The Pennsylvania State University
The Graduate School
Eberly College of Science

GRAVITATIONAL WAVE ASTRONOMY WITH LIGO: FROM DATA TO SCIENCE

A Thesis in
Physics
by
Tiffany Z. Summerscales

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2006

The thesis of Tiffany Z. Summerscales was reviewed and approved* by the following:

Lee Samuel Finn
Professor of Physics
Thesis Advisor, Chair of Committee

Abhay V. Ashtekar
Professor of Physics

Pablo Laguna
Professor of Astronomy and Astrophysics

James J. Whitmore
Professor of Physics

Jayanth R. Banavar
Professor of Physics
Chairman, Department of Physics

*Signatures are on file in the Graduate School.

Abstract

This thesis describes three LIGO data analysis projects. The first project is development of a procedure that whitens LIGO data (white data have the same power at all frequencies) and breaks it into several non-overlapping frequency bands. The data can then be used by gravitational wave search algorithms that assume a white noise background. Breaking the data into frequency bands not only simplifies the whitening process but allows a rough frequency spectrum to be assigned to candidate events. A whitening figure of merit is also described and, in the case of data from LIGO's third science run, it is shown that the whitening procedure nearly always improves the whiteness of the data.

The second project described in this thesis is development of a computationally inexpensive test that can be run quickly over LIGO data to flag times where data have been corrupted by a nonlinear coupling. The test is applied to LIGO data and is shown to flag segments whose bispectra contains similar features to the bispectra of data produced with a nonlinear model.

Finally, the third project seeks to address two problems that one would confront if one tried to do core-collapse supernova astronomy with gravitational waves. The first problem involves extracting a short-duration gravitational waveform from the data produced by a network of detectors. The maximum entropy method is proposed as a solution to this deconvolution problem. The second problem involves deducing properties of the source from the recovered waveform when our source models are incomplete. We propose calculating the cross correlation between a recovered waveform and a catalog of waveforms associated with models having varying properties. The catalog waveform having the highest cross correlation with the recovered waveform is assumed to be associated with a model whose properties most closely resemble those of the source. The maximum entropy method is used to recover supernova waveforms from simulated LIGO data which are created assuming detector responses and white noise having amplitudes

typical of recent LIGO science runs. Next, the recovered waveform is cross correlated with a catalog of waveforms and it is shown that the recovered waveform carries information about the type of bounce the core undergoes as well as the progenitor mass, angular momentum and degree of differential rotation for supernova occurring less than a few kpc away. Supernova waveforms are also recovered using maximum entropy from simulated data using actual LIGO data for noise and from hardware injections. Recovering signals from these data show that maximum entropy can successfully handle colored noise and imperfect knowledge of the LIGO detector responses.

Table of Contents

List of Figures

List of Tables

Preface

The work described in this thesis was carried out with the assistance of many people.

The data conditioning work described in Chapter 2 was aided by Mike Ashley, Bethany Bongiorno, Eric Rotthoff, Matthew Tibbits, Daniel Tomcavage, and Kristina Zaleski who helped measure some of the line properties and filter parameters necessary to remove lines. Keith Thorne later automated the finding of lines and their parameters and assisted in debugging the Kalman filtering code. I carried out all other aspects regarding development and testing of the data conditioning pipeline.

The supernova study in Chapter 4 was performed in collaboration with Adam Burrows and Christian Ott who supplied the gravitational waveforms produced by their core-collapse supernova models as well as excellent advice on the direction of the project. Initial development of the maximum entropy code was carried out by Lee Samuel Finn. I assisted with later development of the maximum entropy code and carried out all of the tests described in this chapter.

Acknowledgments

There are many people who deserve great thanks in helping me along the way from starting to finally finishing this thesis.

Thanks first of all to Gabriella for getting me started with LIGO and grad school and to Sam for taking over and seeing me through to the end. Thanks to the postdocs: Patrick, Mike, John, Keith, and Shantanu for all their help and good advice. Thanks to Richard Tutwiler for introducing me to the world of deconvolution and to Adam Burrows and Christian Ott for all of their help with the supernova project.

Thanks to Amber for help in preservation of sanity by always being there and always understanding.

A big thanks to family for love and support, especially to my parents, Rodney's parents and Sis and Morris, our virtual parents. Biggest thanks to Rodney, my best friend and hubby extrordinare.

Of course those mentioned above are only a small percentage of all those who make the past seven years go so quickly and so well. I have learned lots. Thanks to all.

> The heavens declare the glory of God; the skies proclaim the work of his hands. Ps 19:1 (NIV)

Chapter 1

Introduction

1.1 Gravitational Waves

In 1916 Einstein showed that one of the predictions of his theory of general relativity is the existence of gravitational waves. General relativity states that spacetime curvature is determined by the distribution of matter and energy. As that distribution changes, the curvature must also change and that curvature change travels at the speed of light as ripples in spacetime. As they travel, gravitational waves stretch and squeeze space in a direction transverse to their propagation. The magnitude of these disturbances is specified by the unitless quantity strain (h) which is the change in distance between two free masses divided by twice the distance between them or $h = \Delta L/2L$.

Gravitational waves are best produced by matter distributions with a large variation in quadrupole moment. In cases where matter motion is slow compared with the speed of light, the strongest possible component of strain is given by

$$h_{\mu\nu} = \frac{2G}{Rc^4}\frac{d^2\left[I_{\mu\nu}(t - R/c)\right]}{dt^2}, \tag{1.1}$$

where G is the gravitational constant, c the speed of light, R the distance to the source, and $I_{\mu\nu}$ the reduced quadrupole moment. One can see from this expression that the strain produced by gravitational waves is extremely small. Only the most dense, massive objects and most extreme events in the universe such as the big bang, supernovae, and the coalescence of black holes and neutron stars produce

gravitational waves of significant strength. Even so, the gravitational waves from these events are expected to produce strains at the Earth of no more than on the order of 10^{-21}.

Gravitational waves may be weak but their influence on the systems that emit them has already been measured. In 1974 Hulse and Taylor discovered the pulsar PSR 1913+16 whose periodically varying pulsation rate indicated that it and a companion neutron star were orbiting about their center of mass [6]. Subsequent measurements of this system have shown that the orbit is decaying and the orbital frequency increasing in exact agreement with predictions of how much energy should be lost through gravitational waves [7, 8].

PSR 1913+16 provides strong evidence that gravitational waves exist, but there are still many incentives for attempting to measure gravitational waves themselves. Gravitational waves can provide novel tests of general relativity. For example, general relativity predicts that gravitational waves are transverse quadrupolar, having only two polarizations (see Figure 1.1) whereas other theories of gravity predict more. Gravitational waves also have the potential to open a new window on the universe. The objects and events that produce them are sources of some very perplexing questions. Gravitational waves from supernovae and rotating, asymmetric neutron stars may contain valuable information about neutron star material. Gravitational waves produced by black holes could give information about these most mysterious objects whose presence can now only be inferred. In short, the measurement of gravitational waves could be an invaluable new tool for astronomy.

The attempt to measure gravitational waves directly was pioneered by Joseph Weber in the 1960s with the development of his "Weber bars" [9]. These were cylindrical bars of aluminum with quartz strain gauges that could measure vibrations caused by a gravitational wave passing through the bars at their resonant frequencies. Gravitational wave resonant bar detectors are still in use in a worldwide network that includes the detectors ALLEGRO [10], AURIGA [11], EXPLORER [12], NAUTILUS [13], and NIOBE [14]. These detectors are operated at very low temperatures and incorporate other improvements over the first bars, including sophisticated vibration isolation systems and using resonant transducers to measure vibrations. In the future, it is likely that bars will change their shape

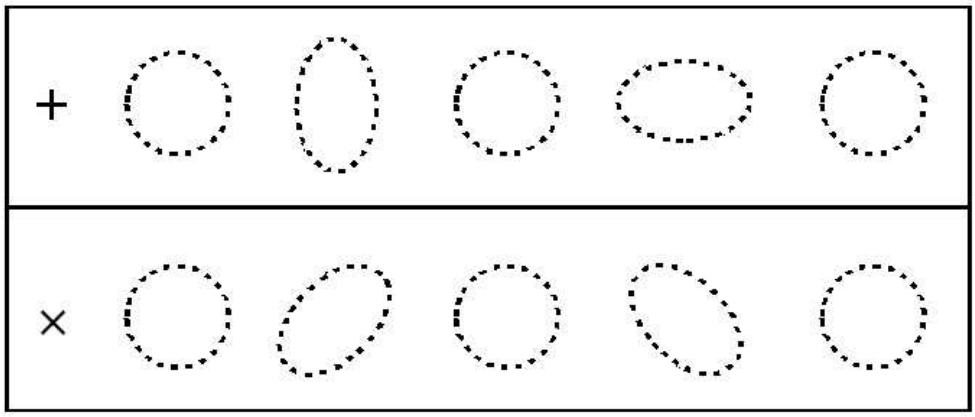


Figure 1.1. Gravitational wave polarizations. A set of free masses arranged in a circle would follow the deformations of the circles shown above as a gravitational wave passed through them. The top row shows the successive deformations that would be caused by a gravitational wave with plus polarization traveling out of the page. The second row shows the deformations caused by the cross polarization which are rotated 45 degrees from the plus.

to spheres which are sensitive to gravitational waves coming from all directions and will be sensitive to much wider frequency ranges as transducer designs are improved [15].

1.2 Gravitational Wave Interferometers

Within the past few years, several large detectors have been built that are based on an interferometric approach. These include the three LIGO detectors [16], two in Hanford, Washington and one in Livingston, Louisiana, GEO 600 [17] near Hannover, Germany, TAMA300 [18] in Tokyo, Japan, and VIRGO [19] in Cascina, Italy.

In a gravitational wave interferometer, a laser beam is separated into two beams by a beamsplitting mirror and sent down the perpendicular arms of the detector (see Figure 1.2). At the end of each arm, the beam is reflected by a mirror and sent back to interfere at the beamsplitter. The amount of interference is measured at

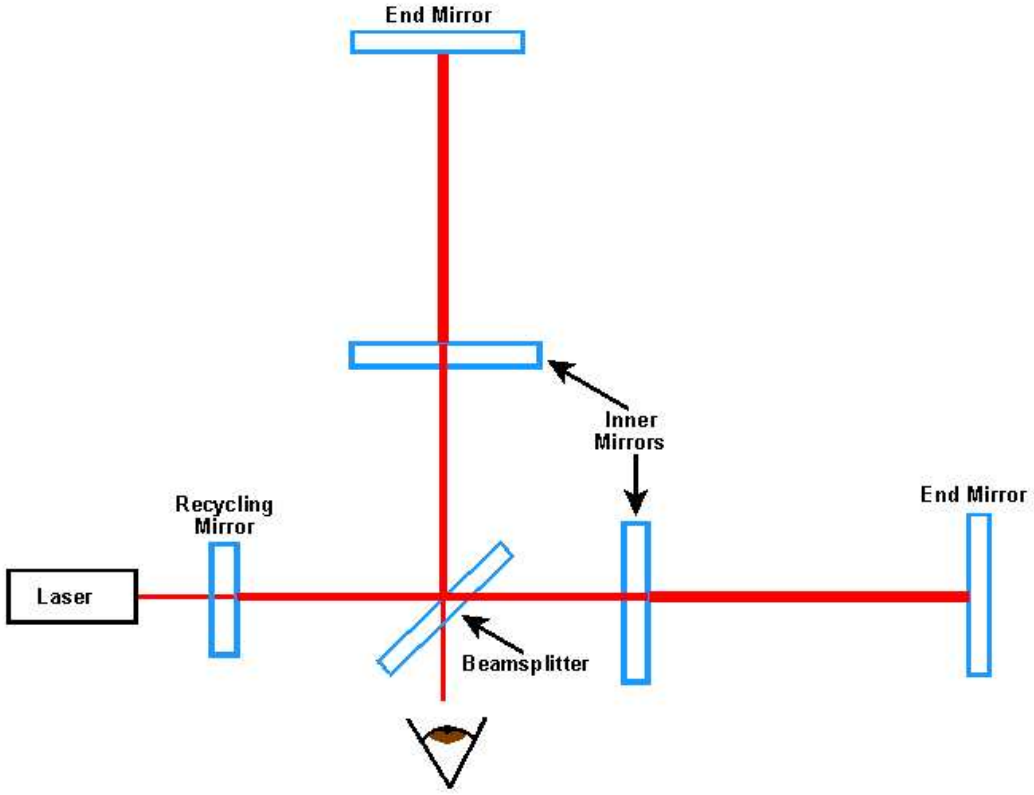


Figure 1.2. Basic layout of an interferometric gravitational wave detector like LIGO.

the detector's output port. A passing gravitational wave will change the distances between mirrors (which are hung as pendula and can be treated as free masses above their pendulum frequencies) and hence a change in the amount of light reaching the output port.

In order to increase the amount of interaction time between the wavefronts traveling down the detector arms and the gravitational waves and therefore increase the effective length of the arms, additional mirrors are placed after the beamsplitter to make the arms into Fabry-Perot cavities. This approach is used with LIGO, TAMA300 and VIRGO. GEO 600 uses a different approach of folding the light path within its arms. In addition, all detectors use a mirror between the laser and beam splitter to reflect light headed toward the laser back into the interferometer and thus increase the effective power of the laser in a technique known as power

recycling. GEO 600 also employs signal recycling, where a mirror is placed in front of the output port, increasing sensitivity at a signal frequency.

1.3 LIGO Science

The operation of LIGO and extraction of scientific results from LIGO data is carried out by two groups: the LIGO Lab run by the California Institute of Technology (Caltech) and the Massachusetts Institute of Technology (MIT) and by the LIGO Scientific Collaboration (LSC) whose members include both LIGO Lab members and scientists from other institutions (the GEO Collaboration are all members of the LSC). The LIGO Lab is in charge of the administration of the observatories while the LSC carries out the scientific activities associated with LIGO. These activities are split between several LSC working groups and each LSC member is usually a member of more than one working group. Some of the working groups carry out investigations into new technologies to be incorporated into future versions of the detectors. The Detector Characterization Working Group evaluates the performance of the detectors, tracks down noise sources, provides calibration information, and works on improving current detector sensitivity. The Data Analysis Working Group is in charge of computing issues. Finally, there are four upper limit groups that carry out searches for gravitational waves produced by four different types of gravitational wave sources. The Stochastic Upper Limit Group searches for gravitational wave backgrounds caused by either the Big Bang or sources like binary star systems which could be so numerous that signals from individual sources blend together and cannot be distinguished. The Pulsar Upper Limit Group searches for single frequency, continuous gravitational waves from spinning neutron stars. The Inspiral Upper Limit Group looks for neutron star-neutron star, neutron star-black hole and black hole-black hole inspirals. Finally, the Burst Upper Limit Group searches for any bursts of gravitational waves less than a second in duration. These bursts could be produced by supernovae, the coalescence of objects at the end of inspiral or any other source of short-duration gravitational waves.

At the time of writing, LIGO, GEO 600 and TAMA300 have taken data used in scientific analysis. (VIRGO will soon complete the final stages of commissioning.)

Run	Start Date	End Date
S1	23 Aug 2002	09 Sep 2002
S2	14 Feb 2003	14 Apr 2003
S3	31 Oct 2003	09 Jan 2004
S4	22 Feb 2005	23 Mar 2005

Table 1.1. Start and end dates for the four LIGO science runs (S1-S4) when data were taken for scientific analysis purposes.

While no gravitational waves have yet been measured, these studies have set upper limits on the rates of gravitational waves of different strains for different sources [20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34]. In particular, LIGO has undergone four science runs (S1, S2, S3, and S4) interspersed with commissioning work (see Table 1.1). With each science run, there was a dramatic improvement in the sensitivity and stability of the instruments [2, 3, 4, 5].

LIGO currently is embarking on a year-long run of data taking. Now that the instruments are near their design sensitivities, the upper limit groups are transitioning their focus from setting upper limits to making detections of gravitational waves. Hopefully, the direct measurement of gravitational waves is just around the corner and we are on the edge of the new era of gravitational wave astronomy.

1.4 Thesis Outline

This thesis is concerned with three aspects of the problem of gravitational wave detection: "conditioning" the data as a step in the detection of burst gravitational waves, characterization and diagnosis of detector performance based on the data and devising and implementing a method for recovering gravitational waveforms from detector data as well as relating these observations to the astrophysics of the source.

The first project forms a part of the search for unmodeled gravitational wave bursts. Searches for unmodeled bursts are very difficult to conduct. In other types of searches such as those for inspiral, periodic, and stochastic gravitational waves, one is looking for a signal of known character. In the case of searching for bursts, however, one has no prior information and is simply looking for something in the

data that is not noise. All untriggered burst searches involve identifying times when the data are inconsistent with the general statistics of the detector noise. A first step in these searches is to characterize the detector noise statistics and remove all of the correlations and instrumental artifacts identified, also known as conditioning the data.

Chapter 2 describes a data conditioning procedure that was developed to remove instrumental and environmental artifacts from the data, break the data into frequency bands and remove correlations from, or whiten, each band. (White data contain equal power at all frequencies.) Data conditioning is essential for gravitational wave search algorithms that assume that only white noise accompanies the sought-after signals and can provide a means of identifying a rough spectrum of events.

The second project involves characterization of the LIGO detectors' performance. These detectors are the first of their kind in terms of both size and complexity. Consequently, there is no prior experience to be drawn from when dealing with these instruments. Evaluating performance, identifying problems and making improvements are areas of intense activity by the Detector Characterization Working Group and are essential for the scientific success of LIGO.

One detector characterization issue that is of concern is the possible presence of bilinear couplings which are the modulation of one noise source by another. When this happens a noise source that is at frequencies outside the band of frequencies at which LIGO is most sensitive to gravitational waves may be converted to in-band frequencies and decrease sensitivity. For instance, seismic noise at low frequencies (< 10 Hz) could couple to resonances of the wires that suspend the mirrors and thus be converted to noise about these resonance frequencies at a few hundred Hz, where LIGO is most sensitive to gravitational waves. These nonlinear couplings are difficult to find since they are invisible to second-order measures like the power spectrum. Higher-order spectra are sensitive to these couplings but are very computationally expensive to calculate if one is trying to detect these couplings in large amounts of data. Chapter 3 introduces a computationally inexpensive test for identifying sections of data that have been influenced by nonlinear couplings.

Finally, the third project pioneered development of an analysis capable of retrieving astrophysical information borne by gravitational waves from LIGO data.

Core-collapse supernovae were used as the prospective source in this study because of the potential for astrophysical discovery through the gravitational waves they produce. Details of the core-collapse cannot be obtained from electromagnetic observation. However, gravitational waves which can travel freely from the center of the collapsing star may be able to carry this information. In order to discover what the gravitational wave signature tells us about a supernovae we must be able to recover that signature from the data and then compare it with waveforms associated with theoretical models of supernovae.

Chapter 4 presents maximum entropy (a regularization approach with a significant history in the field of astronomical image reconstruction) as a solution to the problem of retrieving gravitational waveforms from data produced by networks of two or more detectors. The second problem of identifying source properties is addressed by using cross correlation between a recovered signal and a catalog of waveforms from supernova simulations to see what waveform the recovered signal most resembles. That waveform is associated with a simulation with properties that likely mimic those of the source. It was found that the recovered signal carried information about source properties such as progenitor mass, angular momentum and angular momentum distribution as well as bounce type. This result is particularly exciting as it is the first demonstration that the promise of gravitational wave astronomy - the use of gravitational wave observations as a tool of astronomical discovery - can be realized.

Chapter 2

Data Conditioning

2.1 Introduction

In the search for gravitational waves, we do not want to be restricted to those solely from well modeled sources like neutron star inspirals. We also want to keep our eyes open to the unexpected and to sources that are not as well modeled, such as core-collapse supernovae and black hole mergers. The Burst Upper Limit Group has been charged with this difficult task of looking for gravitational waves of unknown character. The only strategy capable of finding any short-duration signal that may be present is to search for places where the detector data differ in a statistical fashion from the noise.

The LIGO data analysis group at Penn State, which participates in the search for unmodeled bursts, has developed an algorithm for finding sections of data that have different statistics from the noise. The algorithm is called BlockNormal [35]. BlockNormal assumes that the data are normally distributed and breaks the data into blocks of time where the data can be described by a single mean and variance. If the mean or variance of a block of data exceeds a certain threshold, those data are flagged as possibly containing a gravitational wave event.

It is assumed that the noise that BlockNormal is characterizing has no correlations, or in other words is white. White data are data that have equal power at all frequencies. However, the data produced by the LIGO interferometers are far from white (see Figure 2.1). There are many sharp features in frequency (lines) due to instrumental resonances. For example, the wires suspending each of the

mirrors vibrate at their resonant frequencies (also known as violin modes since the wires are vibrating like the strings of a violin). Consequently, the mirrors are moved at these violin modes, changing the arm lengths and injecting power at these frequencies into the data. There are features due to the 60Hz oscillations of the U.S. power grid and there are mirror oscillations that are intentionally injected for calibration purposes. In addition, the general shape of the noise in frequency space is dominated by seismic noise which increases with low frequency and an increase in noise with high frequency that is related to shot noise (caused by the discrete nature of photons).

In order to remove these instrumental and environmental artifacts and produce whitened data suitable for BlockNormal, we have designed a data conditioning procedure. This procedure not only removes lines and whitens the data but breaks the data into frequency bands. Breaking the data into frequency bands simplifies the whitening process by reducing the number of features that must be dealt with at any one time. It also means that when BlockNormal identifies a possible event, we know the range of frequencies to assign to the event based on the band it was found in.

The data conditioning procedure is described in detail in the following sections along with the ordering necessary to simplify successive steps. The procedure was applied in its entirety to LIGO data from both the second and third science runs (S2 and S3). A shortened version that excluded line removal was applied to the data from the fourth science run (S4). The conditioned LIGO data from these science runs was analyzed by BlockNormal.

2.2 Downsampling

LIGO data are sampled at 16384 Hz, resulting in very large amounts of data that are computationally expensive to whiten. Therefore, the first step in our data conditioning procedure is to downsample the data. Downsampling reduces the sampling rate (number of data samples per second) by a constant factor. The downsampling applied to the S2-S4 data reduced the sample rate from 16384Hz to 4096Hz. This significantly reduces the size of the data which can greatly reduce the computational cost of the next, Kalman filtering step as well as simplify

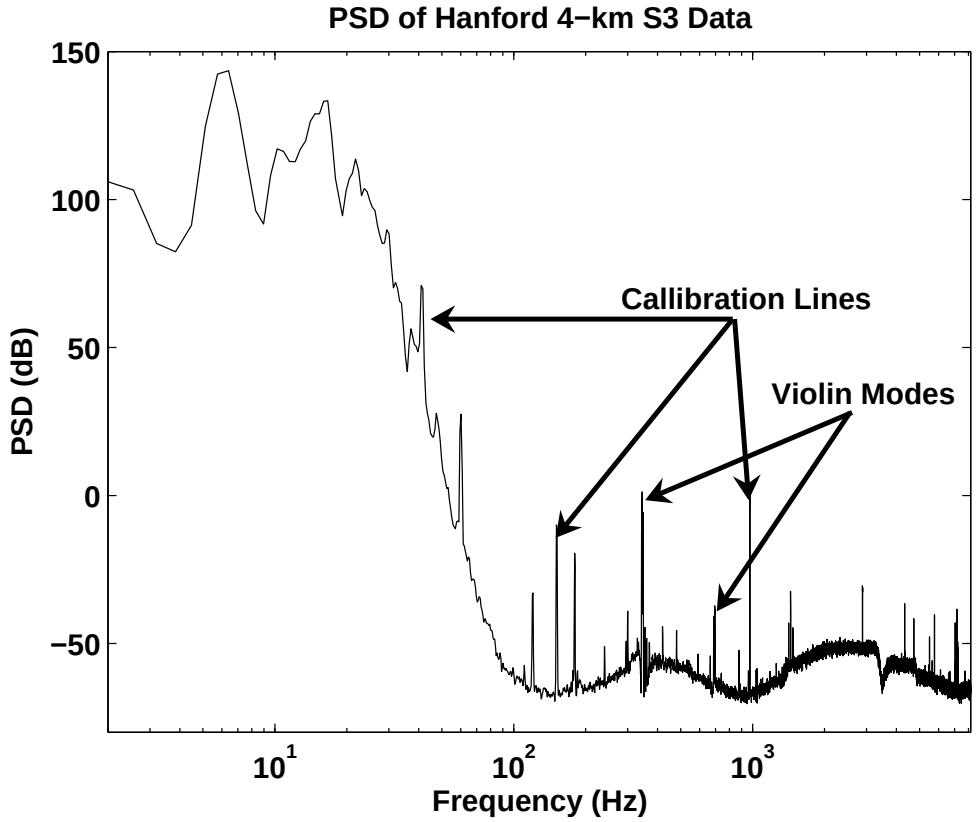


Figure 2.1. Power Spectral Density (PSD) of data from the Hanford, 4-km detector taken during the third science run (S3). There is a large increase in power at low frequencies caused by seismic noise as well as many narrow frequency features. Some of the narrow "lines" are caused by violin modes, which are mirror suspension wire resonances and their harmonics. Some lines are the result of the mirrors being moved at certain frequencies for calibration purposes. Most of the remaining lines are power lines which are harmonics of the 60 Hz oscillations of the U.S. power grid.

the highpass shaping filter described in Section 2.4. Since discretely sampled data contains information about frequencies up to half the sample rate (half the sample rate is called the Nyquist frequency), downsampling will throw away some high frequency information. We therefore select the highest frequency that is of interest (2048 Hz) and downsample the data to set the new Nyquist to that highest frequency. Discarding information at very high frequencies does not reduce our chances of seeing gravitational waves much because the detectors are less sensitive at these frequencies [2, 3, 4, 5].

The process of downsampling involves more than throwing away data points

to achieve a smaller number of samples. First, a lowpass filter that suppresses the frequencies above the new Nyquist frequency must be applied to prevent aliasing. Aliasing can occur when a high frequency component of a signal is under-sampled. When this happens, there are not enough samples to distinguish between a signal at a frequency less than or equal to Nyquist and one that is above Nyquist so that the power at above-Nyquist frequencies is erroneously attributed to lower frequencies (see Figure 2.2). Once the above-Nyquist frequencies have been suppressed, it is safe to discard samples or interpolate between them to achieve the new sample rate [36].

2.3 Line removal with a Kalman filter

For the next step in our process, we wish to remove some of the narrow frequency features (lines) present in the data. Some lines are caused by processes whose effects are measured by environmental or other monitors. For example, the 60 Hz power oscillations that cause the power lines at multiples of 60 Hz in the data are measured by voltmeters and magnetometers positioned around the detectors. Removal of these lines, based on these measurements, is discussed in Section 2.6. The majority of lines, however, are caused by processes that cannot be directly measured and may not even be understood. To remove these lines we use a Kalman filter which gives us a way of modeling the process creating each line.

The Kalman filter was originally developed in 1960 by R. E. Kalman [37]. The filter is a set of equations that allow us to use a simple model of the process to predict what part of the data is due to the line [38, 39, 40]. It is assumed that the part of the data to be captured by the filter is caused by a system which is not directly observed. (We assume that the system causing each line can be approximated as a weakly damped harmonic oscillator that is excited by random noise.) At each time step k the system can be described by a state vector $\boldsymbol{\psi}[k]$. (In the model of the harmonic oscillator, the state vector could describe the position and momentum of the oscillator.) The state evolves from one time step to the next via

$$\boldsymbol{\psi}[k] = \boldsymbol{A} \cdot \boldsymbol{\psi}[k-1] + \boldsymbol{w}[k], \tag{2.1}$$

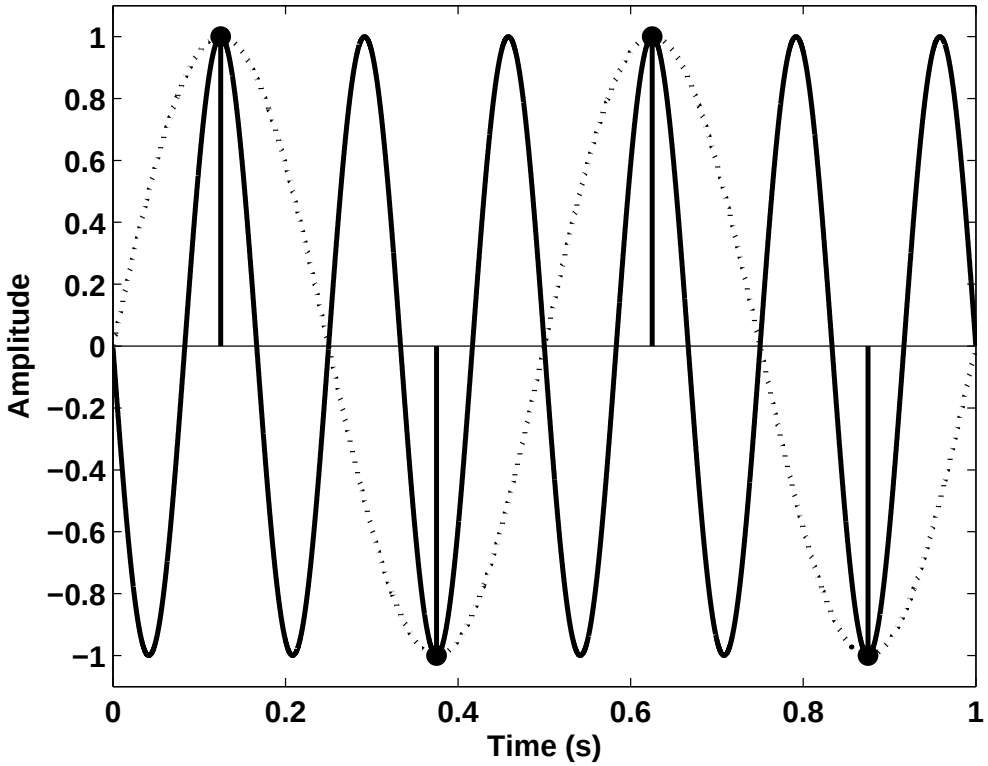


Figure 2.2. Aliasing example. Suppose a signal is sampled at a rate of 4 samples per second (black dots in the figure above). The sampled data only contain information on frequencies up to the Nyquist frequency which in this example is 2 Hz. If the signal being sampled is a 6 Hz sine wave, as shown above, the samples will have exactly the same values as if the signal were a 2 Hz sine wave. The power in the 6 Hz sine will be aliased or mistakenly attributed to the 2 Hz frequency since there is no way to distinguish the two frequencies once the sampling has taken place.

where $\boldsymbol{A}$ is the evolution operator matrix determined by the nature of the system and $\boldsymbol{w}[k]$ is the white noise that is exciting the system. This noise is also called the process noise.

The state of the system is unknown, but affects the output of the detector which is observed. The observable (data) at time step (sample) k is expressed as the vector $\boldsymbol{z}[k]$ The dependence of the observable on the state can be expressed as:

$$\boldsymbol{z}[k] = \boldsymbol{C} \cdot \boldsymbol{\psi}[k] + \boldsymbol{v}[k], \tag{2.2}$$

where $\boldsymbol{C}$ is a linear transformation called the measurement operator and $\boldsymbol{v}[k]$ is an additional, white noise background. This noise background is known as the measurement noise, which in the case of LIGO data is primarily caused by seismic noise at lower frequencies and laser shot noise at higher frequencies. Approximating the measurement noise background $\boldsymbol{v}[k]$ as white would not be valid if the observable $\boldsymbol{z}[k]$ was the data as produced by the detectors. (LIGO data minus a line are far from white.) However, if we take as our observable the data restricted to a small region in frequency about the line, as described in Section 2.3.1, the background is approximately white. The $\boldsymbol{C} \cdot \boldsymbol{\psi}[k]$ term of Equation 2.2 is the part of the observable due to the line and is the part that we wish to isolate and remove.

The process and measurement noise, $\boldsymbol{w}[k]$ and $\boldsymbol{v}[k]$, are both white and zero-mean and are therefore completely characterized by their covariance matrices

$$\boldsymbol{W} := E[\boldsymbol{w} \cdot \boldsymbol{w}^T] \tag{2.3}$$

$$\boldsymbol{V} := E[\boldsymbol{v} \cdot \boldsymbol{v}^T], \tag{2.4}$$

where $E[\cdot]$ denotes the expectation operator.

Suppose we have an estimate $\hat{\boldsymbol{\psi}}[k-1]$ of the system's state at the sample $k-1$. The error in this estimate is

$$\hat{\boldsymbol{e}}[k-1] := \boldsymbol{\psi}[k-1] - \hat{\boldsymbol{\psi}}[k-1]. \tag{2.5}$$

The variance of the error is

$$\hat{\boldsymbol{P}}[k-1] := E\left[\hat{\boldsymbol{e}}[k-1] \cdot \hat{\boldsymbol{e}}^T[k-1]\right]. \tag{2.6}$$

The state estimate can then be evolved to the next time sample k by using the evolution operator

$$\tilde{\boldsymbol{\psi}}[k] := \boldsymbol{A} \cdot \hat{\boldsymbol{\psi}}[k-1]. \tag{2.7}$$

Because we started with an estimate of the state that had some error, there will be error in $\tilde{\boldsymbol{\psi}}[k]$ due to the error in the previous state estimate and due to process

noise. The total error in $\tilde{\boldsymbol{\psi}}[k]$ is defined to be

$$\tilde{\boldsymbol{e}}[k] := \boldsymbol{\psi}[k] - \tilde{\boldsymbol{\psi}}[k], \tag{2.8}$$

with a second moment of the error

$$\begin{aligned} \tilde{\boldsymbol{P}}[k] &:= E\left[\tilde{\boldsymbol{e}}[k] \cdot \tilde{\boldsymbol{e}}^T[k]\right] \\ &= \boldsymbol{A} \cdot \hat{\boldsymbol{P}}[k-1] \cdot \boldsymbol{A}^T + \boldsymbol{W}. \end{aligned} \tag{2.9}$$

Now we can calculate the observable $\tilde{\boldsymbol{z}}[k]$ based on the estimated, evolved state $\tilde{\boldsymbol{\psi}}[k]$

$$\tilde{\boldsymbol{z}}[k] = \boldsymbol{C} \cdot \tilde{\boldsymbol{\psi}}[k]. \tag{2.10}$$

Note that, while $\tilde{\boldsymbol{z}}[k]$ is derived from an estimate of the state and therefore contains some error, it is free of the background measurement noise present in the measured observable $\boldsymbol{z}[k]$. Thus $\tilde{\boldsymbol{z}}[k]$ is an estimate of the part of the observable due to the line, the part we wish to isolate and remove.

We can now calculate a corrected estimate of the state $\hat{\boldsymbol{\psi}}[k]$ based on the difference between $\tilde{\boldsymbol{z}}[k]$ and $\boldsymbol{z}[k]$

$$\hat{\boldsymbol{\psi}}[k] = \tilde{\boldsymbol{\psi}}[k] + \boldsymbol{K}[k] \cdot (\boldsymbol{z}[k] - \tilde{\boldsymbol{z}}[k]), \tag{2.11}$$

where $\boldsymbol{K}$ is the Kalman gain, described below. This corrected state estimate has an associated error and second moment of the error $\hat{\boldsymbol{P}}[k]$ where, according to [41],

$$\begin{aligned} \hat{\boldsymbol{P}}[k] &= (\boldsymbol{I} - \boldsymbol{K}[k] \cdot \boldsymbol{C}) \cdot \tilde{\boldsymbol{P}}[k] \cdot (\boldsymbol{I} - \boldsymbol{K}[k] \cdot \boldsymbol{C})^T + \boldsymbol{K}[k] \cdot \boldsymbol{V} \cdot \boldsymbol{K}[k]^T \\ &= (\boldsymbol{I} - \boldsymbol{K}[k] \cdot \boldsymbol{C}) \cdot \tilde{\boldsymbol{P}}[k]. \end{aligned} \tag{2.12}$$

The Kalman gain is a linear operator that minimizes the error in the state. It is found by minimizing $\mathrm{tr}\hat{\boldsymbol{P}}[k]$ which is a measure of the total error in the corrected estimate of the state. According to [39], this results in

$$\boldsymbol{K}[k] = \tilde{\boldsymbol{P}}[k] \cdot \boldsymbol{C}^T / (\boldsymbol{V} + \boldsymbol{C} \cdot \tilde{\boldsymbol{P}}[k] \cdot \boldsymbol{C}^T). \tag{2.13}$$

The procedure for finding the portion of the observable due to the line, $\tilde{\boldsymbol{z}}[k]$, is

as follows:

Begin with an estimate of the initial state $\hat{\boldsymbol{\psi}}[0]$ and initial second moment of the error $\hat{\boldsymbol{P}}[0]$. A reasonable choice is to set $\hat{\boldsymbol{\psi}}[0]$ equal to an array of zeros and $\hat{\boldsymbol{P}}[0] = \boldsymbol{W}$. Now for $k = 1$ to the length of $\boldsymbol{z}$:

1. Evolve the estimate of the state via Equation 2.7 and find the error second moment using Equation2.9.

2. Calculate $\tilde{\boldsymbol{z}}[k]$ using Equation 2.10.

3. Find the Kalman gain via Equation 2.13.

4. Update the state and error second moment with Equations 2.11 and 2.12.

We apply a Kalman filter to the lines that are within the frequency band that will be selected when the data are broken into frequency bands. In some cases, lines just outside the frequency band are also removed so that they will not be aliased into the band of interest in the process of restricting the data to a band (see Appendix A). Kalman filtering is applied separately to each line or small group of lines if they are closely clustered in frequency.

2.3.1 Application to Narrow-Band LIGO Features

As mentioned in Section 2.3 the detector data are not appropriate to use as the observable $\boldsymbol{z}$ in the Kalman filtering process. The data minus the line being removed are not white as is assumed for the background process noise. In order to get an observable that consists only of the line and an approximately white background, it is necessary to restrict attention to a narrow frequency range about the line.

For each line or set of lines, a small frequency band is chosen and designated by two numbers: f_c, the central frequency, and Δf, the width of the band. The band then spans the frequencies $f_c - \Delta f/2$ to $f_c + \Delta f/2$. The band must include the line frequency or frequencies f_0 (f_0 is a vector of frequencies in the case of multiple lines) and should be chosen so that the power spectrum within the band is relatively flat and does not include any lines other than those being removed (see Figure 2.3). Typical values we use for Δf are 8-16 Hz. A list of Kalman filtered lines along with the chosen values of f_c and Δf is given in Appendix A

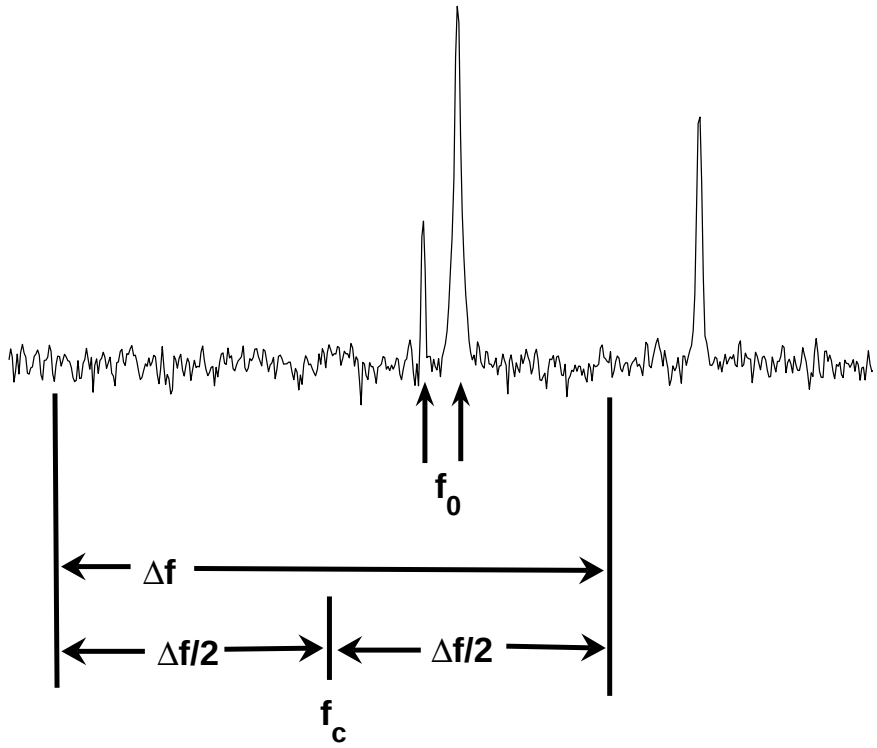


Figure 2.3. Diagram showing the specification of frequency-related variables needed for Kalman filtering. In this case, we want to filter the two lines near the center of the diagram. f_0 is a two element vector that contains the central frequencies of these two lines. f_c and Δf are used to specify a frequency band that includes the lines of interest and some of the white background but excludes any other frequency features.

To restrict the data to the chosen band, the data are shifted down in frequency (heterodyned) by f_c Hz. This shift is done by multiplying the full data, $g[k]$, by a complex exponential with frequency f_c. That is for $k = 1$ to the length of $g[k]$

$$g'[k] = 2 \cdot \exp\left(-2\pi\imath k f_c / f_s\right) g[k], \tag{2.14}$$

where $g'[k]$ are the shifted data and f_s is the sample rate of the original data. After the shift, f_c is at zero frequency and both $f_c + \Delta f/2$ and $f_c - \Delta f/2$ are shifted to $\Delta f/2$.

Next, the shifted data are downsampled to a new sample rate of Δf. This

downsampling throws out all frequencies outside the band of interest. The downsampled data, labeled z, are complex. We construct the real vector series $\boldsymbol{z}[k]$:

$$\boldsymbol{z}[k] = \begin{pmatrix} \Re(z)[k] \\ \Im(z)[k] \end{pmatrix}, \tag{2.15}$$

which is the observable that undergoes the Kalman filtering procedure.

After Kalman filtering is done and the line component of the observable, $\tilde{\boldsymbol{z}}[k]$, is obtained it must be put through a process inverse to that described above before it can be subtracted from the data. First, $\tilde{\boldsymbol{z}}[k]$ is turned into a complex series:

$$\tilde{z}[k] = \tilde{\boldsymbol{z}}[1,k] + \imath \cdot \tilde{\boldsymbol{z}}[2,k], \tag{2.16}$$

where $\tilde{\boldsymbol{z}}[1,k]$ is the element from the first row, kth column of $\tilde{\boldsymbol{z}}$ and $\tilde{\boldsymbol{z}}[2,k]$ is from the second row, kth column (see Equation 2.15).

Then $\tilde{z}[k]$ is resampled to a new sample rate f_s, where f_s is the sample rate of the data prior to Kalman filtering. These resampled data, denoted $\tilde{g}'[k]$, are shifted up in frequency by the amount f_c, the central frequency of the band used to find the observable:

$$\tilde{g}[k] = \exp\left(2\pi\imath k f_c / f_s\right) \tilde{g}'[k]. \tag{2.17}$$

Finally, the real part of the shifted data is taken:

$$\tilde{g}[k] = \Re\tilde{g}[k]. \tag{2.18}$$

$\tilde{g}[k]$ is now the part of the data due to the line. Subtracting $\tilde{g}[k]$ from the data, $g[k]$ removes the line from the data, as shown in Figure 2.4.

2.3.2 Calculating $\boldsymbol{A}$, $\boldsymbol{C}$, $\boldsymbol{W}$, and $\boldsymbol{V}$

Before we can apply the Kalman filter, we must find the values of the matrices $\boldsymbol{A}$, $\boldsymbol{C}$, $\boldsymbol{W}$, and $\boldsymbol{V}$. Detailed derivations of the expressions given below may be found in [39].

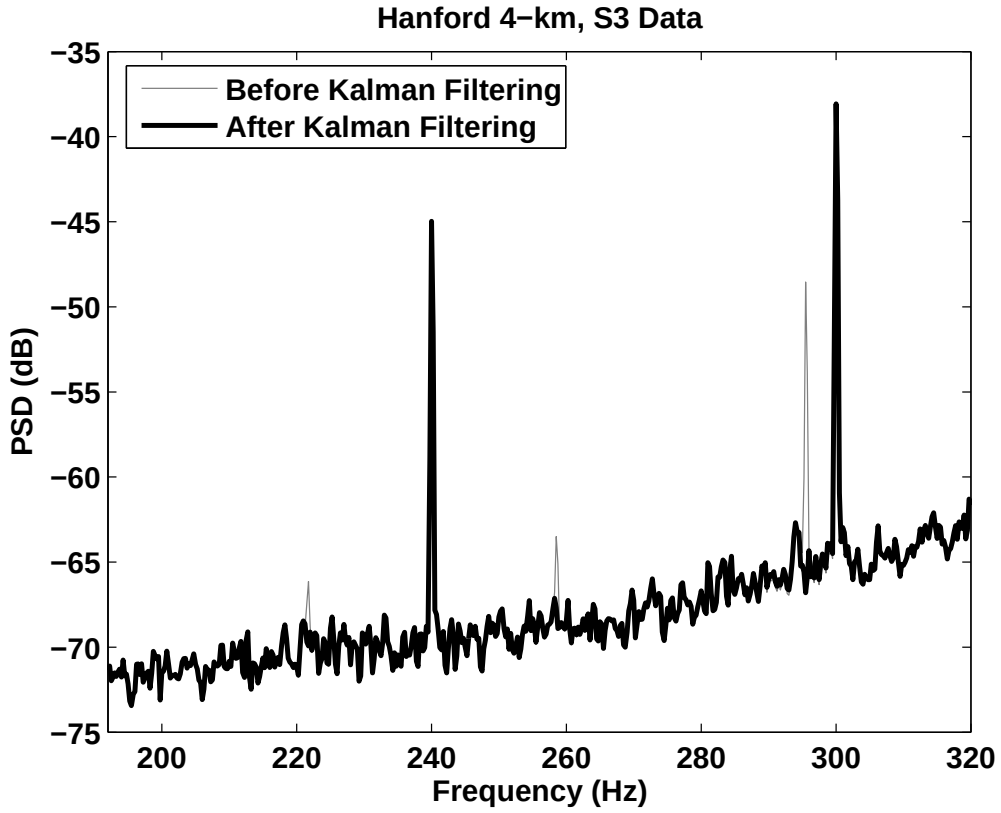


Figure 2.4. Line removal via Kalman filtering in the 192-320 Hz band of data taken during the third science run (S3) with the Hanford 4-km detector. The two lines at 240 and 300 Hz that are not removed are power lines that will be dealt with in the later, line regression step.

We can write the state vector at each time sample as a real vector such that

$$\boldsymbol{\psi}[k] = \begin{pmatrix} \Re(\psi[k]) \\ \Im(\psi[k]) \end{pmatrix}. \tag{2.19}$$

The evolution operator $\boldsymbol{A}$, which controls the evolution of the state vector according to

$$\boldsymbol{\psi}[k+1] = \boldsymbol{A} \cdot \boldsymbol{\psi}[k] + \boldsymbol{w}[k], \tag{2.20}$$

can be calculated via $\boldsymbol{A} = \exp(\boldsymbol{A'}/\Delta f)$ where exp denotes matrix exponentiation

and

$$\boldsymbol{A'} = \begin{pmatrix} -\omega_0/(2Q) & \omega_c - \omega_0\sqrt{1-1/(4Q^2)} \\ -\omega_c + \omega_0\sqrt{1-1/(4Q^2)} & -\omega_0/(2Q) \end{pmatrix}. \tag{2.21}$$

The angular frequencies $\omega_0 = 2\pi f_0$ and $\omega_c = 2\pi f_c$ are the frequency of the line and central frequency of the band about the line that we use in creating the observable, while Q is a damping constant. Q and f_0 are found for each line by calculating high resolution power spectra for each detector. f_0 is the frequency corresponding to the maximum of the line and Q is measured as f_0 divided by the width of the line at half the maximum. f_0 and Q for each Kalman filtered line are given in Appendix A.

If multiple lines are being removed together, the expressions for the state and evolution operator become

$$\boldsymbol{\Psi} = \begin{pmatrix} \Re(\psi_1) \\ \Im(\psi_1) \\ \vdots \\ \Re(\psi_N) \\ \Im(\psi_N) \end{pmatrix} \tag{2.22}$$

and

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A_1} & & \\ & \ddots & \\ & & \boldsymbol{A_N} \end{pmatrix}, \tag{2.23}$$

where $\psi_1 \ldots \psi_N$ and $\boldsymbol{A_1} \ldots \boldsymbol{A_N}$ are the state and evolution operators, respectively, for each of the individual lines.

Next, is the measurement operator $\boldsymbol{C}$ which relates the state to the observable

$$\boldsymbol{z}[k] = \boldsymbol{C} \cdot \boldsymbol{\psi}[k] + \boldsymbol{v}[k]. \tag{2.24}$$

We let the state be the shifted amplitude of the mode causing the line so the measurement operator $\boldsymbol{C}$ is just the identity matrix. For multiple lines, $\boldsymbol{C}$ is a

matrix of repeats of the identity matrix

$$\boldsymbol{C} = (\boldsymbol{C_1} \ldots \boldsymbol{C}_N) = \begin{pmatrix} 1 & 0 & \ldots & 1 & 0 \\ 0 & 1 & \ldots & 0 & 1 \end{pmatrix}. \tag{2.25}$$

The measurement noise covariance matrix $\boldsymbol{V}$ is found by measuring the average value of the power spectrum of the observable $\boldsymbol{z}$ away from the line. This average value, $\bar{P}_m$, is equivalent to the total measurement noise tr$\boldsymbol{V}$. We find $\bar{P}_m$ by summing the power spectrum over frequencies more than 2 Hz away from the line or lines being removed. Also, since the observable $\boldsymbol{z}$ is a vector with real and imaginary parts separated, we take the power spectrum of the two parts separately and then add the means of the two spectra to get $\bar{P}_m$. The measurement noise covariance matrix is then

$$\boldsymbol{V} = \frac{1}{2}\bar{P}_m \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.26}$$

The process noise covariance matrix $\boldsymbol{W}$, on the other hand, is related to the power spectrum at frequencies dominated by the line P_z via the expression [41]

$$P_z(\omega) = \frac{4Q^2 \mathrm{tr}\boldsymbol{W}}{\omega_0^2 + (2Q(\omega + \omega_c) - \omega_0(4Q^2 - 1)^{1/2})^2}. \tag{2.27}$$

The process noise covariance matrix is found by integrating both sides of Equation 2.27 over the range of frequencies where the line is dominant. The expression on the right hand side is so strongly peaked about ω_0 that integrating over the line frequencies is approximately equivalent to integrating over the entire range from $\omega = -\infty$ to $\omega = \infty$. Also, since P_z is the power spectrum of the observable $\boldsymbol{z}$, which are data that have been shifted down in frequency by the amount ω_c, the value of ω_c in Equation 2.27 is zero. The integration of the right side is thus simplified and results in

$$\int P_z(\omega) d\omega = (2\pi Q/\omega_0)\, \mathrm{tr}\boldsymbol{W}. \tag{2.28}$$

When evaluating this expression, it is important to remember that the frequency shift used in producing the observable changes the values of ω_0 and Q. (ω_0(shifted) $= \omega_0 - \omega_c$, Q(shifted) $= (f_0 - f_c)/\Delta f$). It is also important to make sure that one

calculates a power spectrum that preserves the absolute magnitude of the signal creating the line.

Once the power spectra for the real and imaginary parts of the observable $\boldsymbol{z}$ have been found, each is integrated over the line-dominated frequencies. This is done by multiplying the magnitude of the power spectra at each frequency by the frequency bin width and then summing. P_I is computed as the mean of the real and imaginary integrated power spectra. tr$\boldsymbol{W}$ is found via Equation 2.28 and the process noise covariance matrix, $\boldsymbol{W}$ is

$$\boldsymbol{W} = \frac{\omega_0 P_I}{4\pi Q} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.29}$$

If multiple lines are being filtered at once, $\boldsymbol{W}$ becomes

$$\boldsymbol{W} = \begin{pmatrix} \boldsymbol{W_1} & & \\ & \ddots & \\ & & \boldsymbol{W_N} \end{pmatrix}, \tag{2.30}$$

where $\boldsymbol{W_1 \ldots W_N}$ are the process noise covariance matrices for each of the individual lines.

It is important to note that $\boldsymbol{W}$, as a measure of the power contained in a line, is a quantity that can vary greatly over time. The resonances that produce the lines can be excited over short timescales and lines that have large amplitudes at some times can be completely absent at others. To have an accurate value for the process noise it is important to measure it periodically. For the S2 and S3 data, following collaboration specified protocols [42, 43], the process noise was calculated on segments of data whose start times were spaced 6370 seconds apart and were 300 seconds long. When Kalman filtering the lines from a segment of data, we select the $\boldsymbol{W}$ measurement from the segment with a midpoint closest to the midpoint time of the data we are filtering.

2.4 Shaping Filter

As can be seen in Figure 2.1, there is a huge amount of power at low frequencies. These frequencies are lower than the lowest edges of any of the frequency bands we are interested in (128 Hz for S2, S3 and 96 Hz for S4) but this high power can still cause problems. During the basebanding step when the data are restricted to a frequency band, as described in Section 2.5, the low frequency part of the data gets shifted to high frequencies and then the data are downsampled. If the power at low frequencies is not reduced before basebanding, aliasing can occur when this downsampling takes place. The next step in our data conditioning procedure is application of a filter that reduces power at low frequencies.

Generally speaking, a filter is a system that modifies the frequency structure of a signal, amplifying or suppressing some frequencies relative to others. A filter can be expressed as either a series of numbers $x[n]$, where $-\infty < n < \infty$, or in the frequency domain as a frequency response $X(\omega)$, where $-\pi \leq \omega \leq \pi$. The series and frequency response representations are related to each other via the discrete Fourier and inverse Fourier transforms:

$$X(\omega) = \sum_{n=-\infty}^{\infty} x[n]e^{-i\omega n}, \tag{2.31}$$

and

$$x[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(\omega)e^{i\omega n}d\omega. \tag{2.32}$$

Applying the filter in the time domain involves convolving the data $d[n]$ with the filter sequence $x[n]$ to obtain a filtered signal $y[n]$ or

$$y[n] = d[n] * x[n] = \sum_{k=-\infty}^{\infty} d[k]x[n-k]. \tag{2.33}$$

Convolution in the time domain is equivalent to multiplication in the frequency domain so filtering can also be performed by multiplying the filter frequency response by the Fourier transform of the data or

$$Y(\omega) = D(\omega)X(\omega), \tag{2.34}$$

where $Y(\omega)$ and $D(\omega)$ are the Fourier transforms of the filtered data and original data, respectively.

If the data being filtered are zero everywhere but a single sample p where its value is one it is known as a unit impulse signal, written as

$$d[n] = \begin{cases} 1 & n = p \\ 0 & n \neq p \end{cases}, \tag{2.35}$$

or

$$d[n] = \delta[n - p]. \tag{2.36}$$

Filtering the unit impulse signal with the filter $x[n]$ results in

$$y[n] = \sum_{k=-\infty}^{\infty} \delta[k - p]x[n - k] = x[n + p], \tag{2.37}$$

which is simply the filter sequence delayed by p samples. For this reason, $x[n]$ is also called the impulse response of a filter. During filtering, any sample of data will be smeared over a number of samples equal to the length of the impulse response.

The shaping filter that reduces the power at low frequencies should have a frequency profile that changes at the lowest frequency of interest f_l ($f_l = 128$ Hz for S2, for example). Above f_l, the filter should leave the data as unchanged as possible and below f_l, the power of the data should be reduced so that it is less than any frequency above f_l. At the same time, we want to use a filter that is as short as possible to avoid smearing any burst signals over a large number of samples, making them harder to detect.

The best filter for our purposes would be an ideal highpass filter, in other words, a filter whose frequency response is zero below the cutoff f_l and one above it

$$X(f) = \begin{cases} 0 & f < f_l \\ 1 & f \geq f_l \end{cases}. \tag{2.38}$$

As stated above, applying the filter is equivalent to multiplying the filter's frequency response $X(f)$ by the Fourier transform of the data $D(f)$ so that the

filtered data $Y(f) = X(f)D(f)$ and

$$Y(f) = \begin{cases} 0 & f < f_l \\ D(f) & f > f_l \end{cases} . \tag{2.39}$$

The ideal highpass filter completely suppresses frequencies below f_l and leaves those above f_l unchanged.

Unfortunately, this type of ideal filter with its infinitely sharp cutoff is not realizable. Since the impulse response can be calculated from the frequency response via

$$x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\omega) e^{i\omega n} d\omega, \tag{2.40}$$

where $\omega = 2\pi f / f_s$ and f_s is the sample rate, the impulse response of the ideal filter is

$$x[n] = \frac{1}{2\pi} \left[\int_{-\pi}^{-\omega_l} e^{i\omega n} d\omega + \int_{\omega_l}^{\pi} e^{i\omega n} d\omega \right] \tag{2.41}$$

$$x[n] = \frac{-sin(\omega_l n)}{\pi n} \tag{2.42}$$

for all $-\infty < n < \infty$. The impulse response of the ideal filter is infinitely long. What is worse, $x[n]$ is not absolutely summable, that is $|x[n]|$ summed over all n is not less than infinity. This means that when we try to build our ideal filter, it will not converge uniformly for all values of ω. To see this, consider the Fourier transform of our ideal filter impulse response:

$$X(\omega) = \sum_{n=-\infty}^{\infty} x[n] e^{-i\omega n} \tag{2.43}$$

$$= \sum_{n=-\infty}^{\infty} \frac{-sin(\omega_l n)}{\pi n} e^{-i\omega n}. \tag{2.44}$$

If we let $X_M(\omega)$ be the sum of a finite number of terms, or

$$X_M(\omega) = \sum_{n=-M}^{M} \frac{-sin(\omega_l n)}{\pi n} e^{-i\omega n}, \tag{2.45}$$

Figure 2.5 illustrates the behavior of the sum as M increases. With increasing

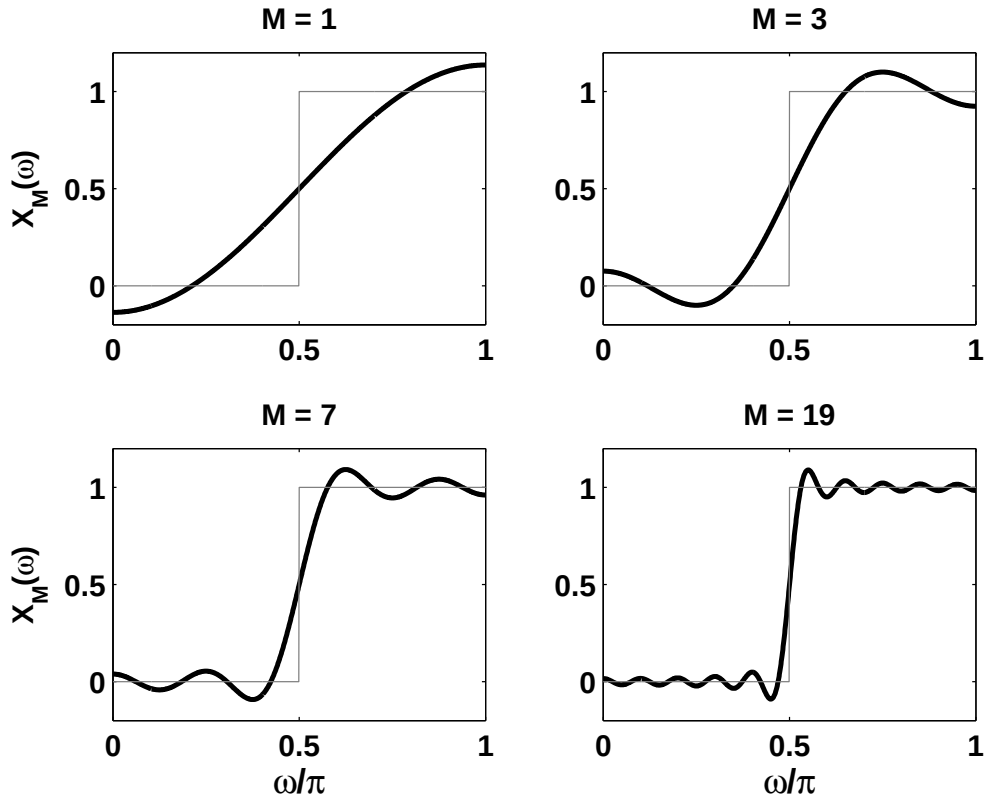


Figure 2.5. Result of summing different number of terms of $X_M(\omega)$ from Equation 2.45 with $\omega_l = 0.5\pi$. As M increases, the frequency response $X_M(\omega)$ (dark line) approaches that of the ideal highpass filter (light line). The transition at $\omega_l = 0.5\pi$ becomes sharper with higher M but the magnitude of the first ripple on either side of the transition does not decrease.

