A Statistical Search for Small-Strain Burst Sources Proliferating in LIGO Time Series Data

Brett Bochner[1], Baokun Gu[2], Stanley Wong[3], Yuanyi Yin[2]

1. Dept. of Physics and Astronomy, Hofstra University, NY, USA; 2. Manhasset High School; 3. Hunter College High School

Introduction

With the profound discovery of a handful of strong-signal Binary Black Hole (BBH) coalescences, these objects are now verified as a primary LIGO gravitational wave (GW) source. These few spectacular events must therefore be the tip of a very large iceberg of a multitude of BBH coalescences which are too weak to be individually identified and detected. As we move from the "big discovery" phase of GW Astronomy to the business of detection productivity, it is crucial to scour the time series data for all recoverable signs of real GW events. In this student-led research project, an algorithm is designed and executed to search for excess coincidence between interferometers due to short-period bursts hiding in the LIGO Online Science Center (LOSC) database. Our algorithm & results so far are presented.

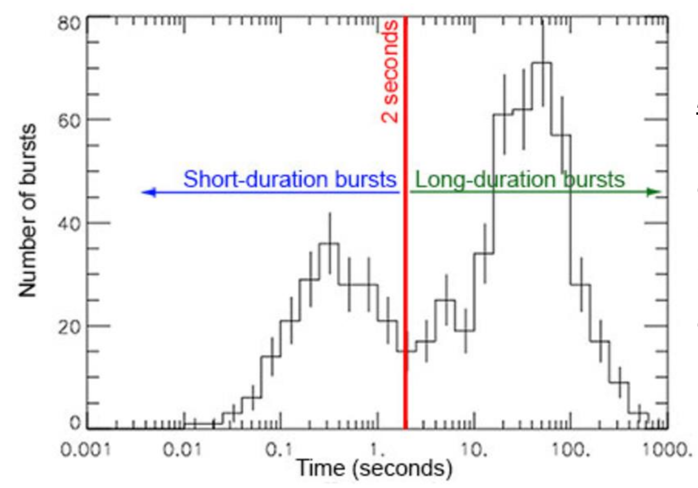


Figure 1: Typical gamma ray burst (GRB) durations detected by the Compton Gamma-ray Telescope[1], with "short bursts" defined as $\Delta t \lesssim 2$ sec.

Objectives and Methods

- Search for statistical evidence of numerous, low GW-amplitude, short-duration burst signals in the LIGO S6 Science run, without requiring identification of individual signals or their waveforms
  - ➢ Such signals may result from various binary system mergers, gamma ray bursts, cosmic string cusps, or unknown sources.[2]
- Only the cleanest stretches of LIGO time series data are used:
  - ➢ There must be coincident data between Hanford (H1) and Livingston (L1), with no artificial injections
  - ➢ Data must pass the most stringent burst flag (BURST_CAT4H)
  - ➢ Only blocks of continuous data with $\Delta t \geq 96$ sec are used (w/16-sec buffers at beginning & end, so $\Delta t_{search} \geq 64$ sec)
  - ➢ An example of time series data passing all flags (for H1, before coincidence w/L1) is in Fig. 2 (GPS Time = 957984768)

