LASER INTERFEROMETER GRAVITATIONAL WAVE OBSERVATORY
-LIGO-
CALIFORNIA INSTITUTE OF TECHNOLOGY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Technical Note **LIGO-T070037- 00- W** 8/06 (Rev. 2/26/07, 3/24/07)

Development of a High Frequency Burst Analysis Pipeline - SURF Project Final Report

Jeff Parker
Mentors: Rick Savage, Greg Mendell, and Malik Rakhmanov

This is an internal working note
of the LIGO Project.

California Institute of Technology
LIGO Project - MS 18-34
Pasadena CA 91125
Phone (626) 395-2129
Fax (626) 304-9834
E-mail: info@ligo.caltech.edu

Massachusetts Institute of Technology
LIGO Project - MS NW17-161
Cambridge, MA 02139
Phone (617) 253-4824
Fax (617) 253-4824
E-mail: info@ligo.mit.edu

WWW: http://www.ligo.caltech.edu/

file /home/name/directory/T070037.pdf

Abstract

LIGO's current searches for gravitational waves take place in the sub-8 kHz regime. Sensitivity drops off for higher frequencies but peaks again at multiples of the free spectral range (FSR) frequency, 37.52 kHz. A channel with sufficient sampling rate has been only recently installed, and analysis at this frequency is uncharted territory. An initial analysis pipeline consisting of two main components has been created. The first component is the frequency and sky position dependent detector response which takes an arbitrary waveform and modifies it to what would be outputted by the detector and added to interferometer noise. For signals with a frequency near the FSR frequency, the peak in sensitivity causes resonance and an exponentially decaying ringing in the modified signal, with decay time of about 2 ms. The second component is an excess power algorithm, used to determine if a signal is present by looking for power greater than expected due to noise alone. The analysis pipeline allows us to inject signals from sources located at various positions in the sky and compute detection efficiency near the FSR frequency, as well as compare the all-sky average efficiency to that at low frequencies.

Introduction

The Laser Interferometer Gravitational-Wave Observatory (LIGO) seeks to detect gravitational radiation from astrophysical sources, primarily focusing on waves with frequency of less than 8 kHz. Most of LIGO's high sensitivity band falls within that region, but there is an additional sensitivity peak at 37.52 kHz. Recently, there have been proposals to extend LIGO's search to higher frequency waves [1]. LIGO's low frequency channels are thoroughly searched, but analysis in the sensitive high frequency region is uncharted territory. Possible sources in these high frequency channels include the stochastic gravitational background [2], high order neutron star vibrational modes, and high frequency components of black hole ring down predicted by string theory [3]. In the absence of detection, upper limits on such sources could be set.

The various elements of an analysis pipeline are examined. At the beginning of the analysis, we would like to take into account the effect of the frequency-dependent detector response on incoming signals. We examine these effects on a sine-Gaussian, a typical waveform, at the frequencies of interest. We then consider an algorithm to determine if a signal is present within noise and discuss an example where it might be used.

Cavity Response

Due to properties of a Fabry-Perot cavity, the LIGO arms' length sensitivity is dependent on the frequency at which the mirrors move. The sensitivity peaks at DC, decays as approximately $1/f$, and then peaks again at multiples of the free spectral range (FSR) frequency, where FSR is the frequency at which a signal's period of oscillation is exactly equal to the time it takes for a photon to make a round trip in one of the arms. For the 4-km-long LIGO

interferometers, the FSR frequency is about 37.52 kHz. The cavity response is not dependent on gravitational waves and occurs for all sources of mirror motion, including vibrational noise. The normalized transfer function is (See [4] for more information)

$$H(s) = \frac{1 - r_a r_b}{1 - r_a r_b e^{-2sT}}, \tag{1}$$

where s is the Laplace domain frequency variable, r_a and r_b are the mirror reflectivities, and $2T$ is the round trip time for a photon in the arm. At integer multiples of the FSR frequency the response peaks (Fig. 1).