M the transition at $\omega = \omega_l$ becomes sharper and the ripples near $\omega = \omega_l$ become more rapid. In fact, the magnitude of the first ripple on either side of ω_l does not decrease with larger M, a behavior known as the Gibbs phenomenon.

We now see that if we are to apply a real-world highpass filter having a finite length, we are going to have to deal with imperfections in its frequency response. The frequency response will have three regions, a stopband where the response is nearly zero, a passband where the response is close to one and a transition band where the response transitions smoothly between the two extremes. A frequency response with these bands is shown in Figure 2.6. Plus, there will be some ripple in both the passband and stopband. The goal is to build a shaping filter whose

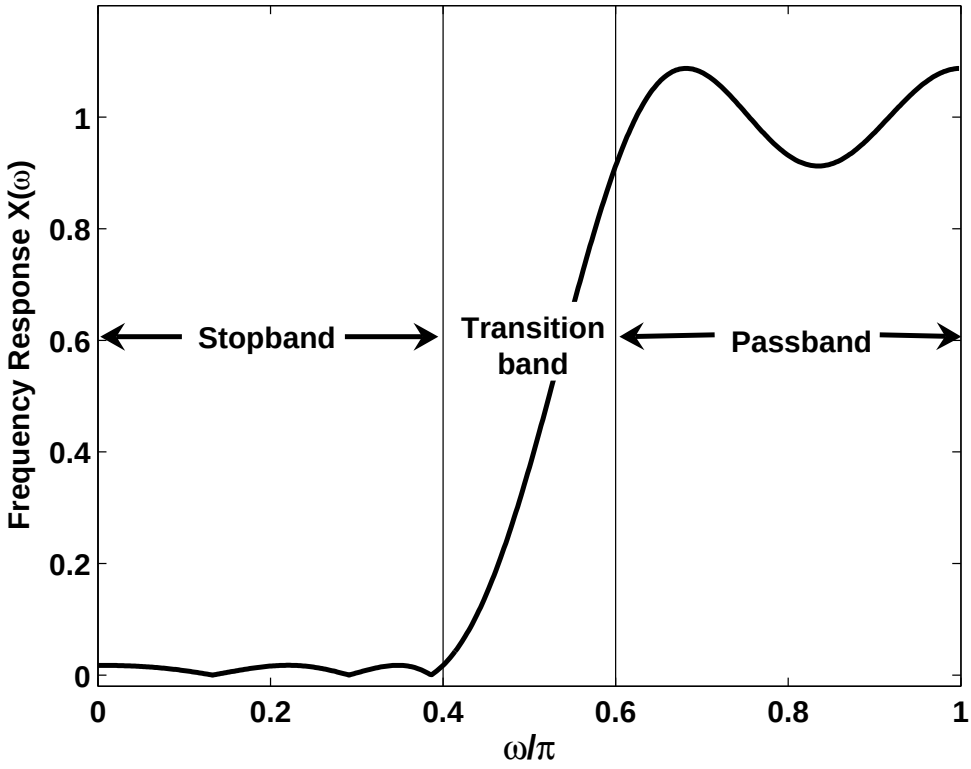


Figure 2.6. Frequency response of realistic highpass filter. The frequency response consists of three frequency regions: a stopband running from 0-0.4π where the response is close to zero, a passband running from 0.6π-π where the response is close to one and a transition band where the response transitions smoothly between the two extremes. The stopband has an allowed error (maximum ripple height) of 0.02 while the passband's allowed error is 0.1. This filter was designed using the Parks-McClellan algorithm.

passband begins at f_l. The transition band should be as wide as possible to lower the filter length but it cannot be too low or the high power at low frequencies will not be suppressed. Also, the filter must be designed so that the ripple is small, especially in the stopband. We design the necessary highpass filter using the Parks-McClellan algorithm. This algorithm returns an optimum filter for a given transition band and given errors in the stopband and passband [36].

The effect of the shaping filter on S3 data from the Hanford 4-km detector is shown in Figure 2.7.

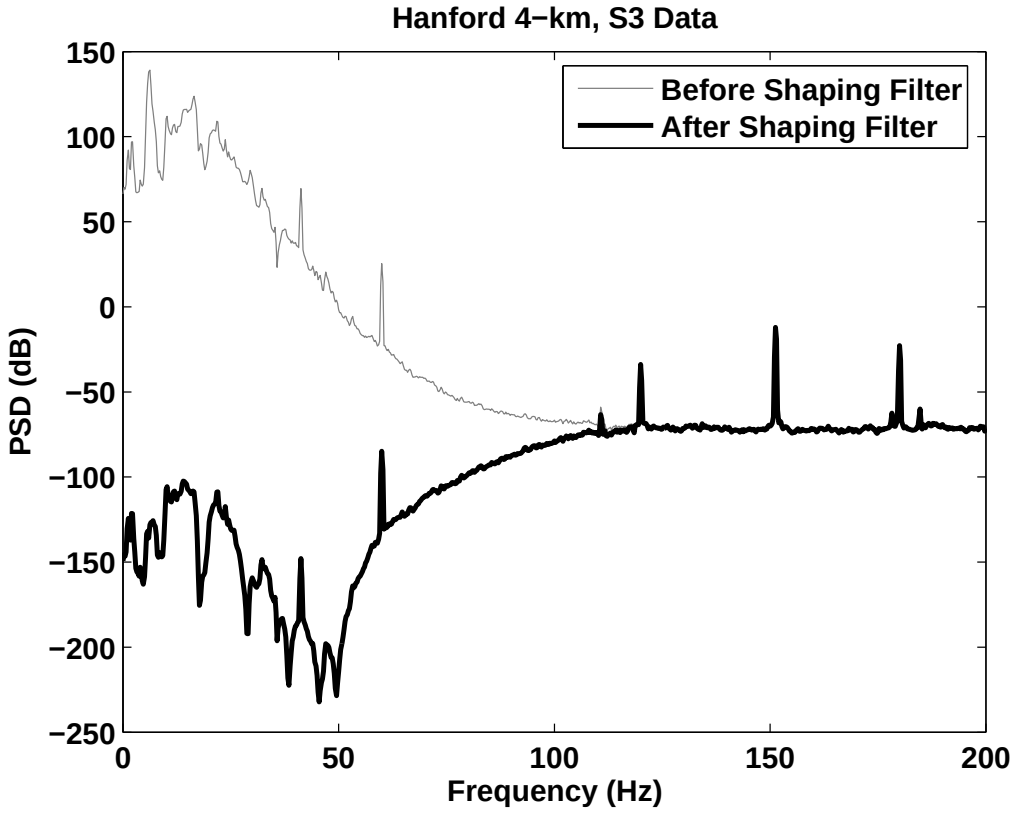


Figure 2.7. The effect of the shaping filter on third science run (S3) data from the Hanford 4-km detector. The power below our lowest frequency of interest (128 Hz) is well suppressed by the filter.

2.5 Basebanding

Trying to remove all of the artifacts and correlations from the data at all of the frequencies where LIGO is sensitive to gravitational waves would be a difficult task and require very long and complex filters. The necessary filters could decrease our sensitivity to gravitational wave bursts by smearing the signals out over a large number of samples. To make the task of whitening simpler, requiring shorter filters, we break the data into frequency bands, each containing a smaller number of features requiring removal. The bands can also be chosen to avoid artifacts that are difficult to remove. For example, the data from S2-S4 were broken into frequency bands that avoided the violin modes (resonances of mirror suspension wires) and their harmonics (see Appendix B). A second advantage to breaking the

data into separate bands is that any candidate gravitational wave events found in a band can be automatically associated with a narrow range of frequencies.

We call the process of breaking the data into frequency bands basebanding. The goal of basebanding is to completely restrict the data in frequency to a band in the original data. After basebanding, the interval of zero frequency to Nyquist frequency is completely taken up by a band of interest. For example, if we want to restrict the data to a band running from 128 Hz to 256 Hz, the basebanding procedure will return data with these frequencies shifted to 0 Hz - 128 Hz and with a sample rate of 256 Hz.

Basebanding has a number of steps and the order in which they are applied is very important. This procedure must be carried out for each frequency band of central frequency f_c and width Δf.

1. Shift by f_c - For each band with central frequency f_c and frequency width Δf, heterodyne (shift down in frequency) the data by f_c. For original data $g[k]$ and shifted data $g'[k]$

$$g'[k] = \exp\left(-2\pi\imath k f_c/\Delta f\right) g[k] \tag{2.46}$$

   for $k = 1$ to the length of $g[k]$. This will have the effect of shifting the data down in frequency by f_c so that the band of interest will be centered at zero frequency. To see how this shifting works, consider data consisting of a single frequency ω_0 or $g[k] = \exp(\imath\omega_0 k)$. Applying Equation 2.46 with $2\pi f_c/\Delta f = \omega_c$ will result in $g'[k] = \exp(-\imath\omega_c k)\exp(\imath\omega_0 k) = \exp(\imath(\omega_0 - \omega_c)k)$ or data shifted to the frequency $\omega_0 - \omega_c$.

2. Downsample - Downsample to move the Nyquist frequency to Δf. This will make the design of the filter used in the next step much easier since now the filter's transition band as a fraction of the Nyquist frequency will be much larger.

3. Lowpass - A lowpass filter (having a frequency response near one below a cutoff and near zero above a cutoff frequency) is applied with a cutoff frequency equal to $\Delta f/2$, removing the frequencies outside the desired band (see Figure 2.8). Therefore, by first shifting the data in frequency, we are

able to effectively bandpass with a lowpass filter, which is easier to design. The lowpass filters used for the S2-S4 data were created using the Parks-McClellan algorithm (see Section 2.4). The filters we used are different for each frequency band, but are the same across detectors. The differences are required by both the differences in frequency width of the bands and the features that lie outside the bands that necessitate different falloff rates so that they will not be aliased into the band. The filter parameters are given in Appendix B.

4. Shift up in frequency by half the bandwidth - Shift the data up by the frequency equal to $\Delta f/2$. This shift will move the low end of the desired band to zero frequency and the high end to the frequency equal to Δf. For $g[k]$ equal to the data before shifting and $g'[k]$ equal to the data after

$$g'[k] = \exp(2\pi\imath k/4)g[k]. \tag{2.47}$$

   Now the entire range of available frequencies, from zero to Nyquist is taken up by the band of interest (see Figure 2.8).

5. Take the real part - Finally, take the real part of the complex data. The complex part will be approximately zero after the previous step. Taking the real part will result in a real time series just as the data before basebanding was real.

2.6 Line Regression

The next step in our data conditioning procedure is to remove the lines that were not removed via Kalman filtering. These lines can be removed by building filters that capture the correlations between the detector output channel and other channels such as environmental channels that monitor the interferometers' surroundings. Power lines and calibration lines, for example, can often be regressed from the data using filters applied to other channels and then subtracting the filtered channel data from the detector output data. The calibration lines are, of

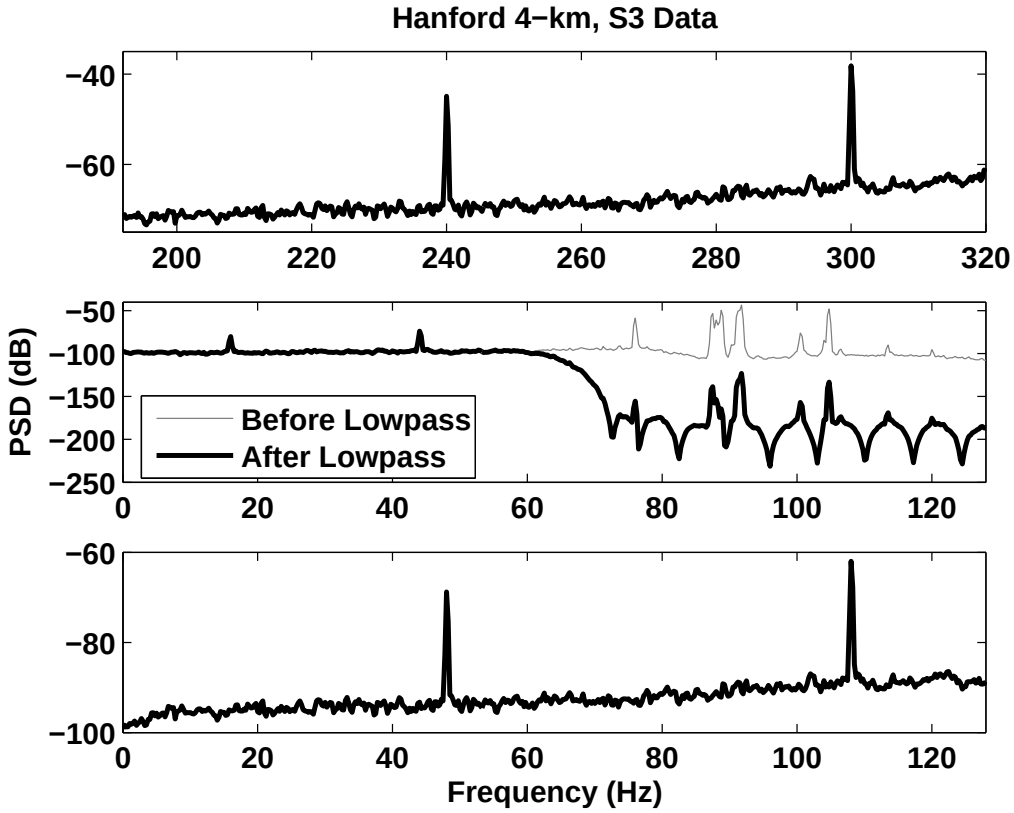


Figure 2.8. The above three plots show the basebanding process performed on the 192-320Hz band of Hanford-4km, S3 data. The top plot is a power spectra of the pre-basebanded data from 192-320 Hz. The middle plot shows the effect of the lowpass filter applied to the shifted, resampled data. The lowpass filter suppresses the features outside the desired band. Finally, the bottom plot shows the basebanded data which ranges in frequency from 0-128 Hz.

course, correlated strongly with the channel which measures the excitation applied to the mirrors to create the calibration lines. The power lines at multiples of 60 Hz that are caused by the AC oscillations of the power grid are correlated with channels from environmental monitors like voltmeters and radiometers and sometimes with optical channels.

To remove the power and calibration lines, the environmental or excitation channel is first resampled to a Nyquist frequency of 2048 Hz and then put through the same data conditioning steps as the detector output including application of the shaping filter and basebanding but excluding Kalman line removal.

For each band, an ARX model is found that relates the basebanded detector data to the basebanded environmental or excitation data. An ARX model has the form $A(q)y(t) = B(q)u(t-k) + e(t)$ or

$$y[t]+a[1]y[t-1]+\ldots+a[n]y[t-n] = b[1]u[t-k]+\ldots+b[m]u[t-k-m]+e[t], \quad (2.48)$$

where $y[t]$ is the output data, $u[t]$ is the input data, $e[t]$ is white noise, k is a delay parameter, the $a[1]\ldots a[n]$ relate current output data to past output samples and the $b[1]\ldots b[m]$ relate input data to output data [44]. In the case of line removal, $y[t]$ is the basebanded detector data and $u[t]$ is the basebanded environmental or excitation data. Since we are only interested in how detector data are related to excitation or environmental data and not in how they are related to past detector data samples $n = 0$, there are no a coefficients. We also assume there is no delay between the channels so $k = 0$.

Once the data have been fit to the ARX model and the least squares estimate of the b coefficients have been found, the power or calibration lines can be removed by first filtering the basebanded environmental or excitation data with the b coefficients and then subtracting the result from the detector data or $d[t] = y[t]-b[n]*u[t]$ where $d[t]$ is the basebanded detector data with the lines removed.

Appendix C gives a list of the lines removed via regression for the three interferometers along with the regression channel used and the filter order m. As with the process noise measurements, the necessary regression filters change with time. Therefore, these filters were also calculated on 300 seconds of data spaced every 6370 seconds and the filters from the closest segment were used for a given data stretch.

The effect of line regression on the 192-320 Hz band of Hanford 4-km, S3 data is shown in Figure 2.9.

2.7 Final Whitening

Now that we have mostly removed the prominent lines and other artifacts from the data, all that remains is to smooth out any frequency features that have survived thus far and remove any general increases or decreases in the noise with frequency.

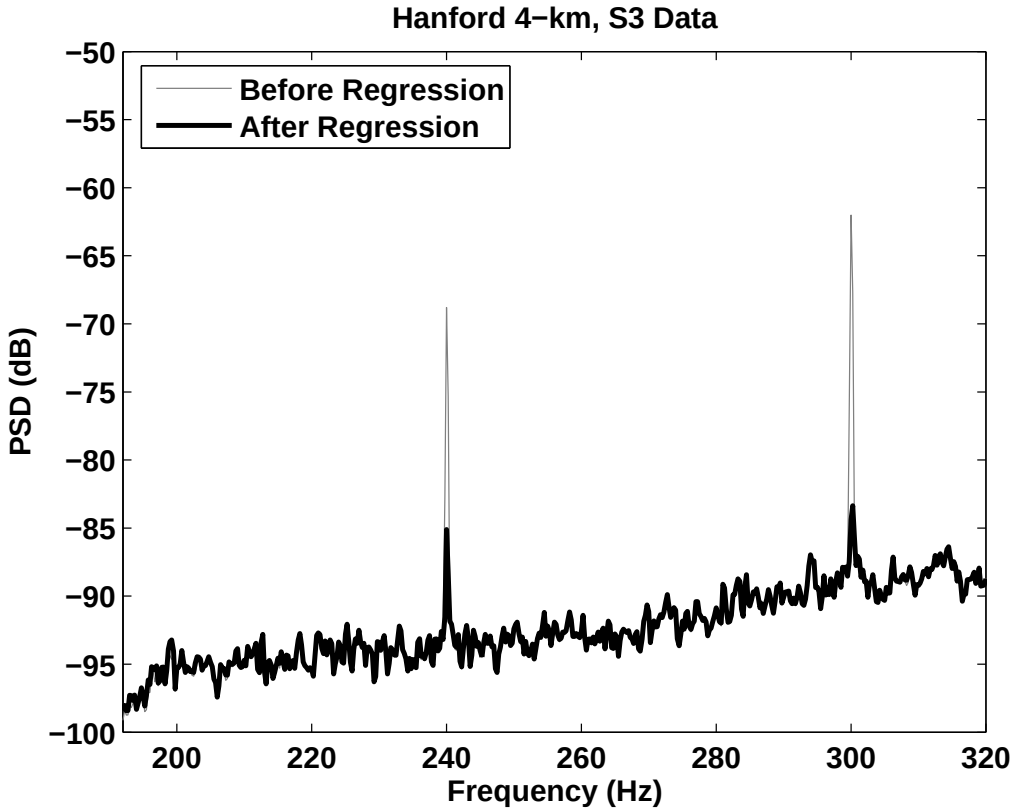


Figure 2.9. The effect of line regression on the 192-320 Hz band of Hanford 4-km, S3 data using a voltmeter channel. The two power lines at 240 and 300 Hz are greatly reduced. This plot shows basebanded data but the frequencies have been mapped to corresponding frequencies in non-basebanded data for ease of comparison.

This last step in our data conditioning procedure is accomplished with a final whitening filter.

The final whitening filter is created by finding an AR model for the basebanded, line-filtered data and using the model as a filter to remove remaining features. AR is a linear difference equation which has the form $A(q)x(t) = e(t)$ or

$$x[t] + a[1]x[t-1] + a[2]x[t-2] + ... + a[n]x[t-n] = e[t] \tag{2.49}$$

[44]. The use of this model assumes that elements of the time series are made up of a component which can be predicted from past behavior and a white noise component. If the time series is filtered with the AR coefficients, the result is just

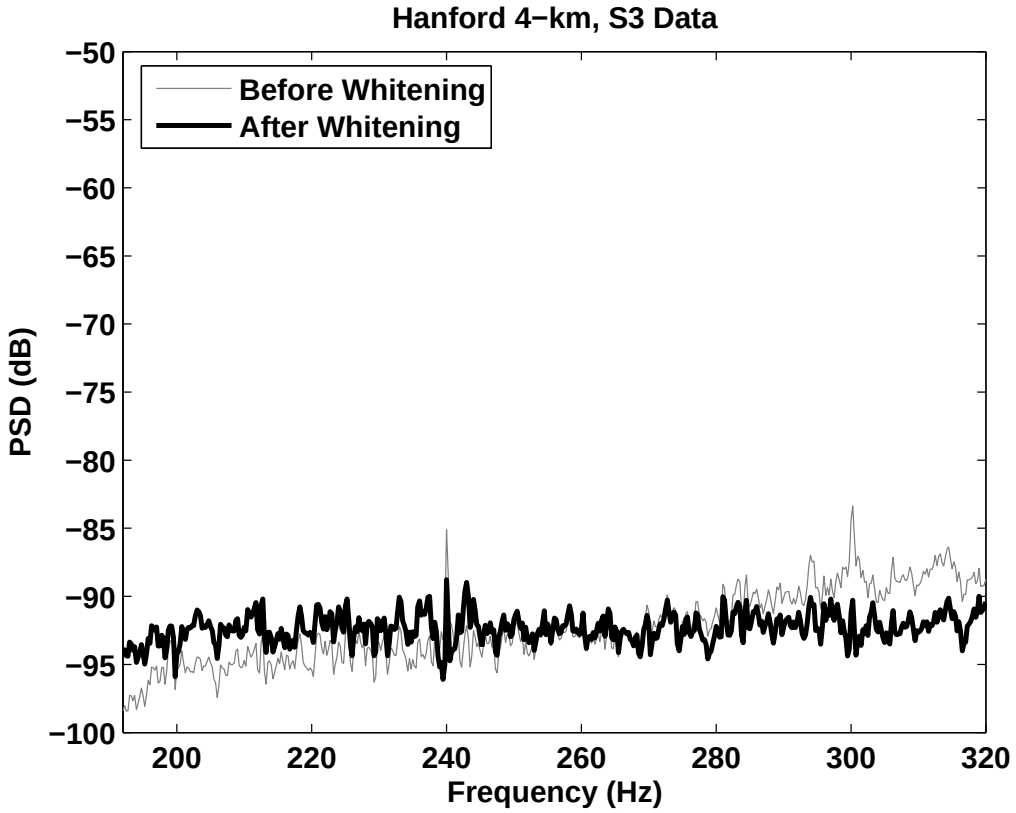


Figure 2.10. The effect of whitening on the 192-320 Hz band of Hanford 4-km, S3 data. The general increase in power with frequency has been flattened. This plot shows basebanded data but the frequencies have been mapped to corresponding frequencies in the non-basebanded data for ease of comparison.

the white noise component, $e(t)$. For the filter applied to the S1-S4 LIGO data, we set the number of a coefficients, n to 60 for each band. This choice of order was made by noticing that an increase in order above 60 did not result in substantially better whitening.

Before calculating the whitening filter, it is important to remove a small piece of data from the beginning and end of the data to avoid startup and ending transients (described in Section 2.8). These are created by previous steps in the data conditioning process.

Once again, the whitening filter was found using 300 seconds of data every 6370 seconds and the filter from the nearest segment is used to whiten data. An example of whitening is shown in Figure 2.10.

2.8 Startup and Ending Transients

Several of the data conditioning steps involve filtering the data. Filtering usually involves zero padding, or adding zeros to the beginning of the data so that the filtered data is of the same length as the initial data. While zero padding preserves length, it can cause the first samples of the filtered data to behave strangely. (For example, these samples can have abnormally large or small amplitudes). As a result, it is a good idea to discard these affected samples at the end of the data conditioning procedure.

To find the length of this startup transient, we calculate the length of the impulse response of the data conditioning procedure. The impulse response is found by applying the data conditioning procedure to a time series with a one for its first sample and all other samples equal to zero. The data conditioning filters will smear out the one in the first sample. The number of nonzero samples in the conditioned time series will be equal to the number of samples influenced by the zero padding. There is also an end transient caused by filtering during basebanding. The length of this transient can be found by applying data conditioning to a vector with a unit impulse for its last sample. When calculating the impulse responses, Kalman filtering and line regression are turned off since these steps do not involve direct filtering of the data but instead apply filters to copies of the data or other channels that are then subtracted from the data.

The startup and ending impulse responses for S3 data conditioning on the 192-320 Hz band are shown in Figures 2.11 and 2.12. The lengths of the transients for all bands for S3 are given in Table 2.1. According to this table, removing a few tenths of a second from the beginning and end of conditioned data should be sufficient to avoid data conditioning transients.

2.9 Effectiveness of Data Conditioning

Data conditioning involves many choices, from what lines will be removed with a Kalman filter to what environmental or interferometer channels will be used in regression to filter orders. Poor choices could result in ineffective data conditioning. Plus, line strength measurements as well as filter calculations are done on data

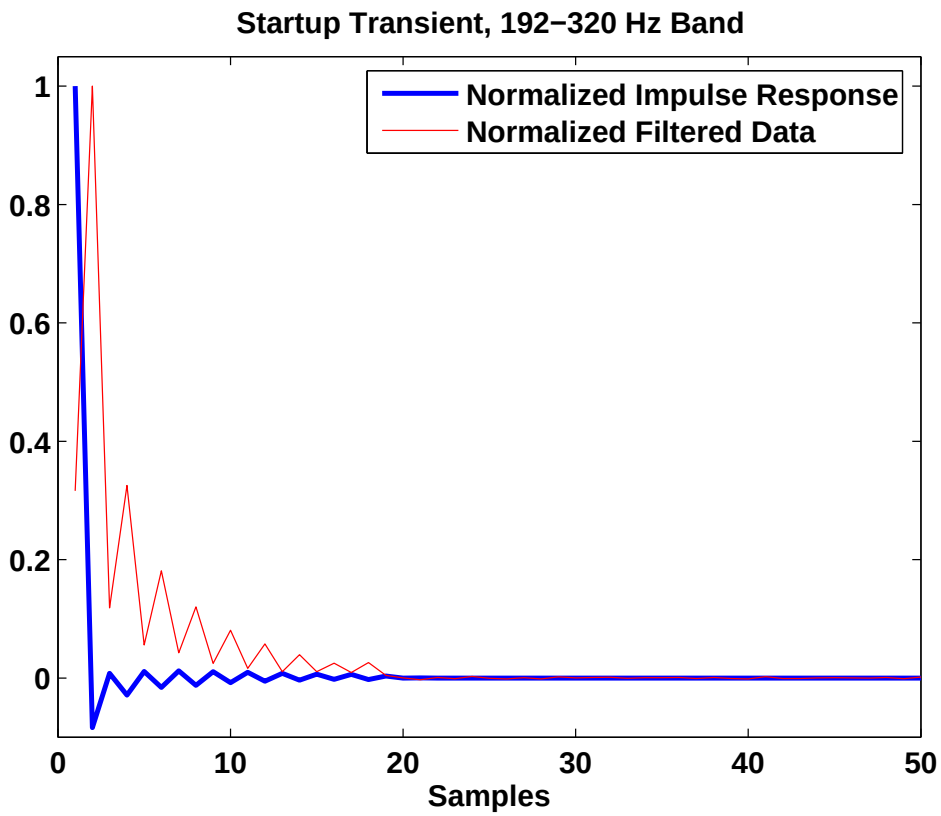


Figure 2.11. Startup impulse response for S3 data conditioning on the 192-320 Hz band as well as startup transient for conditioned data. The impulse response was found by applying data conditioning to a vector with one as its first element and zeros for all other elements. Kalman filtering and line regression were both turned off. The impulse response is the same length as the startup transient for the filtered data. Both the impulse response and filtered data have been normalized so that their largest samples are equal to one for ease of comparison.

segments spaced every 6370 seconds and these filters are applied to data away from these segments where the noise characteristics may have changed. It is important to know if applying the data conditioning procedure is actually working and producing nearly white data or if there is a failure either as a result of a poor parameter choice or a bad section of data that should not be used in an analysis.

The simplest way to see if data are being whitened is to look at a power spectral density (PSD) plot and see if the conditioned data has a PSD that has nearly the same magnitude for all frequencies. Plots showing the effectiveness of data conditioning on a segment of S3 data from the Hanford 4-km detector and all

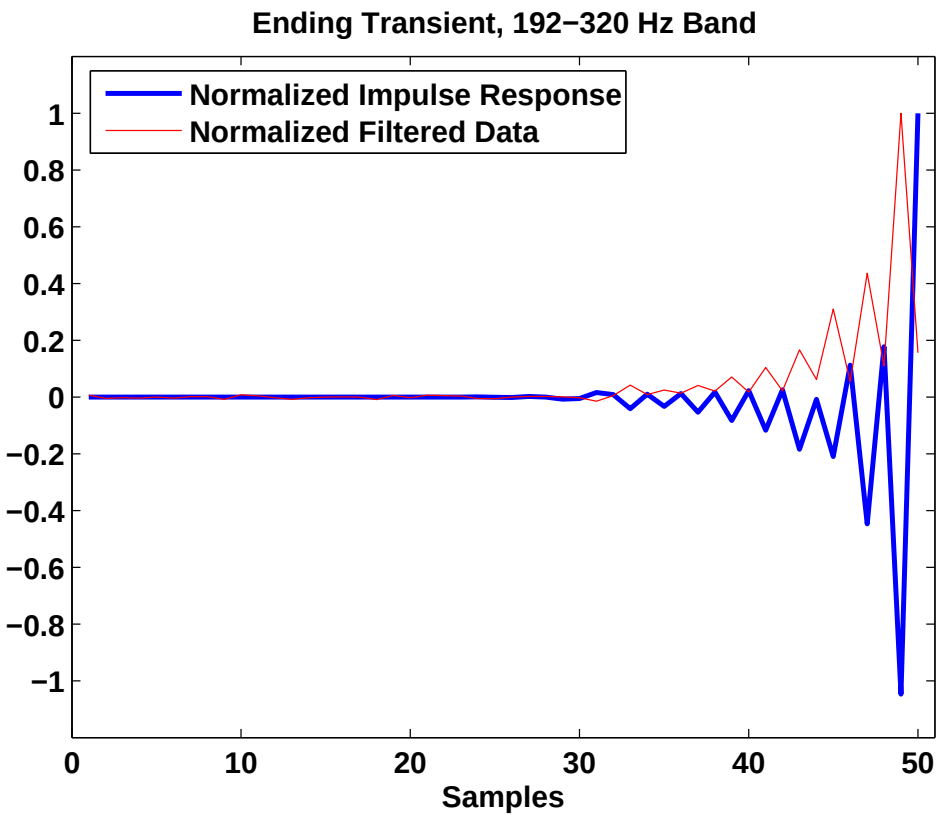


Figure 2.12. Ending impulse response for S3 data conditioning on the 192-320 Hz band as well as ending transient for conditioned data. The impulse response was found by applying data conditioning to a vector with one as its last element and zeros for all other elements. Kalman filtering and line regression were both turned off. The impulse response is the same length as the ending transient for the filtered data. Both the impulse response and filtered data have been normalized so that their largest samples are equal to one for ease of comparison.

frequency bands are shown in Appendix D. Looking at PSDs by eye, besides being merely qualitative, would be impossibly time consuming if applied to a significant fraction of the huge amount of data produced by the three LIGO detectors. We need a quantitative measure of data conditioning performance that can be applied quickly.

A simple measure of data conditioning effectiveness has been proposed in [45]. For this method, the Rayleigh statistic $\mathcal{R}$ is calculated before and after data conditioning is applied. The Rayleigh statistic is a measure of the flatness of the PSD and is calculated as $\mathcal{R}_{PSD} = \sigma_{PSD}/\mu_{PSD}$ where σ_{PSD} and μ_{PSD} are the vari-

Freq. Band (Hz)	Startup Imp. Resp.		Ending Imp. Resp.	
	Samples	Seconds	Samples	Seconds
128 - 192	27	0.2109	26	0.2031
192 - 320	36	0.1406	35	0.1367
384 - 512	34	0.1328	33	0.1289
512 - 640	34	0.1328	33	0.1289
704 - 832	43	0.1680	42	0.1641
832 - 1024	59	0.1536	58	0.1510
1065 - 1365	64	0.1067	63	0.1050
1408 - 1708	64	0.1067	63	0.1050
1758 - 2048	62	0.1069	61	0.1052
890 - 940	23	0.2875	22	0.2200

Table 2.1. Startup and ending impulse response lengths for S3 data conditioning and all bands. The impulse responses were found by applying data conditioning to a vector of all zeros except for a unit impulse at the first (startup) or last (ending) sample. Kalman filtering and line regression were both turned off.

ance and mean of the PSD, respectively. The flatter the PSD, the smaller $\mathcal{R}_{PSD}$. Therefore, a figure-of-merit (FOM) for the data conditioning is

$$\text{FOM} = \frac{\mathcal{R}_{PSD'}}{\mathcal{R}_{PSD}} = \frac{\sigma_{PSD'}/\mu_{PSD'}}{\sigma_{PSD}/\mu_{PSD}}, \tag{2.50}$$

where PSD is the power spectral density calculated on data that has only been treated with the shaping filter and then basebanded and PSD' is data that has undergone the entire data conditioning process. If full data conditioning including Kalman filtering, line regression and whitening produces data that are whiter than data that have only been restricted to the same frequency band, then FOM < 1.

The FOM was calculated for data from all three LIGO detectors taken during S3 and for all frequency bands used in S3. For each detector, the data that passed quality cuts recommended by the Detector Characterization Working Group [46, 47, 48, 49] (the cuts ensure the detector was in a good, operating state) were broken into 300 second long non-overlapping segments. Any segment less than 300 seconds long was discarded. Data conditioning was applied to each segment and the FOM calculated. The percentages of S3 segments for which data conditioning was successful are shown in Table 2.2. For most segments, the FOM was less than one.

Freq. Band (Hz)	% FOM < 1		
	Hanford 4-km	Hanford 2-km	Livingston 4-km
128 - 192	100.0	100.0	99.9
192 - 320	99.6	100.0	99.8
384 - 512	99.6	99.7	98.2
512 - 640	99.6	100.0	98.6
704 - 832	96.1	97.0	98.7
832 - 1024	100.0	100.0	100.0
1065 - 1365	100.0	100.0	100.0
1408 - 1708	100.0	99.9	99.0
1758 - 2048	100.0	98.5	87.5
890 - 940	81.4	99.6	100.0

Table 2.2. Percentage successful data conditioning. Shown above are the percentages of S3 data segments for each detector and frequency band where the figure-of-merit (FOM) was less than 1. A FOM less than 1 indicates that the Kalman filtering, line regression and whitening portions of data conditioning were successful in producing whiter data.

Figures 2.13 and 2.14 show the FOM per segment and a histogram of FOM values for the 192-320 Hz band of the Hanford 4-km detector. Corresponding plots for all frequency bands and detectors are given in Appendix E. In the 192-320 Hz band for the Hanford 4-km detector, the FOM was less than one for 9,814 (or 99.6%) of the 9,854 segments tested in S3. Looking closely at the remaining 40 segments with high FOM values, data conditioning was found to fail for two reasons. In the first 31 segments, the regression of power lines failed and the strength of a power line was increased. Usually this occurred when the 240 Hz line disappeared or mostly disappeared from the data but did not disappear from the regression channel. The disappearance of the line may have been caused by fans, motors, or heaters switching on or off and changing the magnetic environment around the mirrors, for example. This change in relationship between the data and regression channels occurred sometime between these segments and the playground segment on which the regression filter was calculated. To fix the power line regression, filters would have to be calculated closer to these segments. The remaining 9 segments with high FOM values had "glitches" which are very brief (usually just a few samples) very high amplitude pulses in the data of instrumental origin. Most data segments containing glitches are excluded by data quality cuts but the ones in our 9 segments had evaded these cuts. The data conditioning procedure cannot

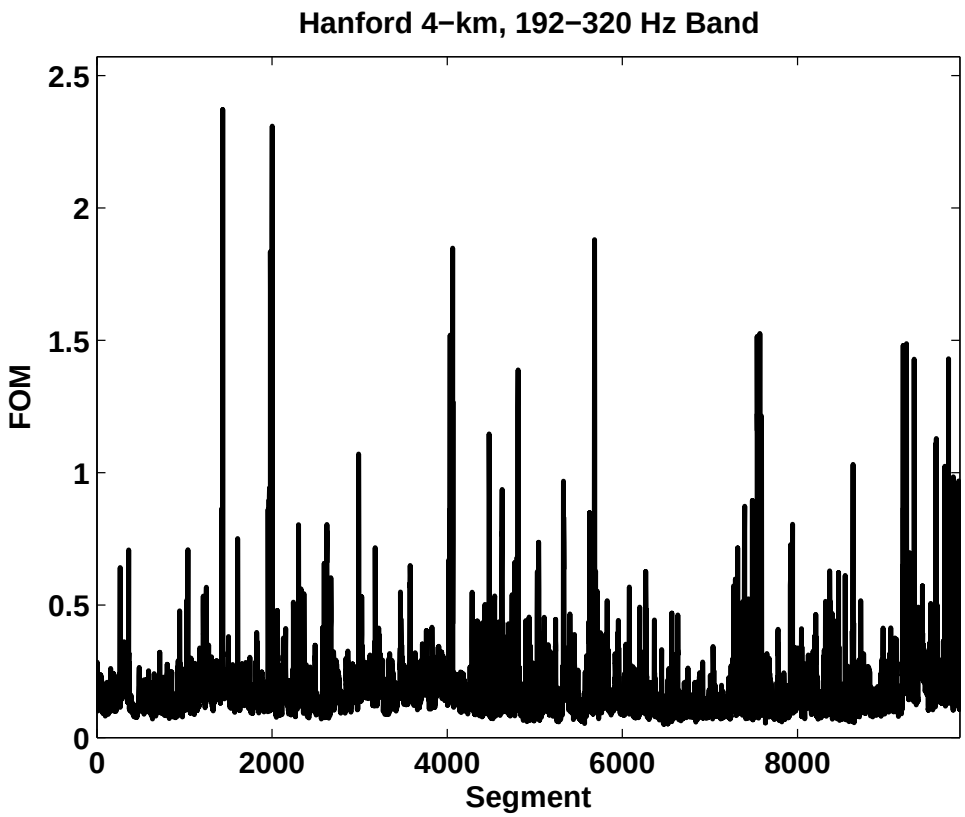


Figure 2.13. FOM per segment for the Hanford 4-km, S3 data, 192-320 Hz band. The segments for which FOM > 1 are distributed throughout the run.

remove the effects of glitches and these segments should be avoided (or vetoed) by any analysis looking for gravitational wave bursts.

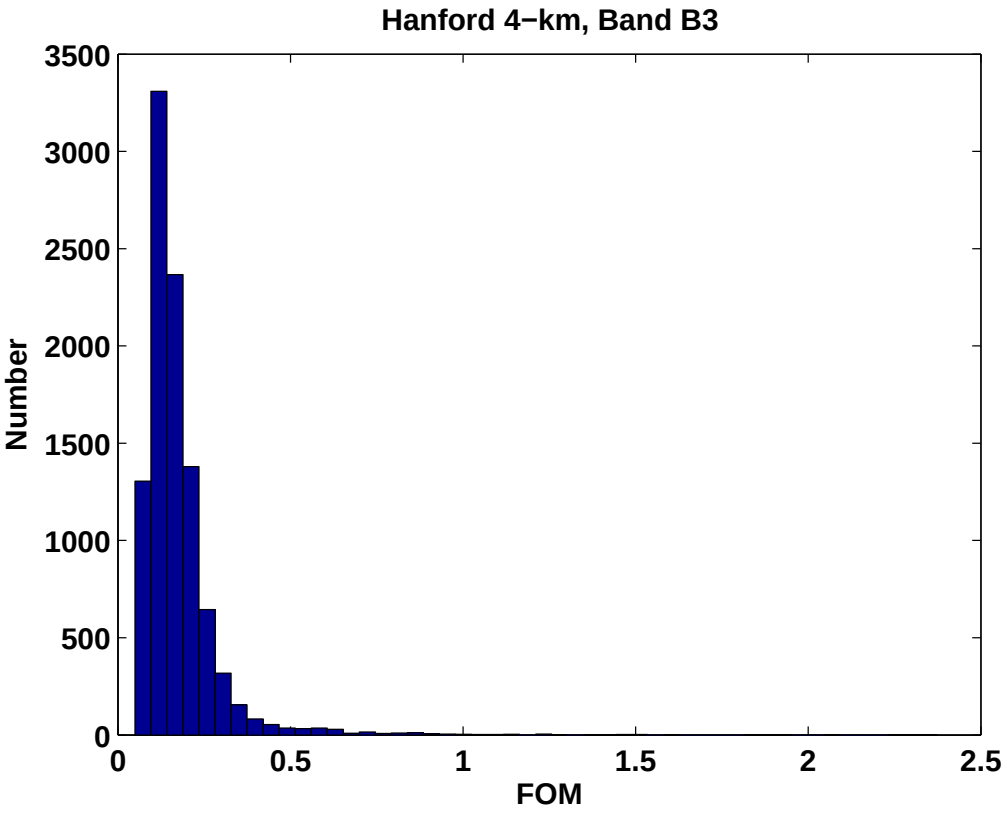


Figure 2.14. Histogram of FOM values for the Hanford 4-km, S3 data, 192-320 Hz band. Most segments have FOM < 1 but there is a tail of FOM values that extend above 1.

Chapter 3

A Test for Non-Linear Couplings

3.1 Introduction

While second order tools such as the power spectral density (PSD) function are useful tools for discovering many interesting features present in data, there are many important features to which these tools are blind. As an example, consider an input, such as seismic noise, n_j affecting the output x_j of a system. The relationship between the input and output may be a simple linear one. On the other hand, the size of the coupling, or gain, between input and output may depend upon the magnitude of the input itself. Furthermore, the dependence may not involve the size of the present input but the size of the input in the past. A model for this nonlinear situation is given by

$$x_j = \frac{1}{(1+\tan^2\theta)^{1/2}}(n_j + n_j n_{j-k}\tan\theta) \tag{3.1}$$

where x_j and n_j are the current outputs and inputs respectively, n_{j-k} is an input that occurred k samples into the past, and $\theta \in [0, \pi/2]$ is a parameter that controls the size of the nonlinearity. The $(1+\tan^2\theta)^{-1/2}$ term is included to maintain unit variance.

If the n_j are assumed to be white, the x_j will be white also since the auto-correlation function $C_l = 1/T\sum_{j=0}^{T} x_j x_{j-l}$ is zero for all values of $l \neq 0$. Therefore, second order measures like the PSD contain no information about this nonlinearity. What is more, when θ is close to $\pi/2$ so that $\tan\theta$ becomes large, the third order

moment also goes to zero. This results in a problem: can one discover this type of nonlinearity without calculating many, computationally expensive, higher order moments?

Despite the fact that the power spectral density of x is white, it is still the case that x_j is related to x_{j-k} and x_{j+k}. If there were *no* correlation between the x_j and a threshold X applied, the above threshold points would be Poisson distributed. In other words, the number of above-X points N within any interval T will be given by

$$P(N|T,\tau) = \frac{(T/\tau)^N}{N!} e^{-T/\tau} \tag{3.2}$$

where $1/\tau$ is the event rate parameter. The time correlations present in the data produced by Equation 3.1 will lead to clustering in time of the above-X points and therefore they will not be Poisson distributed. This observation suggests a test: determining whether the type of nonlinearity described by Equation 3.1 is present is equivalent to determining if the data points above a threshold X are or are not Poisson distributed.

3.2 Description of Test

Given some data, we want to discover if there is evidence of non-linear couplings. To accomplish this, we test the null-hypothesis that the time distribution of events above a threshold X, drawn from whitened data, are Poisson distributed. If the null hypothesis is rejected, we conclude that there is evidence for non-linear couplings.

First, the data must be whitened. A procedure for whitening is described in Chapter 2. Once the data are white, the first step in determining if the time series is free of inter-sample correlations is to pick a threshold X and select the points of the time series with values greater than X. As previously stated, if the series is indeed made up of independent samples, the probability of the number of above-X points N within any interval T will be Poisson and given by Equation 3.2.

The second step in the test is to consider the set of intervals between successive above-X points (Δt). If these points are Poisson distributed, the probability that

there are no above-X points in the time interval Δt is

$$P(0|\Delta t, \tau) = e^{-\Delta t/\tau}. \tag{3.3}$$

Let $Q(t|\tau)$ be the probability that the interval between threshold crossings is t. The probability that there are no threshold crossings in an interval Δt is equal to the probability that the interval between threshold crossings is Δt or larger:

$$P(0|\Delta t, \tau) = \int_{\Delta t}^{\infty} dt\, Q(t|\tau), \tag{3.4}$$

or

$$\frac{d}{d(\Delta t)} P(0|\Delta t, \tau) = -Q(\Delta t|\tau). \tag{3.5}$$

Therefore, the intervals between above-X points will have an exponential distribution $Q(\Delta t|\tau) = \exp(-\Delta t/\tau)/\tau$.

The problem of determining if a given time series is without correlations is now simplified from calculating many moments of high order to determining if the set of Δt for a given X follows an exponential distribution. If we bin the Δt, we can use the χ^2 test [50] to compare the number of Δt in each bin to what would be predicted based on an exponential distribution.

The intervals are first histogrammed into N_B bins. We set the lower edge of the first bin b_0 to zero. The other bin edges b_n are chosen so that if the Δt follow an exponential distribution, there will be an equal number in each bin (each bin is equally probable) or

$$\int_{b_{n-1}}^{b_n} dt\, e^{-\Delta t/\tau} = \frac{1}{N_B} \tag{3.6}$$

so that $b_n = -\tau \ln(\exp(-b_{n-1}/\tau) - N_B^{-1})$. The expected number of Δt in each bin (i), written as E_i, is equal to $1/N_B$ times the total number of Δt. On the other hand, the number of Δt observed to fall in each bin (i) is referred to as O_i. Finally, the quantity χ^2 is calculated via:

$$\chi^2 = \sum_{i=1}^{N_B} \frac{|O_i - E_i|^2}{|E_i|} \tag{3.7}$$

If the set of Δt intervals agrees with the expected, exponential, probability dis-

tribution, χ^2 will be approximately equal to the number of degrees of freedom ν. The number of degrees of freedom is defined as the number of independent observations of a sample minus the number of parameters which are estimated from observations of the sample. In our case, ν is equal to $N_B - 2$ where N_B is the number of bins. ν is approximately equal to N_B since our observations are the number of Δt that fall in each bin. ν is two less than N_B for a couple of reasons. One degree of freedom is lost because we know the total number of Δt so that once we know the number of Δt in all but the last bin, we know the number in the last bin without counting, so the number in the last bin is not independent of the others. The other degree of freedom is lost via our method of determining the rate τ. Before calculating our bins of equal probability via Equation 3.6 we use bins of equal width (except for the last bin which has its upper edge at infinity and lower edge set so that there are 5 Δt observed in it) to calculate χ^2 for differing τ and choose the τ that minimizes χ^2.

3.3 Application to Model

3.3.1 Apply to Data Without Correlations

The Poisson test described above was first applied to a white time series without any inter-sample correlations, generated by a normally distributed random number generator. A distribution of the test results was created by running the test 10,000 times on randomly generated data segments 100,000 points in length with a threshold $X = 2.6$ and 10 bins ($N_B = 10$). The mean of this distribution was 8.14 which is close to the number of degrees of freedom $\nu = N_B - 2 = 8$. Since these data have no inter-sample correlations, the Poisson test results should follow a χ^2 probability distribution with $\nu = 8$. The χ^2 probability distribution is given by

$$f(x) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{(\nu-2)/2} e^{-x/2}, \tag{3.8}$$

where $\Gamma(\alpha)$ is the gamma function and

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy. \tag{3.9}$$

Applying the χ^2 test again to see if the distribution of Poisson test results are described by a χ^2 distribution with $\nu = 8$ we get a χ^2 test result equal to 8.676. We applied the test using 10 bins and knew ν independently from the Poisson test results so the number of degrees of freedom for the χ^2 test is 9. $\chi^2 = 8.676$ is less than 9 and the probability of getting a χ^2 result of this value or greater is 48.8%. Therefore, we can say with confidence that the Poisson test applied to uncorrelated data gives results that follow a χ^2 probability distribution with ν equal to the number of degrees of freedom used in the test.

3.3.2 Apply to Model Data with Correlations

The next task is to apply the Poisson test to a time series generated by a nonlinear model which is white but possesses higher order moments. One such model is $x_j = (1+\tan^2\theta)^{-1/2}(n_j + n_j n_{j-k}\tan\theta)$. The series produced by this model is white since the auto-correlation function $C_l = 1/T\sum_{j=0}^{T} x_j x_{j-l}$ is zero for all values of $l \neq 0$. The higher order moments, however, are nonzero. Physically, this model may be thought of as representing a type of nonlinear echo where the repeated part is delayed by k timesteps but its coupling depends upon the current state of the signal. The model may also represent a system where impulses travel along two different paths and then are coupled linearly and nonlinearly into the signal.

When the test is applied to this model, progressively higher values of χ^2 are obtained for higher values of θ up to about 0.3π where the χ^2 values start decreasing somewhat but always remain significantly higher than the degrees of freedom. Figure 3.1 shows the means of distributions of χ^2 results calculated for values of θ ranging from zero to $\pi/2$, corresponding, on one hand, to no contribution from non-linear couplings, and on the other hand, to no contribution to the uncorrelated signal.

3.3.3 False Alarm / Efficiency

It is important to find a quantitative measure of how well the Poisson test detects the presence of higher-order moments beyond simply noting the increase in the average χ^2 value for larger θ values. One way of making this evaluation is to compare the probability that data having non-linear couplings will be correctly

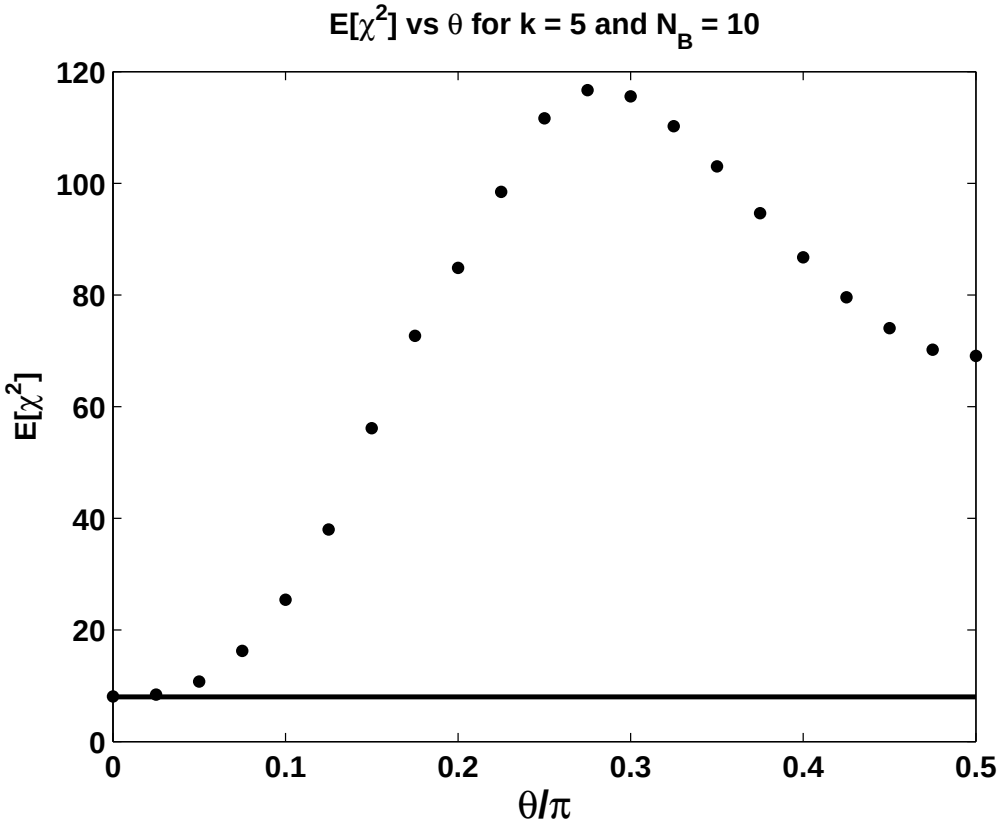


Figure 3.1. Averages of χ^2 test result distributions using the model of Equation 3.1 and different values of θ. Ten bins were used in the test, the threshold $X = 2.6$ and a value of 5 was chosen for k. The solid line shows the number of degrees of freedom.

identified by the Poisson test (efficiency probability) with the probability that data having no couplings will be erroneously flagged as having couplings by the Poisson test (false alarm probability). We want a test that is highly sensitive to couplings and flags data containing couplings as such as often as possible (high efficiency) but that incorrectly flags data as having couplings when they actually do not as infrequently as possible (low false alarm). Comparing the efficiency and false alarm probabilities for Poisson tests having different parameters such as threshold X and number of bins N_B will allow us to design a test that is best at correctly categorizing data that do and do not contain non-linear couplings.

Data are identified as containing correlated samples by setting a threshold on χ^2 labeled χ^2_T. The Poisson test is applied to the data and if the χ^2 resulting

from the test is greater than χ^2_T the data are identified as containing non-linear couplings. If, on the other hand, the Poisson test's χ^2 is less than χ^2_T we say that the data are free of non-linear couplings.

Now consider applying the Poisson test to many sets of data generated by the model $x_j = (1 + \tan^2\theta)^{-1/2}(n_j + n_j n_{j-k} \tan\theta)$ for some nonzero θ that contain a non-linear coupling. One will obtain a distribution of Poisson test χ^2 values from the different data sets. A certain fraction of these χ^2 results will be above the threshold χ^2_T. The fraction of above-threshold results is equal to the efficiency probability (ϵ) and it gives the probability of a test result generated from the model being correctly identified with the threshold χ^2_T.

Next, if we apply the Poisson test to many sets of data without any couplings, we will obtain a different distribution of Poisson test χ^2 values. A fraction of these χ^2 results will also be above the threshold χ^2_T. This fraction of above-threshold results is equal to the false alarm probability (α) and gives the probability that data without couplings will be incorrectly labeled as having couplings if one uses the threshold χ^2_T.

We can let the χ^2_T threshold vary and calculate pairs of ϵ and α probabilities for a range of χ^2_T. The probability pairs may then be plotted against one another as is done in Figure 3.2. A straight line from $\epsilon = 0$, $\alpha = 0$ to $\epsilon = 1$, $\alpha = 1$ indicates that at any threshold χ^2_T, the probability of correctly identifying couplings with the Poisson test is the same as incorrectly flagging data without couplings. The Poisson test in this case would be no better then flipping a coin to decide if data contained couplings. If, on the other hand, the efficiency vs false alarm curve lies above this straight line it means that for any threshold χ^2_T, the Poisson test is more likely to correctly identify couplings than it is to incorrectly flag data without couplings. In fact, the more the ϵ vs α curve departs from the straight line, the better the test will perform.

Figure 3.2 gives ϵ vs α curves made with distributions of Poisson test results calculated using the non-linear coupling model with different values of θ. The curves associated with higher value of θ depart most from a straight line, showing that the Poisson test is most successful at identifying couplings in data from models with higher values of θ. This is what one would expect, since the model data with the highest θ is the least like data without couplings.

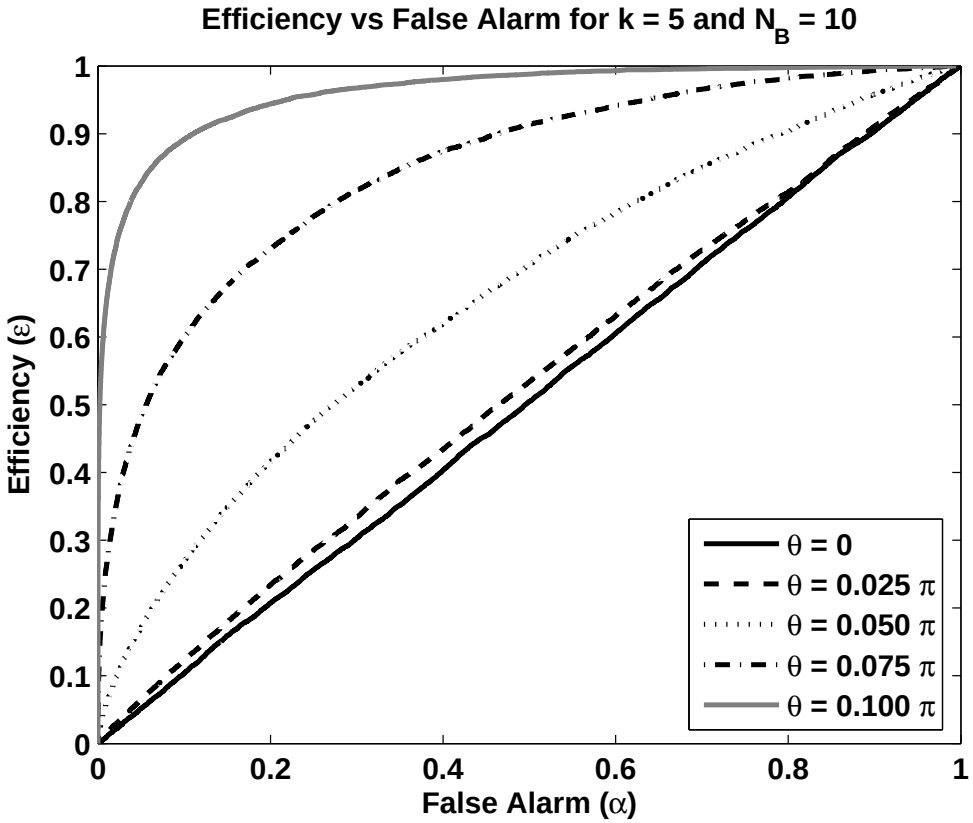


Figure 3.2. False alarm / efficiency probability pairs for the model data and different values of θ. At any false alarm value, the corresponding efficiencies increase for greater θ. This shows that the test works best when the portion of the model due to correlated samples is large.

3.4 Choosing Parameters

When performing the Poisson test, there are two parameters whose values must be chosen. These parameters are the threshold X applied to the data and the number of bins N_B used in calculating χ^2. It is important to select these parameters carefully as they not only affect the sensitivity of the test but are only valid for ranges of values that depend on the number of samples making up the data being tested.

Number of Samples	Threshold X Range
1,000 samp	1.5 - 1.9
5,000 samp	1.7 - 2.5
10,000 samp	1.9 - 2.8
50,000 samp	2.2 - 3.2
100,000 samp	2.2 - 3.4

Table 3.1. Threshold X ranges for various numbers of samples. The lower bound in each range is the lowest X that results in $E[\chi^2]/\nu$ less than 1.3 when the Poisson test is applied to uncorrelated data of unit variance using 10 bins. The upper bound is the highest X that usually produces enough threshold crossings to perform the Poisson test.

3.4.1 Threshold

The Poisson test threshold X needs to be set high enough so that the intervals between threshold crossings (Δt) are large enough to hide their discrete nature as multiples of the time series' timestep; otherwise the set of Δt for uncorrelated data cannot be accurately modeled as obeying an exponential distribution, which is a continuous probability distribution. On the other hand, X cannot be chosen too high, or the total number of above-X samples will be so small that the test will lack reasonable sensitivity. (The expected number in each bin must be at least 5 for the assumptions underlying the χ^2 test to be valid [51].) Finally, the upper and lower bounds on the range of good threshold values both depend on the number of samples making up the data being tested. Figure 3.3 shows the mean value of χ^2 per degree of freedom ν (whose expectation value is unity for large ν) for different X and numbers of samples. We denote mean χ^2 per degree of freedom as $E[\chi^2]/\nu$. The test was applied to uncorrelated data of unit variance and 10,000 trials were performed for each threshold and number of samples. Ten bins were used in each test. As can be seen from this figure, $E[\chi^2]/\nu$ is close to one only over a certain range of X. Table 3.1 gives X ranges for different numbers of samples. These intervals are bounded below by the smallest X for which $E[\chi^2]/\nu$ was less than 1.3. Table 3.1 may be used as a guide for setting thresholds as long as the data are normalized to have unit variance.

