

Searching for gravitational waves from the coalescence of high mass black hole binaries

First Progress Report

Ka-Lok, Lo¹

Mentors: Surabhi Sachdev², Kent Blackburn², Alan Weinstein²

¹ Department of Physics, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong SAR

²LIGO Laboratory, California Institute of Technology, Pasadena, California 91125, US

July 6, 2016

Program: LIGO SURF 2016

Version 2

DCC Number: LIGO-T1600261

Abstract

We aim to detect gravitational wave signals from the coalescence (inspiral, merger and final black hole ringdown) of compact binary systems (neutron stars and/or black holes) with data from the advanced detectors (LIGO, Virgo, KAGRA). The merger signal from the coalescence of low-mass systems (binary neutron stars) tends to lie above the LIGO frequency band; for most events, only the inspiral phase is detectable. For higher mass systems (involving black holes, each of mass greater than 5 solar masses), the merger and final ringdown are also detectable. We search for these signals using analysis pipelines which filter all the data, identify triggers of interest, form coincident triggers between multiple detectors in the network, and attempt to optimally separate signal from detector background noise fluctuations. We use simulated signal injections to evaluate the sensitivity of the search pipeline. The analysis pipeline has numerous parameters that can be tuned to improve the sensitivity. In this project, we will run high-statistics simulations to evaluate the search sensitivity as the analysis parameters are tuned, to arrive at optimal settings under different anticipated noise fluctuation conditions.

1 Motivation: searching for gravitational wave signals

The advanced LIGO (aLIGO) detectors are capable of measuring a strain as tiny as 10^{-23} in order to detect gravitational wave signals from various astrophysical sources such as binary black hole coalescences as in the GW150914 and GW151226 events. However, under such a high sensitivity, the detectors are subject to tremendous amount of noises, such as seismic noise from the ocean wave and near-by traffic; dark noise from the electronic fluctuation in photo-diodes; radiation pressure noise from laser, etc. Figure 1 shows the amplitude spectrum density of both aLIGO detectors near the GW150914 event which provides an indication of instrument noises.

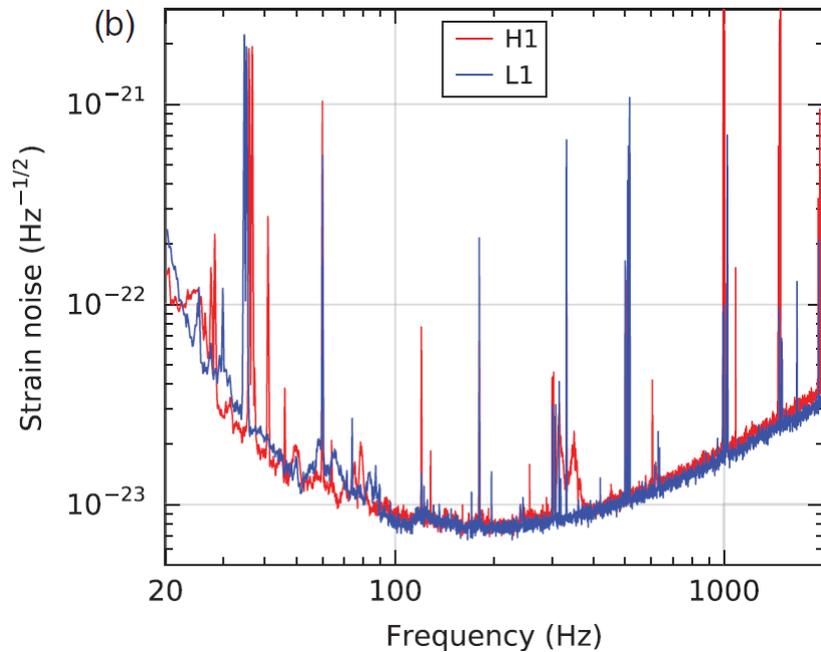


Figure 1: The amplitude spectrum density (ASD) as an indication of instrument noises of the two detectors, H1 in Hanford and L1 in Livingston, during the GW150914 event, at different frequencies. (Figure taken from [7])

The top row in Figure 2 shows the bandpass-filtered strain data from detectors in Hanford and Livingston during the GW151226 boxing day event. It is impossible to use human eyes to look for gravitational wave signals from the raw data and we need to apply appropriate data analysis techniques to

search for gravitational waves in a more systematic and reliable way, which will be discussed in the next sub-section.

