

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

Technical Note	LIGO-T1400342-v2	2014/05/23
<b>Recovering S5 Burst Injections</b>		
Alexander E. Cole ( <a href="mailto:aec02011@mymail.pomona.edu">aec02011@mymail.pomona.edu</a> ), Jonah Kanner ( <a href="mailto:jkanner@caltech.edu">jkanner@caltech.edu</a> )		

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](mailto:info@ligo.caltech.edu)

Massachusetts Institute of Technology  
LIGO Project, Room NW22-295  
Cambridge, MA 02139  
Phone (617) 253-4824  
Fax (617) 253-7014  
E-mail: [info@ligo.mit.edu](mailto:info@ligo.mit.edu)

LIGO Hanford Observatory  
Route 10, Mile Marker 2  
Richland, WA 99352  
Phone (509) 372-8106  
Fax (509) 372-8137  
E-mail: [info@ligo.caltech.edu](mailto:info@ligo.caltech.edu)

LIGO Livingston Observatory  
19100 LIGO Lane  
Livingston, LA 70754  
Phone (225) 686-3100  
Fax (225) 686-7189  
E-mail: [info@ligo.caltech.edu](mailto:info@ligo.caltech.edu)

## 1 Abstract

During LIGO’s 5th Science Run (S5), signals simulating various astrophysical phenomena were added to the LIGO detectors for testing and calibration. These simulated signals are known as hardware injections, because they were created by manipulating mirrors in the arms of the interferometers. In this paper, we review the process of employing a matched filter search and detail the results of our adaption of this process to finding hardware injections simulating bursts in S5 data. Of the approximately 60,000 burst injections, there are no remaining unexplained outliers.

## 2 Introduction

LIGO took data for its S5 period from the Fall of 2005 to the Fall of 2007. More recently, the LIGO Scientific Collaboration has committed to the eventual public release of this data. This task, an affirmation of LIGO’s commitment to open science, falls under the purview of the LIGO Open Science Center (LOSC). In order to ensure that LIGO data releases are accessible and useful to the public, LOSC archives and maintains LIGO data sets and develops software and tutorials to teach the public how to meaningfully interact with public LIGO data. Prior to the public release of S5 data, it is desirable to have a detailed understanding of every hardware injection that took place during the S5 period, with a particular eye toward how “successful” each injection was. We can then emphasize to the public that the signals from these injections are *not* the results of gravitational waves.

The data files under consideration for this paper give the strain during S5 at the Livingston and Hanford interferometers at a sampling rate of 4096 Hz. We use a log detailing various parameters of each injection (time, scale, and waveform shape) to perform an optimal matched filter search looking for the corresponding signal. The output of a matched filter search[1] is a signal to noise ratio (SNR), which gives a sense of how well a particular section of data matches a template. For each injection, we can then compare the recovered SNR to the SNR we expect from the data’s noise and the power spectrum of the template. This analysis has been performed for S5 injections simulating compact-binary-coalescences (CBCs) in a previous work[2].

## 3 Burst Injections

The beginning of our search for burst injections begins with the injection logs for each detector. Each line of the log file corresponds to one injection, and lists the time, scale, and waveform type of the injection, in addition to a log message describing whether the injection ought to have been successful given our knowledge of interferometer conditions around the time of injection.

In general, burst waveforms are not very well-known. When they are known, they are often crudely modeled. As a result, the category of burst injections includes a diverse range of waveforms. Some waveforms have no astrophysical motivation, but were injected because they are easy to describe and have well-defined frequency bands and durations, making them

good calibration tools. These waveforms are sine-gaussians (Fig. 1) and gaussians (Fig. 2). Our sine-gaussians are sine functions of various frequencies with a gaussian envelope that has a Q factor of 9. Our gaussians are characterized by  $\tau$  in  $h(t) = Ae^{-\left(\frac{t-t_{\text{peak}}}{\tau}\right)^2}$ , where  $h(t)$  is the strain as a function of time and  $A$  is a normalization constant.

