

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

<b>Publication</b>	<b>LIGO-P1400168-v1</b>	2014/09/29
<b>Searching for Gravitational Waves from the Coalescence of High-mass Black Hole Binaries</b>		
Sophia Xiao Mentors: Alan Weinstein, Tjonnie Li, Surabhi Sachdev		

**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, Room NW22-295**  
**Cambridge, MA 02139**  
Phone (617) 253-4824  
Fax (617) 253-7014  
E-mail: 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

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

**Abstract**

We search for gravitational waves from the coalescence (inspiral, merger and ring-down) of binary black holes with data from the Laser Interferometer Gravitational-Wave Observatory (LIGO). Provided with well-described waveform models from General Relativity, matched filtering is employed in the GSTLAL analysis pipeline as the optimal detection technique for weak signals in Gaussian noise. The GSTLAL analysis pipeline filters data with waveform template banks, identifies triggers with SNR greater than 4, forms coincident triggers between multiple detectors in the LSC-Virgo Collaboration, and attempts to optimally separate signal from detector background noise fluctuations using a Chi-squared test. We analyze high-statistics simulations of binary merger waveforms injected into LIGO recolored S6 data to evaluate the pipeline search sensitivity and to test the readiness of the pipeline for Advanced LIGO. With Advanced LIGO fully in operation by 2015 and the upgraded analysis pipelines, the expected detection rate is increased to as much as 100 events/year or more as compared to 0.01-1 events/year in Initial LIGO. Our work will make it possible to detect gravitational waves from binary black hole coalescence in Advanced LIGO data with high confidence.

# 1 Introduction

General Relativity predicts the presence of gravitational waves as the ripples of space-time curvature. Gravitational waves (GWs) propagate outwards at the speed of light from the sources of changing gravitational fields. As GWs pass by objects, the observer on earth would detect that the distances between objects increase and decrease periodically, known as the effects of strain. Astrophysical sources of GWs include compact binary coalescences (CBCs), pulsars, supernovae, gamma ray bursts, the Big Bang, etc [1]. The amplitudes of GWs are estimated to be of the order of  $10^{-21}$  [1]. Direct detection of GWs has not been made yet. However, measurements of the Hulse-Taylor binary pulsar system provided indirect evidence of the existence of GWs. Orbital decay of the Hulse-Taylor system agrees with the prediction of General Relativity that the energy loss is due to emission of GWs [2].

LIGO Scientific Collaboration (LSC) aims to directly detect gravitational waves using several Michelson interferometers located across the globe (Livingston in LA, Hanford in WA, GEO in Germany, and a planned LIGO in India) in association with Virgo in Italy and KAGRA in Japan. A schematic of the advanced LIGO interferometer is shown in Figure 1. Successful detection of GWs will provide a direct test of General Relativity and a probe into the most exotic, least understood objects in the universe, such as black holes, neutron stars, supernovae, and the Big Bang [1]. The most promising source of ground-based detections is the CBCs, where the frequencies of GWs lie in the high frequency band (i.e.  $10^1$  to  $10^4$  Hz) that LIGO is sensitive to. Initial LIGO has searched for CBCs (NS-NS, NS-BH, and BH-BH systems) as they spiral in towards each other and merge with an expected detection rate of 0.01–1 events/year from 2002 to 2012 but reported no unambiguous detections [3]. Soon in 2015-2017, Advanced LIGO detectors will increase the amplitude sensitivity at least ten times, and will thereby increase the expected detection rate to as much as 100 events/year or more. With the upgraded sensitivity, the probability of detecting CBCs is greatly increased [4].

The CBC-BBH working group in LIGO searches for GW signals coming from binary black hole (BBH) systems with the use of matched filtering detection technique. BBH systems involve two black holes orbiting around each other, radiating off GWs. As BBH systems give off GWs, they go through three phases of coalescence – inspiral, merger, and final ringdown (IMR) [5]. While for low-mass CBC systems only inspiral lies in the LIGO frequency band, all three coalescence phases are detectable in LIGO for high-mass BBH systems (each of mass greater than  $M_{tot} = (m_1 + m_2) > 10M_{\odot}$ ).

Previously in Initial LIGO, the iHOPE analysis pipeline was implemented to search for non-spinning BBH systems, with lower mass region and higher mass region being searched separately [6]. Nonetheless, BBH systems are predicted to contain substantial spins, either aligned spins from the death of high-mass stars, or precessing spins from dynamical captures. We hereby introduce the GSTLAL analysis pipeline that is currently under development to search for BBH systems with aligned spins (work to include precessing spins is underway), as well as a larger mass range. Unlike other pipelines in LIGO, the GSTLAL pipeline has low latency, which means it can initiate other astrophysical searches using electromagnetic telescopes immediately after flagging a detection in the pipeline. This is helpful in short gamma ray bursts and supernovae searches.

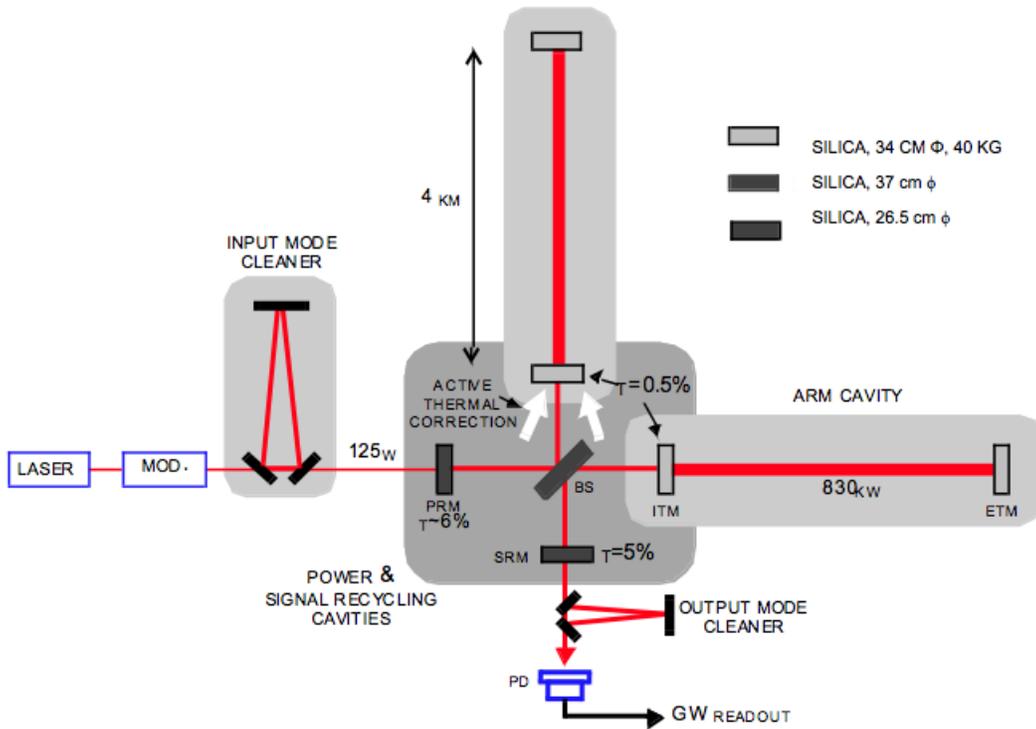


Figure 1: A schematic of the advanced LIGO interferometer. ETM = end test mass; ITM = input test mass; PRM = power recycling mirror; SRM = signal recycling mirror; BS = 50/50 beam splitter; PD = photodetector; MOD = phase modulation [4].

## 2 The GSTLAL Analysis Pipeline

Gravitational waveforms from BBH systems are believed to be well described by General Relativity. Along with the simplification of the detector noise being stationary and Gaussian, matched filtering searches for GW signals are employed as the optimal detection technique [7]. Matched filtering correlates a known signal, a template, with an unknown signal to detect the presence of the template in the data stream. In implementing matched filtering, however, the research problem is significantly complicated by the facts that (a) the exact parameters of the BBH system giving off the signal are not provided to us ahead of time, which requires the use of template banks to cover a large parameter space, and that (b) detector noise is by no means stationary and Gaussian, which is why we utilize a  $\chi^2$  statistic to suppress the influence of non-Gaussian “glitches” in the data stream.

In the GSTLAL analysis pipeline implementation, we first construct stochastic banks of discretely sampled template waveforms within specified spin range and mass range. To make a stochastic bank, waveform parameters are proposed randomly and the template is admitted to the bank if satisfying certain criteria. Then data are matched filtered with the same template bank separately in different detectors, generating signal-to-noise ratio (SNR) for all possible filters/templates. Simultaneously, because of the non-Gaussian noise existing in GW detectors, a  $\chi^2$  statistic is computed which compares the data to potential signal plus

Gaussian noise. This  $\chi^2$  is useful in distinguishing potential signals from triggers arising from non-Gaussian detector noise fluctuations, known as “glitches”. Events exceeding a certain predetermined SNR cut ( $\text{SNR} = 4$  in the current case) are recorded as triggers of interest. Then triggers close in time are clustered together and only the one with the largest SNR is recorded. We then compare triggers from different detectors and look for coincident triggers within a certain tolerance in time and the exact match in masses and spins. Subsequently, a list of coincident triggers ranked by a likelihood ratio, computed from the SNR and  $\chi^2$ , is obtained. For the remaining coincident triggers, we further compute the false alarm probability (FAP) and the inverse false alarm rate (IFAR), to determine how likely it is to mistake noise as signals. We claim a detection if a trigger has an IFAR value larger than a threshold [9]. A flowchart of the GSTLAL analysis pipeline is shown in Fig. 2.

In this section, we discuss the waveform template banks and some important detection statistics, i.e., SNR,  $\chi^2$ , and FAP, in detailed derivations.

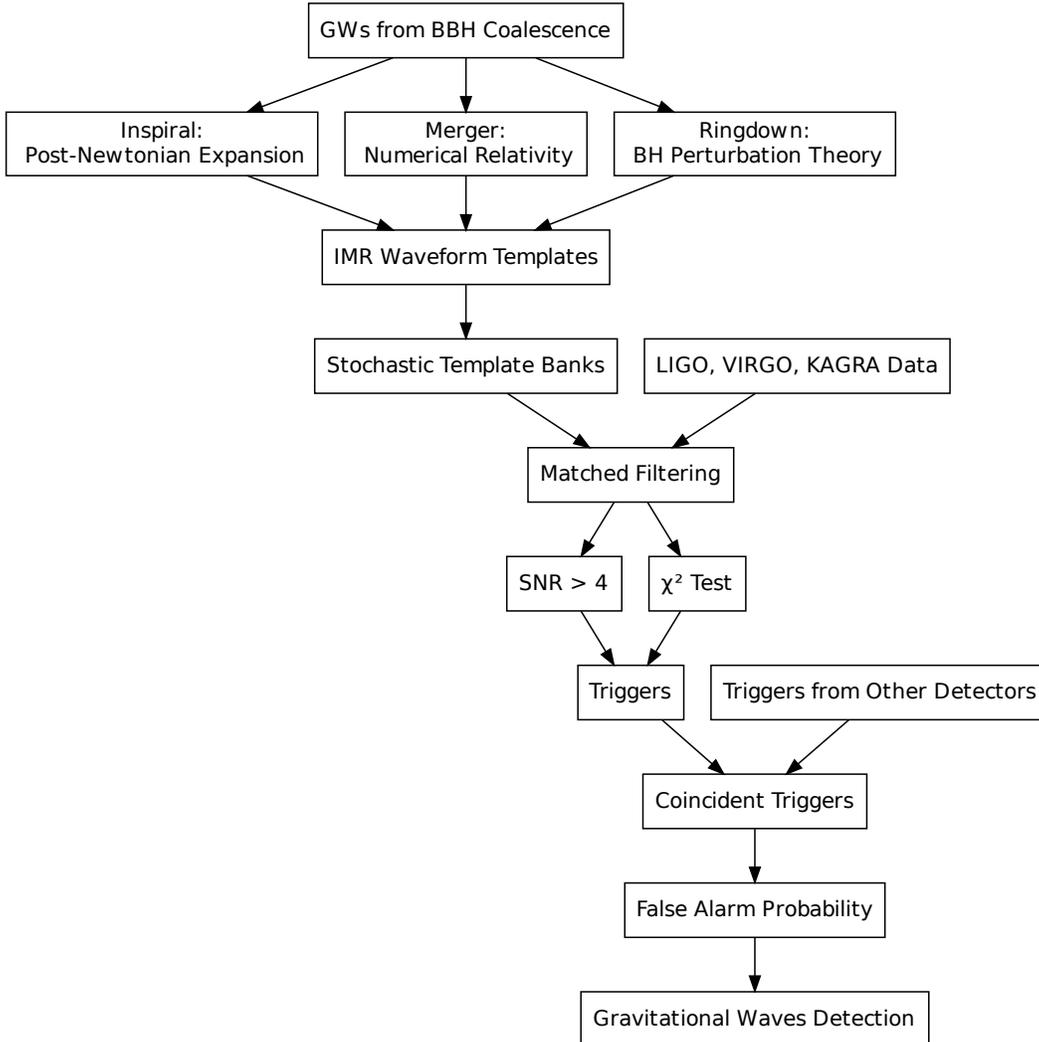


Figure 2: Flow chart of the GSTLAL analysis pipeline.