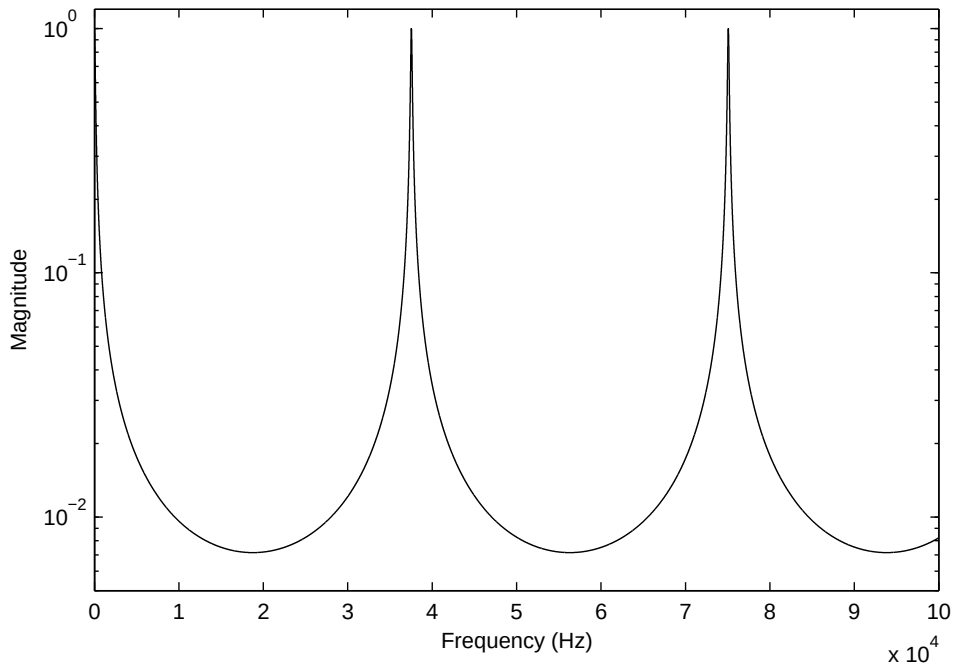

Figure 1: Cavity response of the interferometer. Peaks occur at multiples of the FSR frequency (37.5 kHz).

Directional Sensitivity

There is also a directional dependence to LIGO's strain sensitivity. A gravitational wave coming from straight overhead, polarized along the arms, is called *optimally oriented*, because it provides the maximum response at DC. On the other hand, LIGO has no strain sensitivity to a gravitational wave that comes in the plane of the detector along the bisector of the arms. At these low frequencies, the period of the gravitational wave is much longer than the round trip time in the arm, or $1/f_{gw} >> 2T$. In this regime, the wave oscillates slowly and does not change much in one round trip of a photon. The effects of the propagation time are small and

neglected. The directional sensitivity, called an antenna pattern, is currently calculated at DC and used for all frequencies up to 8 kHz.

As f_{gw} increases to the FSR frequency, the wave oscillates significantly in one round trip of a photon. The effect of the photon propagation can no longer be neglected at higher frequencies. The exact frequency dependent directional sensitivity has been worked out [5]. It is important to note that antenna patterns are not frequency independent in this regime, or equivalently, for a given orientation, the strain sensitivity depends on frequency. For example, the *optimal orientation* for DC provides no strain sensitivity at the FSR frequency, and some orientations provide sensitivity at the FSR frequency but not at DC (Fig. 2). For plots and a more detailed discussion of the antenna patterns, see [6].

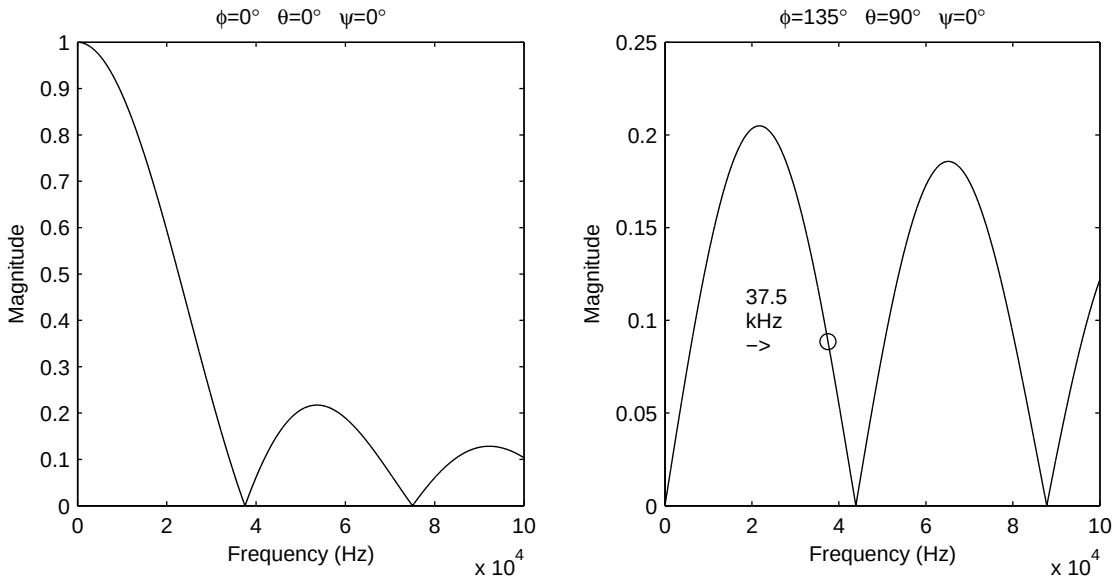

Figure 2: Sensitivity vs. frequency at two source locations, where θ and ϕ are standard spherical coordinates in the detector frame and ψ refers to orientation. Left: A source directly overhead the interferometer (ϕ=0°, θ=0°, ψ=0°)is optimal for DC but provides no sensitivity at the FSR frequency. Right: A source at this location (ϕ=135°, θ=90°, ψ=0°) provides no sensitivity at DC but some at the FSR frequency.

A Sine-Gaussian

The detector response, consisting of the cavity response and directional sensitivity, determines how a gravitational wave signal would be seen by LIGO. Before conducting long-running searches, it would be useful to know how the detector response affects an incoming gravitational wave at the FSR frequency. As the exact shape of a signal waveform is unknown, a sine-Gaussian burst of .5 ms is used as one possible model [7]. The signal is given a location

and orientation in the sky, described by ϕ, θ, which are the standard spherical coordinates in the detector reference frame, and ψ, an orientation angle. (See [5] for more detail).