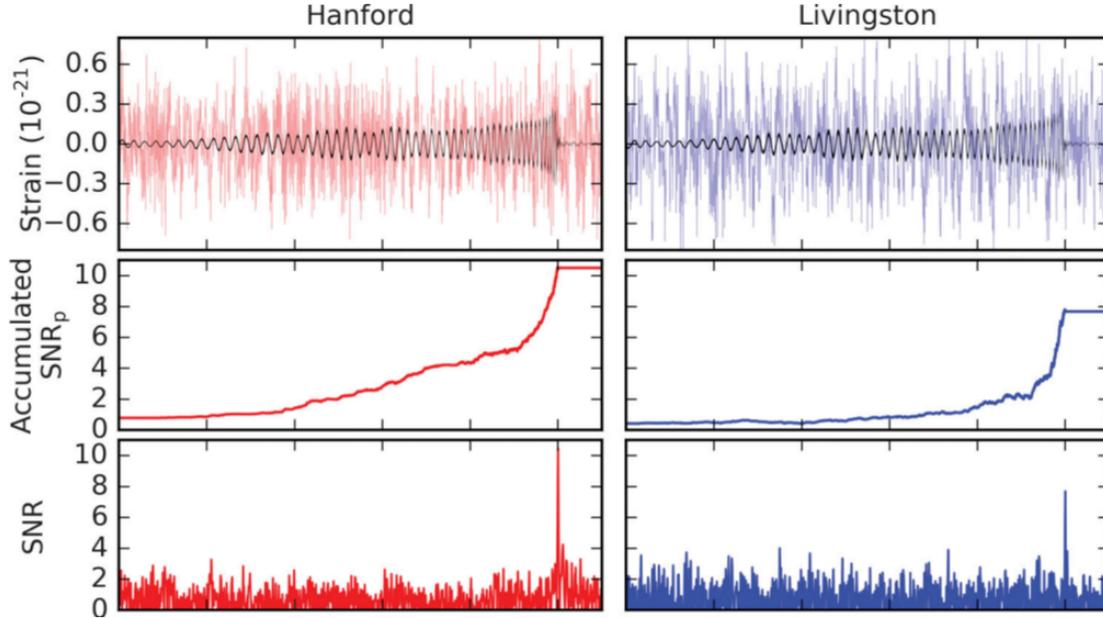


Figure 2: *Top row:* Bandpass-filtered (30-600Hz) strain data from Hanford (Left, Red in color) and Livingston (Right, Blue in color) during the GW151226 event, where gravitational wave signal can hardly be seen. The black curve on both plots is the best-match template found. *Middle row:* Time series of accumulated SNR as calculated with the strain data and the best-match template. *Bottom row:* Time series of SNR as calculated with the strain data and the best-match template. The SNR in both sites peak near the merger phase. (Figure taken from [6])

1.1 Technique for searching gravitational waves from compact binary coalescence: Matched filtering

In order to search for gravitational wave signals buried with noises, we can deploy matched filtering. The basic idea behind matched filtering is that we slide a pre-generated template across the signal and compare the two and repeat the above process for a large number of templates. For each comparison we compute a value known as **Signal to Noise Ratio** (SNR) ρ which can be intuitively defined as

$$\text{SNR} = \rho = \frac{\sqrt{\text{Power}_{\text{signal}}}}{\sqrt{\text{Power}_{\text{noise}}}}.$$

Mathematically, if we define the inner product between two functions $a(t)$ and $b(t)$ in time domain to be [3]

$$\langle a, b \rangle = 2 \int_0^\infty \frac{df}{S_h(f)} [\tilde{a}(f)\tilde{b}^*(f) + \tilde{a}^*(f)\tilde{b}(f)], \quad (1)$$

where $\tilde{a}(f)$ is the Fourier Transform of $a(t)$ and $*$ denotes the complex conjugate.

Suppose there is a gravitational wave signal $h(t)$ arriving at a detector at time $t = t_a$ and we match-filter it with a template $q(t)$, the signal-to-noise ratio is calculated by

$$\rho^2 = \frac{\langle h e^{2\pi i f(\tau - t_a)}, q \rangle}{\sqrt{\langle q, q \rangle}}, \quad (2)$$

where $S_h(f)$ is the one-sided power spectrum density (PSD) and τ is the lag, which is the time duration that the filter lags behind the output of detector. [3].

Higher the SNR, higher the similarity between the signal and the template. Since the signals that we are looking for are not known in advance, as a result we have to prepare a collection of hundreds of thousands of templates, called the template bank, that covers the parameter space we are searching for.

component masses	m_1, m_2
component spins	\vec{S}_1, \vec{S}_2

Table 1: Parameters in generating the template banks for the search.

