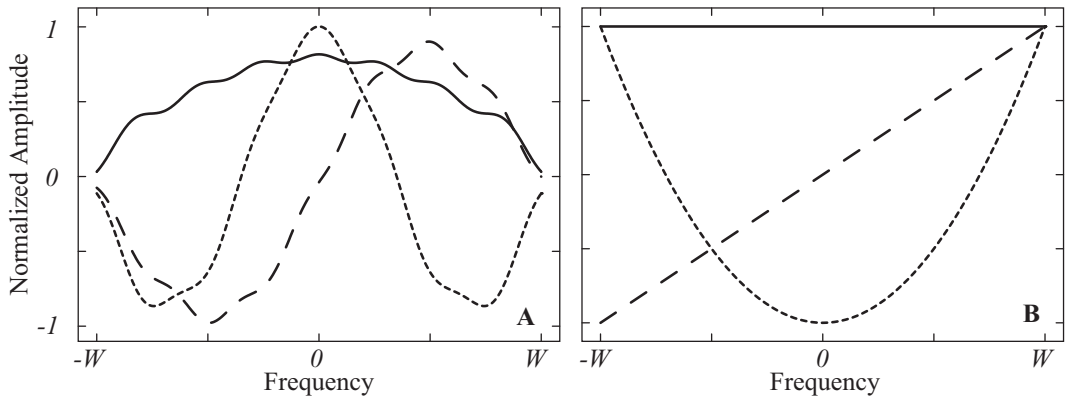

Figure 5.3: Comparison between first 3 basis functions used in Thomson (1990) (A) and Chebyshev polynomials (B) used in this study. Note that in (A), the basis functions always tend to zero when getting close to either $-W$ or W and are not sensitive to structure at the boundaries. The Chebyshev polynomials in (B) are also very simple approximations of a constant, slope, and quadratic terms.

where the matrices

$$
\begin{aligned}
H_{jk}^{(0)} &= \int_{-W}^{W} G_{jk}(f')\, T_0(f')\, df' = \delta_{jk} \\
H_{jk}^{(1)} &= \int_{-W}^{W} G_{jk}(f')\, T_1(f')\, df' \\
H_{jk}^{(2)} &= \int_{-W}^{W} G_{jk}(f')\, T_2(f')\, df'
\end{aligned}
$$

describe the zero, first, and second derivative basis matrices. We can then obtain the Chebyshev coefficients, $\alpha_0, \alpha_1, \alpha_2$, by solving the least squares problem where we use the observed values of $d_j d_k Y_j Y_k^*$ to approximate the left side.

The Chebyshev coefficients are estimates of the derivatives of the spectrum

$$
\begin{aligned}
\alpha_0 &\approx S(f) \\
\alpha_1 &\approx S'(f) \\
\alpha_2 &\approx S''(f)
\end{aligned}
$$

around the center frequency on the interval $(f - W, f + W)$.

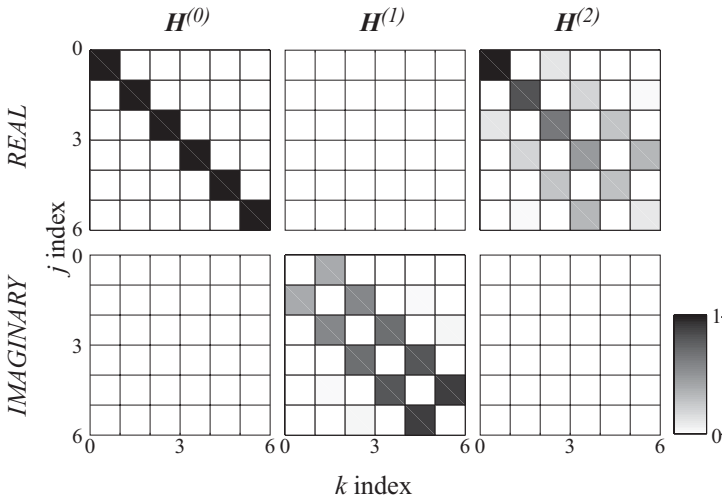

Figure 5.4: Comparison between the different basis matrices for zero (left), first (middle), and second (right) order coefficients. The absolute values are shown for simplicity. Note that the completely white matrices show that the real part of the covariance matrix is insensitive to slopes, while the imaginary part is insensitive to a constant and quadratic structure of the spectrum.

In practice, the calculation of the integrals in $H_{jk}^{(1)}$ and $H_{jk}^{(2)}$ is done numerically using a trapezoidal quadrature. In Figure (5.4) we plot the case of the three matrices in (5.30). The absolute values are plotted. Note that both the zero and second order terms are only present in the real part of the covariance matrix, while the first order term is present only in the imaginary part of the covariance matrix.

It is clear that the constant term will result in a diagonal covariance matrix, while the effects of the first and second order terms are quite different. $H_{jk}^{(1)}$ has no effect on the diagonal terms, suggesting that this term does not bias the spectrum estimate (5.25). In contrast, $H_{jk}^{(2)}$ has an important contribution to the diagonal but is also present in the off-diagonal terms, showing the dependence of the eigencomponents in spectra that are highly variable. A slight correlation between the estimates of α_0 and α_2 is present.

5.5 Quadratic Multitaper

Up to now, the literature (e.g., Thomson (1982); Park (1992); Percival and Walden (1993), and many others) has assumed the spectrum varies slowly in this interval

and can be taken out of the integral in equation (5.27). Now, within $(-W, W)$ we can try to find further information on the structure of the spectrum and relax the assumption of a constant spectrum inside the interval.

Assume the spectrum has a Taylor series expansion on the interval $(f - W, f + W)$ of the form:

$$S(f) = S(f') + (f - f')S'(f') + \tfrac{1}{2}(f - f')^2 S''(f') + O(f - f')^3$$

We know that the estimate of the spectrum $\hat{S}(f)$ at any frequency f is an average over the interval $(f - W, f + W)$:

$$\begin{aligned} E[\hat{S}(f)] &= \frac{1}{2W} \int_{-W}^{W} \left[S(f') + \tfrac{1}{2}(f - f')^2 S''(f') \right] df' \\ &= S(f) + \frac{1}{6} W^2 S''(f) \end{aligned}$$

where the term associated with the first derivative does not contribute to the integral due to symmetry. Note that in Figure 5.4 the matrix $\mathbf{H}^{(1)}$, associated with the slope, is zero in the main diagonal and does not bias $E[\hat{S}(f)]$.

We can obtain the Quadratic multitaper estimate of the spectrum $\tilde{S}(f)$ at frequency f by applying the correction:

$$\tilde{S}(f) = \hat{S}(f) - \frac{1}{6} W^2 \hat{\alpha}_2 \tag{5.31}$$

where we assume $\hat{\alpha}_2 \approx S''(f)$ obtained by solving (5.30). Note that we apply the correction to the multitaper estimate $\hat{S}(f)$ in (5.25).

Applying the quadratic correction in (5.31) will increase the variance of the overall estimate, because $\hat{\alpha}_2$ is also uncertain. We propose to implement a mean-square error criteria instead of directly applying (5.31) to avoid exacerbating the uncertainties of the Quadratic multitaper:

$$\tilde{S}(f) = \hat{S}(f) - \mu \frac{1}{6} W^2 \hat{\alpha}_2 \tag{5.32}$$

where μ is a weight:

$$\mu = \frac{{\alpha_2}^2}{({\alpha_2}^2 + var\{\alpha_2\})} \tag{5.33}$$

In Appendix B the derivation of the weight μ in (5.33) is explained as well as the approximate estimation of the variance of α_2.

The Quadratic multitaper is an approximately unbiased estimate of the PSD of the signal analyzed. As will be demonstrated in the next section with different examples, the Quadratic multitaper provides a reduction of curvature bias while at the same time generating smooth estimates.

5.6 Examples

As a demonstration of the benefits of the Quadratic spectrum algorithm (QMT) we show a number of synthetic examples. The main features we would like to concentrate on are the resolution close to significant structure in the spectrum, smoothness of the resultant estimates and the overall spectral leakage properties.

5.6.1 Random signal

The first signal to be analyzed is a simple pseudo-random number $r(t)$ with a normal distribution and standard deviation σ=1. The number of data points for this example is N=1000, For all plots in this paper we compute 6 tapers (K=6) with 3.5 as the time-bandwidth product.

A visual inspection of the results of the two different methods in Figure 5.5 shows that the algorithm presented here generates a smoother spectrum than the original Thomson method (TMT). A more quantitative comparison is provided in Table 5.2, where 10 random realizations of a 1000 sample long random time series were analyzed and two different measures were used to assess the smoothness of the resultant spectrum: the norm of the second (numerical) derivative of the spectrum and a count of the number of maxima in each spectra.

Results of these measures of smoothness in Table 5.2 demonstrate that the QMT generates smoother estimates of the spectrum. In all 10 realizations, both measures of smoothness where lower when using the Quadratic algorithm.