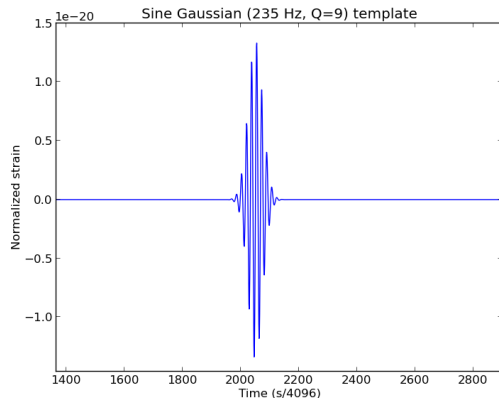


Figure 1: A sine-gaussian with a frequency of 235 Hz and a Q factor of 9.

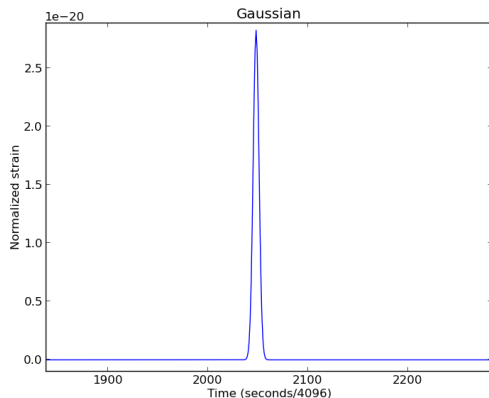


Figure 2: A gaussian with a characteristic time of  $\tau = 1$  ms.

Other models are based on astrophysical sources. Waveforms simulating supernovae are based on simulations described in Zwerger and Mueller [3] (Fig. 3) (catalog waveform A3B3G1, with minor tweaks to remove discontinuities). Another waveform simulating astrophysical phenomena is the cosmic string cusp (Fig. 4), which was injected with a cutoff frequency of 220 Hz.

The final two waveforms falling under the burst category are ringdowns (Fig. 5) and white noise bursts (Fig. 6). The ringdown is a damped sinusoid at 2600 Hz with a characteristic decay time of 300 ms. The white noise bursts are designed so that there is a fixed coherence of 0.7 between waveforms injected at the Hanford and Livingston interferometers.

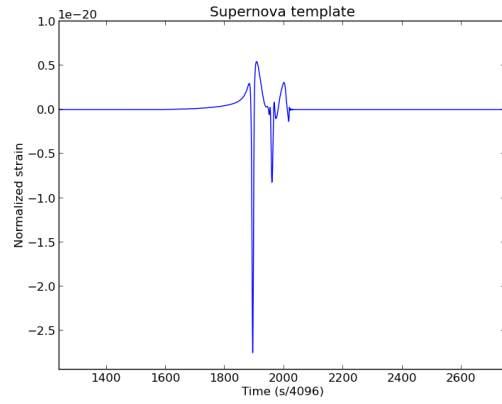


Figure 3: A Zwerger-Mueller supernova gravitational wave simulation.

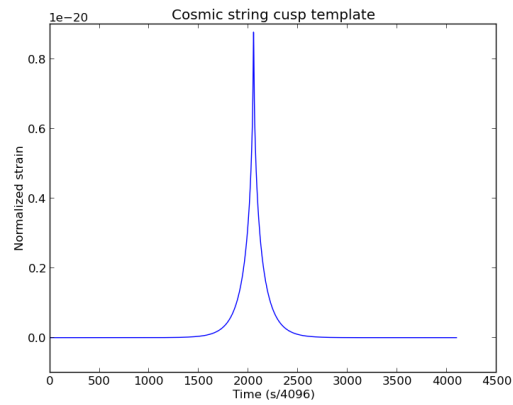


Figure 4: A cosmic string cusp with a cutoff frequency of 220 Hz.

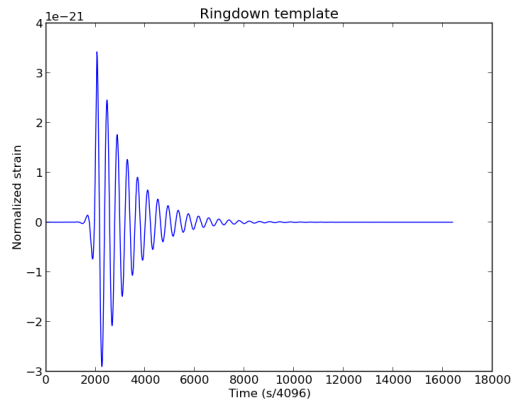


Figure 5: A ringdown at 2600 Hz with a decay time of 300 ms.

