Searching for Gravitational Waves from the Coalescence of High Mass Black Hole Binaries

Ka-Lok Lo
Department of Physics, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong SAR

Surabhi Sachdev, Kent Blackburn, and Alan Weinstein
LIGO Laboratory, California Institute of Technology, Pasadena, California 91125, US

(LIGO Scientific Collaboration)

(Dated: November 18, 2016)

`gstlal` is a search pipeline used by LIGO to search for gravitational waves from the coalescence of compact binaries (neutron stars and/or black holes), which can be operated in on-line and/or off-line configuration. Currently, the parameter space covered by `gstlal` consists of binaries with component masses $m_1, m_2 \geq 1M_\odot$ ($M_\odot$ is the solar mass) and total mass $2M_\odot \leq m_1 + m_2 \leq 100M_\odot$. In this project, we aim to increase the upper bound of the total mass to $600M_\odot$ by including intermediate mass black holes (IMBH) region into the parameter space. Simulated signal injections with real noises were used to evaluate the sensitivity of `gstlal` when decreasing the lower frequency cut-off so as to observe signals from IMBH. Also, the drop in performance of the search pipeline was measured as we added more templates in the IMBH region and we experimented various techniques such as different bank splitting and $h(t)$ gating schemes to increase the performance and sensitivity. This would enable us to detect more gravitational wave signals from the coalescence of black holes with higher masses in real time gravitational wave detection.

LIGO DCC Number: LIGO-T1600261

I. BACKGROUND

A. Introduction to Gravitational Waves

The first theoretical prediction of the existence of gravitational waves was made by renowned physicist Albert Einstein in 1916. In his elegant Einstein's field equation, he realized that under the weak field approximation, there exist wave solutions, namely transverse waves of oscillating gravitational field travelling at the speed of light [1]. Under Einstein's theory of general relativity, there are two types of polarization, namely plus + and cross × polarization, which stretch and squeeze the spacetime into the shapes as shown in Figure 1 [2].

Currently, scientists propose that there are four classes of gravitational waves (GWs), namely continuous GWs from a massive spinning object with imperfection in symmetry (Figure 2); compact binary inspiral GWs from the coalescence of two massive and dense objects, which will be the focus of this study (Figure 3); stochastic GWs from the background, possibly from the big bang (Figure 4); burst GWs, from an unexpected source (Figure 5). In general, the magnitude of gravitational wave is very small, of the order of 10^{-21} to 10^{-28}. This makes the detection of gravitational waves seemingly an impossible mission.

B. Detections of Gravitational Waves

The first indirect detection of gravitational waves was made by Russell Hulse and Joe Taylor where PSR B1913+16, a pulsar and a neutron star orbiting around

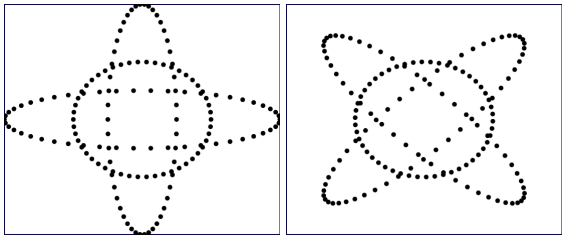


FIG. 1: (Left) The effect brought about by plus-polarized gravitational waves to a circular ring. It stretches the ring into an ellipse of equal area horizontally for the first half of a period and stretches the ring into an ellipse vertically for the second half of a period; (Right) The effect brought about by cross-polarized gravitational waves. The effect is similar to that of plus-polarized waves by rotating an angle of $\frac{\pi}{4}$. (Figure taken from [2])

the center of mass of the system lose energy due to the emission of gravitational waves [1]. However, direct detection of gravitational waves was not possible before LIGO, which stands for Laser Interferometer Gravitational-wave Observatory.

A LIGO detector, as shown in Figure 6, is a Fabry-Pérot Michelson interferometer consisting of two orthogonal arms with lengths 4 km each. A laser beam passes through a beam splitter and is separated into two beams travelling in two arms respectively. They are then reflected by mirrors on the test masses and eventually rejoin and the signal is recorded by the photo-detector. If the length of the two arms are identical, namely

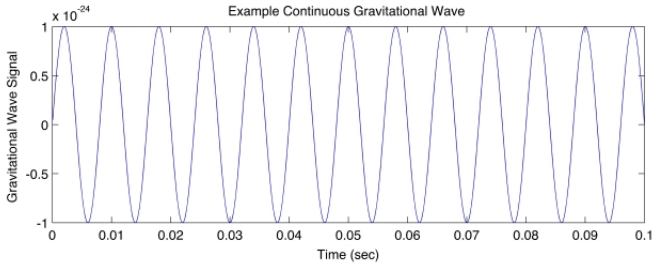


FIG. 2: Example of continuous gravitational wave. One can see that the frequency and the amplitude are constant in time. (Figure taken from [3])

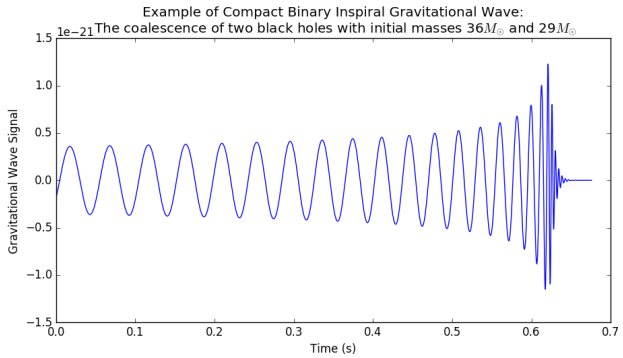


FIG. 3: Example of compact binary inspiral gravitational waves. In particular, the signal is generated by the merger of two black holes with initial masses $36M_\odot$ and $29M_\odot$.