For signals with center frequency far from the FSR frequency, there is no qualitative difference between the source waveform and the detected waveform; there is an amplitude change but the general shape remains the same. Near 35.5 kHz, a slight tail appears on the waveform. At the FSR frequency, there is a significant change in the overall shape of the waveform (Fig. 3). An exponentially decaying tail dominates the Gaussian envelope and drops off with a decay time $\tau = 2$ ms. The signal 'rings' due to the resonance induced by the cavity response. For bursts with duration less than 2 ms, the ringing is a significant effect. The detector signal also exhibits a delay in rise to maximum amplitude, caused by the time necessary for the energy to build up in the cavity at resonance. The ringing effect is due solely to the peaks in the length response and is not qualitatively changed at different sky locations.

This example was calculated with a center frequency equal to the FSR frequency, but similar effects occur at 2 $\times$ FSR, 3 $\times$ FSR, and so on.

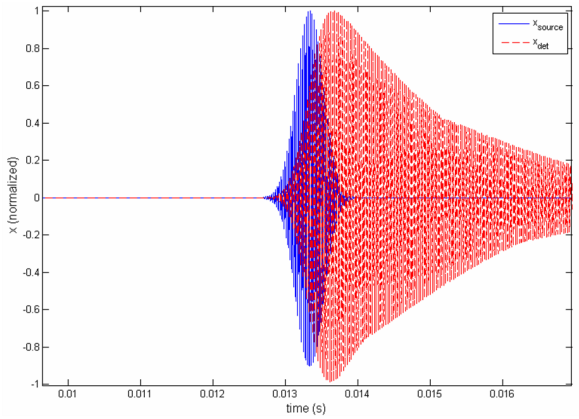

Figure 3: Sine-Gaussian signal of the source (blue) and detector (red). Plots have been normalized to compare shapes. Note the exponentially decaying tail of the detector signal with decay time $\tau = 2$ ms.

Analysis Pipeline

The sequence of events leading to a gravitational wave detection can be stated simply as follows:

1. A gravitational wave is emitted from a source.

2. The gravitational wave passes through LIGO, producing a phase change in the light that corresponds to an 'effective' length change of the arms. The signal at the photodetector is modified by the detector response – the cavity length response and the directional sensitivity.

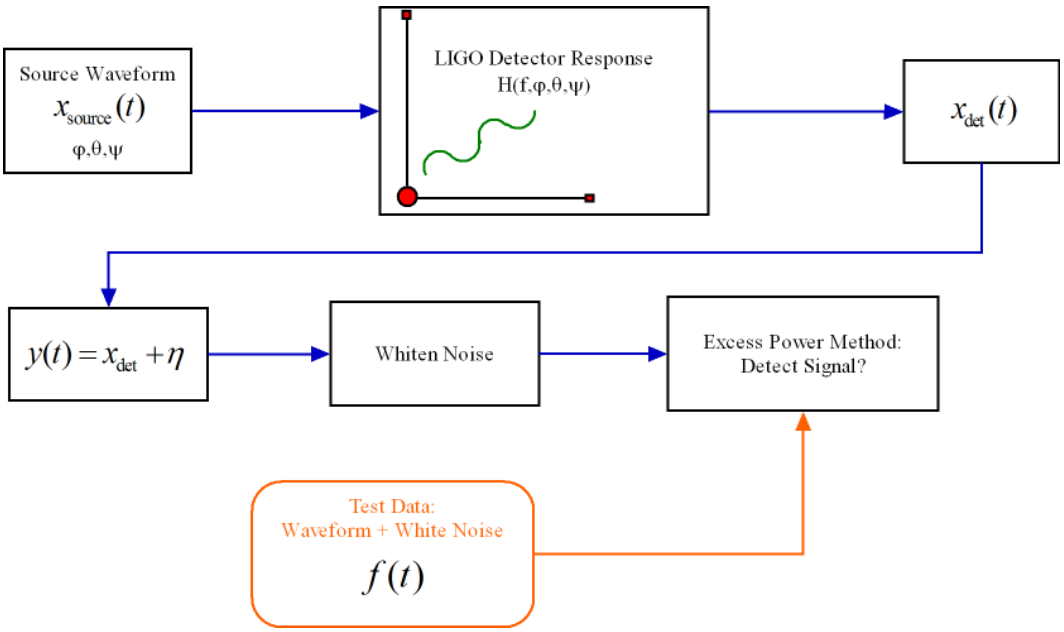

Figure 4: Analysis Pipeline. η represents interferometer noise.

3. The gravitational wave signal is joined with noise: shot noise, seismic noise, vibrational modes, etc.

4. LIGO attempts to detect the gravitational wave signal in the noise.

With the detector response in place, we can construct an analysis pipeline to mimic these steps (Fig. 4). A generated signal is given a sky location and modified by the detector response and then added to LIGO noise, at which point detection is attempted through some algorithm. The detection algorithm used here, called the excess power method [8], is described in the next section, Methods. It can, for simplicity, be tested on a signal added to white noise rather than interferometer data. The analysis pipeline may be used to perform simulations with injected signals or to run searches on interferometer data.