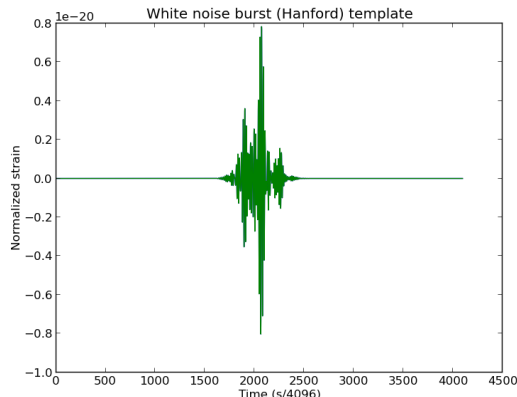


Figure 6: A Hanford white noise burst.

## 4 Searching for Hardware Injections

One of the main challenges for LIGO is finding very small signals buried in noise. In the case of recovering hardware injections, however, we are handed significant advantages.

### 4.1 Template Generation

We generate templates for our waveforms in various ways. For waveforms that are easily expressed with basic mathematical functions, we use Python’s built-in mathematical functions to generate our templates. This is our approach for gaussians, sine-gaussians, and ringdowns. For the other waveforms (Zwenger-Mueller supernovae, cosmic string cusps, and white noise bursts), we downsample the original template files to 4096 Hz. The templates are normalized to have a root-sum-square strain amplitude of  $h_{rss} = 10^{-21} \text{ Hz}^{-1/2}$ . We comment on the possibility of using the second approach to implement an automated system to perform our overall task in the “Conclusions and Future Work” section of this paper.

### 4.2 Matched Filter Search

Once we have a template, it is possible to perform a matched filter search[1]. We use our injection list to choose a 64-second interval centered on the time of injection. The output of a matched filter is given by

$$z(t) = \int_0^\infty \frac{\tilde{s}(f)\tilde{h}^*(f)}{S_n(f)} e^{2\pi ift} df, \quad (1)$$

where  $\tilde{s}(f)$  is the data in the frequency domain,  $\tilde{h}^*(f)$  is the complex conjugate of the template in the frequency domain, and  $S_n(f)$  is the power spectral density of noise near the signal. An example noise PSD is given in Fig. 7. We are thus performing a weighted cross-correlation where the weight of a frequency bin is inversely related to the noise in that frequency bin. The matched filter search takes advantage of the fact that we know *when* to expect the injection (which also gives us a good estimate of when the data is just noise)

as well as *how* the injection ought to look. We also compute a normalization constant that describes the sensitivity of the instrument at a given time:

$$\sigma_m^2 = \int_0^\infty \frac{|\tilde{h}(f)|^2}{S_n(f)} df, \quad (2)$$

where  $\tilde{h}(f)$  and  $S_n(f)$  are defined as before. Dividing our matched filter output by  $\sigma_m$  gives the amplitude SNR at which we recover the signal (Fig. 8). We say that the recovered SNR is the maximum SNR we get from a matched filter search of the data around the time of injection. From the normalization of our templates, we have that the expected SNR for a particular injection is given by  $A\sigma_m$ , where  $A$  is the scale factor given in the log file. We can then compare how well we recover an injection to how well we expect to be able to recover it.

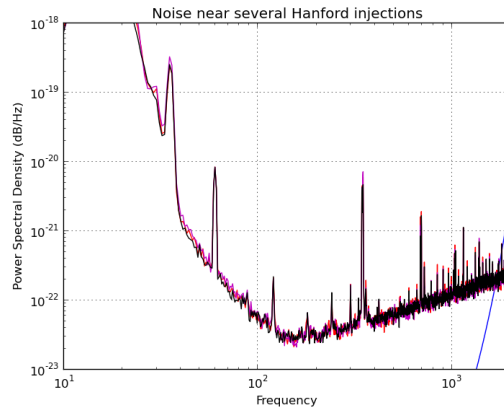


Figure 7: The power spectral densities for noise near several Hanford injections.