$L = L_X = L_Y$, then destructive interference occurs and the signals from the two arms cancel each other out perfectly. However, if there is a slight difference in the two lengths $\Delta L = |L_X - L_Y| \neq 0$, the two signals will not destructively interfere completely, and thus a signal can be registered by a photo-detector and we can calculate the strain $h = \frac{\Delta L}{L}$.

In reality, LIGO detectors are subject to different types of noises, such as ground vibration, thermal noise, quantum noise and gravity gradient noise [2]. Various efforts, such as installing seismic sensors around the detector and data analysis techniques, have been made to facilitate the observation of gravitational waves and to distinguish gravitational wave signals from noises. These noises limit the sensitivity of the LIGO detectors [1] to observe gravitational waves. Fortunately, some types of signals emitted from a compact binary star system are detectable by the LIGO detectors. Possible combinations of a compact binary star system include:

1. Binary Black Holes system (BBH)
2. Neutron Star-Black Hole system (NSBH)
3. Binary Neutron Stars system (BNS)

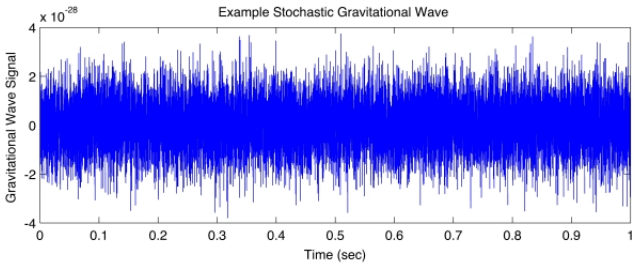


FIG. 4: Example of signal of stochastic gravitational waves. (Figure taken from [4])

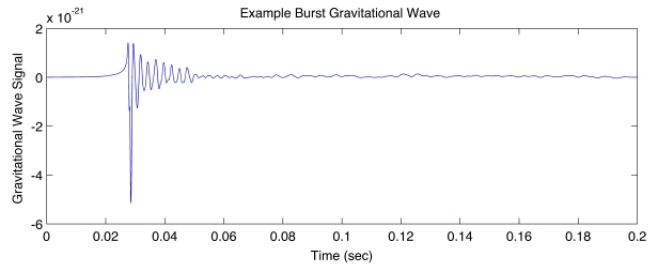


FIG. 5: Example of signal of burst gravitational waves. (Figure taken from [5])

In particular, the signals emitted in the inspiral phase of binary neutron stars system, the inspiral, merger and ring-down phases of the coalescence of compact binary black holes with each black hole of mass greater than 5 solar mass, lie in the detection band of the LIGO detectors.

During the run from 2002-2010, the initial LIGO (iLIGO) was not able to detect any gravitational wave signals. The two detectors in Livingston and Hanford were upgraded to advanced LIGO (aLIGO), with improved sensitivity to different frequencies of signals as shown in Figure 7. Shortly after the upgrade, on 14 September, 2015, LIGO first directly observed a gravitational wave signal emitted by the coalescence of a binary black hole system with with initial masses $36M_\odot$ and $29M_\odot$, which is also known as GW150914 event. Another detection GW151226 was also made during the first observing run (O1). Figure 8 shows the waveform of the signal detected in GW150914. It is worthwhile to mention that the GW150914 event could not have been detected by iLIGO because of the insufficient sensitivity. This observation directly proves the existence of gravitational waves and shows the capability of aLIGO detectors to observe gravitational waves. Also, it opens up a new chapter in gravitational-wave astronomy.

Low latency on-line searches, which can detect a signal within a few minutes of data acquisition, allow electromagnetic counterparts such as gamma-ray telescopes to observe the binary star system within a short notice and also provide feedback when the search sensitivity has

dropped, implying that there might be a problem in the detector [6].

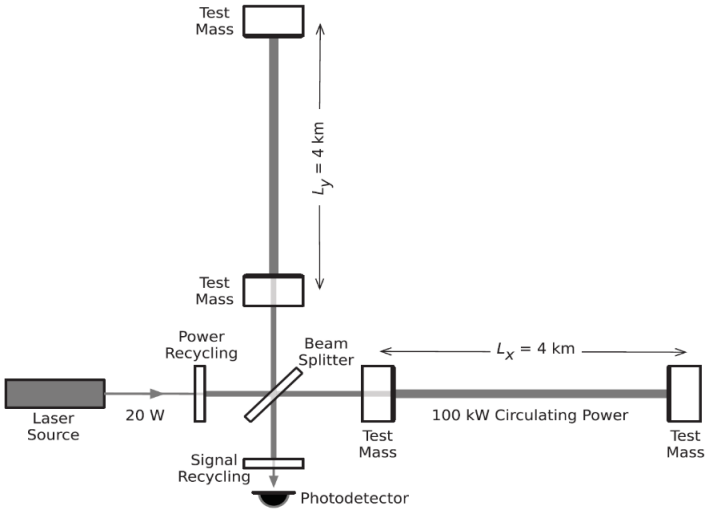


FIG. 6: A schematic diagram of a LIGO detector. LIGO detector is more sensitive to plus-polarized GWs than those cross-polarized. (Figure taken from [1])