## 2.1 Waveform Template Banks

Since the parameters of BBH systems are unknown to us, we cover a large parameter space in the searches by constructing template banks. In GSTLAL, we make use of a phenomenological model, IMRPhenomB, to compute waveform templates that constitute template banks. Computing waveforms of BBH coalescence in different phases is done by different methods [9]. The inspiral phase is well described by post-Newtonian expansion. The merger phase can be calculated through numerical relativity. The ringdown phase can be calculated by black-hole perturbation theory [5]. Nonetheless, for a particular BBH system, 15 parameters are needed to specify the expected incoming waveform. Luckily, most extrinsic parameters only affect the overall amplitudes of signals, leaving us only intrinsic parameters to consider in order to cover the whole parameter space. These intrinsic parameters include component masses  $m_1$ ,  $m_2$ , the total mass of the system  $M_{\text{tot}} = m_1 + m_2$ , mass ratio of the component masses  $q = \max(m_2/m_1, m_1/m_2)$ , and spins of the component masses in the binary system  $\vec{S}_1, \vec{S}_2$ .

IMRPhenomB incorporates all three (IMR) phases of the waveform, and includes the effect of aligned spin using a single parameter  $\vec{\chi}_{\text{eff}} = (m_1\vec{S}_1 + m_2\vec{S}_2)/M_{\text{tot}}$ . Using this waveform model, the time domain and frequency domain waveforms for an example BBH system with  $M_1 = 30M_\odot$ ,  $M_2 = 30M_\odot$ , distance = 100 Mpc,  $s_{1z} = 0$ ,  $s_{2z} = 0$  are shown in Fig. 3. The three stages (i.e. inspiral, merger, ringdown) of coalescences are observable in both domains.

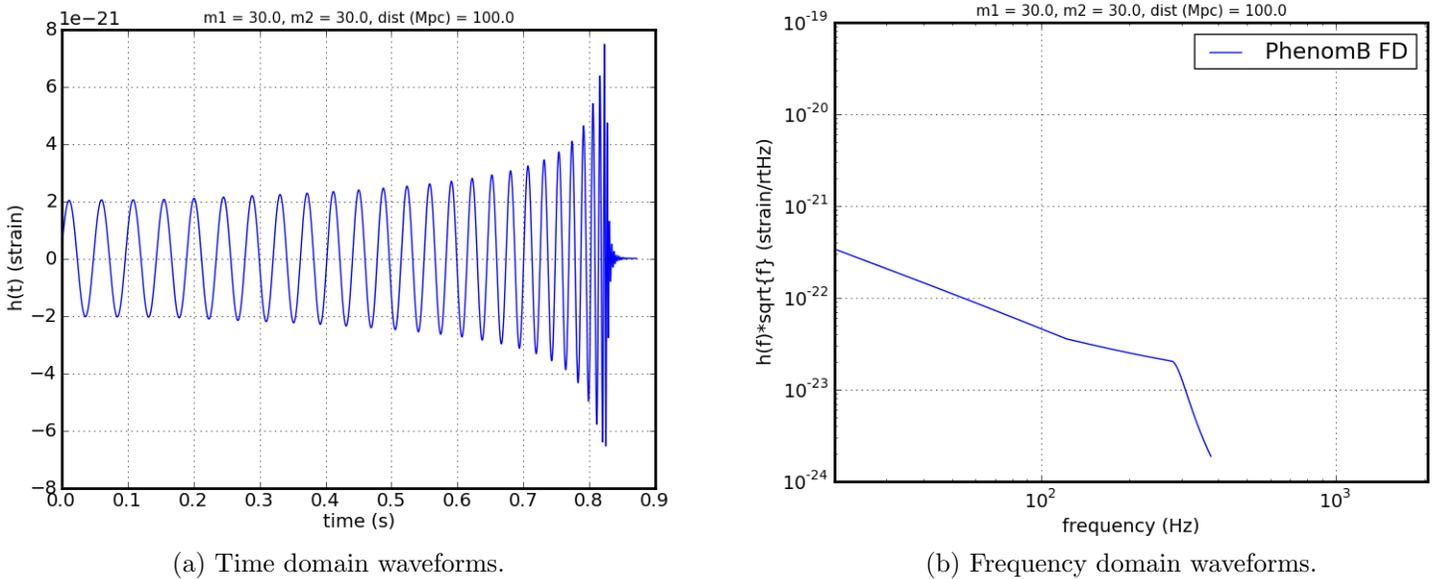


Figure 3: Example time domain and frequency domain waveforms of a BBH system. The frequency domain is the Fourier Transform of the time domain.

To generate template banks covering the parameter space (here, a 3-dimensional space comprising two component masses and one aligned spin parameter), lattice and stochastic template placement techniques could be used. In the GSTLAL pipeline implementation, stochastic banks are used. To generate a stochastic template bank, mass and spin parameters are

proposed randomly first. If the proposed template’s fitting factor, which describes how well a waveform “fits” another waveform, against the initial bank does not exceed a chosen minimum value (currently  $FF_{\min} = 0.97$ ), this template is accepted into the bank. Otherwise, the proposed template is discarded. This process is iterated until the rejection rate of proposed templates reaches a set value [9].

## 2.2 Signal-to-noise Ratio

The signal-to-noise ratio  $\rho$  statistic used in the GSTLAL analysis pipeline, which characterizes the maximum detection probability for an event, is described as the optimal detection statistic for a single event [8].

For a data set  $s(t) = h(t) + n(t)$ , where  $h(t)$  is the signal and  $n(t)$  is the Gaussian stationary noise, we define the SNR time series as

$$\rho(t) \equiv \frac{|z(t)|}{\sigma}, \quad (1)$$

where the cross-correlation of two waveforms is defined by their Hermitian inner product

$$z(t) = \langle \tilde{s} | \tilde{h} \rangle = 4 \int_0^\infty \frac{\tilde{s}(f)\tilde{h}^*(f)}{S_n(f)} e^{2\pi i f t} df, \quad (2)$$

and the measure of the sensitivity of detector  $\sigma$  is

$$\sigma = \sqrt{\langle \tilde{h} | \tilde{h} \rangle} = \sqrt{4 \int_0^\infty \frac{|\tilde{h}(f)|^2}{S_n(f)} df}. \quad (3)$$

The tilde represents the Fourier Transform of the waveform, which is given by

$$\tilde{h}(f) = \int e^{-2\pi i f t} h(t) dt. \quad (4)$$

And the noise power spectral density (PSD), which gives the average intensity at a given frequency, is defined as

$$S_n(f) = 2\langle |\tilde{n}(f)|^2 \rangle, \quad (5)$$

where  $\langle |\tilde{n}(f)|^2 \rangle$  here denotes the average of  $|\tilde{n}(f)|^2$ .

A “trigger” is a peak exceeding 4 in the SNR time series,  $\rho(t_o) > 4$ .

### 2.3 $\chi^2$ Test

A  $\chi^2$  test is used to determine whether the waveform template added to Gaussian noise is a good fit for the trigger [11]. We compute an autocorrelation  $\chi^2$  by comparing the SNR time series to the autocorrelation of the template.

In a data stream  $s = Ah + n$ ,  $n$  is the detector noise and  $Ah$  is the signal, where  $A$  is the amplitude of the signal and  $\langle h, h \rangle = 1$ . We measure the SNR time series

$$\rho(\tau) = \langle n, he^{2\pi if\tau} \rangle + A \langle h, he^{2\pi if\tau} \rangle, \quad (6)$$

where  $\tau$  is the time difference between the actual signal coalescence time and the time-shifted template. Since we do not know the BBH system's actual coalescence time, we define  $\tau = 0$  to be the time that gives the maximum SNR.

We define the autocorrelation of the whitened template as

$$\alpha(\tau) \equiv \langle h, he^{2\pi if\tau} \rangle. \quad (7)$$

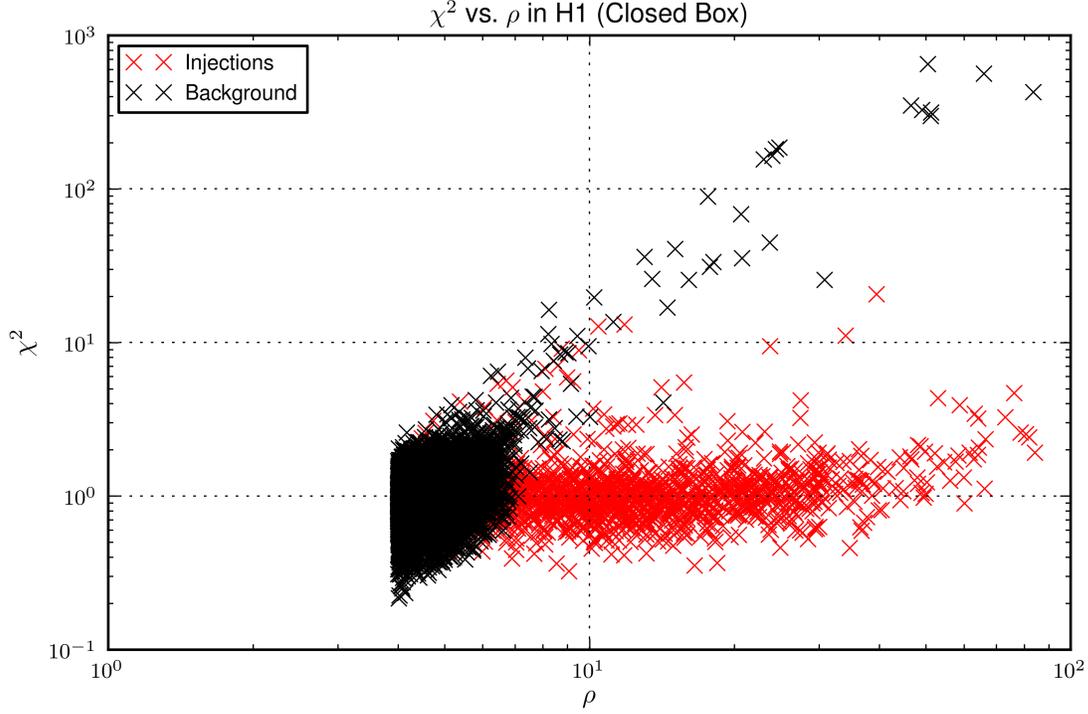
If we maximize Eqn. (6) over time and take the ensemble average, the noise term will disappear, which then will give us  $A \approx \langle \rho_{\max} \rangle$ . And thus,

$$\rho(\tau) - \langle \rho_{\max} \rangle \alpha(\tau) = \langle n, he^{2\pi if\tau} \rangle. \quad (8)$$

Quantities  $\rho(\tau)$ ,  $\langle \rho_{\max} \rangle$ , and  $\alpha(\tau)$  are easily measurable from data and templates. We use these quantities to compute the  $\chi^2$  statistic for an event,

$$\chi^2 = \int_0^{T_{\max}} |\rho(\tau) - \langle \rho_{\max} \rangle \alpha(\tau)|^2 d\tau. \quad (9)$$

If we only have Gaussian noise, SNR is the optimal detection statistic. However, with non-Gaussian detector noise, SNR statistic only is not good enough to separate signal and noise. Shown in Fig. 4 is an example  $\chi^2$  vs. SNR plot for triggers arising from simulated signals injected into recolored S6 data. We can see that some backgrounds also have high SNR values, and thus, without  $\chi^2$  statistic, these backgrounds would blend into signals. Because noise triggers tend to have higher  $\chi^2$  values than those from signal triggers,  $\chi^2$  is very effective in differentiating signal from noise when non-Gaussian noise fluctuations are present.

Figure 4: Example  $\chi^2$  vs. SNR plot.

## 2.4 False Alarm Probability

In the pipeline, it is also crucial to estimate the probability of observing a coincident event that is purely noise at the multi-detector level. False alarm probability  $P(\mathcal{L}|n)$  is the probability of identifying noise as signal at a likelihood  $\mathcal{L}$  when the data contain only noise. We aim to establish a FAP that does not depend on the time shifts between detectors, but depends on intrinsic parameters  $\vec{\theta}$  (here, the template parameters  $m_1$ ,  $m_2$  and  $\bar{\chi}_{\text{eff}}$ ) of BBH systems, matched filter SNR  $\rho$ , and  $\chi^2$  statistic [10].