Figure 3.3 may lead one to think that the best threshold is always the highest one possible since higher X result in χ^2 values closest to ν. However, high thresholds give the Poisson test few threshold crossing intervals to work with and

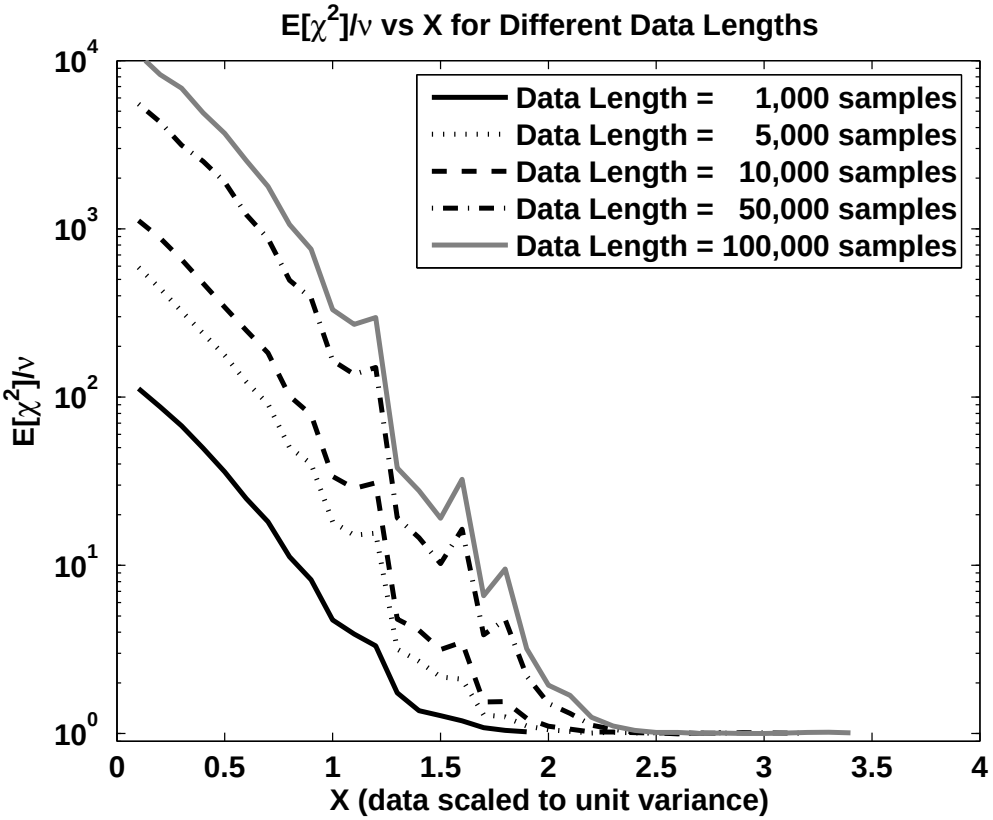


Figure 3.3. Mean χ^2 per degree of freedom ($E[\chi^2]/\nu$) for various thresholds X and numbers of samples. Each mean was calculated over 10,000 trials where the Poisson test was applied to uncorrelated data of unit variance using 10 bins. For each number of samples, there is a range of values where X is high enough that $E[\chi^2]/\nu$ is close to one but still low enough to produce a sufficient number of threshold crossings to perform the test.

its ability to find correlations is reduced. Figure 3.4 shows efficiency versus false alarm rates for the test applied to model data with $k = 5$, $\theta = 0.1\pi$, and 100,000 samples using 10 bins and varying X. The lower X, the higher the efficiency for any false alarm rate. Thus, a lower X increases the Poisson test's ability to detect correlations.

3.4.2 Number of Bins

When choosing the number of bins N_B to use in the Poisson test, one encounters a similar trade-off as when choosing the threshold X. Figure 3.5 shows $E[\chi^2]/\nu$ for

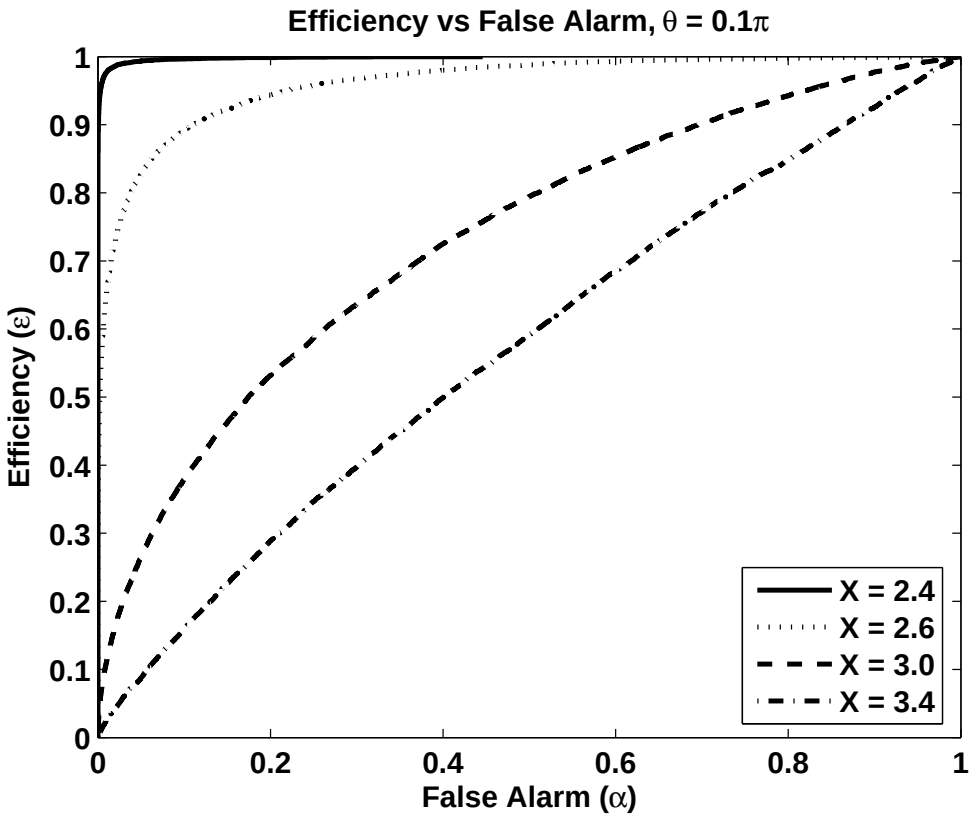


Figure 3.4. False alarm / efficiency probability pairs for the Poisson test using 15 bins and different thresholds. The test was applied to model data with $k = 5$, $\theta = 0.1\pi$, and 100,000 samples in duration.

different N_B. The test is applied to uncorrelated data 100,000 samples in length, using X values between the bounds given in Table 3.1 for this length. As can be seen in Figure 3.5, choosing a smaller N_B will result in a $E[\chi^2]/\nu$ closer to one for uncorrelated data. On the other hand, Figure 3.6 shows that the greater N_B is, the more effective the Poisson test. However, in Figure 3.5, one can see that the increase in χ^2 with more bins is not very dramatic, especially with higher X. This slower increase allows for more flexibility in choosing the N_B than with choosing the Poisson test threshold X.

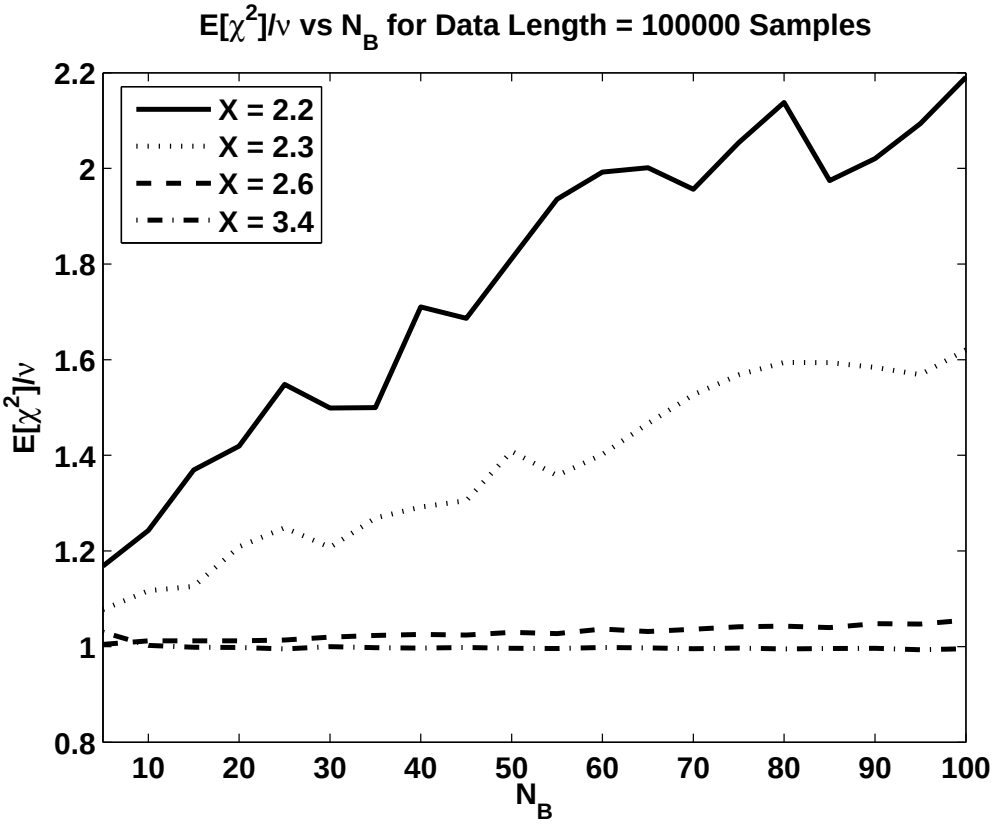


Figure 3.5. $E[\chi^2]/\nu$ for different numbers of bins N_B and thresholds X. Each mean was calculated over 10,000 trials where the Poisson test was applied to uncorrelated data of unit variance and length of 100,000 samples. The smaller N_B, the closer $E[\chi^2]/\nu$ is to one. However, the increase in χ^2 with more bins is not very dramatic, especially with higher X. This allows for more flexibility in choosing N_B than with choosing the test threshold X.

3.5 Setting Confidence Intervals

The confidence interval method, which is discussed in detail by Feldman and Cousins in [52], allows one to compare a given time series with the model in Equation 3.1 to determine the θ most representative of the time series and thus the magnitude of its non-linear couplings. The distribution of Poisson test χ^2 results can be found for model data with different θ values and a k chosen to be representative of the delays in the data (ie. the center of the bin that contributes most to χ^2). These distributions can then be used to find an interval $[\chi_1^2, \chi_2^2]$ such that the probability $P(\chi^2 \in [\chi_1^2, \chi_2^2]|\theta) = \delta$, where δ is the desired confidence. The

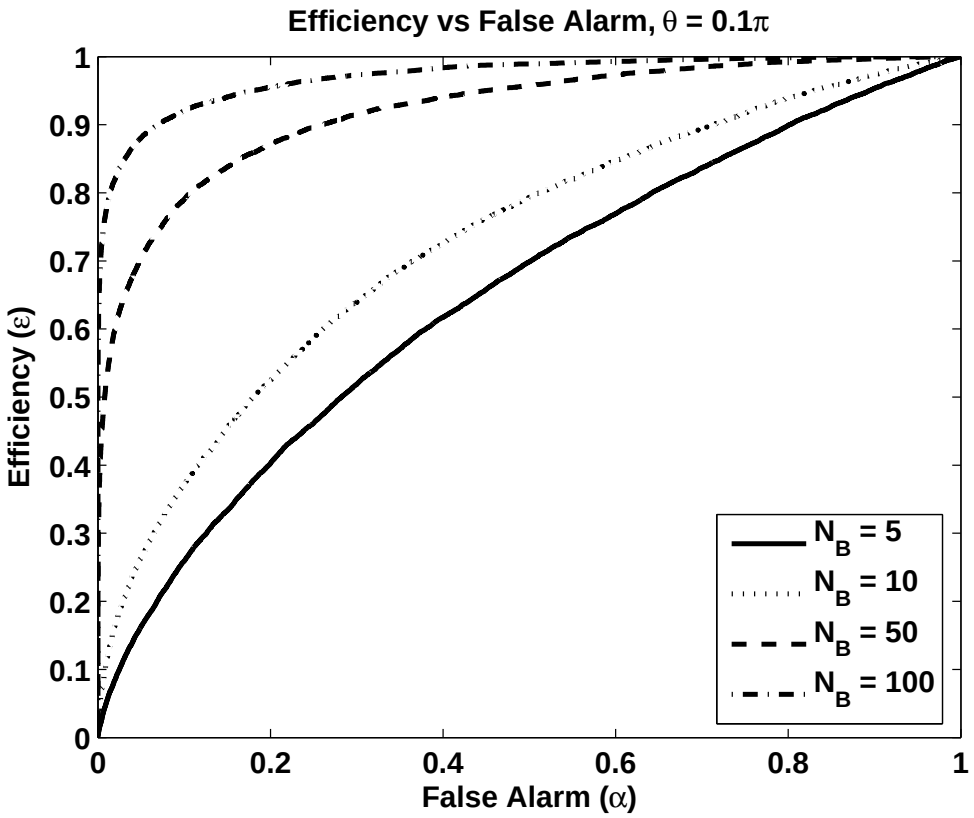


Figure 3.6. False alarm / efficiency probability pairs for the Poisson test using a $X = 3.0$ and different N_B. The test was applied to model data with $k = 5$, $\theta = 0.1\pi$, and 100,000 samples in duration.

interval is centered so that $P(\chi^2 \in [0, \chi_1^2]) = P(\chi^2 \in [\chi_2^2, \infty]) = (1 - \delta)/2$.

For very small values of θ, a different approach must be used, however, since intervals centered in this way will not include $\chi^2 = 0$. A zero χ^2 result does not indicate that the data have correlated samples and therefore should be included in the confidence intervals when θ is very small. For the small θ distributions, a second method for finding intervals should be used which is discussed in [52]. First the probabilities $P(\chi^2|\theta)$ must be found. This can be done by histogramming the distribution of test results, using the same set of bins for each value of θ. Then for each bin (χ^2 value), the maximum probability among the different values of θ

is called $P(\chi^2|\theta_{max})$. Next, the ratio

$$R(\chi^2, \theta) = \frac{P(\chi^2|\theta)}{P(\chi^2|\theta_{max})} \tag{3.10}$$

is calculated. Finally, for each θ, the confidence intervals are found by finding the value R_o such that

$$\delta = \sum_{\chi^2 \ni (R > R_o)} P(\chi^2|\theta). \tag{3.11}$$

The confidence bounds are then the maximum and minimum values of χ^2 for which $R > R_o$.

A plot of θ versus confidence intervals for model data is shown in Figure 3.7 for $N = 10$ and $k = 5$. Now when χ^2 is calculated for a time series with an unknown θ, a vertical line drawn on the confidence interval plot at the calculated value of χ^2 will intersect the confidence interval boundaries at different values of θ. The θ that best describes the time series will fall within this new interval with a probability of δ, the desired confidence.

3.6 Application to LIGO Data

After the test was developed, it was applied to time series data from the 4-km LIGO Livingston detector, acquired during the third science run (S3), which took place from October 31, 2003 to January 9, 2004. The Poisson test cannot be applied to the raw output of the LIGO detectors as it is very nonwhite. Besides having higher power at lower frequencies, there are many narrow-frequency features in the data caused by instrumental (such as resonances) and environmental (such as 60 Hz power lines) effects. The data must first be whitened. The whitening procedure, described in detail in Chapter 2, involves several steps which include breaking up the broadband detector output (with sample rate of 16384 Hz) into smaller frequency bands that are easier to whiten and removing narrow frequency artifacts. The data in these frequency bands are heterodyned or shifted so that the low edge of the band is at zero frequency.

We applied the Poisson test to Livingston detector data in the 128-192 Hz band at times when all three interferometers were in "science mode" and producing

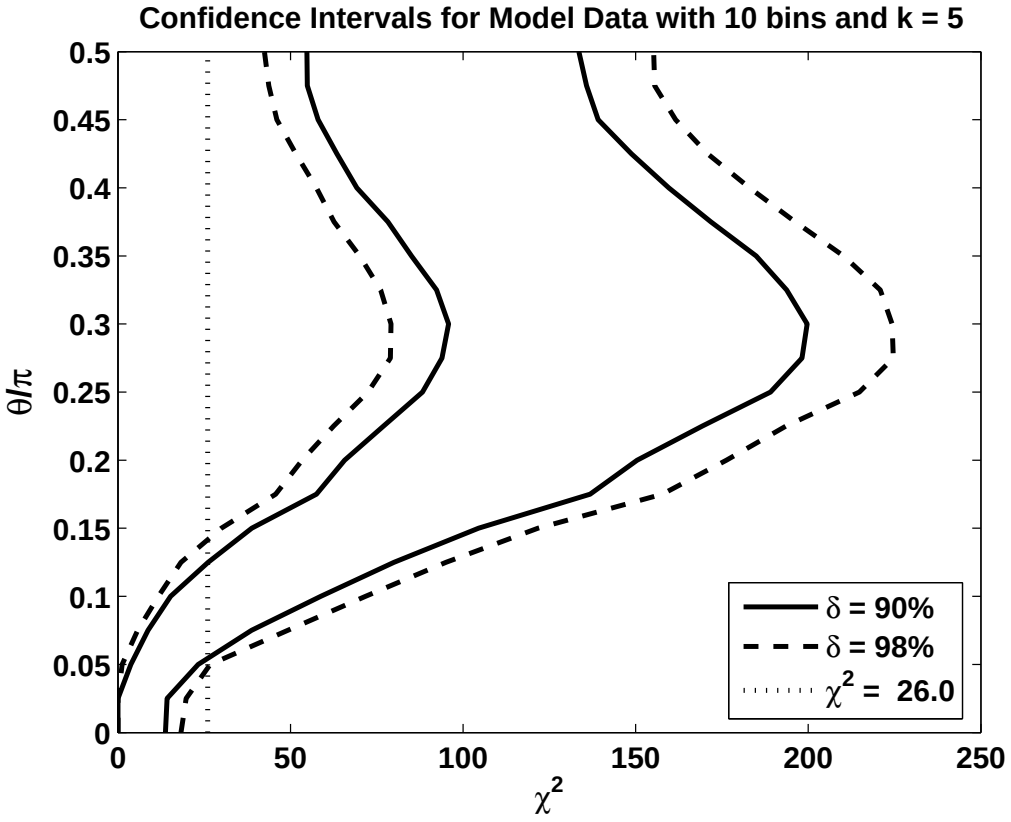


Figure 3.7. Confidence intervals calculated from model data using 10 bins, threshold $X = 2.5$ and $k = 5$ for two different levels of confidence. A vertical line positioned at a χ^2 value will intersect the intervals, giving upper and lower bounds on θ which describes the magnitude of the correlations in the data. For example, if χ^2 of 26 is obtained, the 90% confidence interval for θ is 0.0544π to 0.1247π and the 98% confidence interval is 0.0473π to 0.1415π.

data that passed all data quality tests. Data from these times are those used for scientific analysis. These data were broken up into minute long (7680 samples long) segments. After whitening, the Poisson test was applied to each segment using a threshold $X = 2.3$ and 60 bins. Then, confidence intervals were calculated using model distributions with a k equal to the median bin edge from the bin that contributed most to χ^2 over all segments (in this case $k = 9$).

Out of a total of 4,637 segments, 44.4% had 90% confidence intervals on θ that did not include zero and 36.5% had 98% confidence intervals that did not include zero. The vast majority of these segments had a high χ^2 that was obviously due

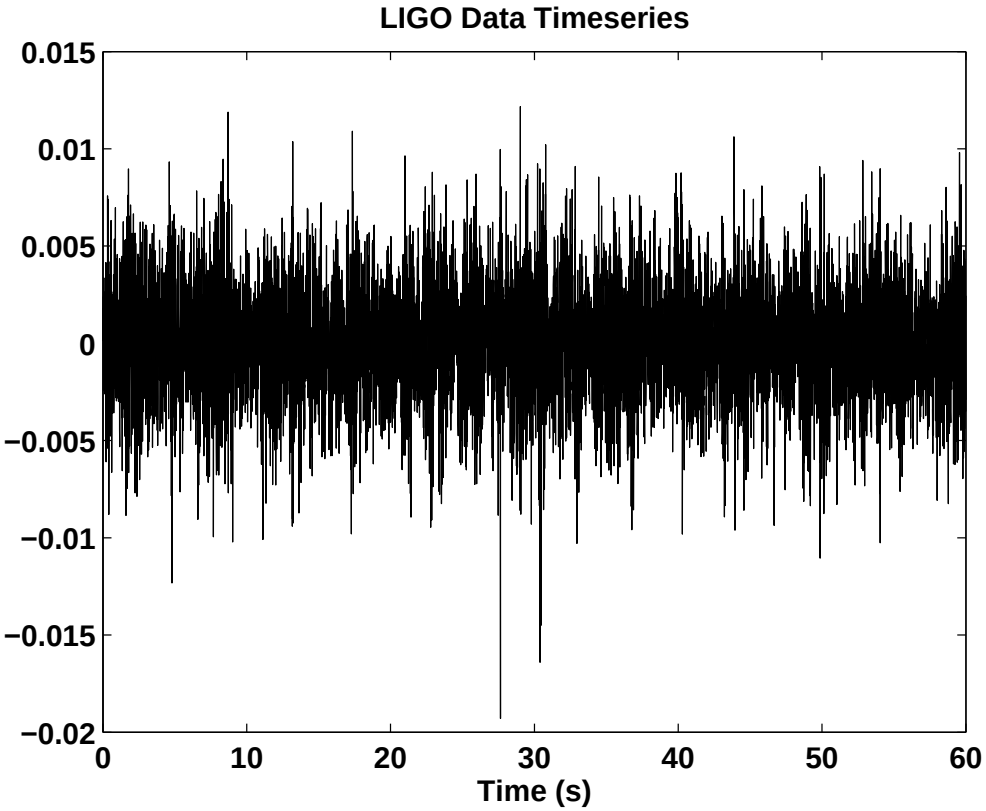


Figure 3.8. Time series of LIGO Livingston data. Data shown are from the frequency band 192-320Hz that has been heterodyned to 0-64Hz and whitened.

to either "glitches" which are high amplitude pulses in the data of instrumental origin or to transient noise features which interfered with the data conditioning procedure. There should be fewer of these types of noise events in the future, as the stability of the detectors improves. There were also, however, some segments with high χ^2 with no glitches or problems with data conditioning. As an example, consider the minute long segment beginning at GPS time 757336870. The time series of the Livingston, 128-192 Hz, whitened data is shown in Figure 3.8 and power spectral densities (PSDs) of unwhitened and whitened data are shown in Figure 3.9. This segment had a $\chi^2 = 108.59$ ($\chi^2/\nu = 1.81$), a 90% confidence interval of 0.0785π to 0.5π on θ and a 98% confidence interval of 0.0548π to 0.5π on θ. A confidence interval plot is shown in Figure 3.10. As shown below, this segment has more than just a high χ^2 value to link it to our nonlinear model.

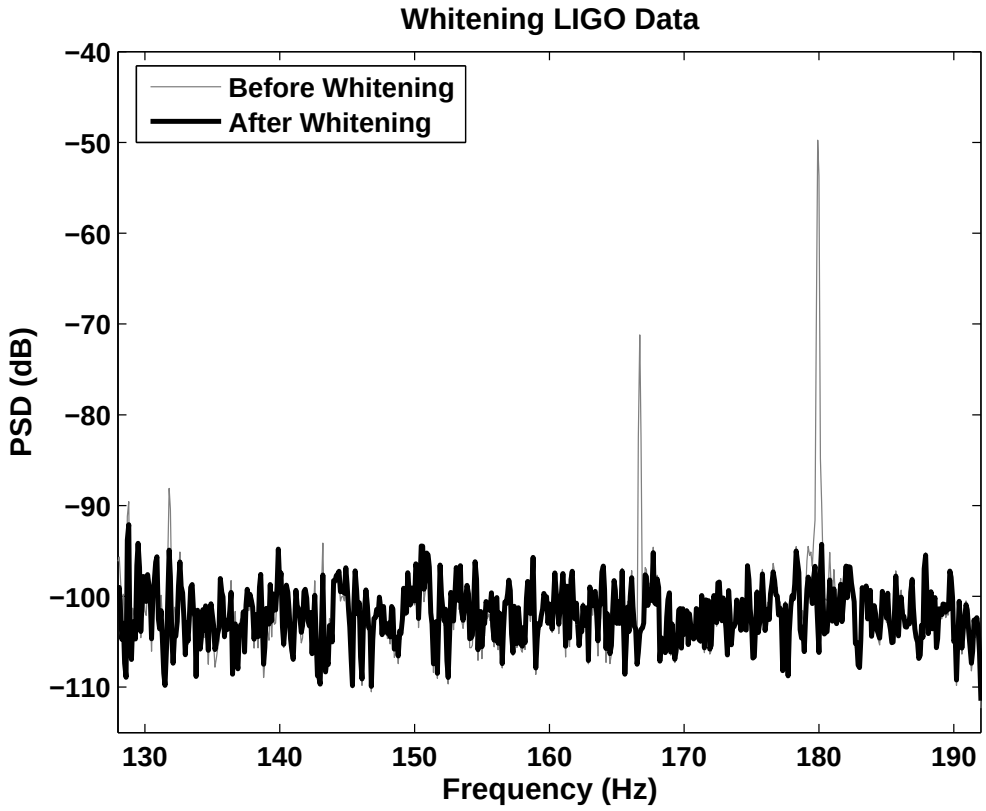


Figure 3.9. Power spectral density (PSD) of LIGO data in the 128-192Hz band before and after whitening. The "before whitening" PSD is from data that have only been restricted to the 128-192Hz band and heterodyned while the "after whitening" PSD is from data that have undergone the full data conditioning procedure including removal of narrow features (lines) and being passed through a whitening filter (see Chapter 2). The most prominent features removed are the line at 166.70 Hz which is caused by intentionally moving mirrors at this frequency for instrument calibration purposes and the line at 180 Hz which is a harmonic of the 60 Hz power oscillations. The heterodyned data actually go from 0-64 Hz but the frequency axis has been relabeled to make identifying frequency features easier.

3.6.1 Bispectra

A common graphical approach for determining the higher-order nature of a time series is to take its higher order spectra [53], [54], [55]. These spectra are simply the higher dimensional analogs of the power spectrum. The most commonly used is the third order bispectrum since the difficulty of calculation increases with order. The power spectrum can be calculated by taking the Fourier transform of the

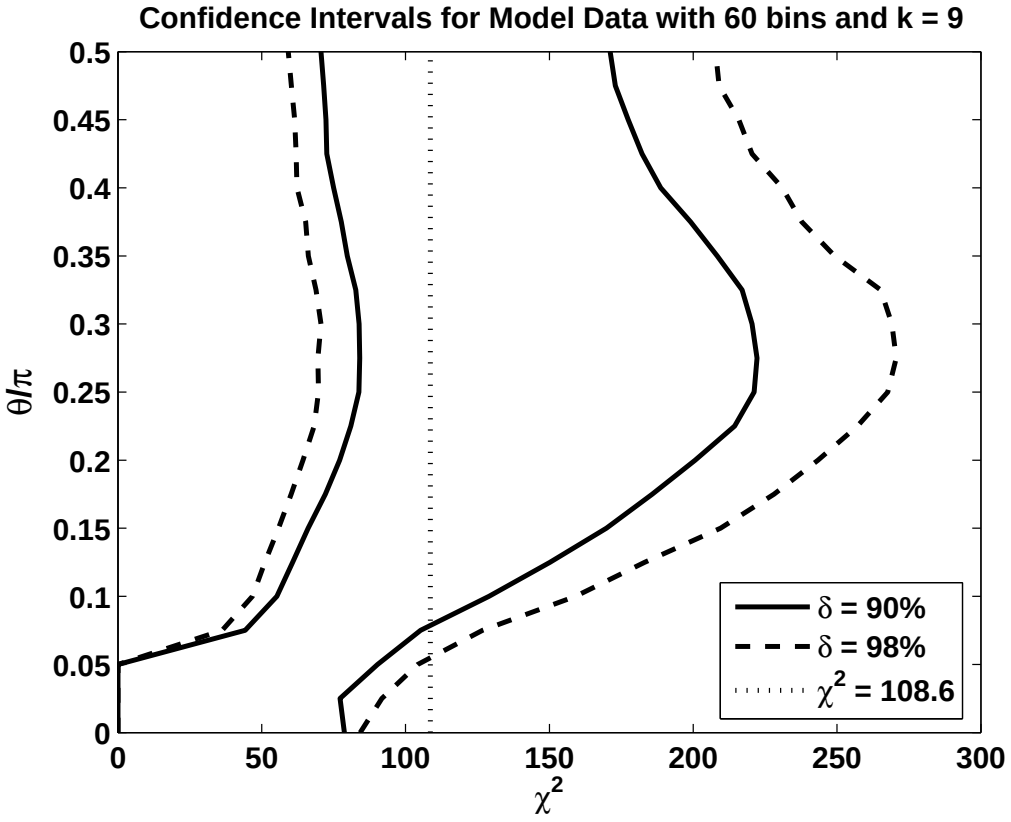


Figure 3.10. Confidence intervals calculated from model data with $k = 9$ using 60 bins. The χ^2 value of an example segment of LIGO data is 108.59 which corresponds to bounds on θ of 0.0785π - 0.5π at 90% confidence and 0.0548π - 0.5π at 98% confidence.

autocorrelation function. Likewise, the bispectrum can be calculated from the third order cumulant function:

$$C(\tau_1, \tau_2) = \sum_t x(t)x(t+\tau_1)x(t+\tau_2), \tag{3.12}$$

where $x(t)$ is the mean-subtracted time series. The bispectrum is then the two dimensional Fourier transform of the third order cumulant.

$$B(\omega_1, \omega_2) = \int_{\tau_1}\int_{\tau_2} C(\tau_1, \tau_2)e^{-i\omega_1\tau_1}e^{-i\omega_2\tau_2}d\tau_1 d\tau_2. \tag{3.13}$$

Using the above expressions, we can analytically calculate the bispectrum of our model time series. A non-zero bispectrum confirms that the model data deviate

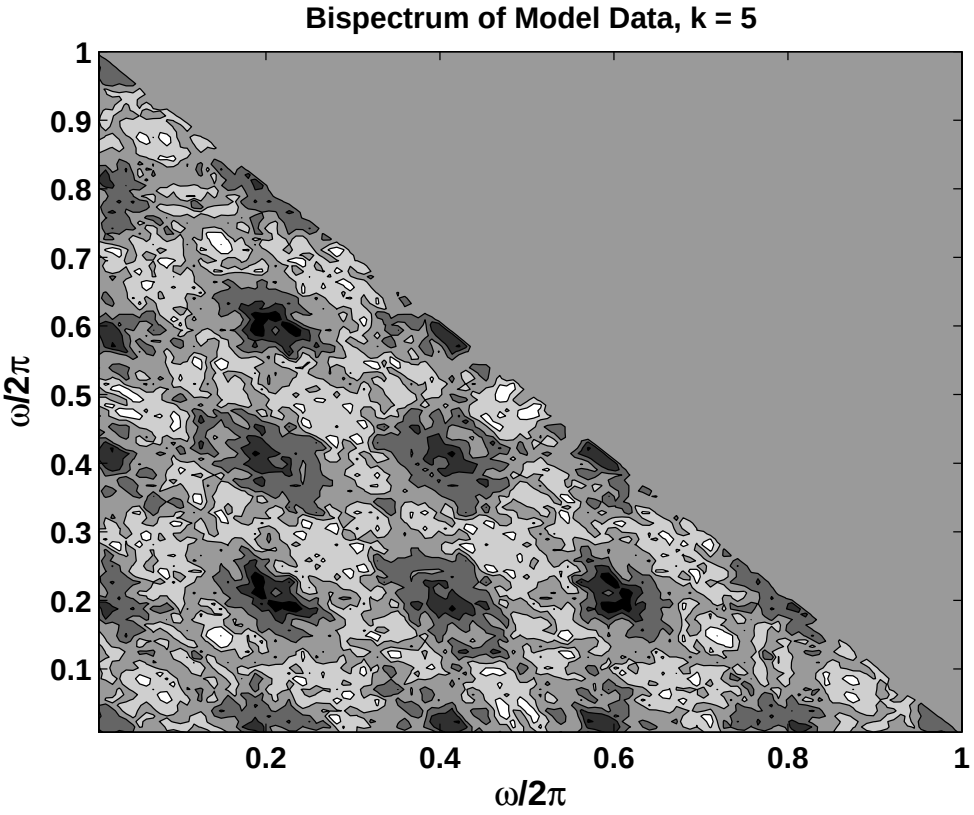


Figure 3.11. Real part of the bispectrum of model data with $\theta = \pi/4$ and $k = 5$. A frequency of 5 can be clearly seen.

from being uncorrelated. Furthermore, a similarity between the bispectrum of the model data and the bispectrum of the LIGO data would indicate that the model accurately describes the third order behavior of the LIGO data. Inserting the model of Equation 3.1 into Equations 3.12 and 3.13 results in the expression:

$$\begin{aligned} B(\omega_1, \omega_2) &= 2 \sin\theta \cos^2\theta [\cos\omega_1 k + \cos\omega_2 k + \cos(\omega_1 + \omega_2)k \\ &\quad -i(\sin\omega_1 k + \sin\omega_2 k - \sin(\omega_1 + \omega_2)k)]. \end{aligned} \tag{3.14}$$

In a plot of the bispectrum of randomly generated model data, one should see the sines and cosines with frequencies equal to k. Such a plot is shown in Figure 3.11 for $k = 5$.

The bispectrum of the LIGO data from the 128-192 Hz band of the Livingston

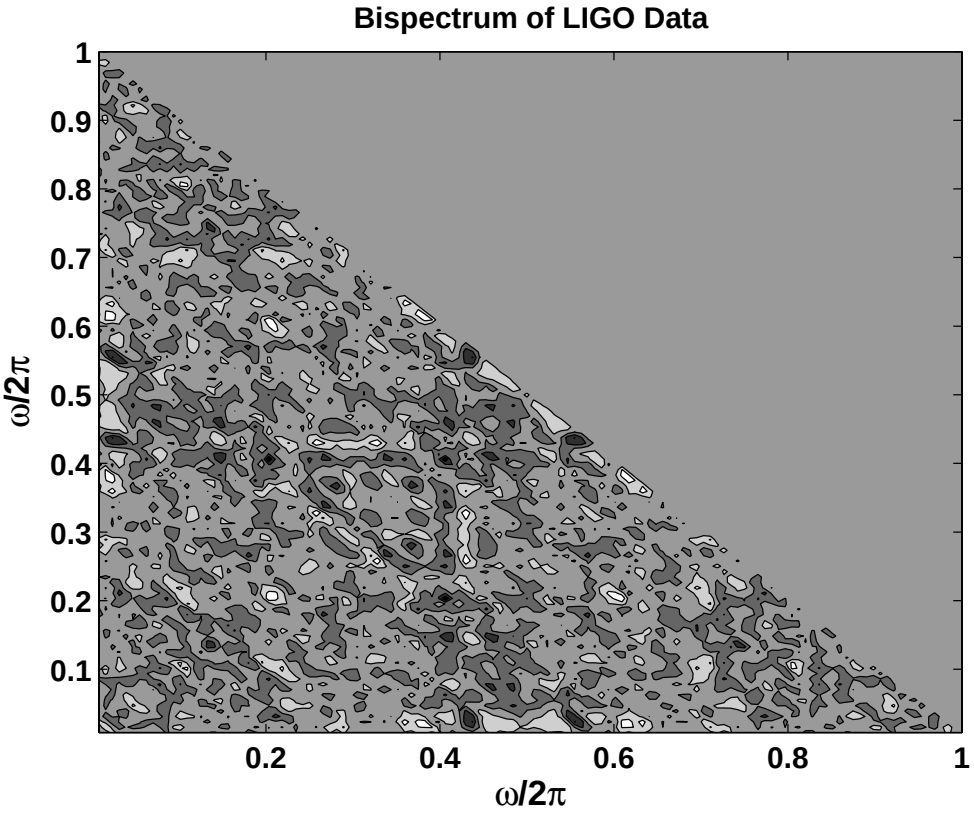


Figure 3.12. Real part of the bispectrum of 128-192 Hz data from the LIGO Livingston detector. There is a frequency of 2 present in the bispectrum, just as there is in the bispectrum of model data with $k = 2$, shown in Figure 3.13.

detector taken at GPS time 757556870 and analyzed above is shown in Figure 3.12. Looking carefully, one can see a pattern with a frequency of 2. The bispectrum of model data with $k = 2$ shown in Figure 3.13 has the same pattern. Thus we conclude that these data contains a nonlinear component with a delay of two.

The bispectra shown in Figures 3.11, 3.12, and 3.13 give far more detail than is present in the single χ^2 value of the Poisson test for correlations. However, this test has an important advantage over the bispectra besides a low computational cost. One can see from Equation 3.14 when $\theta \to \pi/2$ that $B(\omega_1, \omega_2) \to 0$. Therefore, when $\tan\theta$ is large, the bispectrum cannot give any information about the nonlinearity. The test for correlations, on the other hand, is sensitive to correlations of any type.

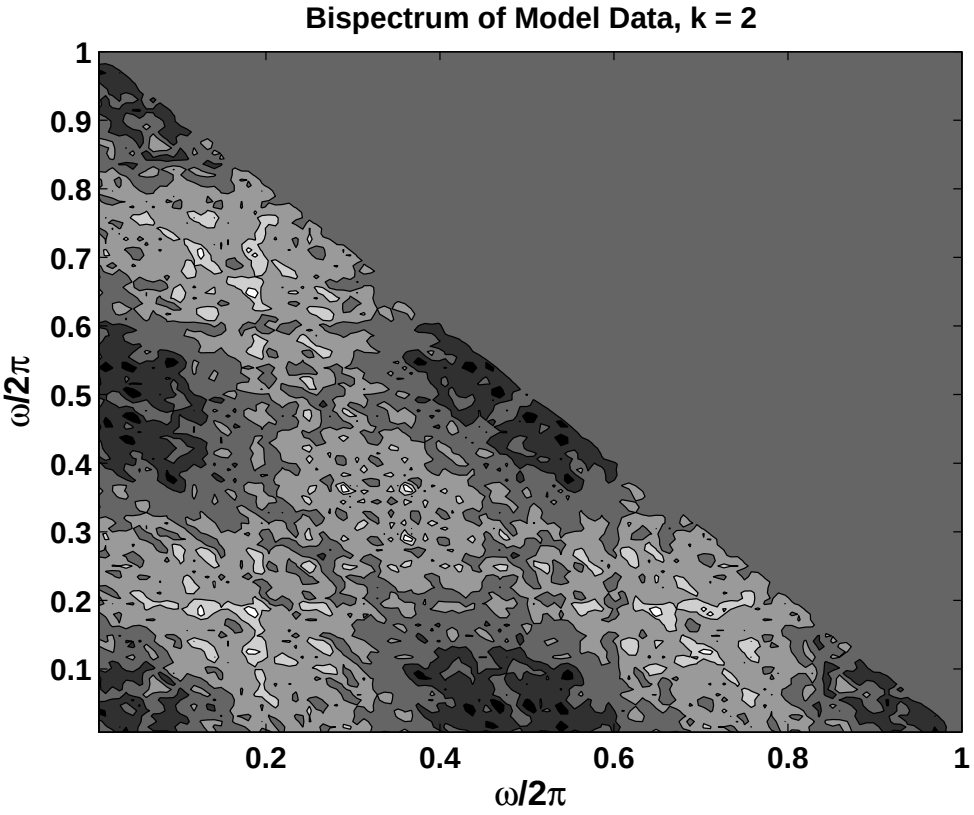


Figure 3.13. Real part of the bispectrum of model data with $\theta = \pi/4$ and $k = 2$. A frequency of 2 can be clearly seen.

3.7 Conclusion

Using the Poisson test to determine the statistical nature of above threshold data points and so determining if there are correlations present is a good way of detecting certain non-linear couplings. Higher order tools such as the bispectrum may give more detail concerning the nonlinearity but are computationally expensive. The Poisson test is quite inexpensive and therefore can be applied often and used to provide warning that non-linear couplings are present which merit further investigation.

Chapter 4

Supernovae

4.1 Introduction

The gravitational waves that we expect to observe in large detectors like LIGO [16] and VIRGO [19] reflect the coherent evolution of the most compact part of the source mass distribution. From the observed waves, we have the potential to decode the forces that govern that evolution.[1] For example, the evolution of the collapsing stellar core in a type II supernova is determined by the progenitor mass, angular momentum and its distribution throughout the core, and magnetic field. None of these properties of the progenitor can be determined by the electromagnetic radiation that we observe when the shock emerges from the stellar envelope. In addition, the forces that govern the evolution of the core depend upon many things that are unknown, or poorly modeled, by us today: for example, the matter equation of state, the role played by neutrinos and neutrino transport, general relativity, convection and non-axisymmetry. The gravitational waves emitted during the collapse and its aftermath carry the signature of all these properties of the collapsing star and to read that signature, we must learn to recover the waveform from the noisy detector observations. Here we develop a maximum entropy based method for recovering the gravitational radiation waveform from the noisy data from two or more detectors, demonstrate its effectiveness when applied to signals

[1] Or, if we assume we understand the forces — as in the case of a coalescing neutron star or black hole binary system — then we can discern the bulk properties of the mass distribution itself: e.g., the component masses and spins of the binary system.

arising from core-collapse simulated supernova buried in simulated detector noise, and show how the recovered waveform can be used to gain insight into the properties of the progenitor star and the forces acting on the collapsing stellar core.

The first problem addressed in this study is the problem of extracting the gravitational wave signal from the data produced by the detectors (also known as the deconvolution problem). The detection process not only significantly modifies the waveforms due to the projection of the gravitational wave polarizations onto each detector and a detector's varying sensitivity at different frequencies, but also adds a significant amount of noise. It is a problem to be approached with care. A naïve inversion could cause the noise to completely overwhelm a faint signal, especially for those frequencies at which the responses of the detectors are weak. The problem of inverting the responses of the detectors to recover the original gravitational wave signal is discussed in detail in Gürsel & Tinto [56], where a filter is developed which inverts the projection of the gravitational wave onto the detectors but not the varying frequency responses. The maximum entropy method that we apply to the problem of deconvolution has been often employed in astronomy. Maximum entropy is commonly used in astronomical image reconstruction, a recent example being its application to CMB maps [57, 58, 59] including maps based on data from WMAP [60] and COBE [61, 62]. We extend the method from image analysis to recovering signals from time series data produced by gravitational wave interferometers such as LIGO.

Once a gravitational wave signal is deconvolved from the data, a second problem is to be able to deduce the properties of the source from the signal. One way of associating gravitational waveforms with supernova properties is to compare a detected signal with theoretical waveforms from simulations using varying physical parameters. One would expect the detected signal to have the most in common with a waveform from a collapse model with similar character to the actual supernova. One example of this type of study, described in [63], creates a simulation waveform vector subspace with which detected signals are compared. We use the correlation between a recovered signal and a catalog of waveforms resulting from models with differing progenitor characteristics as a qualitative measure of the properties of the source.

This chapter describes a study that investigates the types of information that

can be obtained from a LIGO detection of gravitational waves emitted by a core-collapse supernova. Section 4.2 describes the maximum entropy method chosen to recover a common signal detected by multiple interferometers and Section 4.3 compares maximum entropy with other methods. Section 4.4 describes a recently produced catalog of core-collapse supernova waveforms based on computer simulations and the physics that went into these simulation models. We simulate a LIGO detection using one of these waveforms and demonstrate that, using maximum entropy and comparing the recovered signal with the waveform catalog, we are able to determine several physical parameters of the simulation model that produced the detected waveform, provided that the gravitational waves are of sufficient strength. Finally, in Section 4.7 we summarize what was learned as well as directions for further study.

4.2 Maximum Entropy Signal Extraction

Focus attention on a gravitational wave burst, with time dependent polarizations $\mathbf{h}_+(t)$ and $\mathbf{h}_\times(t)$, incident from a direction $\mathbf{s}$ on a set of detectors. (Bold upper and lowercase letters represent matrices and vectors, respectively.) The response of each detector is a linear transformation of the $\mathbf{h}_+$ and $\mathbf{h}_\times$. The details of the transformation depend on the direction of propagation of the waves relative to the orientation of the detector, the details of the detector, and is different for the different wave polarizations. We can write the output $\mathbf{d}_k$ of detector k as the sum of detector noise $\mathbf{n}_k$ with the convolution of the signals $\mathbf{h}_+$ and $\mathbf{h}_\times$ with the (time independent) response functions $\mathbf{R}_{k,+}$ and $\mathbf{R}_{k,\times}$:

$$\mathbf{d}_k(t) = \mathbf{R}_{k,+}(\mathbf{s})\mathbf{h}_+(t) + \mathbf{R}_{k,\times}(\mathbf{s})\mathbf{h}_\times(t) + \mathbf{n}_k(t). \tag{4.1}$$

Symbolically, then, we can write the output $\mathbf{d}$ of a network of detectors as the sum of detector noise $\mathbf{n}$ with a linear transformation of an incident gravitational wave signal $\mathbf{h}_I$:

$$\mathbf{d} = \mathbf{R}(\mathbf{s})\mathbf{h}_I + \mathbf{n}. \tag{4.2}$$

Our goal is to infer $\mathbf{h}_I$ given $\mathbf{d}$, $\mathbf{R}(\mathbf{s})$ and the statistical properties of the detector noise $\mathbf{n}$.

Deconvolution of a noisy signal is a challenging problem [64]. The response function $\mathbf{R}$ is often ill-conditioned, and may not even be invertible. One very successful approach to the problem of deconvolution follows from Jaynes' principle of maximum entropy [65, 66]. Let $P(\mathbf{h}|\mathbf{d},\mathbf{R},I)$ be the probability density that $\mathbf{h}$ is the incident signal given our knowledge of $\mathbf{d}$, $\mathbf{R}$, and other prior information I, which includes the statistical properties of the noise $\mathbf{n}$. Using Bayes' Law we can write this probability density as

$$P(\mathbf{h}|\mathbf{d},\mathbf{R},I) \propto P(\mathbf{d}|\mathbf{h},\mathbf{R},I)P(\mathbf{h}|I). \tag{4.3}$$

The probability density $P(\mathbf{d}|\mathbf{h},\mathbf{R},I)$ is the probability of the data under the hypothesis $\mathbf{h}$. The probability density $P(\mathbf{h}|I)$ describes how likely we think a given $\mathbf{h}$ before we have taken any data.

Without loss of generality, assume that the noise has zero mean and covariance $\mathbf{C}$. If the noise is Gaussian, or if the covariance $\mathbf{C}$ is all the knowledge we have of the noise statistics, then we can write

$$P(\mathbf{d}|\mathbf{h},\mathbf{R},I) \propto \exp\left[-\frac{1}{2}\left(\mathbf{d}-\mathbf{R}\mathbf{h}\right)^T \mathbf{C}^{-1}\left(\mathbf{d}-\mathbf{R}\mathbf{h}\right)\right] \tag{4.4a}$$

$$= \exp\left[-\frac{1}{2}\chi^2(\mathbf{h},\mathbf{R},\mathbf{C},\mathbf{d})\right]. \tag{4.4b}$$

In general terms, Gürsel & Tinto [56] take as their estimate $\mathbf{h}_E$ of $\mathbf{h}_I$ the $\mathbf{h}$ that maximizes $\chi^2(\mathbf{h},\mathbf{R},\mathbf{C},\mathbf{d})$: i.e., that maximizes the likelihood given the data. In addition to sometimes being poorly behaved (owing to the properties of $\mathbf{R}$) the resulting $\mathbf{h}_E$ will generally over-fit the data: i.e., the resulting χ^2 will be lower than expected, given the number of degrees of freedom $\dim(\mathbf{d}) - \dim(\mathbf{h})$.

Maximizing the likelihood amounts to finding the $\mathbf{h}$ for which $\mathbf{d}$ is the most probable detector output. Our interest, on the other hand, is in the $\mathbf{h}$ that is most probable given the detector output $\mathbf{d}$. To find that $\mathbf{h}$ we need to maximize $P(\mathbf{h}|\mathbf{d},\mathbf{R},I)$, which requires specifying $P(\mathbf{h}|I)$.

Jaynes' principle of maximum entropy [65, 66] motivates a particular choice of $P(\mathbf{h}|I)$. Suppose that $\mathbf{h}_I$ is composed by distributing a collection of positive and negative "quanta" of amplitude $1/\alpha$ among the different elements h_i of $\mathbf{h}$, with the value of h_i, the i^{th} component of $\mathbf{h}$, equal to the number of positive quanta minus

the number of negative quanta. If we suppose that the probability of a positive or negative quanta being distributed into bin i is independent of i, then it is natural to assume that the probability $P(\mathbf{h}|I)$ of observing signal $\mathbf{h}$ is proportional to the number of different ways in which the quanta can be distributed and yield $\mathbf{h}$. The log of $P(\mathbf{h}|I)$ is then related to the Shannon information entropy [67] associated with the probability distribution $P(\mathbf{h}|I)$ [68]:

$$P(\mathbf{h}|I) \propto \exp\left[\alpha S(\mathbf{h},\mathbf{m})\right] \tag{4.5a}$$

$$S(\mathbf{h},\mathbf{m}) = \sum_i s(h_i, m_i) \tag{4.5b}$$

$$s(h_i, m_i) = \Psi_i - 2m_i - h_i \log \frac{\Psi_i + h_i}{2m_i}, \tag{4.5c}$$

where $h_i(m_i)$ are the scalar components of $\mathbf{h}(\mathbf{m})$,

$$\Psi_i = \left(h_i^2 + 4m_i^2\right)^{1/2}, \tag{4.5d}$$

and m_i is the expected number of positive and negative quanta in bin i. (I.e., on average we expect m_i positive quanta and m_i negative quanta, for a total h_i of zero.)

Setting aside for the moment the assignment of $\mathbf{m}$ and α, the maximum entropy choice of $P(\mathbf{h}|I)$ leads us to choose as the best estimate $\mathbf{h}_E$ of $\mathbf{h}_I$ the $\mathbf{h}$ that minimizes

$$F(\mathbf{h}|\mathbf{d},\mathbf{C},\alpha,\mathbf{m}) = \frac{1}{2}\chi^2(\mathbf{h},\mathbf{R},\mathbf{C},\mathbf{d}) - \alpha S(\mathbf{h},\mathbf{m}). \tag{4.6}$$

Stripped of its Bayesian statistical motivation, we recognize that maximum entropy deconvolution amounts to the solution of the inverse problem using regularization methods with a specific regularization function ($\alpha S(\mathbf{h},\mathbf{m})$).

Different choices of α and $\mathbf{m}$ will lead to different $\mathbf{h}$ that minimize the function $F(\mathbf{h}|\mathbf{d},\mathbf{C},\alpha,\mathbf{m})$. Intuitively, too small an α will lead to over-fitting of the noisy data and too large an α will lead to a solution $\mathbf{h}$ that approaches 0 (i.e., the $\mathbf{h}$ that maximizes the entropy functional $S(\mathbf{h},\mathbf{m})$). Gull [69] describes a Bayesian approach to the determination of α that relates the curvature of F at its minimum (which is in turn related to the uncertainty in the parameter values h_i) to the number of parameters being sought (i.e., $\dim \mathbf{h}$). MacKay [70] finds a simple

approximation, which is to seek α, such that, when F is minimized, $F = \dim(\mathbf{h})/2$.

An appropriate choice of the m_i can be made by following a procedure described by [70]. Using Bayes' Law, we can write the probability of $\mathbf{m}$ given $\mathbf{d}$ as

$$P(\mathbf{m}|\mathbf{d}, I) \propto P(\mathbf{d}|\mathbf{m}, I)P(\mathbf{m}|I), \tag{4.7}$$

where I is any other prior information, as before. If we assume no prior preference for any particular choice of $\mathbf{m}$, $P(\mathbf{m}|I)$ is constant so the most probable $\mathbf{m}$ will be the one that maximizes $P(\mathbf{d}|\mathbf{m}, I)$. At the same time, from Bayes' Law

$$P(\mathbf{h}|\mathbf{d}, \mathbf{m}, I)P(\mathbf{d}|\mathbf{m}, I) = P(\mathbf{d}|\mathbf{h}, \mathbf{m}, I)P(\mathbf{h}|\mathbf{m}, I). \tag{4.8}$$

Integrating this expression over all $\mathbf{h}$ gives

$$P(\mathbf{d}|\mathbf{m}, I) = \int D\mathbf{h} P(\mathbf{d}|\mathbf{h}, \mathbf{m}, I)P(\mathbf{h}|\mathbf{m}, I). \tag{4.9}$$

Now, according to Equations 4.4b and 4.5a,

$$P(\mathbf{d}|\mathbf{h}, \mathbf{m}, I) = \frac{\exp\left(-\frac{1}{2}\chi^2\right)}{\int D\mathbf{d} \exp\left(-\frac{1}{2}\chi^2\right)} \tag{4.10a}$$

$$P(\mathbf{h}|\mathbf{m}, I) = \frac{\exp\left(\alpha S\right)}{\int D\mathbf{h} \exp\left(\alpha S\right)}. \tag{4.10b}$$

Therefore, the most probable $\mathbf{m}$ is found by evaluating

$$P(\mathbf{d}|\mathbf{m}, I) = \frac{\int D\mathbf{h} \exp\left(-\frac{1}{2}\chi^2 + \alpha S\right)}{\int D\mathbf{d} \exp\left(-\frac{1}{2}\chi^2\right) \int D\mathbf{h} \exp\left(\alpha S\right)} \tag{4.11a}$$

$$= \frac{Z_F}{Z_d Z_h}. \tag{4.11b}$$

When evaluating Z_F, it is helpful to realize that $(-1/2)\chi^2 + \alpha S$ is the negative of the functional F in Equation 4.6 which we minimize to find the best estimate $\mathbf{h}_E$ of $\mathbf{h}_I$. Therefore, we can simplify the integration by performing a Taylor expansion about $\mathbf{h}_E$

$$Z_F = \int D\mathbf{h} \exp\left[-F(\mathbf{h}_E) + \frac{1}{2}(\mathbf{h} - \mathbf{h}_E)^T \cdot H(-F)|_{\mathbf{h}_E} \cdot (\mathbf{h} - \mathbf{h}_E) + \ldots\right], \tag{4.12}$$

where $H(F)|_{\mathbf{h}_E}$ is the Hessian of F with respect to $\mathbf{h}$, evaluated at $\mathbf{h} = \mathbf{h}_E$. Ignoring higher order terms and performing the multivariate Gaussian integration then gives

$$Z_F \approx \exp\left[-F(\mathbf{h}_E)\right](2\pi)^{N_h/2}\left(\det\left\|H(F)|_{\mathbf{h}_E}\right\|\right)^{-1/2}, \tag{4.13}$$

where N_h is the number of elements of $\mathbf{h}$. The direct calculation of the determinant of a large matrix is prone to being overwhelmed by numerical errors so we instead calculate the log of Z_F

$$\log Z_F \approx -F(\mathbf{h}_E) + \frac{N_h}{2}\log(2\pi) - \frac{1}{2}\mathrm{logdet}\left\|H(F)|_{\mathbf{h}_E}\right\|. \tag{4.14}$$

To calculate the log determinant, we decompose the matrix into the form $\mathbf{U}^T\mathbf{A}(\mathbf{I}-\mathbf{D})\mathbf{A}\mathbf{U}$, where $\mathbf{U}$ and $\mathbf{A}$ are upper-triangular and diagonal matrices, respectively, $\mathbf{I}$ is the identity matrix, and $\mathbf{D}$ is a sparse matrix with eigenvalues in $[-1,1]$. We then evaluate the log determinant of $(\mathbf{I}-\mathbf{D})$ using the Monte Carlo estimate method of [71].

Following a similar approach in evaluating Z_d we find

$$Z_d \approx \exp\left[-\frac{1}{2}\chi^2(\mathbf{d}_0)\right](2\pi)^{N_d/2}\left(\det\left\|H\left(\chi^2/2\right)|_{\mathbf{d}_0}\right\|\right)^{-1/2}, \tag{4.15}$$

where $H\left(\chi^2/2\right)|_{\mathbf{d}_0}$ is the Hessian of $\chi^2/2$ with respect to $\mathbf{d}$, evaluated at $\mathbf{d} = \mathbf{d}_0$ and N_d is the number of elements in $\mathbf{d}$. $\mathbf{d}_0$ is the $\mathbf{d}$ that maximizes $-\chi^2/2$, or $\mathbf{d}_0 = \mathbf{R}\mathbf{h}$ and $\chi^2(\mathbf{d}_0) = 0$. Also, $H\left(\chi^2/2\right) = \mathbf{C}^{-1}$, the inverse of the covariance matrix $\mathbf{C}$. If we compute the Cholesky decomposition of the covariance matrix $\mathbf{C} = \mathbf{C}_d^T\mathbf{C}_d$, where $\mathbf{C}_d$ is an upper triangular matrix, the log of Z_d can be simply calculated as

$$\log Z_d \approx \frac{N_d}{2}\log(2\pi) + \mathrm{sum}(\log(\mathrm{diag}(\mathbf{C}_d))), \tag{4.16}$$

where the second term on the right is found by taking the log of each element along the main diagonal of $\mathbf{C}_d$ and summing.

Finally, evaluating Z_h, we find

$$Z_h \approx \exp\left[\alpha S(\mathbf{h}_0)\right](2\pi)^{N_h/2}\left(\det\left\|H\left(-\alpha S\right)|_{\mathbf{h}_0}\right\|\right)^{-1/2}, \tag{4.17}$$

where $H\left(-\alpha S\right)|_{\mathbf{d}_0}$ is the Hessian of $-\alpha S$ with respect to $\mathbf{h}$, evaluated at $\mathbf{h} = \mathbf{h}_0$.

$\mathbf{h}_0$ is the $\mathbf{h}$ that maximizes αS, or $\mathbf{h}_0 = \mathbf{0}$ and $S(\mathbf{h}_0) = 0$. At this value of $\mathbf{h}$, the Hessian $H\left(-\alpha S\right)|_{\mathbf{h}_0}$ is equal to a diagonal matrix. If we assume all of the elements of $\mathbf{m}$ are the same ($m_i = m$ for all i), each nonzero element of this diagonal matrix is equal to $\alpha/2m$ and

$$\log Z_h \approx \frac{N_h}{2}\left[\log(2\pi) - \log\left(\alpha/2m\right)\right]. \tag{4.18}$$

We assume no prior knowledge about the signal and, therefore, have no information about the distribution of the positive and negative quanta that make up the signal and determines m. The maximum entropy best estimate $\mathbf{h}_E$ does not change quickly with changing m but remains essentially the same as m ranges over a few orders of magnitude. We determine the appropriate model by finding $\mathbf{h}_E$ for m of different orders of magnitude and select the one for which the log evidence, $\log P(\mathbf{d}|\mathbf{m}, I) = \log Z_F - \log Z_d - \log Z_h$, is greatest.

4.3 Comparing Maximum Entropy With Other Methods

To demonstrate the effectiveness of the maximum entropy method, we compare its ability to reconstruct a signal with that of two simple deconvolution approaches: zero-forcing (ZF) and minimum mean square error (or Weiner) filtering (MMSE). ZF and MMSE assume that the signal to be recovered has been convolved with a response function and noise added to create the data from each detector (d_k)

$$\mathbf{d}_k = \mathbf{r}_k * \mathbf{h}_I + \mathbf{n}, \tag{4.19}$$

where $*$ denotes convolution and $\mathbf{r}_k$ is a vector describing the response of a single detector. ZF ignores the presence of noise entirely and in the frequency domain can be implemented by simply dividing the data by the response:

$$\tilde{h}_{ZF}(\omega) = \frac{\tilde{d}_k(\omega)}{\tilde{r}_k(\omega)}, \tag{4.20}$$

where tilde denotes Fourier transform. MMSE, on the other hand, takes the amount of noise into account and recovers the signal via

$$\tilde{h}_{MMSE}(\omega) = \left[\frac{1}{\tilde{r}_k(\omega)}\frac{|\tilde{r}_k(\omega)|^2}{|\tilde{r}_k(\omega)|^2 + K}\right]\tilde{d}_k(\omega), \tag{4.21}$$

where K is a constant that we set equal to the rms of the noise divided by the variance of the signal [64]. This form of MMSE assumes that one knows the variance of the signal but does not have access to detailed information about the power spectral density (PSD) of the signal. It is therefore optimal for white signals only, but is the best we can do with limited prior information about the signal.