As discussed in the research project proposal, the intrinsic parameters of a gravitational wave that determine the waveform are listed in Table 1. The dimension of the full parameter space would be $1 + 1 + 3 + 3 = 8$. Currently, the parameter space covered by `gstLAL` and `pyCBC` (the two CBC search pipelines) during O1 search is that each component mass is at least one solar mass $m_{1,2} \geq M_\odot$ and the total mass $2M_\odot \leq m_1 + m_2 \leq 100M_\odot$ [5]. Furthermore, we restrict the component angular momentum to be aligned with the orbital angular momentum, namely $S_{1,x} = S_{1,y} = S_{2,x} = S_{2,y} = 0$ and this reduces the dimension of the parameter space to just 4. The parameter space covered is graphically shown in Figure 3.

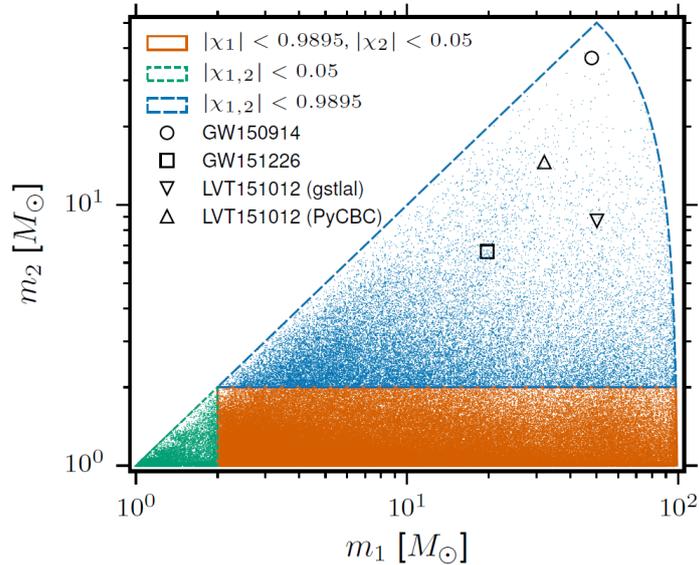


Figure 3: Parameter space currently covered in the first observing run (O1). Note that the convention $m_1 \geq m_2$ is adopted. (Figure taken from [4])

An intuitive approach to improve a gravitational wave search pipeline is to expand the parameter space covered by the template bank during the search, for instance, increasing the limit for total mass. However, we need to consider

1. Ability of waveform models to produce such a template for matched filtering
2. Extra computation cost in both preparing the template bank and matched filtering
3. Compatibility of existing codes

The main objective of this research is to investigate the possibility of including Intermediate-Mass Black Hole (IMBH) region into the parameter space of gstLAL for the second observing run. In the next sub-section, we will discuss various waveform models used by LIGO.

1.2 Generation of a template bank

To generate the templates necessary for matched filtering to search for gravitational wave signals from compact binary coalescence (CBC), different models are used by LIGO, as tabulated in Table 2.

Time domain	Frequency domain
SpinTaylorT2/T4/T5	IMRPhenomB
SEOBNR v2	IMRPhenomC
SEOBNR v3	SEOBNRv2_ROM_DoubleSpin
	SEOBNRv2_ROM_DoubleSpin_HI
	IMRPhenomD
	IMRPhenomPv2
	TaylorF2
	TaylorF2Redspin

Table 2: Different waveform models used for the construction of templates. SEOBNR stands for Spin Effective One Body Numerical Relativity; IMR-Phenom stands for Phenomenological Inspiral-Merger-Ringdown Model

Of course, the analytical waveform can be obtained, in theory, by solving the Einstein’s field equation analytically, which is very difficult. An alternative way is to solve the equation numerically (i.e. Numerical Relativity) but still it is computationally expensive. Post-Newtonian (PN) approximation, which is an expansion in a small parameter like v/c , is another option. However, PN approximation is only valid in the inspiral regime but in merger and ringdown regimes and hence it is not ideal to use PN approximation for high mass binary systems where the merger and ringdown phases contribute to the SNR significantly. Using Post-Newtonian approximation, the primary factor to the waveform of gravitational waves from a binary system of component masses m_1 and m_2 respectively is the chirp mass $\mathcal{M}_{\text{chirp}}$, which is defined as