Table 5.2: Comparison of smoothness for the multitaper methods

	Thomson	Quadratic
Number of realizations	10	10
Second derivative norm	216.8	49.9
Standard deviation	22.2	5.99
Count of maxima	123.3	67.3
Standard deviation	5.07	3.62

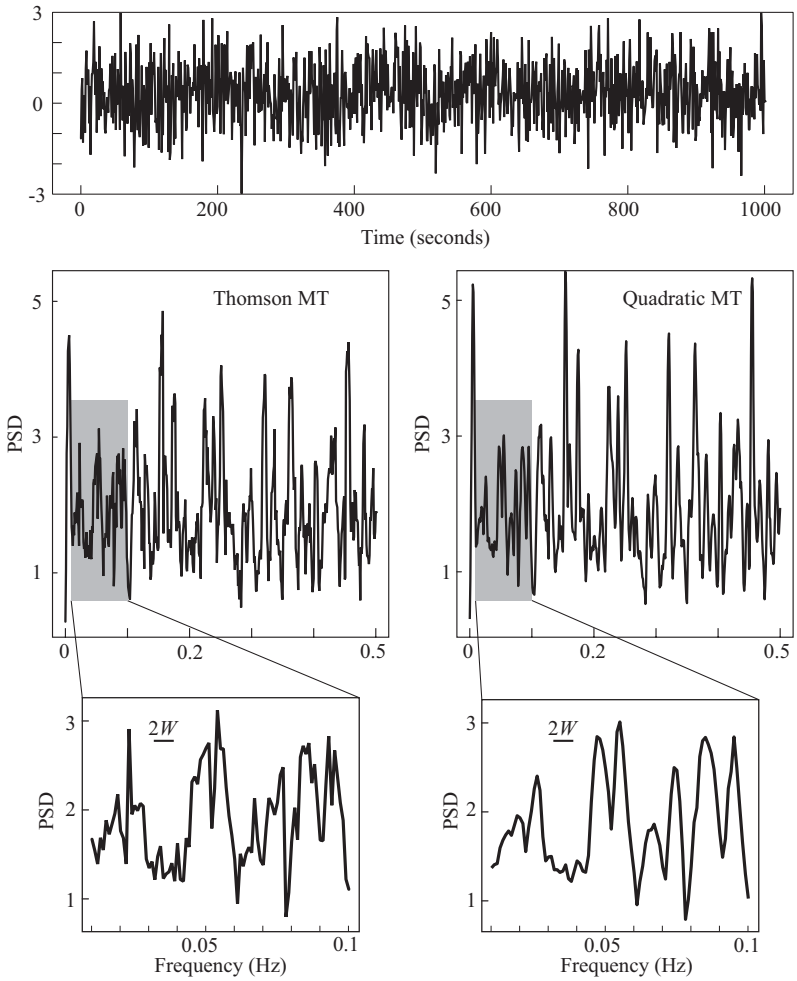

Figure 5.5: Spectrum estimation of pseudo-random number signal. The signal is a normally distributed random vector, with N=1000 samples. The top panel shows the time-series random signal. The middle panels show the TMT (left) and the QMT (right) estimates of the spectrum. Bottom panels show a detailed view of the estimates between 0.01 and 0.1 Hz and the $2W$ bandwidth for reference. The QMT generates a smoother spectral estimate.

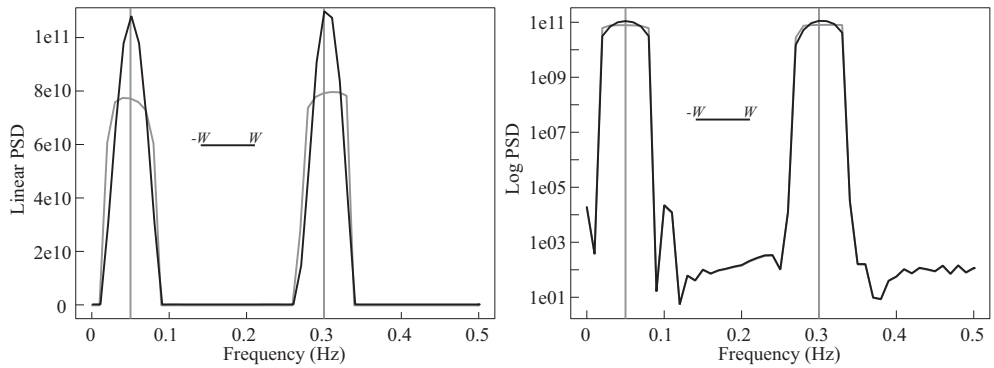

Figure 5.6: Spectrum estimation for a high dynamic range signal with two periodic components at 0.05Hz and 0.3Hz. N=100 points, time-bandwidth product NW=3.5 and K=6 . The linear scale spectrum (top panel) shows the improved performance of the QMT in describing the periodic components, while the logarithmic scale (lower panel) shows the similar spectral leakage properties of both algorithms.

5.6.2 Periodic Components

We test the effectiveness of the new algorithm with two signals; we examine a random signal $r(t)$ with σ=1.0 and a pair of periodic components. The number of data points is reduced to N=100 in order to have a comparison of spectral leakage around the linear components.

Our first test is to see whether this algorithm represents the periodic components in the signal better, without introducing additional spectral leakage. For this, we take the signal:

$$x(t) = A_0 \sin(2\pi f_1 t) + A_0 \sin(2\pi f_2 t) + r(t) \tag{5.34}$$

where $f_1 = 0.05$, $f_2 = 0.3$, and the amplification factor $A_0 = 10^5$. Two questions arise here. How well can we describe the periodic components; effectively, line features in the spectrum, and is there any spectral leakage introduced due to the Quadratic algorithm? Figure 5.6 shows the results of spectral analysis from both TMT (gray) and QMT, on linear and log axes. The linear plot clearly shows the more accurate description of the periodic components provided by the QMT, the logarithmic plot shows that no additional spectral leakage is introduced. Note how both methods overlap at very low amplitudes. The signal has 8 to 10 orders of magnitude dynamic range and both methods behave similarly in terms of spectral leakage.

The second test signal has a much smaller signal-to-noise ratio. In this case we let the amplification factor be $A_0 = 1.0$ and the standard deviation σ of $r(t)$ remains fixed. Figure 5.7 shows the result of spectral analysis on this signal. We would like to stress two important features that can be seen from these results. First, the linear components are better described by the QMT. Second, the information outside the range of the periodic components is smoother, as shown in the random signal example (Table 5.2).

Additionally, we have also obtained extra information about the spectrum. Figure (5.7) shows the estimate of the first and second derivatives of the spectral contents of the signal. As expected, the first derivative should be very close to zero, when getting close to the periodic component. Similarly, the second derivative of the spectrum should have a large negative value, showing that the line represents a local maxima of the spectrum.. These two features are clearly present in the estimates of the derivatives. Given the randomness of the signal and also uncertainties due to the non-uniqueness of the problem in our example, the second derivative around the 0.3Hz component is not exactly the largest negative value, but is rather off by a frequency bin. This shows that still some uncertainties remain in all estimates, including the spectrum and its derivatives. Nevertheless, the extra information that is gained from the derivatives could certainly be relevant.

Note that the estimates of the derivatives are not computed by a numerical differentiation of the spectrum estimate $\hat{S}(f)$, but rather by the steps described in the previous sections.

A method for the detection of periodic components using the multitaper algorithm was developed by Thomson (1982), known as the F-test for spectral lines. For a complete description of the methodology the reader is refered to Thomson (1982) and Percival and Walden (1993). The method can be applied for reshaping the spectrum near spectral lines (e.g., Park *et al.*, 1987a; Thomson, 1990; Denison *et al.*, 1999) or even for removal of these periodic signals embedded in a colored spectrum (Lees, 1995; Percival and Walden, 1993, Chapter 10). For high signal-to-noise ratios as in Figure 5.6, F-test or other line detection algorithms are preferred for harmonic analysis.

The general idea of spectral reshaping is to subtract the effect of the statistically

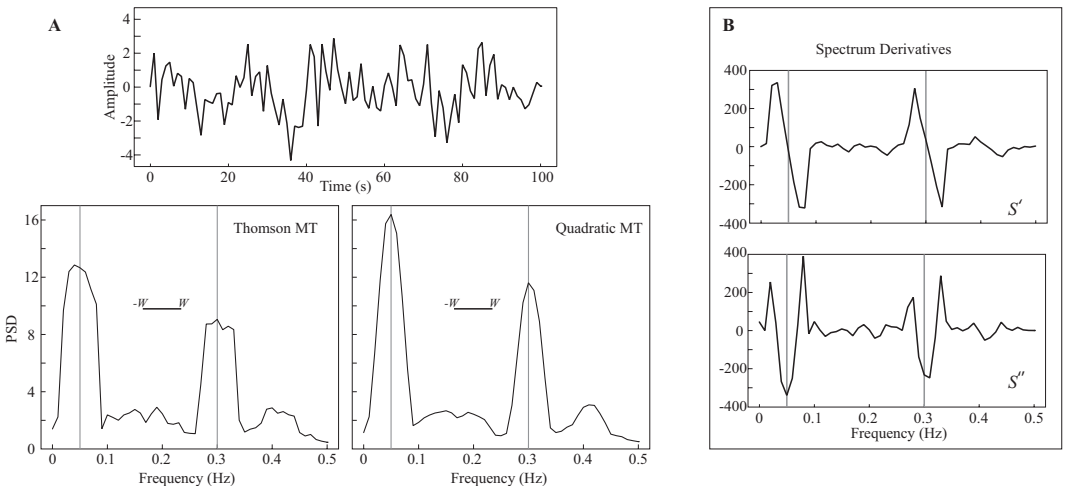

Figure 5.7: Spectrum estimation of a normally distributed random signal with two sinusoids at 0.05 Hz and 0.3 Hz. Parameters as in Figure 5.6. Plots below the time series contain the estimates of the spectrum using the TMT (left) and the QMT (right). The figures in the bounded box (right) show the first and second derivatives as estimated from the covariance matrix. Vertical gray lines represent the location of the line components. Note how the derivative estimates help in pinning where the line components are located.

significant lines from the eigencomponents Y_k (see equation 5.22). This subtraction is done before the adaptive weighting in equations (5.24) and (5.25), meaning that it is possible to perform either TMT or the QMT estimates on the remaining stochastic part of the spectrum (without the deterministic periodic components, just removed), with similar improvements as presented in the examples above, if the Quadratic algorithm is applied.