Assume that we have  $N$  independent detectors. For each detector  $i$ , we rank the coincident events by standard likelihood ratio  $\mathcal{L}$ , which characterizes the likelihood of an event being signal as opposed to being noise,

$$\mathcal{L}(\rho_1, \chi_1^2, \dots, \rho_N, \chi_N^2, \vec{\theta}) = \frac{P(\rho_1, \chi_1^2, \dots, \rho_N, \chi_N^2, \vec{\theta} | s)}{P(\rho_1, \chi_1^2, \dots, \rho_N, \chi_N^2, \vec{\theta} | n)}, \quad (10)$$

where  $P(\dots | s)$  is the probability of observing the event when it is signal, and  $P(\dots | n)$  is the probability of observing the event when it is noise.

Because the noise in the different detectors is independent, the total likelihood ratio is the product of likelihood ratios from the  $N$  detectors,

$$\mathcal{L}(\rho_1, \chi_1^2, \dots, \rho_N, \chi_N^2, \vec{\theta}) \simeq \prod_{i=1}^N \mathcal{L}_i(\rho_i, \chi_i^2, \vec{\theta}). \quad (11)$$

The FAP is then, again due to the independence of different detectors,

$$P(\mathcal{L}, \vec{\theta} | n) = \int_{\Sigma} P(\mathcal{L}_1, \dots, \mathcal{L}_N, \vec{\theta} | n) d^{N-1}\Sigma = \int_{\Sigma} \prod_{i=1}^N P(\mathcal{L}_i, \vec{\theta} | n) d^{N-1}\Sigma, \quad (12)$$

where  $\Sigma$  is the contour of constant likelihood  $\mathcal{L} = \prod_{i=1}^N \mathcal{L}_i$ . And  $P(\mathcal{L}_i, \vec{\theta} | n)$  is the result of marginalization of  $P(\rho_i, \chi_i^2, \vec{\theta} | n)$  over  $\rho_i$  and  $\chi_i^2$ ,

$$P(\mathcal{L}_i, \vec{\theta} | n) = \int_{\delta} P(\rho_i, \chi_i^2, \vec{\theta} | n) d\delta, \quad (13)$$

where  $\delta$  is the contour of constant likelihood  $\mathcal{L}_i$  in the space of  $\{\rho_i, \chi_i^2\}$  at constant  $\vec{\theta}$ .

Finally, to determine a FAP over a broad range of parameter space  $\vec{\theta}$ , we average (marginalize) over  $\vec{\theta}$  (the template parameters  $m_1, m_2$  and  $\vec{\chi}_{\text{eff}}$ ):

$$P(\mathcal{L} | n) = \int P(\mathcal{L}, \vec{\theta} | n) d\vec{\theta}. \quad (14)$$

In application, the FAP of observing one noise event having likelihood ratio larger than a threshold  $\mathcal{L}^*$  is

$$P(\mathcal{L} \geq \mathcal{L}^* | n) = \int_{\mathcal{L}^*}^{\infty} P(\mathcal{L} | n) d\mathcal{L}. \quad (15)$$

Thus, by binomial expansion, the FAP of observing at least one noise event having likelihood ratio greater than a threshold  $\mathcal{L}^*$  within a sample of  $M$  coincident triggers is,

$$\begin{aligned} P(\mathcal{L} \geq \mathcal{L}^* | n_1, \dots, n_M) &= 1 - \binom{M}{0} P(\mathcal{L} \geq \mathcal{L}^* | n)^0 [1 - P(\mathcal{L} \geq \mathcal{L}^* | n)]^M \\ &= 1 - [1 - P(\mathcal{L} \geq \mathcal{L}^* | n)]^M. \end{aligned} \quad (16)$$

For a noise event, it is expected that the noise event follows Poisson Statistics [10]. Thus, the probability of observing one noise event with a Poisson mean  $\lambda$  is defined as,

$$P(N = 1 | n) = 1 - e^{-\lambda} \sum_{j=0}^{N-1} \frac{\lambda^j}{j!}. \quad (17)$$

Combining Eqn. (16) and Eqn. (17), we obtain the mean number of noise events with likelihood ratio  $\mathcal{L} \geq \mathcal{L}^*$ ,

$$\lambda(\mathcal{L} \geq \mathcal{L}^*) = -\ln[1 - P(\mathcal{L} \geq \mathcal{L}^*)]. \quad (18)$$

Substitute the mean  $\lambda$  from Eqn. (18) back into Eqn. (17), we get the probability of observing  $N$  noise events with likelihood ratio  $\mathcal{L} \geq \mathcal{L}^*$ ,

$$P(N | n) = 1 - e^{-\lambda(\mathcal{L} \geq \mathcal{L}^*)} \sum_{j=0}^{N-1} \frac{\lambda(\mathcal{L} \geq \mathcal{L}^*)^j}{j!}. \quad (19)$$

In practice, nevertheless, the pipeline computes the FAP from the observed distribution in likelihood  $\mathcal{L}$  of triggers. The event likelihood is calculated by locating its  $\rho$ ,  $\chi^2$  values in a 2-dimensional histogram for signal and for background; a typical background histogram is shown in Fig. 5a. The resulting FAP distribution for candidate triggers falls off with increasing likelihood value shown in Fig. 5b, as expected.

Furthermore, traditionally a quantity derived from FAP is utilized to estimate the significance of coincident triggers, which is the inverse false alarm rate, defined using Eqn. (18) as,

$$\text{IFAR} = \frac{T}{\lambda}, \quad (20)$$

where  $T$  is the total run time of the observation.

Coincident triggers with large IFAR are our detection candidate events.

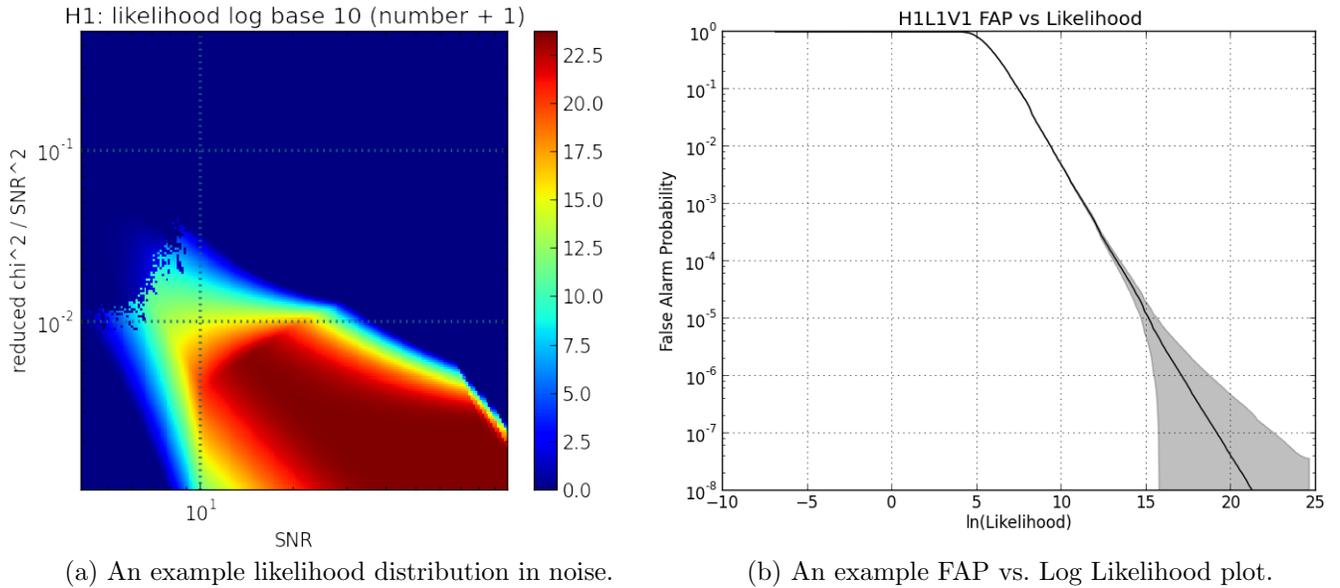


Figure 5: Likelihood distribution in noise and FAP in an example pipeline run.

### 3 Testing the GSTLAL Pipeline

In order to have the optimal analysis pipeline ready when Advanced LIGO (aLIGO) turns on, we test the GSTLAL analysis pipeline by running it with a collection of stochastic banks and injections of simulated signal added to recolored S6 data. Recolored S6 data means real Science Run 6 data, but strips out the PSDs from iLIGO and recolors the data with the Early aLIGO noise curve. Shown in Fig. 6 is the aLIGO Amplitude Spectral Density (ASD), the square root of the PSD. We then evaluate the pipeline efficiency through injection accuracy, missed and found plots, horizon distance, SNR vs.  $\chi^2$ , FAP, pipeline run time, etc. In this section, we are only concerned with the pipeline search sensitivity, characterized by different detection statistics, setting aside the computational run time. In addition to search sensitivity analysis, in the next section, we will analyze the pipeline run time for one day of data, trying to minimize the run time without jeopardizing the detection efficiency.

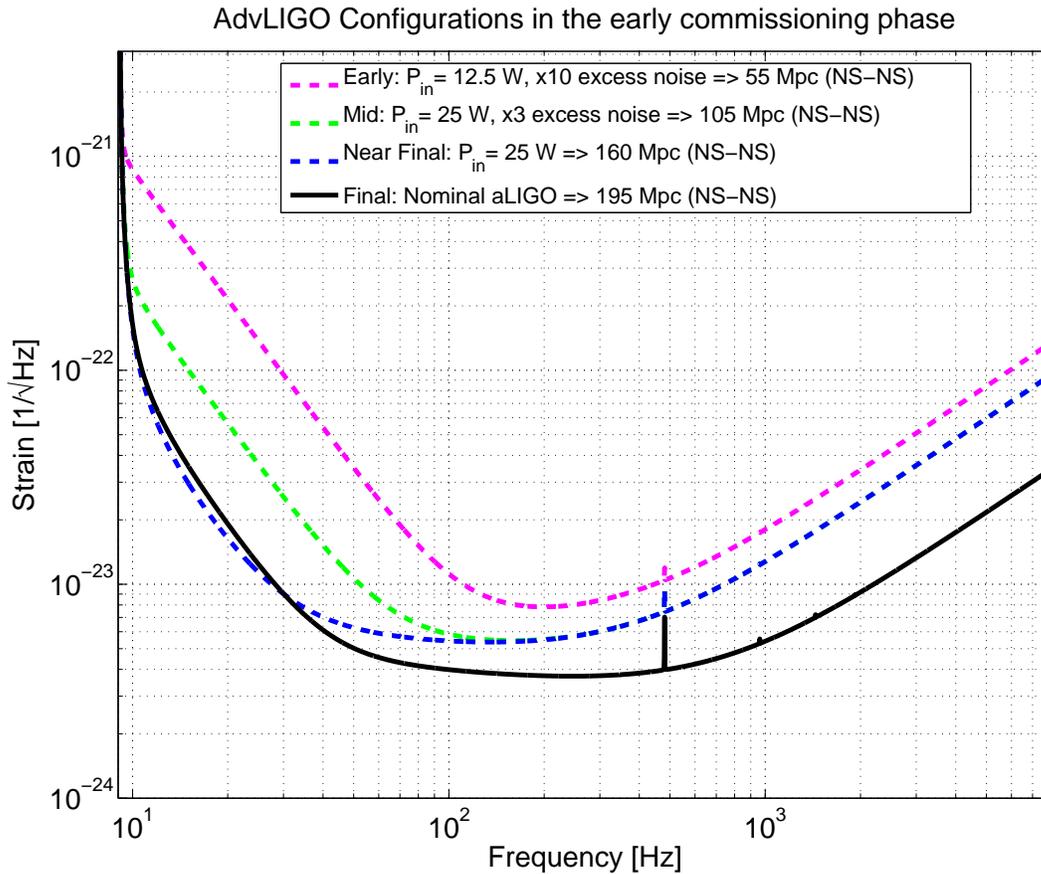


Figure 6: aLIGO Amplitude Spectral Density curves. The pink dash line is the ASD model of Early aLIGO.

### 3.1 Current Runs Status

We investigate different aspects of the GSTLAL analysis pipeline with several stochastic template banks, varied by spin range, minimum integrating frequency, bank shape, etc. We have finished generating a collection of stochastic template banks, with parameter space listed in Table 1. In Table 1, the last two banks, LgSq and LgTrape result from stitching dense low mass banks together with sparse high mass banks. Adding in higher mass region does not significantly increase the total number of templates, since BBH systems with higher mass tend to have shorter waveforms as shown in the simple example in Fig. 7, leading to wider coverage of mass parameter space for a specified minimal match. Simultaneously, other group members are trying to push down to lower mass region with a minimum component mass of  $3M_{\odot}$ . We will then combine these efforts to generate banks that cover from  $6M_{\odot}$  to  $110M_{\odot}$  in total mass, and run the pipeline with them. In addition to mass, by comparing Bank Sq2 and Bank Sq3, we demonstrate the significant contribution of spins to the number of templates in a bank. The effects of including spins on number of templates in banks is also shown in Fig. 8.