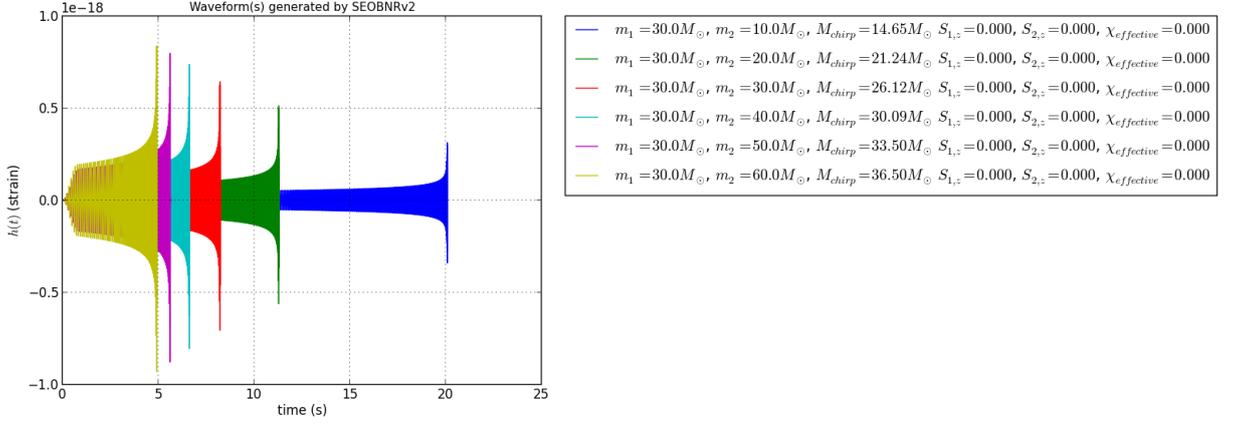
$$\mathcal{M}_{\text{chirp}} = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}, \quad (3)$$

and the secondary effect is from a quantity called effective spin parameter $\chi_{\text{effective}}$, which is defined as [2]

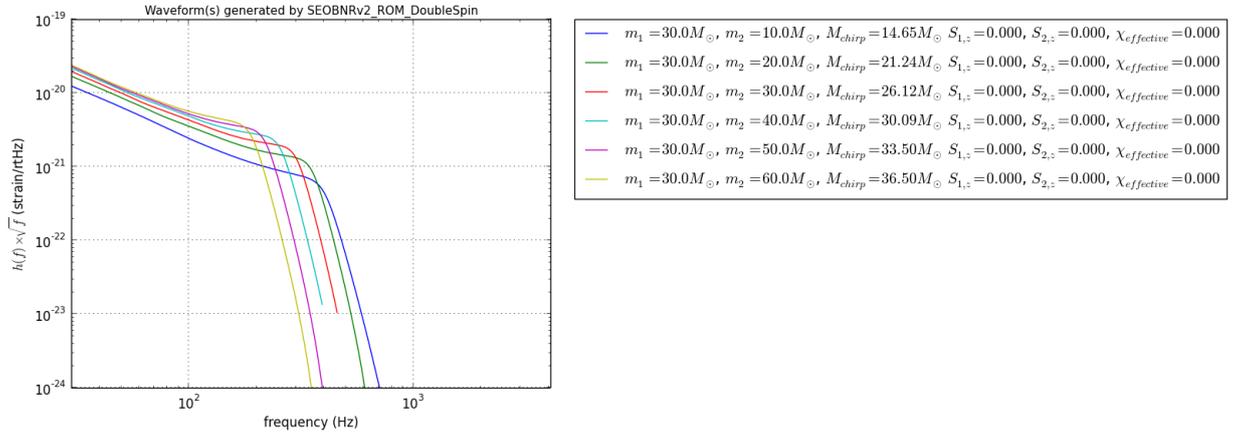
$$\chi_{\text{effective}} = \frac{m_1 \chi_1 + m_2 \chi_2}{m_1 + m_2}, \quad (4)$$

where the dimensionless spin parameter χ_i , where $-1 \leq \chi_i \leq 1$ can be calculated from the component spin \vec{S}_i and orbital angular momentum \vec{L}

$$\chi_i = \frac{c}{G} \frac{\vec{S}_i \cdot \vec{L}}{m_i^2} \quad (i = 1, 2).$$



(a) Generated by SEOBNRv2 in time domain



(b) Generated by SEOBNRv2_ROM_DoubleSpin in frequency domain

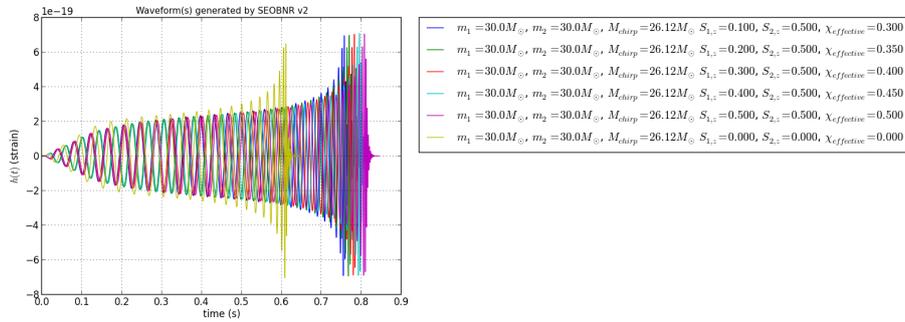
Figure 4: Waveforms from systems with zero spin and increasing chirp mass.