5.6.3 Resolution test and the choice of multitaper parameters

One important question that remains unanswered in spectral analysis using multitaper methods is, what is the optimal choice of the time-bandwidth product NW (the averaging bandwidth) and the number of tapers K (the more tapers, the smoother the estimate)? Or, having a chosen bandwidth W, what is the ideal number of tapers that should be used?. Riedel and Sidorenko (1995) invented the sine multitaper method to get around this problem by choosing an optimal number of tapers iteratively at each frequency.

In the multitaper literature, it has been proposed that $K = 2NW - 1$ as

an appropriate choice, since the eigenvalues λ_k are all close to unity. This choice is essentially based on the leakage properties of the tapers, but does not take into account the particular shape of the spectrum of the signal under analysis. In this subsection, we present a comparison of the effect of the choices of NW and K on the resolution of the spectra around a periodic component. We show how the QMT is less dependent on these choices compared with TMT.

In Figure 5.8 and Table 5.3 we compare the resolution of TMT and the Quadratic multitaper. A useful criterion is that of the width of the half-power points, also known as the 3-dB bandwidth. This criterion reflects the fact that two equal-strength periodic components separated by less than the 3-dB bandwidth will show in the spectrum as a single peak instead of two (Harris, 1978). We use the signal in the previous section (Figure 5.6) and plot the spectrum centered on one of the periodic components on a dB scale defined as:

$$dB = 10 \log_{10} \left(S(f)/S(f_0) \right) \tag{5.35}$$

where f_0 is the frequency of the periodic component. We vary the time-bandwidth NW and the number of tapers K to investigate the effect of these choices. We also present in Table 5.3 the result of the 3-dB and 9-dB ($\frac{1}{8}th$ power) bandwidths for the different choices of NW and K by applying a linear interpolation.

The QMT always outperforms Thomson's algorithm given the same parameters as shown in Table 5.3. Note that at the 9-dB line in Figure 5.8 (see Table 5.3 as well) both methods provide similar results, with the method introduced here being slightly better.

An important result obtained by conducting this test is the fact that the 3-dB bandwidth is less sensitive to the choice of NW for the QMT (see Figure 5.8A). Once we reach the 9-dB bandwidth, a larger value of NW decreases the resolving power. On the other hand the choice of K is directly proportional to the resolution bandwidth for both algorithms (this is also evident from Figure 5.2), with the Quadratic algorithm having narrower 3-dB and 9-dB bandwidths in all cases. A final observation from Table 5.3 indicates that a comparatively better resolution is achieved with the QMT even if one more taper K is used compared to TMT (compare also red and gray lines in Figure 5.8B), leading to smoother estimates due to the increased degrees of freedom.

Table 5.3: Comparison between multitaper methods by their 3-dB and 9-dB bandwidths in Rayleigh frequency f_R units with different choices of time-bandwidth product NW and number of tapers K for a periodic component in the spectrum in Figure 5.6.

		3-dB		9-dB	
K	NW	Thomson	Quadratic	Thomson	Quadratic
7	3.0	3.07	2.85	3.22	3.16
7	3.5	3.11	2.96	3.32	3.22
7	4.0	3.33	2.84	4.01	3.65
7	4.5	3.65	2.75	4.16	4.03
7	5.0	4.03	2.81	4.28	4.10

Constant number of tapers, varying time-bandwidth

		3-dB		9-dB	
K	NW	Thomson	Quadratic	Thomson	Quadratic
4	4.0	2.47	1.79	3.08	2.68
5	4.0	3.01	2.05	3.25	3.08
6	4.0	3.17	2.54	3.53	3.29
7	4.0	3.33	2.84	4.01	3.65
8	4.0	4.01	3.23	4.17	4.09

Varying number of tapers, constant time-bandwidth

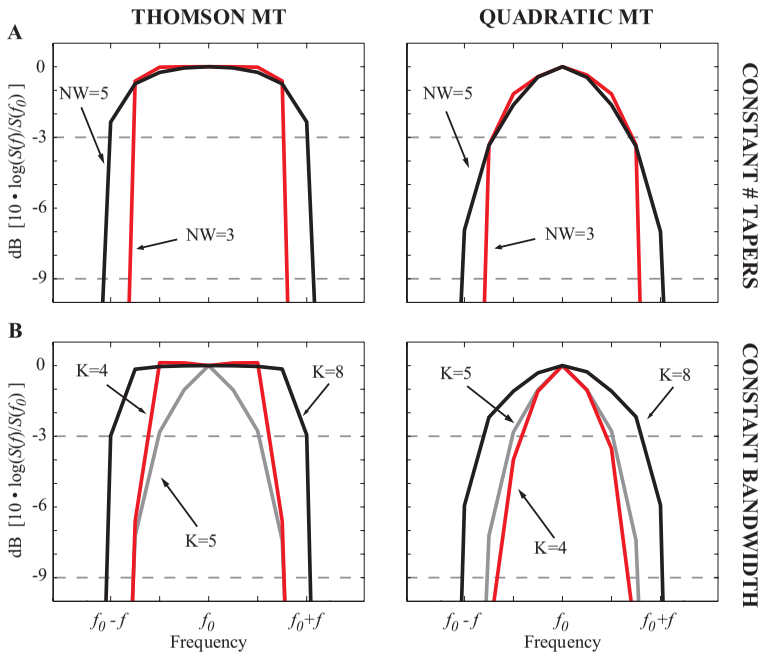

Figure 5.8: Resolution test comparison between TMT and the QMT with different choices of time-bandwidth product NW and number of tapers K for a periodic component in the spectrum in Figure 5.6. In A we plot the spectral shape around a periodic component at frequency f_0 for two choices of time-bandwidth product NW with $K = 7$ tapers using the two methods. Two dashed lines represent the 3-dB (half-power) and 9-dB lines. The QMT has a 3-dB bandwidth that is narrower than the equivalent multitaper and is relatively independent of the choice of NW. In B we fix the time-bandwidth product $NW = 4.0$ and vary the number of tapers. For the same choice of parameters the QMT outperforms the TMT with a narrower 3-dB bandwidth. We added the Quadratic estimate using $NW = 4.0$ and $K = 5$ on both plots for reference (gray line), resulting in a similar resolution to that of the $K = 4$ TMT (red line on left plot) showing that the QMT can effectively provide smoother estimates (by using more tapers) without the loss of resolution.

5.6.4 Synthetic earthquake signal

In geophysical applications many signals have spectral shapes with large dynamic range (red spectra) but rarely with deterministic components (periodic signals). The spectra are continuous, for example the Earth's background seismic noise (Berger *et al.*, 2004), medium and small sized earthquake sources (e.g., Abercrombie, 1995; Prieto *et al.*, 2004), the crustal magnetic field (Korte *et al.*, 2002), and many others.

Consider the spectrum of an earthquake, which follows the Brune (1970) model:

$$\dot{u}(f) = \frac{2\pi f M_0}{1 + (f/f_c)^2} \tag{5.36}$$

where $\dot{u}(f)$ is the velocity amplitude source spectrum associated with the earthquake, M_0 is the seismic moment (related to the size of the earthquake) and f_c is the corner frequency. The corner frequency represents the predominant frequency content of the radiated seismic energy from the earthquake rupture. The spectrum from this model has a triangular shape if plotted in log-log axes with a slope of two in power.

In this synthetic example, we generate a pseudo-random time series with 1000 samples whose spectra follow the source model in Equation (5.36). Even though TMT possesses good spectral leakage reduction, the spectrum may be biased due to the quadratic effects we discussed previously. This is especially true when the corner frequency gets extremely close to the sampling frequency, so that the curvature around f_c is described by a small number of spectral points.

Figure (5.9) shows the spectral estimates from a realization of a synthetic source model with $f_c = 0.005$Hz using TMT and the QMT. The triangular shape that is expected from source spectra is better constraint using the new algorithm.

In addition to the standard spectrum, it is also possible to obtain an estimate of the derivative of the spectrum. The derivative estimate of two different source models are shown in Figure (5.10), taken from an average of 100 random realizations and corresponding standard errors. The two cases presented have corner frequencies close to the Rayleigh frequency f_R. Whenever the corner frequency is close to the sampling frequency, its curvature is represented by few spectrum bins.