Methods

Detector Response

The detector response determines how a gravitational wave signal would be seen by LIGO. The signal that LIGO sees, $x_{det}(t)$, is a function of the source signal, $x_{source}(t)$, and the angles ϕ, θ, and ψ.

Application of the frequency-dependent detector response must be done in the frequency domain [6]. A fast Fourier transform function (FFT) is used to obtain the spectrum $\tilde{X}_{source}(f)$

of the time signal.

$$\tilde{X}_{source}(f) \Leftrightarrow x_{source}(t) \tag{2}$$

The cavity response and directional sensitivity, represented here by $H(f)$ and $D(f, \phi, \theta, \psi)$, respectively, are then applied for each frequency component to calculate the spectrum of the detector signal.

$$\tilde{X}_{det}(f) = H(f)D(f, \phi, \theta, \psi)\tilde{X}_{source}(f) \tag{3}$$

The new spectrum $\tilde{X}_{det}(f)$ is then transformed back to a time signal $x_{det}(t)$.

$$x_{det}(t) \Leftrightarrow \tilde{X}_{det}(f) \tag{4}$$

Excess Power

The excess power method is useful for detecting signals in the presence of white Gaussian noise [8]. It has the advantage of not requiring knowledge of the duration or frequency bandwidth of the waveform, which is often unknown, especially for high frequencies. The essence of the excess power method is to look for power greater than one would expect from noise alone; that is, to look for *excess* power.

Let detector output be $y(t)$, noise $n(t)$, and a (possibly absent) signal $x(t)$. In the absence of a signal,

$$y = n. \tag{5}$$

Power is proportional to y^2, and thus, letting $<>$ denote average value,

$$< y^2 > \;\; = \;\; < n^2 > \tag{6}$$

If a signal is present,

$$< y^2 > \;\; = \;\; < n^2 > + < x^2 > + 2 < nx > . \tag{7}$$

If we assume that n and x are uncorrelated, then $< nx > = 0$, and

$$< y^2 > \;\; = \;\; < n^2 > + < x^2 > . \tag{8}$$

If a signal is present, $< x^2 >$ is positive definite, and thus the expected power in the detector output will be greater than in the absence of a signal.

Because LIGO's high frequency sensitivity peak is centered in a narrow band around the FSR frequency, it makes sense to focus our search within that band. Fourier analysis is used to break the power in the signal up into different frequency bins. We could imagine taking the Fast Fourier Transform (FFT) of 1 second of time domain data and having a frequency resolution of 1 Hz, but then relatively short duration bursts could be drowned out by the noise power. Instead, we create a *Time-Frequency Plane* (TF-Plane), or a spectrogram, where we divide the time data into many time bins and FFT each time bin separately. In this way, we have better resolution in time, but the frequency resolution is worse (Fig. 5).

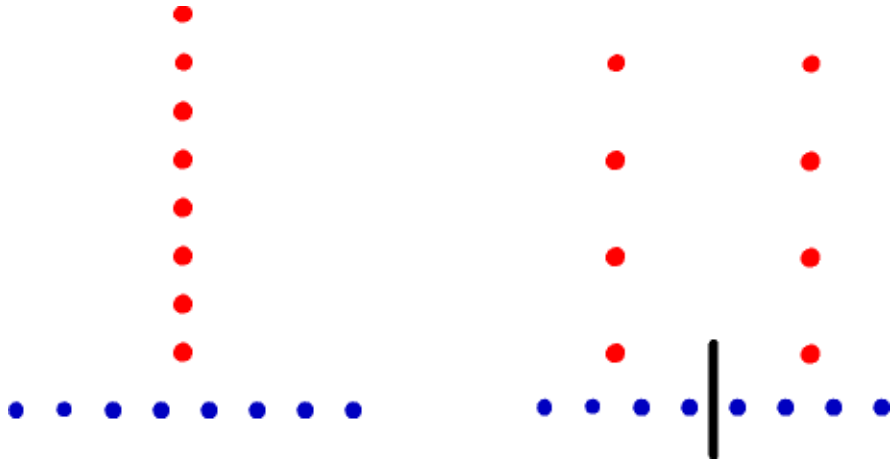

Figure 5: Left: spectrum (red) of 8 points of time domain data (blue). Right: time domain data has been divided up into two bins of 4 points each. Spectrum of each time bin is shown, yielding higher resolution in time, but lower resolution in frequency.

We cannot achieve arbitrarily high resolution in both time and frequency due to the uncertainty principle. If we denote time resolution by Δt and frequency resolution by Δf, the best we can achieve is $\Delta t \Delta f = 1$. Instead, we can settle for producing several TF-planes, with successively higher time resolution and successively lower frequency resolution in each. The procedure described in the LIGO Scientific Collaboration Algorithm Library (LAL) has been followed and adapted [9]. The power contained within a single time and frequency bin (a single pixel) would be described by

$$P = (\text{Re}(\tilde{Y}_{jJ\Sigma}))^2 + (\text{Im}(\tilde{Y}_{jJ\Sigma}))^2, \tag{9}$$

where j denotes the TF-plane, J denotes the time index, and Σ denotes the frequency index.