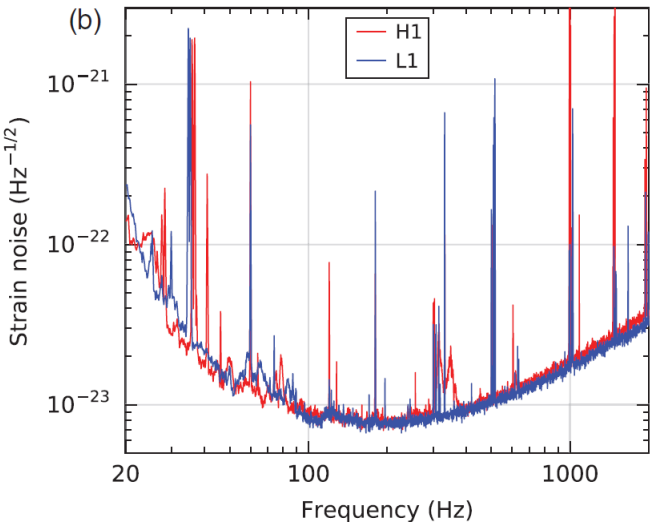


FIG. 7: Sensitivity of the LIGO detector to different frequencies of signal. (Figure taken from [1])

C. `gstlal`: the Search Pipeline

`gstlal`, which is derived from the multi-media library GStreamer and LIGO Algorithm Library [6], is one of the search pipelines for detecting compact binary coalescences (CBC) designed to operate at low latency for rapid EM follow-up. `gstlal` looks for the coalescence of compact binary systems mentioned above, BBH, NSBH and BNS.

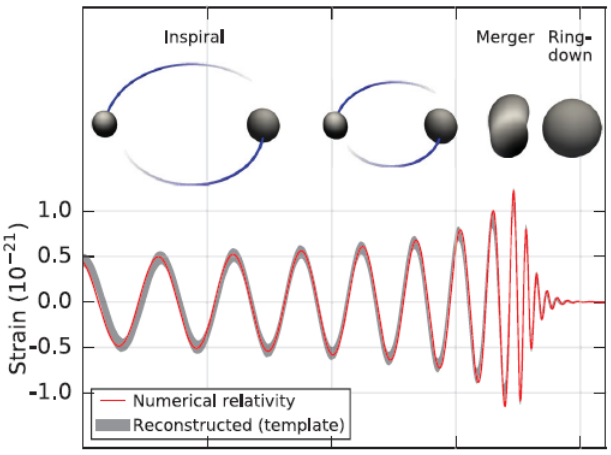


FIG. 8: Inspiral, merger and ring-down phases of the coalescence of compact binary star system. In particular, this waveform comes from GW150914. (Figure taken from [1])

Figure 9 outlines the `gstlal` search pipeline. First, the data from Hanford (H1) and Livingston (L1) enter the pipeline. The power spectral density (PSD) is estimated for different portions of time $t_0, t_1, ..., t_N$ and using the results obtained for median PSD estimation. Matched filtering, which is the most computationally expensive part of the search, is then performed on the data received against the pre-calculated template bank of gravitational wave waveforms. It will be discussed in section I D 1. Currently, the template bank used by `gstlal` (the uberbank) consists of simulated signals from the coalescence of compact binary systems with component masses $m_1, m_2 \geq 1M_\odot$ and total mass $2M_\odot \leq m_1 + m_2 \leq 100M_\odot$. Table I lists the intrinsic parameters of a gravitational wave that determine its waveform and thus are used in generating the template banks for the search and Figure 10 shows the parameter space covered by `gstlal`. The dimension of the full parameter space would be $1+1+3+3 = 8$ [7]. We restrict the component angular momentum to be aligned with the orbital angular momentum along z direction (by convention), which gives us $S_{1,x} = S_{1,y} = S_{2,x} = S_{2,y} = 0$ and this reduces the dimension of the parameter space to just 4. It is also worthwhile to mention that the parameters found by `gstlal` may not coincide with the signal parameters, for which full parameter estimation analysis will be carried out afterwards, due to the discrete nature of the template bank [8]. Table II lists the extrinsic parameters of a gravitational wave that determine its amplitude as the amplitude observed depends on the observer's location respect to the binary system. Both intrinsic and extrinsic parameters will be estimated by a complete parameter estimation at a later stage.

Singular Value Decomposition (SVD) is performed before matched filtering on the waveforms in template bank to save computational costs and it will be discussed in

section I F. Signal-to-Noise Ratio (SNR, ρ) is calculated during the process of matched filtering. If a candidate has a SNR value greater than a threshold value $\rho_{\rm threshold}$, then it is categorized as a "trigger". Also, the chi-squared χ^2 of the candidate will be calculated to evaluate the statistical significance of the event.

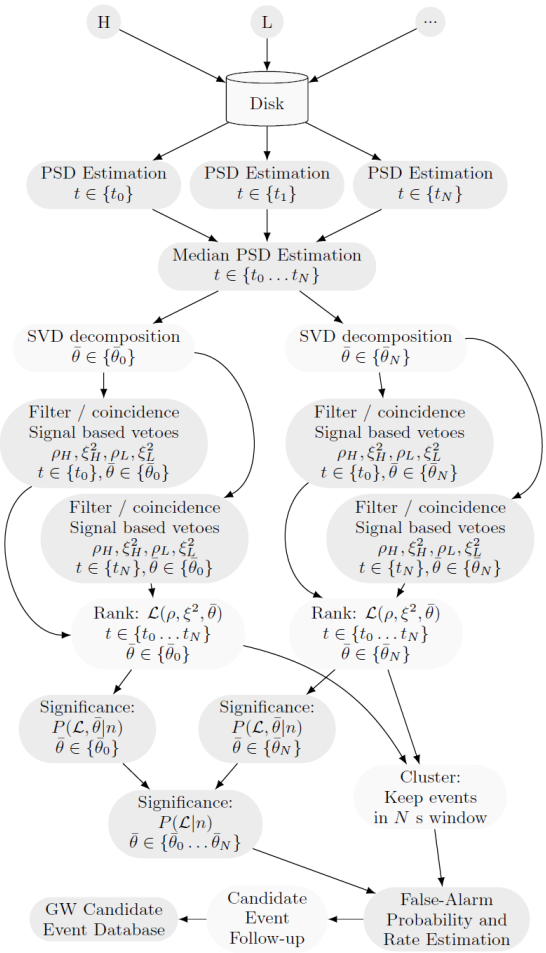


FIG. 9: A flow chart describing the work-flow of `gstlal` search pipeline. (Figure taken from [6])

component masses	m_1, m_2
component spins	$\vec{S}_1 = (S_{1,x}, S_{1,y}, S_{1,z})$, $\vec{S}_2 = (S_{2,x}, S_{2,y}, S_{2,z})$