The uncertainties of the derivative estimate are, similar to the uncertainties of the PSD, proportional to the amplitude of the spectrum. Using the derivative provides

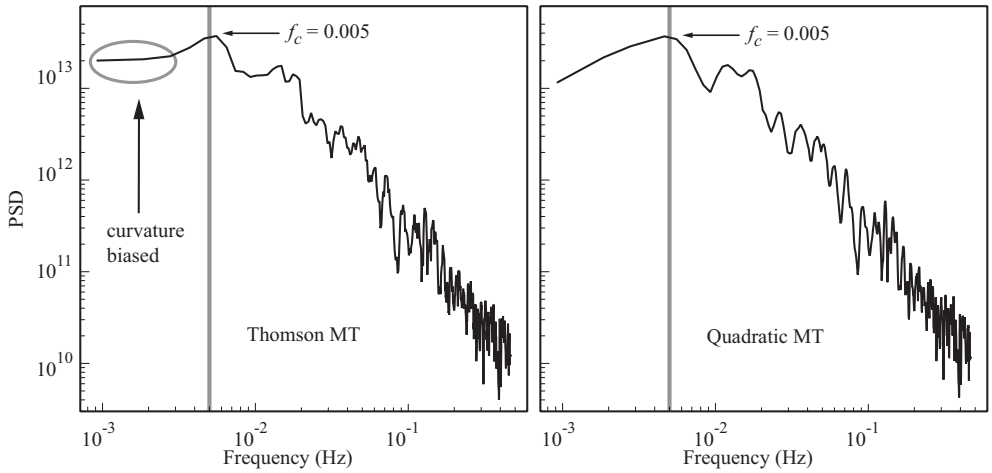

Figure 5.9: Spectrum analysis of synthetic earthquake model. The time series (not shown) has 1000 samples ($N = 1000$) and we use time-bandwidth product of 3.5. The corner frequency $f_c = 0.005$Hz (shown by an arrow) is to be estimated from the computed spectrum. Note how Thomson's algorithm biases the lower frequency part of the spectrum and does not resemble the triangular shape expected for these kind of signals. The QMT reduces the bias at lower frequencies considerably and is a better description of the shape of the spectrum around f_c.

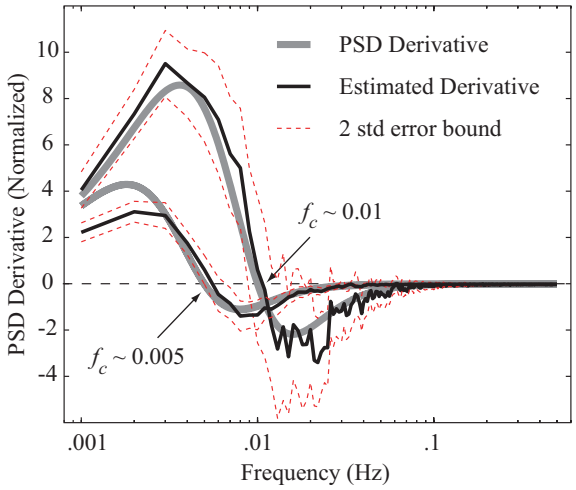

Figure 5.10: Mean derivative of the spectrum from 100 random realizations and two standard error bounds for two different earthquake models (see equation 5.36), with corner frequencies $f_c = 0.005$ and $f_c = 0.01$Hz. Time series analyzed is 1000 samples long. Note how the estimate closely resembles the model. In the case of the lower corner frequency, there is considerable bias at the lower frequency band.

additional degrees of freedom for estimating parameters from the spectrum.

5.6.5 Bathymetry profiles

A simple, isotropic, three-parameter model for the power spectrum of marine topography has been proposed of the form (Goff and Jordan, 1988):

$$S(k) = \frac{a^4}{(1 + (k/k_c)^2)^\mu} \tag{5.37}$$

where a is the amplitude of the total root-mean-square roughness of the topography, $k = |\mathbf{k}|$ is the wavenumber, $\mu > 1$ is the slope in the roll-off in the short wavelength part of the spectrum, and k_c is the corner wavenumber. We decided to use bathymetry data, given their more closely isotropic behavior (stationarity in terms of time series) and the availability of very high quality data sets.

We used bathymetry data obtained in the Central Pacific region (See Figure 5.11), drawn from ship multibeam data (Macdonald *et al.*, 1992, see also Marine Geoscience Data System, http://www.marine-geo.org). From the available data we chose 5 profiles (east-west directions), four of them crossing the mid-ocean ridge, the other along the transform fault.

The idea of this example is to show what extra information can be extracted via the QMT. The bathymetry profiles have a sampling rate of 500 samples per deg and a total of 1251 samples per profile. The location of the profiles is shown in Figure (5.11).

Figure (5.12) shows the QMT analysis of the selected profiles. The spectrum is shown to have a large dynamic range (about 7 orders of magnitude) over the entire frequency range. We focus our attention at the lower frequency range, where the corner wavenumber is expected from the model in Equation (5.37). The spectral shapes are very similar for all profiles and the different k_0's are hardly distinguishable. An independent observation of the different behavior between profiles can be drawn from the derivative estimates. See Figure (5.12) and caption for discussion.

From the methods described above, we can obtain estimates of the derivative of the spectrum, and by normalizing by the spectrum,

$$\frac{S'(f)}{S(f)} = \frac{d}{df} \{\log S(f)\} \tag{5.38}$$

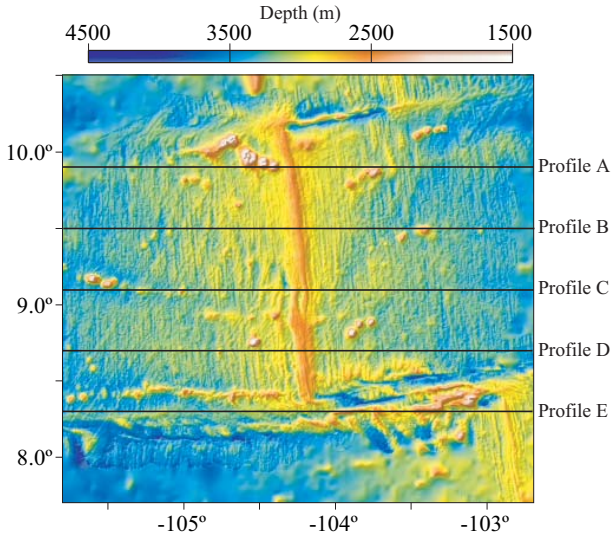

Figure 5.11: Location of the study area, where the profiles were taken from. Four of these profiles run across the mid-ocean ridge, and one is paralel to the transform fault. Location of the profiles is shown as thick black lines.

we have an estimate of the derivative of the log spectrum (Thomson, 1994). The derivatives also provide the means for comparing the different profiles, and clearly show the presence of two groups with particular spectral characteristics. This suggests that the profiles sampling the transform fault posses a lower corner frequency than the profiles sampling the mid-ocean ridge structure.

5.7 Conclusions

Multitaper spectral analysis can be obtained by multiplying the data by a set of orthogonal (in time and frequency) sequences, all having good spectral leakage properties. The sequences have the property to concentrate within a band $2W$ the frequency content of the spectral estimate. A simple average of the eigenspectra $\hat{S}_k$ is not ideal, given the large dynamic ranges of some signals, and an adaptive weighting function is necessary, especially in regions where the spectrum has low amplitudes, and thus is prone to leakage from frequencies that have much larger amplitudes.

As noted by Riedel and Sidorenko (1995) and confirmed in this study, in regions where spectral leakage is not expected, corresponding to the regions of the spectrum with large amplitudes, the local or quadratic bias can have an important effect on the shape