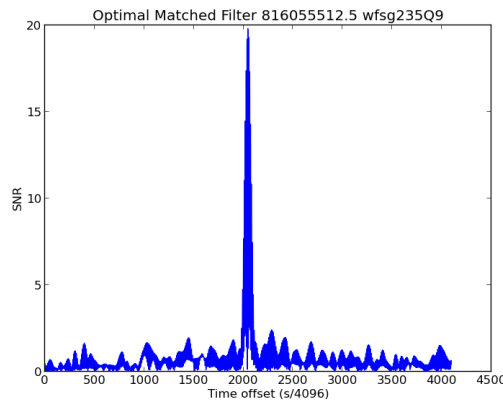


Figure 8: The normalized matched filter output for a sine-gaussian injection.

### 4.3 Efficiency

While previous work [2] provided a foundation for this project, an update to the code needed to be made for efficiency’s sake. Whereas S5 only included around 1,500 successful compact binary coalescence injections per detector, there were approximately 60,000 burst injections during S5. The updated code takes advantage of the fact that many injections happened close enough together that they are contained within the same data file. Specifically, the code begins by sorting the injections into groups. Each injection in a group happens within 100 seconds of its nearest neighbor. This allows us to decrease the number of times we need to read in data files, which becomes an important computational bottleneck when there are 60,000 injections. Not only was this efficiency scheme not necessary in [2], but the compact binary coalescence injections also happened far enough apart that the scheme would have barely helped.

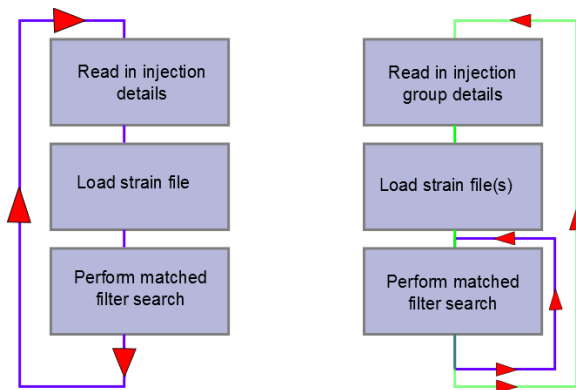


Figure 9: Updates to the flow of code for adaption to burst injections in S5. The blue lines represent loops over injections while the green lines represent loops over injection groups. Previously (left side), the code would read in a data file for each injection. In the updated efficiency scheme (right side), the injections are first grouped by proximity so that the same data file only needs to be read in once. These groups range in size from 3 to 36, improving the runtime for the code.

## 5 Results

This section provides an overview of the results of our investigation. We discuss all injection types and provide a representative sample of plots. All of the plots, along with annotated log files, can be found at <https://losc-dev.ligo.org/s5hwburst/>

### 5.1 Successful Injections

Successful injections are those whose log notes read “Successful,” as opposed to “Not-in-Science-Mode,” “Injection-Process-Off,” “GRB-Alert,” or “Injection-Compromised.” A good way to understand which injections were truly successful is to plot the recovered SNR versus

the expected SNR for each injection. We do this for each waveform type and analyze the results.

Our first group of waveform types produces plots that look as expected, in that the vast majority of points fall very close to the line where expected SNR equals recovered SNR. These waveform types include gaussians (Fig. 10), low-frequency sine-gaussians (Fig. 11), and Zwerger-Mueller supernovae (Fig. 12). There are some points which fall far away from this line. We discuss these injections in the "Outliers" section of this report. Note that the clumping of points is due to the discrete nature of scale factors applied to each injection.

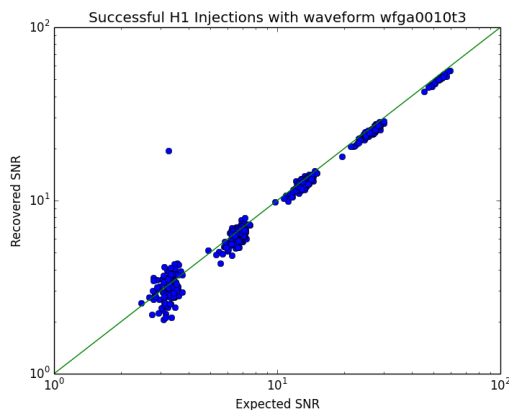


Figure 10: The results of the matched filter search for gaussians with  $\tau = 1$  ms. Note that for most points, Expected SNR  $\approx$  Recovered SNR.