The excess power method works by summing up the power within the detector output and comparing that value with the expected power from noise. As previously stated, if power from the *entire* output were summed, then a signal could be easily lost in the noise power. Ideally, power from just the times and frequencies at which the signal existed could be summed, and the rest of the noise discarded, allowing the highest chance of detection. However, because the duration and bandwidth of the signal are unknown, we must search over all possible *TF-tiles*, where a TF-tile is a subset of a TF-plane and is specified by a start time, end time, start frequency, and end frequency. Searching over all TF-tiles is equivalent to placing a rectangular window on a TF-plane and only looking at the power within that window, and then allowing that window to change its location and dimensions. The total power in each TF-tile is summed:

$$P = \sum_{J=J_1,\Sigma=\Sigma_1}^{J=J_2,\Sigma=\Sigma_2} \left[(\text{Re}(\tilde{Y}_{jJ\Sigma}))^2 + (\text{Im}(\tilde{Y}_{jJ\Sigma}))^2 \right] \tag{10}$$

Chi-Square Distribution

In the presence of white Gaussian noise, the power follows a χ^2 distribution. A χ^2 distributed variable is the sum of squares of independent, zero mean, unit variance, normally distributed variables. The number of variables being summed is called the number of degrees of freedom, ν, which is equal to the expected value of the χ^2 variable. In a TF-plane, each pixel contributes two degrees of freedom: one each for the real and imaginary parts of the FFT.

The χ^2 distribution is useful because its probability distribution function (PDF) and cumulative distribution function (CDF) are analytic.

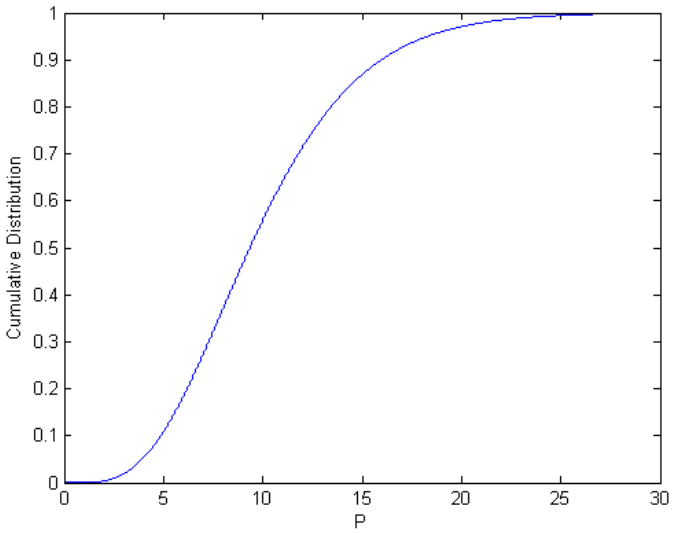

Figure 6: Cumulative Distribution Function for a χ^2 distribution with $\nu = 10$.

Figure 6 shows an example CDF for a χ^2 distribution with $\nu = 10$. At $P = 20$, the value of the CDF is $\sim$.97, meaning that 97% of observed values of P can be expected to be < 20. Then 3% of observed values can be expected to be > 20. One could also say that the *probability* of observing $P > 20$ due to noise alone is .03.

TF-Tile Power

For each TF-tile, the probability α of obtaining the summed power due to white Gaussian noise alone can be computed using the χ^2 CDF. A lower α means it is less probable that the total summed power is due to noise alone. The number of degrees of freedom is twice the number of pixels in each tile. One way of deciding if a signal is present is to simply compare the probability level α against a threshold level, α_{thresh}. If $\alpha < \alpha_{\text{thresh}}$, then a candidate event (detection) is marked. The threshold is set to achieve a desired false alarm rate; e.g., $\alpha_{\text{thresh}} = 10^{-6}$ yields one false detection per one million windows checked, on average.

Looping through TF-Tiles

To save computational time, we calculate the power only for a fraction of all possible TF-tiles. Some subset of all possible tiles would be presumed satisfactory coverage for detecting a signal. A quadruple loop over the time width, time start, frequency width, and frequency start determines the tile. The algorithm for determining which tiles are selected is as follows (the reasoning will be shown for the time dimension; it is identical for frequency):

Let J_{start} represent the starting time index and J_{width} represent the width index of the tile. The increment for J_{start} is

$$\Delta J_{\text{start}} = 1 + \left[\frac{J_{\text{width}}}{N_{\text{ov}}} \right], \tag{11}$$

where brackets represents rounding down to the nearest integer, and N_{ov} is an overlap constant. For constant J_{width} but shifting J_{start}, this provides a fractional overlap of tiles $\sim (N_{\text{ov}} - 1)/N_{\text{ov}}$. The worst case scenario for a signal of a given width occurs when its starting time lies just between the starting times of the surrounding tiles, causing neither tile to fully surround the signal (Fig. 7). With this algorithm, the worst case fractional loss of signal power is $\sim 1/2N_{\text{ov}}$.