	Shape	Number of Templates	Component Mass	Total Mass	Mass Ratio	F-low	Spin	Rejection Rate
Sq1	square	13713	[4.9, 22]	[9.8, 44]	[1, 4]	20Hz	[-0.9, 0.9]	100
Sq2	square	27430	[4.9, 22]	[9.8, 44]	[1, 4]	20Hz	[-0.9, 0.9]	1000
Sq3	square	18370	[4.9, 22]	[9.8, 44]	[1, 4]	20Hz	[-0.2, 0.9]	1000
Sq4	square	12237	[4.9, 22]	[9.8, 44]	[1, 4]	30Hz	[-0.9, 0.9]	1000
Trape5	trapezoid	32289	[1, 44]	[9.8, 44]	[1, 4]	30Hz	[-0.9, 0.9]	1000
Trape6	trapezoid	100713	[1, 44]	[9.8, 44]	[1, 4]	20Hz	[-0.9, 0.9]	1000
LgSq	square	13561	[4.9, 55]	[9.8, 110]	[1, 4]	30Hz	[-0.9, 0.9]	1000
LgTrape	trapezoid	32597	[1, 110]	[9.8, 110]	[1, 4]	30Hz	[-0.9, 0.9]	1000

Table 1: Bank parameter space

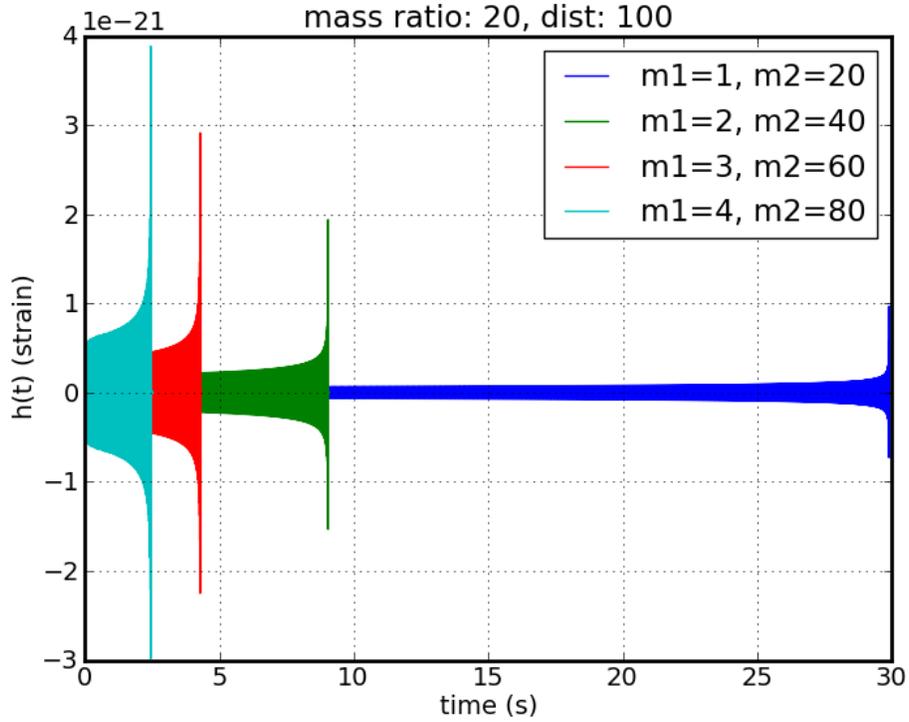


Figure 7: Example GW waveforms. Higher-mass BBH systems have shorter waveform durations above a specified  $f_{\min} = 20\text{Hz}$ .

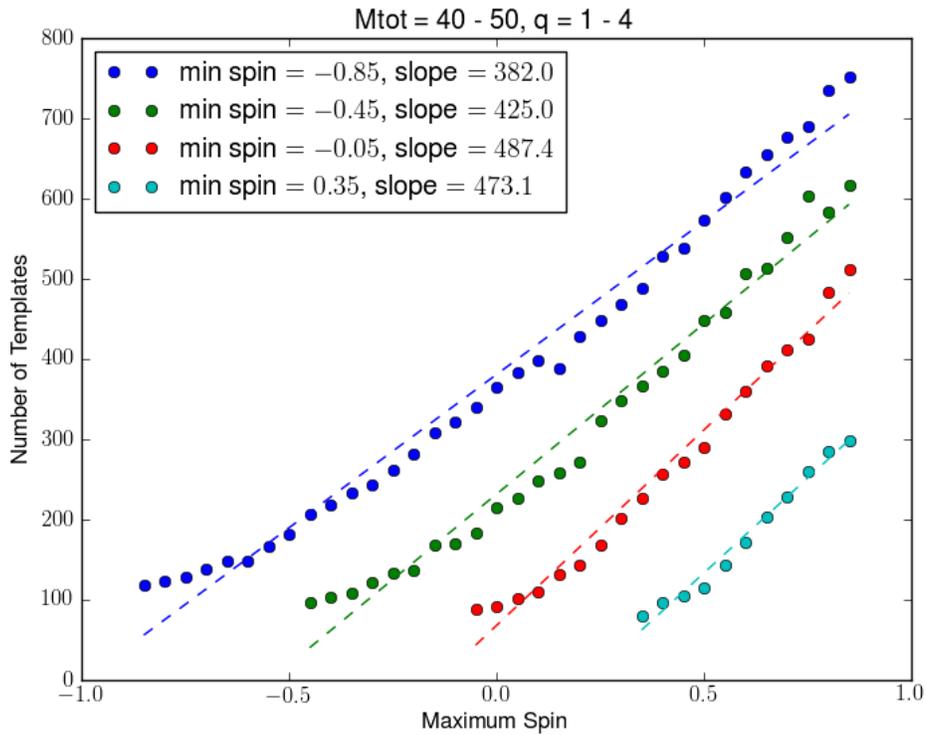


Figure 8: Range of spins vs. Number of Templates

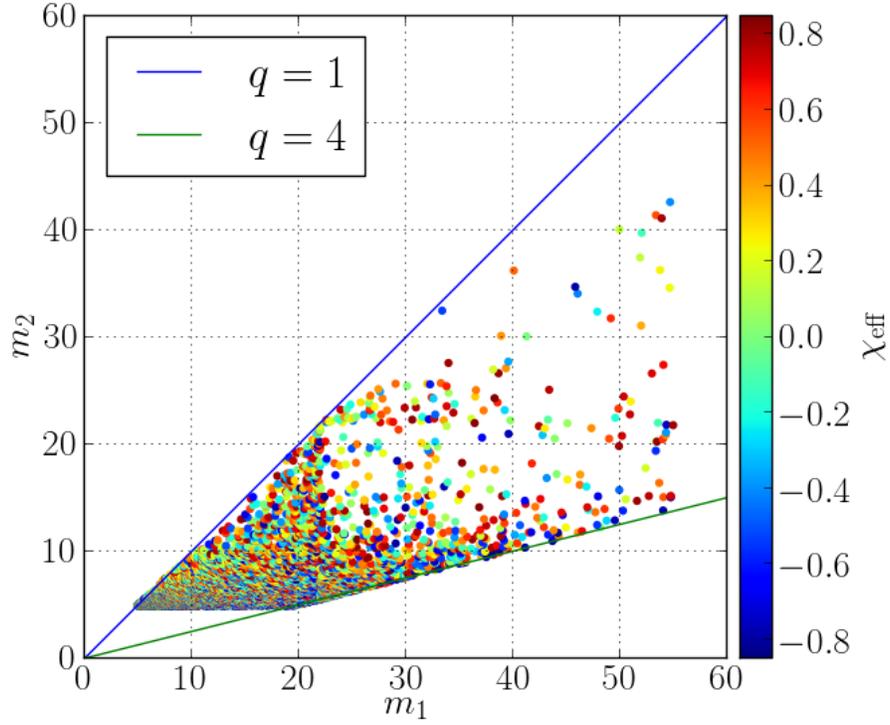
We have finished a number of daily (one-day data analysis) pipeline runs with injections of recolored S6 data using these banks, and the status of the daily pipeline runs is listed in Table 2. The GPS start time of S6 is 966384015 s, and the GPS stop time of S6 is 1077667216 s. Analysis of one week of data with Bank Sq2 and Bank Trape5 is done. Our next step is to combine the daily results into weekly results. Ultimately, we hope to combine weeklies into monthlies, monthlies into yearlies, etc.

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
GPS Start Time	966384015	966470415	966556815	966643215	966729615	966816015	966902415
GPS Stop Time	966470415	966556815	966643215	966729615	966816015	966902415	966988815
Sq1	✓	X	X	X	X	X	X
Sq2	✓	✓	✓	✓	✓	✓	✓
Sq3	✓	X	X	X	X	X	X
Sq4	✓	X	X	X	X	X	X
Trape5	✓	✓	✓	✓	✓	✓	✓
Trape6	X	X	X	X	X	X	X
LgSq	✓	△	△	△	△	△	△
LgTrape	✓	△	△	△	△	△	△

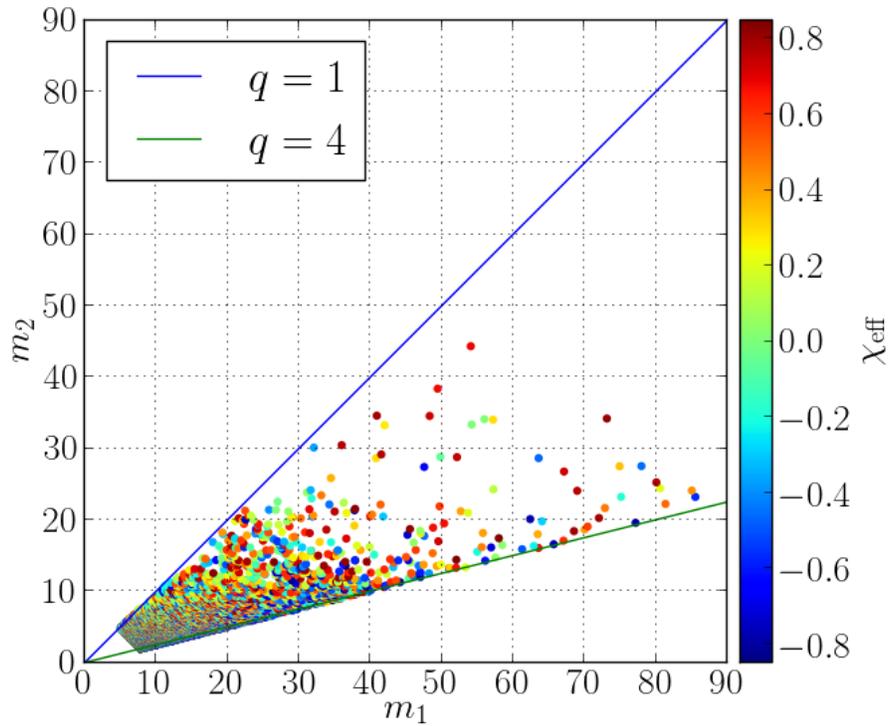
Table 2: GSTLAL daily runs status (✓: Finished; ○: Running; △: Planning; X: No plan to run.)

### 3.2 Pushing up into High Mass Region

We have successfully completed pipeline runs with banks LgSq and LgTrape. These two large banks are shown in Fig. 9. We could see that in the upper right high mass region, both banks are very sparse, as explained before. In Fig. 10, we show histograms of bank parameters vs. number of templates. Fig. 10a and Fig. 10b show that we successfully cover the full spin range, and that spins distribute uniformly throughout the spin range. And both banks have much more templates in the lower mass region than in the higher mass region as shown in Fig. 10e and Fig. 10f. Fig. 10c and Fig. 10d illustrates that templates in square banks concentrate near the  $q = 1$  line, while templates in trapezoid banks gather around the  $q = 4$  line. Our next step is to push down into low mass region, combine into an even larger mass range, and run the pipeline.