1.2.1 Effect of \mathcal{M}_{chirp} to the waveform

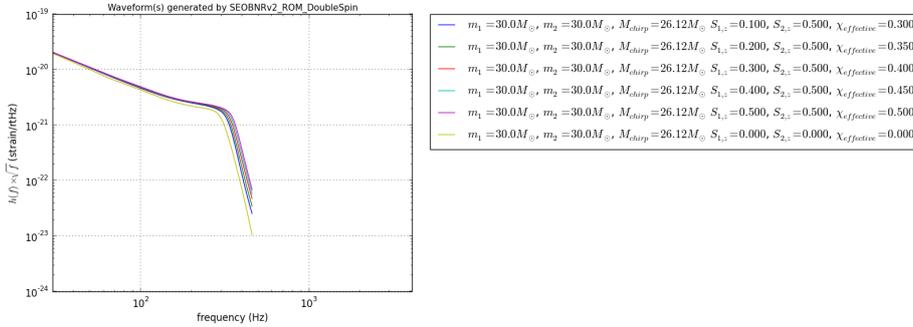
Figure 4a shows waveforms generated by SEOBNRv2 in time domain with zero spin and increasing component and chirp mass. One can see that the effects of increasing the chirp mass are that the amplitude of the wave would increase and the time that the signal would stay in the LIGO frequency band (In this simulation, the frequency band is 30 - 4096 Hz) would decrease. This is because systems with higher mass merge at a lower frequency and hence the time for which the signals stay in LIGO detection band would be shorter. One can see this effect more obviously when the waveforms are in frequency domain instead.

Figure 4b shows waveform generated by SEOBNRv2_ROM_DoubleSpin in frequency domain with zero spin and increasing component and chirp mass. As the chirp mass goes up, the amplitude of the signal increases but the frequency bandwidth decreases. Also note that the turning points in the figure correspond to the merger phase and as the chirp mass increases, the frequency at which the merger occurs decreases, which is consistent with our observation in the time domain.

1.2.2 Effect of $\chi_{\text{effective}}$ to the waveform



(a) Generated by SEOBNRv2 in time domain



(b) Generated by SEOBNRv2_ROM_DoubleSpin in frequency domain

Figure 5: Waveforms from systems with constant component masses and increasing effective spin.

Figure 5a shows waveform generated by SEOBNRv2 in time domain with constant component masses and increasing aligned spins and effective spins. One can see that the amplitude of gravitational waves roughly remains the same but the time in which they stay in the LIGO frequency band increases with the effective spin. Alternatively, we can observe from Figure 5b that shows waveforms generated by SEOBNRv2_ROM_DoubleSpin in frequency

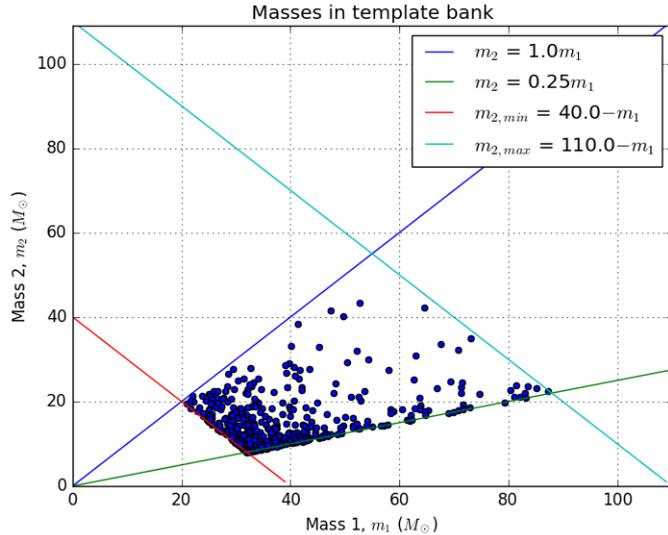


Figure 6: An example of template bank with discrete templates represented by a blue dot. Notice that the templates in low-mass region are denser than that in high-mass region. This is because a change in mass for a low-mass system will cause a more significant change in the waveform and hence more templates are required to maintain the same maximum fractional loss in SNR.

domain. Indeed, the frequency bandwidth of the signals increases with effective spin and the frequency that corresponds to the merger phase also increases with effective spin.

1.2.3 Template bank construction

In the previous sub-section, we discussed various waveform models to produce the model signals we need in matched filtering for searching gravitational waves from CBC. In reality, it is impossible to generate a template bank that completely covers the parameter space because that would require infinitely-many templates. Instead, we use a discrete template bank that contains finite number of templates (e.g. Figure 6) and we place templates into the template bank such that the maximum fractional loss in SNR would be bounded by some number we specify when constructing the bank.