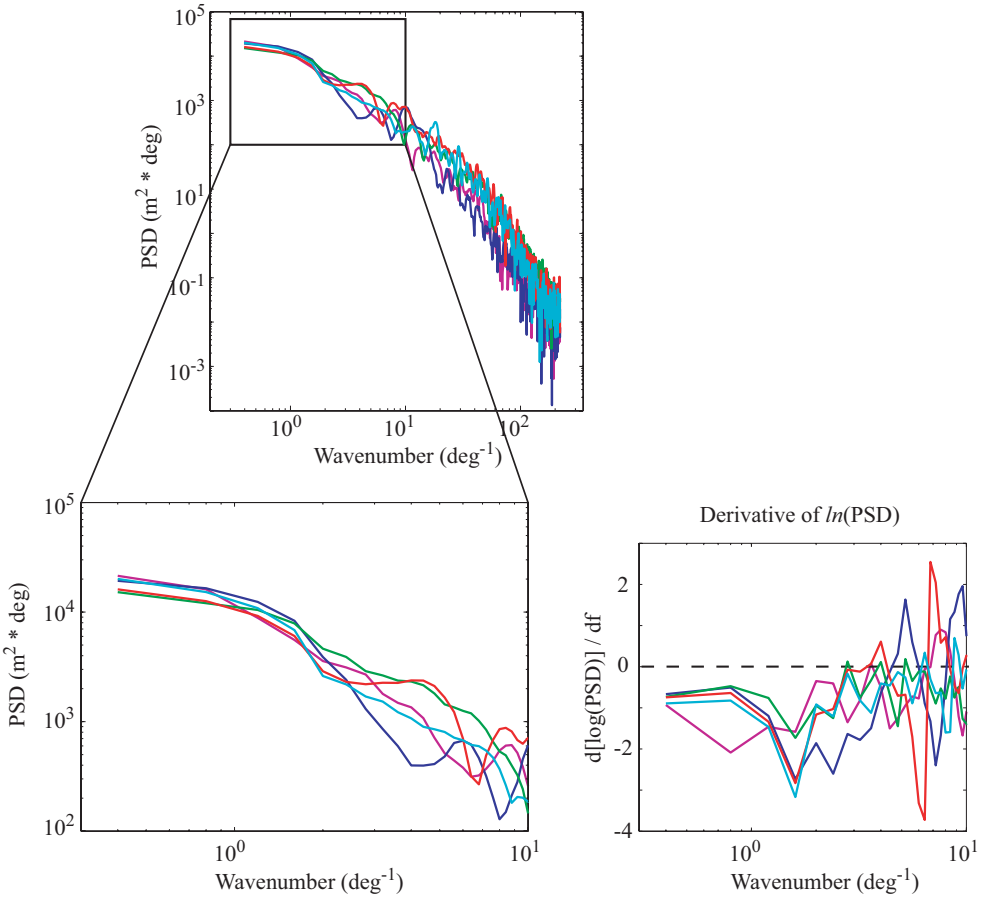

Figure 5.12: Spectral Analysis of 5 selected profiles of bathymetric data in the Central Pacific Ocean around 9°N (see map in Figure 5.11). QMT (left plots) show very similar behavior of the spectra. In addition to the standard spectra, the algorithm provides an estimate of the derivative of the spectra (right plot). We show the scaled derivative (approximately the derivative of ln(PSD). Note that the spectra can be grouped according to the derivatives, with one profile having a quite different behavior (magenta line) showing a lower corner frequency. This derivative corresponds to profile E in Figure 5.11, which samples the transform fault; while the rest of the profiles sample the mid-ocean ridge.

of the spectrum.

We introduce the Quadratic multitaper method, which estimates the derivatives of the spectrum, that is, the slope and curvature of the spectrum on the interval $(-W, W)$, by solving a parameter estimation problem relating the derivatives of the spectrum and the K eigencomponents.

With the estimation of the second derivative (the curvature of the spectrum) we can apply a correction to the spectrum to obtain a new estimate that is unbiased to quadratic structure. This algorithm reduces to the original Thomson (1982) multitaper when the spectrum is locally flat in the interval $(-W, W)$.

We present a variety of examples that indicate that the Quadratic multitaper provides a smoother, less biased spectral estimate of the data, in addition to independent estimates of the derivatives of the spectrum. When the dynamic range of the signal is very large, the improvements are not as striking, but the information contained in the slope estimates can readily be applied in parameter estimation, or as an additional discriminant to compare two signals. No additional spectral leakage was introduced in the examples shown in this study.

We also discuss the effect of chosen multitaper parameters such as the time-bandwidth product and the number of tapers to compute. Even though it is still a user-defined set of parameters, we show that the Quadratic multitaper leads to increased resolution compared to TMT and it is less dependent on the choice of the time-bandwidth in the inner interval. It allows the use of more tapers without the loss of resolution power compared to Thomson's algorithm.

Finally, model parameters can be found by analysing the goodness-of-fit between a Quadratic spectral estimate plus the slope of the spectrum of a data set with a theoretical model of the spectrum and its derivative. Another approach would be to generate from the theoretical model a covariance matrix C_{jk} (as in Equation 5.27) and find the model that best fits the data. In the later case, the information is not restricted to curvature; rather all information from the theoretical models is used.

Acknowledgments

We thank David Sandwell and J.J. Becker for providing the bathymetry data used here. Funding for this research was provided by NSF Grant number EAR0417983.

6

Conclusions

6.1 Main Results

This thesis presents new methods to estimate and reduce the uncertainties and biases that are unavoidable in the analysis of the earthquake rupture process. The methods include spectral stacking (Prieto *et al.*, 2004, Chapter 3), empirical Green function (EGF) analysis (Prieto *et al.*, 2006, Chapter 4), source spectral fitting and uncertainty estimation (Prieto *et al.*, 2007a, Chapter 2), and improving analysis tools to reduce bias and variance in spectrum estimates (Prieto *et al.*, 2007b, Chapter 5). We have presented three basic results:

1. Measuring the uncertainties in earthquake source parameter estimation.
2. Reducing these uncertainties by stacking or averaging spectral estimates.
3. Generating a multitaper spectral estimate with significant bias reduction, with extra independent information from the slope of the PSD.

As discussed in the introduction (see Figure 1.6), in order to compare different studies of the earthquake rupture it is useful to have a measure of some static and dynamic source parameters, as well as their associated uncertainties. The error analysis is necessary if issues regarding scaling relations and source parameter size dependence are to be discussed.

In **Chapter 2** we use a multitaper spectrum algorithm combined with jackknife statistical analysis to obtain, from the seismic spectrum, the source parameters and their

confidence intervals. Analysis of the Cajon Pass Borehole finds a slight increase in stress drop with earthquake magnitude, deviating from the constant stress drop expected if earthquake ruptures are self-similar.

Two additional observations from the results are relevant. First, stations close to each other (i.e., two borehole stations at 2.9 km and 1.5 km depths) may produce very different estimates of the source parameters, indicating the need to address the significance of individual estimates when comparing results. Second, given the assumptions used in the analysis (radiation pattern, earthquake locations, velocity model, etc.), it is likely that the confidence intervals presented here represent a lower bound. This suggests we need to find ways of reducing uncertainties, and the subsequent chapters seek to address this point.

In **Chapter 3**, we study the scaling relationships of source parameters and the self-similarity seen in 400 small earthquakes located in a compact region near the Anza Seismic Network in southern California. By iteratively stacking P and S spectra, we are able to separate source and receiver contributions and, after an EGF correction, obtain a relative earthquake source spectrum. Due to the large number of events and multiple stations, the stacked source spectra are smooth compared to individual earthquake estimates. We show using standard scaling relationships that for the magnitude range $M_L = 1.8$ to 3.4 the earthquake rupture is self-similar. The static scaling relation between the size of the earthquake – given by the seismic moment M_0 – and the length scale of the rupture measured from the corner frequency f_c is $M_0 \propto f_c^{-3}$. The dynamic parameter apparent stress σ_a shows an average value of 1MPa and is constant as a function of M_0. This is confirmed by directly testing for self-similarity with the spectral shapes at various seismic moments. We applied a similar methodology to a comprehensive analysis of over 60,000 earthquakes in southern California in Shearer, Prieto, and Hauksson (2006).

The empirical Green function, where a small earthquake is taken as an approximate impulse response between a source and a receiver, has been shown in multiple cases to be a very useful method. It has a number of drawbacks, however, with one of them being that the uncertainties after performing the EGF deconvolution (in time or frequency domains) tend to be very large.

In **Chapter 4** we take advantage of the large data set available in southern California to average spectral ratios of the 2001 $M5.1$ Anza earthquake and 160 aftershocks. Due to the large number of EGFs used, we can use a weighted averaging of the spectral ratios and are able to control the trade-off between the variance reduction and the bias inherent in the EGF approximation by using a mean-square error criteria. Using propagation of errors, we also obtain the confidence intervals of the source spectrum and dynamic source parameters (radiated seismic energy E_S and apparent stress σ_a).

The results of this study show that in order to reduce uncertainties, it is necessary to use multiple EGFs; otherwise, even if an ideal EGF is available, the large uncertainties would make the result less significant. Note in Figure 1.6 that the error bars of the σ_a estimate for this particular study are almost invisible in the given scale, while this is not so for the majority of the other data points in the same plot.

In **Chapter 5** we present an improvement to the multitaper spectral analysis (TMT) tools introduced by Thomson (1982), which we have used throughout this thesis. As discussed in this chapter, the TMT algorithm suffers from *local bias*, i.e., at frequencies close to spectral lines the spectrum estimate is approximately constant. The Quadratic multitaper (QMT) introduced can be used to estimate the derivatives of the spectrum as a function of frequency. The estimate of the second derivative is used to apply a correction and obtain a spectrum that is unbiased with respect to quadratic structure. In addition, we present the use of the first derivative – the slope – of the spectrum for comparing signals from marine topography profiles.

In order to use the multitaper algorithm (either TMT or QMT), it is necessary to define two basic parameters: a frequency bandwidth (W) over which the user wishes to *average* the spectrum and the number (K) of tapers to compute. If the user chooses a very narrow bandwidth or very few tapers, the spectral estimate will have large uncertainties and will not be smooth. A considerable advantage of the QMT presented in **Chapter 5** is that for the same choice of parameters (W and K), the spectrum has higher resolution and is smoother than TMT estimates. We also show that the user may choose to compute one more taper without reducing the resolution of the estimate compared to TMT and since one more taper is used, an even smoother estimate is expected.

6.2 Future Research Directions

The methodologies developed during this research should provide considerable advantages in studying the physics of earthquake rupture from seismic recordings. They provide improved estimation of some earthquake source parameters such as stress drop and radiated seismic energy. The methods proposed to reduce uncertainties can easily be applied to different data sets in which either lots of earthquakes or lots of stations are available. In the applications we presented mainly small and medium sized earthquakes ($M < 6.0$) analyzed at either local or regional distances.

The stacking or averaging methods are not useful for large teleseismic earthquakes, given that at teleseismic distances it is rare to record many small earthquakes near a large one. Another difficulty at teleseismic distances is that large earthquakes rupture over tens of kilometers, requiring different EGFs for different patches of the fault, and complicating any averaging between these patches. Nevertheless, EGF methodologies have been adapted for prediction or simulation of ground motion from large earthquakes using small earthquakes as EGFs (e.g., Wössner *et al.*, 2002).