TABLE I: Intrinsic parameters of a gravitational wave. The waveform of a GW signal depends on these intrinsic parameters.

luminosity distance	d
time of coalescence	$t_{\rm c}$
phase	ϕ
sky location: right ascension	α
sky location: declination	δ
inclination with respect to the line of sight	ι
polarization	ψ

TABLE II: Extrinsic parameters of a gravitational wave. The amplitude of a GW signal depends on these extrinsic parameters.

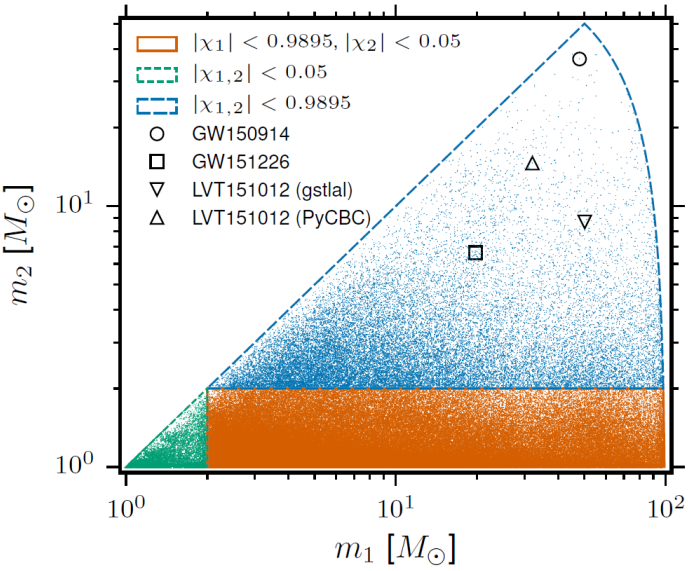


FIG. 10: Parameter space currently covered in the first observing run (O1). Note that the convention $m_1 \geq m_2$ is adopted. (Figure taken from [9])

D. Data Analysis Techniques in Searching Gravitational Waves from Compact Binary Coalescence

The advanced LIGO 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 7 shows the amplitude spectrum density of both aLIGO detectors near the GW150914 event which provides an indication of instrument noises.

The top row in Figure 11 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.

1. 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 pregenerated 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

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

Suppose we are filtering the data $d(t)$ with some normalized templates $\{h_i^c(t)\}$, then the complex matched filtering response $z_i(t)$ is given by [6]

$$\begin{aligned} z_i(t) &= x_i(t) + iy_i(t), \\ &= 4\int_0^\infty df \frac{\tilde{h_i^c}^*(f)\tilde{d}(f)}{S_n(f)} e^{2\pi i f t}, \end{aligned}$$

where $x_i(t)$ and $y_i(t)$ are the matched filtering response to h_+ and $h_\times$ respectively and $S_n(f)$ is the power spectral density of noise. The SNR is simply the modulus of $z(t)$, namely

$$\text{SNR} = \rho = |z| \tag{1}$$

Higher the SNR, higher the similarity between the signal and the template. Since parameters of 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.

E. Generation of Gravitational Wave Waveform

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

Of course, the analytical waveform can be obtained, in theory, by solving the Einstein's field equation analytically, but we do not know how to do that. 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 not in merger and ringdown regimes

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

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

$$\mathcal{M}_{\text{chirp}} = \frac{(m_1 m_2)^{3/5}}{(m_1+m_2)^{1/5}}, \tag{2}$$

and the secondary effect is from a quantity called effective spin parameter $\chi_{\text{effective}}$, which is defined as [6]

$$\chi_{\text{effective}} = \frac{m_1\chi_1 + m_2\chi_2}{m_1+m_2}, \tag{3}$$

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 $\vec{L}$

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

1. Effect of $\mathcal{M}_{\text{chirp}}$ to the waveform

Figure 12a 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, which is known as the chirp time would decrease (In this simulation, the frequency band is 30 - 4096 Hz). 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 is shorter. One can see this effect more obviously when the waveforms are in frequency domain instead.

Figure 12b shows waveforms generated by SEOBNRv2_ROM_DoubleSpin in frequency domain with zero