1.3 gstLAL: The search pipeline

Currently, there are two search pipelines for CBC detection in LIGO, namely pyCBC which can only be used in offline mode and also gstLAL which can be used in both online and offline mode. Figure 7 shows a flow chart describing the work flow of gstLAL search pipeline.

In order to achieve a low latency for online search, various methods have been deployed to reduce the computational time. One of the methods adopted by gstLAL is called LLIOD (Low Latency Inspiral Online Detection) method that involves singular value decomposition (SVD) and critical sampling [2] so as to reduce the number of templates (filters) required for matched filtering and hence lower the computational time, which is essential for online search. In the next section, we will discuss singular value decomposition and its application by gstLAL.

2 Low-Rank Matrix Approximation Using Singular Value Decomposition

2.1 A Brief Introduction to Singular Value Decomposition

Singular Value Decomposition (SVD) decomposes any matrix $A_{m \times n}$ into a product of three matrices with nice properties, namely

$$A = U\Sigma V^T. \quad (5)$$

The matrix U is a $m \times m$ matrix and the columns of U are known as the **left singular vectors** \vec{u}_i ($i = 1, 2, \dots, m$). The matrix Σ is a $m \times n$ diagonal matrix with non-negative real diagonal elements known as the **singular values** σ of matrix A . The matrix V is a $n \times n$ matrix and the columns of V are known as the **right singular vectors** \vec{v}_j ($j = 1, 2, \dots, n$).

For the ease of discussion, we require that $m \geq n$. This requirement is merely arbitrary. It is easy to see that if a matrix A has a dimension of $m \times n$, then there will be n singular values σ_i ($i = 1, 2, \dots, n$). In addition, one can order these singular values

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0.$$

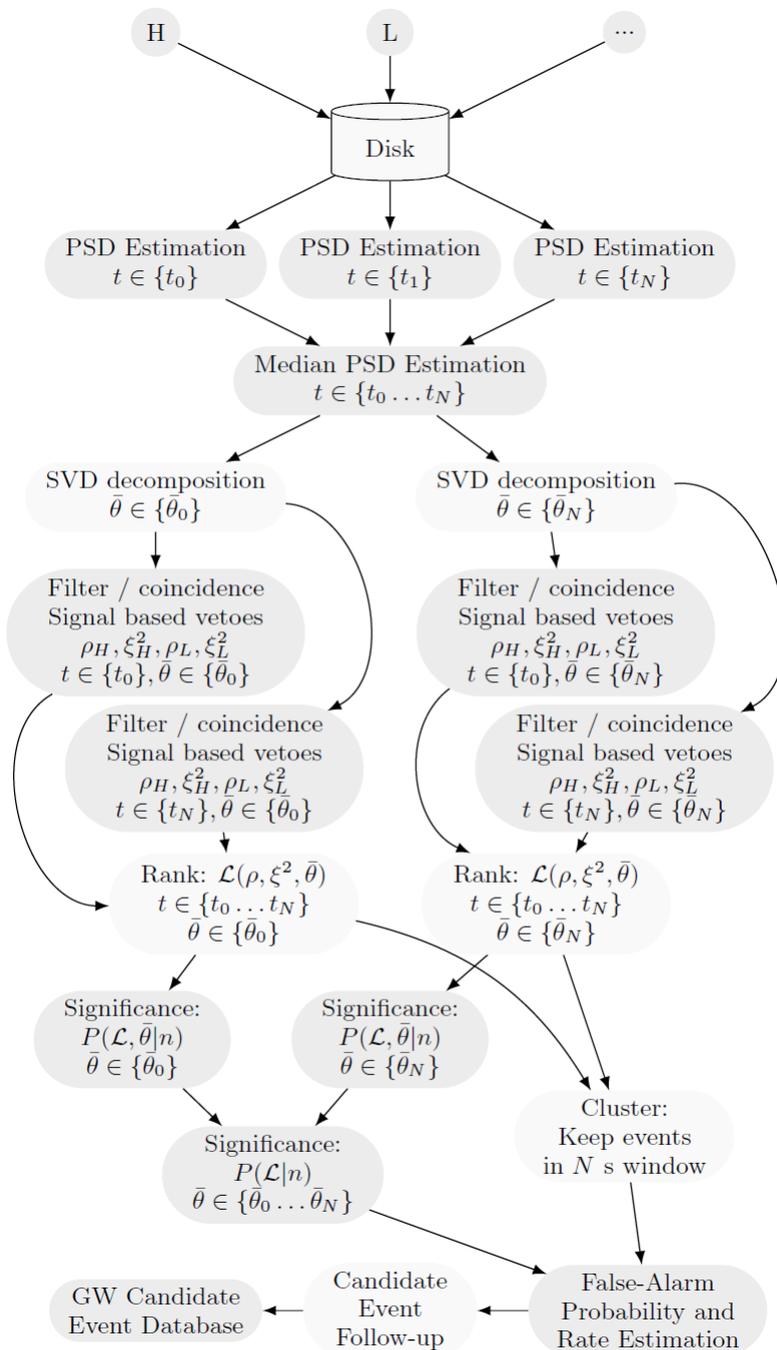


Figure 7: A flow chart describing the work-flow of gstLAL search pipeline. (Figure taken from [2])

