LIGO-SURF Summer 2006 Project Proposal for

"Searching for Gravitational-Wave Bursts of Arbitrary Waveform"

Research Mentor: Dr. Shourov K. Chatterji

Primary Researcher: Rubab Khan, Columbia University

Abstract:

One class of signal LIGO is searching for consists of short duration gravitational wave bursts of a priori unknown waveform. Potential sources include core collapse supernovae and the coalescence of binary black holes. To detect such events, existing search algorithms project the LIGO data stream onto various time-frequency bases and then search for regions of excess signal energy. One of these search algorithms, the Q Pipeline, determines the statistical significance of events based solely on the peak signal observed in the time-frequency plane. This project will investigate extensions to this approach that also consider the statistical significance of arbitrarily shaped regions in the time-frequency plane. Such approaches offer the prospect of improved performance for a variety of sources of both known and unknown waveform.

Background:

Einstein's General Theory of Relativity describes gravity as a space-time curvature due to the presence of mass-energy. One prediction of GR is that as concentrations of mass-energy rapidly changes shape (i.e. supernovae explosion, merger of astronomical binary systems, star-quake etc.) they create dynamically changing space-time warpage or ripples in spacetime that propagates throughout the universe at the speed of light [1,2]. When they reach Earth, gravitational waves are extremely weak perturbations of local flat spacetime, and the detection of these elusive waves to gain knowledge of their sources becomes a tremendous challenge. Unlike electromagnetic waves which propagate through spacetime after being created by the incoherent motion of atoms and molecules and have a wavelength much smaller than their sources, gravitational waves are propagated as perturbation of spacetime itself after being created by massive astronomical sources and have wavelengths similar to the size of their sources. Most significantly, while electromagnetic waves interact with most objects, the universe is mostly transparent gravitational waves. This is both a blessing and a curse since it ensures that the gravitational waves that reach us have not been meddled with since they were created, while this also makes their direct observation more difficult [3]. While the existence of gravitational waves have been observationally confirmed through the discovery of binary stars' (PSR 1913+16) spiraling together just at the rate predicted by GR due energy loss through gravitational radiation [2], efforts to directly detect gravitational waves are yet to achieve decisive success. While the very first direct detection of gravitational waves will be a very significant, the payoff will come when analyzing the detected waves from which it will be possible to extract information about the physics of the extreme conditions where the waves had originated – such as very strong gravity and nuclear densities – information that is not easily accessible or totally inaccessible from electromagnetic observations.

Current efforts of building gravitational wave observatories focus on using interferometry to make extremely precise observations of the distance between two sets of test masses. A global network of interferometric detectors is currently at various stages of data-collection, commissioning, construction, or planning phase. They include the Laser Interferometer Gravitational-wave Observatory (LIGO; three detectors; two in Hanford, Washington and one in Livingstone, Louisiana), Virgo (Pisa, Italy), GEO600 (Hanover, Germany), TAMA300 (Tokyo, Japan), and ACIGA (Perth, Australia) [3]. Each LIGO detector (or every interferometric detector) is basically an ultra-sensitive giant Michelson interferometer having two arms in L-shaped structure. Naively, it can be imagined that as a gravitational wave passes through stretching one arm and squeezing the other, and then reversing this effect periodically, LIGO measures this change of length and thus convey information about the passing waves. If waves of similar nature are detected at multiple detectors nearly simultaneously, it further enhances the level of confidence about the detection and value of the information collected. Moreover, the time difference (<10ms) between detection at two sites give the LIGO detectors directional sensitivity that helps to better determine exactly from which direction of sky a wave is coming in from. The LIGO detectors have now reached their design sensitivity as a result of relentless efforts on the part of numerous engineers and scientists pushing the limits of technology, and currently LIGO is collecting data as it is undergoing its yearlong science run.