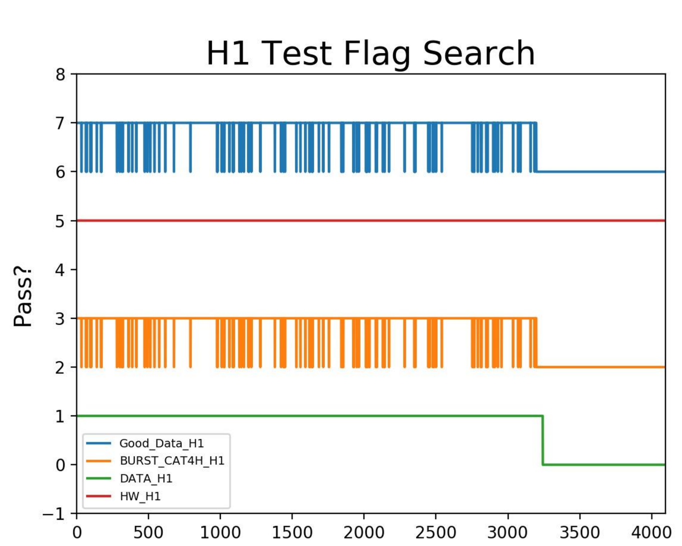


Figure 2: Binary flag data for one time series file (flag value hi = pass, lo = fail), with the positive intersection of all 3 tests in blue. The acceptable H1 data and L1 data must then overlap in $\Delta t \geq 96$ sec durations.

- The acceptable stretches of data within each 4096-second LOSC file are divided into 2-second blocks (each block overlapping 50% with the previous time block, i.e., 1 sec overlap), with the data correlated between the H1 and L1 data for each block

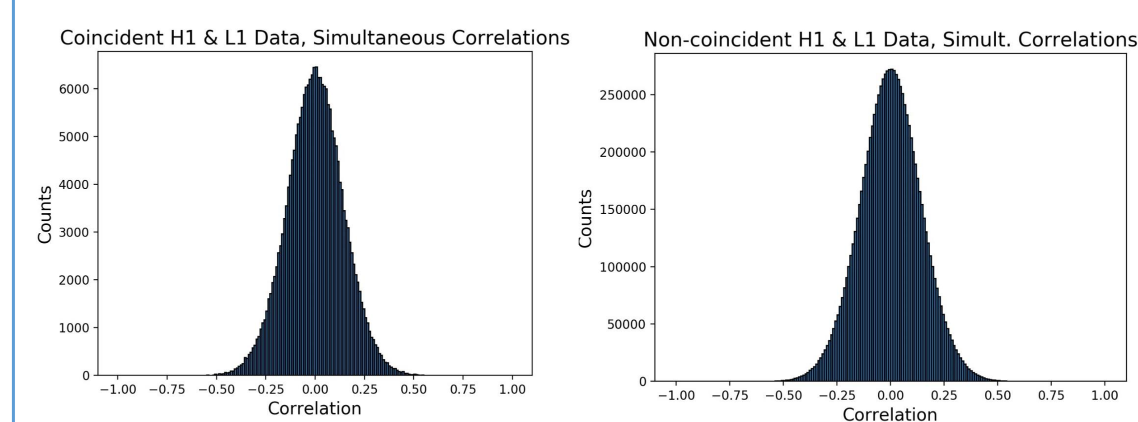
Algorithm Details

- The accepted H1 and L1 data blocks are filtered with a 10th-order Butterworth band-pass filter between 100 Hz and 800 Hz
  - ➢ Dominant low-frequency noise sources (especially Seismic and Thermal Noise) are strongly suppressed
  - ➢ Low-frequency instrumental and environmental lines are partly suppressed (particularly the 60 Hz background)
  - ➢ Filtering this range results in a flatter noise spectrum much closer to white noise, resulting in more Gaussian backgrounds
- The filtered ($\Delta t_{search} \geq 64$ sec) H1 and L1 data ranges are divided into 2-second blocks (i.e., 8192 data points at the 4096 Hz LOSC sampling rate), used for short-burst searches ($\Delta t_{burst} \lesssim 2$ sec), and H1 vs. L1 data are cross-correlated in 4 different ways:
  - ➢ ***Coincident*** correlations are done using H1 and L1 data blocks at the *same time*, for which positive correlations may be due to real Gravitational Waves (GW's) acting on both detectors
  - ➢ ***Non-coincident*** correlations are done using 2-second blocks of H1 and L1 data in the same GPS-time file, but ***not*** at the same time ($\Delta t_{offset} \geq 2$ sec), thus correlations are not due to GW's, but due to noise only; used as a "control" for searches
    - ❖ For both *Coincident* & *Non-coincident* correlations, we do:
  - ➢ ***Simultaneous*** correlations, done using H1 and L1 data blocks centered at the *exact same time*, which leaves the background of "false GW correlations" due to noise as mostly Gaussian – with equal likelihood of "positive" and "negative' correlations – but does not factor in the likely earlier arrival time of a GW at H1 or L1, due to various arrival directions
  - ➢ ***Maximized*** correlations, which factors in the possible *earlier arrival* of a GW at H1 or L1, due to the light transit time between the detectors ($\Delta t \lesssim 10$ ms, for various directions)
    - ❖ For 4096 Hz sampling rate, 10 ms ≈ 41 data points
    - ❖ Shifting the L1 time series data (relative to H1) back & forth by up to ± 41 data points, 83 different correlation values are computed; the *max positive value* is saved, and taken as the correlation best indicating a hidden burst GW
    - ❖ While this helps optimize the search for GW burst signals (hopefully maximizing their contributions to the total correlation values), as a side-effect this also biases the *noise* to produce excess positive (over negative) correlations, *even* in the Non-coincident case
- All 4 correlation analyses are done for each accepted 2-second block in a particular GPS data file, with blocks defined with 50% overlap – the "next" block starts 1 second after the last block *starts* – thus if a particular GPS-time file contains N seconds of accepted data, this results in N Coincident correlations, and (slightly less than) N^2 Non-coincident correlations. (Both done Simultaneous and Maximized.)
- The current goal is to complete this analysis for all time series files in LIGO Science Run S6; and then to broaden the analysis to triple correlations (L1, H1, & H2) for LIGO Science Run S5