In fact, we can expand the matrix A as an outer product, namely

$$A = \sum_{i=1}^n \sigma_i \vec{u}_i \vec{v}_i^T. \quad (6)$$

However, if the matrix A is row-deficient, namely $r = \text{rank}(A) < m$, there are only r non-zero singular values, with the remaining singular values all equal to zero. Mathematically,

$$\sigma_{r+1} = \sigma_{r+2} = \dots = \sigma_n = 0.$$

So we can simplify the equation 6 into

$$A = \sum_{i=1}^r \sigma_i \vec{u}_i \vec{v}_i^T. \quad (7)$$

We can already see that singular value decomposition helps us to eliminate redundant information during the matrix operations [1].

2.2 Approximating a matrix with a lower rank matrix

Suppose we seek for a matrix M_k that best approximates the matrix A with a lower rank k , namely $k = \text{rank}(M_k) < \text{rank}(A)$. A theorem in linear algebra (not proved here) says that such a matrix M_k is related to the singular value decomposition of A . In fact, the matrix M_k is merely the truncated SVD (cf Equation 6) up to the k^{th} term. Mathematically,

$$M_k = \sum_{i=1}^k \sigma_i \vec{u}_i \vec{v}_i^T, \quad (8)$$

and we call \vec{u}_i the orthonormal basis vector and $\sigma_i \vec{v}_i^T$ the reconstruction matrix.

2.3 Application of SVD in gstLAL

Before applying singular value decomposition to the templates, we first split the template bank into ‘split-banks’ that contain similar waveforms so as to maximize the compression and efficiency of SVD since SVD in itself is computationally expensive. Currently, we split the template bank according to the effective spin parameter (cf Equation 4) first and then sort by $\mathcal{M}_{\text{chirp}}$ (cf Equation 3). An example of split template bank is shown in Figure 8. This project will also investigate another way to split the template bank, namely

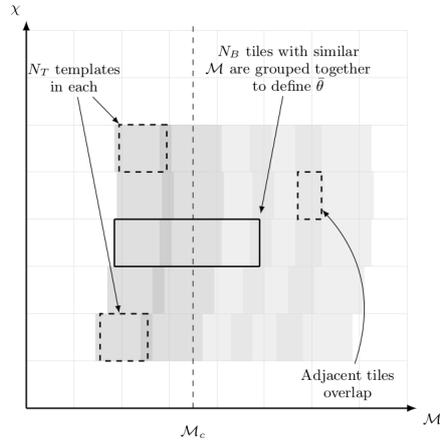


Figure 8: An example of how a template bank is split into ‘split-banks’ according to effective spin and chirp mass. (Figure taken from [2])

first splitting templates by chirp mass and then sorting by effective spin parameter as chirp mass is the primary factor that determines the waveform as discussed, so as to optimize the compression and background estimation. After splitting the template bank, for each ‘split-bank’ we divide the bank into many different time slices and perform singular value decomposition to different time slices with different appropriate frequencies to avoid over-sampling the low-frequency region [2].

3 Progress Review

3.1 Modifying the summary page of gstLAL

1.2 Search Information

Table 1.2.1. WALL TIME

	GPS Time	UTC MM/DD/YYYY HH:MM:SS
Start	1134450017	12/18/2015 05:00:00
End	1135054817	12/25/2015 05:00:00

Table 1.2.1. This table gives the start and end time of the search in GPS Time and UTC

Table 1.2.2. COINC EVENTS SUMMARY

DETECTORS	# COINC EVENTS
H1L1	20864

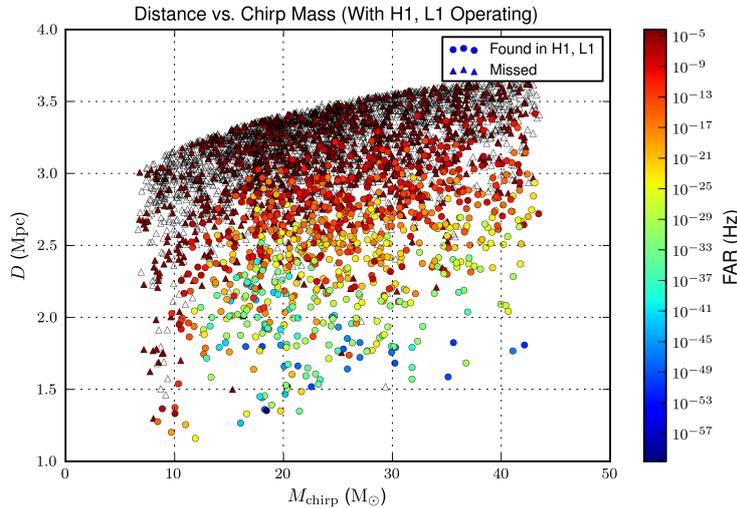
Table 1.2.2. This table gives the number of coincident events