It is expected that observation of signals at LIGO will occur near the limit of detector sensitivity, and searching for and identifying such small signals in the presence of various detector noise is a daunting task [4]. Depending on what algorithm we use to search for gravitational waves in data collected at detectors, sources of gravitational waves are classified into four major groups. The inspiraling of a star into its compact binary partner (i.e. neutron star or black hole) causes chirp signals. This process is sufficiently well understood and there are special tools to search for gravitational waves produced by them. Signals from very short-lived events of which we do not have sufficient understanding to predict an expected waveform, such as merger of binary objects, core collapse supernovae, gamma ray bursts, and even unexpected sources, are classified as Burst Sources. Stochastic signals can be either relic gravitational-waves from the very early universe or the cumulative effect of many low amplitude sources that can give rise to a correlated random noise in multiple LIGO detectors as co-incident events. Periodic signals from spinning compact objects (pulsars) can produce a signal if they have asymmetric shape or mass distribution [5].

If the waveform of GW burst is known, then matched filtering can be used which first whitens the data under test by a filter whose magnitude response is the inverse of the detector noise spectrum, next forms the projection of the data onto the waveforms that are to be detected, and then looks for times when the projection is large. However, for bursts of unmodeled waveform, the data under test are typically projected onto a convenient basis of abstract waveforms that are chosen to cover a targeted region of signal space, and one looks for large projections. These searches can be time-domain searches in which the primary basis consists of delta functions in time, and time-frequency searches in which the typical basis consists of windowed complex exponentials or wavelets [6]. The focus of this project is the expansion of the gravitational wave burst search algorithm titled Q-pipeline that looks for unmodeled bursts. The Q-pipeline is a comprehensive end-to-end analysis pipeline for the detection of gravitational-wave bursts in data from a single interferometric detector. It consists of whitening by zero-phase linear prediction, application of the discrete Q transform, thresholding on the white noise significance of Q transform coefficients, identification of the most significant set of non-overlapping time-frequency tiles, and a final stage that excludes all but the most significant time-frequency tile within a specified time window in order to prevent the redundant reporting of candidate events [7]. The analysis tool of Q-pipeline is the Q-transform which is a modification of the standard short time Fourier transform in which the analysis window duration varies inversely with frequency such that the time frequency plane is covered by tiles of constant “Q” [7] which can be naively interpreted as a dimensionless quality factor for bursts which is the ratio of the center frequency to the bandwidth of a burst [6].

Objectives:

As the Q-pipeline projects the data into small regions of the time-frequency plane, a Heisenberg like uncertainty relation applies and bursts cannot have both a well-defined frequency and a well-defined time. The minimum uncertainty signal is a "sine-Gaussian", that is a sinusoid with a Gaussian amplitude envelope:
h(t) = exp(-(t - t0)^2 / (4 sigmat^2)) * sin(2 * pi * f * t). The algorithm is therefore optimal for signals that have this waveform. For signals that are less localized in the time frequency plane, their detectability is currently determined by their maximum projection onto the space of sinusoidal Gaussians. Q-pipeline's treatment has been somewhat limited for bursts that are poorly localized in the time-frequency plane. In particular, it only considers the statistical significance of single most significant tile with minimum time-frequency uncertainty. In searching for statistically significant events, methods of clustering the measurements from neighboring or overlapping basis functions to more optimally detect signals that are not well represented by the particular choice of basis can be employed [6]. An improvement in the detectability of poorly localized bursts should be possible if it would consider the combined statistical significance of clusters of time-frequency tiles by clustering together projections that are nearby in time and frequency[8]. In addition, as described in [7], when evaluating the statistical significance of clusters of time-frequency tiles or testing for time-frequency coincidence between detectors, a more accurate treatment of the overlap between time-frequency tiles can be obtained by applying the mismatch formalism [8]. This should significantly improve the detectability of bursts with arbitrary waveform.

This project will investigate extensions to Q-pipeline that would consider the statistical significance of arbitrarily shaped regions in the time-frequency plane by utilizing the advantages of clustering algorithms, and thus would significantly improve detectability of bursts with arbitrary waveform that are less localized in time frequency plane.

Approach:

There has been some work done already on clustering algorithms in the past, and the initial stage of this project will be to identify among them an appropriately promising one, which can be utilized as an extension to Q-pipeline. Some algorithms that should be considered can include:

1. Windowed Clustering: The simplest approach would be to slide a "window" of duration T and bandwidth F over the mosaics and record whenever the total significance inside the window exceeds a given threshold. This is of course sensitive to the choice of T and F. It may not be as powerful as some other approaches, but it is conceptually and computationally very simple.