We use these three signal recovery methods: maximum entropy, ZF, and MMSE to recover two different signals. The first signal is a Gaussian distributed, white, random signal and the second is a supernova waveform signal described in Section 4.4.2. The two signals were convolved with two simple response filters $\mathbf{r}_1 = [\ 1 \quad -2 \quad 2\ ]^T$ and $\mathbf{r}_2 = [\ 1 \quad 2 \quad 2\ ]^T$ and added to Gaussian distributed, random noise. For each initial signal, the two data streams can be written as

$$\begin{aligned} \mathbf{d}_1 &= \mathbf{r}_1 * \mathbf{h}_I + \mathbf{n}_1 \\ \mathbf{d}_2 &= \mathbf{r}_2 * \mathbf{h}_I + \mathbf{n}_2, \end{aligned} \tag{4.22}$$

or in terms of the maximum entropy response function array $\mathbf{R}$

$$\mathbf{d} = \begin{pmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \end{pmatrix} = \begin{pmatrix} \mathrm{diag}(\mathbf{r}_1) \\ \mathrm{diag}(\mathbf{r}_2) \end{pmatrix}(\mathbf{h}_I) + \begin{pmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \end{pmatrix} = \mathbf{R}\mathbf{h}_I + \mathbf{n}, \tag{4.23}$$

where, for this example,

$$\mathrm{diag}(\mathbf{r}_1) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots \\ -2 & 1 & 0 & 0 & 0 & \cdots \\ 2 & -2 & 1 & 0 & 0 & \cdots \\ 0 & 2 & -2 & 1 & 0 & \cdots \\ 0 & 0 & 2 & -2 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{4.24}$$

Both ZF and MMSE invert only one data stream at a time to recover the initial

signal $\mathbf{h}_I$. The estimates based on $\mathbf{d}_1$ and $\mathbf{d}_2$ separately are added together to form the final estimates $\mathbf{h}_{ZF}$ and $\mathbf{h}_{MMSE}$. The maximum entropy algorithm, on the other hand, returns its estimate $\mathbf{h}_{MaxEnt}$ directly.

The Gaussian distributed, random noise added to form the simulated data is added at several different signal-to-noise ratios (SNRs). We calculate the SNR as the ratio of the "energy" of the response function times the signal to the "energy" of the noise or

$$\text{SNR} = \frac{\sum_i (\mathbf{R}\mathbf{h}_I)_i^2}{\sum_i n_i^2}. \tag{4.25}$$

Once we have an estimate of the initial signal using maximum entropy, ZF, or MMSE, we would like to compare the estimate with the initial signal. We use maximum cross correlation as a measure of similarity. To calculate the maximum cross correlation between two vectors $\mathbf{x}$ and $\mathbf{y}$ we first normalize the two vectors such that their autocorrelations at zero lag are equal to one. We denote these normalized vectors as $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$. If the two vectors are both of length L we then calculate the cross correlations $C(i)$ for lags i between $-L/2$ and $L/2$, where, if $i \geq 0$,

$$C(i) = \sum_{j=1}^{L-i} \tilde{x}_{j+i}\tilde{y}_j \tag{4.26a}$$

or, if $i < 0$,

$$C(i) = \sum_{j=1}^{L-|i|} \tilde{x}_j\tilde{y}_{j+|i|}. \tag{4.26b}$$

The maximum cross correlation is the maximum value that C obtains over all lags i.

Figures 4.1 and 4.2 show the maximum cross correlations between the initial and reconstructed signals for various SNRs and with the different deconvolution methods. Figure 4.1 shows the results of the methods for the random, Gaussian signals. For low SNR, maximum entropy and MMSE are much better at deconvolving the signal than ZF, which does not account for the noise amplitude. This poor performance of ZF highlights the need for regularization of the response $\mathbf{R}$ inverse. At low SNR, MMSE performs slightly better than maximum entropy because MMSE is given information about the signal amplitude. The slightly lower performance of MMSE and ZF compared with maximum entropy at high SNR

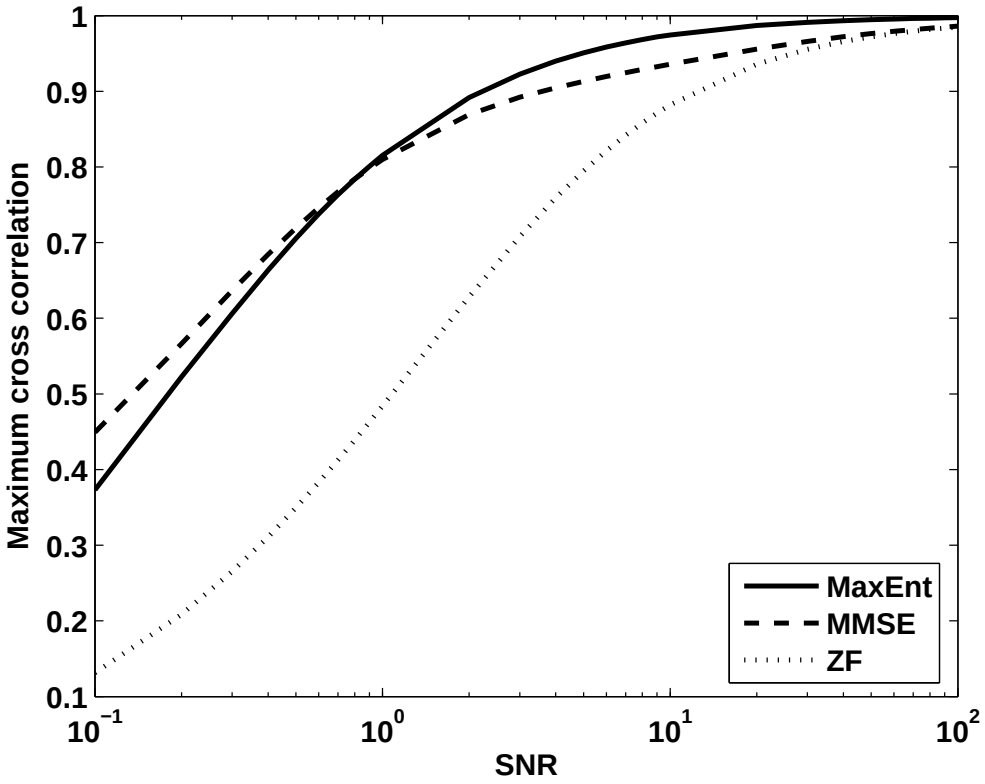


Figure 4.1. Maximum cross correlation between the recovered and initial, random, Gaussian signal at different SNR using different deconvolution methods. Maximum entropy and MMSE perform better than ZF at low SNR. At very low SNR, MMSE performs slightly better than maximum entropy because it is given more information about the signal.

is most likely due to the addition of the separate channel estimates (each with a small amount of noise) to form final estimates. Figure 4.2 compares the methods for the case of a supernova signal [1]. Notice that for this nonwhite signal, maximum entropy outperforms MMSE for low SNR. Finally, Figs. 4.3 and 4.4 show the estimated supernova signal for low and high SNR for both maximum entropy and MMSE.

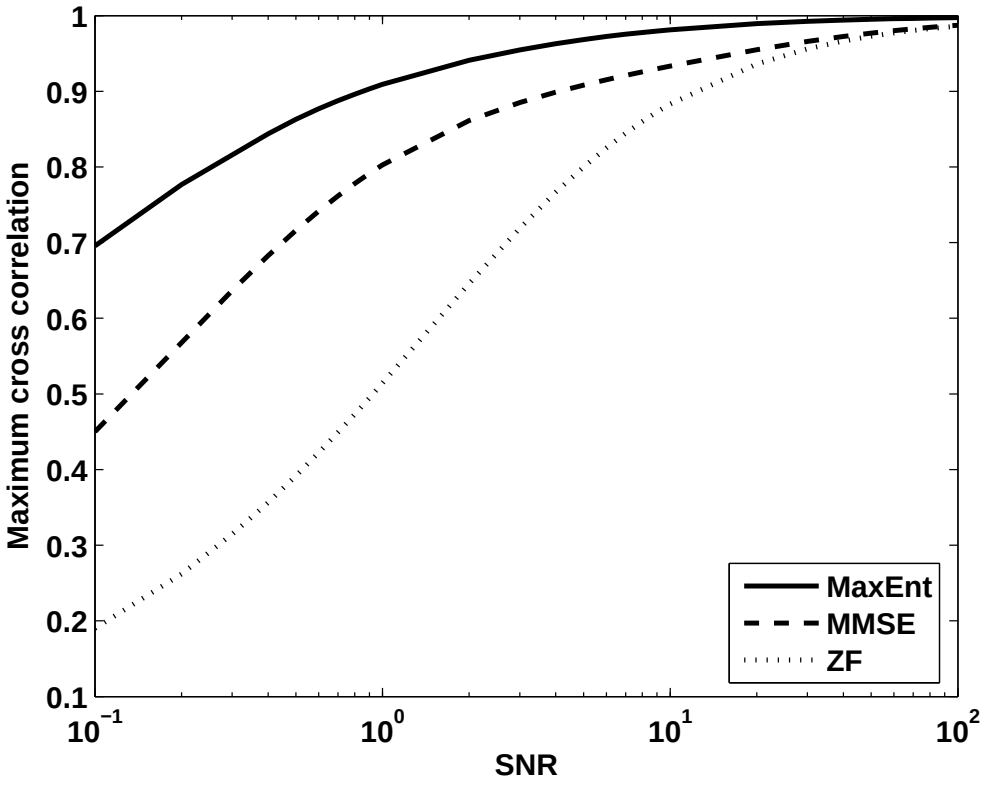


Figure 4.2. Maximum cross correlation between the recovered and initial supernova signal at different SNR using different deconvolution methods. Maximum entropy outperforms MMSE and ZF for the case of this nonwhite signal.

4.4 Application

The objective of LIGO goes beyond waveform recovery to discovery of the properties of the waveform source, which in the case of this chapter is a collapsing stellar core. To make this discovery, we propose to recover the waveform, using maximum entropy, and compare the recovered waveform with model waveforms, created by simulation, and see which models best match what we have observed. A measure of the "match" between model waveforms and recovered waveform is given by the maximum cross correlation between the two signals, as defined in Equations (4.26a) and (4.26b). We show that the catalog waveforms that result from models with character similar to the model that produced the waveform used as the simulation's initial signal have the highest cross correlation with the estimated signal, provided

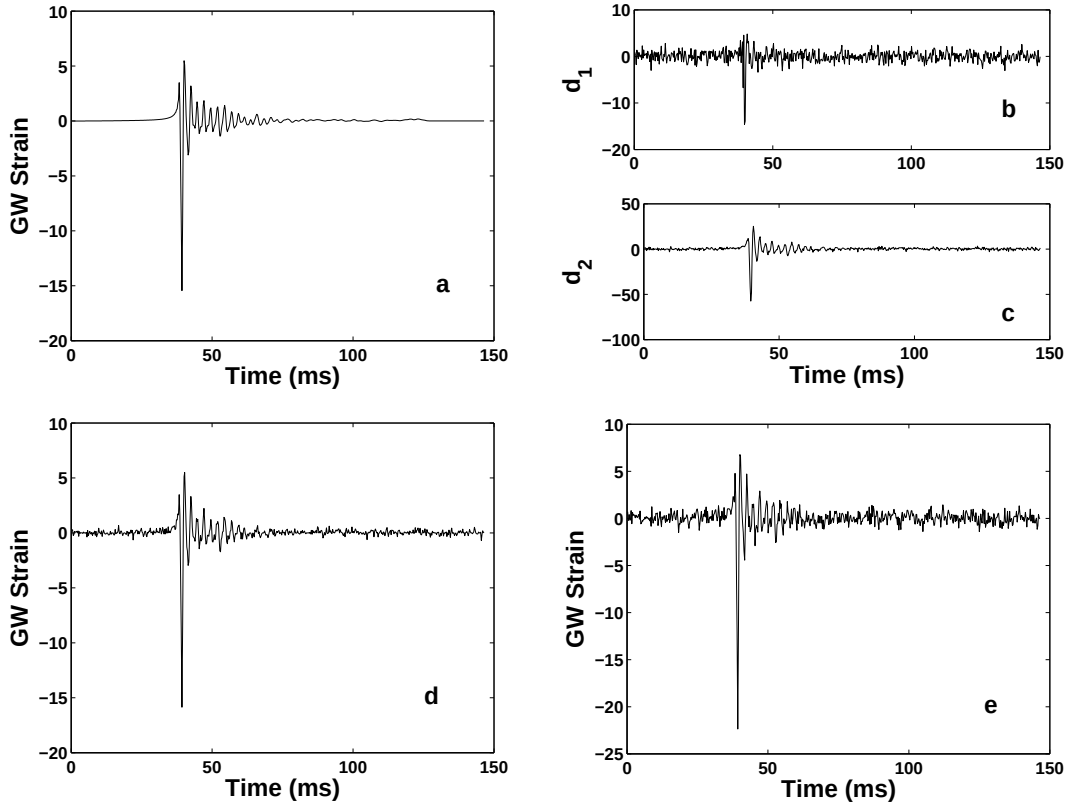


Figure 4.3. Estimation of supernova signal for high SNR of 10. (a) Portion of waveform ([1] s15A1000B0.1) used as initial signal. The waveform is scaled to correspond to a supernova at 10 kpc. (b) Data stream $\mathbf{d}_1$ formed by convolving the initial signal with the simple filter [1 -2 2] and adding white, Gaussian noise. (c) Data stream $\mathbf{d}_2$ formed by convolving the initial signal with the simple filter [1 2 2] and adding white, Gaussian noise. The convolved signal created using the filter [1 2 2] has a higher amplitude than the signal created with [1 -2 2]. The amplitude of the noise is chosen so that the average SNR for the two, convolved signals is 10 and the same amplitude noise is added to both signals. (d) Maximum entropy estimated signal. (e) MMSE estimated signal. Both maximum entropy and MMSE do a good job of estimating the initial signal.

sufficient signal strength. Therefore, the maximum cross correlation provides a qualitative indication of the physical properties of the source.

4.4.1 Description of Core-Collapse Models

This study used gravitational waveforms computed from a set of 2D, axisymmetric core collapse simulations by Ott et. al. [1] which focused on the dynamics of ro-

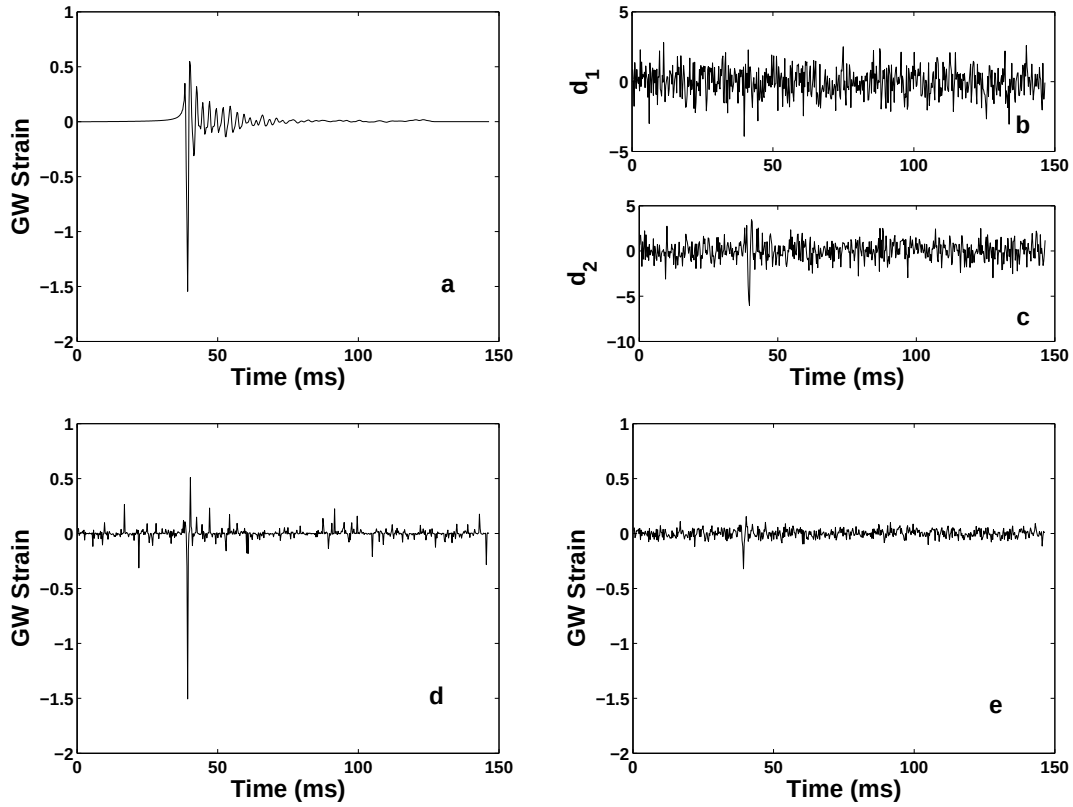


Figure 4.4. Estimation of supernova signal for low SNR of 0.1. (a) Portion of waveform ([1] s15A1000B0.1) used as initial signal. The waveform is scaled to correspond to a supernova at 10 kpc. (b) Data stream $\mathbf{d}_1$ formed by convolving the initial signal with the simple filter [1 -2 2] and adding white, Gaussian noise. (c) Data stream $\mathbf{d}_2$ formed by convolving the initial signal with the simple filter [1 2 2] and adding white, Gaussian noise. The convolved signal created using the filter [1 2 2] has a higher amplitude than the signal created with [1 -2 2]. The amplitude of the noise is chosen so that the average SNR for the two, convolved signals is 0.1 and the same amplitude noise is added to both signals. (d) Maximum entropy estimated signal. (e) MMSE estimated signal. Maximum entropy is able to estimate the signal more accurately than MMSE

tational collapse and bounce. These models used stellar progenitors with various masses: 11, 15, 20, and $25M_{\odot}$ calculated in [72]. The simulations neglected the effects of neutrinos and general relativity, but used the realistic, finite-temperature equation of state of [73]. A small number of simulations also investigated stellar progenitor models from [74] and [75], which were evolved to the onset of iron core collapse with an approximate treatment of rotation [74, 75] and angular momentum redistribution by magnetic torques [75]. The gravitational wave signature extrac-

tion was performed using the Newtonian quadrupole formalism (see e.g. [76]).

The effects of rotation were investigated very thoroughly in [1]. The initial rotation of the progenitor was controlled by two parameters: the rotation parameter β where

$$\beta = \frac{E_{rot}}{\mid E_{grav} \mid} \tag{4.27}$$

and the differential rotation scale parameter A, which is the distance from the rotational axis at which the rotational velocity drops to half that at the center. A is defined as

$$\Omega(r) = \Omega_0 \left[1 + \left(\frac{r}{A}\right)^2\right]^{-1}, \tag{4.28}$$

where r is the distance from the axis of rotation and $\Omega(r)$ is the angular frequency at r.

When the progenitor is rotating slowly and β is small (zero to a few tenths of a percent), the collapse is halted when the inner core reaches supranuclear densities. The core bounces rapidly and then quickly rings down. When the progenitor rotates more rapidly and β is larger, the core collapse is halted by centrifugal forces and the core bounces at subnuclear densities. The core then undergoes multiple damped, harmonic oscillator-like expansion-collapse-bounce cycles. The initial degree of differential rotation affects the value of β at which this bounce type transition occurs. A progenitor with a smaller value of A experiences a greater amount of differential rotation and hence, a more rapidly rotating inner core. As a result, the transition from a supranuclear to a subnuclear bounce occurs for a lower value of β.

4.4.2 Extracting Source Information from a Simulated Detection

How well can we hope to characterize a supernova progenitor from the gravitational wave signature we observe? To study this question, we have simulated the noisy data stream from two 4-km LIGO gravitational wave detectors observing a simulated supernova. For this study, we used the parameter study of gravitational wave signatures generated by supernovae in [1] and the noise performance and detector response of the LIGO gravitational wave detectors. In all simulations the actual

signal embedded in the noise came from the model s15A1000B0.1 of [1], which corresponds to a 15 $M_{\odot}$ progenitor from [72] with rotation parameter β equal to 0.1% and differential rotation scale parameter A equal to 1000 km. The gravitational wave signal arising from this model was scaled to different distances and projected onto the 4-km Hanford, WA (LHO) and Livingston, LA (LLO) LIGO detectors as if the supernova were directly above the Hanford site. Since the supernova models are axisymmetric, the gravitational radiation is linearly polarized. The polarization angle was chosen to maximize the response of the LHO detector [77]. We used response functions for the LHO and LLO detectors characteristic of four different science runs (S1, S2, S3 and S4) [78, 79, 80, 81] and simulated the detector noise by adding white noise with power spectral density amplitude approximately equal to the noise amplitude at 100 Hz in the corresponding science run [2, 3, 4, 5]. We used maximum entropy to find the best estimate of the embedded signal and cross-correlated this estimate signal with different signals drawn from the parameter survey of [1]. The subsections below describe our observations based on this study.

4.4.2.1 Improvement with Science Run

Figure 4.5 shows the maximum cross correlation between the estimated signal and initial waveform used in the data simulation versus supernova distance for the four science runs. There is a steady improvement, from S1 to S4, in maximum entropy's ability to recover signals at greater distances, corresponding to improving detector sensitivity [2, 3, 4, 5].

By S4, maximum entropy is able to recover the gravitational waveform from a supernova that occurs as far as a few kpc away, whereas for S1 the cross correlation between the initial and recovered signals drops off at ten parsecs. The investigations into the source information present in the recovered, estimated signals described in the following sections, use the data simulated for the most sensitive S4 science run.

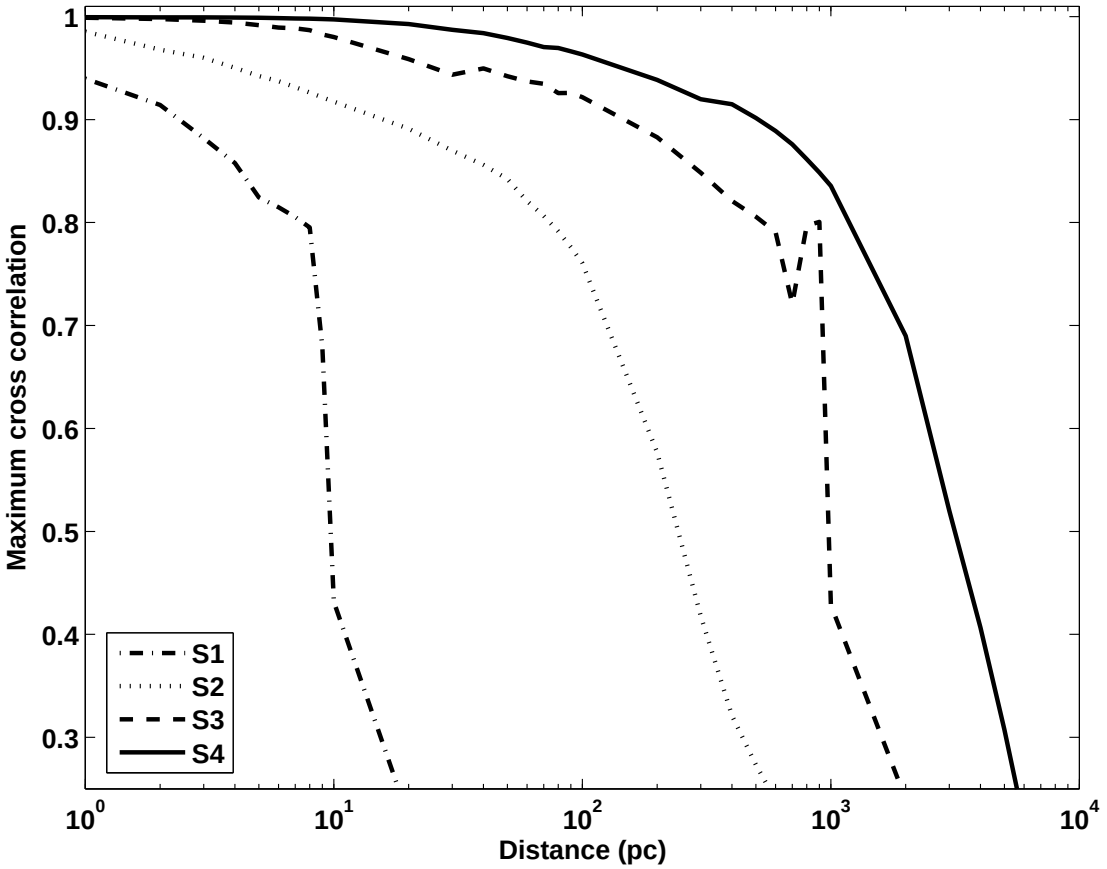


Figure 4.5. Maximum cross correlation between estimated and initial signals versus supernova distance for data simulated using different science run detector impulse responses and noise levels. The estimated signal is recovered using maximum entropy from simulated detections that use Ott et.al. [1] model s15A1000B0.1 as the initial signal waveform. There is a steady improvement in maximum entropy's ability to reconstruct fainter, more distant signals as the sensitivity of the detectors improved [2, 3, 4, 5].

4.4.2.2 Bounce Type

We can classify the models of [1] into those that bounce at supranuclear densities, subnuclear densities, and transitional densities. Figure 4.6 shows the maximum cross correlation between the recovered waveform and the catalog waveforms of the three bounce varieties that have the greatest maximum cross correlation at small distance. It is clear that the recovered waveform has the most in common with that generated from a model with the same, supranuclear, bounce type and that our ability to distinguish bounce type fails for supernovae more than 2-3 kpc

away for S4 detector sensitivities. Figure 4.7 shows the different waveforms that are compared with the recovered signal in Fig. 4.6. The waveform amplitudes are scaled to correspond to optimally oriented supernovae at a distance of 10 kpc.

4.4.2.3 Mass

Figure 4.8 shows the cross correlation between the recovered waveform and the waveforms associated with supernova models that differ by progenitor mass, but share the same rotational parameters, while Fig. 4.9 shows these waveforms. The recovered waveform for distances up to 2 kpc most closely resembles the waveform from the model with the same mass.

4.4.2.4 Rotation

Figures 4.10 and 4.11 show the cross correlation between the recovered waveforms and the waveforms associated with models that differ only by rotation parameter and by differential rotation parameter. Figures 4.12 and 4.13 show these waveforms. As the rotational parameters of the models depart from those associated with the original waveform's model the magnitude of the maximum cross correlation decreases.

4.4.3 Adding a Detector

In order to see how the performance of maximum entropy changed with the addition of the third LIGO detector, we repeated the supernova detection simulations described in Section 4.4.2. The simulated data were created for not only the two interferometers with 4 km long arms, but for the Hanford detector with 2 km long arms as well. The maximum cross correlation between the estimated and initial signals versus distance for the four science runs with the three detectors is shown in Fig. 4.14 and compared with the performance using only the two 4-km detectors. As can be seen in this figure, adding the 2-km detector sometimes does and sometimes does not improve maximum entropy's ability to recover supernova signals. In the case of the S3 simulations, using three detectors worsens performance significantly. During the S3 run, the 2-km detector was much less sensitive than the 4-km detectors below a few hundred Hz [4]. During S4, however, the

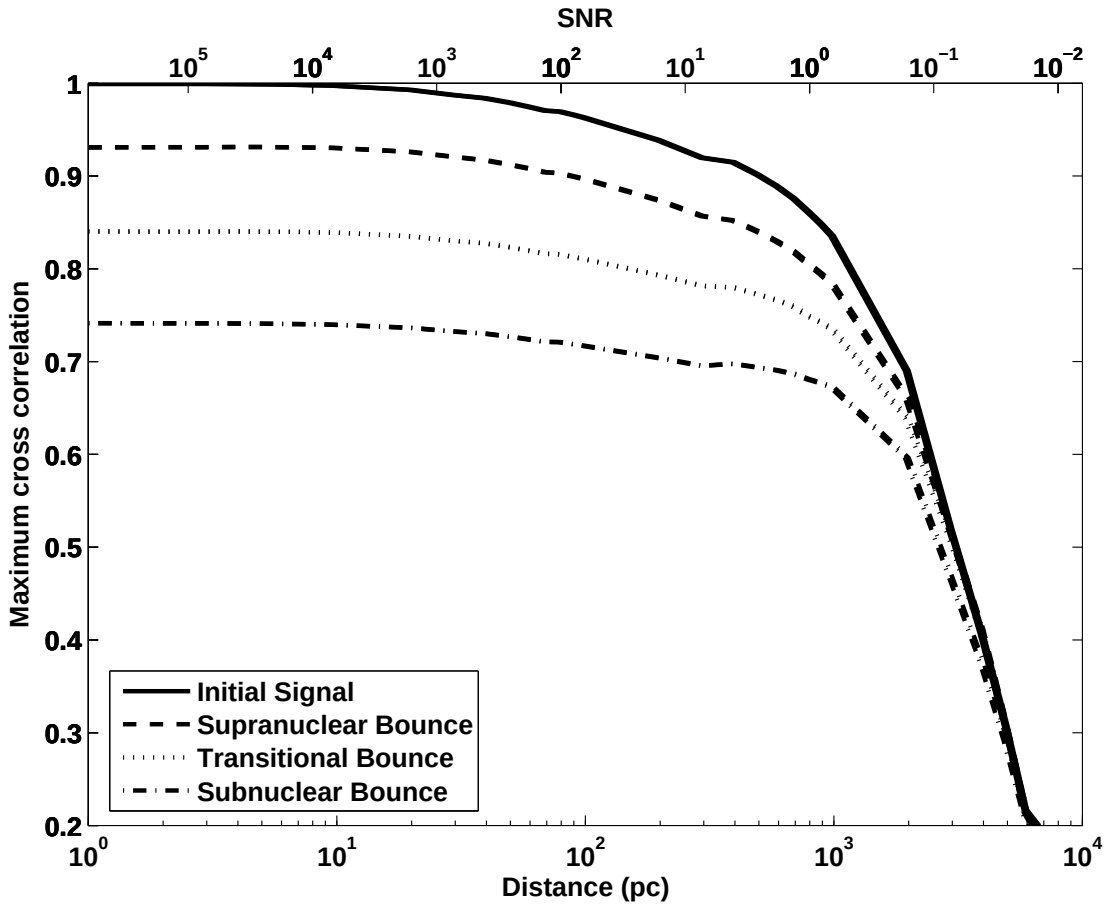


Figure 4.6. Maximum cross correlation between reconstructed waveforms and waveforms associated with models that differ by bounce type versus supernova distance and SNR. The reconstructed waveform is recovered using maximum entropy from simulated detections that use the waveform from the Ott et.al. [1] model s15A1000B0.1 as the initial signal waveform as well as detector responses and noise levels from the fourth science run (S4). The solid line represents the maximum cross correlation between the reconstructed signal and the initial signal waveform. The other lines represent the maximum cross correlations between the recovered waveforms and the waveforms resulting from each bounce type for which the maximum cross correlation at 1 pc is greatest, excluding that used for the initial signal. The recovered waveform is most similar to those generated by models with the same, supranuclear bounce type as the initial signal waveform, for the simulations corresponding to supernovae that occur less than 2-3 kpc away.

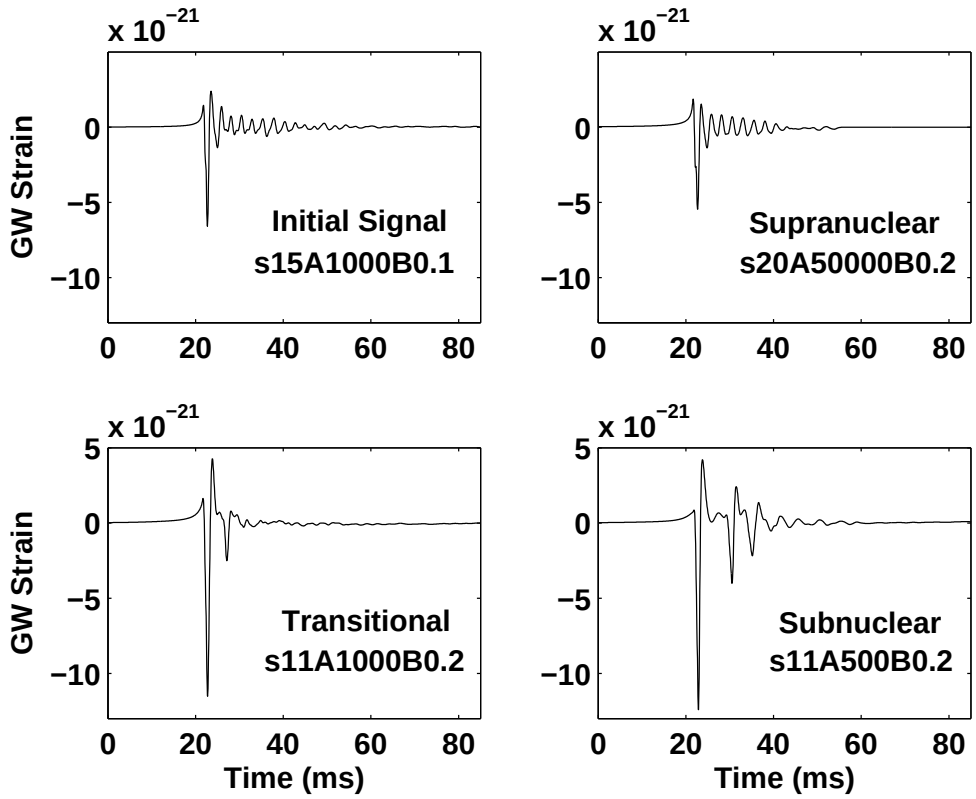


Figure 4.7. The waveforms associated with various bounce types that are compared with the recovered signal in Fig. 4.6. The upper left plot shows the waveform (from Ott et al. [2004] model s15A1000B0.1) that was used as the initial signal in the simulated detection. The three other waveforms shown are those that are most similar to this initial signal waveform, for each bounce type. The waveform from the [1] s20A50000B0.2 model looks much like the initial signal waveform which is of the same, supranuclear bounce type. The subnuclear bounce waveform shows the effects of multiple damped, harmonic oscillator-like expansion-collapse-bounce cycles. The zero points of the time axes are chosen so that the minima of the waveforms occur at the same time for ease of comparison. The waveform amplitudes are scaled to correspond to supernovae at 10 kpc.

2-km detector was much closer to the expected factor of two less sensitive than the 4-km detectors [5] and adding data from the 2-km detector improved maximum entropy's performance. As all three detectors continue to improve in stability and reach their design sensitivities we expect the relative detector sensitivities to be closest to those in S4 and that using the data from all three interferometers will give the best results.

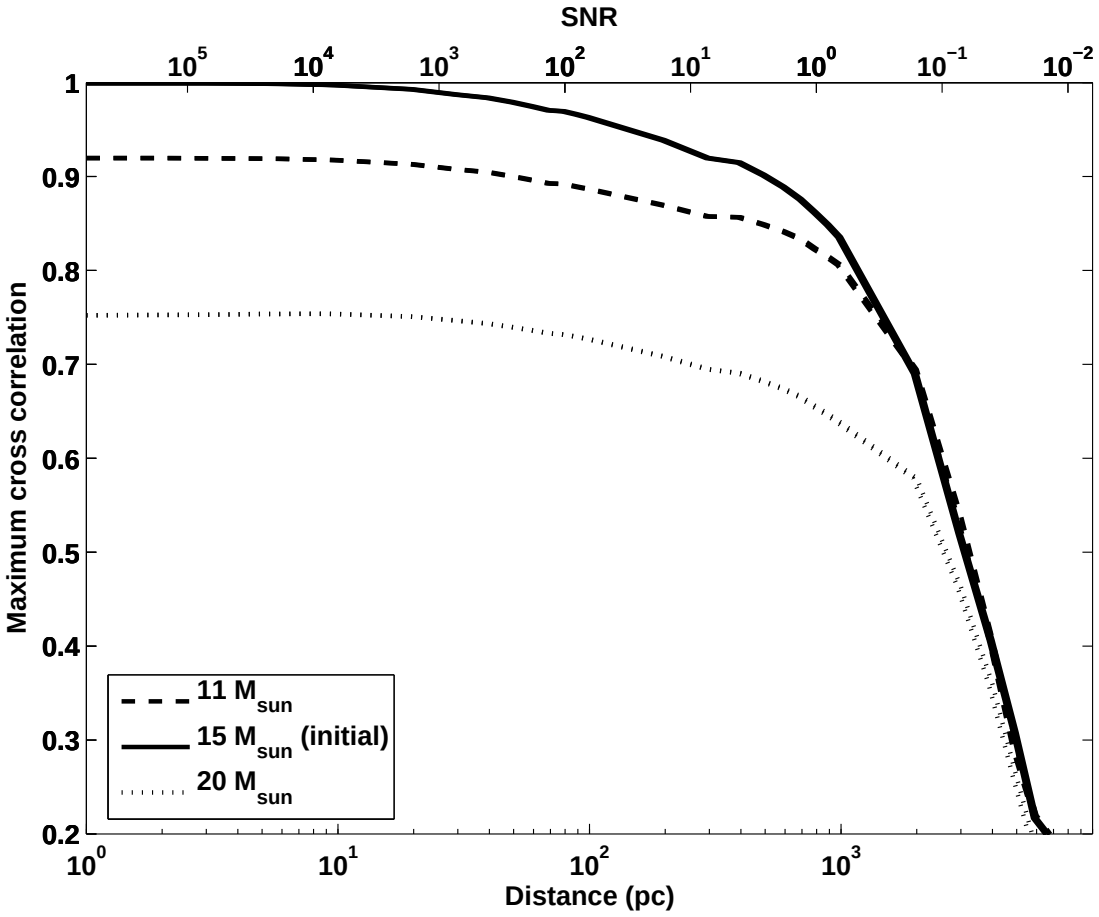


Figure 4.8. Maximum cross correlation between reconstructed waveforms and waveforms associated with models that differ only by progenitor mass versus supernova distance and SNR. The reconstructed waveform is recovered using maximum entropy from simulated detections that use a waveform from a model with a progenitor mass of 15 solar masses (Ott et al. [1] model s15A1000B0.1) as the initial signal waveform as well as detector responses and noise levels from the fourth science run (S4). The recovered waveform is most similar to that generated by the model with the same progenitor mass for the simulations corresponding to supernovae that occur less than 2-3 kpc away.

In addition to the three LIGO detectors, the maximum entropy method could be applied to data from other gravitational wave interferometric detectors such as GEO600 [17], TAMA300 [18], and VIRGO [19]. For the simulation described above, adding one of these detectors would most likely not improve maximum entropy's estimate of the signal, since we have chosen the optimal sky location and source orientation for Hanford. The absolute amplitude of the gravitational wave

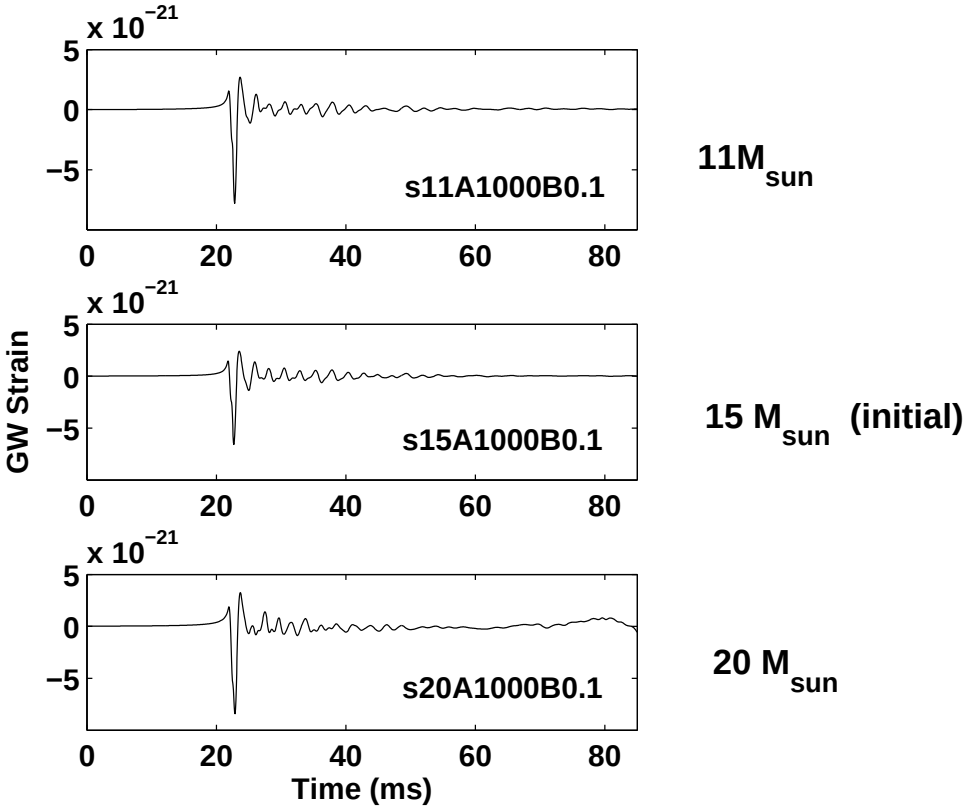


Figure 4.9. Waveforms from models that differ only by progenitor mass. The waveform corresponding to model s15A1000B0.1 was used as the initial signal in the detection simulations. The zero points of the time axes are chosen so that the minima of the waveforms occur at the same time for ease of comparison. The waveform amplitudes are scaled to correspond to supernovae at 10 kpc.

strain from a source at this sky location and orientation projected onto the LIGO Livingston detector is 89% that for the Hanford detectors. The strain amplitude for GEO300 is 42%, for TAMA300 is 22%, and for VIRGO is 1% that at the Hanford detectors. For signals coming from sky locations that are more favorably located for these, non-LIGO detectors, including data from these instruments will be essential for waveform recovery and analysis.

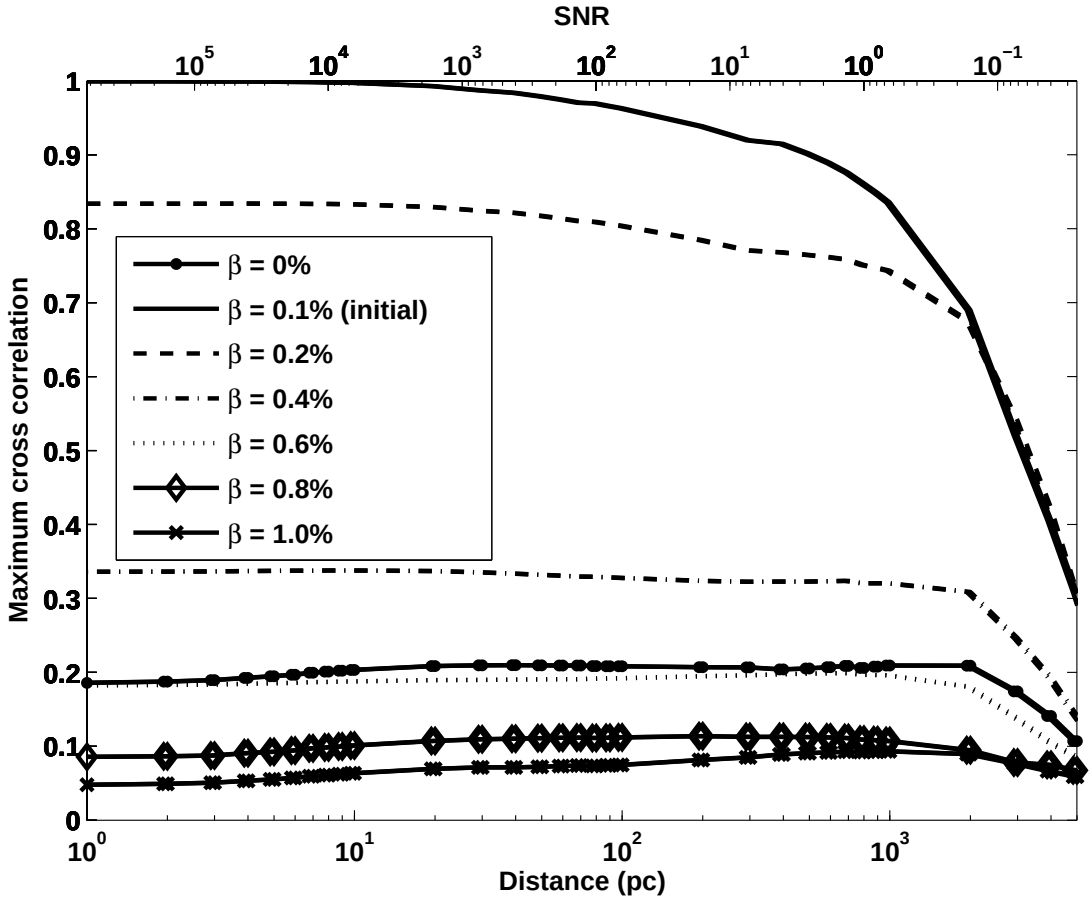


Figure 4.10. Maximum cross correlation between reconstructed waveforms and waveforms associated with models that differ only by rotation parameter β, which is defined in Equation (4.27), versus supernova distance and SNR. The reconstructed waveform is recovered using maximum entropy from simulated detections that use a waveform from a model with a rotation parameter of $\beta = 0.1\%$ (Ott et al. [1] model s15A1000B0.1) as the initial signal waveform as well as detector responses and noise levels from the fourth science run (S4). The recovered waveform is most similar to that generated by the model with the same β for the simulations corresponding to supernovae that occur less than 2-3 kpc away.

4.5 Realistic Noise

Using white noise to simulate LIGO noise is unrealistic since LIGO noise is far from white (see Chapter 1). One may be concerned that in the case of an actual supernova detection, the maximum entropy method may be unable to recover the gravitational wave signal due to the presence of highly colored noise. In order

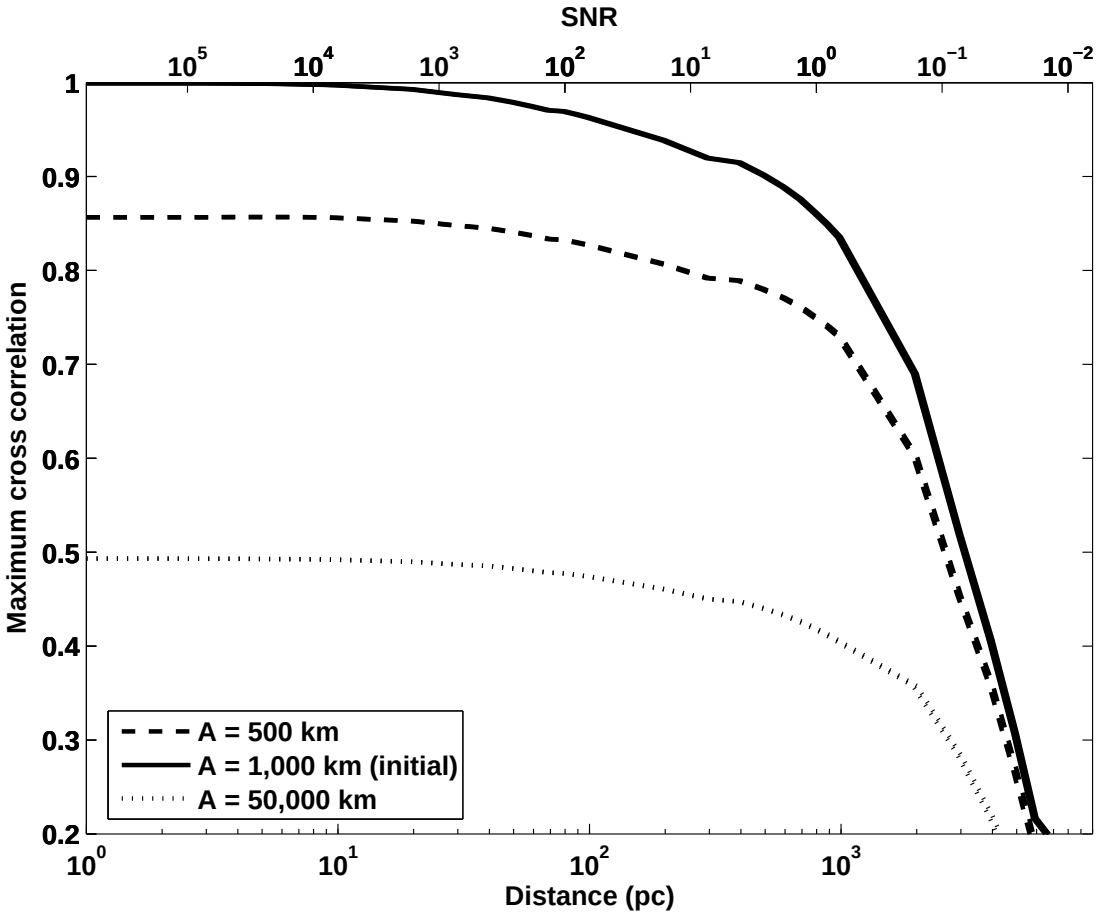


Figure 4.11. Maximum cross correlation between reconstructed waveforms and waveforms associated with models that differ only by initial degree of differential rotation as parameterized by A, which is defined in Equation (4.28), versus supernova distance and SNR. The reconstructed waveform is recovered using maximum entropy from simulated detections that use a waveform from a model with a differential rotation parameter of $A = 1{,}000$ km (Ott et al. [1] model s15A1000B0.1) as the initial signal waveform as well as detector responses and noise levels from the fourth science run (S4). The recovered waveform is the most similar to that generated by the model with the same initial degree of differential rotation.

to address this concern, we repeated the detection simulation using S4 detector responses and S4 data as noise. The maximum cross correlation between the original and estimated signal for both white and realistic detector noise is shown in Figure 4.15. Maximum entropy's ability to recover the signal from the data created with nonwhite noise is actually a little better than for white noise. This

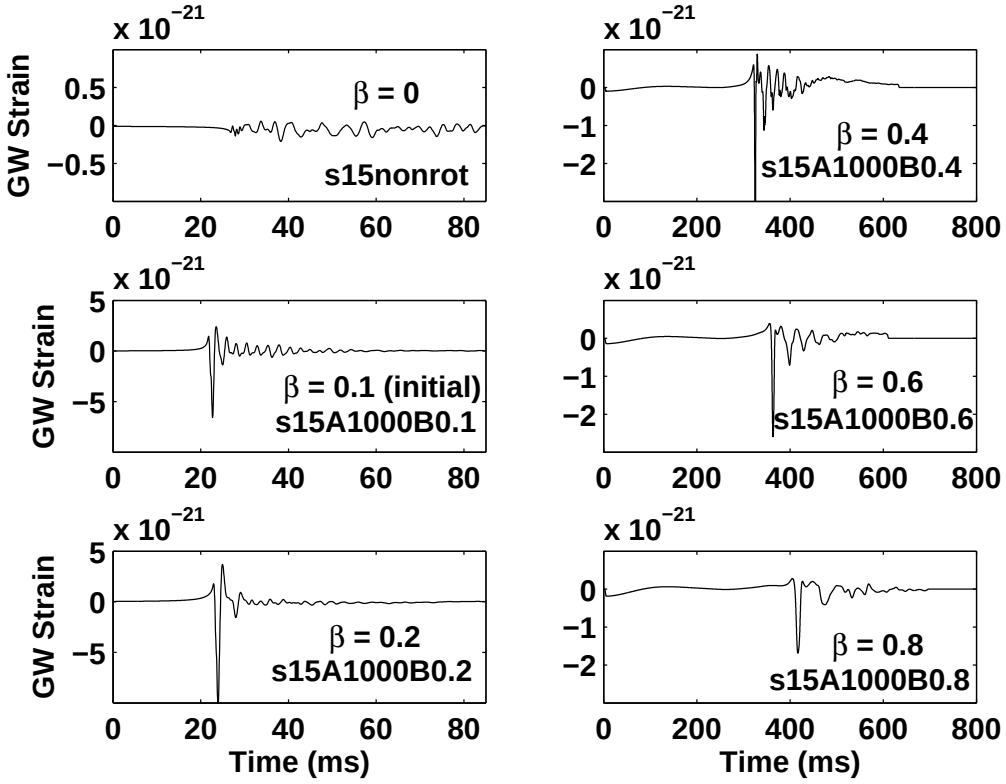


Figure 4.12. Waveforms from models that differ only by rotation parameter β, defined by Equation (4.27). The waveforms for larger β ($\geq 0.4\%$) have significant amplitude over durations of hundreds of ms, while low β waveforms last only for tens of ms. The $\beta = 0\%$ waveform has very low amplitude as the non-rotating collapse is nearly spherically symmetric. The waveform corresponding to model s15A1000B0.1 was used as the initial signal in the detection simulations. The zero point of the time axes for the plots on the right is chosen so that the onset of significant gravitational wave amplitude occurs at roughly the same time while the plots on the left show the first 800 ms of the waveform. The waveform amplitudes are scaled to correspond to supernovae at 10 kpc.

not only shows that maximum entropy is well able to handle colored noise but that choosing to scale the white noise by LIGO noise power at 100 Hz for the simulated detections is a conservative choice.

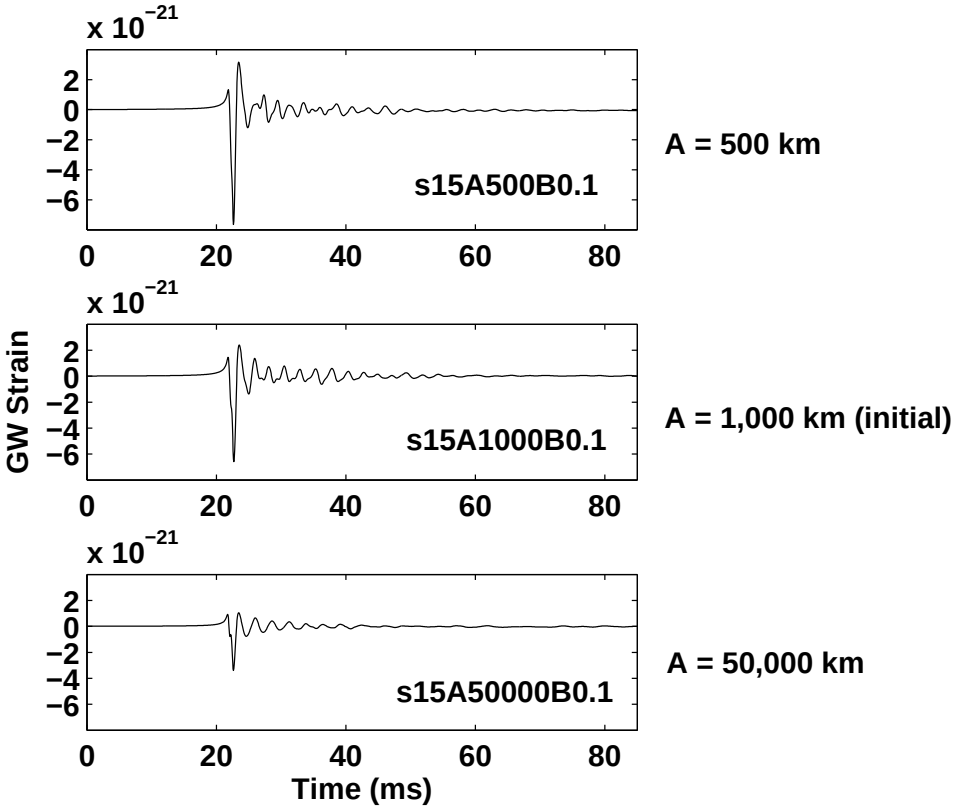


Figure 4.13. Waveforms from models that differ only by initial degree of differential rotation. The differential rotation parameter, A, is the distance at which the rotational velocity of the progenitor drops to half the rotational velocity at its center. As A decreases the differential rotation of the progenitor becomes more extreme and the amplitudes of the gravitational waves increase. The center plot shows the waveform corresponding to model s15A1000B0.1 which was used as the initial signal in the detection simulations. The zero points of the time axes are chosen so that the minima of the waveforms occur at the same time for ease of comparison. The waveform amplitudes are scaled to correspond to supernovae at 10 kpc.

4.6 Hardware Injection Recovery

The maximum entropy method assumes that the response matrix $\mathbf{R}$ is known perfectly. In reality, the portion of $\mathbf{R}$ due to the projection of the gravitational waves onto the arms of the interferometers is well known while the portion due to the varying frequency responses of the detectors is measured. This measurement process is described in [78, 79, 80, 81, 82].

In all of the simulations described in previous sections, the $\mathbf{R}$ used by maximum

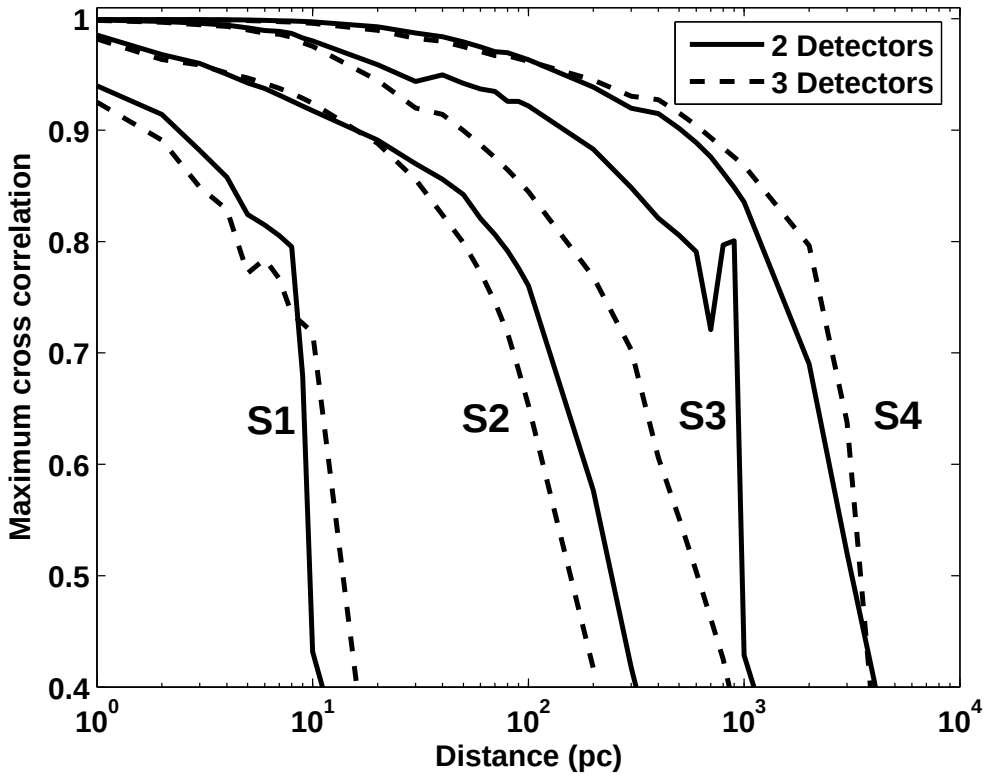


Figure 4.14. Comparison of maximum entropy performance with simulated data from two and three interferometers. Adding the third, 2-km interferometer does or does not help depending on its sensitivity relative to the other two detectors. During S3 the Hanford 2-km interferometer was much less sensitive than the two 4-km detectors below a few hundred Hz [4]. During S4, however, the 2-km detector was much closer to the expected factor of two less sensitive than the 4-km detectors [5] and adding the 2-km data improved maximum entropy's results.

entropy is the same as that used to create the simulated data. In order to test our ability to recover a signal under more realistic conditions, where $\mathbf{R}$ is measured, we apply maximum entropy to data containing hardware injections. Hardware injections are performed by moving the mirrors of the interferometers so that signals are "injected" into the data. Studying hardware injections is the closest one can get to data from actual gravitational wave detections.