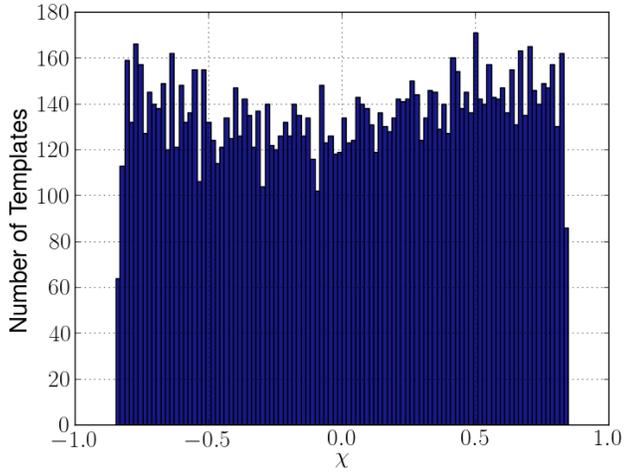


(a) Large square bank

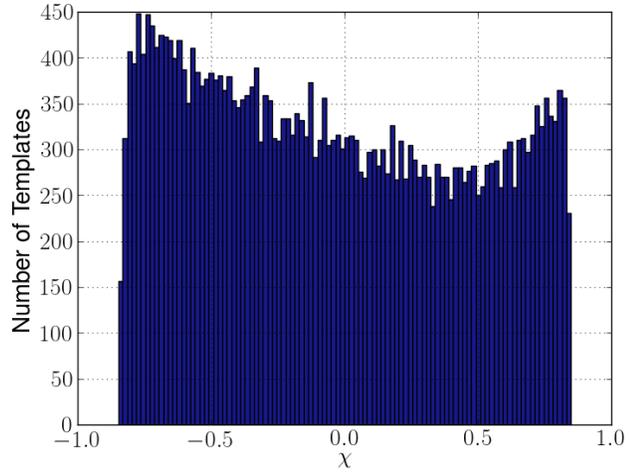


(b) Large trapezoid bank

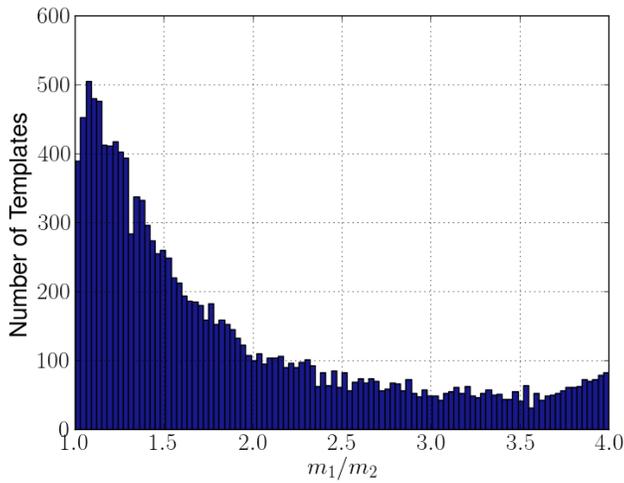
Figure 9: Visualization of large mass span banks.



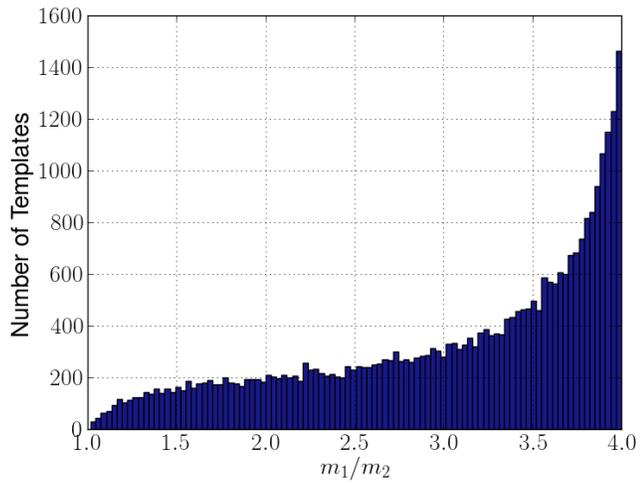
(a) Square bank:  $\chi$  vs. Number of Templates



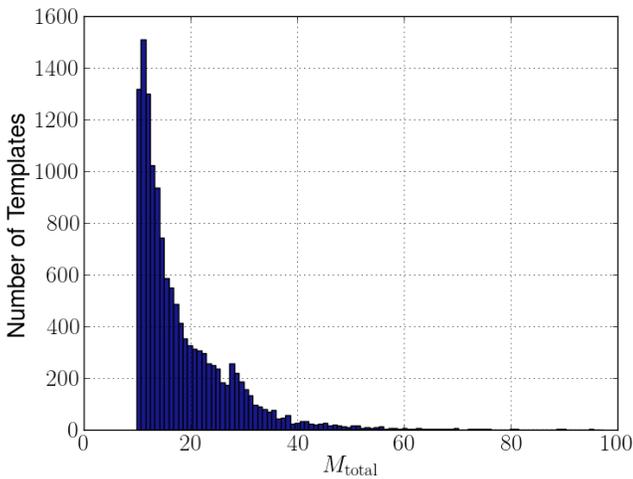
(b) Trapezoid bank:  $\chi$  vs. Number of Templates



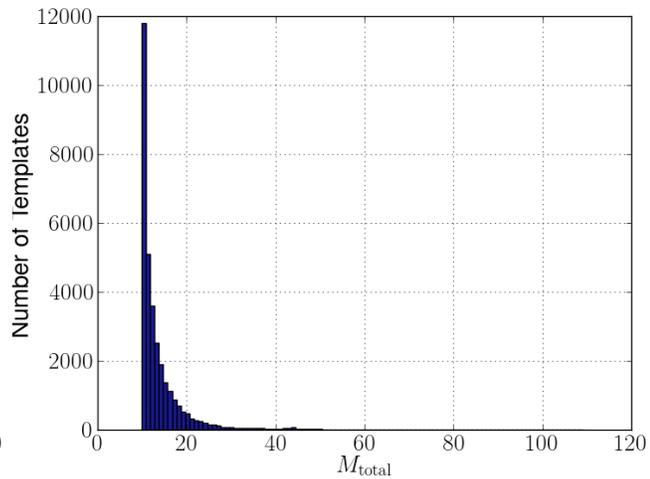
(c) Square bank: mass ratio vs. Number of Templates



(d) Trapezoid bank: mass ratio vs. Number of Templates



(e) Square bank: total mass vs. Number of Templates



(f) Trapezoid bank: total mass vs. Number of Templates

Figure 10: 1-D histograms of bank parameters.

### 3.3 Evaluate the Pipeline Efficiency

This subsection briefly shows how we check the pipeline’s performance. We demonstrate ways to evaluate the pipeline efficiency through injection accuracy, missed and found plots, horizon distance, SNR vs.  $\chi^2$ , and FAP using Day1 run on Bank Trape5 as an example. The parameters of Bank Trape5 are noted in Table 1. The shape of the bank as well as certain parameters is shown in Fig. 11.

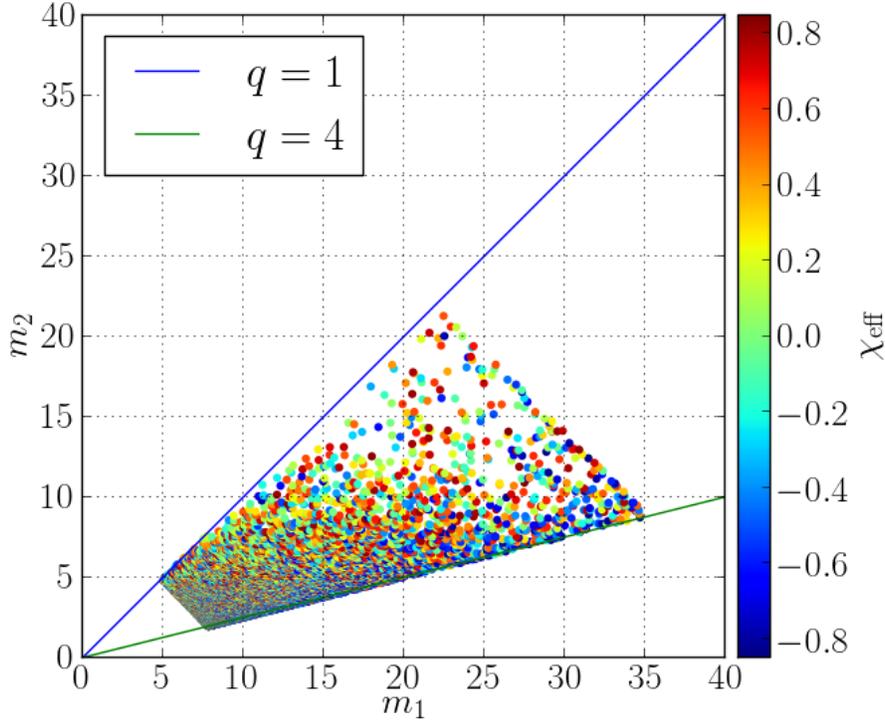
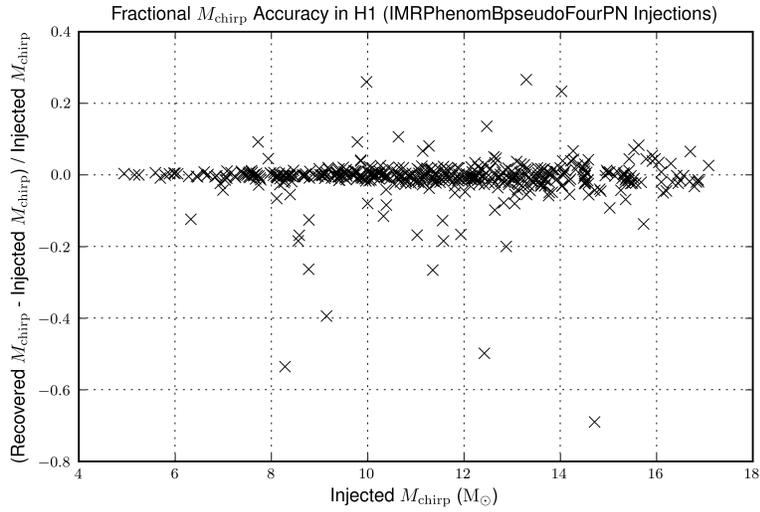


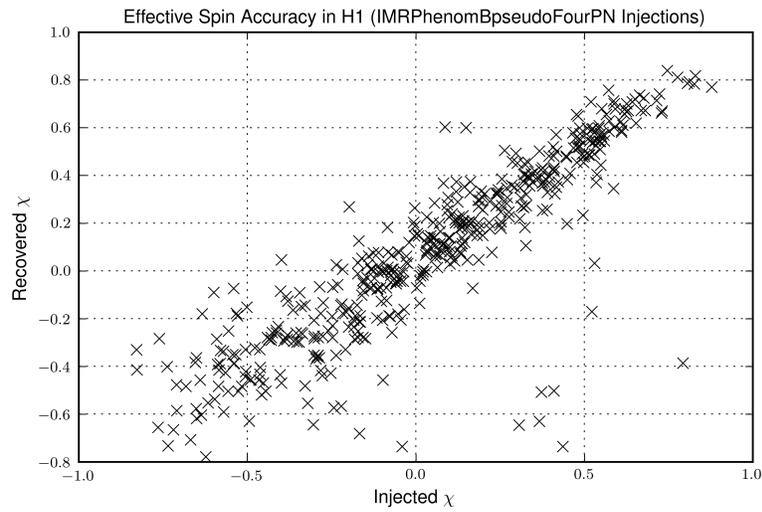
Figure 11: Visualization of Bank Trape5

#### 3.3.1 Injection Accuracy

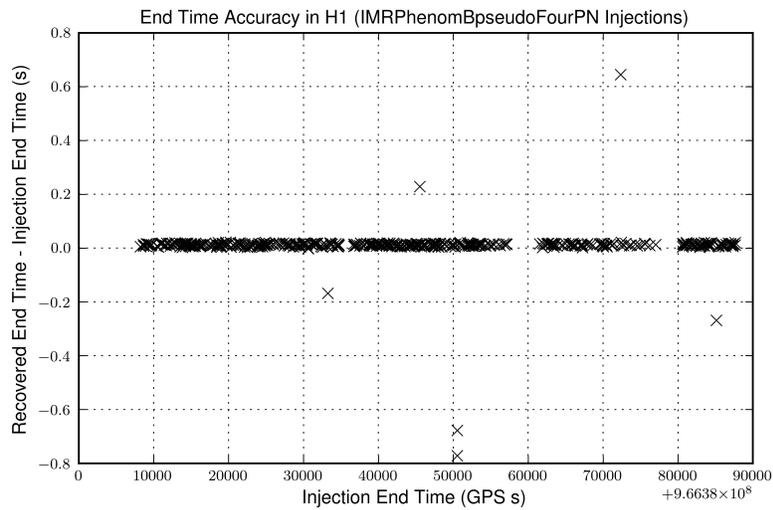
We examine the pipeline’s ability to recover BBH systems’ parameters by checking the injection accuracy. The accuracy of recovered parameters aids in diagnosing the performance of the pipeline as well as hinting at the possibility of measuring certain parameters. As illustrated in Fig. 12, we see the full spin range, the chirp mass  $M_{\text{chirp}} = (m_1 m_2)^{3/5} / M_{\text{tot}}^{1/5}$ , and the end time can be mostly recovered. Mass parameters for the low mass search are often recovered better than mass parameters for the high mass search because high mass waveforms are shorter as explained before. Degeneracy with chirp mass causes some outliers in recovering spins. Generally, most trigger parameters are well recovered, but we do have outliers that return very badly recovered parameters. At least some of these are due to “glitches” in the data masking the (weaker) signal injections. Our next step is to examine the output trigger database (“sqlite”) files to dig into these events, identify the problems and try to resolve them.



