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

Technical Note **LIGO-T1000148** 2010/03/17

Generalized Flare Pipeline Detection Statistic

Peter Kalmus

Final Version

This is an internal working note
of the LIGO Project

California Institute of Technology
LIGO Project - MS 51-33
Phone (626) 395-2129
Fax (626) 304-9834
E-mail: info@ligo.caltech.edu

Massachusetts Institute of Technology
LIGO Project - MS 20B-145
Phone (617) 253-4824
Fax (617) 253-7014
E-mail: info@ligo.mit.edu

http://www.ligo.caltech.edu

File printed: 2010/05/29

Abstract

We describe the generalized Flare pipeline detection statistic used in the S5y2/VSR1/A5 magnetar search.

I. INTRODUCTION

For the purposes of this document, a "detection statistic" is a method for reducing gravitational wave (GW) data from one or more GW detectors into "analysis events" which may be used to quantify the probability that a GW signal exists in the noisy data. Flare pipeline analysis events are each ascribed a loudness (the principle quantity derived by the detection statistic), a central frequency and bandwidth, and a central time and duration. As each analysis event is a cluster of pixels in a time-frequency map, each analysis event can also be considered a "picture" in the time-frequency plane.

The original Flare pipeline detection statistic is described in [3], for one or two detectors. The one-detector statistic has not changed. Section II describes a slightly modified general statistic used in the S5y2/VSR1/A5 magnetar search for two or more detectors. The generalized statistic allows detectors with arbitrary alignment and sensitivity to be combined effectively.

II. GENERALIZED N-DETECTOR STATISTIC

N time-frequency spectrograms are created from conditioned data for N individual detectors from a series of Blackman-windowed discrete Fourier transforms, of time length δt set by the target signal duration. For details on the data conditioning procedure, see [3]. A *tile* is an estimate of the short-time Fourier transform of the data at a specific time and frequency. Each column in the tiling corresponds to a time bin of width δt and each row corresponds to a frequency bin of width δf, both linearly spaced, with $\delta f \delta t = 1$. Adjacent time bins overlap by $0.9\delta t$ to guard against mismatch between prospective signals and tiling time bins. Larger overlaps require more computation and do not noticeably improve sensitivity (see [3]).

We then have N complex-valued time-frequency tilings, one for each detector, from which we calculate the real-valued one-sided PSD for every time bin:

$$P^{\mathrm{D}}_{tf} = \mathrm{Re}\left[T^{\mathrm{D}}_{tf} T^{\mathrm{D}^*}_{tf} e^{-i2\pi f \Delta t}\right]. \quad (1)$$

Here T represents a tiling matrix and t and f are time and frequency bin indices, D denotes the detector, and Δt is the gravitational wave crossing time difference between the detector and the geocenter. The Δt term takes care of applying the appropriate time difference between detector data streams in the Fourier space, with the advantage of permitting sub-sample time delays, which significantly increases the sensitivity at higher frequencies. The result is normalized to account for sampling frequency and windowing function. We discard frequency bins outside of the chosen search band.

We then determine a scalar weighting for each detector matrix by determining the rms antenna factor for each detector at the time of the center of the signal region. We also determine a frequency-dependent weighting by taking the inverse square of the PSD of each individual detector. These weightings are multiplied together and applied to the detector tilings. Then the N weighted detector tilings are added together.

The rest of the procedure is the same as in [1] and [2]. We next use off-source data to remove the background noise power from each element of the PSD time-frequency tiling. The elements are fit to a gamma distribution, and outliers above a threshold (typically four standard deviations) are discarded. This process repeats until no outliers remain.

The resulting estimate on the mean is subtracted from each element of the corresponding frequency bin in the PSD matrix, giving a matrix of excess power. We also normalize each frequency bin element in the excess power matrix by the resulting estimate on the standard deviation.

We then apply a density-based clustering algorithm [4] which allows retention of signal energy which might otherwise be fragmented in the case of extended signals in the time-frequency plane. Only the brightest pixels are sent to the clustering algorithm; the number of pixels is a fixed percentage of the pixels in the tiling; the percentage of pixels is tuned on injections in the background and depends on the target injection waveform. The analysis events correspond to discrete clusters found by the algorithm, and include information on cluster central frequency, central time, bandwidth, duration, and so forth. The statistic in this case is the sum over the cluster of tile significance.

This statistic reduces to the one-detector statistic in the case where $N = 1$.

---

[1] B. Abbott et al. Search for gravitational-wave bursts from soft gamma repeaters. *Phys. Rev. Lett.*, 101(21):211102, 2008.
[2] B. P. Abbott et al. Stacked Search for Gravitational Waves from the 2006 SGR 1900+14 Storm. *ApJ Lett.*, 701:L68–L74, August 2009.
[3] P. Kalmus. *Search for Gravitational Wave Bursts from Soft Gamma Repeaters*. PhD thesis, Columbia University, 2008. arXiv:0904.4394.
[4] R. Khan and S. Chatterji. Enhancing the capabilities of LIGO time-frequency plane searches through clustering. *in progress*, 2007.