We applied maximum entropy to two hardware injections carried out during S4. Both used the waveform A3B3G1 from the Zwerger-Müller catalog as the injection waveform [83]. The Zwerger-Müller waveforms come from core-collapse

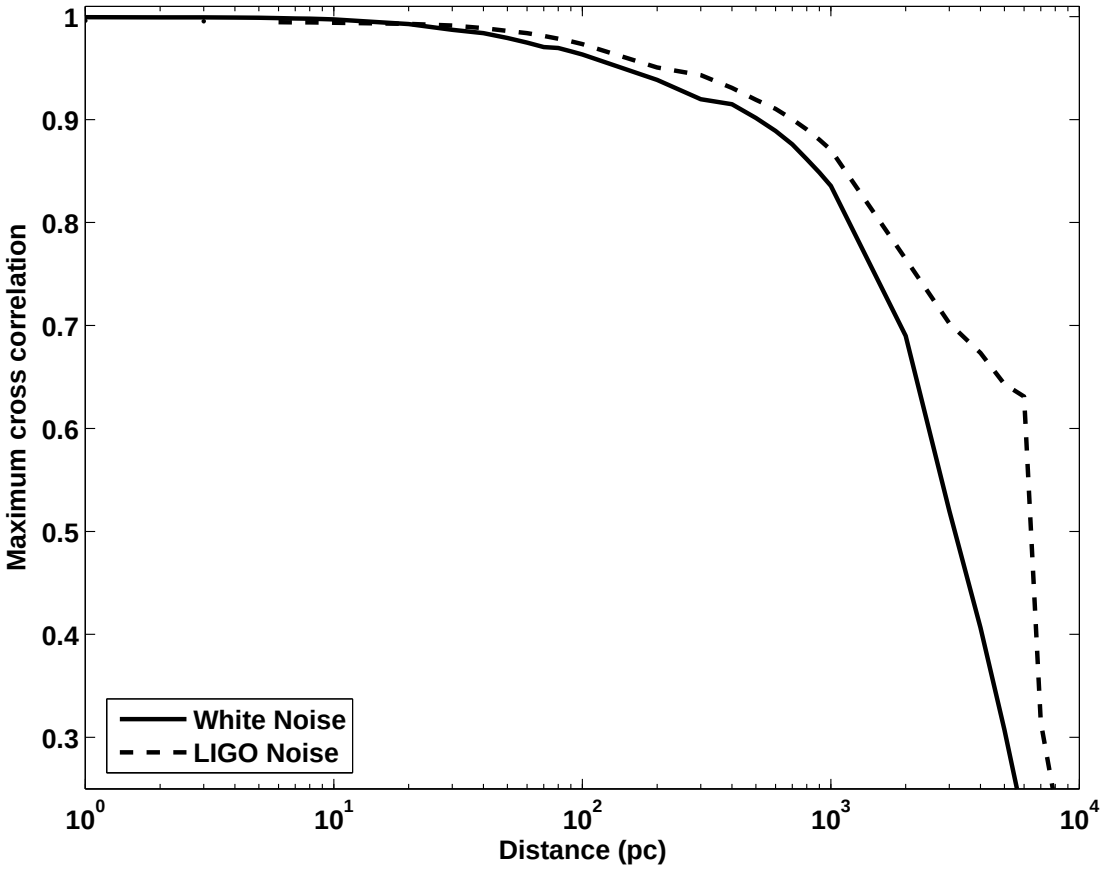


Figure 4.15. Comparison of maximum entropy performance with simulated data using white noise and using actual LIGO data for noise. The LIGO data was acquired during the S4 data run and the white noise was scaled to have the same power as the LIGO data at 100Hz. Maximum entropy is able to recover signals better from the data with the LIGO data noise showing that the method is well able to deal with colored noise

simulations that use many of the same parameters as Ott et. al. [1] which provide the waveforms used in the simulations described above. The main differences between the two sets of simulations are that the Zwerger-Müller simulations assume an equation of state (EOS) that is polytropic, that is, part of the pressure can be described by $P = K\rho^{\Gamma}$ where K is a constant that depends on the electron number fraction (which is set to one half), ρ is the density and $\Gamma = 4/3$. Collapse is initiated by suddenly reducing Γ from 4/3 to a new value Γ_r [83]. Models by Ott et. al., on the other hand, use realistic progenitors instead of polytropes, use a tabular, realistic equation of state, and collapse on their own [1]. The A3B3G1

Zwerger-Müller model has a differential rotation parameter $A = 500$ km, rotation parameter $\beta = 0.9$ % and $\Gamma_r = 1.325$ [83].

The two hardware injections that we studied were done at two different signal strengths. The stronger hardware injection was scaled such that its integrated strain $h_{rss} = 8.0 \times 10^{-21}$ while the weaker injection had $h_{rss} = 0.5 \times 10^{-21}$. The hardware injections were performed in all three detectors.

The response matrix, $\mathbf{R}$, was calculated assuming a source directly above and optimally aligned for each detector (this nonphysical arrangement was used in the injections), a 0.002 second delay between Livingston and Hanford and the measured calibration of each detector.

Figures 4.16 and 4.17 show the signals recovered from the strong and weak hardware injection data from all three LIGO detectors using maximum entropy. The comparison with the original injection signal shows that maximum entropy is successful in recovering signals from real LIGO data. Figures 4.18 and 4.19 show the maximum cross correlation between the recovered signal and waveforms from Zwerger-Müller models that differ only by A and β. These figures show that these signals recovered from actual data using an imperfect $\mathbf{R}$ still contain information about the rotation of the progenitor.

4.7 Discussion/Conclusions

In this chapter we have introduced the use of the maximum entropy method to recover supernova gravitational wave signals from detector data. We have shown that maximum entropy compares favorably with other methods and can even reconstruct waveforms at low SNR. Once the signals have been recovered with maximum entropy, calculating the maximum cross correlation between the recovered signals and a catalog of waveforms from models incorporating a variety of physical parameters, provides a qualitative indication of the properties of the source.

We have also demonstrated that the promise of gravitational wave astronomy can be realized: that gravitational waves associated with supernovae contain significant information about the systems that generate them, including progenitor mass, rotation parameter, initial degree of differential rotation, and that this information can be extracted from gravitational wave observations. With current

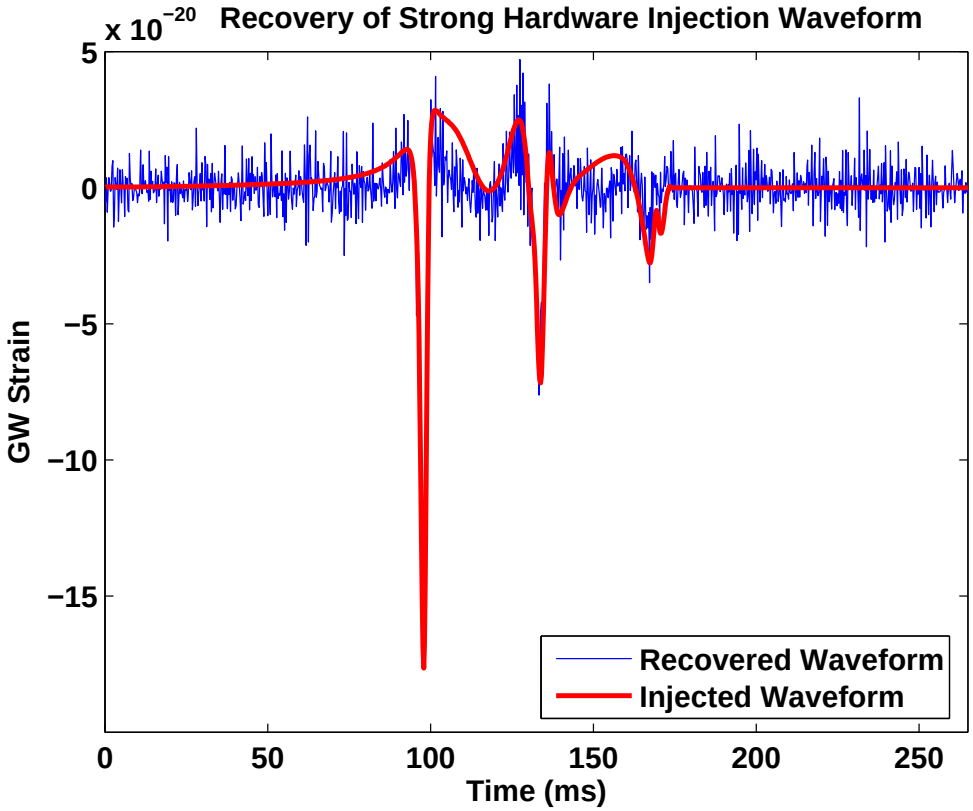


Figure 4.16. Strong hardware injection recovery using maximum entropy. The recovered waveform has many of the features of the original signal used to create the hardware injection

detector sensitivities and supernova modeling, the details of observed supernovae can be recovered for supernovae as far away as a few kpc.

Further studies will investigate more sophisticated measures of waveform information content than the simple cross correlation. We will also be investigating waveforms from new simulations, as they become available, to see how our ability to determine physical parameters changes with the addition of new physics (such as neutrinos) in the models.

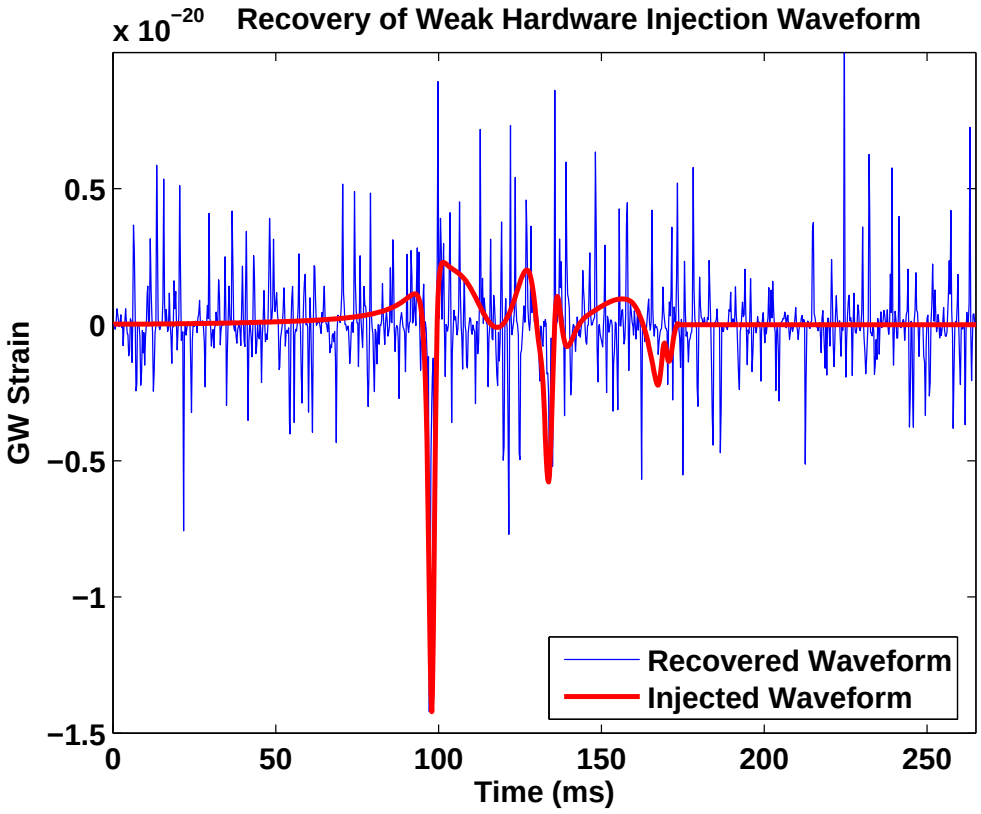


Figure 4.17. Weak hardware injection recovery using maximum entropy. The recovered waveform is very noisy but still retains the strongest features of the original signal used to create the hardware injection.

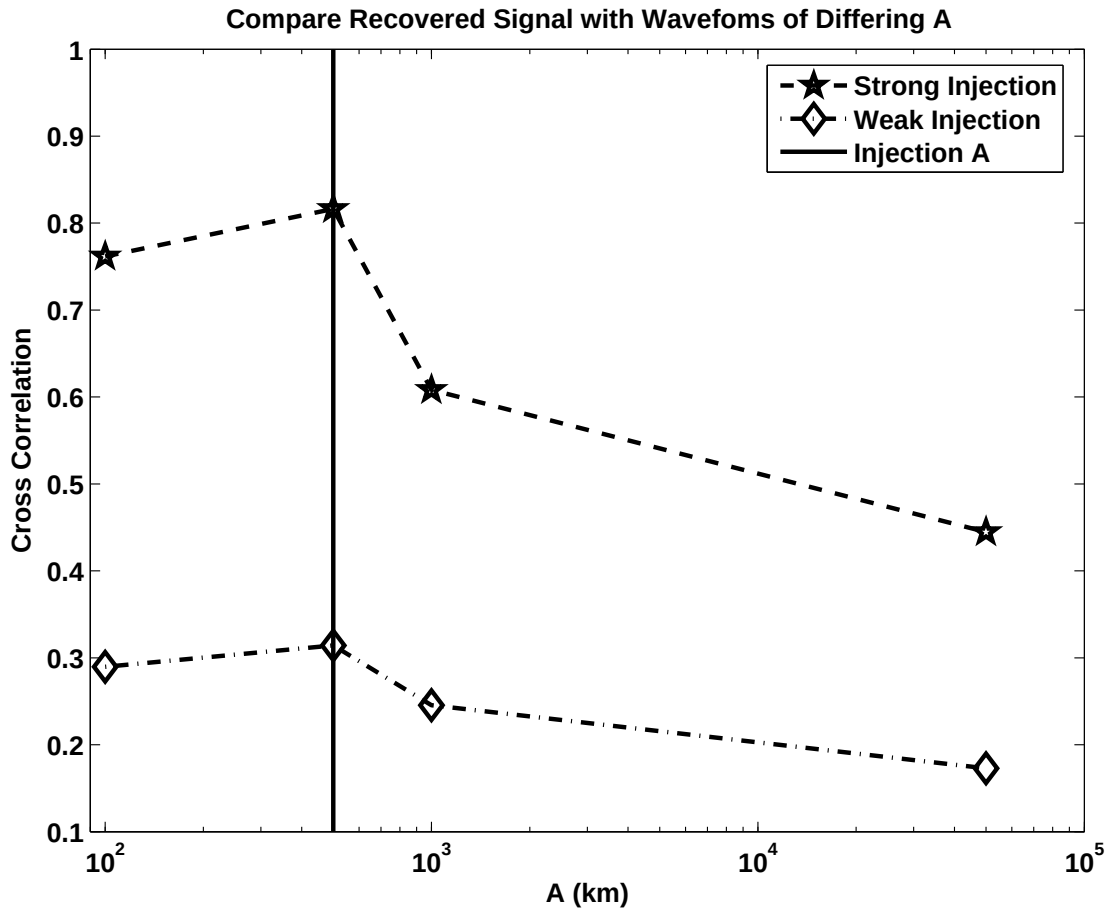


Figure 4.18. Maximum cross correlation between the recovered waveform and waveforms from Zwerger-Müller models that differ only by the amount of differential rotation A. Even for the weaker hardware injection, the recovered waveform most resembles that from the model with the same A as used in the injection.

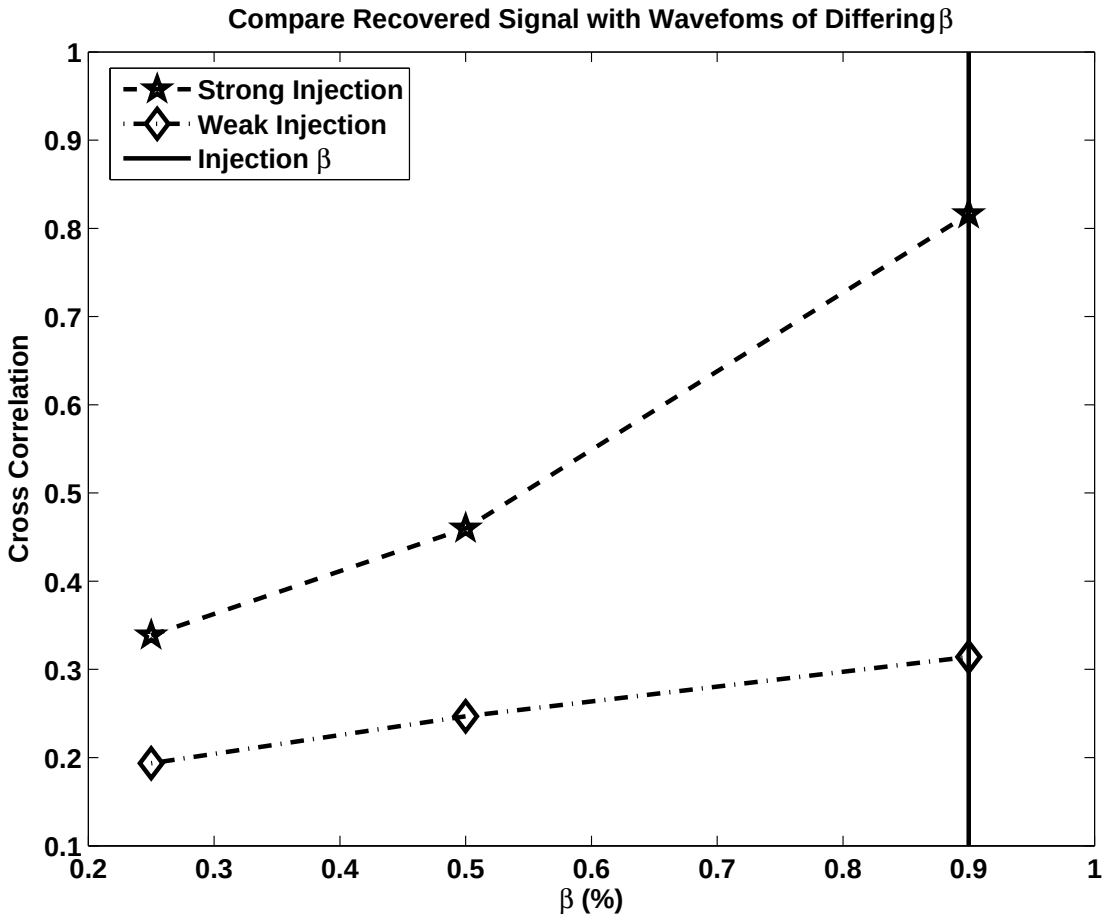


Figure 4.19. Maximum cross correlation between the recovered waveform and waveforms from Zwerger-Müller models that differ only by rotation parameter β. Even for the weaker hardware injection, the recovered waveform most resembles that from the model with the same β as used in the injection.

Chapter 5

Conclusions

In order for the potential of LIGO to be realized and gravitational waves become an important astronomical tool, there are many data analysis challenges that must be met. The three projects described in this thesis make contributions toward addressing some of these challenges.

Chapter 2 introduced a procedure for conditioning LIGO data. Such a procedure is essential for gravitational wave burst search algorithms that require the data be as free of artifacts and correlations as possible. In addition the procedure breaks the data into frequency bands. Candidate events found within a band can automatically be associated with the band's range of frequencies. It was shown that this data conditioning procedure is successful in making data whiter most (81% - 100%) of the time. Failures of the data conditioning are caused by instrumental "glitches" and by changes in the character and/or amplitude of the noise. The performance of the data conditioning should continue to improve as the stability of the detectors improves.

Chapter 3 introduced the Poisson test which can be used to flag segments of data contaminated by nonlinear couplings. The test was applied to LIGO data and it was demonstrated that along with glitches and data conditioning failures, the test was able to pick out segments with correlations between data points. One segment in particular that had a high Poisson test value had frequencies present in its bicoherence that indicated that data samples were correlated with those advanced or delayed by two samples.

Finally, Chapter 4 tackled the problems of recovering a gravitational wave

signal from a network of detectors as well as discovering source information from the recovered signal when the source is imperfectly modeled. Maximum entropy is proposed as a method for signal recovery and is shown to perform better than a simple inversion of the detector response or minimum mean square error (MMSE) filtering for colored signals. This despite the fact that MMSE filtering assumes more prior information about the signal than maximum entropy.

Maximum entropy was used to recover the signals from simulated supernova detection using a waveform from an Ott et. al. [1] core-collapse model. The simulations were done with signal strengths that corresponded to supernovae occurring at various distances and with detector responses and white noise of amplitudes characteristic of the LIGO detectors. The recovered signals were cross-correlated with the waveforms in the Ott et.al. catalog and the cross-correlation taken as an indication of which models most resembled the supernova that produced the original signal. It was found that the recovered signals contained information about the type of bounce the core underwent as well as information about the progenitor mass, rotational energy, and differential rotation. It will be especially interesting to repeat this analysis when waveforms from models incorporating more physical effects become available. For example, if waveforms from models that include the effects of neutrinos have waveforms that have a high cross-correlation with waveforms from models without neutrinos but with the same progenitor mass and rotation characteristics it would indicate that waveforms from incomplete models may still be very useful in analyzing a gravitational wave detection. If, on the other hand, including more physics completely changes the waveforms then only very complete models would be useful.

Maximum entropy was also tested under more realistic conditions such as using real detector data for noise and recovering hardware injections where the response of the detectors is measured. It was found to perform well under these more realistic situations.

Gravitational wave astronomy holds the exiting potential of opening an entirely new window on the universe. The projects described in this thesis represent a few steps toward realizing that potential. Hopefully, the next few years will see the first direct detection of gravitational waves and the beginning of a new era in astronomy.

Appendix A

Kalman filtered line parameters

The following tables specify the Kalman filtering parameters used during the S2 and S3 runs. The first column of each table gives the frequency band for which each line was removed. Frequency band definitions are given in Appendix B. The second column gives the line number. Lines with the same band and line number are removed together with the same Kalman filter. As for the other columns, f_0 is the frequency and Q the quality factor of each line. f_c and Δf are the central frequency and size of the frequency band upon which the filter acts (see Figure 2.3). v is the measurement noise of the background close to the line such that the measurement noise covariance matrix

$$\boldsymbol{V} = v \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{A.1}$$

Band	Index	f_0	f_c	Q	Δf	v
A2	1	184.718	184	31000	8	3.312e-04
B2	1	258.606	258	32000	8	1.706e-06
B2	2	295.550	295	49000	8	1.154e-05
C2	1	406.380	406	51000	8	1.063e-06
C2	2	465.694	465	78000	8	5.161e-07
D2	1	546.068	546	91000	8	5.282e-07
D2	2	564.070	566	94000	16	9.738e-07
D2	2	566.102	566	142000	16	9.738e-07
D2	2	568.050	566	95000	16	9.738e-07
D2	2	568.102	566	142000	16	9.738e-07
D2	3	591.098	589	74000	8	3.995e-07
E2	1	775.816	775	78000	8	5.773e-07
E2	2	812.760	812	68000	8	3.443e-07
E2	3	886.646	886	148000	8	3.106e-07
E2	4	931.386	931	233000	8	3.247e-07
F2	1	1366.803	1369	10000	8	3.543e-07
F2	2	1108.308	1108	277000	8	2.657e-07
F2	3	1182.194	1182	296000	8	2.833e-07
F2	4	1256.082	1254	628000	8	3.850e-07
F2	5	1329.970	1330	332000	8	3.385e-07
G2	1	1403.856	1403	351000	8	3.273e-07
G2	2	1551.632	1551	388000	8	3.630e-07
G2	3	1625.482	1628	406000	8	3.952e-07
G2	4	1699.404	1699	425000	8	3.948e-07
H2	1	1773.292	1773	443000	8	3.725e-07
H2	2	1847.180	1847	462000	8	3.256e-07
H2	3	1862.750	1866	465000	8	3.435e-07
H2	3	1862.772	1866	465000	8	3.435e-07

Table A.1. Hanford 4-km, S2 Kalman line parameters

Band	Index	f_0	f_c	Q	Δf	v
A2	1	162.000	162	405000	16	2.186e-05
B2	1	295.800	295	49000	16	6.849e-07
C2	1	406.724	406	102000	8	4.927e-07
C2	2	465.708	465	78000	8	5.676e-07
C2	3	486.000	486	122000	8	2.859e-07
D2	1	510.720	510	800	4	1.500e-06
D2	2	591.598	591	99000	8	1.902e-07
E2	1	698.454	698	134000	8	1.202e-07
E2	1	698.565	698	152000	8	1.202e-07
E2	1	698.643	698	152000	8	1.202e-07
E2	1	699.378	698	175000	8	1.202e-07
E2	2	931.418	931	155000	8	8.416e-08
E2	3	961.848	961	160000	16	6.777e-08
E2	4	972.000	972	162000	16	7.122e-08
E2	5	988.080	988	35000	8	6.957e-08
F2	1	1257.138	1254	314000	8	5.746e-08
G2	1	1719.500	1719	4000	8	8.180e-08
G2	2	1504.790	1506	5000	8	5.541e-08
G2	3	1528.240	1528	5000	8	7.736e-08
G2	4	1496.400	1496	5000	8	8.103e-08
G2	5	1453.790	1453	182000	8	7.085e-08
G2	5	1453.802	1453	103000	8	7.085e-08
G2	6	1463.540	1463	10000	8	5.882e-08
G2	7	1472.570	1472	5000	8	6.129e-08
G2	8	1481.470	1481	5000	8	6.348e-08
G2	9	1626.890	1626	406000	8	7.015e-08
G2	10	1689.160	1689	5000	8	7.923e-08
H2	1	1830.470	1830	5000	8	6.234e-08
H2	2	1862.834	1866	465000	8	6.956e-08
H2	3	1996.644	1996	499000	8	4.234e-08
H2	4	1976.196	1974	494000	8	7.671e-08
H2	5	1902.600	1903	5000	8	7.636e-08
H2	5	1903.500	1903	5000	8	7.636e-08

Table A.2. Hanford 2-km, S2 Kalman line parameters

Band	Index	f_0	f_c	Q	Δf	v
A2	1	140.605	140	2200	8	1.062e-07
A2	2	187.580	187	19000	8	3.299e-08
B2	1	221.206	221	55000	8	7.070e-08
B2	2	234.474	234	20000	8	1.649e-08
B2	3	281.370	281	20000	8	2.263e-08
B2	4	294.940	294	37000	8	8.962e-08
D2	1	516.146	516	86000	8	7.346e-08
D2	2	589.882	589	147000	16	4.128e-08
E2	1	737.352	737	184000	16	7.254e-08
E2	2	811.088	811	135000	16	8.782e-08
E2	3	884.822	884	147000	16	8.561e-08
E2	4	958.558	958	240000	16	7.763e-08
F2	1	1079.750	1080	3000	8	3.251e-08
F2	2	1106.028	1106	553000	8	3.191e-08
F2	3	1139.650	1140	3000	8	3.150e-08
F2	4	1179.764	1180	295000	8	3.249e-08
F2	5	1199.650	1200	3000	8	3.144e-08
F2	6	1243.700	1243.5	300000	8	5.670e-08
F2	7	1247.160	1247	300000	8	2.503e-08
F2	8	1260.000	1260	3000	8	3.449e-08
F2	9	1319.630	1319.63	3000	8	3.656e-08
F2	10	1327.234	1327	313000	8	3.532e-08

Table A.3. Livingston 4-km, S2 Kalman line parameters

Band	Index	f_0	f_c	Q	Δf	v
G2	1	1400.940	1400.97	350000	8	4.026e-08
G2	2	1407.100	1407.1	3500	8	4.177e-08
G2	3	1412.070	1412.07	300000	8	4.170e-08
G2	4	1415.020	1415	300000	8	3.095e-08
G2	5	1428.000	1428	500000	8	7.034e-08
G2	6	1431.500	1432	600000	8	2.771e-10
G2	7	1437.600	1437.6	359000	8	4.829e-08
G2	8	1439.530	1440	3000	8	4.316e-08
G2	9	1474.704	1478	737000	8	4.863e-08
G2	10	1500.000	1500	3000	8	4.743e-08
G2	11	1548.440	1548	387000	8	4.658e-08
G2	12	1559.500	1559.5	3000	8	5.006e-08
G2	13	1612.320	1612.32	50000	8	8.461e-10
G2	14	1615.880	1616	60000	8	3.588e-10
G2	15	1619.950	1619.95	80000	8	5.179e-10
G2	16	1620.200	1620.2	3000	8	5.463e-08
G2	17	1622.174	1622.17	405000	8	9.248e-08
G2	18	1680.000	1680	300	8	8.957e-08
G2	19	1695.910	1695.91	847000	8	9.165e-08
H2	1	1769.646	1769.65	442000	8	5.394e-08
H2	2	1799.500	1799.5	6000	8	5.335e-08
H2	3	1843.380	1843.38	921000	8	4.973e-08
H2	4	1860.100	1860.1	3000	8	4.797e-08
H2	5	1917.116	1917.12	479000	8	4.021e-08
H2	6	1920.000	1920	3000	8	4.027e-08
H2	7	1979.500	1979.5	3000	8	3.024e-08
H2	8	1990.850	1990.85	497000	8	3.012e-08
H2	9	2031.414	2028	1013000	8	2.911e-08
H2	10	2032.600	2034	1013000	8	2.142e-08
H2	11	2039.500	2039.5	3000	8	2.754e-08

Table A.4. Livingston 4-km, S2 Kalman line parameters continued

Band	Index	f_0	f_c	Q	Δf	v
A3	1	184.721	184.50	57439	8	8.5812e-07
A3	2	178.344	178.00	5360	8	7.8817e-07
B3	1	295.553	294.50	90460	12	3.3275e-06
B3	1	296.286	294.50	59248	12	3.3275e-06
B3	1	294.820	294.50	59754	12	3.3275e-06
B3	2	258.609	258.25	78708	8	9.3666e-07
B3	3	221.665	221.50	67922	8	7.1649e-07
C3	1	406.385	405.50	121361	16	2.6095e-06
C3	1	407.119	405.50	69779	16	2.6095e-06
C3	1	410.098	405.50	27171	16	2.6095e-06
C3	1	405.653	405.50	71470	16	2.6095e-06
C3	2	443.330	443.00	132723	8	1.7300e-06
C3	3	465.693	465.50	121331	8	1.7060e-06
D3	1	591.106	590.25	177343	12	1.8413e-06
D3	1	590.373	590.25	124736	12	1.8413e-06
D3	1	591.840	590.25	122069	12	1.8413e-06
D3	2	517.218	517.00	151213	8	1.5553e-06
D3	3	628.050	627.75	183616	8	1.0631e-06
D3	4	546.052	545.75	104658	8	1.4594e-06
D3	5	566.102	563.75	132537	16	2.4088e-06
D3	5	564.067	563.75	111928	16	2.4088e-06
D3	5	568.101	563.75	139885	16	2.4088e-06
D3	6	554.162	554.00	145531	8	1.4378e-06
D3	7	535.056	534.75	32872	8	1.5853e-06
D3	7	534.900	534.75	81605	8	1.5853e-06
E3	1	738.883	738.00	220593	12	1.3913e-06
E3	1	739.616	738.00	134574	12	1.3913e-06
E3	1	738.150	738.00	142383	12	1.3913e-06
E3	2	775.827	775.50	230807	8	7.9951e-07
E3	3	812.771	812.50	241657	8	7.3486e-07
F3	1	886.660	885.75	245345	12	1.0426e-06
F3	1	887.393	885.75	165479	12	1.0426e-06
F3	1	885.926	885.75	169739	12	1.0426e-06
F3	2	931.385	931.00	249564	8	7.0324e-07
F3	3	997.491	997.25	283184	8	7.3332e-07
F3	4	923.601	923.25	257940	8	6.9436e-07
F3	5	947.750	947.50	295644	8	7.1309e-07
F3	6	849.715	849.50	176666	8	7.0975e-07
F3	7	902.250	902.00	226767	8	6.9790e-07
F3	8	880.339	880.00	147200	8	7.8983e-07

Table A.5. Hanford 4-km, S3 Kalman line parameters

Band	Index	f_0	f_c	Q	Δf	v
G3	1	1256.100	1255.25	353120	12	1.7324e-06
G3	1	1256.834	1255.25	244113	12	1.7324e-06
G3	1	1255.368	1255.25	251989	12	1.7324e-06
G3	2	1329.989	1329.00	370711	12	1.9393e-06
G3	2	1330.723	1329.00	248832	12	1.9393e-06
G3	2	1329.256	1329.00	256971	12	1.9393e-06
G3	3	1182.212	1182.00	334731	8	1.0211e-06
G3	4	1108.324	1108.00	315056	8	8.6505e-07
G3	5	1293.045	1292.75	347216	8	1.2347e-06
G3	6	1071.380	1071.00	297310	8	7.9759e-07
G3	7	1145.268	1145.00	308879	8	9.2550e-07
G3	8	1219.156	1219.00	317695	8	1.0936e-06
G3	9	1360.250	1360.00	340112	8	1.5423e-06
G3	10	1239.750	1239.50	359985	8	1.1568e-06
H3	1	1442.000	1441.75	534074	12	3.2950e-06
H3	1	1442.733	1441.75	300611	12	3.2950e-06
H3	2	1625.542	1625.25	432878	8	2.5889e-06
H3	3	1699.430	1699.25	441708	8	2.8208e-06
H3	4	1551.653	1551.50	419329	8	2.2518e-06
H3	5	1663.750	1662.25	514580	12	4.0169e-06
H3	5	1662.486	1662.25	395461	12	4.0169e-06
H3	6	1477.765	1477.50	370571	8	5.0368e-06
H3	7	1672.750	1672.50	531144	8	2.6860e-06
H3	8	1577.250	1577.00	487712	8	2.3030e-06
H3	9	1426.684	1426.50	153660	8	2.2018e-06
H3	10	1486.204	1486.00	274980	8	2.1738e-06
I3	1	1773.318	1773.00	462004	8	3.7692e-06
I3	2	1847.207	1847.00	471483	8	2.7871e-06
I3	3	1994.983	1994.75	512935	8	1.8392e-06
I3	4	1862.771	1862.50	426324	8	2.7065e-06
R3	1	931.386	931.25	294878	12	9.6572e-07
R3	1	933.750	931.25	200081	12	9.6572e-07
R3	2	923.608	923.50	241796	16	1.3367e-06
R3	2	927.250	923.50	217876	16	1.3367e-06

Table A.6. Hanford 4-km, S3 Kalman line parameters continued

Band	Index	f_0	f_c	Q	Δf	v
A3	1	184.870	184.50	59067	8	1.7559e-06
A3	2	137.625	137.25	8695	8	2.5259e-06
A3	2	137.439	137.25	8575	8	2.5259e-06
A3	2	137.784	137.25	8364	8	2.5259e-06
A3	3	155.642	155.25	7833	8	3.2137e-05
A3	4	150.355	150.00	11641	8	2.5948e-06
B3	1	295.792	295.50	88803	8	1.7635e-06
B3	2	258.819	257.50	69907	12	1.6850e-06
B3	2	257.620	257.50	15151	12	1.6850e-06
B3	2	259.403	257.50	19570	12	1.6850e-06
B3	3	221.844	221.50	64724	8	2.9281e-06
B3	4	267.000	266.75	98617	8	1.0981e-06
C3	1	432.999	431.50	81801	16	8.5591e-07
C3	1	433.173	431.50	24919	16	8.5591e-07
C3	1	432.829	431.50	25124	16	8.5591e-07
C3	1	432.231	431.50	30828	16	8.5591e-07
C3	1	433.761	431.50	30817	16	8.5591e-07
C3	1	434.398	431.50	23383	16	8.5591e-07
C3	1	431.597	431.50	24312	16	8.5591e-07
C3	2	406.715	406.50	122006	8	4.4874e-07
C3	3	443.689	443.50	130522	8	4.1669e-07
C3	4	400.002	399.75	81434	8	4.6002e-07
D3	1	591.585	591.25	176834	8	4.3434e-07
D3	2	517.637	517.50	123820	16	7.5280e-07
D3	2	520.865	517.50	121268	16	7.5280e-07
D3	3	554.611	554.25	140740	8	4.8240e-07
D3	4	512.726	512.50	132403	8	3.9184e-07
D3	5	529.003	528.75	111044	8	4.0161e-07
D3	6	537.142	536.75	125030	8	3.1701e-06
D3	7	545.279	545.00	141549	8	4.8657e-07
D3	8	569.695	569.50	127039	8	4.5525e-07
D3	9	534.000	533.75	189520	8	4.2397e-07
D3	10	551.299	551.00	44174	8	4.3647e-07
D3	11	570.772	570.50	118490	8	4.5411e-07

Table A.7. Hanford 2-km, S3 Kalman line parameters

Band	Index	f_0	f_c	Q	Δf	v
E3	1	739.481	739.25	211902	12	1.4206e-07
E3	1	740.604	739.25	190191	12	1.4206e-07
E3	2	776.455	773.00	213934	16	3.3788e-07
E3	2	773.158	773.00	192801	16	3.3788e-07
E3	3	813.429	813.25	212025	8	7.6753e-08
E3	4	716.189	716.00	150390	8	2.5970e-07
E3	5	708.050	707.75	168302	8	1.0928e-07
E3	6	765.020	764.75	197130	8	8.9535e-08
E3	7	724.327	724.00	150323	8	2.4429e-07
E3	8	732.467	732.25	175248	8	9.8033e-08
E3	9	801.000	800.75	269610	8	7.8696e-08
F3	1	866.000	865.75	267239	8	8.1725e-08
F3	2	1000.000	998.00	361785	12	7.7418e-08
F3	2	998.299	998.00	263271	12	7.7418e-08
F3	3	887.377	887.00	251798	8	6.2372e-08
G3	1	1331.066	1330.75	379092	8	5.0779e-08
G3	2	1298.999	1298.75	359094	8	4.9757e-08
G3	3	1257.118	1256.75	357617	8	8.9992e-08
G3	4	1072.247	1072.00	297874	8	5.5602e-08
G3	5	1183.169	1183.00	325128	8	5.1539e-08
H3	1	1626.858	1626.50	467533	8	6.0395e-08
H3	2	1700.806	1700.50	461920	8	6.6547e-08
H3	3	1552.910	1552.75	426357	8	5.8778e-08
H3	4	1478.962	1478.75	372932	8	5.4200e-08
H3	5	1587.099	1586.75	139230	8	7.8578e-07
H3	6	1512.723	1512.50	67244	8	8.8186e-08
H3	7	1474.468	1474.25	250750	8	6.6713e-08
H3	8	1418.361	1418.00	273790	8	5.9116e-08
R3	1	903.373	903.00	194739	8	2.6961e-07
R3	2	931.418	927.50	250128	16	1.1510e-07
R3	2	927.790	927.50	243060	16	1.1510e-07
R3	3	911.514	911.25	166713	8	7.2944e-08
R3	4	895.236	895.00	230290	8	6.5898e-08
R3	5	935.928	935.75	227913	8	6.8618e-08
R3	6	919.652	919.50	196327	8	5.9984e-08

Table A.8. Hanford 2-km, S3 Kalman line parameters continued

Band	Index	f_0	f_c	Q	Δf	v
A3	1	143.147	142.75	44455	8	1.6043e-06
A3	2	190.862	190.50	50761	8	2.0054e-06
A3	3	131.857	131.50	21600	8	1.7299e-06
B3	1	237.788	237.50	21486	16	4.0554e-06
B3	1	241.544	237.50	13344	16	4.0554e-06
B3	1	241.960	237.50	26395	16	4.0554e-06
B3	2	251.756	251.50	14331	8	2.3387e-06
B3	3	294.937	294.75	79712	8	2.2760e-06
B3	4	221.203	221.00	54617	8	2.2223e-06
B3	5	242.169	242.00	39484	8	2.4708e-06
B3	6	258.070	257.75	56511	8	2.3401e-06
C3	1	405.538	405.25	135179	8	1.7829e-06
C3	2	442.405	442.25	118199	8	1.6988e-06
D3	1	516.139	515.75	118847	8	1.3204e-06
D3	2	589.874	589.50	174962	8	1.1622e-06
D3	3	626.742	626.50	181664	8	1.2686e-06
D3	4	553.007	552.75	146169	8	1.2159e-06
E3	1	737.342	737.00	215957	8	1.0476e-06
E3	2	811.076	810.75	240573	8	9.4177e-07
E3	3	774.210	774.00	202976	8	9.8769e-07
F3	1	884.803	884.50	213849	8	9.1640e-07
F3	2	995.410	995.25	269029	8	9.7339e-07
F3	3	921.672	921.50	253206	32	2.8254e-06
F3	3	925.606	921.50	112878	32	2.8254e-06
F3	3	929.264	921.50	101837	32	2.8254e-06
F3	3	929.726	921.50	36147	32	2.8254e-06
F3	3	933.609	921.50	193828	32	2.8254e-06
F3	4	847.944	847.75	250623	8	9.4388e-07
F3	5	926.090	925.75	35257	8	1.5033e-06
F3	5	926.049	925.75	142469	8	1.5033e-06
F3	6	939.556	939.25	65398	8	9.5486e-07
F3	7	915.825	915.50	43819	8	9.1476e-07
F3	8	958.543	958.25	225380	8	9.6067e-07

Table A.9. Livingston 4-km, S3 Kalman line parameters

Band	Index	f_0	f_c	Q	Δf	v
G3	1	1106.013	1105.75	302425	8	1.1182e-06
G3	2	1327.216	1327.00	388724	8	1.8624e-06
G3	3	1253.482	1253.25	378205	8	1.5780e-06
G3	4	1179.747	1179.50	342665	8	1.3382e-06
G3	5	1142.880	1142.50	312506	8	1.2131e-06
G3	6	1216.615	1216.25	299869	8	1.4450e-06
G3	7	1364.084	1363.75	231200	8	1.9808e-06
G3	8	1069.147	1068.75	311402	8	1.0640e-06
G3	9	1334.164	1334.00	115762	8	1.9289e-06
H3	1	1695.887	1691.75	418000	16	6.2473e-06
H3	1	1692.000	1691.75	143086	16	6.2473e-06
H3	1	1692.439	1691.75	131536	16	6.2473e-06
H3	2	1474.684	1474.50	355958	8	2.4964e-06
H3	3	1548.419	1548.25	399960	8	2.7818e-06
H3	4	1622.153	1622.00	445296	8	2.8605e-06
H3	5	1511.551	1511.25	345779	8	2.6414e-06
H3	6	1437.819	1437.50	359454	8	2.6093e-06
H3	7	1585.289	1585.00	332578	8	2.8504e-06
I3	1	1769.621	1769.25	463945	8	2.4232e-06
I3	2	1843.356	1843.00	492499	8	1.9138e-06
I3	3	1917.090	1916.75	530420	8	1.7582e-06
I3	4	1957.703	1957.50	431212	8	1.6844e-06
I3	5	2031.437	2031.25	447453	8	1.7804e-06
I3	6	1867.219	1867.00	369746	8	1.7910e-06
I3	7	1806.487	1806.25	470439	8	2.1561e-06
I3	8	1990.826	1990.50	530886	8	1.6808e-06
I3	9	1761.500	1758.25	338749	16	4.9036e-06
I3	9	1758.472	1758.25	254851	16	4.9036e-06
R3	1	921.679	921.50	237196	8	9.2760e-07
R3	2	929.316	926.00	64017	16	1.8036e-06
R3	2	926.096	926.00	49790	16	1.8036e-06
R3	3	933.610	933.25	229576	8	9.6617e-07
R3	4	939.562	939.25	27715	8	9.5488e-07
R3	5	915.814	915.50	42268	8	9.1521e-07

Table A.10. Livingston 4-km, S3 Kalman line parameters continued

Appendix B

Frequency Band Definitions and Filters

The following tables give the band definitions and specify parameters for the lowpass filters used in basebanding. Each band covers the frequency range from B_{lo} to B_{hi} Hz. There are gaps in frequency coverage around the violin modes and violin mode harmonics. The violin modes have a lot of power and can vary considerably in power over short timescales, making them difficult to remove. Therefore, we have chosen to avoid them. The R band is a narrow band chosen to correspond to the frequency range of some bar detectors, allowing comparison with their results.

The lowpass filter used in the basebanding step of data conditioning is created using MATLAB's [84] `firpm` function which uses the Parks McClellan algorithm to find an optimal filter. The filter `b` is calculated via the command `b = firpm(n,f,a,w)` where n is the length of the filter, f = [0, T_{lo}, T_{hi}, 1]' is a vector of frequencies (as fraction of Nyquist) and a = [1, 1, 0, 0]' are the amplitudes of the filter's frequency response at these frequencies. w = [1, 1]' specifies the allowed amount of passband and stopband ripple (see [36]). n, T_{lo}, and T_{hi} are given in the tables.

Band	B_{lo}	B_{hi}	n	T_{lo}	T_{hi}
A2	128	192	52	0.4688	0.5469
B2	192	320	56	0.4688	0.5391
C2	384	512	32	0.5000	0.6250
D2	512	640	56	0.4688	0.5391
E2	704	1024	70	0.5000	0.5563
F2	1065	1365	48	0.4833	0.5667
G2	1408	1708	48	0.4833	0.5667
H2	1758	2048	40	0.5000	0.5862
R2	900	930	40	0.5000	0.6000

Table B.1. S2 Bands

Band	B_{lo}	B_{hi}	n	T_{lo}	T_{hi}
A3	128	192	26	0.4375	0.5938
B3	192	320	36	0.4531	0.5625
C3	384	512	32	0.4531	0.5781
D3	512	640	32	0.4531	0.5781
E3	704	832	50	0.4688	0.5469
F3	832	1024	76	0.4792	0.5312
G3	1065	1365	74	0.4800	0.5333
H3	1408	1708	74	0.4800	0.5333
I3	1758	2048	72	0.4793	0.5345
R3	890	940	20	0.4200	0.6200

Table B.2. S3 Bands

Band	B_{lo}	B_{hi}	n	T_{lo}	T_{hi}
A4	96	192	38	0.4583	0.5625
B4	192	320	32	0.4531	0.5781
C4	384	512	28	0.4531	0.5938
D4	512	640	28	0.4531	0.5938
E4	704	832	50	0.4688	0.5469
F4	832	1024	76	0.4792	0.5312
G4	1065	1365	74	0.4800	0.5333
H4	1408	1708	74	0.4800	0.5333
I4	1758	2048	72	0.4793	0.5345
R4	890	940	20	0.4375	0.5833

Table B.3. S4 Bands

Appendix C

Line Regression Parameters

The following tables give the parameters used in line regression for S2 and S3 for each detector. The parameters are given by frequency band and frequency band definitions may be found in Appendix B. There are two types of lines that are regressed: power lines at 60 Hz and harmonics caused by the AC oscillations of the US power grid and calibration lines which are injected into the interferometers by wiggling mirrors and used for calibration purposes. Parameters used to remove power lines are marked as "power" in the table while those used to remove calibration lines are designated "calib".

When line regression is performed, data from an environmental or interferometer channel is put through the same conditioning and bandlimiting procedure (minus Kalman filtering) as the data from the gravitational wave channel of the detector. An ARX model is then found that relates the regression channel data to the detector data. The regression channel is filtered using the ARX model and the result subtracted from the detector data. If the bandlimited detector data is stored in the variable `band_data` and the bandlimited regression data in `band_reg`, the lines can be regressed using the following MATLAB [84] commands:

```
data= iddata(band_data, band_reg);
model = arx(data, orders);
b = model.b;
filt_data = band_data-filter(b,1,band_reg);
```

`orders` is a vector whose three elements are given in the tables below.

Two channels are listed for some of the power line regression parameters because, in those cases, regression based on no single channel could remove both the even and odd harmonics. When two channels are listed, the power lines are subtracted as much as possible with the first channel and then the filtered data is used to find the b coefficients for the second channel:

```
data1 = iddata(band_data', band_reg1');
model1 = arx(data1, orders1);
b1 = model1.b;
filt_data1 = band_data-filter(b1,1,band_reg1);
data2 = iddata(filt_data1', band_reg2');
model2 = arx(data2, orders2);
b2 = model2.b;
```

When the data are filtered, both sets of channels and filters are used:

```
filt_data = band_data-filter(b1,1,band_reg1) ...
- filter(b2,1,band_reg2);
```

This regression approach, where the filter for the second channel is dependent on regression applied with the first channel is specified by the column labeled Dep. in the tables below. A nonzero Dep. value indicates that dependent regression is used and the number given is the Index value of the regression parameters that should be applied first.

As can be seen from Tables C and C, regression is not used to remove power lines from Livingston 4-km data above 1024 Hz. Unfortunately, no suitable channel could be found. L1:LSC-POB_Q was not correlated enough with the power lines to be used to remove them and the environmental channels contain no information at those frequencies because they are sampled at 2048Hz. Instead, these lines are treated with a Kalman filter which does not perform as well as regression on power lines but is able to accomplish some reduction of power.

Band	Index	Type	Channel	orders			Dep.
A2	1	power	H0:PEM-LVEA2_V1	0	6	0	0
A2	2	calib	H1:LSC-ETMX_EXC_DAQ	0	6	0	0
B2	1	power	H0:PEM-LVEA2_V1	0	20	0	0
C2	1	power	H0:PEM-LVEA2_V1	0	16	0	0
D2	1	power	H0:PEM-LVEA2_V1	0	8	0	0
E2	1	power	H0:PEM-LVEA2_V1	0	90	0	0
E2	2	calib	H1:LSC-ETMX_EXC_DAQ	0	8	0	0
F2	1	power	H1:LSC-POB_Q	0	90	0	0
G2	1	power	H1:LSC-POB_Q	0	90	0	0
H2	1	power	H1:LSC-POB_Q	0	90	0	0
R2	1	power	H0:PEM-LVEA2_V1	0	10	0	0

Table C.1. Hanford 4-km, S2 line regression parameters

Band	Index	Type	Channel	orders			Dep.
A2	1	power	H0:PEM-LVEA2_V3	0	6	0	0
A2	2	calib	H2:LSC-ETMX_EXC_DAQ	0	6	0	0
B2	1	power	H0:PEM-LVEA2_V3	0	40	0	0
C2	1	power	H0:PEM-LVEA2_V3	0	50	0	0
D2	1	power	H0:PEM-LVEA2_V3	0	40	0	0
E2	1	power	H0:PEM-LVEA2_V3	0	90	0	0
E2	2	calib	H2:LSC-ETMX_EXC_DAQ	0	8	0	0
F2	1	power	H2:LSC-POB_Q	0	90	0	0
G2	1	power	H2:LSC-POB_Q	0	90	0	0
H2	1	power	H2:LSC-POB_Q	0	90	0	0
R2	1	power	H0:PEM-LVEA2_V3	0	10	0	0

Table C.2. Hanford 2-km, S2 line regression parameters

Band	Index	Type	Channel	orders			Dep.
A2	1	power	L0:PEM-LVEA_V1	0	6	0	0
A2	2	calib	L1:LSC-ETMX_EXC_DAQ	0	6	0	0
B2	1	power	L0:PEM-RADIO_LVEA	0	50	0	0
B2	2	power	L0:PEM-LVEA_V1	0	50	0	1
C2	1	power	L0:PEM-RADIO_LVEA	0	50	0	0
C2	2	power	L0:PEM-LVEA_V1	0	50	0	1
D2	1	power	L0:PEM-RADIO_LVEA	0	70	0	0
D2	2	power	L0:PEM-LVEA_V1	0	70	0	1
E2	1	power	L0:PEM-RADIO_LVEA	0	90	0	0
E2	2	power	L0:PEM-LVEA_V1	0	90	0	1
E2	3	power	L1:LSC-ETMX_EXC_DAQ	0	8	0	1
R2	1	power	L0:PEM-LVEA_V1	0	10	0	0
R2	2	calib	L1:LSC-ETMX_EXC_DAQ	0	10	0	0

Table C.3. Livingston 4-km, S2 line regression parameters

Band	Index	Type	Channel	orders			Dep.
A3	1	power	H0:PEM-LVEA2_V1	0	2	0	0
A3	2	calib	H1:LSC-DARM_CTRL_EXC_DAQ	0	10	0	0
B3	1	power	H0:PEM-LVEA2_V1	0	150	0	0
C3	1	power	H0:PEM-LVEA2_V1	0	16	0	0
D3	1	power	H0:PEM-LVEA2_V1	0	8	0	0
E3	1	power	H0:PEM-LVEA2_V1	0	150	0	0
F3	1	calib	H1:LSC-DARM_CTRL_EXC_DAQ	0	4	0	0
F3	2	power	H0:PEM-LVEA2_V1	0	50	0	1
G3	1	power	H1:LSC-POB_Q	0	90	0	0
H3	1	power	H1:LSC-POB_Q	0	90	0	0
I3	1	power	H1:LSC-POB_Q	0	10	0	0
R3	1	power	H0:PEM-LVEA2_V1	0	10	0	0

Table C.4. Hanford 4-km, S3 line regression parameters

Band	Index	Type	Channel	orders			Dep.
A3	1	power	H0:PEM-LVEA2_V3	0	2	0	0
A3	2	calib	H2:LSC-DARM_CTRL_EXC_DAQ	0	6	0	0
B3	1	power	H0:PEM-LVEA2_V3	0	50	0	0
C3	1	power	H0:PEM-LVEA2_V3	0	50	0	0
D3	1	power	H0:PEM-LVEA2_V3	0	40	0	0
E3	1	power	H0:PEM-LVEA2_V3	0	50	0	0
F3	1	calib	H2:LSC-DARM_CTRL_EXC_DAQ	0	10	0	0
F3	2	power	H0:PEM-LVEA2_V3	0	150	0	0
G3	1	power	H2:LSC-POB_Q	0	150	0	0
H3	1	power	H2:LSC-POB_Q	0	150	0	0
H3	2	calib	H2:LSC-DARM_CTRL_EXC_DAQ	0	10	0	0
R3	1	power	H0:PEM-LVEA2_V3	0	50	0	0

Table C.5. Hanford 2-km, S3 line regression parameters

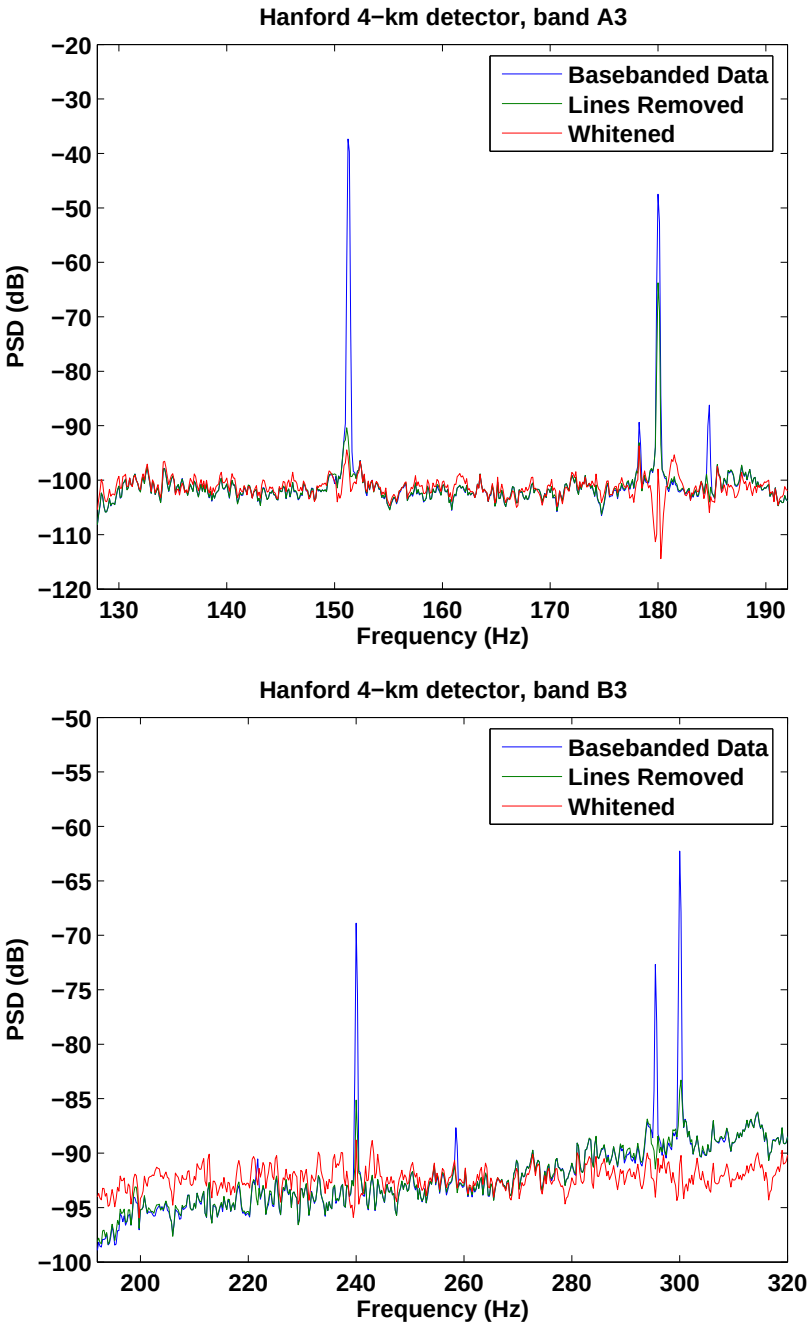
Band	Index	Type	Channel	orders			Dep.
A3	1	power	L0:PEM-LVEA_V1	0	2	0	0
A3	2	calib	L1:LSC-DARM_CTRL_EXC_DAQ	0	22	0	0
B3	1	power	L0:PEM-RADIO_LVEA	0	2	0	0
B3	2	power	L0:PEM-LVEA_V1	0	2	0	1
C3	1	power	L0:PEM-RADIO_LVEA	0	2	0	0
C3	2	power	L0:PEM-LVEA_V1	0	50	0	1
D3	1	power	L0:PEM-RADIO_LVEA	0	8	0	0
D3	2	power	L0:PEM-LVEA_V1	0	50	0	1
E3	1	power	L0:PEM-RADIO_LVEA	0	90	0	0
E3	2	power	L0:PEM-LVEA_V1	0	90	0	1
F3	1	power	L0:PEM-RADIO_LVEA	0	90	0	0
F3	2	power	L0:PEM-LVEA_V1	0	90	0	1
F3	3	calib	L1:LSC-DARM_CTRL_EXC_DAQ	0	4	0	0
R3	1	power	L0:PEM-LVEA_V1	0	10	0	0
R3	2	calib	L1:LSC-DARM_CTRL_EXC_DAQ	0	10	0	0

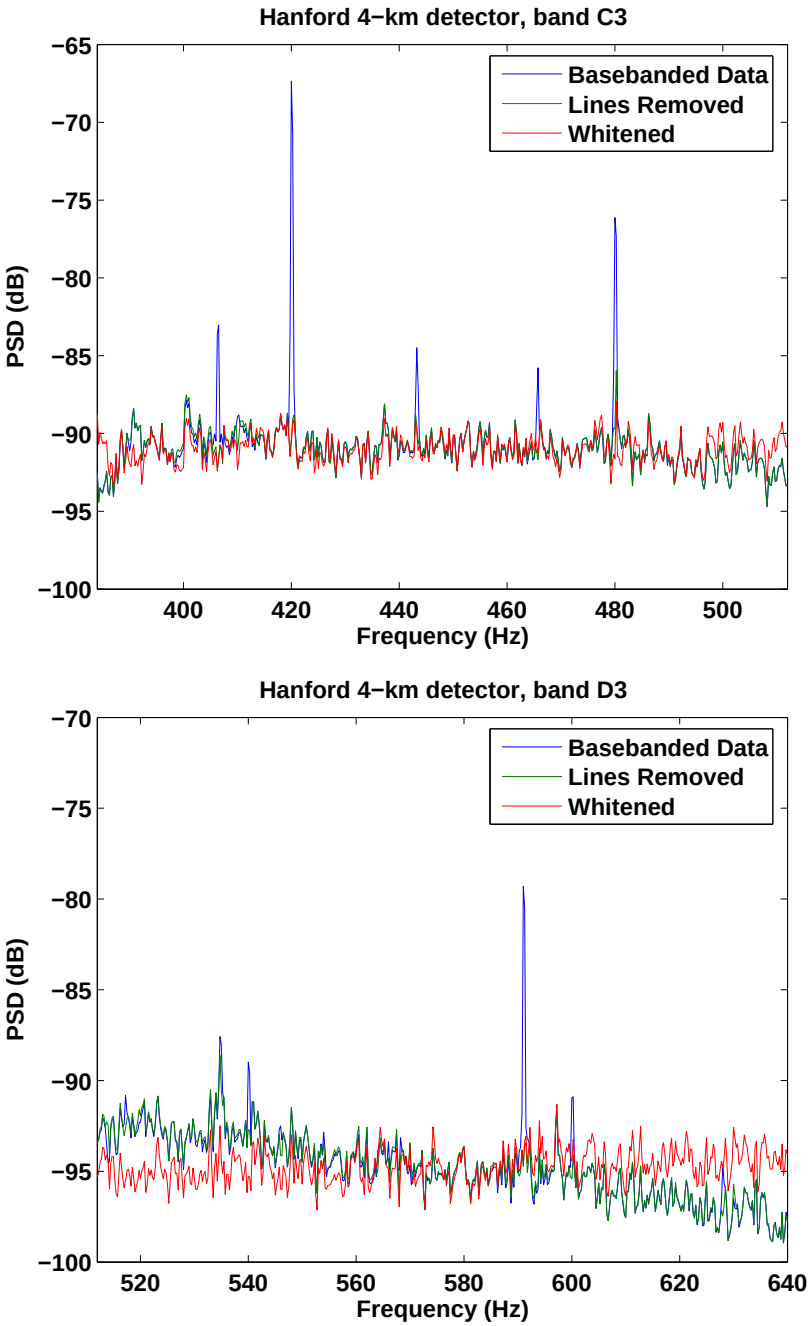
Table C.6. Livingston 4-km, S3 line regression parameters

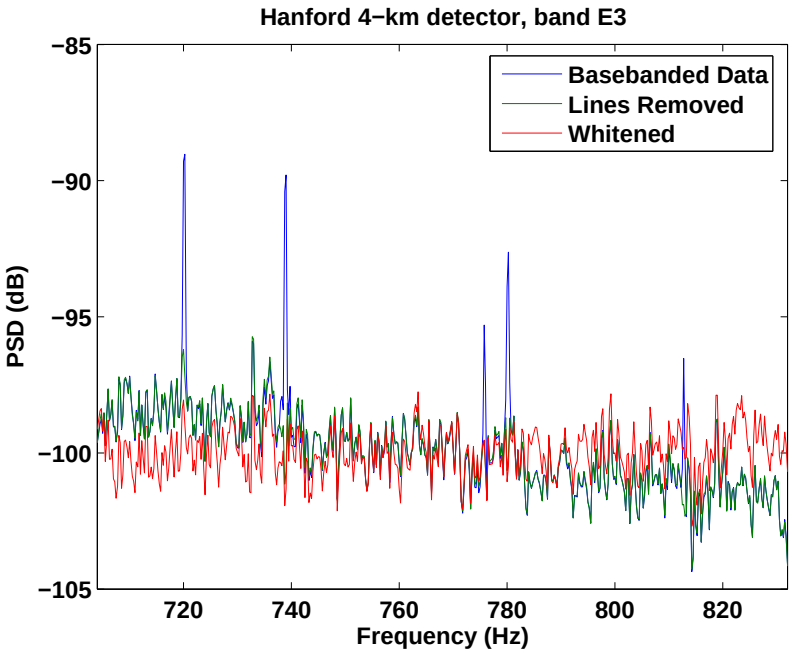
Appendix D

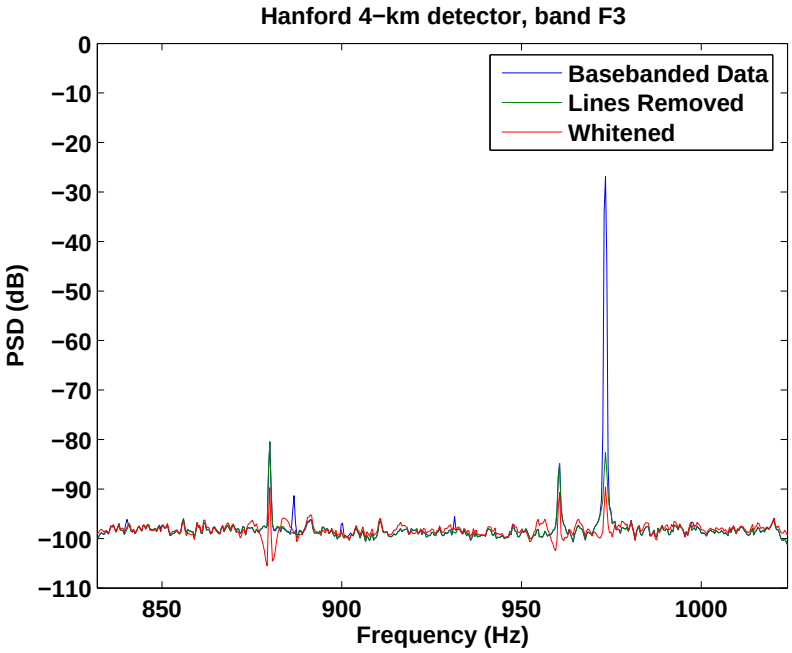
Data conditioning PSDs

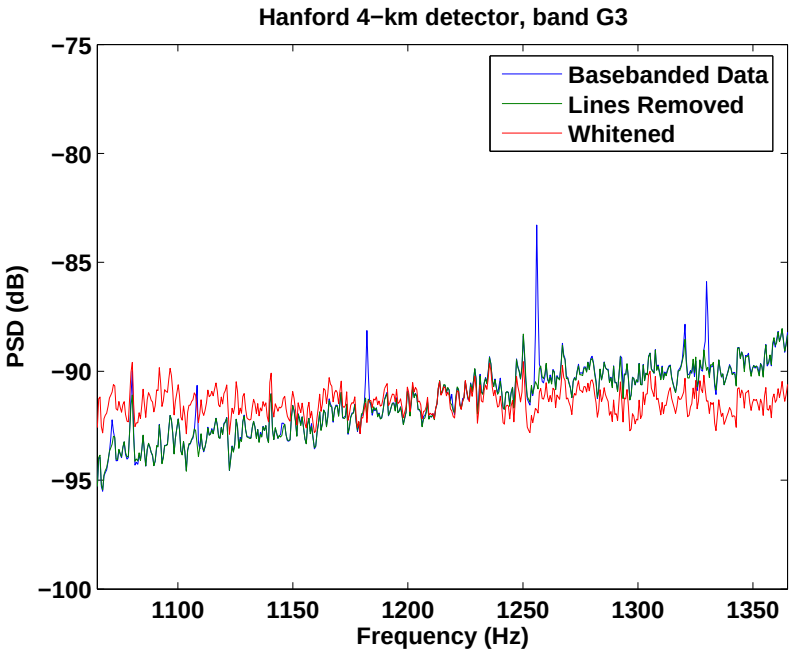
Below are power spectral density (PSD) plots that show the effects of different data conditioning steps. Each plot shows the effects of data conditioning for a different frequency band and detector. The frequency band definitions are given in Appendix B. The plots are made with data from GPS time 756715873 which is 300 seconds in duration. The data conditioning uses filters calculated on the previous 300 seconds. In each plot, the blue curve is the PSD of data that has been basebanded only, the green curve shows the effect of line removal including Kalman filtering and line regression, and the red curve is the PSD of the fully conditioned data including whitening.