The increment for J_{width} is exactly the same:

$$\Delta J_{\text{width}} = 1 + \left[\frac{J_{\text{width}}}{N_{\text{ov}}} \right]. \tag{12}$$

A tile with the same J_{start} but larger J_{width} covers a larger area than the previous tile. The worst case scenario would be a signal that has a time width that lies just in between the two tiles, which would cause a large loss of signal power in the smaller tile as well as a large excess of noise power in the larger tile (Fig. 7). The worst case fractional loss of signal power / gain of noise power is $\sim 1/2N_{\text{ov}}$.

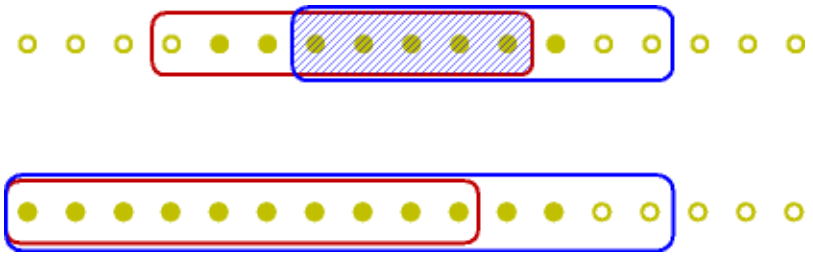

Figure 7: Top: shifting start position for $J_{\text{width}} = 7$, $N_{\text{ov}} = 3$, corresponding to a tile width of 8 points and a fractional overlap of about $2/3$. Overlap area is shaded – in this case, the overlap is 5/8. Bottom: Same start position, but different widths, with $N_{\text{ov}} = 3$. Filled circles indicate a signal located in the worst-case-scenario position.

Results of a Search

After all tiles have been searched, the excess power algorithm has produced a list of candidate events and their associated tiles. A strong signal might produce many candidate events.

For visualizing the results, it would be useful to display all the tiles on one graph, so that areas of high concentration of candidate events are noticeable. A 'fill plot' places the location of each tile on a time-frequency plot in a light color. Areas with overlapping tiles become darker, whereas areas with only one or two tiles are light-colored (Fig. 8).

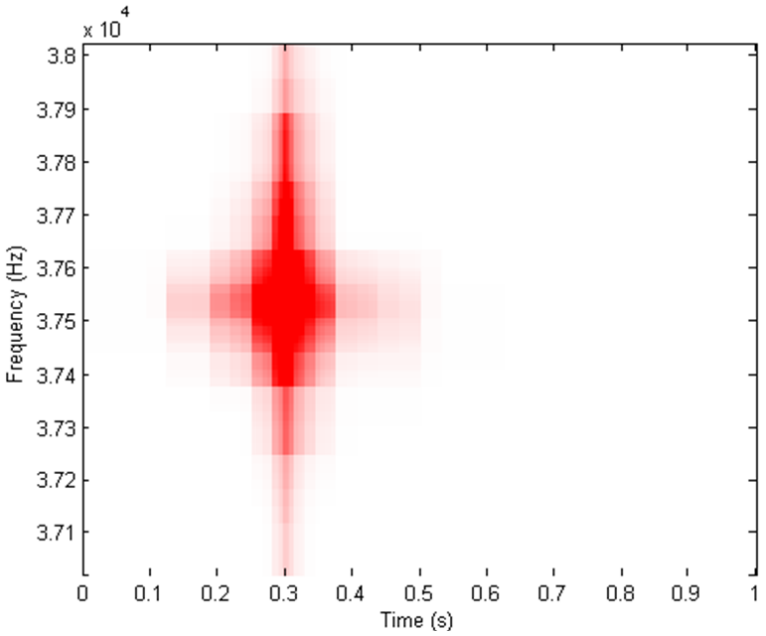

Figure 8: A fill plot of candidate events. Areas with overlapping tiles are displayed in a darker color.

Signal Estimation

As a short and simple method for estimating the original signal, we can take a TF-tile that has a candidate event and produce a time series from it. Zeroing all points on the TF-plane except for those within the TF-tile eliminates most of the noise power while leaving most of the signal power. The frequency domain data can then be inverse Fourier transformed to produce a signal estimate (Fig. 9). The TF-tile selected to estimate the signal could be based on criteria such as the lowest probability level or the largest time-frequency area. Or, we could take an average of the estimated signals from every TF-tile with a candidate event.

Power at the FSR frequency

One analysis that can be performed is to look at the detector output over stretches of time and check if it matches with trends in power corresponding to the movement of a source across the sky. During a single day, the location of a distant source such as the center of our galaxy will remain relatively constant. Due to the earth's rotation, the galactic center will appear to move in the local coordinate system of a point on the surface of the earth. As the source moves