(a) injected chirp mass vs. recovered chirp mass



(b) injected effective spin vs. recovered effective spin



(c) injected effective time vs. recovered effective time

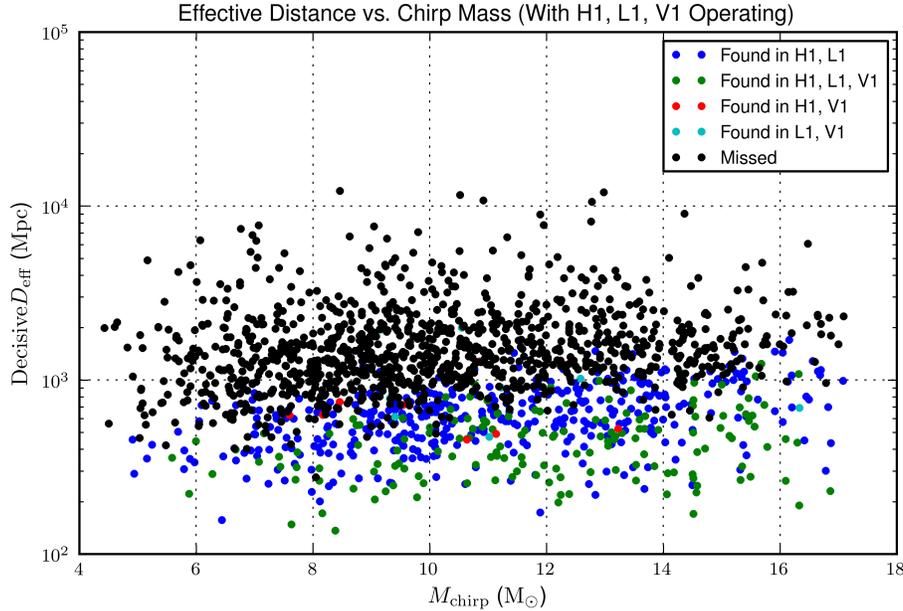
Figure 12: Injection accuracy.

### 3.3.2 Missed and Found

A table listing numbers of missed and found signal injections is shown in Fig. 13a. Measuring the missed and found injections as a function of various parameters aids in diagnosing the pipeline as well as providing the expected sensitivity of the pipeline to real signals. We measure the missed and found injections as a function of chirp mass, total mass, time, and spin. We here show the plot of effective distance of found and missed injections as a function of chirp mass in Fig. 13b. We can see in Fig. 13b that more distant signal injections are usually missed, but the dividing line between missed and found injections increases in distance with chirp mass, because the higher chirp mass sources produce higher amplitude GW signals (see also Fig. 16, below).

ON INSTRUMENTS	PARTICIPATING INSTRUMENTS	MISSED/FOUND
H1V1	H1V1	FOUND: 33
H1V1	---	MISSED: 376
H1L1V1	H1L1	FOUND: 379
H1L1V1	H1L1V1	FOUND: 149
H1L1V1	H1V1	FOUND: 9
H1L1V1	L1V1	FOUND: 4
H1L1V1	---	MISSED: 1125
H1L1	H1L1	FOUND: 14
H1L1	---	MISSED: 28
L1V1	L1V1	FOUND: 20
L1V1	---	MISSED: 212

(a) Summary of missed and found injections broken up by detector time



(b) Effective distance of found and missed injections as a function of chirp mass

Figure 13: Missed and found table and plot

### 3.3.3 Horizon Distance

The horizon distance, which is the furthest distance at which you could see a compact binary system at a given SNR, is another useful pipeline sanity check. Horizon distance grows with total mass, since the gravitational waves are louder for higher mass systems; but above a certain total mass, it will “turn over”, ie, start falling (to zero) at higher masses, as the merger frequency falls below the LIGO frequency band ( $f_{\min}$ ). Example horizon distance plots generated from pipeline runs are shown in Fig. 14, which gives the distance at which you should see a  $1.4M_{\odot} + 1.4M_{\odot}$  binary neutron star (BNS) system at SNR 8.

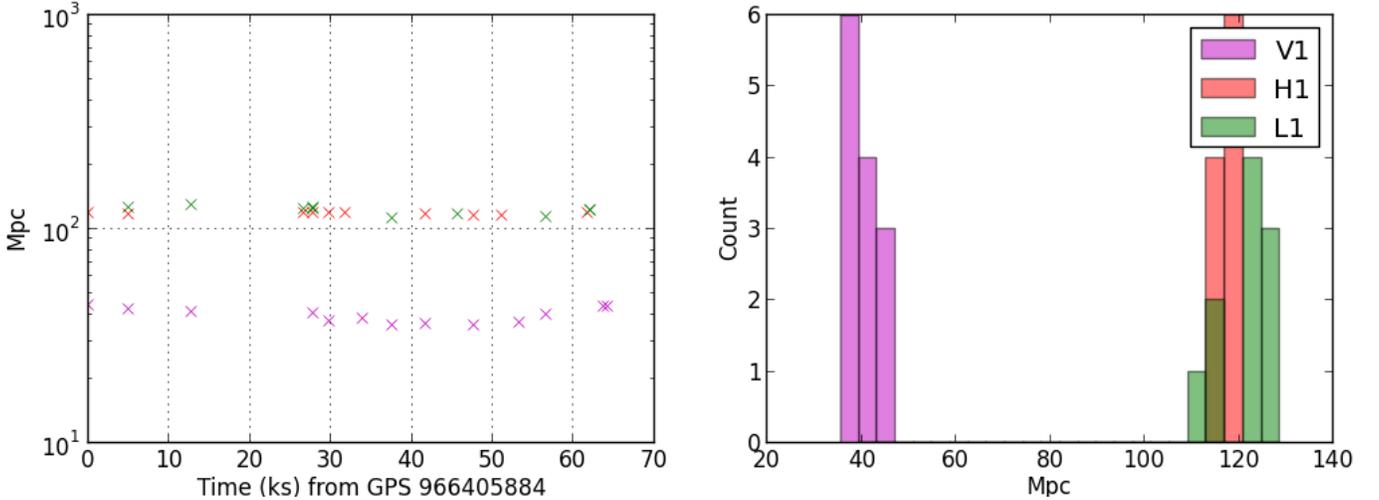


Figure 14: Horizon distance for a  $1.4M_{\odot} + 1.4M_{\odot}$  BNS system at SNR=8. The first panel shows the time evolution of the horizon distance and the second panel shows the same information in histogram for the two LIGO detectors participating in S6 and the VIRGO detector participating in VSR3.

Since horizon distance information for high-mass BBH systems is missing, efforts were put into producing plots that are more informative for higher total mass BBH systems. We use an equivalent quantity, mean sensitive distance, which is the horizon distance averaged over all sky position and orientation angles, to characterize the overall sensitivity of the search. From this averaging, the mean sensitive distance is equal to the horizon distance divided by 2.26. To evaluate the sensitivity, we first compute the theoretical mean sensitive distance using the simulated real detector Amplitude Spectral Density (ASD) curve generated from Early aLIGO ASD model. Then we compare this theoretical mean sensitive distance with the search range computed in the pipeline. The Early aLIGO ASD model and an example simulated real detector Early aLIGO ASD curve are shown in Fig. 6 and Fig. 15. Using the simulated detector ASD better models the theoretical mean sensitive distance than using the Early aLIGO ASD model because it takes into account the detector non-stationarity, non-Gaussianity, and instrumental spectral lines.

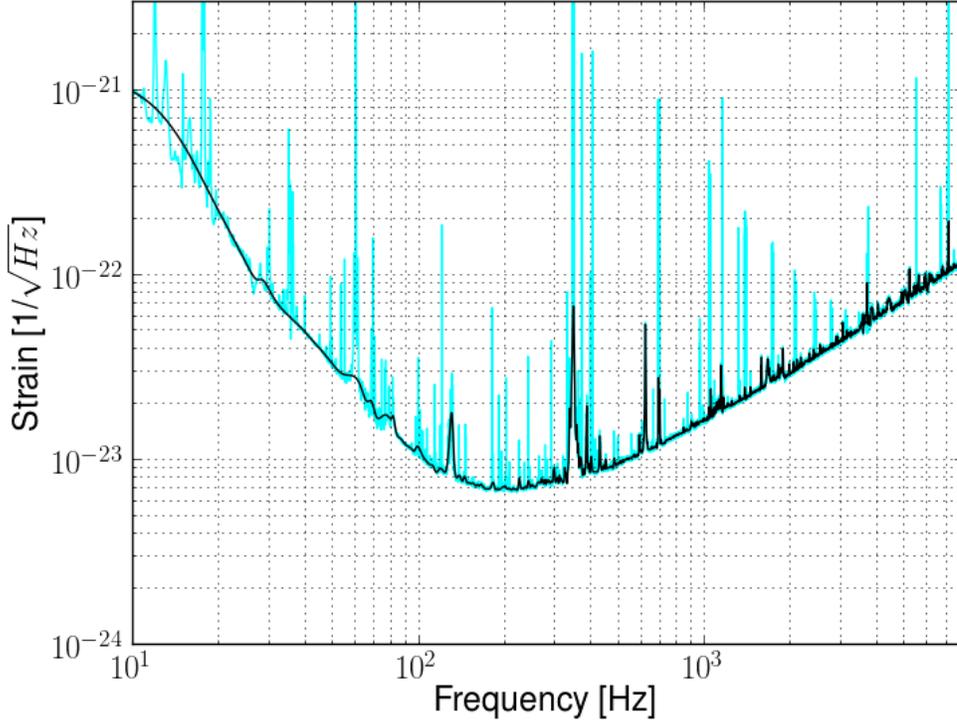
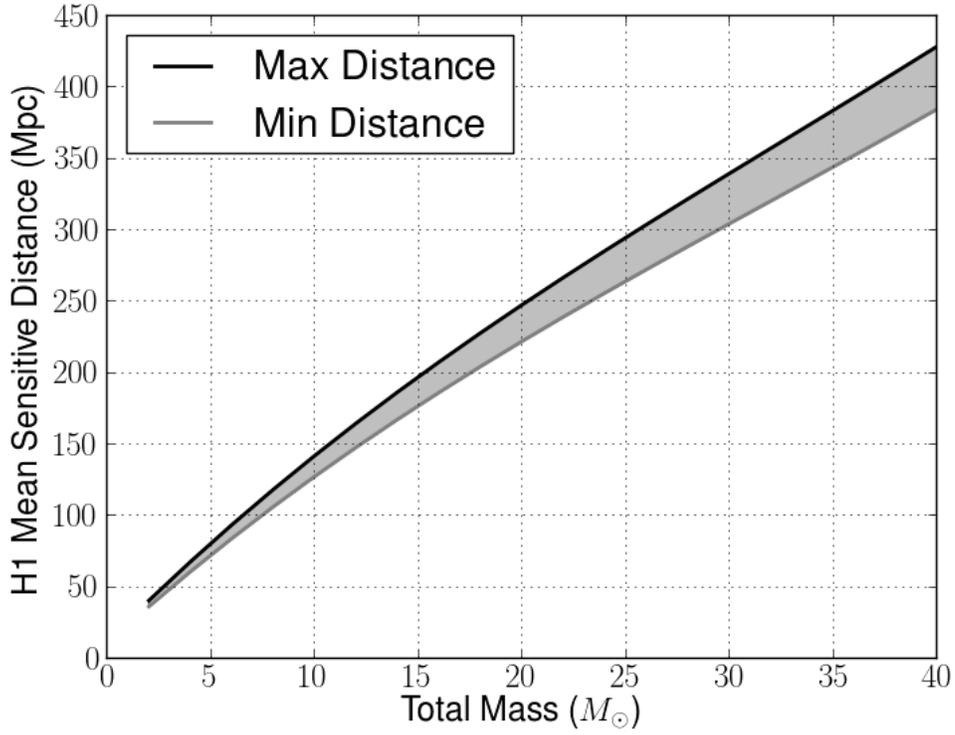