Table 1.2.3. LIVE TIME SUMMARY

DETECTORS ON	LIVETIME (s) (d) (yr)	Duty Cycle
H1L1	210900.00 2.44 0.01	0.35

Table 1.2.3. This table gives the summary of the live time

(a) Modified search information page



(b) A sample of modified missed/found plots

Figure 9: *Top*: The duty cycle and wall time in GPS time and UTC are reported in the summary page. *Bottom*: The missed/found plots are colored according to the false alarm rate (FAR) associated with the injections.

Figure 9a shows the changes made to the search information of the summary page of `gstLAL`, with the wall time of the search as well as the duty cycle being reported. The modification to the code has been formatted as a git patch. In addition, Figure 9b shows an example of the modified missed/found plots in the `gstLAL` summary page. Previously, the missed injections were plotted as black dots and found injections were plotted as blue dots. Now, found injections are plotted as colored dots with colors according to their false-alarm rate (FAR) and missed injections are plotted as upright triangles. An injection is said to be missed if either the FAR assigned is higher than a given threshold value or the pipeline did not recover it at all. For the former case, the triangle is also coloured according to the false-alarm rate with the same color scale whereas for the latter case, the triangle is white-coloured.

4 Schedule

Table 3 shows the tentative schedule for the summer research project.

Date	Task
	Pre-arrival to Caltech
May 15, 2016	Project Proposal
May 15, 2016 - June 13, 2016	Learn more about gravitational waves and LIGO detectors
	Review Python, Numpy and Scipy
	Learn more about Linux/Unix environment
	Summer research in Caltech
June 14, 2016	Arrive at Caltech
Week 1 - 3	Learn how to run GstLAL on the LIGO computing cluster, understand the structure of the analysis pipeline, and the inputs, outputs and tunable parameters
July 06, 2016	First Progress Report
Week 4 - 6	Explore modifications to the analysis pipeline to improve its sensitivity, efficiency, speed and the quality of its outputs
Week 7 - 8	Perform high-statistics GstLAL runs on simulated noise, real noisy data, and simulated signals, to evaluate the results of the modifications
August 03, 2016	Second Progress Report
August 03, 2016	Abstract
Week 9 - 10	Compile and document all results; prepare final presentation and paper

Table 3: Summer research schedule

References

- [1] Kipp Cannon, Adrian Chapman, Chad Hanna, Drew Keppel, Antony C. Searle, and Alan J. Weinstein. Singular value decomposition applied to compact binary coalescence gravitational-wave signals. *Phys. Rev. D*, 82:044025, Aug 2010.
- [2] Cody Messick, Kent Blackburn, Patrick Brady, Patrick Brockill, Kipp Cannon, Sarah Caudill, Sydney J Chamberlin, Ryan Everett, Chad Hanna, G F Li, Duncan Meacher, Stephen Privitera, Hong Qi, B Sathyaprakash, Laleh Sadeghian, Gareth Thomas, Les Wade, Maddie Wade, and Alan Weinstein. Low-latency analysis framework for the prompt discovery of compact binary mergers in gravitational wave data. pages 1–15.
- [3] B. S. Sathyaprakash and Bernard F. Schutz. Physics, astrophysics and cosmology with gravitational waves. *Living Reviews in Relativity*, 12, 2009.
- [4] The LIGO Scientific Collaboration, the Virgo Collaboration, B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, and et al. Binary Black Hole Mergers in the first Advanced LIGO Observing Run. *ArXiv e-prints*, June 2016.
- [5] The LIGO Scientific Collaboration, the Virgo Collaboration, B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, and et al. Gw150914: First results from the search for binary black hole coalescence with advanced ligo. *Phys. Rev. D*, 93:122003, Jun 2016.
- [6] The LIGO Scientific Collaboration, the Virgo Collaboration, B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, and et al. Gw151226: Observation of gravitational waves from a 22-solar-mass binary black hole coalescence. *Phys. Rev. Lett.*, 116:241103, Jun 2016.
- [7] The LIGO Scientific Collaboration, the Virgo Collaboration, B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, and et al. Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett.*, 116:061102, Feb 2016.