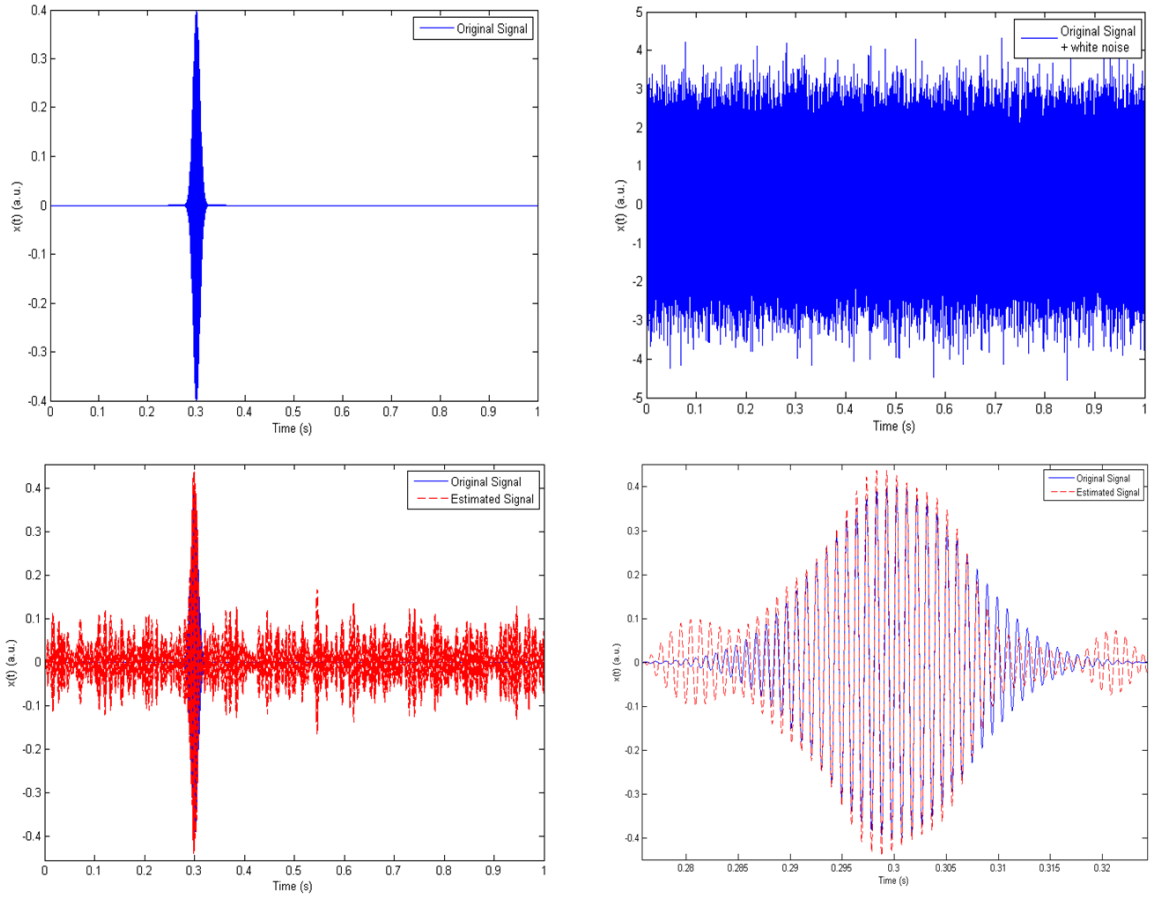

Figure 9: Top left: original signal, a sine-Gaussian. Top right: white noise has been added to the original signal. Bottom left: the estimated signal (red) has been plotted atop the original signal (blue). Noise artifacts are still present. Bottom right: closeup of the estimated signal and the original signal.

through LIGO's sky, it moves across the antenna pattern, through areas of higher and lower sensitivity. One could correlate trends in the sensitivity with trends in the signal. LIGO's antenna patterns at DC and at the FSR frequency for plus, cross, and average polarizations along with an overlay of the galactic center's trajectory is shown in Fig. 10. The excess power algorithm could be used as well to track the number of 'events' that occur during time periods of higher sensitivity.

Conclusion

An analysis pipeline has been developed to begin the search for gravitational waves at the FSR frequency. Its two main components are the cavity response and the excess power algorithm. The cavity response can have a dramatic effect on some signals near the FSR frequency, altering their waveform. A sine-Gaussian distorts to have a ringdown time of