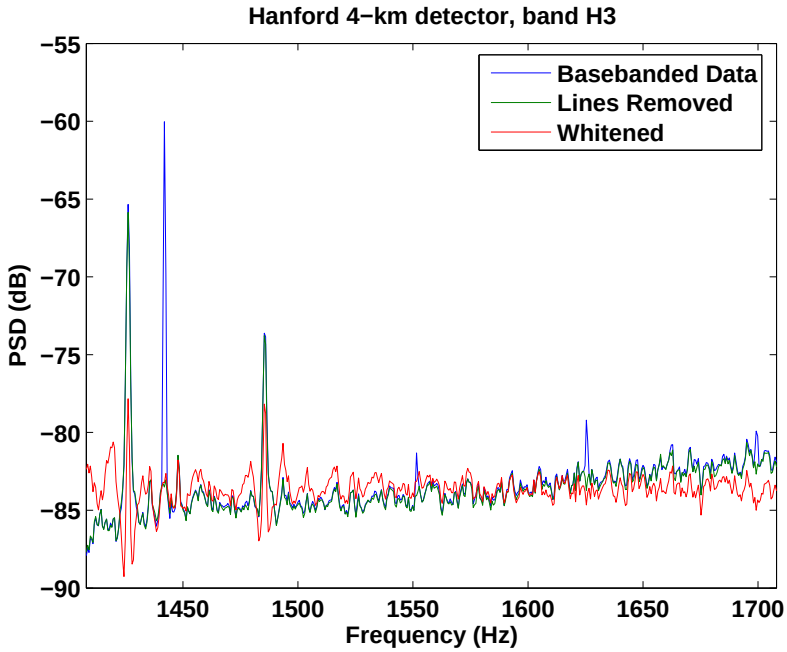

Hanford 4−km detector, band A3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−20
−30
−40
−50
−60
−70
−80
−90
−100
−110
−120
130
140
150
160
170
180
190
Frequency (Hz)
Hanford 4−km detector, band B3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−50
−55
−60
−65
−70
−75
−80
−85
−90
−95
−100
200
220
240
260
280
300
320
Frequency (Hz)

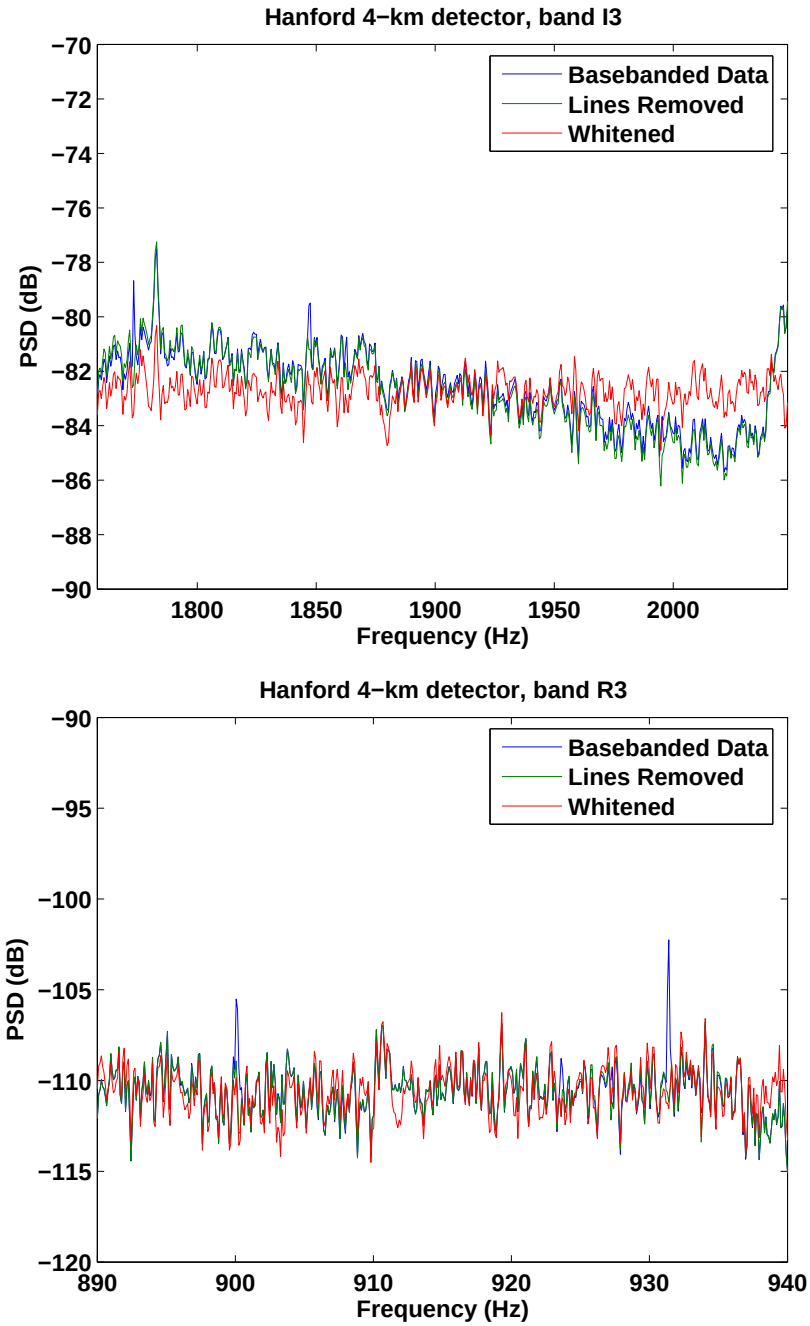

Hanford 4−km detector, band C3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−65
−70
−75
−80
−85
−90
−95
−100
400
420
440
460
480
500
Frequency (Hz)
Hanford 4−km detector, band D3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−70
−75
−80
−85
−90
−95
−100
520
540
560
580
600
620
640
Frequency (Hz)

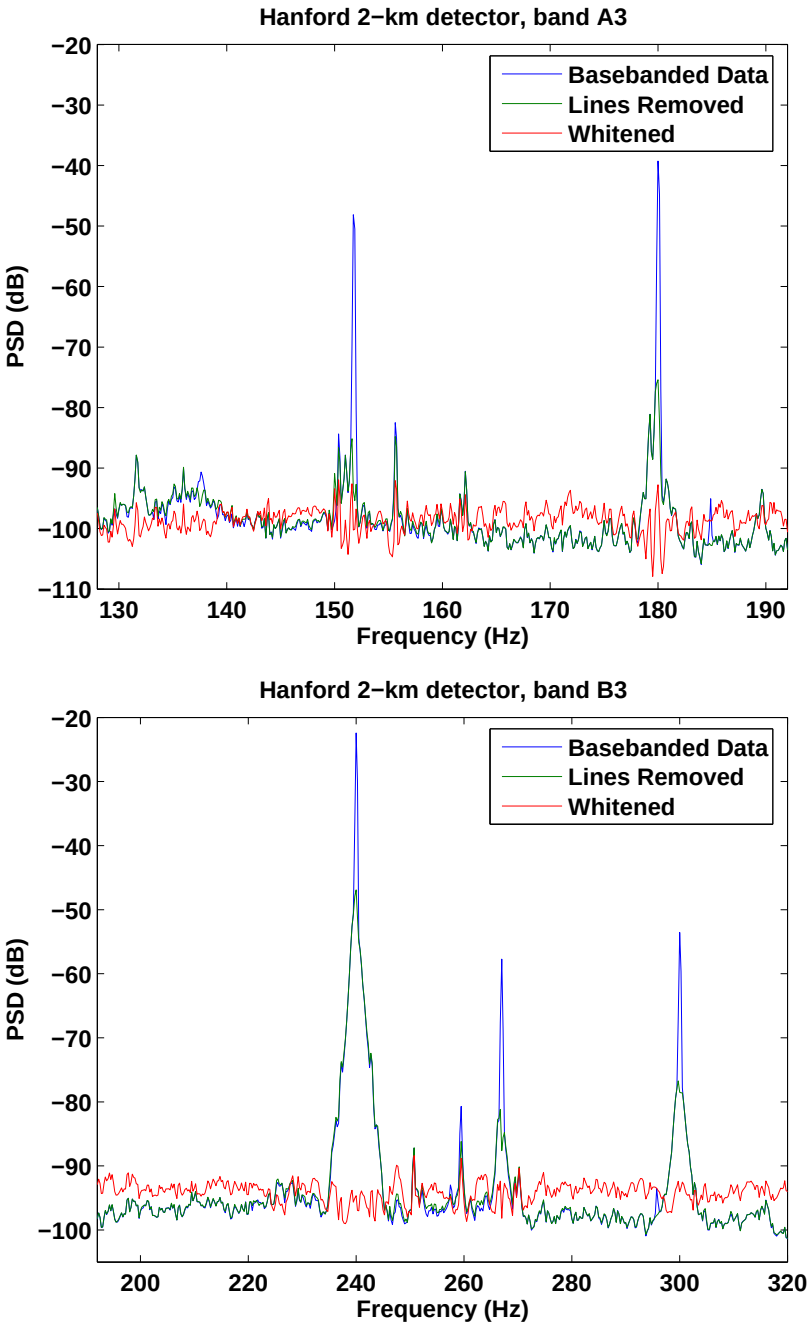

Hanford 4−km detector, band E3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−85
−90
−95
−100
−105
720
740
760
780
800
820
Frequency (Hz)

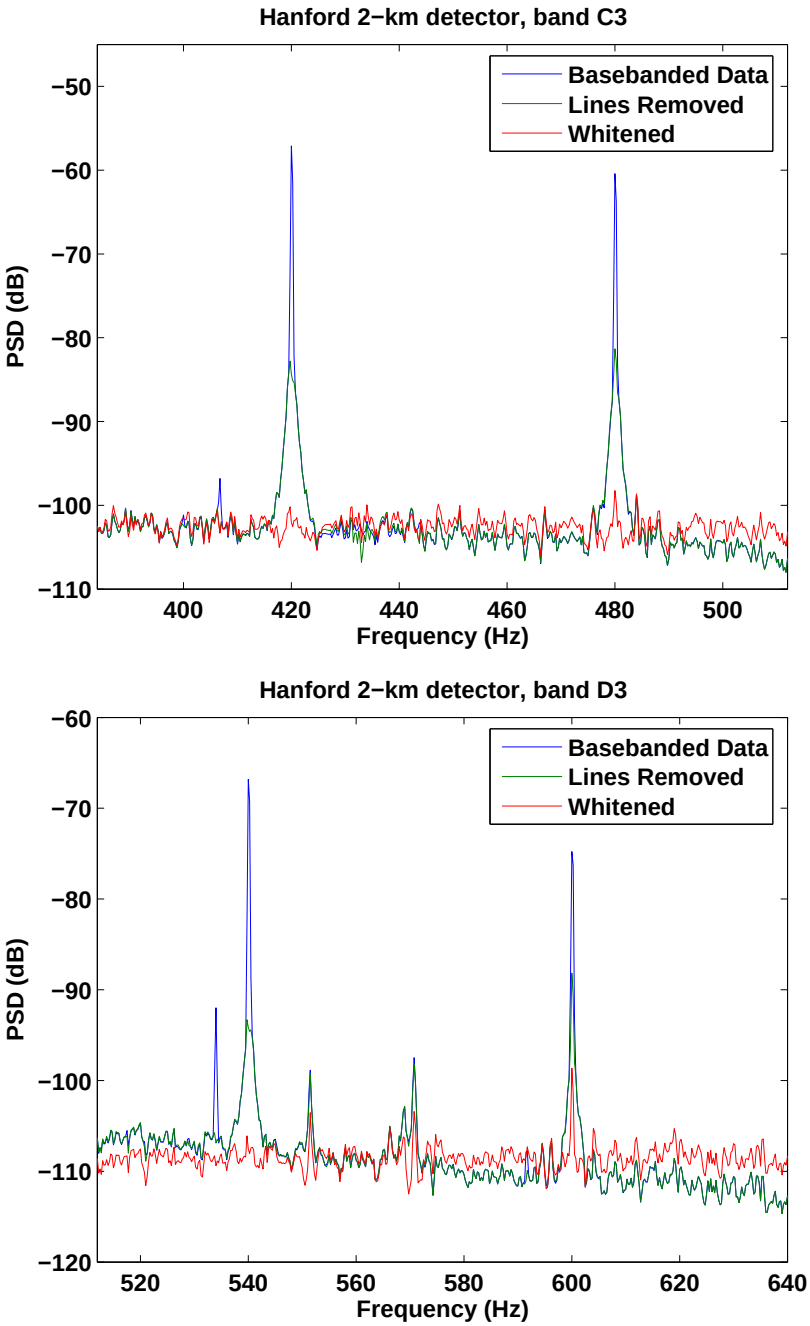

Hanford 4−km detector, band F3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
0
−10
−20
−30
−40
−50
−60
−70
−80
−90
−100
−110
850
900
950
1000
Frequency (Hz)

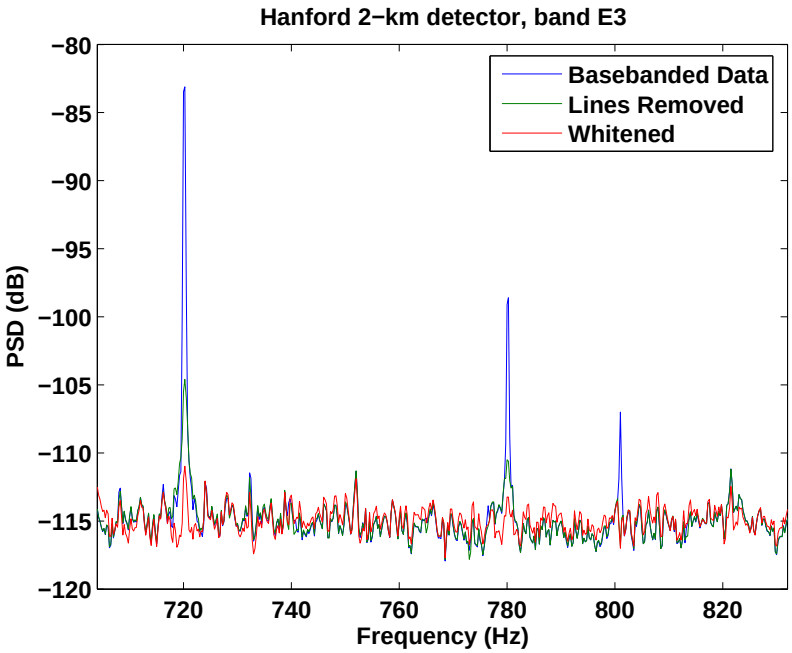

Hanford 4−km detector, band G3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−75
−80
−85
−90
−95
−100
1100
1150
1200
1250
1300
1350
Frequency (Hz)

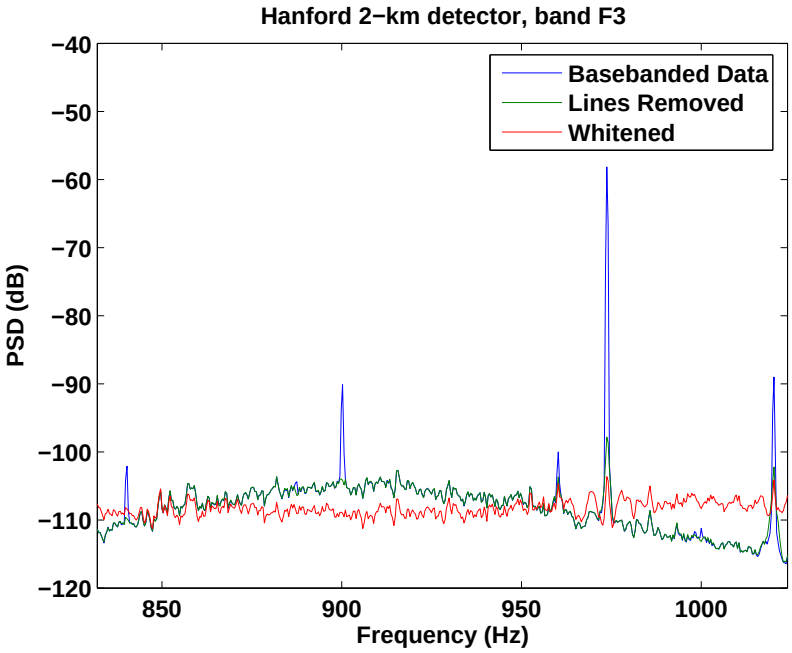

Hanford 4−km detector, band H3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−55
−60
−65
−70
−75
−80
−85
−90
1450
1500
1550
1600
1650
1700
Frequency (Hz)

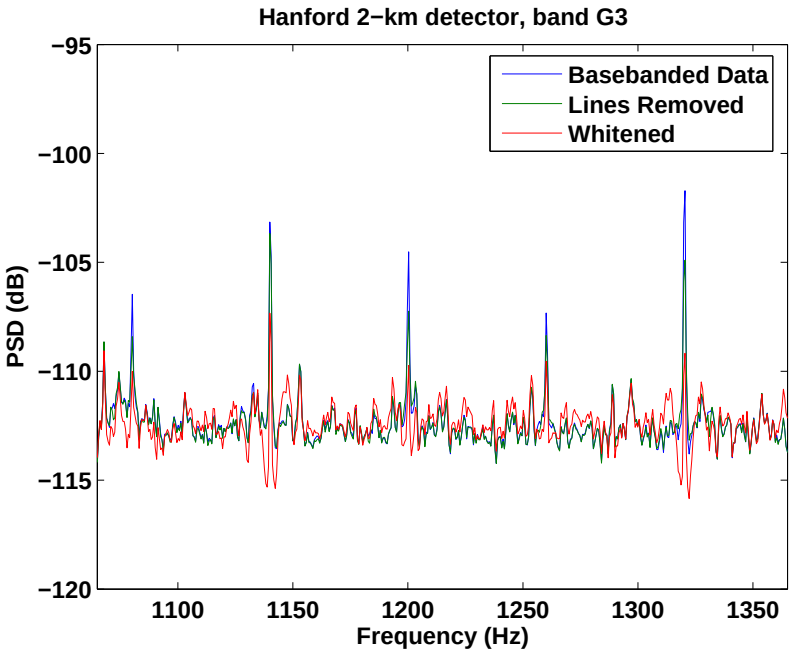

Hanford 4−km detector, band I3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−70
−72
−74
−76
−78
−80
−82
−84
−86
−88
−90
1800
1850
1900
1950
2000
Frequency (Hz)
Hanford 4−km detector, band R3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−90
−95
−100
−105
−110
−115
−120
890
900
910
920
930
940
Frequency (Hz)

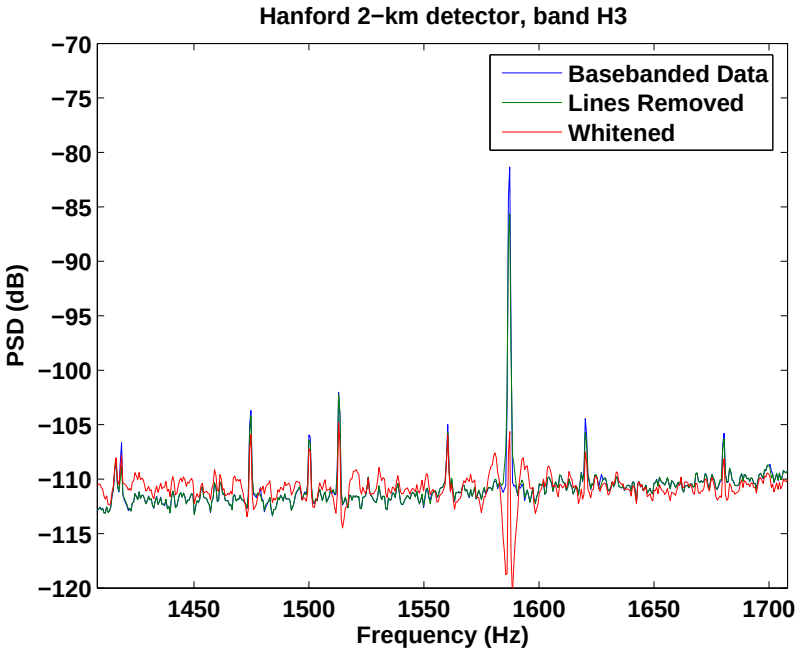

Hanford 2−km detector, band A3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−20
−30
−40
−50
−60
−70
−80
−90
−100
−110
130
140
150
160
170
180
190
Frequency (Hz)
Hanford 2−km detector, band B3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−20
−30
−40
−50
−60
−70
−80
−90
−100
200
220
240
260
280
300
320
Frequency (Hz)

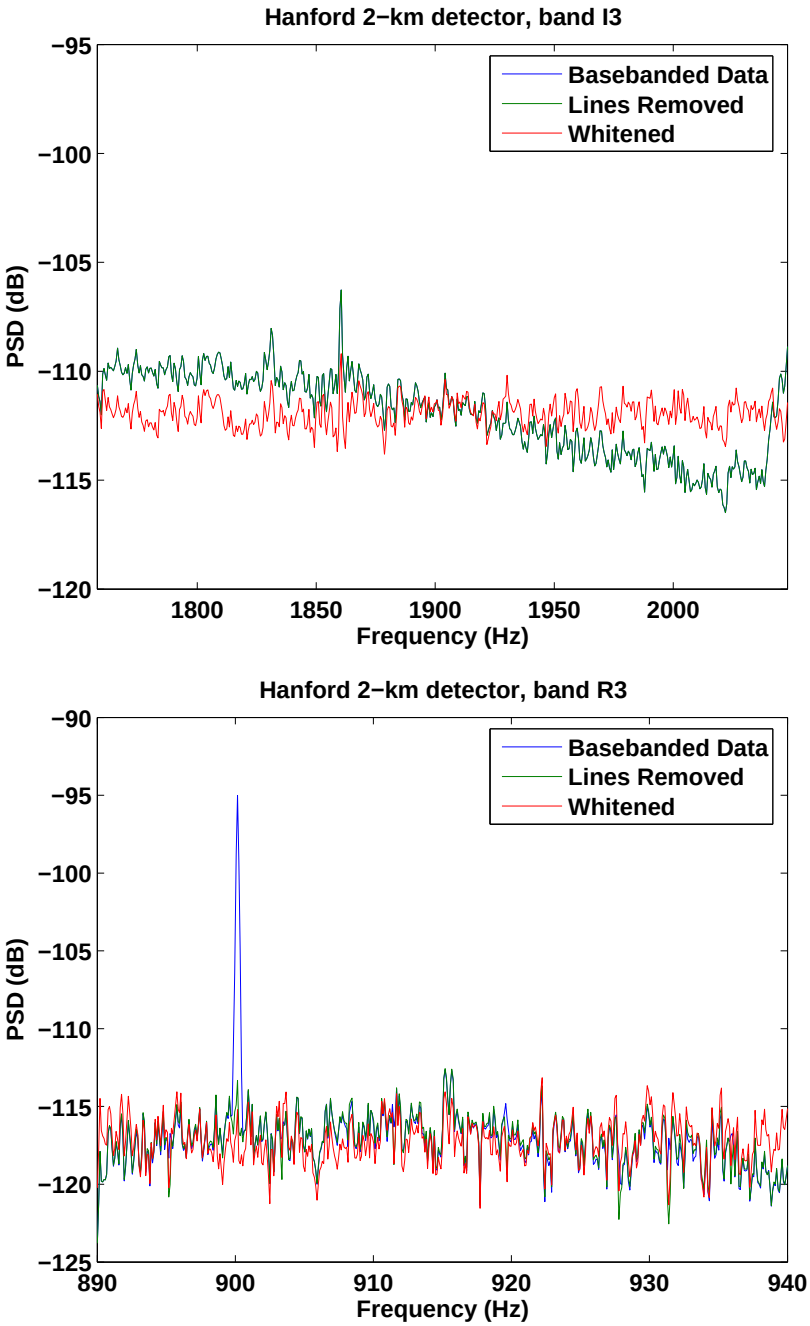

Hanford 2−km detector, band C3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−50
−60
−70
−80
−90
−100
−110
400
420
440
460
480
500
Frequency (Hz)
Hanford 2−km detector, band D3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−60
−70
−80
−90
−100
−110
−120
520
540
560
580
600
620
640
Frequency (Hz)

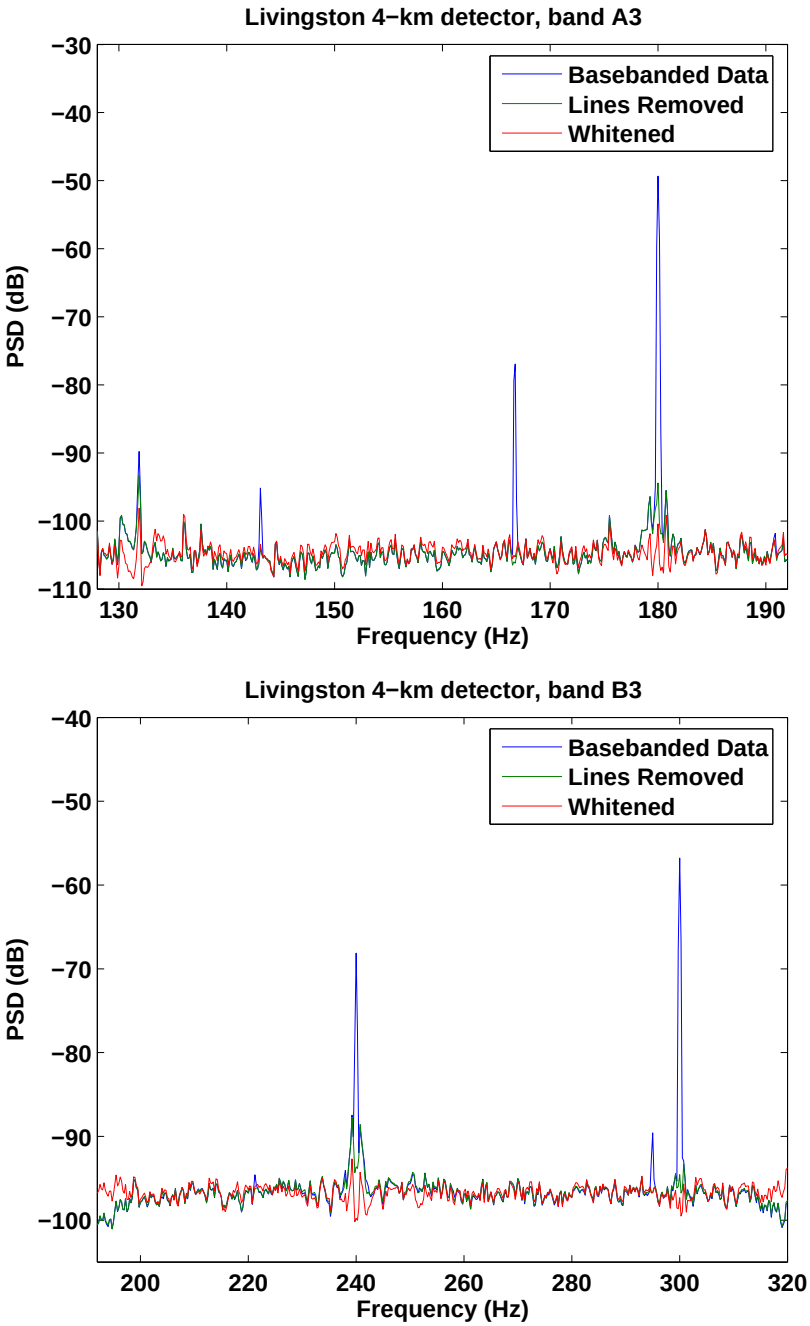

Hanford 2−km detector, band E3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−80
−85
−90
−95
−100
−105
−110
−115
−120
720
740
760
780
800
820
Frequency (Hz)

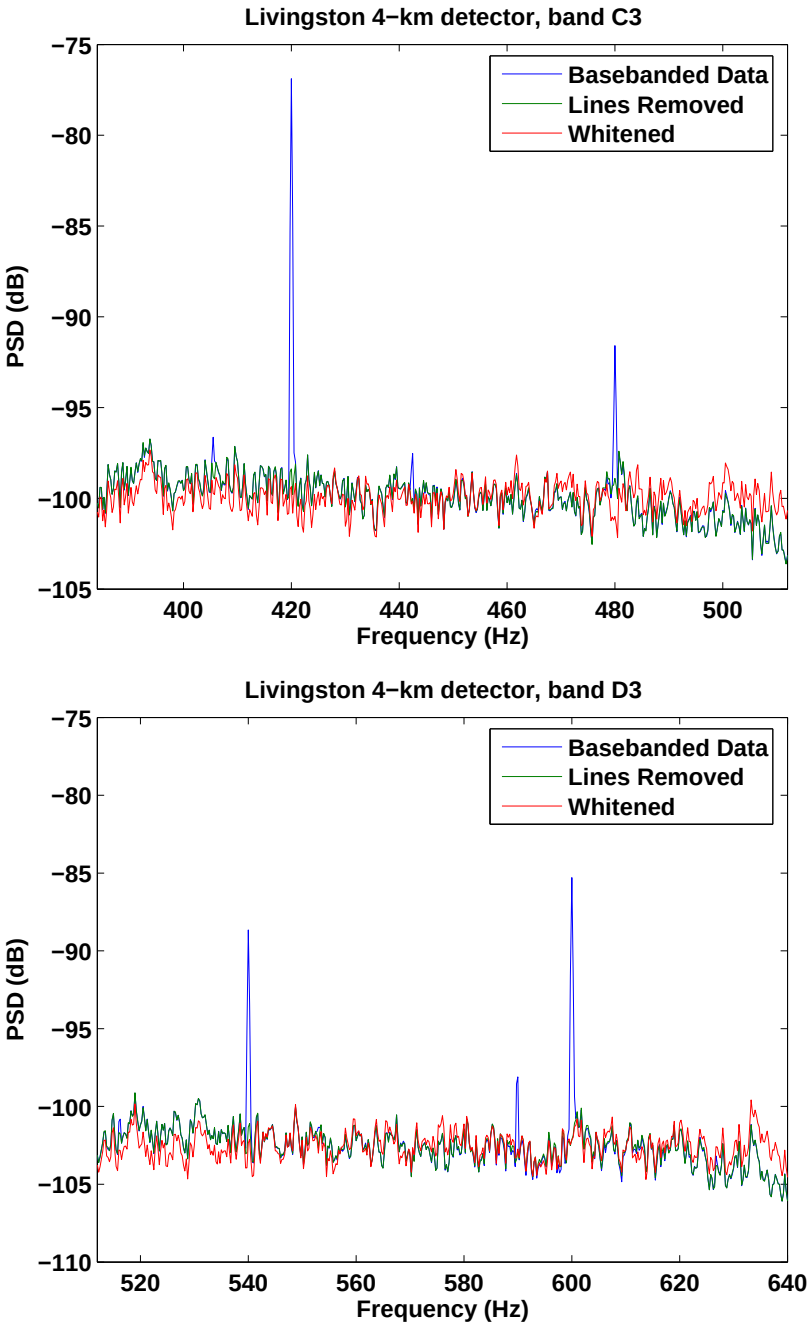

Hanford 2−km detector, band F3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−40
−50
−60
−70
−80
−90
−100
−110
−120
850
900
950
1000
Frequency (Hz)

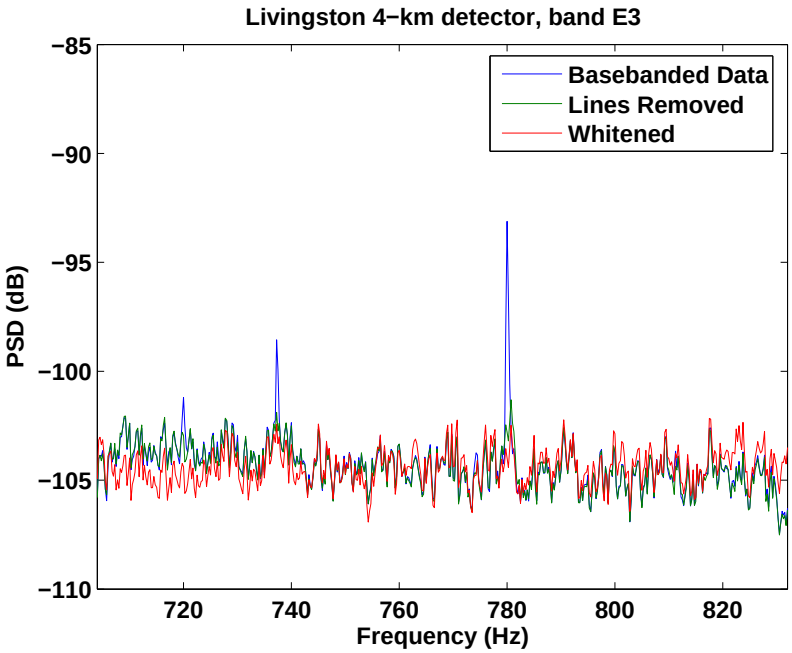

Hanford 2−km detector, band G3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−95
−100
−105
−110
−115
−120
1100
1150
1200
1250
1300
1350
Frequency (Hz)

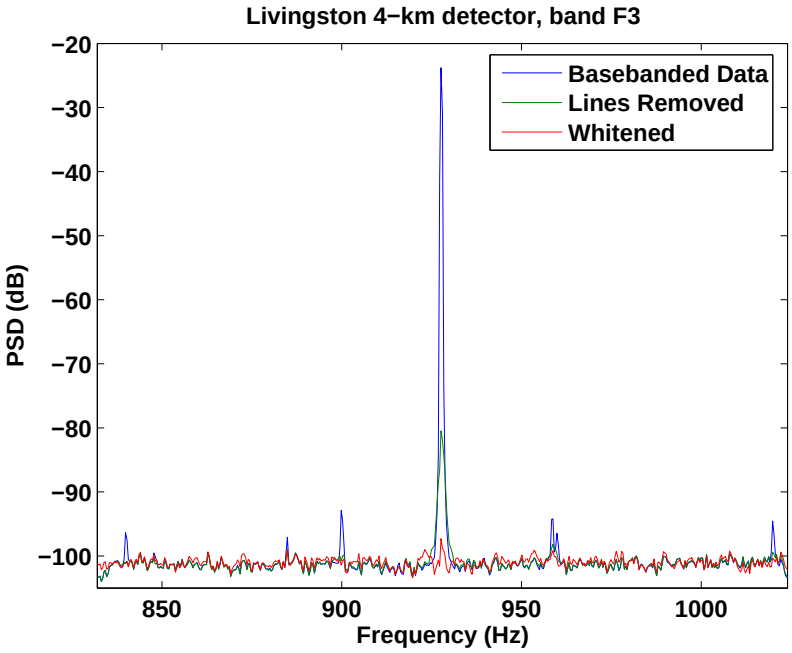

Hanford 2−km detector, band H3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−70
−75
−80
−85
−90
−95
−100
−105
−110
−115
−120
1450
1500
1550
1600
1650
1700
Frequency (Hz)

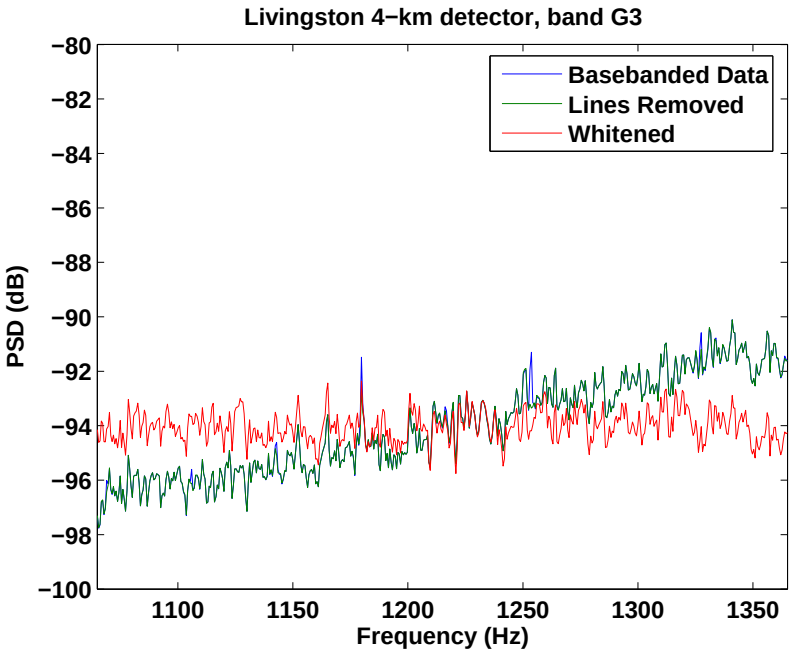

Hanford 2−km detector, band I3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−95
−100
−105
−110
−115
−120
1800
1850
1900
1950
2000
Frequency (Hz)
Hanford 2−km detector, band R3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−90
−95
−100
−105
−110
−115
−120
−125
890
900
910
920
930
940
Frequency (Hz)

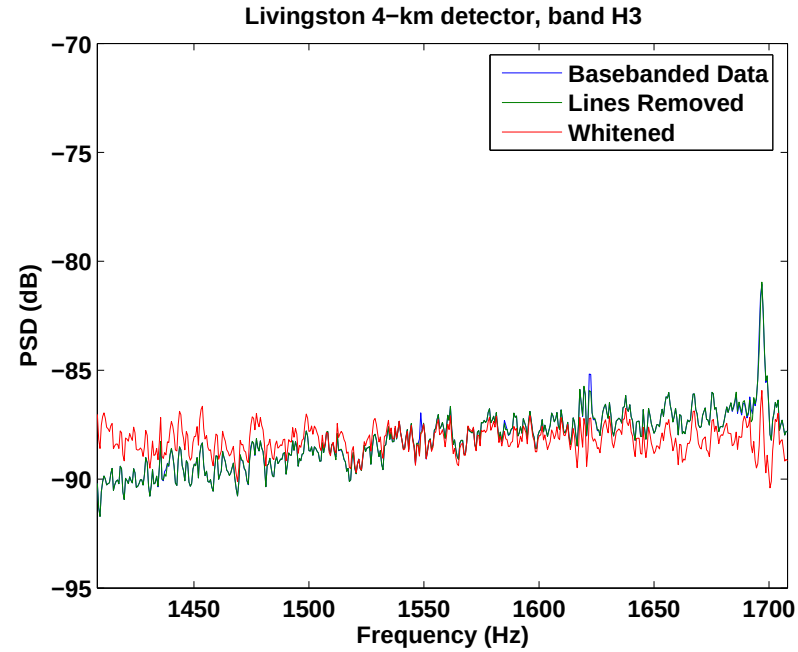

Livingston 4−km detector, band A3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−30
−40
−50
−60
−70
−80
−90
−100
−110
130
140
150
160
170
180
190
Frequency (Hz)
Livingston 4−km detector, band B3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−40
−50
−60
−70
−80
−90
−100
200
220
240
260
280
300
320
Frequency (Hz)

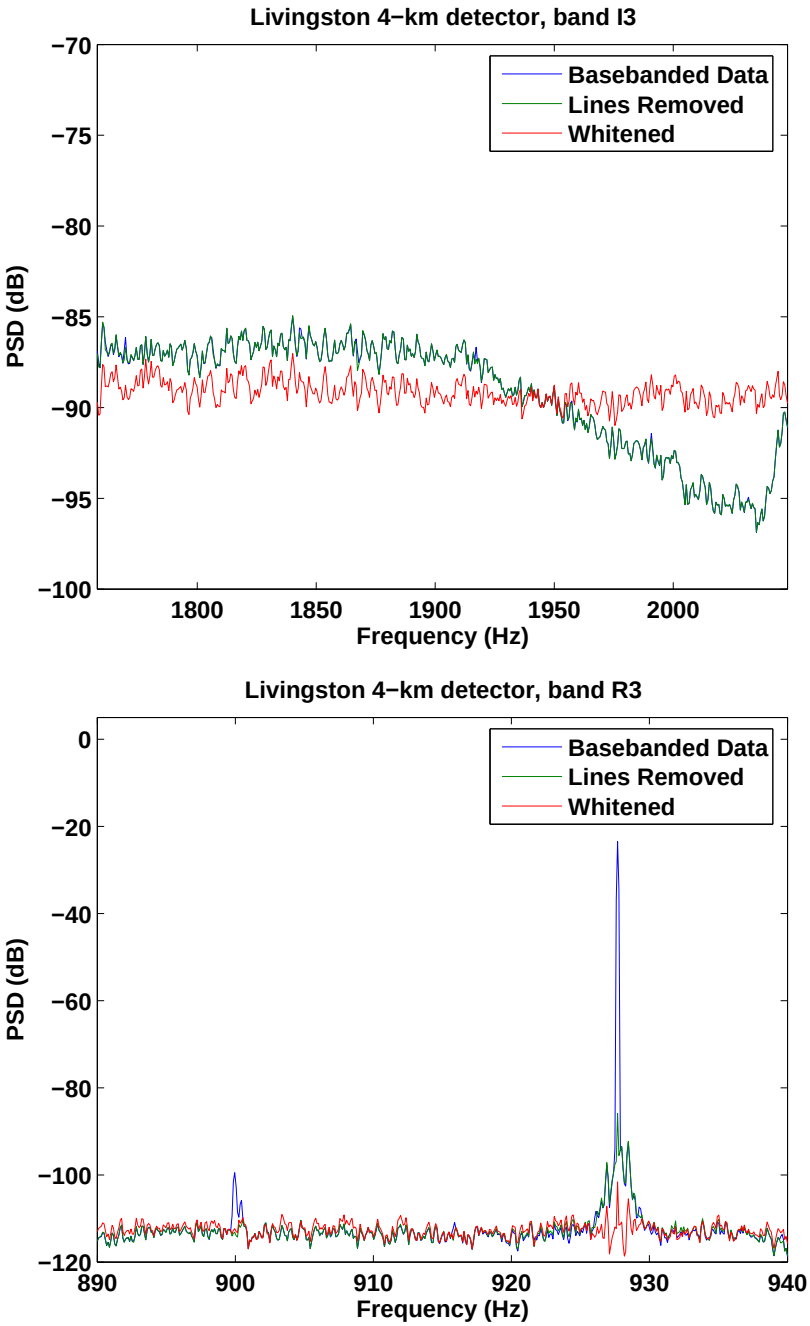

Livingston 4−km detector, band C3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−75
−80
−85
−90
−95
−100
−105
400
420
440
460
480
500
Frequency (Hz)
Livingston 4−km detector, band D3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−75
−80
−85
−90
−95
−100
−105
−110
520
540
560
580
600
620
640
Frequency (Hz)


Livingston 4−km detector, band E3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−85
−90
−95
−100
−105
−110
720
740
760
780
800
820
Frequency (Hz)


Livingston 4−km detector, band F3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−20
−30
−40
−50
−60
−70
−80
−90
−100
850
900
950
1000
Frequency (Hz)


Livingston 4−km detector, band G3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−80
−82
−84
−86
−88
−90
−92
−94
−96
−98
−100
1100
1150
1200
1250
1300
1350
Frequency (Hz)


Livingston 4−km detector, band H3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−70
−75
−80
−85
−90
−95
1450
1500
1550
1600
1650
1700
Frequency (Hz)


Livingston 4−km detector, band I3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
−70
−75
−80
−85
−90
−95
−100
1800
1850
1900
1950
2000
Frequency (Hz)
Livingston 4−km detector, band R3
Basebanded Data
Lines Removed
Whitened
PSD (dB)
0
−20
−40
−60
−80
−100
−120
890
900
910
920
930
940
Frequency (Hz)

Appendix E

Rayleigh FOM Distributions

The Rayleigh figure-of-merit (FOM) is a measure of data conditioning performance. It is calculated via

$$\mathrm{FOM} = \frac{\mathcal{R}_{PSD'}}{\mathcal{R}_{PSD}} = \frac{\sigma_{PSD'}/\mu_{PSD'}}{\sigma_{PSD}/\mu_{PSD}}, \tag{E.1}$$

where PSD is the power spectral density calculated on data that has only been treated with the shaping filter and then basebanded and PSD' is data that has undergone the entire data conditioning process. A FOM < 1 indicates that the line removal and whitening portions of data conditioning have been successful.

Below are plots showing the FOM per segment and distributions of FOM values for each frequency band (frequency band definitions are given in Appendix B) and detector over all non-overlapping 300 second long S3 segments. For most of these segments, the FOM was less than one, indicating that data conditioning was successful.