2. TFClusters: This is one of the first algorithms that were applied to search for gravitational-wave bursts in LIGO data. It is similar to the Q Pipeline in that it identifies regions of statistically significant excess signal energy in the time-frequency plane. However, instead of attempting to test multiple time-frequency resolutions, it clusters the results from a single time-frequency resolution. The clustering algorithm is well documented [9,10] and it may be possible to adapt to the Q Pipeline.

3. Graph Analysis: This is a method developed to identify "glitches" in the auxiliary and environmental data channels from LIGO, but the method may be very applicable to searching for gravitational waves as well [11].

4. Hierarchical Clustering: TFClusters and Graph Analysis fall under the class of hierarchical clustering algorithms. Matlab also provides a heirarchical clustering algorithm that may prove useful.

5. TrackSearch: This algorithm is specifically meant for gravitational-wave bursts that produce ridge-like features in the time-frequency plane. It is currently under development and will be considered if time allows.

6. Density Based Algorithm: If time allows, we will also consider algorithms that cluster based on the density of nearby significant tiles[12].

At first, there shall be a short list of clustering approaches to try out as possibilities. Next, we should write some simple Matlab scripts to experiment with these algorithms, and select a method that seems promising and realistically simple to implement correctly. Then it has to be added as an extension to the Q-pipeline so that we may test its performance on a variety of simulated bursts, and compare it to the performance of the Q-pipeline without clustering. We should then repeat this last step for real LIGO data collected during S5 and compare expanded Q-pipeline's efficiency in identifying-candidate events with that of the Q-pipeline without clustering.

Time-line:

Weeks	*Task*
<1	Prior to project start: i) Familiarize with Q-pipeline algorithm and codes. ii) Familiarize with various clustering methods. iii) Create a shortlist of clustering algorithms to try.
1, 2	i) Write simple Matlab scripts to experiment with some of the short listed algorithms. ii) Evaluate their relative performance and choose one algorithms to implement as the Q-pipeline extension.
3, 4, 5	Add the chosen algorithm as an extension to the Q-pipeline.
6, 7	Test the performance of the extended Q-pipeline on a variety of simulated bursts, and compare it to the performance of the Q-pipeline without clustering.
8, 9	Test the performance of the extended Q-pipeline on real LIGO data collected during S5, and compare its efficiency in identifying candidate-events with that of the Q-pipeline without clustering.
10	i) Document the work and its results. ii) Prepare the presentation.

References:

[1] D. Sigg, “gravitational waves”, LIGO-P980007-00-D, 1998.

[2] K. S. Thorne, “gravitational waves”, arXiv:gr-qc/9506086 v1, 1995.

[3] S. A. Hughes et al. “New physics and astronomy with the new gravitational wave observatories”, arXiv:astro-ph/0110349 v2, 2001.

[4] S. K. Chatterji, “Chapter 1: Introduction”, emvogil-3.mit.edu/~shourov/thesis/chapter1.pdf, 2005.

[5] B. C. Barish and R. Weiss, “LIGO and the Detection of gravitational waves”, LIGO-P99039-00-R, 1999.

[6] S. K. Chatterji, “Chapter 3: Burst Detection”, emvogil-3.mit.edu/~shourov/thesis/chapter3.pdf, 2005.

[7] S. K. Chatterji, “Chapter 5: The Q Transform”, emvogil-3.mit.edu/~shourov/thesis/chapter5.pdf, 2005.

[8] S. K. Chatterji, “Chapter 8: Conclusion”, emvogil-3.mit.edu/~shourov/thesis/chapter8.pdf, 2005.

[9] J. Sylvestre, “Time-frequency detection algorithm for gravitational wave bursts”, arXiv:gr-qc/0210043, 2002.

[10] J. Sylvestre, “Upper Limits for Galactic Transient Sources of Gravitational Radiation from LIGO First Observations”, LIGO-P020007-00-R, 2002.

[11] H. Bantilan, “Graph Analysis”, virgo.physics.carleton.edu/Hans/index.html, 2005.

[12] M. Ester, “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise”, iucaa.ernet.in/~gupchup/DBSCAN.pdf, 1996.