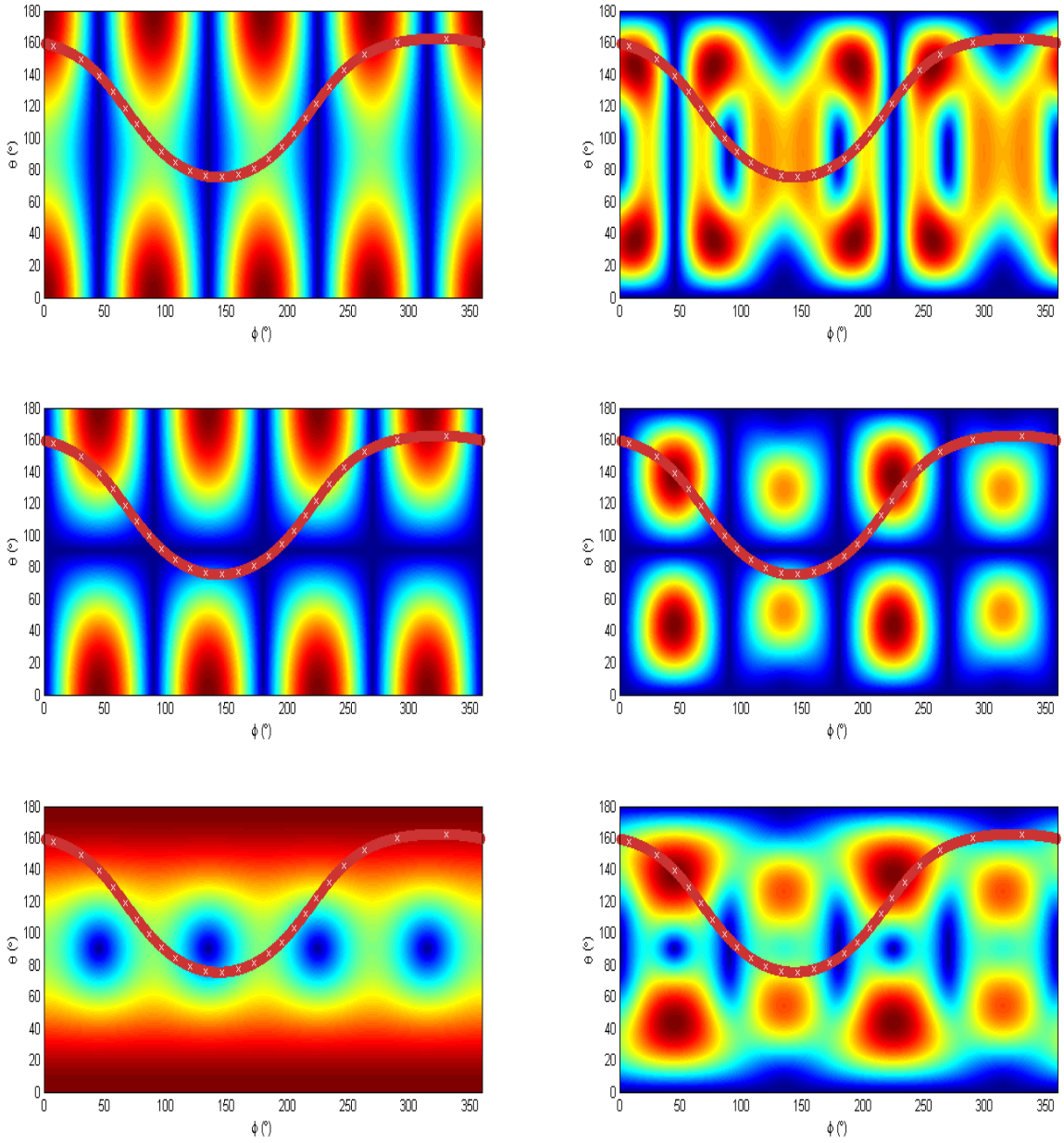

Figure 10: Antenna pattern plots, where red represents high sensitivity and blue represents low. Top left: DC, plus polarization. Top Right: FSR frequency, plus polarization. Middle left: DC, cross polarization. Middle right: FSR frequency, cross polarization. Bottom left: DC, average polarization. Bottom right: FSR frequency, average polarization. The galactic center's path through the sky over one day is plotted, with an x marked at each hour interval.

about 2 ms. The excess power algorithm can detect signals in the presence of white noise. While developed, the algorithm not been tested on a multitude of cases. Future work should

include Monte Carlo simulations of injected signals at various sky locations to determine the success rate of the algorithm. Detailed analysis could be done correlating an astrophysical object's motion through the sky with the strength of a signal, as described with the galactic center example above. Tracking trends over long periods of time may be the only way to detect signals of low amplitude at the FSR frequency. Gravitational waves may exist at the FSR frequency. Further development of the analysis pipeline begun here should ensure that LIGO's sensitivity at the FSR frequency is taken advantage of.

References

[1] D. Sigg and R. L. Savage, Analysis Proposal to Search for Gravitational Waves at Multiples of the LIGO Arm Cavity Free-Spectral-Range Frequency, LIGO Document T030296-00, 2003.

[2] A. C. Melissinos, M. F. Bocko, and S. Giampanis, Search for a High Frequency Stochastic Background of Gravitational Waves, LIGO Document, 2004.

[3] C. Clarkson and R. Maartens, General Relativity and Gravitation **37**, 1681 (2005).

[4] M. Rakhmanov, R. L. Savage, D. H. Reitze, and D. B. Tanner, Phys. Let. A **305**, 239 (2002).

[5] M. Rakhmanov, Response of LIGO to Gravitational Waves at High Frequencies and the vicinity of FSR (37.5 kHz), LIGO Internal Working Note, In Progress, 2005.

[6] H. Elliot, Analysis of the Frequency Dependence of LIGO Directional Sensitivity (Antenna Pattern) and Implications for Detector Calibration, LIGO Document T050136-00-W, 2005.

[7] LIGO Scientific Collaboration, Upper limits of gravitational wave bursts in LIGO's second science run, LIGO Document P040040-07-R, 2005.

[8] W. G. Anderson, P. R. Brady, J. D. E. Creighton, and E. E. Flanagan, Phys. Rev. D **63**, 042003 (2001).

[9] LAL Software Documentation.