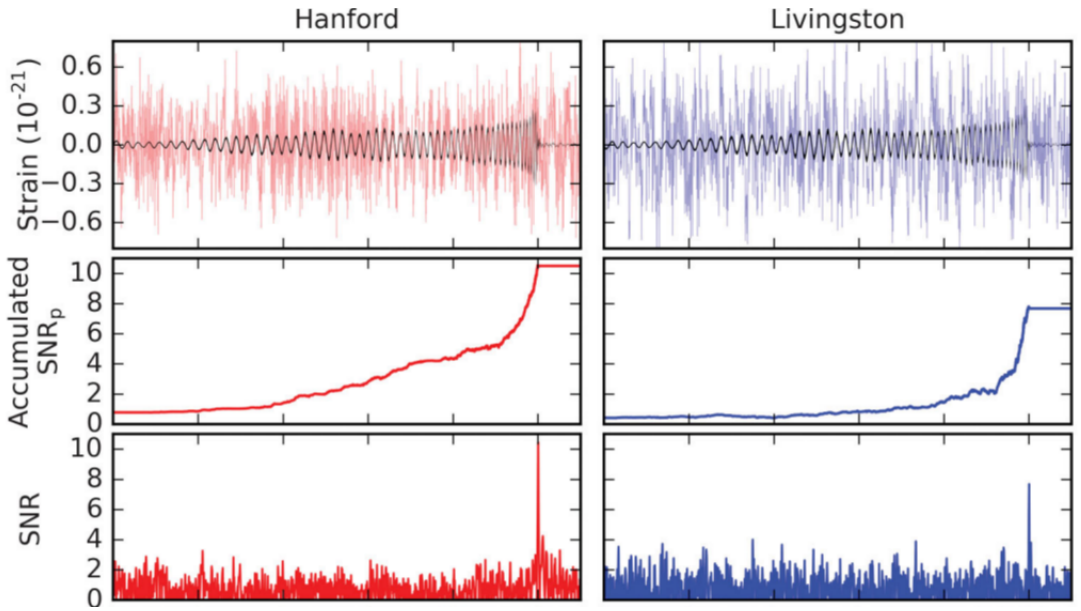


FIG. 11: *Top row:* Bandpass-filtered (30-600Hz) strain data from Hanford (Left, Red in color) and Livingston (Right, Blue in color) during the GW151226 event, where gravitational wave signal can hardly be seen. The black curve on both plots is the best-match template found. *Middle row:* Time series of accumlated 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 [10])

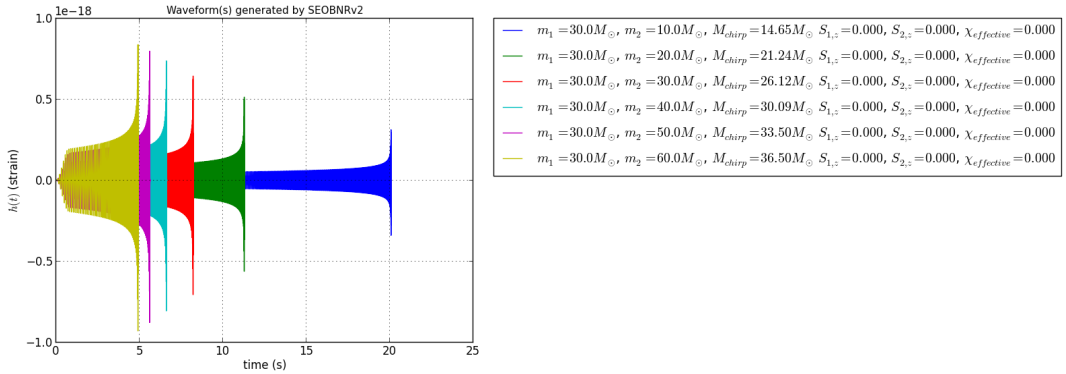
spin and increasing component and chirp mass. As the chirp mass goes up, the amplitude of the signal increases but the frequency bandwidth decreases. Also note that the turning points in the figure correspond to the merger phase and as the chirp mass increases, the frequency at which the merger occurs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

2. *Effect of $\chi_{\text{effective}}$ to the waveform*

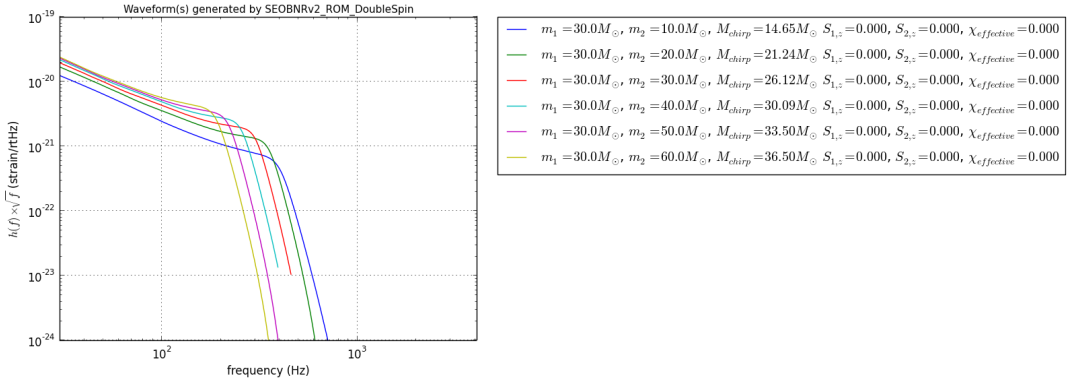
Figure 13a shows waveforms generated by SEOBNRv2 in time domain with constant component masses and increasing aligned spins and effective spins. One can see that the amplitude of gravitational waves roughly remains the same but the chirp time increases with the effective spin. Alternatively, we can observe from Figure 13b that shows waveforms generated by SEOBNRv2_ROM_DoubleSpin in frequency 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.