One particularly interesting approach is the analysis of earthquake data recorded by small aperture arrays. The idea is to look at the consistency of spectral analysis across the array. If two earthquakes – one large and one small that can be used as an EGF – are recorded by the stations of the small aperture array, we can test the stability of the EGF deconvolution. It is assumed that all stations are sampling an identical region of the focal sphere and that the EGF is removing the effects of attenuation, near-site effects, scattering, etc., so the result should be identical for every station. We can then test the uncertainties and resolving power of the stations in the array, given that the Quadratic multitaper (QMT) will provide the least variable, least biased estimate, which will then be used in the spectral division.

With the advent of high-quality borehole seismic networks, it is now possible to investigate the rupture properties of smaller earthquakes. The increasingly available data from the Japanese Hi-net and F-net networks (with over 700 stations), the EarthScope PBO Borehole Strainmeter and Seismometer network with 7 stations near the Anza Seismic Gap, and the SAFOD array, with seismic sensors along the borehole and some

sensors a few hundred meters from known seismic sources in the San Andreas Fault can be exploited using the methodologies described in this thesis.

Using EGF methods, McGuire (2004) showed that it was possible to investigate the fault plane, rupture length, and directivity of M5 earthquakes by using a network of surface stations. Having the borehole data described above, which are expected to have better signal-to-noise ratios and record signals at higher frequencies than surface instruments, we have the opportunity to investigate the rupture properties of even smaller earthquakes (M2 - M4). The network of stations is needed to analyze the variability of source properties as a function of the take-off angle (as a function of the focal sphere sampled at a given station), helpful in solving the fault plane ambiguity for a double-couple source and providing information about directivity of small earthquakes.

Another important application is the use of the methodologies described here for the analysis of attenuation structure, which we have in part neglected. From the stacking procedures or from EGF analysis, one can potentially obtain information about the properties of the medium through which the seismic waves have traveled, i.e., Q structure (attenuation) or near-site effects (see for example Warren and Shearer, 2000; Tsuda *et al.*, 2006; Hauksson and Shearer, 2006). Variations in Q structure may be caused by cracks, chemical composition and temperature variations, which can then be modeled from the spectral analysis.

There are some limitations to using the Quadratic multitaper. As explained in Chapter 5, QMT provides an estimate of the power spectrum and we lose the phase information. This clearly limits our ability to investigate the source time function (STF) and the behavior of the radiated seismic energy in time. For example Mori *et al.* (2003) use the STF to investigate the dynamic stress drops or average frictional stress $ave\{\sigma_f\}$ (see Figure 1.2) by looking at the initial slope of the deconvolved source time functions. Nevertheless, estimation of radiated seismic energy in the spectral domain might be better constrained by using multitaper methods, where uncertainties are easily quantified, and even STF can potentially be extracted if Thomson's multitaper is used (for an example deconvolution applied to receiver functions see: Park and Levin, 2000).

A

Propagation of errors

Some rules of propagation of errors (Taylor, 1997) are listed here. Assume u, v are random variables with associated variance σ_u^2, σ_v^2, and covariance σ_{uv}^2. Constants a, b do not contribute to uncertainties. Then we have:

$$
\begin{aligned}
x = au \pm bv \qquad & \sigma_x^2 = a^2\sigma_u^2 + b^2\sigma_v^2 + 2ab\sigma_{uv}^2 \\
x = auv \qquad & \frac{\sigma_x^2}{x^2} = \frac{\sigma_u^2}{u^2} + \frac{\sigma_v^2}{v^2} + 2\frac{\sigma_{uv}^2}{uv} \\
x = \frac{au}{v} \qquad & \frac{\sigma_x^2}{x^2} = \frac{\sigma_u^2}{u^2} + \frac{\sigma_v^2}{v^2} - 2\frac{\sigma_{uv}^2}{uv} \\
x = au^{\pm b} \qquad & \frac{\sigma_x}{x} = b\frac{\sigma_u}{u} \\
x = ae^{\pm bu} \qquad & \frac{\sigma_x}{x} = b\sigma_u \\
x = a\ln(\pm bu) \qquad & \sigma_x = a\frac{\sigma_u}{u}
\end{aligned}
$$

B

Quadratic mean-square error

In Equation (5.32), a correction for the curvature or quadratic bias is applied to the multitaper estimate. In section 5.5, we defined the expected value of the spectrum as an average over the inner interval $(-W, W)$:

$$E\left[\hat{S}(f)\right] = S(f) + \frac{1}{6}W^2 S''(f) \tag{B.1}$$

and by applying the correction in 5.32, we have the expected value of the Quadratic multitaper

$$E\left[\tilde{S}(f)\right] = S(f) + \frac{1}{6}W^2 S''(f) - \mu\frac{1}{6}W^2 S''(f) \tag{B.2}$$

where we assume that $E[\hat{\alpha}_2] = S''(f)$.

The bias of the Quadratic multitaper is then

$$E\left[\beta\right] = E\left[\tilde{S}(f)\right] - S(f) = \frac{1}{6}W^2 S''(f)(1-\mu) \tag{B.3}$$

and the variance, using the rules of propagation of errors (Taylor, 1997) in Equation 5.32,

$$var\{\tilde{S}\} = var\{\hat{S}\} + \mu^2\frac{W^4}{36}var\{S''\} \tag{B.4}$$

and we can now define the mean square error (bias^2 + variance):

$$L = \left[(1-\mu)\frac{W^2}{6}S''\right]^2 + var\{\hat{S}\} + \mu^2\frac{W^4}{36}var\{S''\} \tag{B.5}$$

where the first term is the bias squared and the two on the right represent the variance. It is assumed that the covariance is insignificant. Taking the derivative with respect to

μ and setting to zero

$$\frac{\partial L}{\partial \mu} = (1-\mu)\left[S''\right]^2 + \mu\, var\{S''\} = 0 \tag{B.6}$$

and rearranging yields

$$\mu = \frac{[S'']^2}{([S'']^2 + var\{S''\})} \tag{B.7}$$

which is the solution shown in 5.33, using α_2 as estimates of S''.

To obtain the variance of the estimates $\hat{\alpha}_0$, $\hat{\alpha}_1$, and $\hat{\alpha}_2$ in the least squares problem (5.30), we compute the covariance matrix of the coefficients, following Lawson and Hanson (1974).

Bibliography

Abercrombie, R. E. (1995). Earthquake source scaling relationships from -1 to 5 M_L using seismograms recorded at 2.5-km depth. *J. Geophys. Res.*, **100**, 24015–24036.

Abercrombie, R. E. (1996). The magnitude-frequency distribution of earthquakes recorded with deep seismometers at Cajon Pass, southern California. *Tectonophysics*, **261**, 1–7.

Abercrombie, R. E. (1997). Near-surface attenuation and site effects from comparison of surface and deep borehole recordings. *Bull. Seism. Soc. Am.*, **87**(3), 731–744.

Abercrombie, R. E. and Leary, P. (1993). Source parameters of small earthquakes recorded at 2.5 km depth, Cajon Pass, southern California: Implications for earthquake scaling. *Geophysics Research Letters*, **20**, 1511–1514.

Abercrombie, R. E. and Rice, J. R. (2005). Can observations of earthquake scaling constrain slip weakening? *Geophys. J. Int.*, **162**, 406–424.

Aki, K. (1967). Scaling law of seismic spectrum. *J. Geophys. Res.*, **72**, 1217–1231.

Aki, K. and Richards, P. (1980). *Quantitative Seismology*. W. H. Freeman, New York.

Andrews, D. J. (1986). Objective determination of source parameters and similarity of earthquakes of different size. In S. Das, J. Boatwright, and C. H. Scholz, editors, *Earthquake Source Mechanics, Geophys. Monogr. Ser.*, pages 259–267. AGU, Washington, DC.

Archuleta, R. J., Cranswick, E., Mueller, C., and Spudich, P. (1982). Source parameters of the 1980 mammoth lakes, california, earthquake sequence. *J. Geophys. Res.*, **87**(B6), 4595–4607.

Bard, Y. (1974). *Nonlinear parameter estimation*. Academic Press, New York.

Berger, J., Davis, P., and Ekström, G. (2004). Ambient earth noise: A survey of the global seismographic network. *J. Geophys. Res.*, **109**(B11307).

Berger, J. L., Baker, L., Brune, J., Fletcher, J., Hanks, T., and Vernon, F. L. (1984). The Anza array: A high dynamic range, broadband, digitally radiometered seismic array. *Bull. Seism. Soc. Am.*, **74**, 1469–1481.

Boatwright, J. A. (1980). A spectral theory for circular seismic sources: Simple estimates of source dimension, dynamic stress drop and radiated energy. *Bull. Seism. Soc. Am.*, **70**, 1–27.

Boatwright, J. A. and Fletcher, J. B. (1984). The partition of radiated energy between P and S waves. *Bull. Seism. Soc. Am.*, **74**, 361–376.

Brodsky, E. E. and Kanamori, H. (2001). Elastohydrodynamic lubrication of faults. *J. Geophys. Res.*, **106**, 16357–16374.

Bronez, T. P. (1992). On the performance advantage of multitaper spectral analysis. *IEEE Trans. Sig. Proc.*, **40**(12), 2941–2946.