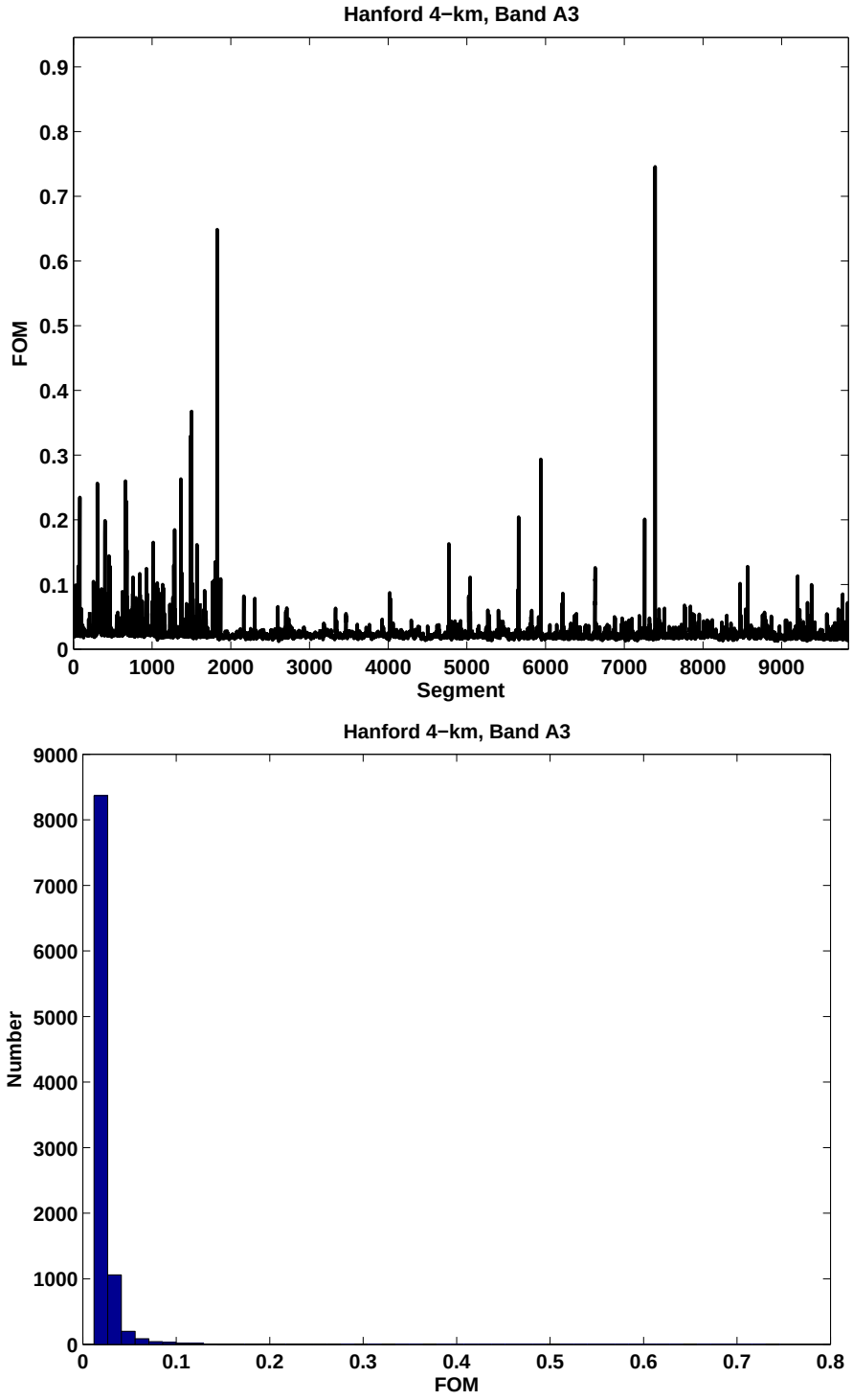


Figure E.1. Data conditioning FOM for Hanford 4-km detector, band A3

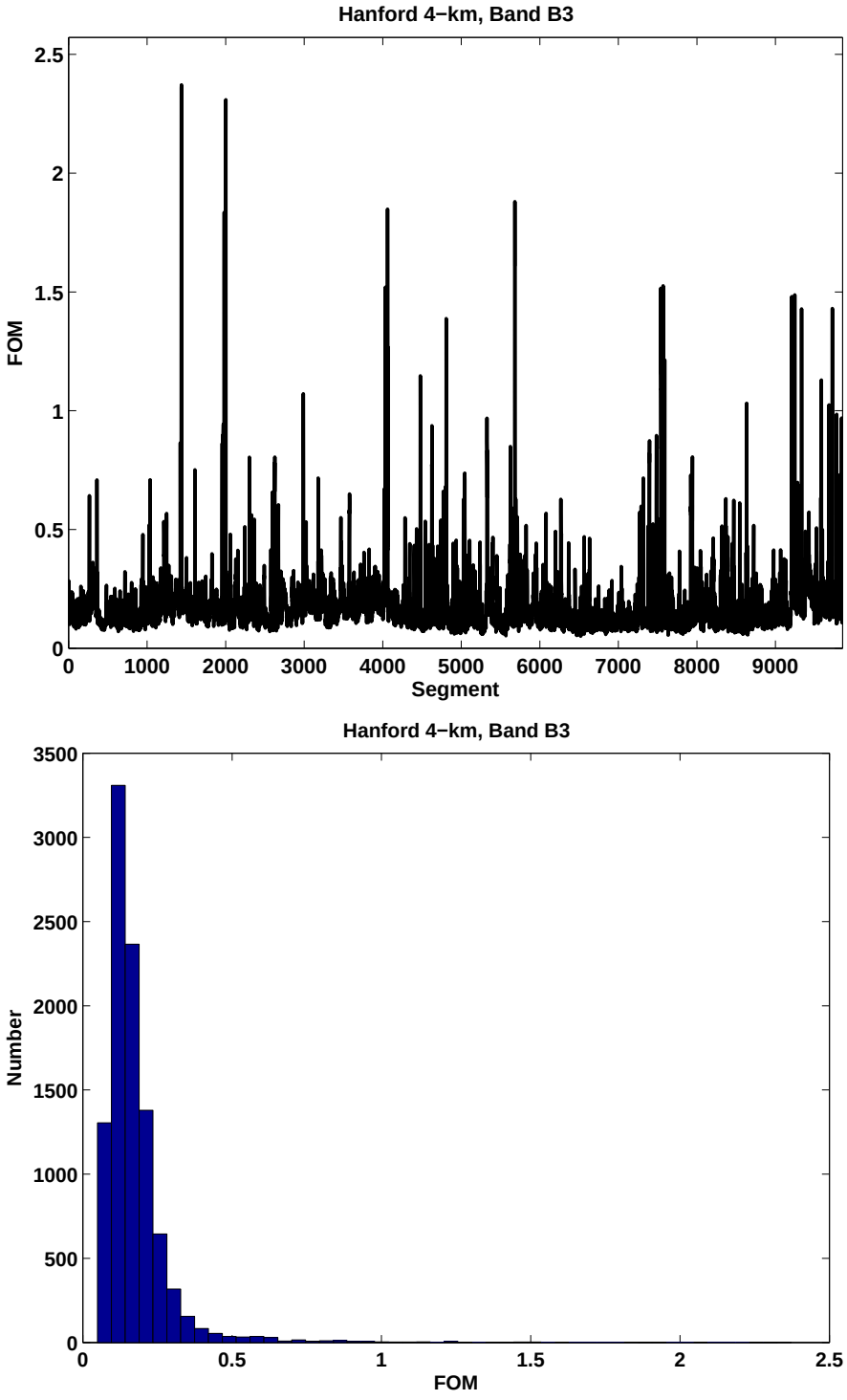


Figure E.2. Data conditioning FOM for Hanford 4-km detector, band B3

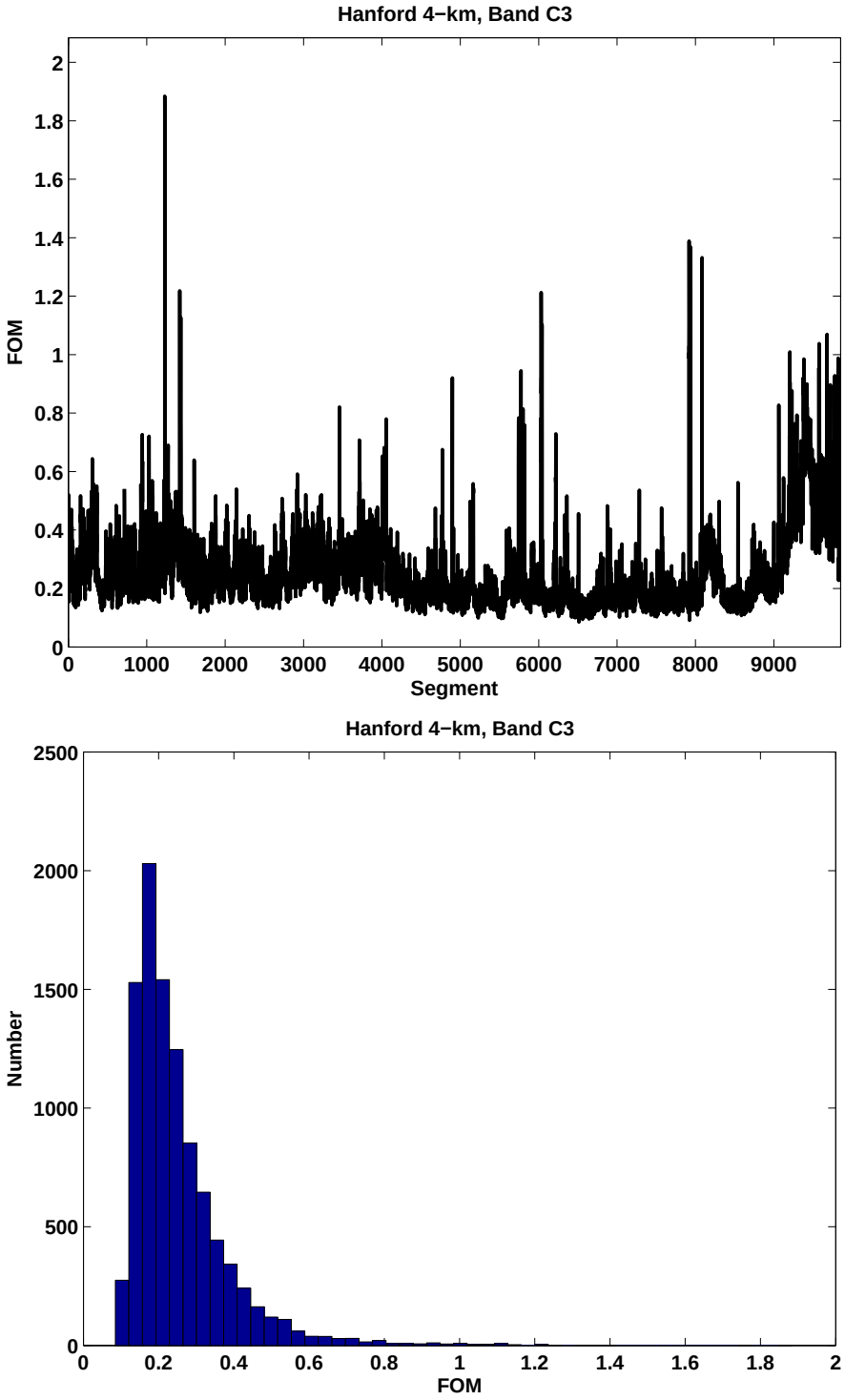


Figure E.3. Data conditioning FOM for Hanford 4-km detector, band C3

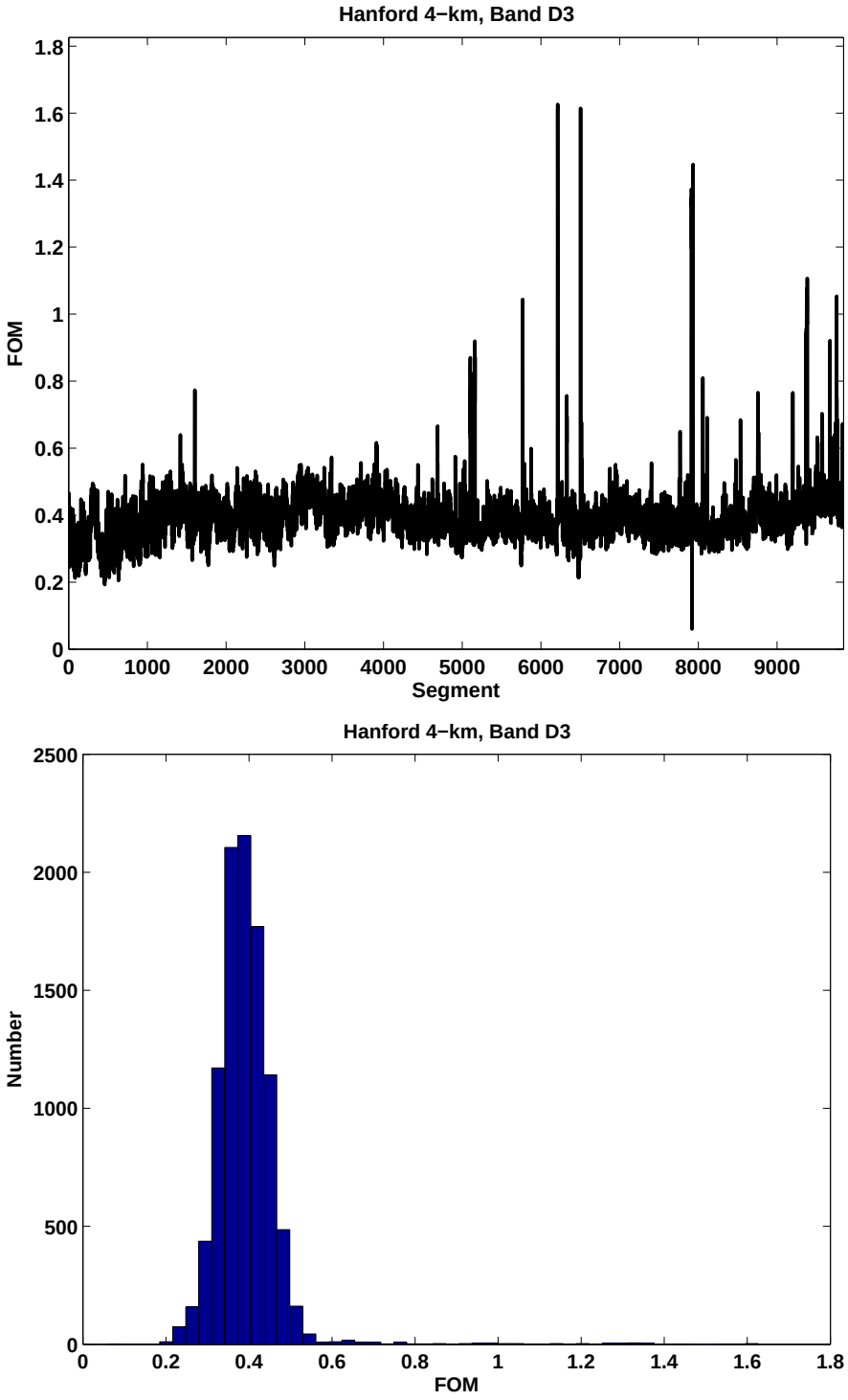


Figure E.4. Data conditioning FOM for Hanford 4-km detector, band D3

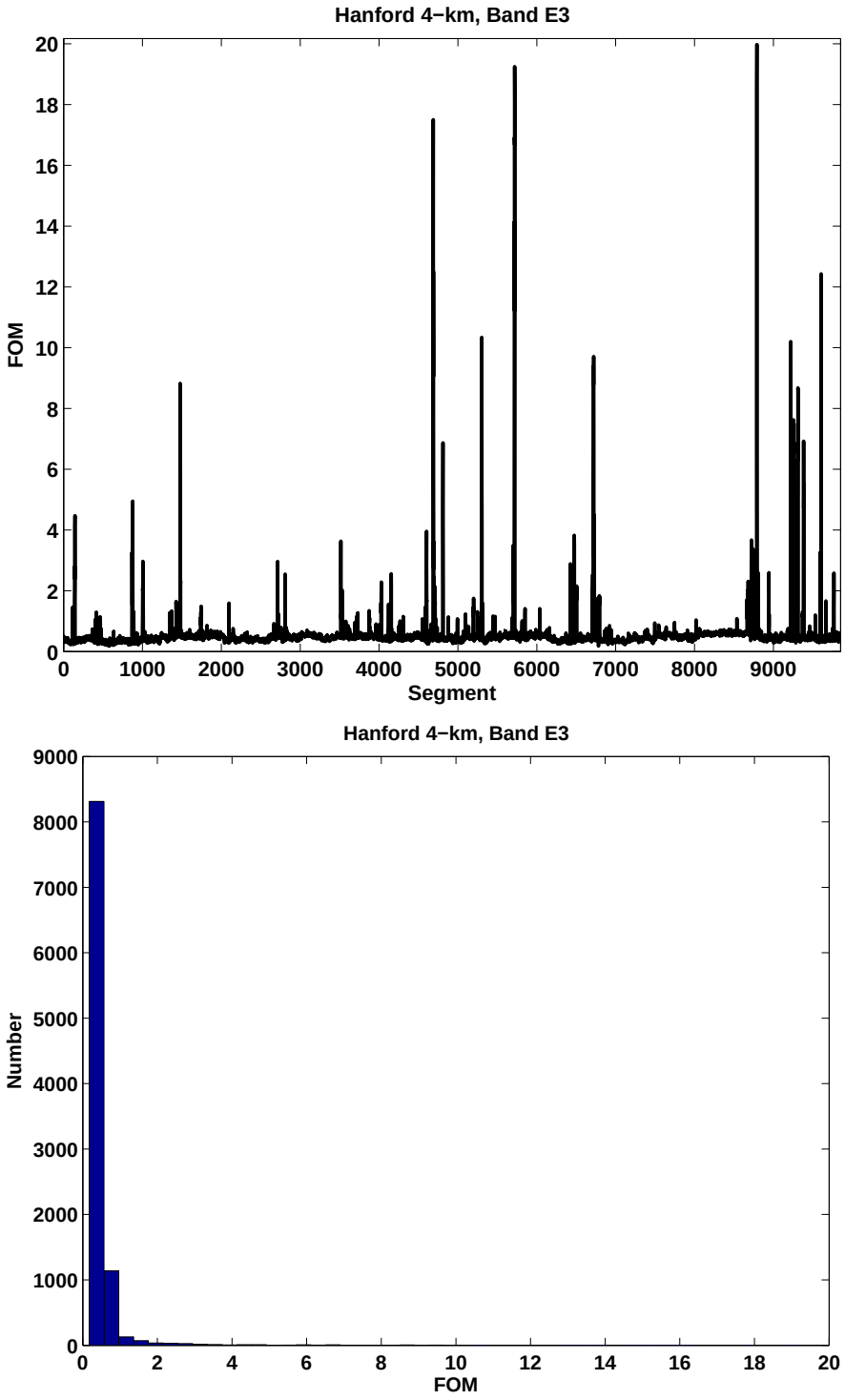


Figure E.5. Data conditioning FOM for Hanford 4-km detector, band E3

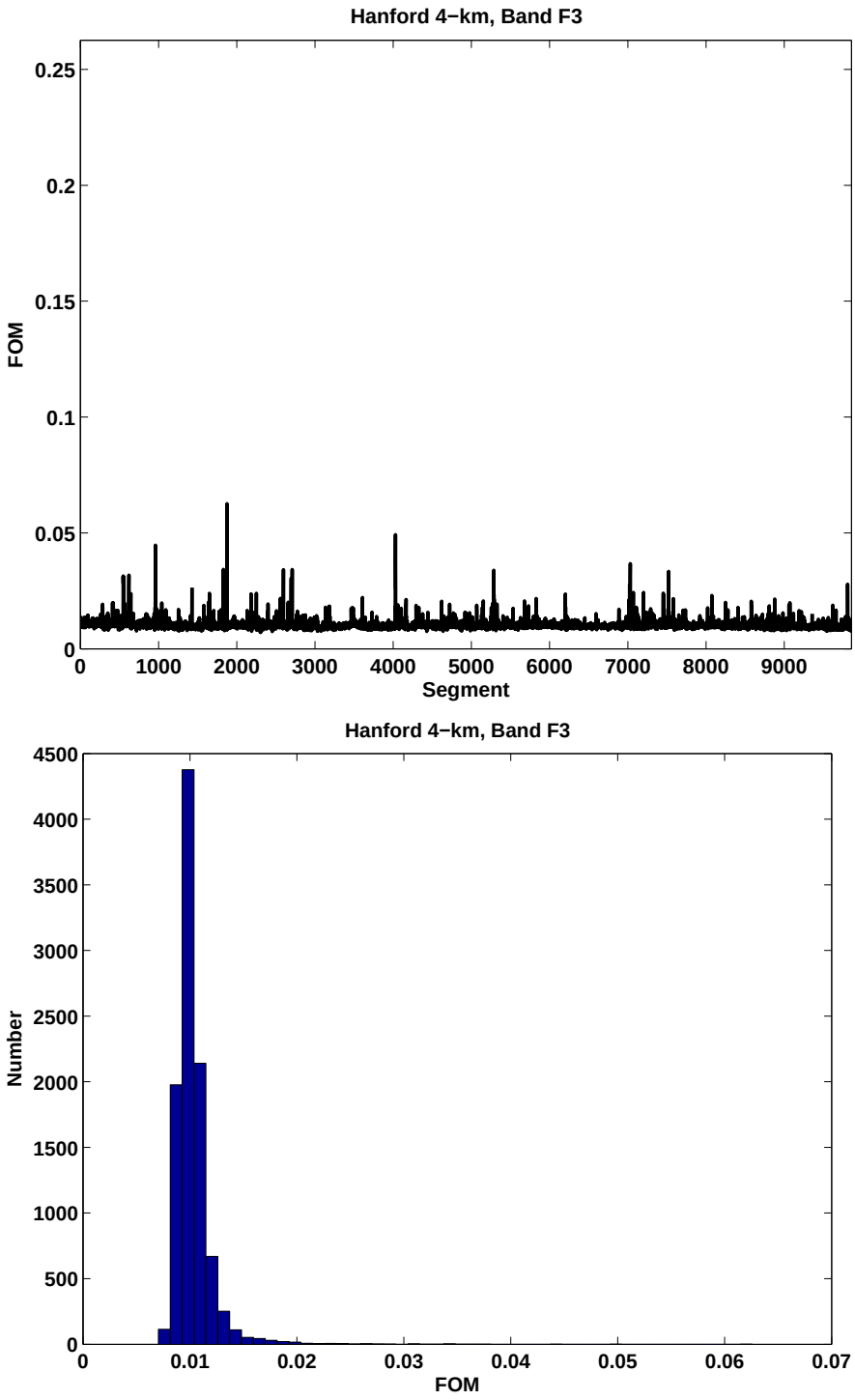


Figure E.6. Data conditioning FOM for Hanford 4-km detector, band F3

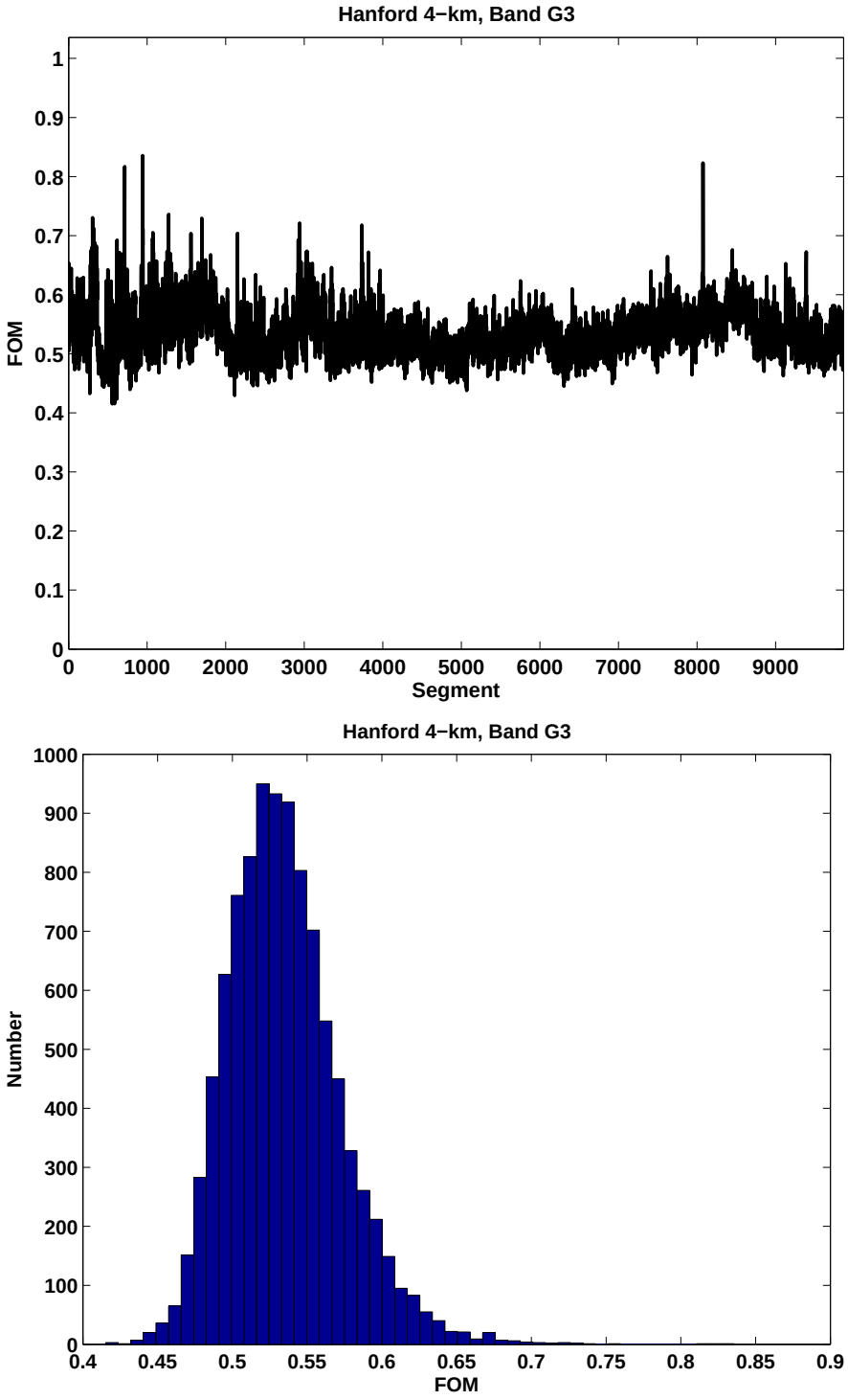


Figure E.7. Data conditioning FOM for Hanford 4-km detector, band G3

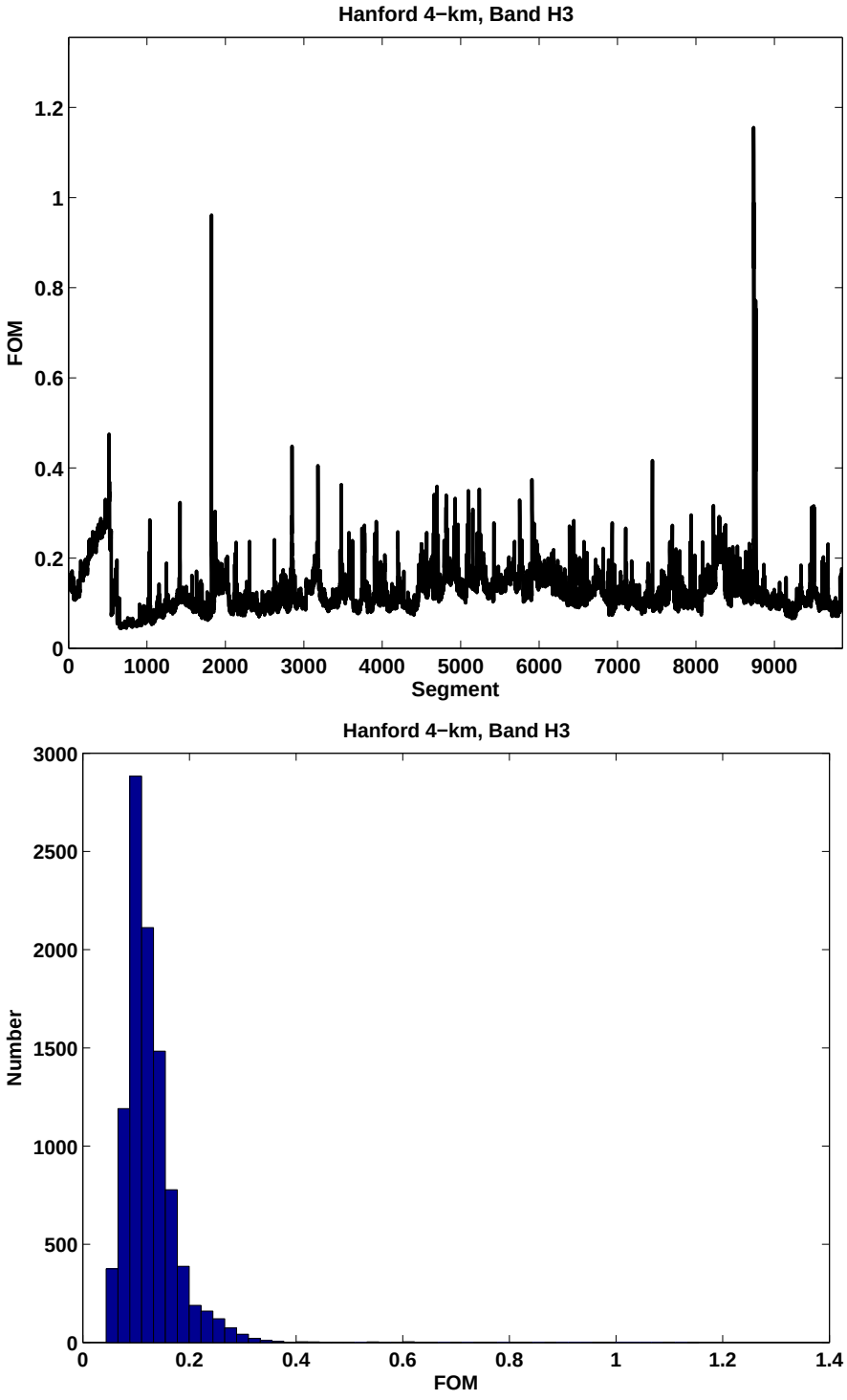


Figure E.8. Data conditioning FOM for Hanford 4-km detector, band H3

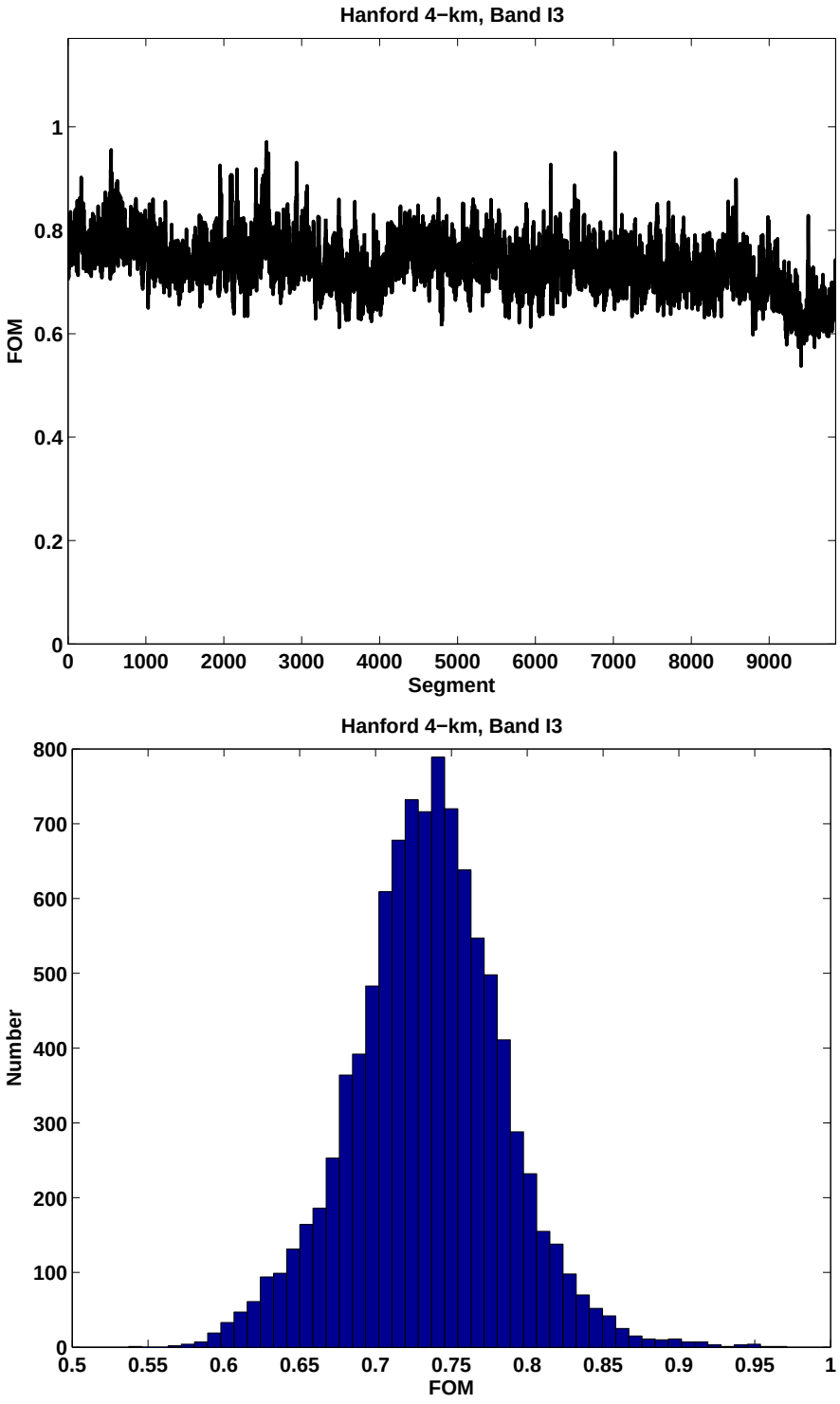


Figure E.9. Data conditioning FOM for Hanford 4-km detector, band I3

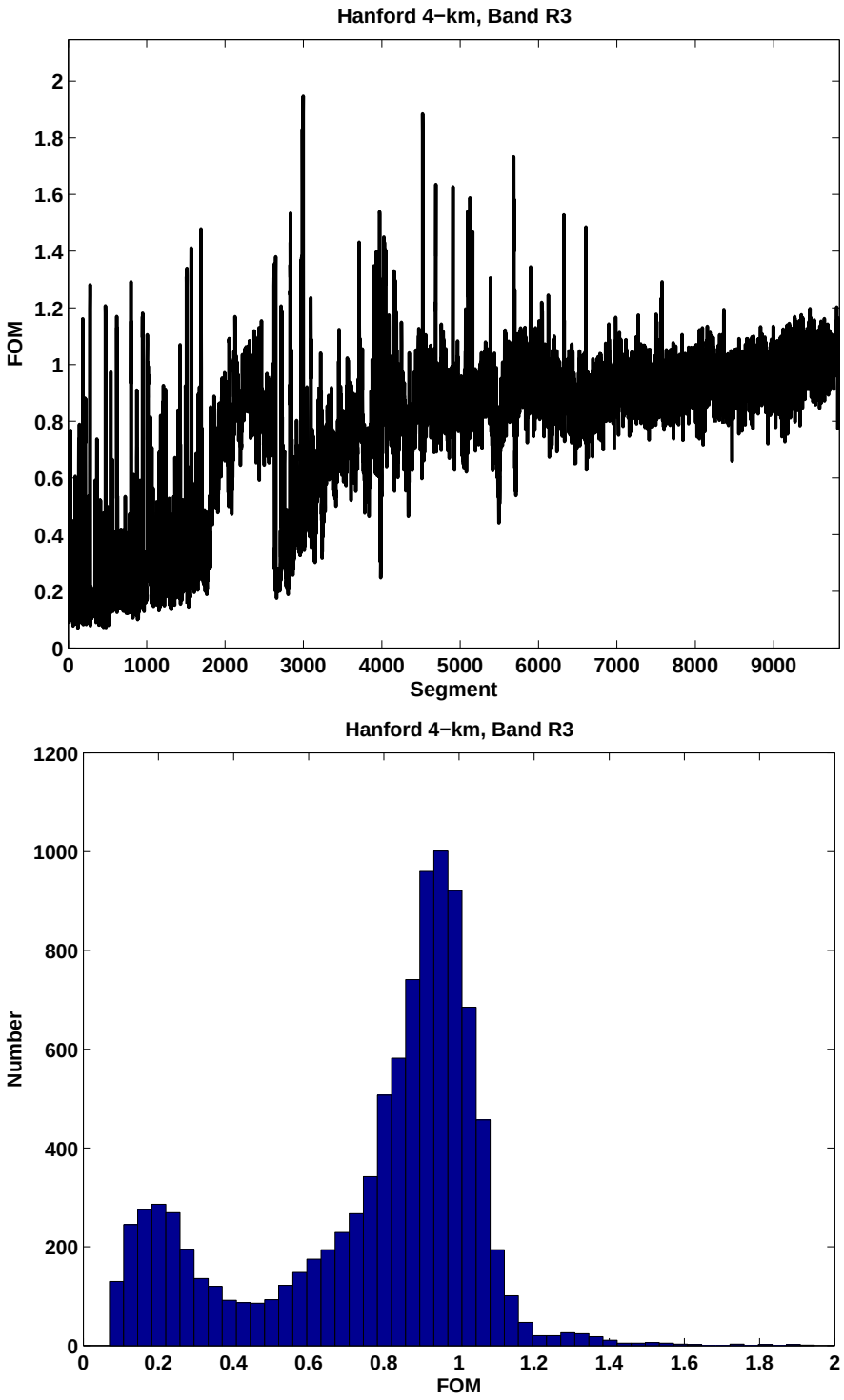


Figure E.10. Data conditioning FOM for Hanford 4-km detector, band R3

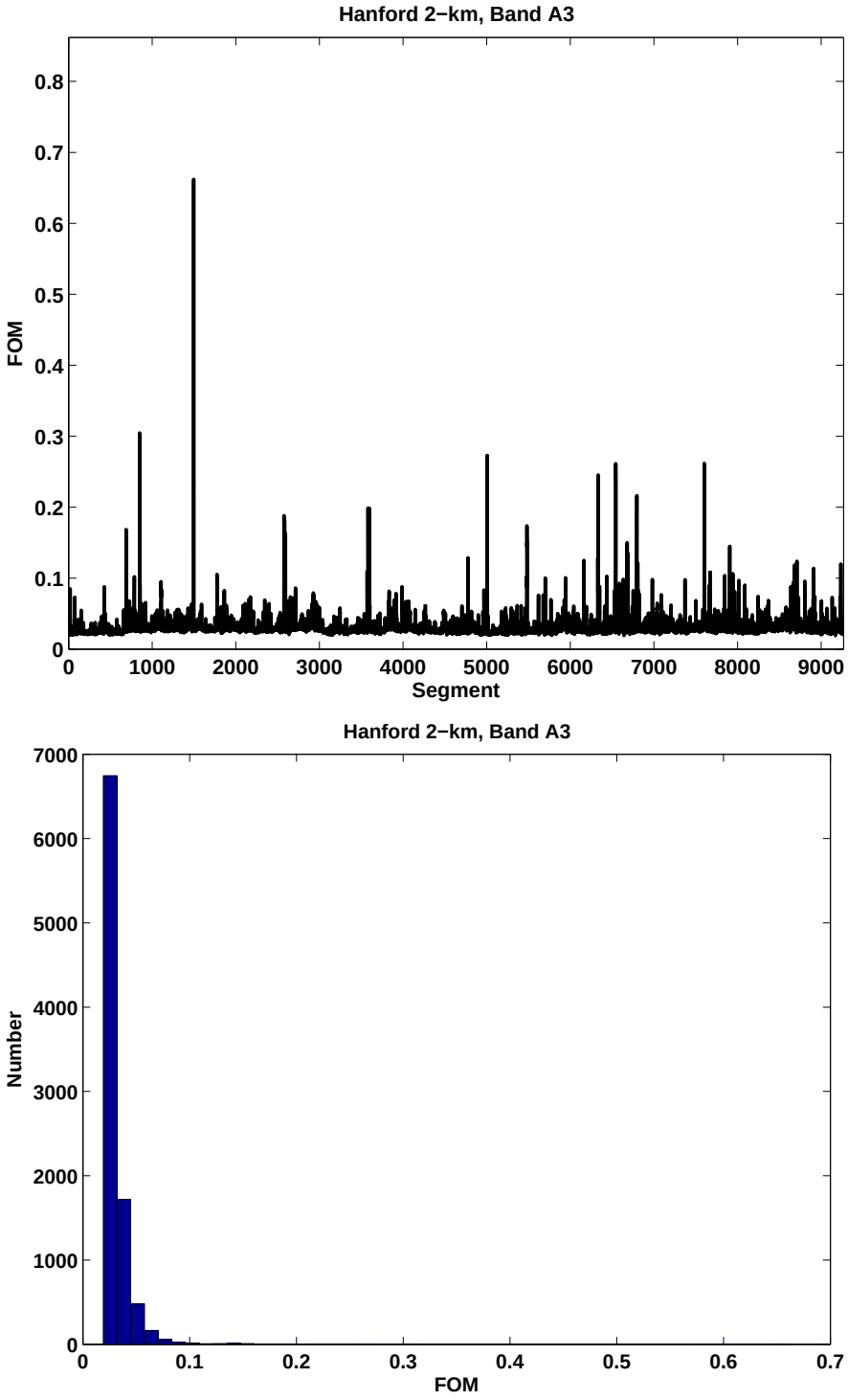


Figure E.11. Data conditioning FOM for Hanford 2-km detector, band A3

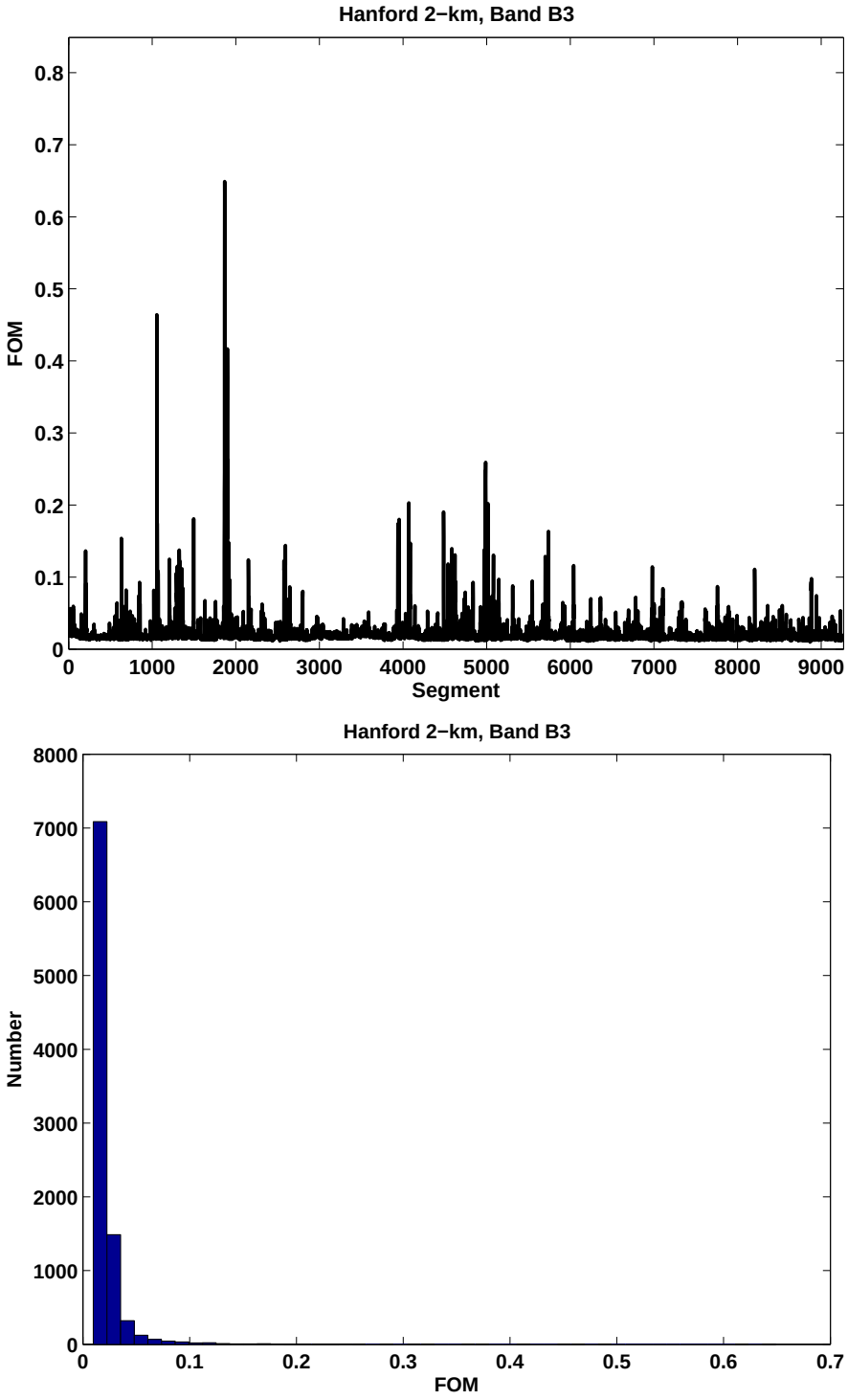


Figure E.12. Data conditioning FOM for Hanford 2-km detector, band B3

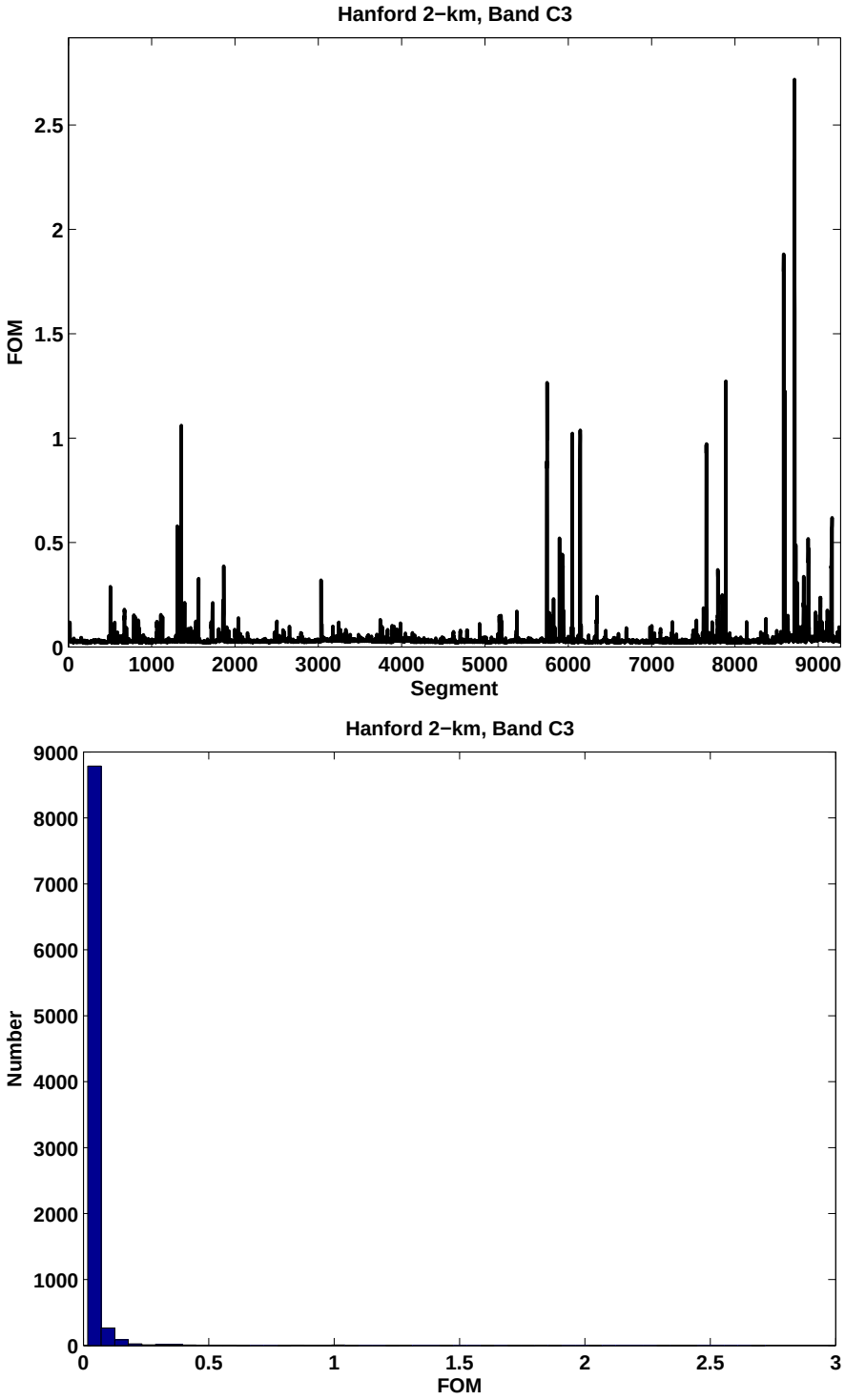


Figure E.13. Data conditioning FOM for Hanford 2-km detector, band C3

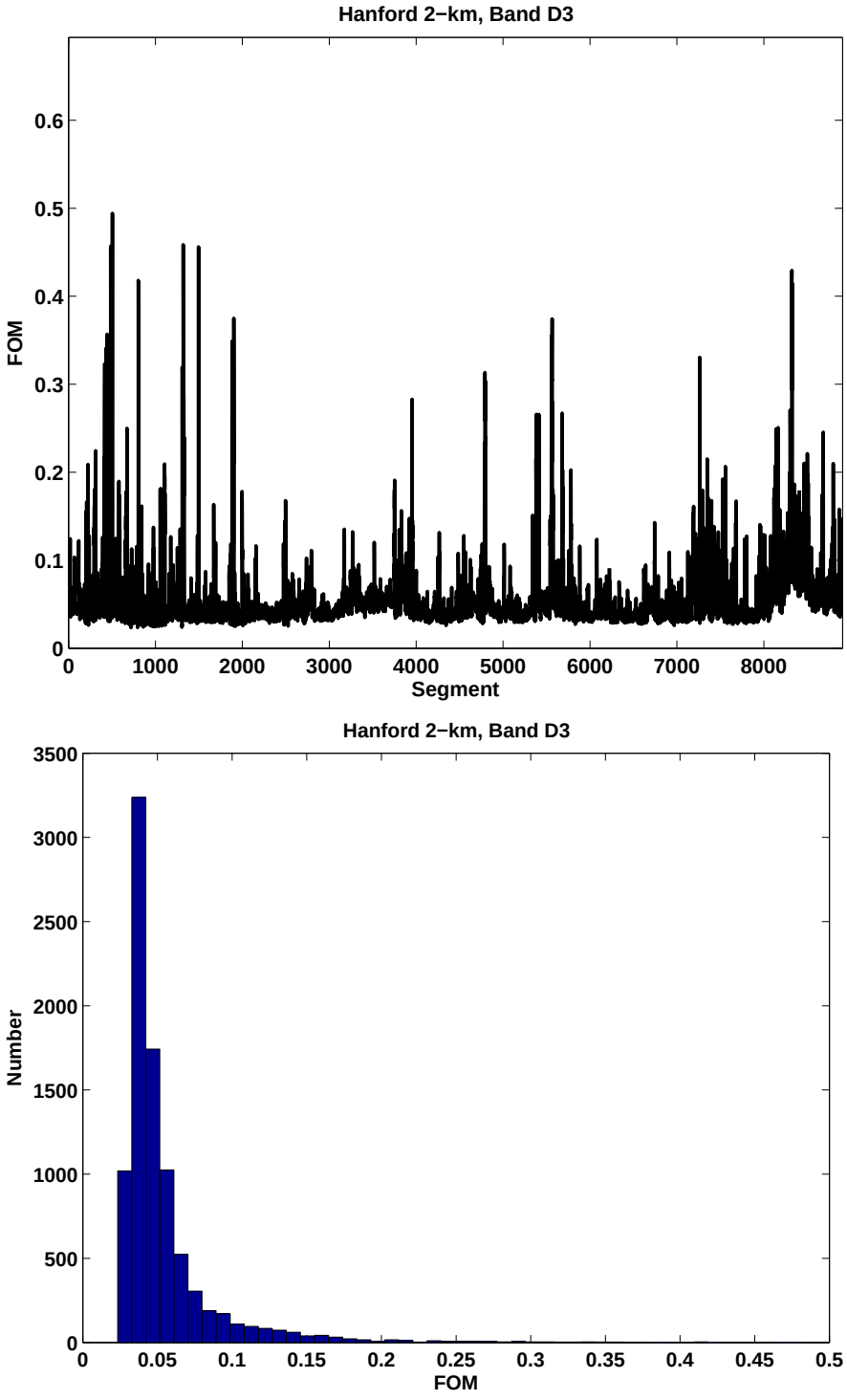


Figure E.14. Data conditioning FOM for Hanford 2-km detector, band D3

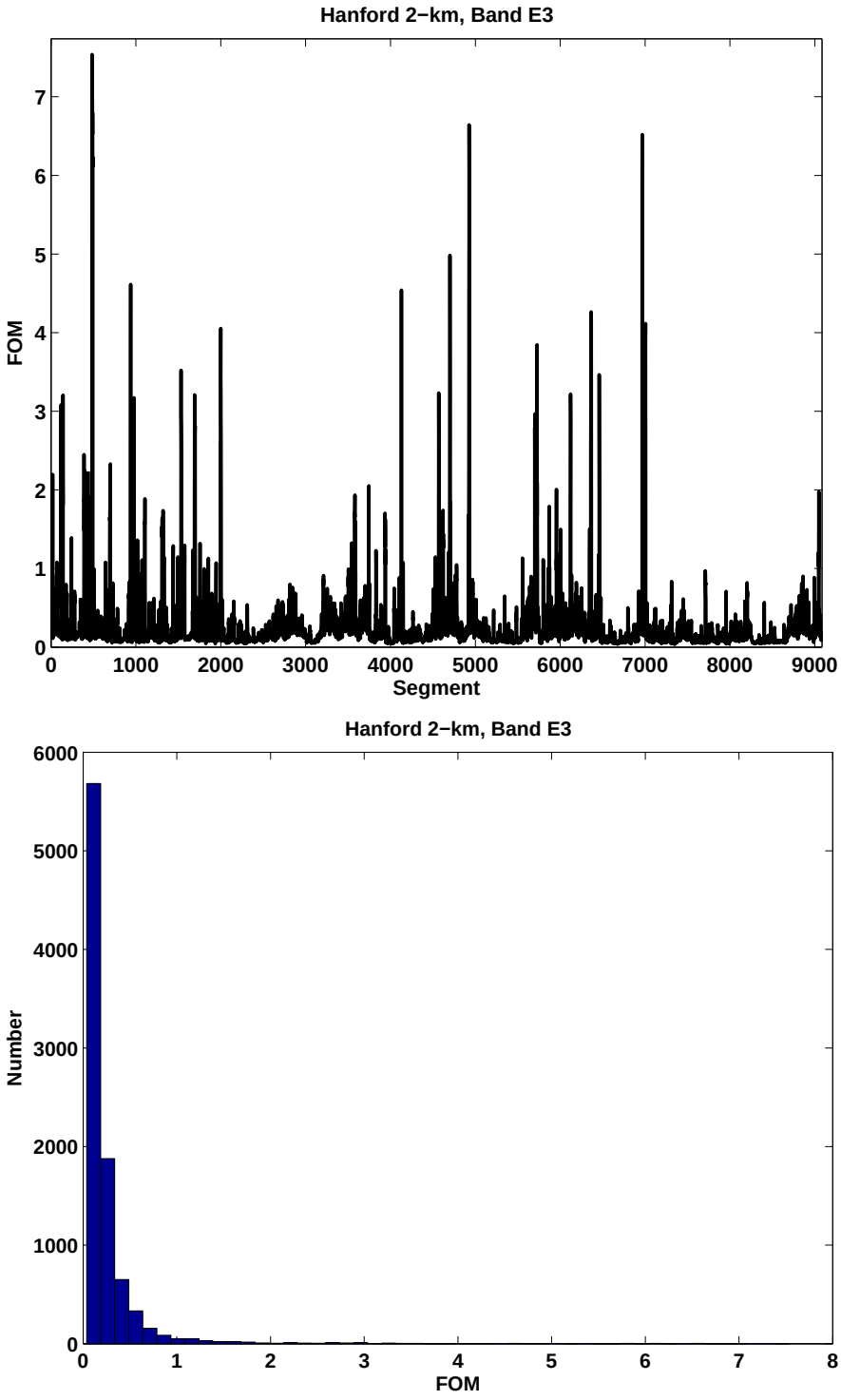


Figure E.15. Data conditioning FOM for Hanford 2-km detector, band E3

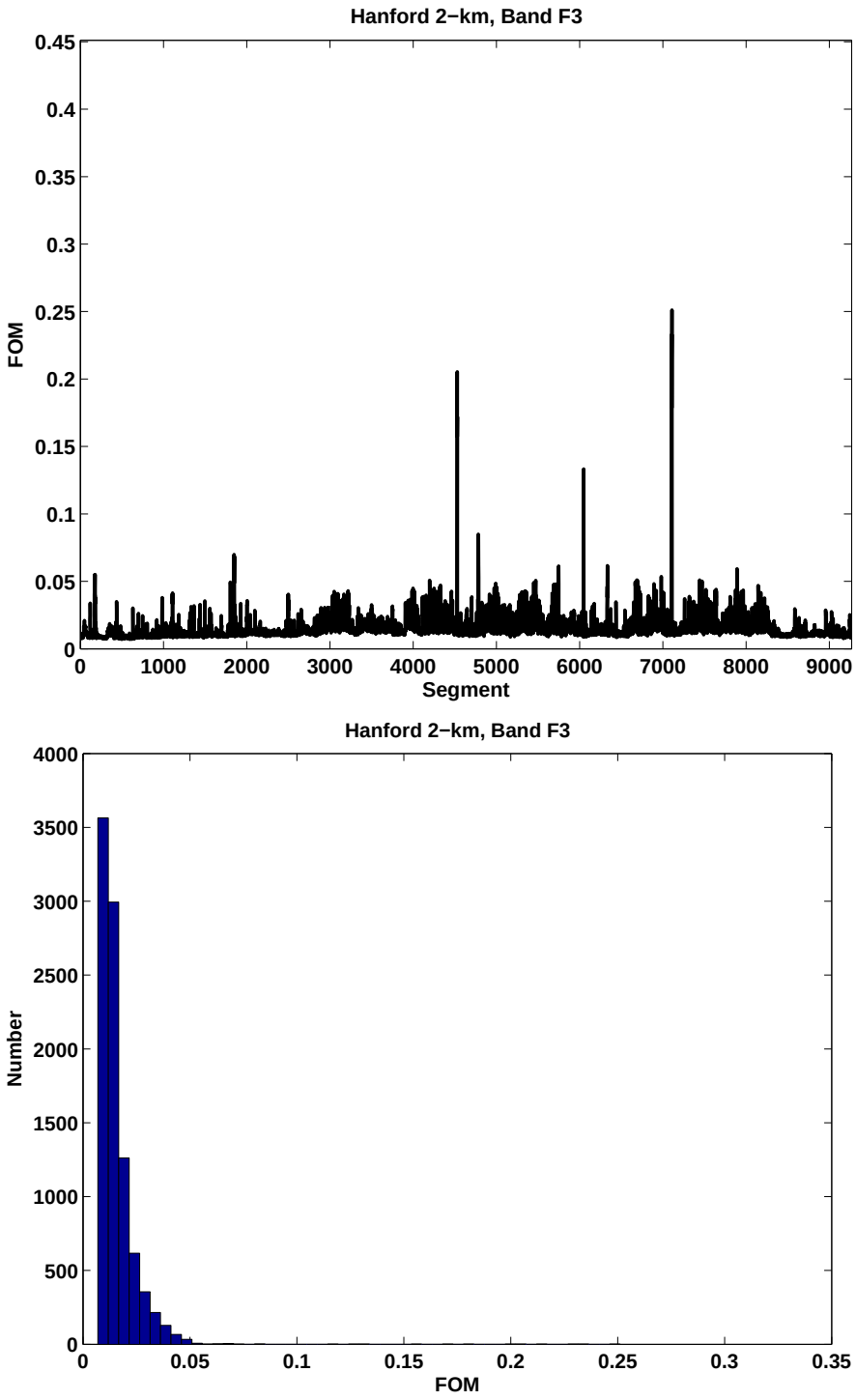


Figure E.16. Data conditioning FOM for Hanford 2-km detector, band F3

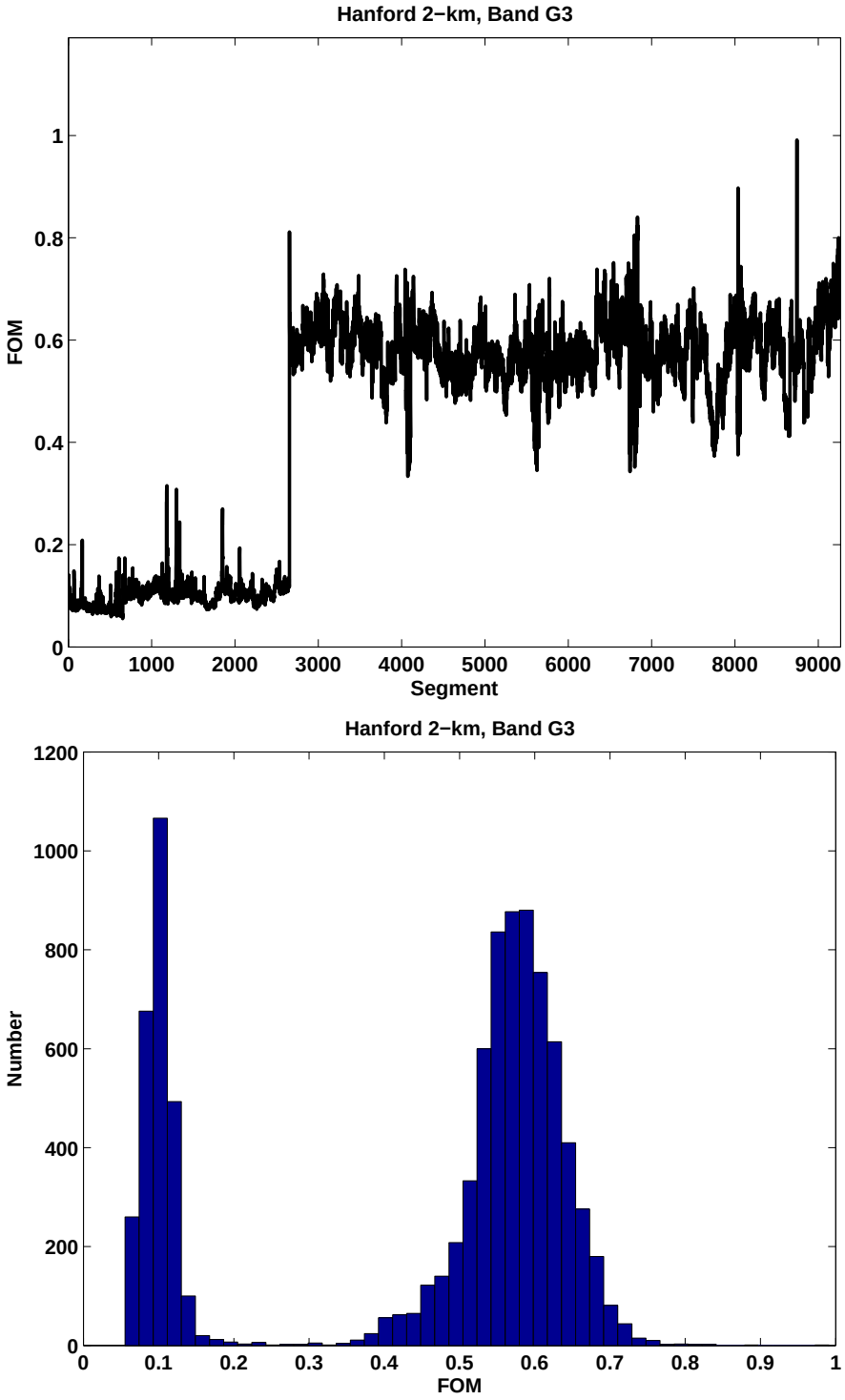


Figure E.17. Data conditioning FOM for Hanford 2-km detector, band G3

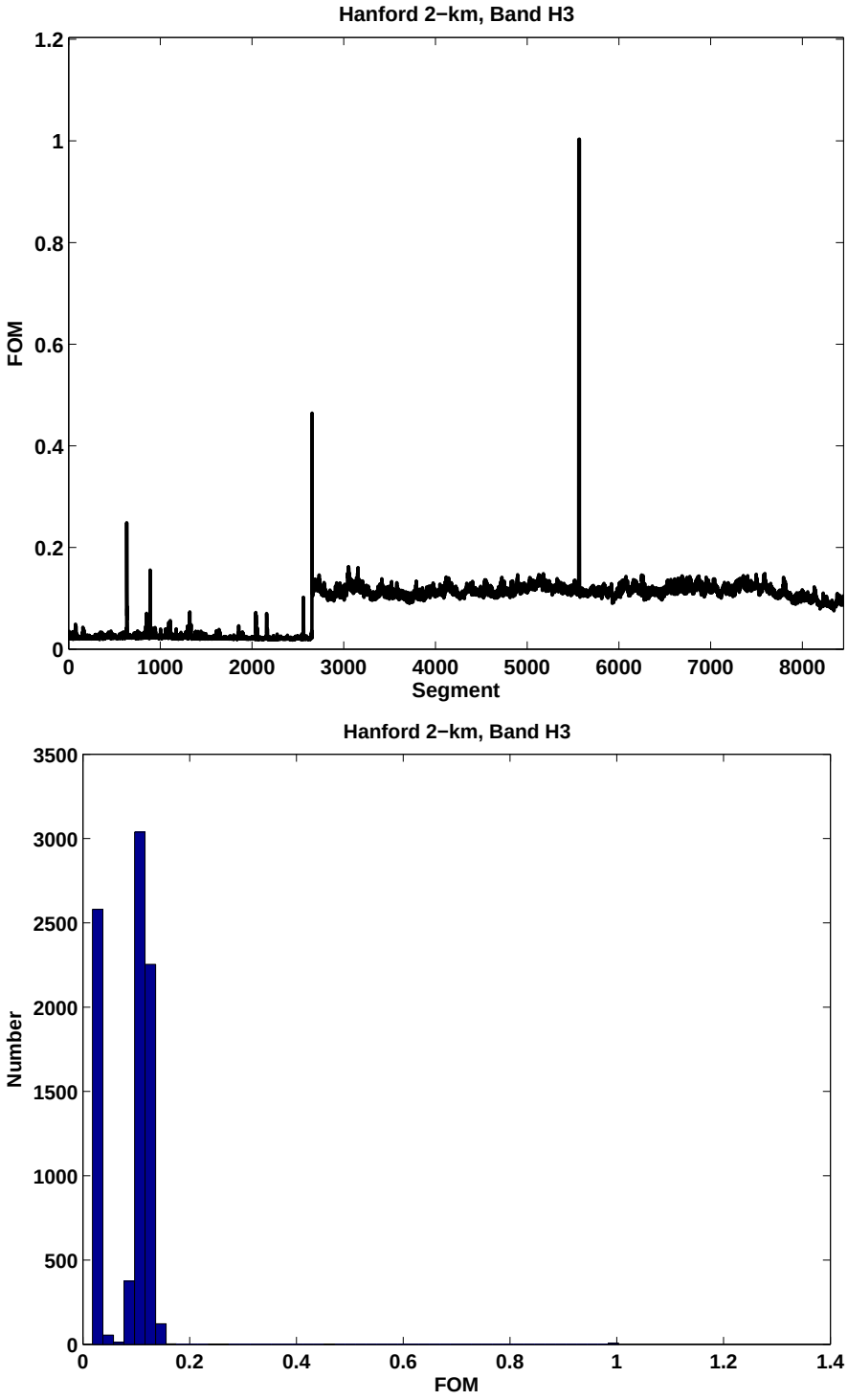


Figure E.18. Data conditioning FOM for Hanford 2-km detector, band H3

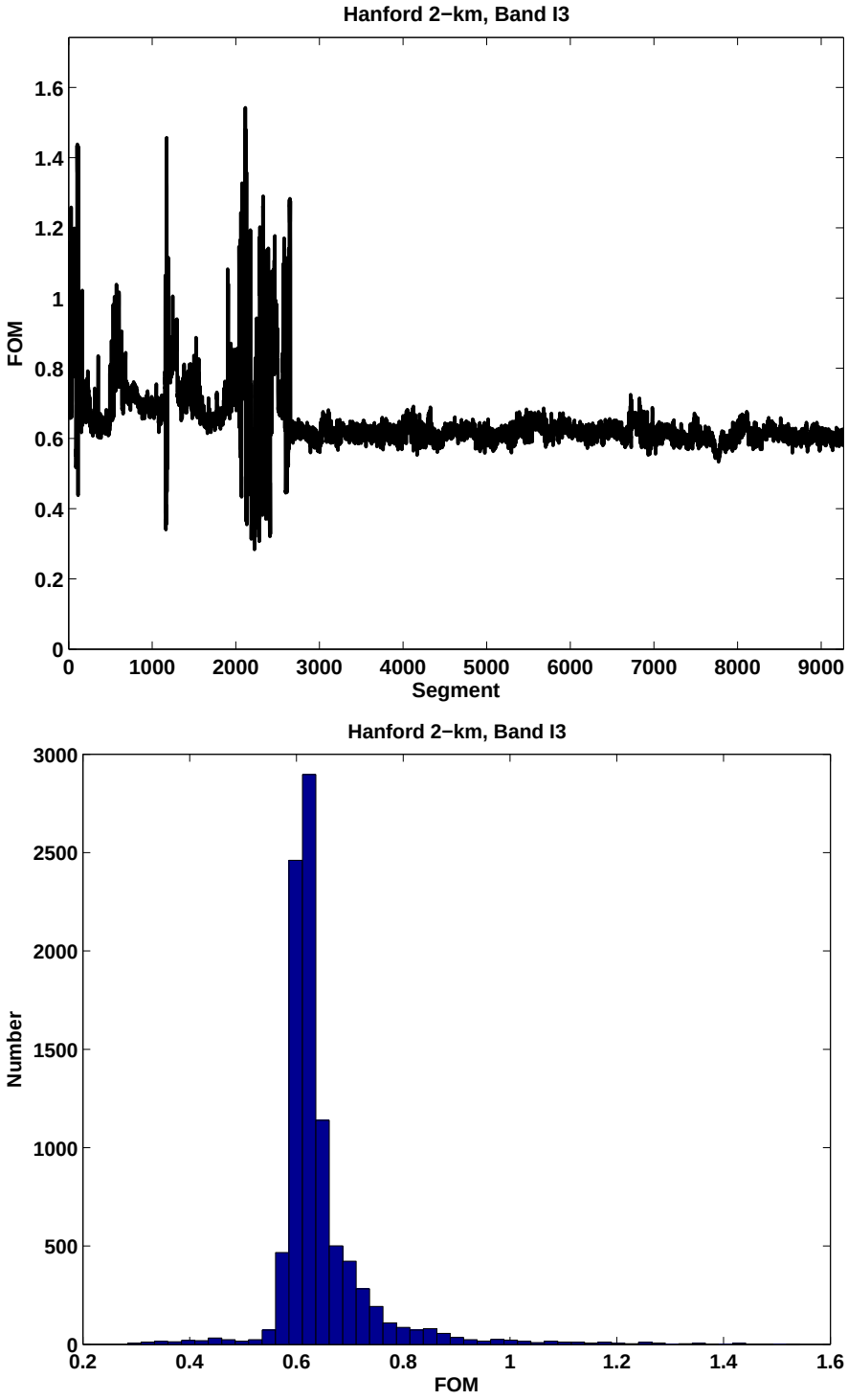


Figure E.19. Data conditioning FOM for Hanford 2-km detector, band I3

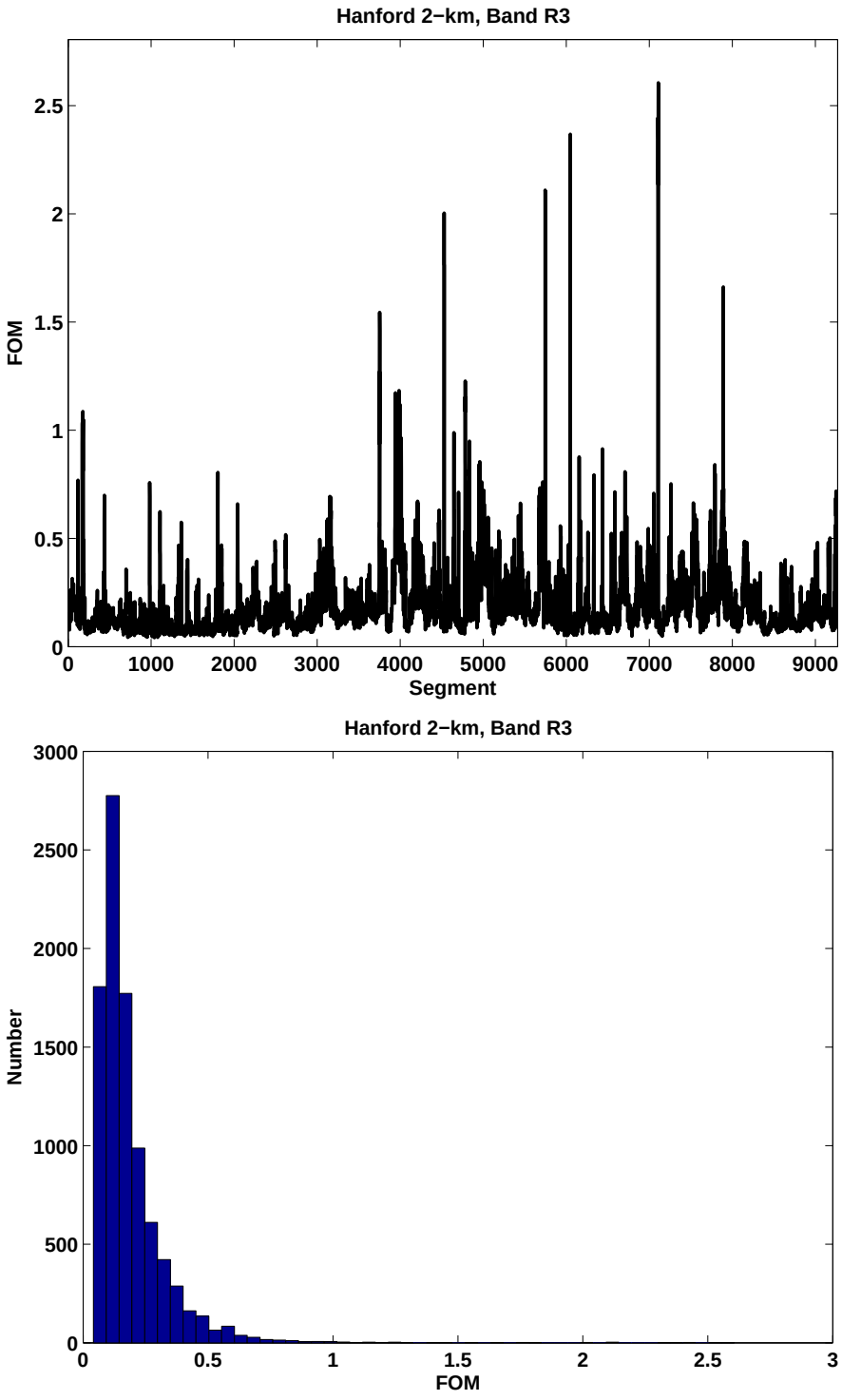


Figure E.20. Data conditioning FOM for Hanford 2-km detector, band R3

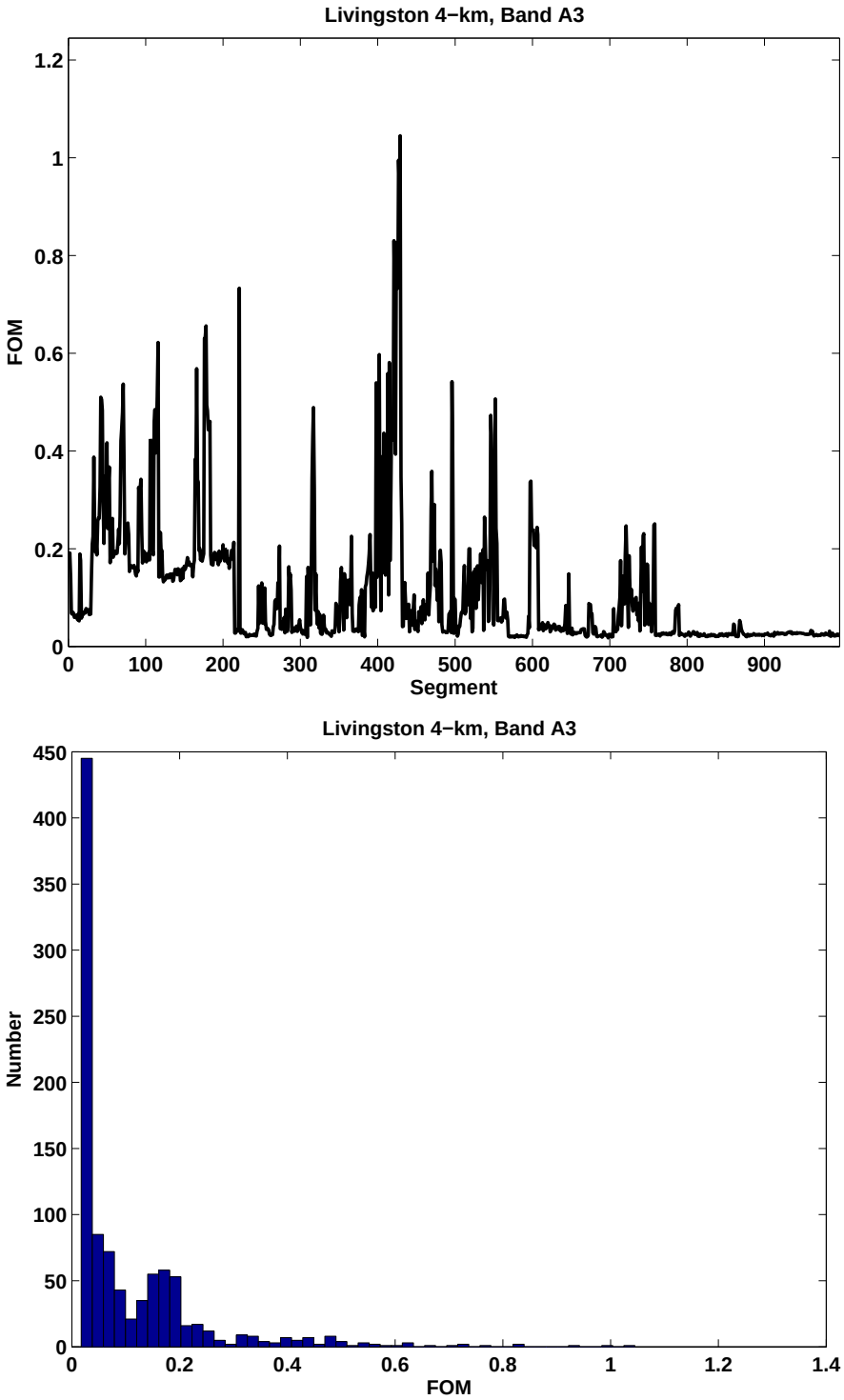


Figure E.21. Data conditioning FOM for Livingston 4-km detector, band A3

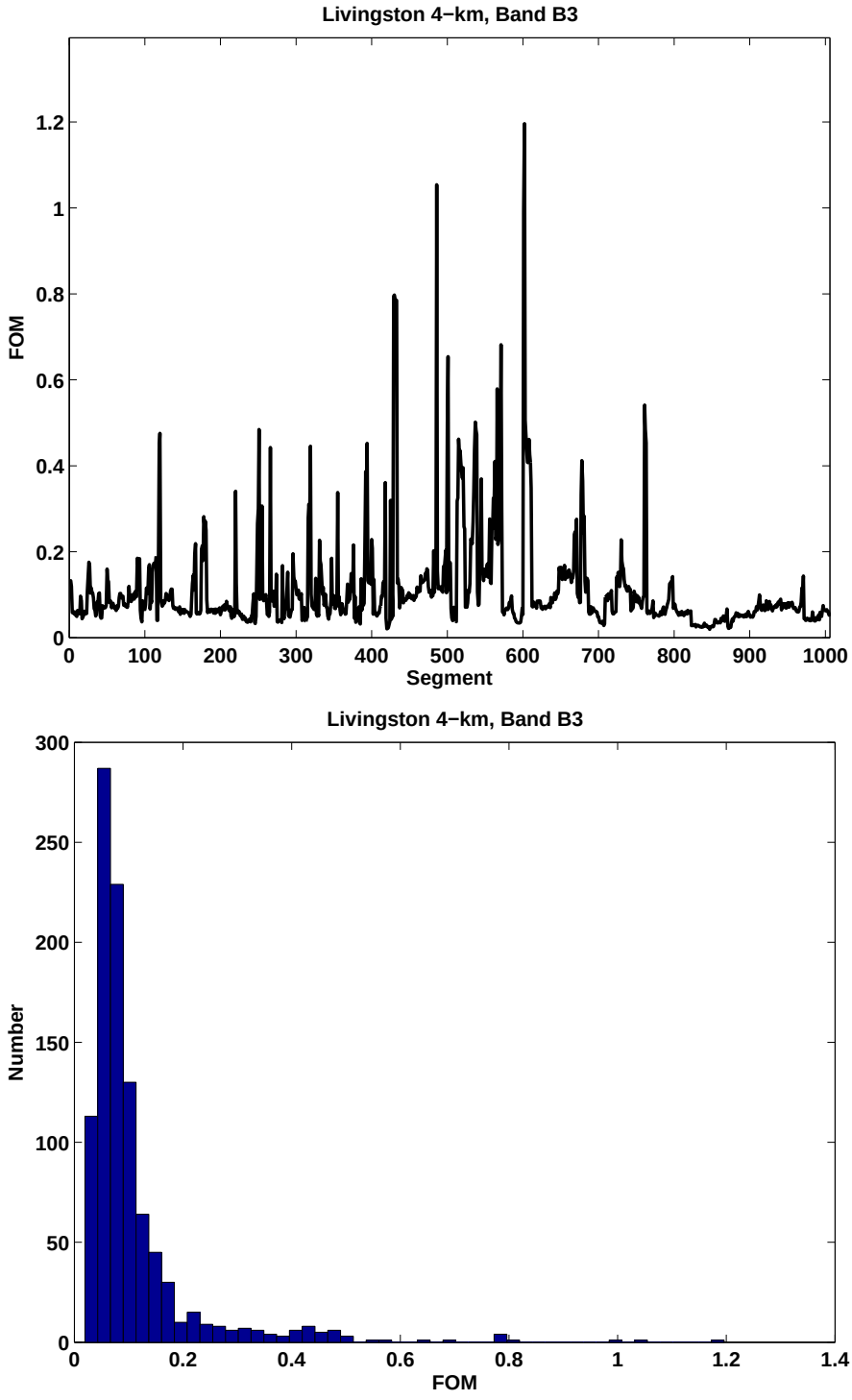


Figure E.22. Data conditioning FOM for Livingston 4-km detector, band B3

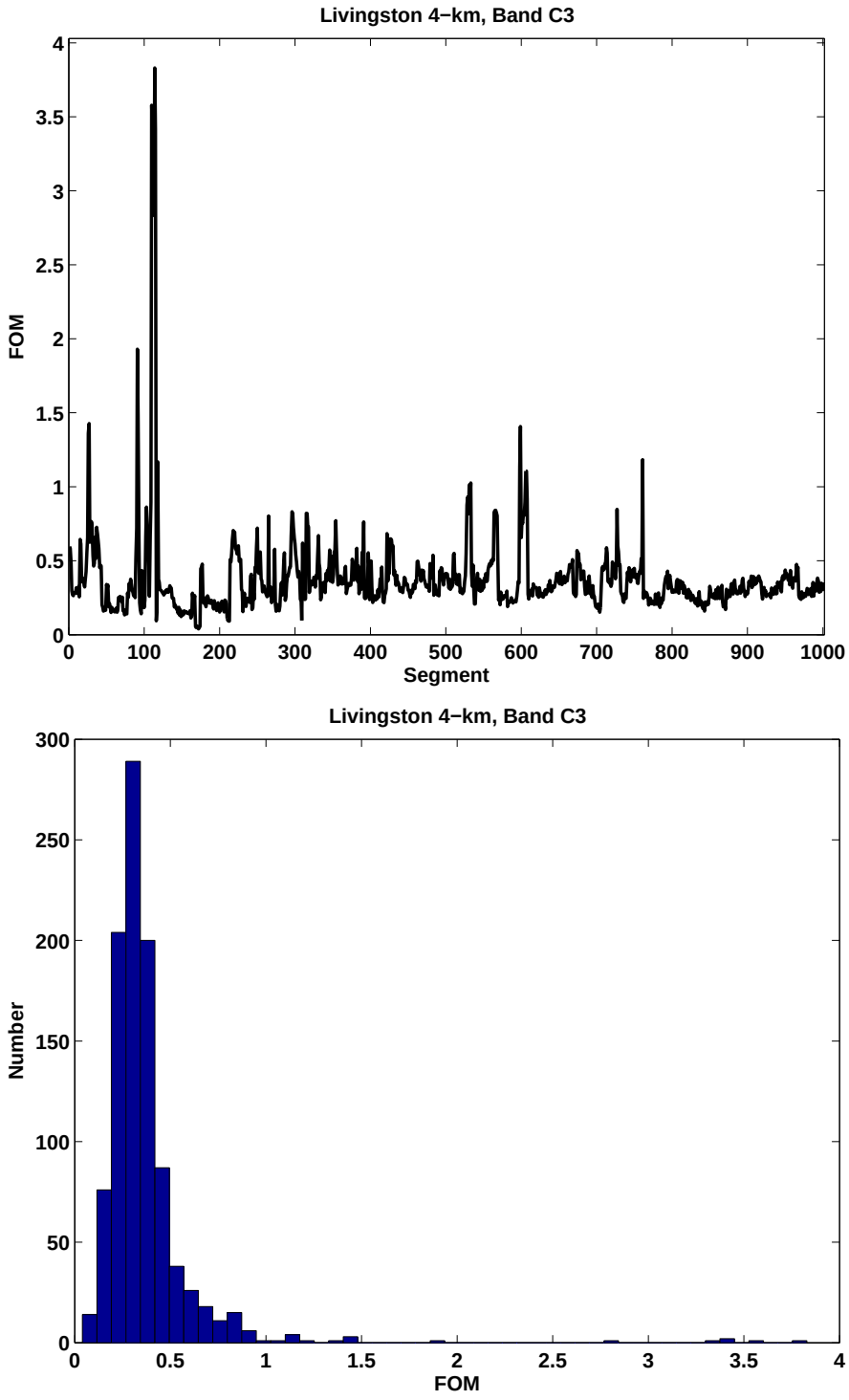


Figure E.23. Data conditioning FOM for Livingston 4-km detector, band C3

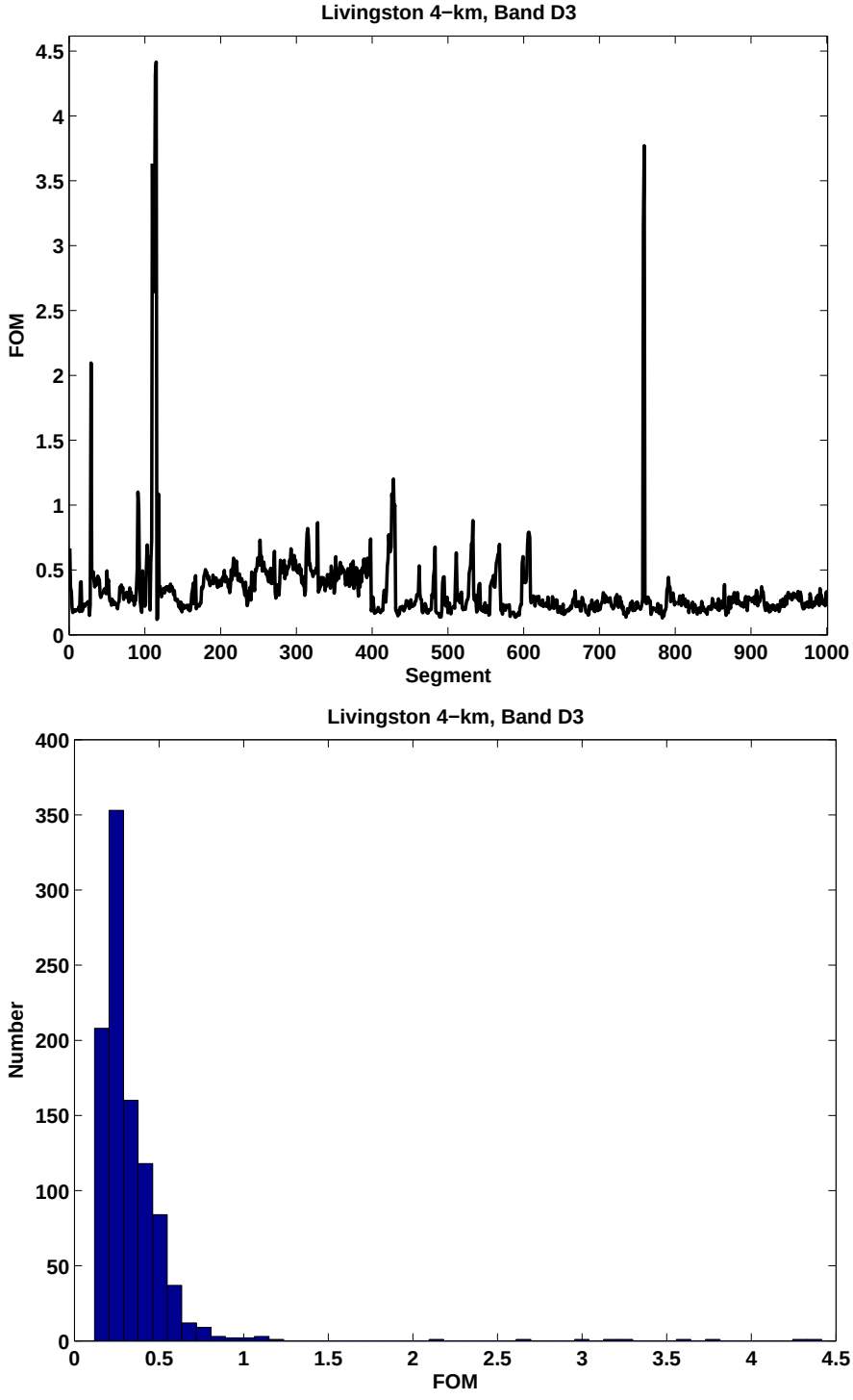


Figure E.24. Data conditioning FOM for Livingston 4-km detector, band D3

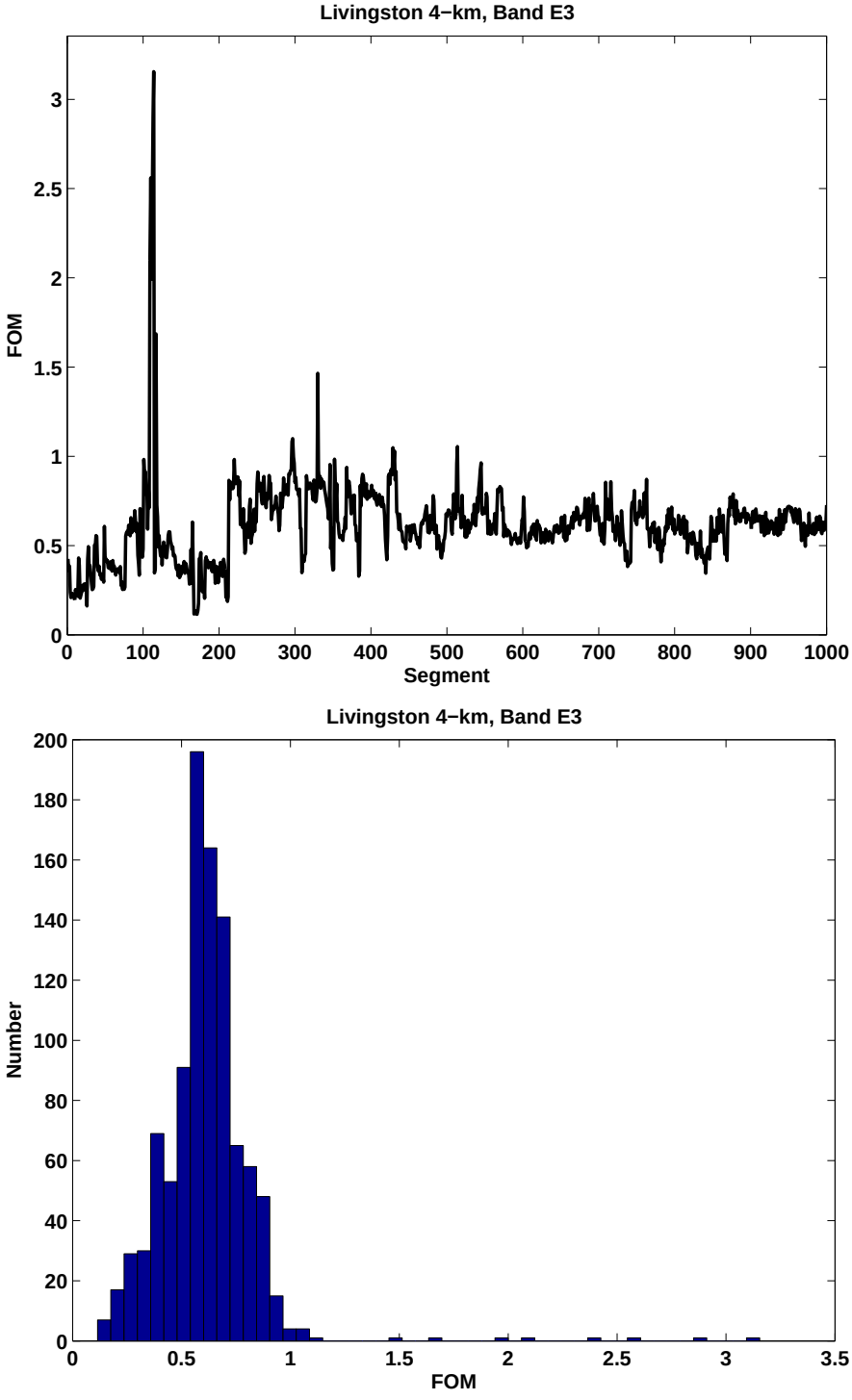


Figure E.25. Data conditioning FOM for Livingston 4-km detector, band E3

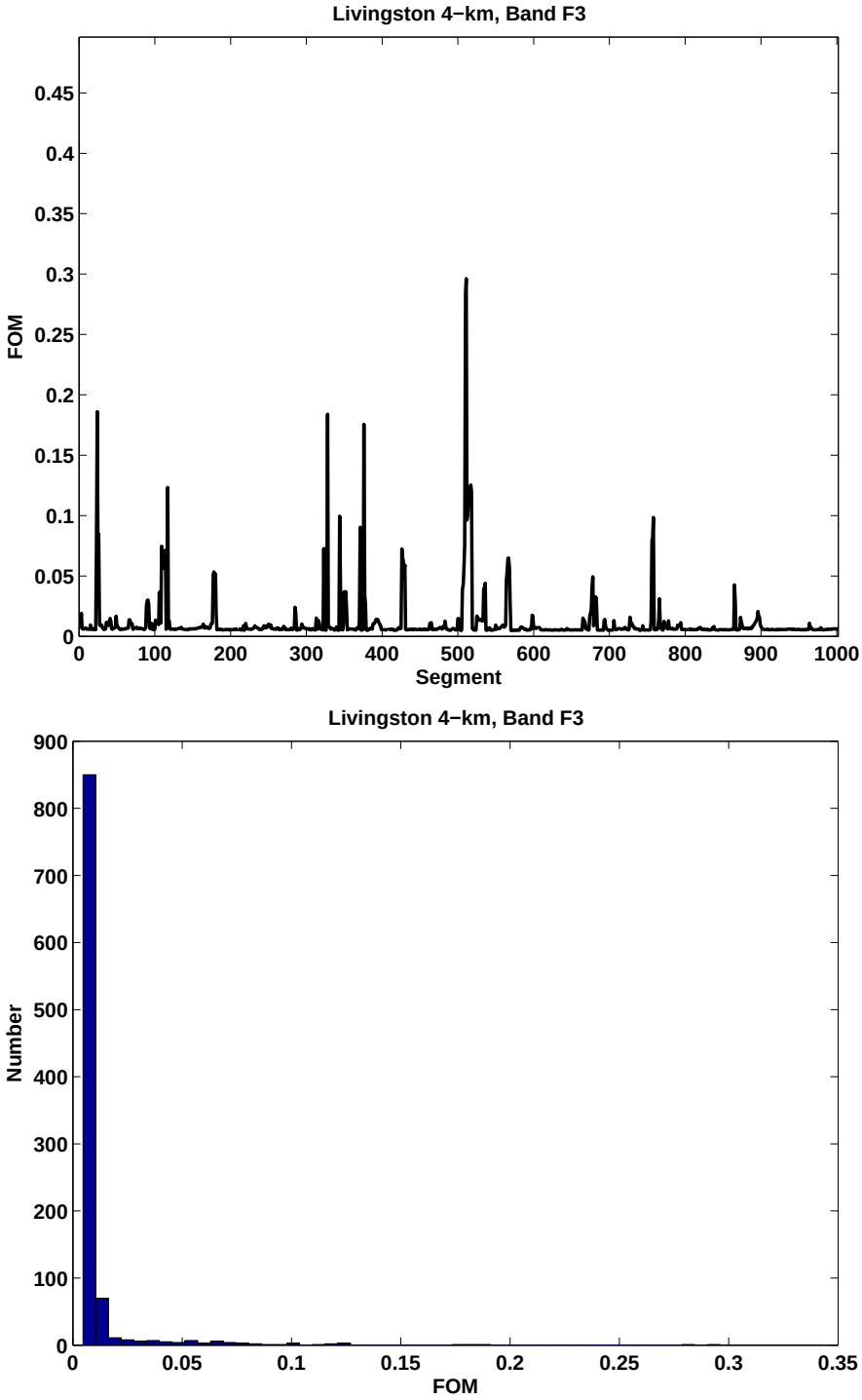


Figure E.26. Data conditioning FOM for Livingston 4-km detector, band F3

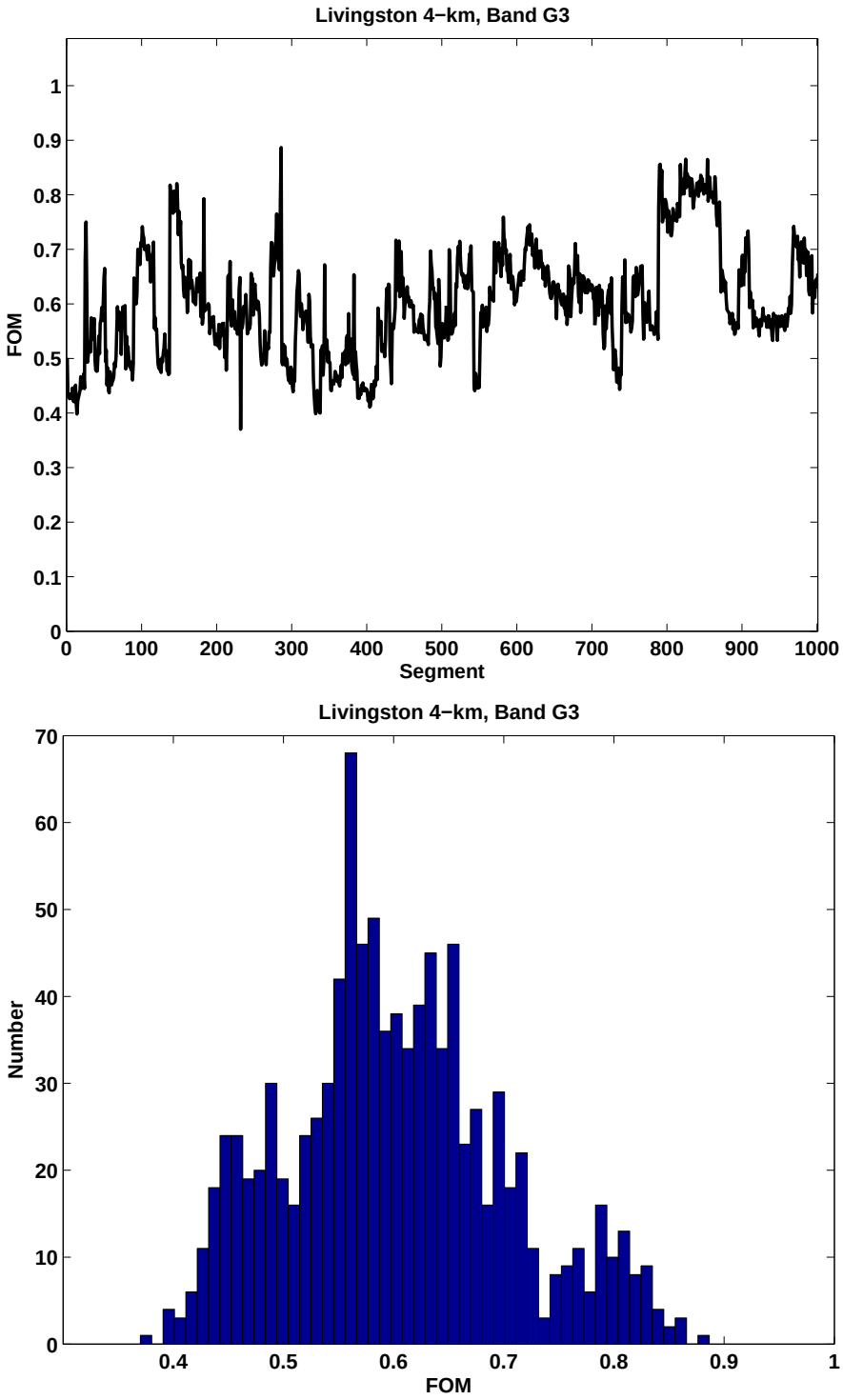


Figure E.27. Data conditioning FOM for Livingston 4-km detector, band G3

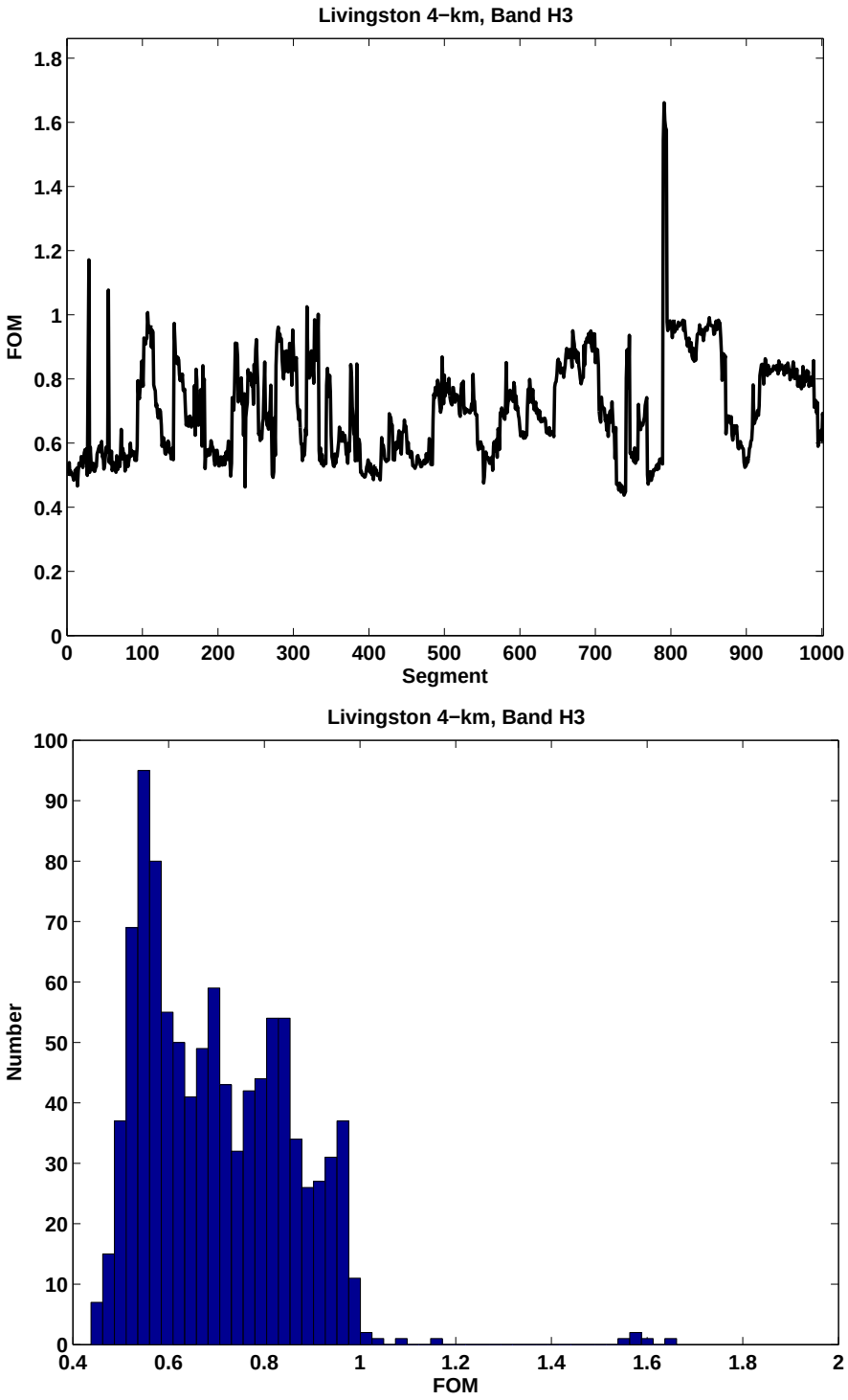


Figure E.28. Data conditioning FOM for Livingston 4-km detector, band H3

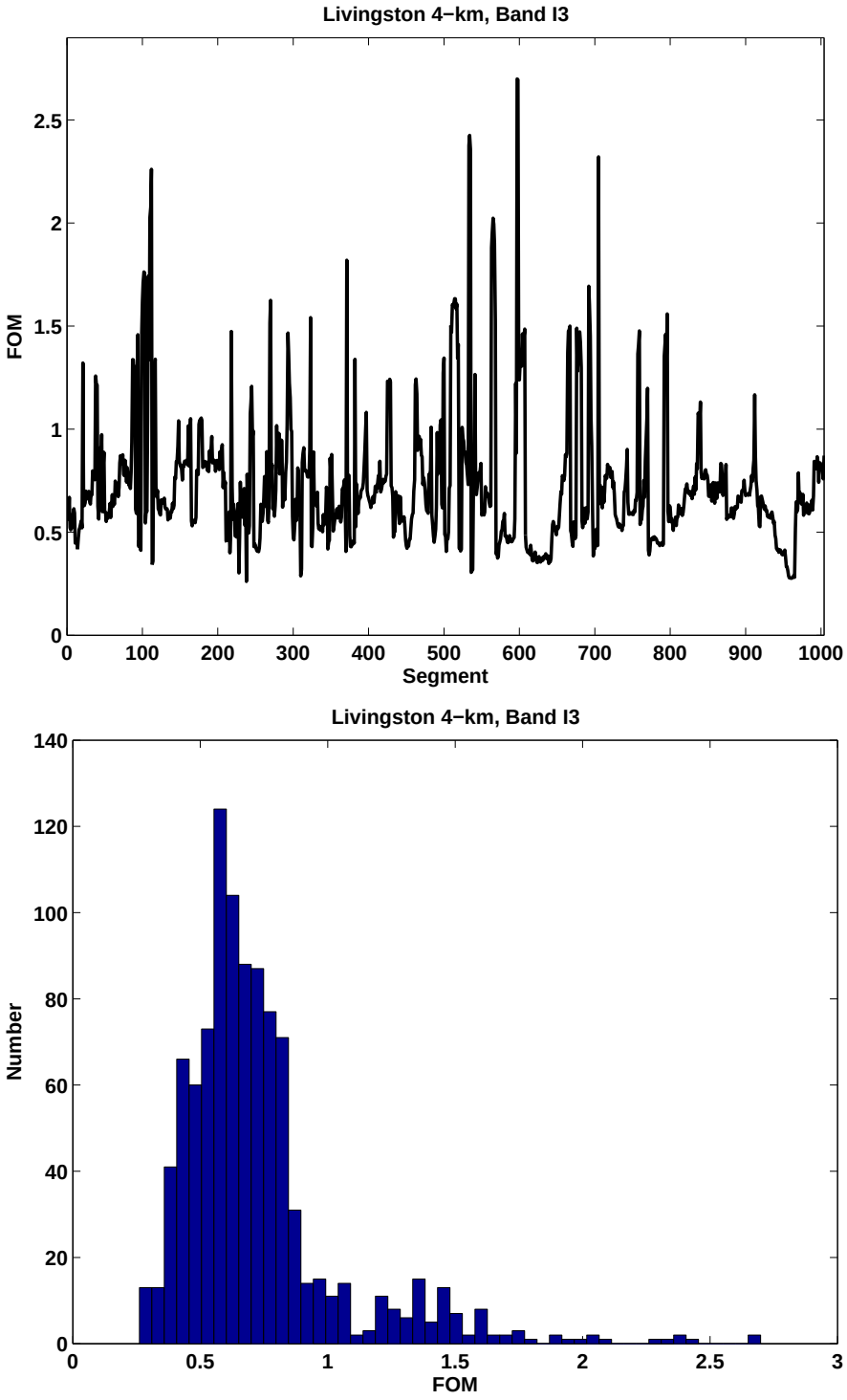


Figure E.29. Data conditioning FOM for Livingston 4-km detector, band I3

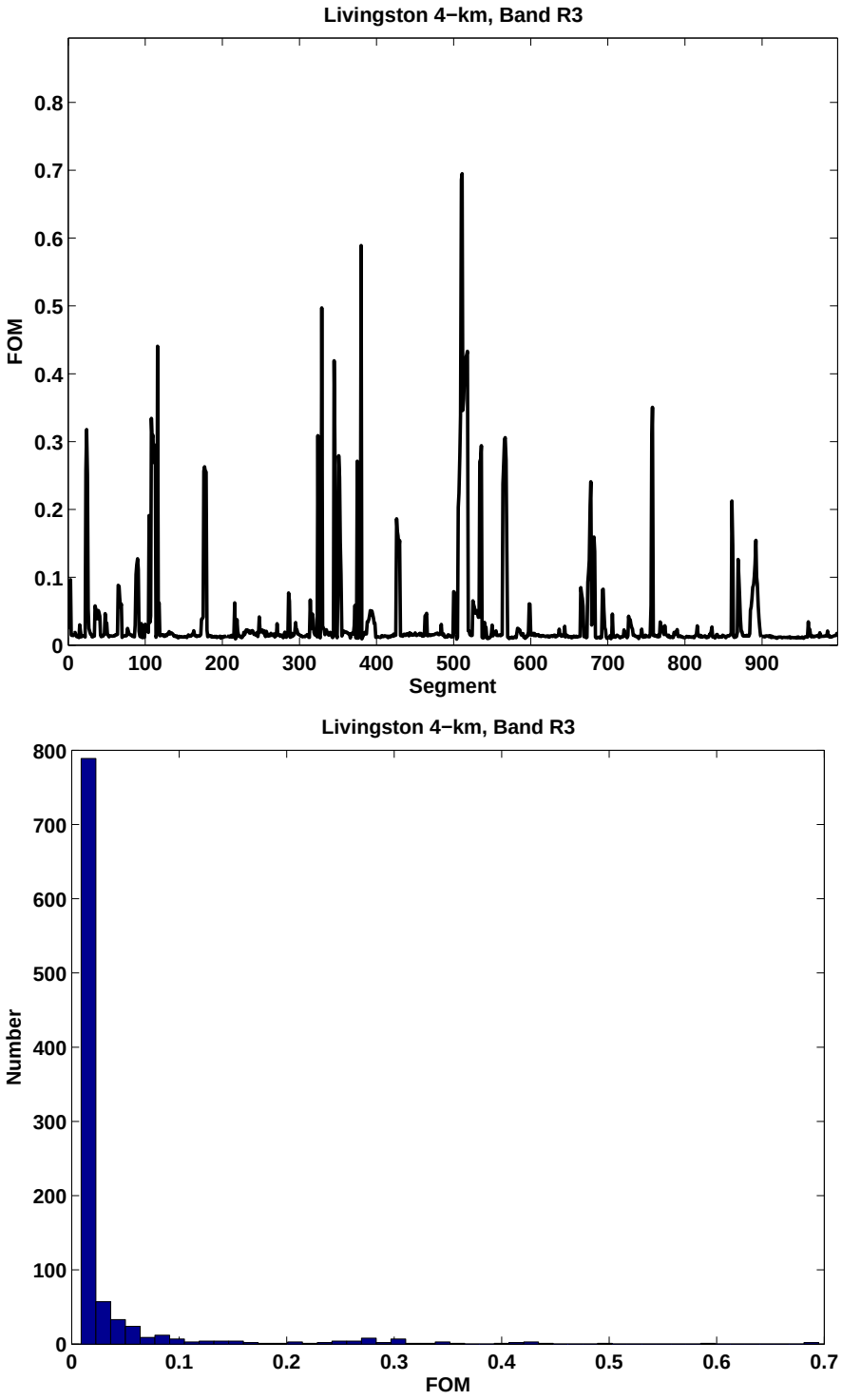


Figure E.30. Data conditioning FOM for Livingston 4-km detector, band R3

Bibliography

[1] C. D. Ott, A. Burrows, E. Livine, and R. Walder, "Gravitational waves from axisymmetric, rotating stellar core collapse," *ApJ*, vol. 600, pp. 834–864, 2004.

[2] A. Lazzarini, "Strain Sensitivities for the LIGO Interferometers for S1," LIGO, Tech. Rep. LIGO-G020461-01-E, 2003. [Online]. Available: http://admdbsrv.ligo.caltech.edu/dcc/

[3] ——, "Strain Sensitivities for the LIGO Interferometers during S2," LIGO, Tech. Rep. LIGO-G030379-00-E, 2003. [Online]. Available: http://admdbsrv.ligo.caltech.edu/dcc/

[4] ——, "S3 Best Strain Sensitivities," LIGO, Tech. Rep. LIGO-G040023-00-E, 2004. [Online]. Available: http://admdbsrv.ligo.caltech.edu/dcc/

[5] ——, "S4 Strain Sensitivities for the LIGO Interferometers," LIGO, Tech. Rep. LIGO-G050230-02-E, 2005. [Online]. Available: http://admdbsrv.ligo.caltech.edu/dcc/

[6] R. A. Hulse and J. H. Taylor, "Discovery of a pulsar in a binary system," *ApJL*, vol. 195, pp. L51–L53, Jan. 1975.

[7] J. H. Taylor and J. M. Weisberg, "A new test of general relativity - Gravitational radiation and the binary pulsar PSR 1913+16," *ApJ*, vol. 253, pp. 908–920, Feb. 1982.

[8] ——, "Further experimental tests of relativistic gravity using the binary pulsar PSR 1913 + 16," *ApJ*, vol. 345, pp. 434–450, Oct. 1989.

[9] J. Weber, "Detection and generation of gravitational waves," *Phys. Rev.*, vol. 117, pp. 306–313, 1960.

[10] E. Mauceli, Z. K. Geng, W. O. Hamilton, W. W. Johnson, S. Merkowitz, A. Morse, B. Price, and N. Solomonson, "The Allegro gravitational wave

detector: Data acquisition and analysis," *Phys. Rev. D*, vol. 54, no. 2, pp. 1264–1275, 1996.

[11] G. A. Prodi, et al., "Initial operation of the gravitational wave detector AURIGA," in *Gravitational Waves - Proceedings of the 2nd Edoardo Amaldi conference on gravitational wave experiments*, E. Coccia, G. Veneziano, and G. Pizzella, Eds. World Scientific, 1998, pp. 148–158.

[12] P. Astone, et al., "Long-term operation of the Rome "Explorer" cryogenic gravitational wave detector," *Phys. Rev. D*, vol. 47, no. 2, pp. 362–375, 1993.

[13] ——, "The gravitational wave detector NAUTILUS operating at T = 0.1 K," *Astropart. Phys.*, vol. 7, pp. 231–243, 1997.

[14] D. G. Blair, E. N. Ivanov, M. E. Tobar, P. J. Turner, F. van Kann, and I. S. Heng, "High sensitivity gravitational wave antenna with parametric transducer readout," *Phys. Rev. Lett.*, vol. 74, pp. 1908–1911, 1995.

[15] P. Astone, et al., "Increasing the bandwidth of resonant gravitational antennas: the case of EXPLORER," *Phys. Rev. Lett.*, vol. 91, no. 11, p. 111101, 2003.

[16] B. Abbott, et al., "Detector Description and Performance for the First Coincidence Observations between LIGO and GEO," *Nucl.Instrum.Meth.A*, vol. 517, p. 154, 2004. [Online]. Available: http://arxiv.org/abs/gr-qc/0308043

[17] B. Willke, et al., "Status of GEO 600," *Class. Quantum Grav.*, vol. 21, pp. S417–S423, 2004.

[18] R. Takahashi, et al., "Status of TAMA300," *Class. Quantum Grav.*, vol. 21, pp. S403–S408, 2004.

[19] F. Acernese, et al., "The present status of the VIRGO central interferometer," *Class. Quantum Grav.*, vol. 19, pp. 1421–1428, 2002.

[20] H. Tagoshi, et al, "First search for gravitational waves from inspiraling compact binaries using TAMA300 data," *Phys. Rev. D*, vol. 63, p. 062001, 2001.

[21] B. Abbott, et al, "Setting upper limits on the strength of periodic gravitational waves from PSR J1939+2134 using the first science data from the GEO 600 and LIGO detectors," *Phys. Rev. D*, vol. 69, p. 082004, 2004.

[22] ——, "Analysis of LIGO data for gravitational waves from binary neutron stars," *Phys. Rev. D*, vol. 69, p. 122001, 2004.

[23] ——, "First upper limits from LIGO on gravitational wave bursts," *Phys. Rev. D*, vol. 69, p. 102001, 2004.

[24] ——, "Analysis of first LIGO science data for stochastic gravitational waves," *Phys. Rev. D*, vol. 69, p. 122004, 2004.

[25] H. Takahashi, et al, "Coincidence analys to search for inspiraling compact binaries using TAMA300 and LISM data," *Phys. Rev. D*, vol. 70, p. 042003, 2004.

[26] M. Ando, et al, "Observation results by the TAMA300 detector on gravitational wave bursts from stellar-core collapses," *Phys. Rev. D*, vol. 71, p. 082002, 2005.

[27] B. Abbott, et al, "Upper limits on gravitational wave bursts in LIGO's second science run," *Phys. Rev. D*, vol. 72, p. 062001, 2005.

[28] ——, "Search for gravitational waves associated with the gamma ray burst GRB030329 using the LIGO detectors," *Phys. Rev. D*, vol. 72, p. 042002, 2005.

[29] ——, "Limits on gravitational-wave emission from selected pulsars using LIGO data," *Phys. Rev. Lett.*, vol. 94, p. 181103, 2005.

[30] B. Abbott, et al., "First all-sky upper limits from LIGO on the strength of periodic gravitational waves using the Hough transform," *Phys. Rev. D*, vol. 72, p. 102004, 2005.

[31] LIGO, TAMA, "Upper limits from the LIGO and TAMA detectors on the rate of gravitational-wave bursts," *Phys. Rev. D*, vol. 72, p. 122004, 2005.

[32] B. Abbott, et al., "Search for gravitational waves from galactic and extra-galactic binary neutron stars," *Phys. Rev. D*, vol. 72, p. 082001, 2005.

[33] ——, "Search for gravitational waves from primordial black hole binary coalescences in the galactic halo," *Phys. Rev. D*, vol. 72, p. 082002, 2005.

[34] ——, "Upper limits on a stochastic background of gravitational waves," *Phys. Rev. Lett.*, vol. 95, p. 221101, 2005.

[35] J. W. C. McNabb, M. Ashley, L. S. Finn, E. Rotthoff, A. Stuver, T. Summerscales, P. Sutton, M. Tibbits, K. Thorne, and K. Zaleski, "Overview of the BlockNormal event trigger generator," *Class. Qant. Grav.*, vol. 21, pp. S1705–S1710, 2004.

[36] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing.* Upper Saddle River, New Jersey: Prentice Hall, 1999.