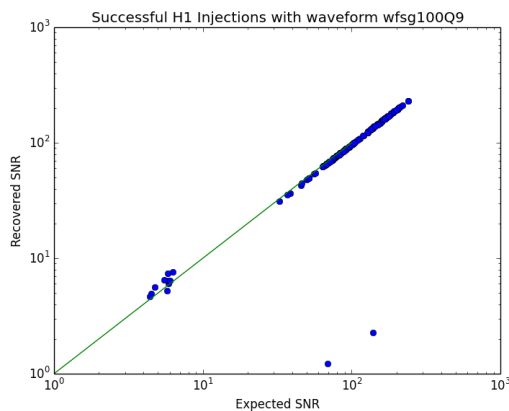


Figure 11: The results of the matched filter search for sine-gaussians at 100 Hz with a Q factor of 9.

Other waveform types produced more complicated plots. For a significant period of time at the beginning of S5, cosmic string cusp waveforms (Fig. 13) were injected with small enough scale factors that their expected SNRs were very low. This led to many injections with Expected SNR  $< 1$ , so the matched filter search picked up noise (which has SNR  $\approx$



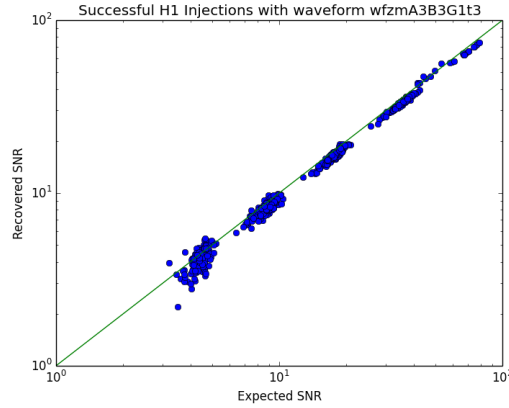


Figure 12: The results of the matched filter search for the Zwerger-Mueller supernova A3B3G1 waveform.

1), giving rise to the groups of points near the bottom left of the plot with higher recovered SNR than expected SNR.

Noting that the data is sampled at 4096 Hz, we expect that waveforms with significant activity above 2048 Hz will not be well-recovered. We sought to adjust our expected SNR calculation according by applying a low-pass filter, but the system proved too delicate for this to be effective. We therefore pick up a shift to the left for mid-frequency sine-gaussians (Fig. 14). In addition, our high-frequency waveforms (Fig. 15, Fig. 16) are not well-recovered. The high-frequency sine-gaussians, however, have many injections with very high recovered SNRs. These are covered in the “Outliers” section of this report.

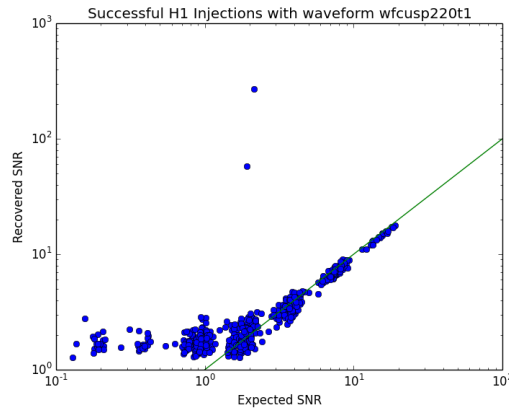


Figure 13: Matched filter search results for cosmic string cusps. The group of point near with Recovered SNR > Expected SNR is the result of many injections being too weak, causing higher-SNR noise to be picked up by the search.