Results

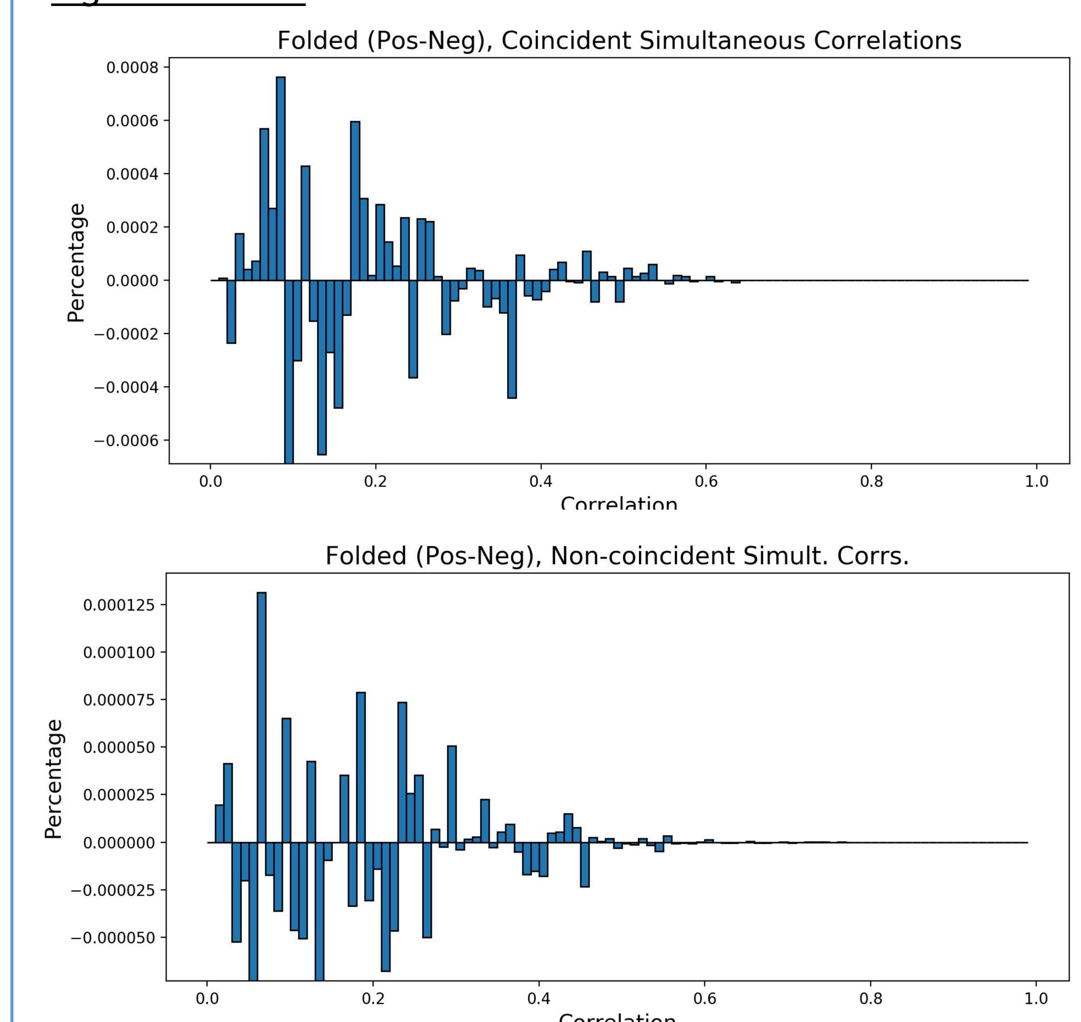
- To date, nearly **20%** of the LOSC S6 data has been analyzed
  - ➢ Approx. **25%** of the time series files contain usable data that passes all of our flags to produce correlation output results.
  - ➢ The aggregate correlation data is combined for analysis here:

Figures 3 and 4:



- The Non-coincident (Simultaneous) correlations are strongly Gaussian and positive/negative symmetric, as expected from false GW "signals" due to close-to-white noise
- The Coincident (Simultaneous) correlations show small deviations from Gaussianity, due either to larger relative fluctuations in smaller data set, to the presence of real GW burst events, or to a combination of both signal and noise

Figures 5 and 6:

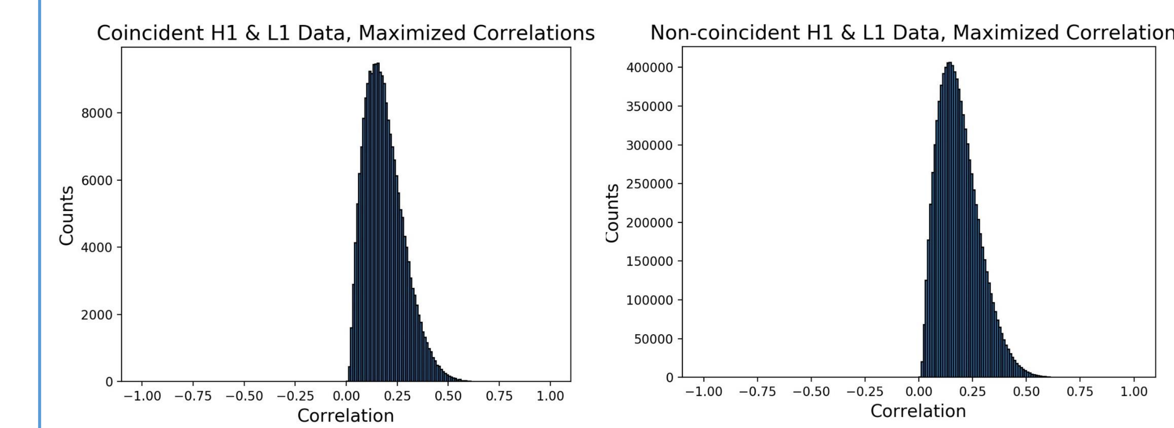


In Figs. 5 & 6 - "Folded" versions of Figs. 3 & 4, with bin values from Negative H1/L1 correlations subtracted from Positive correlation bin values (and normalized by the total # of counts)

- Interpretations of results still limited by statistical fluctuations; will strengthen as remaining S6 data is analyzed
- The Non-coincident (Simultaneous) correlations appear random
- The Coincident (Simultaneous) correlations results may have a slight *excess of positive correlations* in the range $\sim\mathbf{0.2-0.3}$, (and perhaps also in the range $\sim\mathbf{0-0.1}$), counterbalanced by a slight *deficit* of positive correlations (or excess of negative correlations) in the range $\sim\mathbf{0.1-0.2}$
- These are intriguing signs in the Coincidence stats which *may* indicate small burst GW's in the data, but results preliminary