Brune, J., Brown, S., and Johnson, P. (1993). Rupture mechanics and interface separation in foam rubber models of earthquakes: a possible solution to the heat flow paradox and the paradox of large overthrusts. *Tectonophysics*, **218**, 59–67.

Brune, J. N. (1970). Tectonic stress and seismic shear waves from earthquakes. *J. Geophys. Res.*, **75**, 4997–5009.

Chappellaz, J., Barnola, J. M., Raynaud, D., Korotkevich, Y. S., and Lorius, C. (1990). Ice-core record of atmospheric methane over the past 160,000 years. *Nature*, **345**, 127–131.

Choy, G. L. and Boatwright, J. L. (1995). Global patterns of radiated seismic energy and apparent stress. *J. Geophys. Res.*, **100**, 18205–18228.

Constable, S. and Constable, C. (2004). Observing geomagnetic induction in magnetic satellite measurements and associated implications for mantle conductivity. *Geochem. Geophys. Geosyst.*, **5**, Q01006.

Cramér, H. (1940). On the theory of stationary random processes. *Ann. of Math.*, **41**, 215–230.

Daly, E., Brown, C., Stark, C. P., and Ebinger, C. J. (2004). Wavelet and multitaper coherence methods for assesing the elastic tichkness of the Irish Atlantic margin. *Geophys. J. Int.*, **159**, 445–459.

Davidson, A. C. and Hinkley, D. V. (1997). *Bootstrap Methods and their Application*. Cambridge University Press, New York.

Denison, D. G. T., Walden, A. T., Balogh, A., and Forsyth, J. (1999). Multitaper testing of spectral lines and the detection of the solar rotation frequency and its harmonics. *Appl. Statist.*, **48**, 427–439.

Doob, J. L. (1952). *Stochastic processes*. John Wiley and Sons, New York.

Douglas, A., Hudson, J. A., and Pearce, R. G. (1988). Directivity and the Doppler effect. *Bull. Seism. Soc. Am.*, **78**, 1367–1372.

Efron, B. (1982). *The Jackknife, the Bootstrap, and other resampling plans*. SIAM, Philadelphia.

Eshelby, J. D. (1957). The determination of the elastic field of an ellipsoidal inclusion and related problems. *Proc. Roy. Soc. London Series A*, **241**, 376–396.

Fletcher, J., Boatwright, J., Haar, Hanks, and McGarr, A. (1984). Source parameters for aftershocks of the oroville california earthquake. *Bull. Seism. Soc. Am.*, **74**(4), 1101–1123.

Fox, T., Hinkley, D. V., and Larntz, K. (1980). Jackknifing in nonlinear regresion. *Technometrics*, **22**, 29–33.

Frankel, A., Fletcher, J., Vernon, F., Haar, L., Berger, J., and Hanks, T. (1986). Rupture characteristics and tomographic source imaging of $M_L \sim 3$ earthquakes near Anza, Southern California. *J. Geophys. Res.*, **91**, 12633 – 12650.

Gilbert, F. (1970). Excitation of normal modes of the Earth by earthquake sources. *Geophys. J. Royal Astr. Soc.*, **22**, 223–226.

Goff, J. A. and Jordan, T. H. (1988). Stochastic modeling of seafloor morphology: inversion of sea beam data for second-order statistics. *J. Geophys. Res.*, **93**, 13589–13608.

Hanks, T. C. and Boore, D. M. (1984). Moment-magnitude relations in theory and in practice. *J. Geophys. Res.*, **89**, 6229 – 6235.

Harris, F. (1978). On the use of windows for harmonic analysis with the discrete Fourier transform. *Proceedings of the IEEE*, **66**, 51–83.

Hartzell, S. H. (1978). Earthquake Aftershocks as Greens Functions. *Geophysics Research Letters*, **5**, 1–4.

Hauksson, E. and Shearer, P. M. (2006). Attenuation models ($Q_P and Q_S$) in three dimensions of the southern California crust: Inferred fluid saturation at seismogenic depths. *J. Geophys. Res.*, **111**, B05302.

Hauksson, E., Jones, L. M., Perry, S., and Hutton, K. (2002). Emerging from the Stress shadow of the 1992 Mw7.3 Landers southern California earthquake? A preliminary Assessment. *Seismol. Res. Let.*, **73**, 33–38.

Hinkley, D. V. (1977). Jackknife confidence limits using Student t approximations. *Biometrika*, **64**, 21–28.

Hough, S. E. (1997). Empirical Green's function analysis: Taking the next step. *J. Geophys. Res.*, **102**, 5369 – 5384.

Hough, S. E. and Field, E. H. (1996). On the coherence of ground motion in the San Fernando Valley. *Bull. Seism. Soc. Am.*, **86**(6), 1724–1736.

Ida, Y. (1972). Cohesive force across the tip of a longitudinal-shear crack and Griffiths specific surface energy. *J. Geophys. Res.*, **77**, 3796–3805.

Ide, S. and Beroza, G. C. (2001). Does apparent stress vary with earthquake size? *Geophysics Research Letters*, **28**, 3349–3352.

Ide, S., Beroza, G. C., Prejean, S. G., and Ellsworth, W. L. (2003). Apparent break in earthquake scaling due to path and site effects on deep borehole recordings. *J. Geophys. Res.*, **108**(B5 2271).

Izutani, Y. and Kanamori, H. (2001). Scale dependence of seismic energy-to-moment ratio for strike-slip earthquakes in Japan. *Geophysics Research Letters*, **28**, 4007–4010.

Jeffreys, H. (1942). On the mechanics of faulting. *Geol. Mag.*, **79**, 291.

Kanamori, H. (1977). The energy release in great earthquakes. *J. Geophys. Res.*, **82**, 2981–2987.

Kanamori, H. (2005). Real-time seismology and earthquake damage mitigation. *Annu. Rev. Earth Planet. Sci.*, **33**, 195–214.

Kanamori, H. and Anderson, D. L. (1975). Theoretical basis of some empirical relations in seismology. *Bull. Seism. Soc. Am.*, **65**(5), 1073–1095.

Kanamori, H. and Brodsky, E. E. (2004). The physics of earthquakes. *Reports on Progress in Physics*, **67**, 1429–1496.

Kanamori, H. and Heaton, T. H. (2000). Microscopic and macroscopic physics of earthquakes. In J. Rundle, D. L. Turcotte, and W. Kein, editors, *Geocomplexity and the Physics of Earthquakes*, pages 147–155, Washington, D. C. American Geophysical Union. Monograph 120.

Kanamori, H. and Rivera, L. (2004). Static and dynamic scaling relations for earthquakes and their implications for rupture speed and stress drop. *Bull. Seism. Soc. Am.*, **94**, 314–319.

Kanamori, H. and Rivera, L. (2006). Energy partitioning during an earthquake. In Abercrombie, McGarr, Kanamori, and diToro, editors, *Earthquakes: Radiated Energy and the Physics of Faulting*, pages 3–13. AGU Geophys. Monograph 170.

Kanamori, H., Mori, J., Hauksson, E., Heaton, E., Hutton, T., and Jones, K. (1993). Determination of earthquake energy release and M_L using terrascope. *Bull. Seism. Soc. Am.*, **83**(2), 330–346.

Korte, M., Constable, C. G., and Parker, R. L. (2002). Revised magnetic power spectrum of the oceanic crust. *J. Geophys. Res.*, **107**(B9), 2205.

Lachenbruch, A. H. (1980). Frictional heating, fluid pressure and the resistance to fault motion. *J. Geophys. Res.*, **85**, 6097–6112.

Lawson, C. L. and Hanson, R. J. (1974). *Solving Least Squares Problems*. Prentice Hall, Englewood Cliffs, NJ.

Lees, J. (1995). Reshaping spectrum estimates by removing periodic noise: Application to seismic spectral ratios. *Geophysics Research Letters*, **22**(4), 513–516.

Lees, J. and Park, J. (1995). Multiple-taper spectral analysis: a stand-alone C-subroutine. *Compt. Geosci.*, **21**(2), 199–236.

Macdonald, K. C., Fox, P., Miller, S., Carbotte, S., Edwards, M., Eisen, M., Fornari, D., Perram, L., Pockalny, R., Scheirer, D., Tighe, S., Weiland, C., and Wilson, D. (1992). The East Pacific Rise and its flanks 8-18N: History of segmentation, propagation and spreading direction based on SeaMARC II and SeaBeam Studies. *Marine Geophys. Researches*, **14**, 299–344.

Madariaga, R. (1976). Dynamics of an expanding circular fault. *Bull. Seism. Soc. Am.*, **66**, 639–666.

Mason, J. and Handscomb, D. (2003). *Chebyshev Polynomials*. Chapman & Hall/CRC Press, Boca Raton, FL.

Mayeda, K. and Walter, W. R. (1996). Moment, energy, stress drop, and source spectra of Western United States earthquakes from regional coda envelopes. *J. Geophys. Res.*, **101**, 11195–11208.

Mayeda, K., Dreger, D. S., Walter, W. R., and Tajima, F. (2004). Bdsn calibration for northern california earthquakes from coda-derived source spectra: Moment magnitude and radiated energy (abstract). *Seism. Res. Let.*, **75**, 278.

McGarr, A. (1999). On relating apparent stress to the stress causing earthquake fault slip. *J. Geophys. Res.*, **104**, 3001–3003.