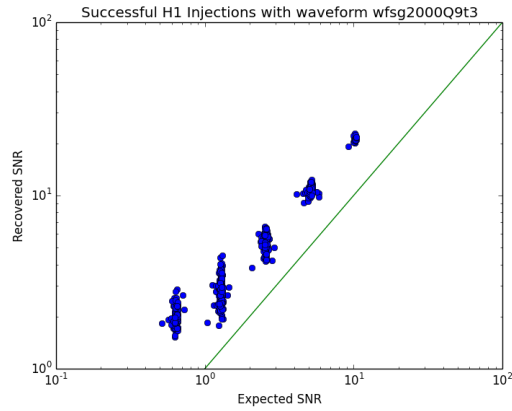


Figure 14: The low-pass filter applied to our expected SNR calculation shifts the mid-frequency sine-gaussian results to the left.

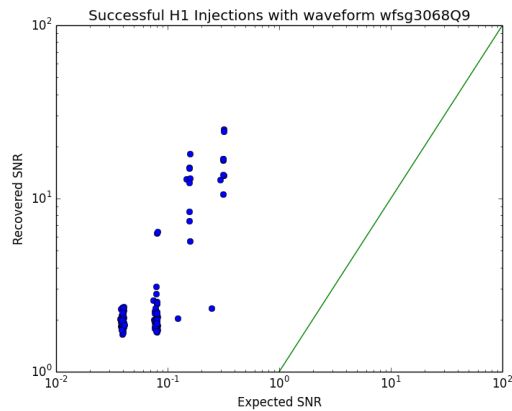


Figure 15: High-frequency sine-gaussians have too much activity above the Nyquist frequency to be recovered by our matched filter search. There are also many outliers with high recovered SNR. These were the results of highly energetic injections that saturated the interferometer's Length Sensing and Control system, causing glitches in the data.

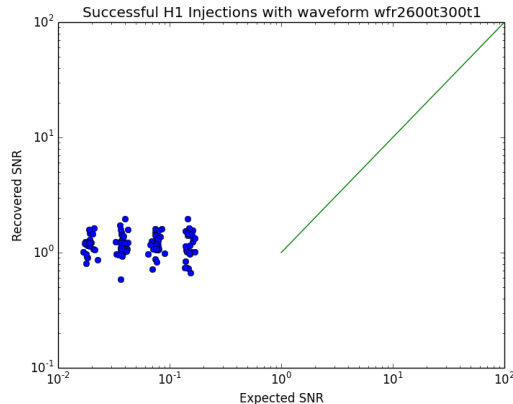


Figure 16: Since our ringdowns were damped sinusoids at 2600 Hz, they were not well-recovered by our matched filter search.

## 5.2 Outliers

In this section, we describe several types of outliers in our results and explain the methods used to understand why their features are not as expected.

One apparent mystery was that of the 3068 Hz sine-gaussians recovered with high SNRs. Since 3068 Hz is well above Nyquist, we expect recovered SNRs to be very low. However, we see in Figure 15 that some of these injections were recovered with SNRs as high as 20. One clue was the omega scans of these injections. An omega scan describes the energy at different frequencies experienced by an interferometer for a given time period. For a typical injection, the omega scan looks like Figure 17. For these high-frequency sine-gaussians recovered with high SNRs, however, the omega scans look like Figure 18. So it is clear that something went wrong with the interferometer at the time of injection. Not only is there a huge amount of energy in the omega scan, but so much of it is at frequencies much lower than 3068 Hz, the primary frequency of the waveform. We then noted that an error message kept popping up for these injections, saying H1 MASTER OVERFLOW LSC 825691756 -1 825691759 +2. This error message indicates that the controls system was overwhelmed by these highly-energetic injections. These injections caused glitches in the interferometer, giving odd SNR results.

The other typical outliers are when successful injections have bad SNRs or when unsuccessful injections have good SNRs. An example of the first case is those injections that stray from the line in Figure 10. We consult the annotated log files (available at <https://losc-dev.ligo.org/s5hwburst/>) and find that for these injections, the injection process was actually off, explaining why no SNR significantly above noise is recovered even though it is expected.