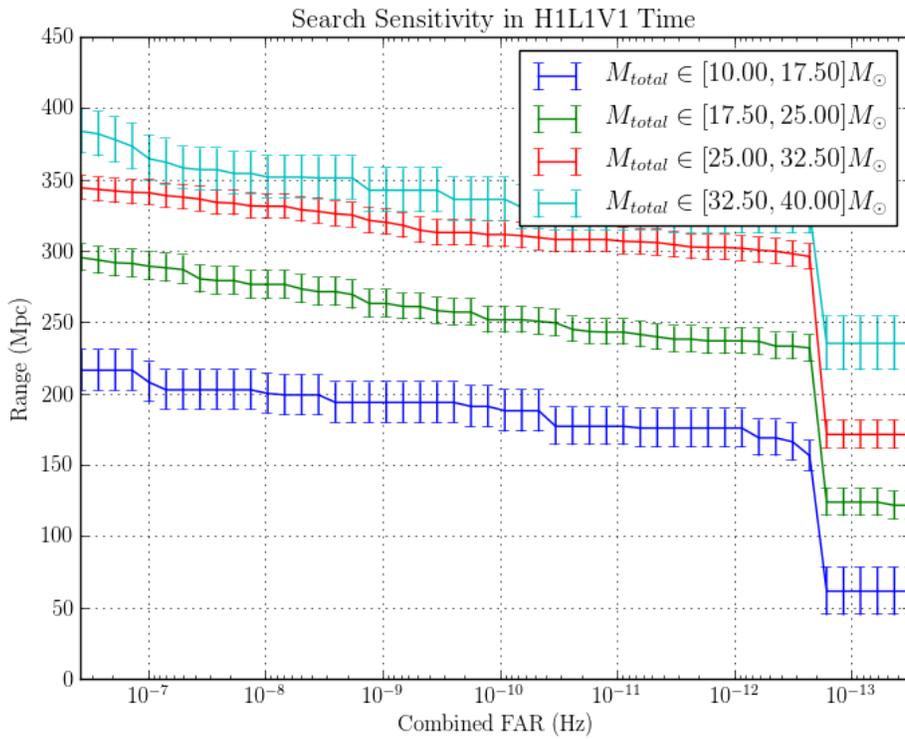
Figure 15: An example smoothed simulated real detector ASD curve generated from Early aLIGO ASD model in a pipeline run.

Finally, with the theoretical mean sensitive distance computed (see Fig. 16a), we estimate the search sensitivity by comparing the theoretical plot with the search sensitivity plot from the pipeline run, as shown in Fig. 16 for the Day1 run on Bank Trape5. If our pipeline is sensitive enough, then the pipeline mean sensitive distance should match with the theoretical one computed from real detector ASDs.

As an example, we see from Fig. 16b that systems with total mass between  $25M_{\odot}$  and  $32.5M_{\odot}$  have a mean sensitive distance between 270 and 350 Mpc. The search sensitivity for that mass range (Fig. 15), at a FAR of  $10^{-9}$  Hz (one false alarm every 30 years) is around 325 Mpc. Thus, the sensitivity with the real search pipeline is quite comparable to expectations.



(a) Theoretical mean sensitive distance



(b) Pipeline search range

Figure 16: Theoretical and real pipeline mean sensitive distance

## 4 Timing Analysis

Timing is another crucial aspect in the GSTLAL analysis pipeline. Undoubtedly, we would like to cover as large a parameter space as possible. However, this idealized case will inescapably introduce a large number of templates into a bank and consequently the pipeline will take a long time to finish. Therefore, it is not possible to enlarge the parameter space indefinitely. With limited resources, we would like to approach a sufficiently large parameter space to increase detection possibility within reasonable search time. Therefore, investigation of template placements in banks and pipeline timing analysis are paramount. In this section, we will discuss our analysis of the execution time of our pipeline. This analysis is still ongoing at an early stage. We will provide the progresses that we had so far.

The entire workflow runs under the Condor job control system, and is constructed as a Condor DAG [12]. Through carefully generating and analyzing program execution time logs, we can access the pipeline run time information after the runs have finished. More importantly, we can scrutinize the jobs that slow down the whole pipeline and try to tune the pipeline parameters so as to speed up the whole process.

For timing analysis, we have three different times recorded, “wall time”, “good time”, and “CPU usage”. “Wall time” is the difference between the time at which a task finishes and the time at which the task started, including time that passes due to artificial delays or waiting for resources to become available. “Good time” is the allocated time which contributes to the completion of the job. “CPU usage” is the CPU time which contributes to the completion of this job. Here we show the “wall time” analysis results from one example pipeline run, which is the Day1 data of S6 using Bank Trape5 referred in Table 1. The analysis of the other two time types is similar, and we include the information in the pipeline runs summary page.

Fig. 17 and Fig. 18 both show the program flowchart of the GSTLAL pipeline. A child process could not start if the information provided by its parent process is not ready. Fig. 17 shows the maximum “wall time” of each process, and Fig. 18 shows the mean “wall time”  $\pm$  standard deviation of each process. Comparing the two flow charts, we see some large discrepancy between the process’s max time and mean time. For example, “gstlal.inspiral” has a mean time of 33 minutes, while having a max time of 3 hours. To better visualize the run time distribution, we generate histograms of different processes, which are shown in Fig. 19. We notice in Fig. 19 that some processes such as “gstlal.inspiral” in Fig. 19a have long tails of run times. “gstlal.inspiral” is one of the core processes that have many jobs (in Fig. 17, there are 1960 gstlal.inspiral jobs) and this spread significantly slows down the whole pipeline run. Ultimately, we would want all the jobs from the same program to finish close in time. In order to achieve this optimization, we will probe the job information by examining the trigger database sqlite files. On one hand, we can reorder the job execution waiting line so that jobs taking the longest time start first; on the other hand, we can split the jobs in a smarter way which will equalize the run time for each job. Subsequently, jobs in line for the next process can start around the same time, greatly minimizing the total run time.

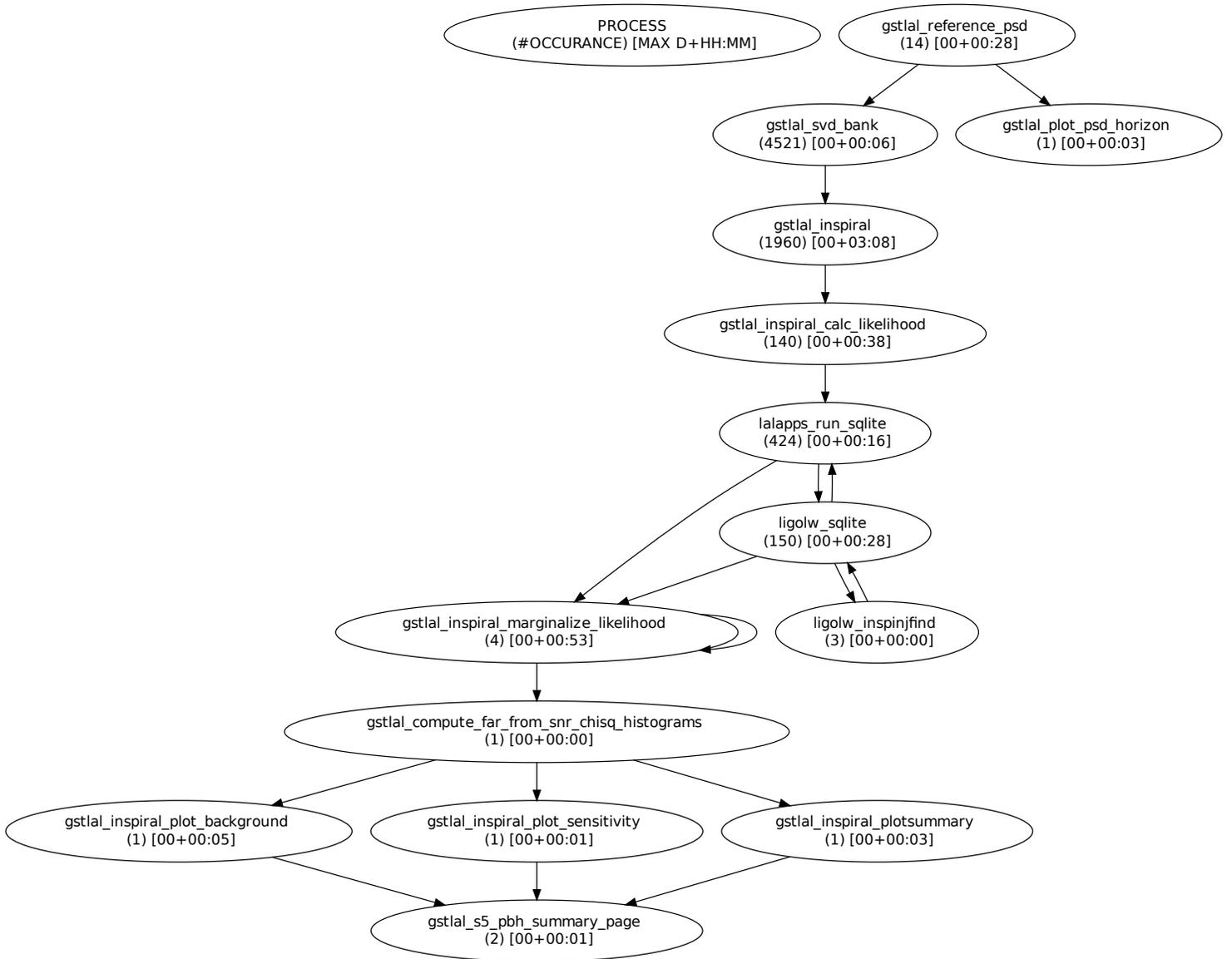


Figure 17: Example flow chart of max process run times.

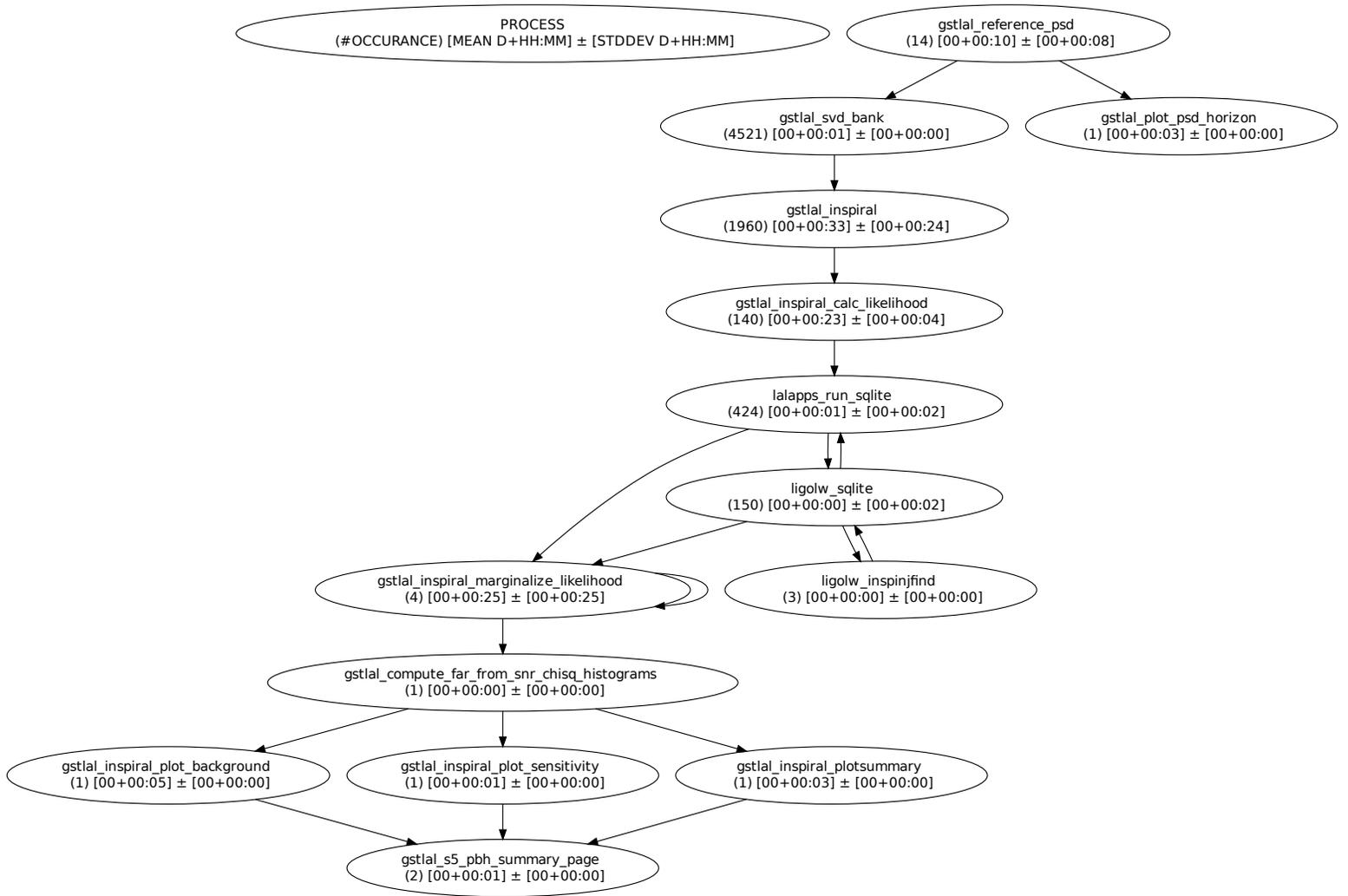
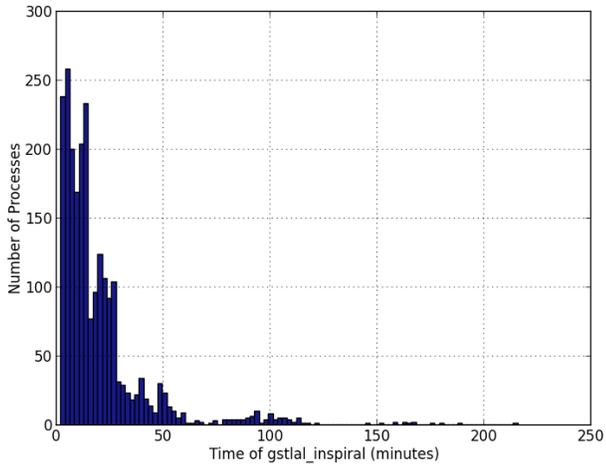
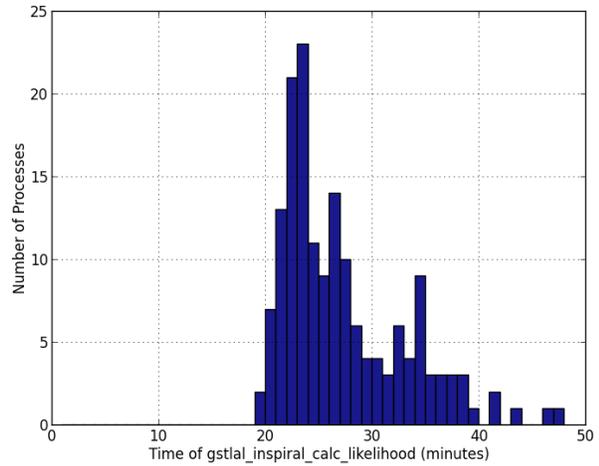