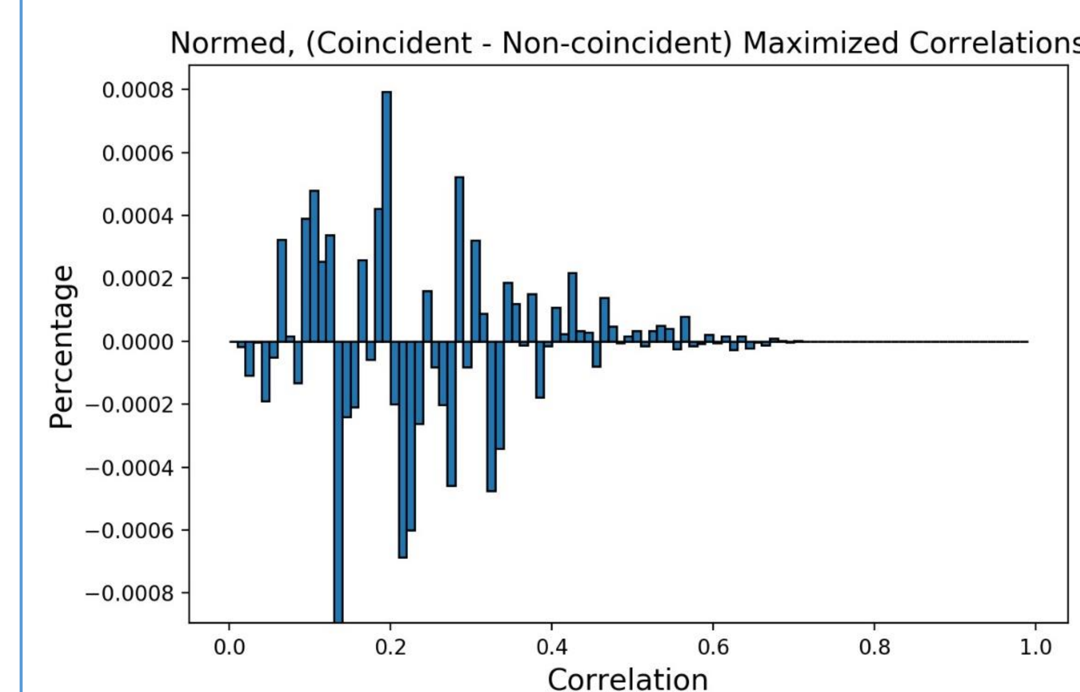
Results *(cont'd.)*

Figures 7 and 8:



- The Maximized correlations, for both the Coincident and (non-GW-signal-bearing) Non-coincident case, show a positive-biased version of the Gaussian distribution, with a tail extending to large correlation values
- Due to the lack of Positive/Negative correlation symmetry, folding the Histograms over the zero-correlation point is not useful here; but we can perform subtraction of the two cases:

Figure 9:



In Fig. 9 - The Maximized correlations, with a subtraction of bin values for the Non-coincident case from the Coincident case (each case normalized for total # of counts before subtraction)

- Comparing the Coincident (vs. Non-coincident) case: There is some evidence apparent of excess positive correlations (or a deficit of negative correlations) in the range $\sim\mathbf{0.35-0.6}$, and the reverse for $\sim\mathbf{0.1-0.35}$
- While intriguing as a sign of small GW bursts in the time series data, no conclusions can be drawn until more data is analyzed

References

❖ All time series data used for this analysis was downloaded from the LIGO Open Science Center (LOSC), at **www.losc.ligo.org**.

1) NASA, Goddard Space Flight Center (2013); https://imagine.gsfc.nasa.gov/science/objects/bursts1.html

2) Chatterji, Shourov K., "The search for grav. wave bursts in data from the second LIGO science run", Ph.D. Thesis, MIT (2005); https://gwic.ligo.org/thesisprize/2006/Chatterji_Thesis.pdf

HOFSTRA UNIVERSITY