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

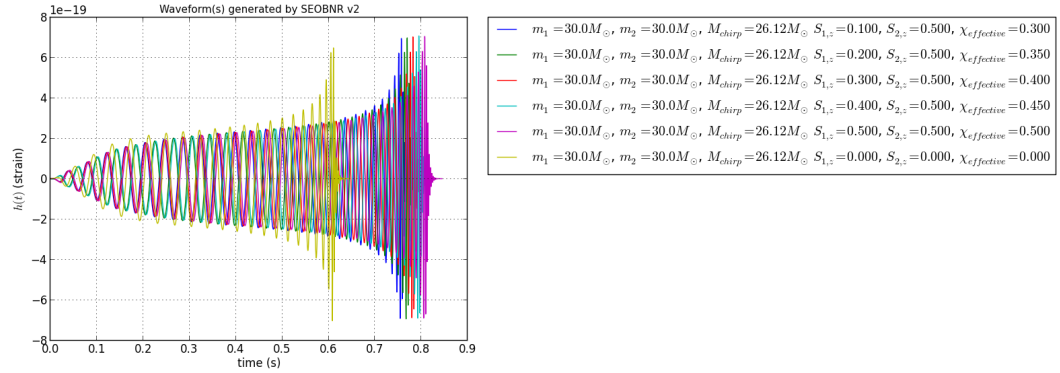


(a) Waveforms generated by SEOBNRv2 in time domain

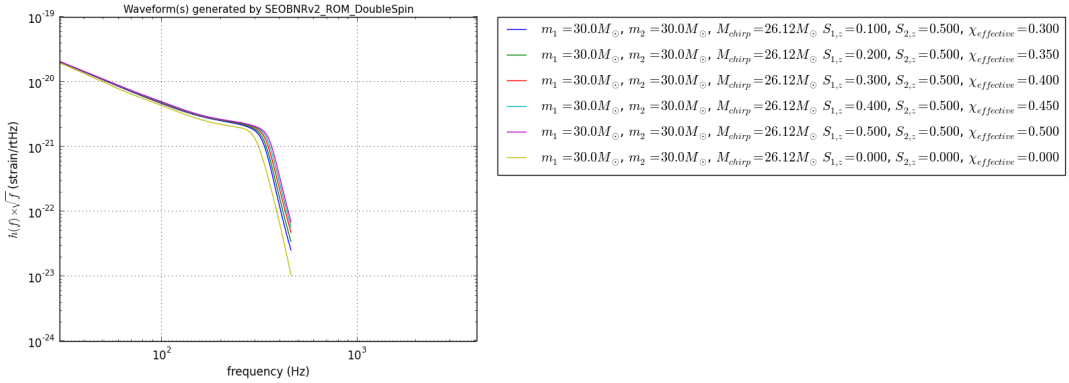


(b) Waveforms generated by SEOBNRv2_ROM_DoubleSpin in frequency domain

FIG. 12: Waveforms from systems with zero spin and increasing chirp mass.



(a) Waveforms generated by SEOBNRv2 in time domain



(b) Waveforms generated by SEOBNRv2_ROM_DoubleSpin in frequency domain

FIG. 13: Waveforms from systems with constant component masses and increasing effective spin.

that would require infinitely-many templates. Instead, we use a discrete template bank that contains finite number of templates (e.g. Figure 14) 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. We call template banks that satisfy this condition effectual template banks.

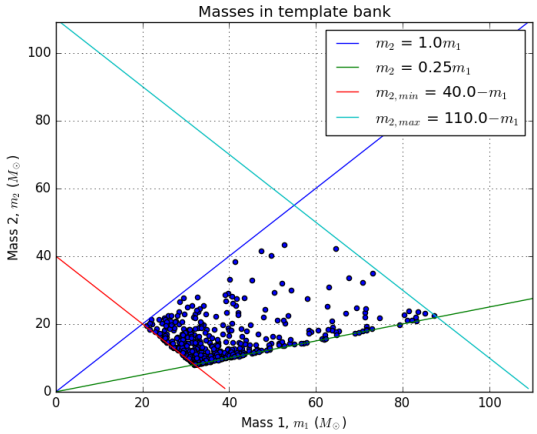


FIG. 14: An example of template bank with discrete templates represented by a blue dot. Notice that the templates in low-mass region are denser than that in high-mass region. This is because low mass systems are in band for a longer duration, therefore slight changes in mass can cause the waveforms to change the phase, resulting in low match with original waveform. As a result, more templates are required to maintain the same maximum fractional loss in SNR

F. Singular Value Decomposition

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 [6] 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 section I F 3, we will discuss singular value decomposition and its application by `gstlal`.

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

$$A = U\Sigma V^T. \tag{4}$$

The matrix U is a $m\times m$ matrix and the columns of U are known as the **left singular vectors** $\vec{u}_i$ $(i = 1, 2, ..., m)$. 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** $\vec{v}_j$ $(j = 1, 2, ..., n)$.

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 $(i = 1, 2, ..., n)$. In addition, one can order these singular values

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

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

$$A = \sum_{i=1}^{n} \sigma_i \vec{u}_i \vec{v}_i^T. \tag{5}$$

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,

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

So we can simplify the equation 5 into

$$A = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^T. \tag{6}$$

We can already see that singular value decomposition helps us to eliminate redundant information during the matrix operations [11].

2. *Low-Rank Matrix Approximation (LRMA) Using Singular Value Decomposition*

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 5) up to th k^{th} term. Mathematically,

$$M_k = \sum_{i=1}^{k} \sigma_i \vec{u}_i \vec{v}_i^T, \tag{7}$$

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

3. *Application of SVD in* `gstlal`

Before applying singular value decomposition to the templates, we first split the template bank into 'split-banks' and put some number of split-banks together into

a SVD sub-bank that contain similar waveforms so as to maximize the compression and efficiency of SVD since SVD in itself is computationally expensive. Currently, the default scheme to splitting a template bank is binning by $\chi_{\text{effective}}$ (cf Equation 3) first and then sorting by $\mathcal{M}_{\text{chirp}}$ (cf Equation 2). An example of split template bank is shown in Figure 15.

After splitting the template bank, for each sub-bank we divide the bank into many different time slices and perform singular value decomposition on different time slices with different appropriate frequencies to avoid oversampling the low-frequency region. This technique is known as multi-banding [6] and this saves computational resources. Figure 16 shows examples of the output of LLOID method applied to templates. Templates are divided into 10 time slices, each sampling at a different sampling rate and singular value decomposition is performed on these slices.