An example of the second case is Figure 19. We see that, even though the injections are marked unsuccessful, there are a few that we recover with expected SNR. Upon consulting the annotated log file, we see that these injections have been marked “Injection Compromised.” This message may be attached to an injection that happens too close to a period in which the interferometer loses lock. However, upon inspection of the annotated log file, we see that this particular injection happened at the beginning of a string of injections marked “Injection-Compromised,” suggesting that this message was attached to this injection out of

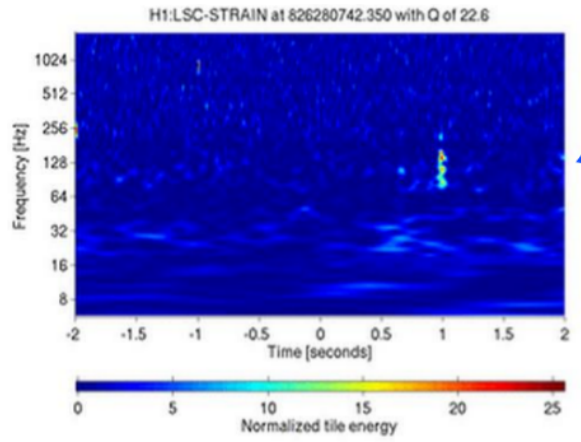


Figure 17: A "good" omega scan.

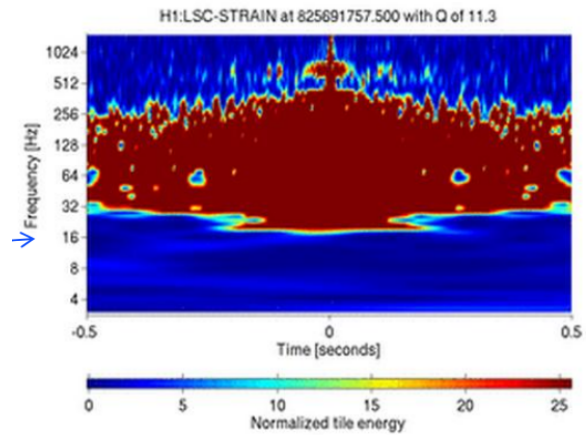


Figure 18: A typical omega scan for a 3068 Hz sine-gaussian recovered with high SNR.

an abundance of caution.

Using these methods, we were able to track down all outlier injections.

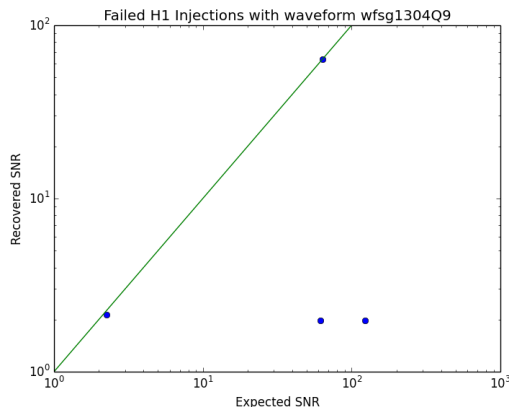


Figure 19: Unsuccessful injections of a 1304 Hz sine-gaussian.

## 6 Conclusions and Future Work

Using our updated matched filter search, we were able to recover almost all burst injections from S5. Out of the 60,000 injections, approximately 5% were outliers, which we tracked down and explained using omega scans and annotated log files. We are confident that we have a good understanding of every burst injection from S5.

This project seems like it could be automated. Instead of waiting years to check hardware injections, we could implement code to run a matched filter search on data a few hours after an injection has happened, or even sooner. The templates for this matched filter search could be easily generated from the files used to perform the hardware injections, as described in section 4.1.

This project fits into the eventual goal of releasing S5 data to the public. The completion of the search for burst injections in S5 data means we are one step closer to completing that goal.

## References

- [1] Bruce Allen, Warren G. Anderson, Patrick R. Brady, Duncan A. Brown, and Jolien D. E. Creighton, *FINDCHIRP: An algorithm for detection of gravitational waves from inspiraling compact binaries*. Phys. Rev. D 85, 122006 Published 19 June 2012.
- [2] A. Disbrow, *Investigations of Consistency in LIGO Open Science Data*. <https://dcc.ligo.org/LIGO-P1300168>.

- [3] T. Zwerger, E. Mueller, *Dynamics and gravitational wave signature of axisymmetric rotational core collapse*. *Astron. Astrophys.* 320, 209227 (1997).