McGuire, J. J. (2004). Estimating the finite source properties of small earthquake ruptures. *Bull. Seism. Soc. Am.*, **94**, 377–393.

Melosh, J. (1979). Acoustic fluidization: a new geologic process? *J. Geophys. Res.*, **84**, 7512–7520.

Miller, R. G. (1974). The jackknife – a review. *Biometrika*, **61**, 1–15.

Mori, J., Kanamori, H., and Abercrombie, R. E. (2003). Stress drops and radiated energies of aftershocks of the 1994 Northridge, California, earthquake. *J. Geophys. Res.*, **108**(B11 2545).

Mueller, C. S. (1985). Source pulse enhancement by deconvolution of an empirical Green's function. *J. Geophys. Res.*, **12**(1), 33–36.

Palmer, A. C. and Rice, J. R. (1973). The growth of slip surfaces in the progressive failure of overconsolidated clay. *Proc. R. Soc. Lond., A*, **332**, 527–548.

Park, J. and Levin, V. (2000). Receiver functions from multiple-taper spectral correlation estimates. *Bull. Seism. Soc. Am.*, **90**, 1507–1520.

Park, J., Lindberg, C. R., and Thomson, D. J. (1987a). Multiple-taper spectral analysis of terrestial free oscillations: part I. *Geophys. J. Royal Astr. Soc.*, **91**, 755–794.

Park, J., Lindberg, C. R., and Vernon, F. L. (1987b). Multitaper spectral analysis of high-frequency seismograms. *J. Geophys. Res.*, **92**, 12675–12684.

Park, J. J. (1992). Envelope estimation for quasi-periodic geophysical signals in noise: a multitaper approach. In A. T. Walden and P. Guttorp, editors, *Statistics in the Enviroment & Earth Sciences*, pages 189–219. Edward Arnold, London.

Percival, D. B. and Walden, A. T. (1993). *Spectral analysis for physical applications: Multitaper and Conventional univariate techniques*. Cambridge Univ. Press.

Pérez-Campos, X. and Beroza, G. C. (2001). An apparent mechanism dependence of radiated seismic energy. *J. Geophys. Res.*, **106**(B6), 11127–11136.

Pérez-Campos, X., Singh, S. K., and Beroza, G. C. (2003). Reconciling teleseismic and regional estimates of seismic energy. *Bull. Seism. Soc. Am.*, **93**(2), 2123–2130.

Prejean, S. G. and Ellsworth, W. L. (2001). Observations of earthquake source parameters and attenuation at 2 km depth in the Long Valley Caldera, eastern California. *Bull. Seism. Soc. Am.*, **95**, 165–177.

Prieto, G. A., Shearer, P. M., Vernon, F. L., and Kilb, D. (2004). Earthquake source scaling and self-similarity estimation from stacking P and S spectra. *J. Geophys. Res.*, **109**(B08310).

Prieto, G. A., Parker, R. L., Vernon, F. L., Shearer, P. M., and Thomson, D. J. (2006). Uncertainties in earthquake source spectrum estimation using empirical Green functions. In Abercrombie, McGarr, Kanamori, and diToro, editors, *Earthquakes: Radiated Energy and the Physics of Faulting*, pages 69–74. AGU Geophys. Monograph 170.

Prieto, G. A., Thomson, D. J., Vernon, F. L., Shearer, P. M., and Parker, R. L. (2007a). Confidence intervals for earthquake source parameters. *Geophys. J. Int.*, **168**, 1227–1234.

Prieto, G. A., Parker, R. L., Thomson, D. J., Vernon, F. L., and Graham, R. L. (2007b). Quadratic multitaper spectrum. Submitted to *Geophys. J. Int.*

Quenouille, M. (1949). Approximate tests of correlation in time series. *J. R. Stat. Soc.*, **B11**, 18–84.

Rice, J. A. (1995). *Mathematical Statistics and Data Analysis*. Duxbury Press, Belmont, CA, 2nd. edition.

Riedel, K. S. and Sidorenko, A. (1995). Minimum bias multiple taper spectral estimation. *IEEE Trans. on Signal Processing*, **43**, 188–195.

Riedel, K. S., Sidorenko, A., and Thomson, D. J. (1993). Spectral estimation of plasma fluctuations i. comparison of methods. *Phys. Plasma*, **1**(3), 485–500.

Scott, J. S., Masters, T. G., and Vernon, F. L. (1994). 3-D velocity structure of the San Jacinto fault zone near Anza, California—I. P waves. *Geophys. J. Int.*, **119**, 611–626.

Shearer, P. M. (1999). *Introduction to Seismology*. Cambridge University Press, Cambridge.

Shearer, P. M., Hauksson, E., and Lin, G. (2005). Southern California hypocenter relocation with waveform cross-correlation, Part 2: results using sourcespecific station terms and cluster analysis. *Bull. Seism. Soc. Am.*, **95**, 904–915.

Shearer, P. M., Prieto, G. A., and Hauksson, E. (2006). Comprehensive analysis of earthquake source spectra in southern California. *J. Geophys. Res.*, **111**(B06303).

Sibson, R. (1973). Interactions between temperature and pore-fluid pressure during earthquake faulting and a mechanism for partial or total stress relief. *Nature*, **243**, 66–68.

Singh, S. K. and Ordaz, M. (1994). Seismic energy release in mexican subduction zone earthquakes. *Bull. Seism. Soc. Am.*, **84**, 1533–1550.

Slepian, D. (1978). Prolate spheroidal wavefunctions, Fourier analysis, and uncertainty V: the discrete case. *Bell System Tech. J.*, **57**, 1371–1429.

Sonley, E. and Abercrombie, R. E. (2006). Variability introduced by methods of attenuation correction: the effect on source parameter determination. In Abercrombie, McGarr, Kanamori, and diToro, editors, *Earthquakes: Radiated Energy and the Physics of Faulting*. AGU Geophys. Monograph 170.

Taylor, J. R. (1997). *An introduction to Error Analysis*. University Science Books, Sausalito, CA.

Thomson, D. J. (1982). Spectrum estimation and harmonic analysis. In *Proceedings of the IEEE*, volume 70, pages 1055–1096.

Thomson, D. J. (1990). Quadratic-Inverse spectrum estimates: applications to paleoclimatology. *Phys. Trans. R. Soc. London A*, **332**, 539–597.

Thomson, D. J. (1994). An overview of multiple-window and quadratic-inverse spectrum estimation methods. In *ICASSP - 94*, volume VI, pages 185–194.

Thomson, D. J. and Chave, A. D. (1991). Jackknife error estimates for spectra, coherences, and transfer functions. In S. Haykin, editor, *Advances in Spectrum Analysis and Array Processing*, volume 1 chapter 2, pages 58–113. Prentice Hall.

Tinti, E., Spudich, P., and Cocco, M. (2005). Earthquake fracture energy inferred from kinematic rupture models on extended faults. *J. Geophys. Res.*, **110**, B12303.

Tsuda, K., Archuleta, R. J., and Koketsu, K. (2006). Quantifying the spatial distribution of site response by use of the Yokohama High-Density Strong-Motion Network. *Bull. Seism. Soc. Am.*, **96**(3), 926–942.

Tukey, J. W. (1958). Bias and confidence in not-quite large samples. *Ann. Math. Stat.*, **29**, 614.

Tukey, J. W. (1960). Conclusions vs decisions. *Technometrics*, **2**(4), 423–433.

Venkataraman, A. and Kanamori, H. (2004). Effect of directivity on estimates of radiated seismic energy. *J. Geophys. Res.*, **109**(B04301).

Venkataraman, A., Kanamori, H., and Rivera, L. (2002). Radiated energy from the 16 October 1999 Hector Mine Earthquake: Regional and Teleseismic Estimates. *Bull. Seism. Soc. Am.*, **92**, 1256–1265.

Venkataraman, A., Boatwright, J., and Beroza, G. C. (2006). A brief review of techniques used to estimate radiated seismic energy. In Abercrombie, McGarr, Kanamori, and diToro, editors, *Earthquakes: Radiated Energy and the Physics of Faulting*, pages –. AGU Geophys. Monograph 170.

Vernon, F. L. (1989). *Analysis of data recorded on the ANZA seismic network*. Ph.D. thesis, University of California, San Diego.

Vossiliou, M. S. and Kanamori, H. (1982). The energy release in earthquakes. *Bull. Seism. Soc. Am.*, **72**, 371–387.

Warren, L. M. and Shearer, P. M. (2000). Investigating the frequency dependance of mantle Q by stacking P and PP spectra. *J. Geophys. Res.*, **105**, 25391–25402.

Warren, L. M. and Shearer, P. M. (2002). Mapping lateral variations in upper mantle attenuation by stacking P and PP spectra. *J. Geophys. Res.*, **107**(B12 2342).

Wössner, J., Treml, M., and Wenzel, F. (2002). Simulation of $M_W = 6.0$ earthquakes in the Upper Rhinegraben using empirical Green functions. *Geophys. J. Int.*, **151**(2), 487–500.

Wu, C. F. J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis (with discussion). *Ann. Stat.*, **14**, 1261–1350.

Wyss, M. and Brune, J. N. (1968). Seismic moment, stress, and source dimensions for earthquakes in the california- nevada region. *J. Geophys. Res.*, **73**, 4681–4694.