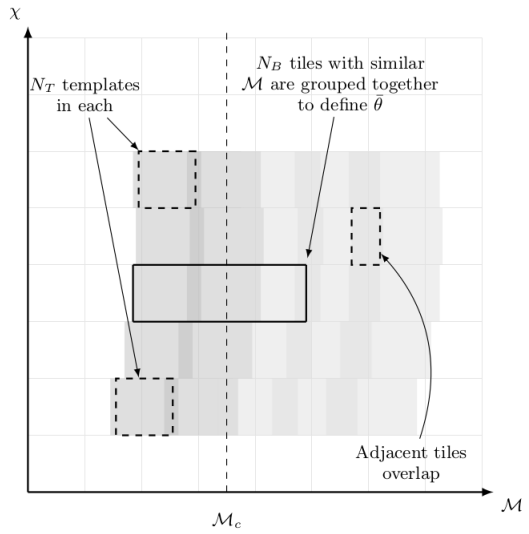


FIG. 15: Default bank splitter (Figure taken from [6])

II. OBJECTIVES

The main objective of this research project is to investigate the viability of combining template bank of intermediate mass black holes (IMBH for short) with the current template bank (the uberbank) used by `gstlal`, the search pipeline that can be operated in on-line or off-line configuration.

Figure 10 shows the parameter space currently covered by `gstlal`, which consists of binary systems (Binary Black Hole, Binary Neutron Star, Neutron Star-Black Hole) with component masses $m_1, m_2 \geq 1M_\odot$ and total mass $2M_\odot \leq m_1 + m_2 \leq 100M_\odot$. Intermediate mass black holes refer to black holes with masses greater than $100M_\odot$ and the searches for IMBH are now separate and done off-line. If it is possible to include IMBH into the

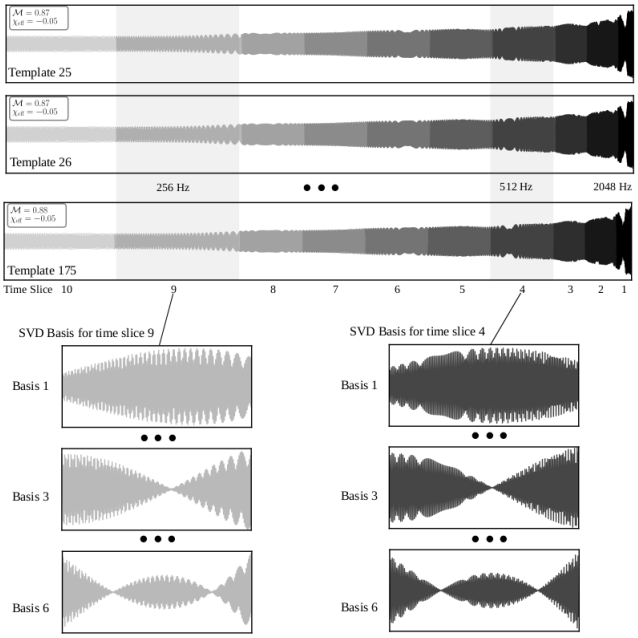


FIG. 16: Output examples of LLIOD (Low Latency Inspiral Online Detection) method. (Figure taken from [6])

parameter space covered by `gstlal`, this would allow us to search for more binary black hole systems in real time.

However, there are a lot of considerations when adding IMBH into the existing template bank of `gstlal`. For example, we need to consider

1. Change in sensitivity of the pipeline for different astrophysical sources
2. Change in performance and extra computation cost in both preparing the template bank and matched filtering
3. Ability of waveform models to produce such a template for matched filtering
4. Compatibility of existing codes

We will address these considerations in the following sections.

III. TUNING SEARCH PARAMETER: f_{min}

A. Motivation of Tuning Search Parameter f_{min}

In the separate IMBH search, the starting frequency is lowered from 30 Hz (used by `gstlal` uberbank search) to 15 Hz. As higher the chirp mass of a binary system, the frequencies that correspond to the inspiral and merger phase decrease, this implies that if we do not lower the starting frequency when searching for IMBH, then most of the signals will be cut-off. As an example, for a binary

system with total mass $M = 600M_\odot$, the ring-down frequency of the gravitational wave emitted from the system is approximately 19.8 Hz and we will not be able to observe the signal if we set the starting frequency as 30 Hz.

When $h(t)$ channel only contains signals, lowering starting frequency will have a gain in SNR simply because we can now observe more parts of the signals which were previously cut off. However, this may not be true in reality because aLIGO detectors are subject to more noise at lower frequencies. Therefore, it is likely that we are adding not enough in SNR but adding a lot to our computational costs when lowering the starting frequency. It is also likely that we would have a loss in SNR when lowering the starting frequency. This is undesirable when we want to expand the parameter space of `gstlal` to IMBH region as this will decrease the sensitivity of the search pipeline for binary systems with lower masses.

We can answer the question of whether the starting frequency should be lowered altogether for uberbank and IMBH regions or use separate starting frequencies, by doing SNR calculations, which compute the SNR of various injections that compare injections with themselves, weighted with a given reference power spectrum density (PSD) [12]. We investigate the SNR of GW signals from binary neutron star systems, binary black hole systems and neutron star-black hole systems with starting frequencies of 15, 20 and 30 Hz.