[37] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Transaction of the ASME – Journal of Basic Engineering*, pp. 35–45, 1960.

[38] G. Welch and G. Bishop, "An Introduction to the Kalman Filter," 2004. [Online]. Available: http://www.cs.unc.edu/ welch/media/pdf/kalman_intro.pdf

[39] L. S. Finn and S. Mukherjee, "Data conditioning for gravitational wave detectors: A Kalman filter for regressing suspension violin modes," *Phys. Rev. D*, vol. 63, 2001.

[40] R. G. Brown and P. Y. Hwang, *Introduction to Random Signals and Applied Kalman Filtering.* New York: John Wiley & Sons, 1997.

[41] L. S. Finn and S. Mukherjee, "Erratum: Data conditioning for gravitational wave detectors: A Kalman filter for regressing suspension violin modes," *Phys. Rev. D*, vol. 67, 2003.

[42] L. S. Finn and P. Saulson, "S2 Playground Selection," LIGO, Tech. Rep. LIGO-T030020-01-Z, 2003. [Online]. Available: http://admdbsrv.ligo.caltech.edu/dcc/

[43] L. S. Finn, "S3 Playground Selection," LIGO, Tech. Rep. LIGO-T030256-00-Z, 2003. [Online]. Available: http://admdbsrv.ligo.caltech.edu/dcc/

[44] L. Ljung, *System Identification Toolbox.* The Mathworks, Inc., 2000.

[45] K. Thorne, "Testing data conditioning used in blocknormal event trigger generator," LIGO, Tech. Rep. LIGO-T060028-00-Z. [Online]. Available: http://admdbsrv.ligo.caltech.edu/dcc/

[46] J. G. Zweizig, "S2/S3 Data Quality Flagging," LIGO, Tech. Rep. LIGO-G030642-00-E, 2003. [Online]. Available: http://admdbsrv.ligo.caltech.edu/dcc/

[47] ——, "S3 Data Quality Investigation," LIGO, Tech. Rep. LIGO-G040136-00-E, 2004. [Online]. Available: http://admdbsrv.ligo.caltech.edu/dcc/

[48] ——, "S4 Data Quality," LIGO, Tech. Rep. LIGO-G050197-00-E, 2005. [Online]. Available: http://admdbsrv.ligo.caltech.edu/dcc/

[49] ——, "S3 Data Quality Review." [Online]. Available: http://www.ligo.caltech.edu/%7Ejzweizig/S3_Data_Quality/index.html

[50] J. R. Taylor, *An Introduction to Error Analysis.* Sausalito, California: University Science Books, 1997.

[51] P. E. Greenwood and M. S. Nikulin, *A Guide to Chi-Squared Testing.* New York: John Wiley & Sons, Inc., 1996.

[52] G. J. Feldman and R. D. Cousins, "Unified approach to the classical statistical analysis of small signals," *Phys. Rev. D*, vol. 57, pp. 3873–3889, 1998.

[53] C. L. Nikias and A. P. Petropulu, *Higher-Order Spectra Analysis.* Englewood Cliffs, New Jersey: PTR Prentice Hall, 1993.

[54] S. McLaughlin, A. Stogioglu, and J. Fackrell, "Introducing Higher Order Statistics (HOS) for the Detection of Nonlinearities," 1995, http://www.amsta.leeds.ac.uk/Applied/news.dir/issue2/hos_intro.html.

[55] M. J. Hinich, "Testing for gaussianity and lineariey of a stationary time series," *Journal of Time Series Analysis*, vol. 3, no. 3, pp. 169–176, 1982.

[56] Y. Gürsel and M. Tinto, "Near optimal solution to the inverse problem for gravitational-wave bursts," *Phys. Rev. D.*, vol. 40, pp. 3884–3938, 1989.

[57] K. Maisinger, M. P. Hobson, and A. N. Lasenby, "A maximum entropy method for reconstructing interferometer maps of fluctuations in the cosmic microwave background radiation," *MNRAS*, vol. 290, pp. 313–326, Sept. 1997.

[58] A. W. Jones, S. Hancock, A. N. Lasenby, R. D. Davies, C. M. Gutiérrez, G. Rocha, R. A. Watson, and R. Rebolo, "10-GHz Tenerife cosmic microwave background observations at 8 deg resolution and their analysis using a new maximum entropy method," *MNRAS*, vol. 294, pp. 582–+, Mar. 1998.

[59] P. Vielva, R. B. Barreiro, M. P. Hobson, E. Martínez-González, A. N. Lasenby, J. L. Sanz, and L. Toffolatti, "Combining maximum-entropy and the Mexican hat wavelet to reconstruct the microwave sky," *MNRAS*, vol. 328, pp. 1–16, Nov. 2001.

[60] C. L. Bennett, R. S. Hill, G. Hinshaw, M. R. Nolta, N. Odegard, L. Page, D. N. Spergel, J. L. Weiland, E. L. Wright, M. Halpern, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. Wollack, "First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Foreground Emission," *ApJS*, vol. 148, pp. 97–117, Sept. 2003.

[61] A. W. Jones, A. N. Lasenby, P. Mukherjee, C. M. Gutierrez, R. D. Davies, R. A. Watson, R. Hoyland, and R. Rebolo, "A multifrequency maximum-entropy joint analysis of COBE and Tenerife data," *MNRAS*, vol. 310, pp. 105–109, Nov. 1999.

[62] R. B. Barreiro, M. P. Hobson, A. J. Banday, A. N. Lasenby, V. Stolyarov, P. Vielva, and K. M. Górski, "Foreground separation using a flexible maximum-entropy algorithm: an application to COBE data," *MNRAS*, vol. 351, pp. 515–540, June 2004.

[63] R. Brady, Patrick and S. Ray-Majumder, "Incorporating information from source simulations into searches for gravitational-wave bursts," *Class. Quant. Grav.*, vol. 21, pp. S1839–S1847, 2004.

[64] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed. Prentice-Hall, 2002.

[65] E. T. Jaynes, "Information Theory and Statistical Mechanics," *Phys. Rev.*, vol. 106, pp. 620–630, 1957.

[66] ——, "Information Theory and Statistical Mechanics II," *Phys. Rev.*, vol. 108, pp. 171–190, 1957.

[67] C. E. Shannon, "A mathematical theory of communication," *Bell System Tech. J.*, vol. 27, pp. 379–457, 1948.

[68] M. P. Hobson and A. N. Lasenby, "The entropic prior for distributions with positive and negative values," *MNRAS*, vol. 298, pp. 905–908, Aug. 1998.

[69] S. F. Gull, "Developments in Maximum Entropy Data Analysis," in *Maximum Entropy and Bayesian Methods*, J. Skilling, Ed. Kluwer Acad. Pub., 1989, pp. 53–71.

[70] D. J. C. MacKay, "Bayesian interpolation," *Neural Comput.*, vol. 4, pp. 415–447, 1991.

[71] R. P. Barry and R. K. Pace, "Monte Carlo estimates of the log determinant of large sparse matricies," *Linear Algebra and its Applications*, vol. 289, pp. 41–54, 1999.

[72] S. Woosley and T. Weaver, "The Evolution and Explosion of Massive Stars. II. Explosive Hydrodynamics and Nucleosynthesis," *ApJS*, vol. 101, pp. 181–235, 1995.

[73] J. M. Lattimer and F. D. Swesty, "A generalized equation of state for hot, dense matter," *Nucl. Phys. A*, vol. 535, pp. 331–376, 1991.

[74] A. Heger, N. Langer, and S. E. Woosley, "Presupernova Evolution of Rotating Massive Stars. I. Numerical Method and Evolution of the Internal Stellar Structure," *ApJ*, vol. 528, pp. 368–396, Jan. 2000.

[75] A. Heger, S. E. Woosley, and N. Langer, "The pre-supernova evolution of rotating massive stars," in *IAU Symposium*, 2003, pp. 357–+.

[76] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*. San Francisco: Freeman, 1973.

[77] A. Stuver, "GravEn Simulation Engine Primer," LIGO, Tech. Rep. LIGO-T040020-01-Z, 2004. [Online]. Available: http://admdbsrv.ligo.caltech.edu/dcc/

[78] R. Adhikari, P. Fritshel, G. González, M. Landry, L. Matone, H. Radkins, A. Takamori, and B. O'Reilly, "Calibration of the LIGO detectors for the First LIGO Scientific Run," LIGO, Tech. Rep. LIGO-T030097-00-D, 2003. [Online]. Available: http://admdbsrv.ligo.caltech.edu/dcc/

[79] G. González, M. Landry, B. O'Reilly, and H. Radkins, "Calibration of the LIGO detectors for S2," LIGO, Tech. Rep. LIGO-T030097-01-D, 2004. [Online]. Available: http://admdbsrv.ligo.caltech.edu/dcc/

[80] G. González, M. Landry, B. O'Reilly, and X. Siemens, "Calibration of the LIGO detectors for S3," LIGO, Tech. Rep. LIGO-T050059-01-D, 2005. [Online]. Available: http://admdbsrv.ligo.caltech.edu/dcc/

[81] A. Dietz, J. Garofoli, G. González, M. Landry, B. O'Reilly, and M. Sung, "Calibration of the LIGO detectors for S4," LIGO, Tech. Rep. LIGO-T050262-00-D, 2005. [Online]. Available: http://admdbsrv.ligo.caltech.edu/dcc/

[82] R. Adhikari, G. González, M. Landry, and B. O'Reilly, "Calibration of the LIGO detectors for the first LIGO science run," *Class. Quantum. Grav.*, vol. 20, pp. S903–S914, 2003.

[83] T. Zwerger and E. Müller, "Dynamics and gravitational wave signature of axisymmetric rotational core collapse." *Astron. Astrophys.*, vol. 320, pp. 209–227, Apr. 1997.

[84] MATLAB is a product of The MathWorks.

Vita

Tiffany Z. Summerscales

Education:

Penn State University, University Park, PA — 1999-present
- PhD. in Physics — May 2006

Andrews University, Berrien Springs, MI — 1995-1999
- B.S. in Mathematics and Physics — June 1999
- Honors Thesis: Searching for Evidence of Multiple Interactions in the 1995 ZEUS Shifted Vertex Data

Honors & Awards:
- David C. Duncan Fellowship — Apr 2001
- Nellie H. and Oscar L. Roberts Fellowship — Aug 1999
- Graduated from Andrews University Summa Cum Laude and J.N. Andrews Scholar — Jun 1999
- Barry M. Goldwater Scholar — Apr 1997

Teaching Experience:

Penn State University, Department of Physics — 1999-present
- Grader, PHYS501 General Relativity I
- Teaching Assistant, PHYS211 Mechanics

Andrews University, Department of Physics — 1995-1999
- Lab Teaching Assistant for Astronomy, Concepts of Physical Science and General Physics

Research Experience:
- Gravitational Wave Astronomy: Member of the LIGO (Laser Interferometer Gravitational-wave Observatory) Collaboration. Projects in detector characterization and data analysis. — 2000-present
- High Energy Physics: Searched for evidence of multiple interactions in dijet events in the ZEUS detector. — 1996-1999