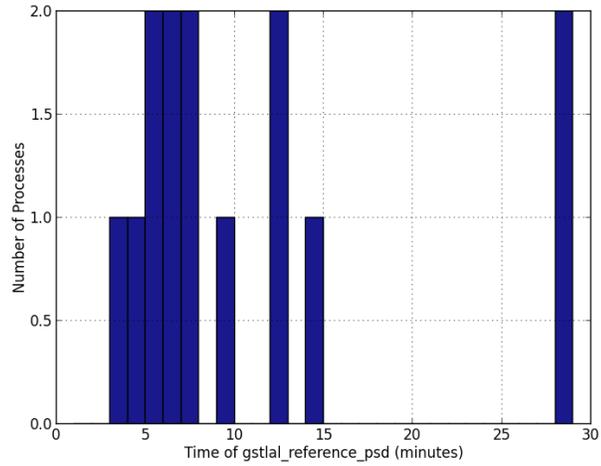
Figure 18: Example flow chart of mean process run times.



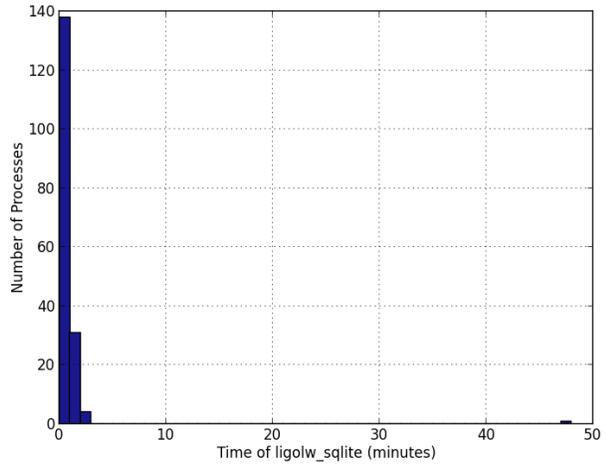
(a) `gstlal_inspiral`



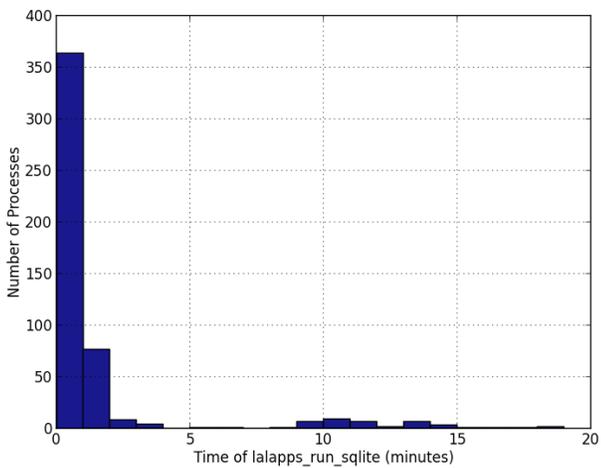
(b) `gstlal_inspiral_calc_likelihood`



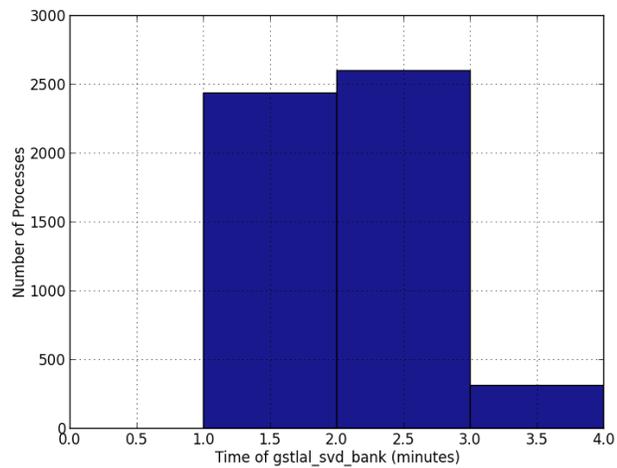
(c) `gstlal_reference_psd`



(d) `ligolw_sqlite`



(e) `lalapps_run_sqlite`



(f) `gstlal_svd_bank`

Figure 19: 1-D histograms of run times of processes which occur more than 10 times in the pipeline runs.

## 5 Future Work

Now that we have a collection of stochastic banks and daily runs, our next steps are to fully debug and tune the GSTLAL pipeline to improve its sensitivity and assemble a complete suite of diagnostic plots and search summary information.

Daily analysis of the first week of data in S6 with Bank Sq2 and Bank Trape5 is done. Up next, we will combine dailies into weeklies, weeklies into monthlies, monthlies into yearlies, etc. To make weeklies, we combine daily coincident triggers together and recalculate the background level. PSDs of GW detectors remain relatively stable over the span of one week. Thus, weeklies enable us to better estimate backgrounds since more data are available to better separate signal from noise. Each weekly summary is a "detection trial"; over many weeks, there are many trials, and thus many opportunities for noise fluctuations to fake signals. Combining them into monthlies and yearlies, and eventually, one summary over the whole run gives us one trial for finding a signal above the noise (no "trials factor").

Pipeline runs using injections containing only Gaussian noise are ongoing now. We want to compare the pipeline efficiency of the existing runs with that of Gaussian runs to see how "glitches" affect the pipeline's efficiency. For further analysis regarding existing runs, we will investigate the outliers that give us bad injection accuracy or exceptionally long job run time. Ultimately, we want to tune the pipeline parameters to improve its performance, in terms of efficiency and pipeline run time so that we can arrive at optimal settings under different anticipated noise fluctuation conditions for aLIGO.

## Acknowledgements

I would like to express my gratitude to my research mentor, Prof. Alan J. Weinstein for giving me the great opportunity of studying and working in the LIGO CBC-BBH working group. I am especially thankful for the invaluable experience of working at LIGO Laboratory at Caltech during the summer in 2014 as a participant of the Caltech LIGO SURF program. I have gained a lot of precious experience and learned a great deal under his guidance during the summer. Without his support and encouragement, I would not have done research on gravitational waves at LIGO or gained important research skills and experience or given talks on the subject that I am enthusiastic about. I would also like to thank Caltech LIGO postdoc Tjonnie G. F. Li and graduate student Surabhi Sachdev for all the help and advice during the summer. I would not have made so much progress without their help. In addition, I am grateful to all the staff at Caltech, the personnel in LIGO Scientific Collaboration, and my colleagues in LIGO CBC-BBH group, Dr. Kent Blackburn, Dr. Stephen Privitera, and Sara Frederick. Finally, I would like to express my greatest appreciation to Caltech SFP Office and NSF for funding.

## References

- [1] B.S. Sathyaprakash, Bernard F. Schutz, *Physics, Astrophysics and Cosmology with Gravitational Waves*, Living Rev. Relativity **12** (2009), [DOI:10.12942/lrr-2009-2].
- [2] Joel M. Weisberg, Joseph H. Taylor, *Relativistic Binary Pulsar B1913+16: Thirty Years of Observations and Analysis*, ASP Conf. Ser. **328** (2005) 25, [astro-ph/0407149].
- [3] **LIGO Scientific and Virgo Collaborations**, J. Abadie *et. al.*, *Search for Gravitational Waves from Binary Black Hole Inspiral, Merger and Ringdown in LIGO-Virgo Data from 2009-2010*, Phys. Rev. **D87** (2013) 022002, [arXiv:1209.6533].
- [4] **Advanced LIGO Team**, *Advanced LIGO reference design*, LIGO technical document LIGO-M060056-10-M, available at <http://www.ligo.caltech.edu/docs/M/M060056-10.pdf>
- [5] **LIGO Scientific and Virgo Collaborations**, J. Abadie *et. al.*, *Search for gravitational waves from binary black hole inspiral, merger and ringdown*, Phys. Rev. **D83** (2011) 122005, Erratum-ibid. **D86** (2012) 069903, [arXiv:1102.3781].
- [6] S. Babak, R. Biswas, P. R. Brady, D. A. Brown, K. Cannon, C. D. Capano, J. H. Clayton, T. Cokelaer, J. D. E. Creighton, T. Dent, A. Dietz, S. Fairhurst, N. Fotopoulos, G. Gonzalez, C. Hanna, I. W. Harry, G. Jones, D. Keppel, D. J. A. McKechn, L. Pekowsky, S. Privitera, C. Robinson, A. C. Rodriguez, B. S. Sathyaprakash, A. S. Sengupta, M. Vallisneri, R. Vaulin, and A. J. Weinstein, *Searching for Gravitational Waves from Binary Coalescence*, Phys. Rev. **D87** (2013) 024033, [arXiv:1208.3491].
- [7] C. W. Helstrom, *Statistical Theory of Signal Detection*, 2nd ed. (Permagon, London, 1968).
- [8] B. Allen, W. Anderson, P. Brady, D. Brown, J. Creighton, *FINDCHIRP: An Algorithm for detection of gravitational waves from inspiraling compact binaries*, Phys. Rev. **D85** (2012) 122006, [DOI:10.1103/PhysRevD.85.122006].
- [9] S. Privitera, S. R. P. Mohapatra, P. Ajith, K. Cannon, N. Fotopoulos, M. A. Frei, C. Hanna, A. J. Weinstein, J. T. Whelan, *Improving the Sensitivity of a Search for Coalescing Binary Black Holes with Non-precessing Spins in Gravitational Wave Data*, Phys. Rev. **D89** (2014) 024003, [arXiv:1310.5633].
- [10] K. Cannon, C. Hanna, D. Keppel. *Method to Estimate the Significance of Coincident Gravitational-wave Observations from Compact Binary Coalescence*, Phys. Rev. **D88** (2013) 024025, [DOI: 10.1103/PhysRevD.88.024025].
- [11] B. Allen, *A  $\chi^2$  time-frequency discriminator for gravitational wave detection*, Phys. Rev. **D71** (2005) 062001, [arXiv:gr-qc/0405045].
- [12] <http://research.cs.wisc.edu/htcondor/dagman/dagman.html>