B. Results of SNR Calculations

Figure 17a shows a plot of the fractional change of SNR ρ with starting frequency lowered from 30 Hz to 20 Hz as a function of chirp mass $\mathcal{M}_{\text{chirp}}$ of the injection. We can see that there is a significant increase in the SNR for systems with chirp mass starting from $\mathcal{M}_{\text{chirp}} \approx 60M_\odot$. The highest fractional increase in SNR observed here is roughly less than 500%. Figure 17b shows a similar plot as Figure 17a with starting frequency lowered from 30 Hz to 15 Hz. A similar observation can be seen from the figure that there is a significant change in the SNR for systems with higher masses and the highest fractional increase in SNR observed is roughly higher than 500%. Figure 17c shows the fractional change in SNR when lowering the starting frequency further from 20 Hz to 15 Hz. We see a smaller change in SNR in this case, with the highest fractional increase in SNR registered around 25%.

C. Conclusions and Future Work for Tuning Search Parameter f_{min}

From the results of SNR calculations, we can conclude that it is worthwhile to lower the starting frequency for the higher mass region ($\mathcal{M}_{\text{chirp}} \gtrapprox 60M_\odot$) in order to get the extra SNR because of the significant increase in SNR. Although the changes in SNR when further lowering the starting frequency from 20 Hz to 15 Hz are diminishing, this will not add a lot of computational cost since there are fewer templates in that region as we have mentioned before. As for the lower mass region ($\mathcal{M}_{\text{chirp}} \lessapprox 60M_\odot$), it is not worthy to lower the starting frequency because this will add a lot more computational costs to the search but we only get tiny increase in SNR.

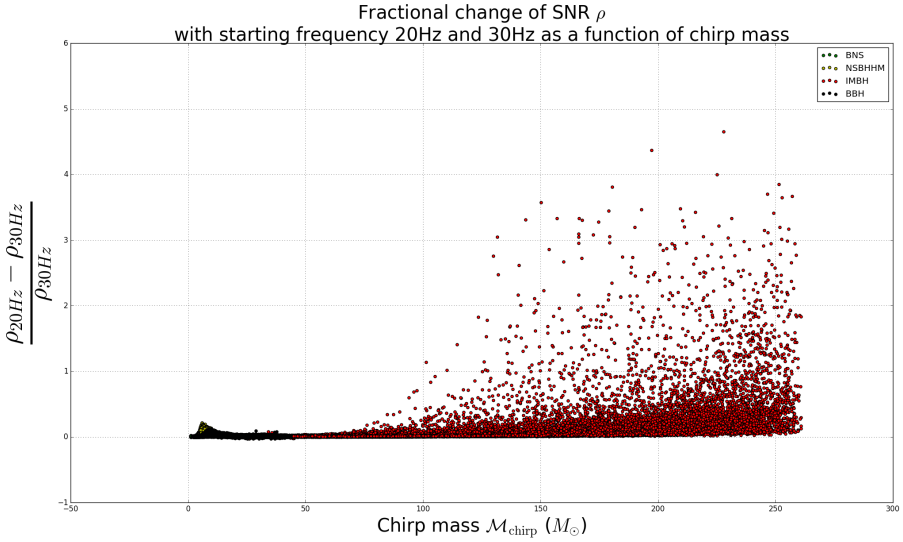
Since we have concluded that it is worthwhile to lower the starting frequency for the higher mass region, `gstlal` has to be modified in order to accept two different starting frequencies for different regions of chirp mass within a SVD sub-bank. Modifications to the `gstlal` codes have been done and we will test the implementation with actual template bank and signals to see the changes in SNR.

IV. $h(t)$ GATING

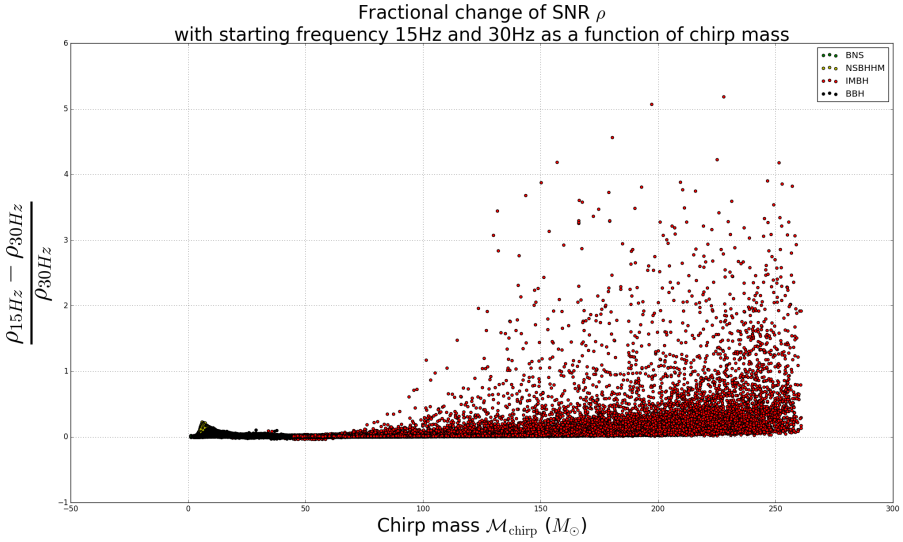
A. Motivation of $h(t)$ gating

Under such a high sensitivity, the advanced LIGO detectors are subject to tremendous amount of noise, where the noise could be fundamental, instrumental or environmental. In off-line analysis, data quality vetoes are available where poor data is flagged. Since the data $h(t)$ is whitened and thus any two segments of $h(t)$ are uncorrelated, segments with poor data are simply replaced by zeros as shown in Figure 17 [6]. However, during on-line analysis, such information is not available. $h(t)$ gating is one of the techniques adopted by `gstlal` to eliminate short transient instrumental noise fluctuations that can cause unreasonably high values of SNR, also known as glitches (e.g. Figure 18a), in the data. These glitches can mimic a real gravitational wave signal and cause a false alarm trigger.

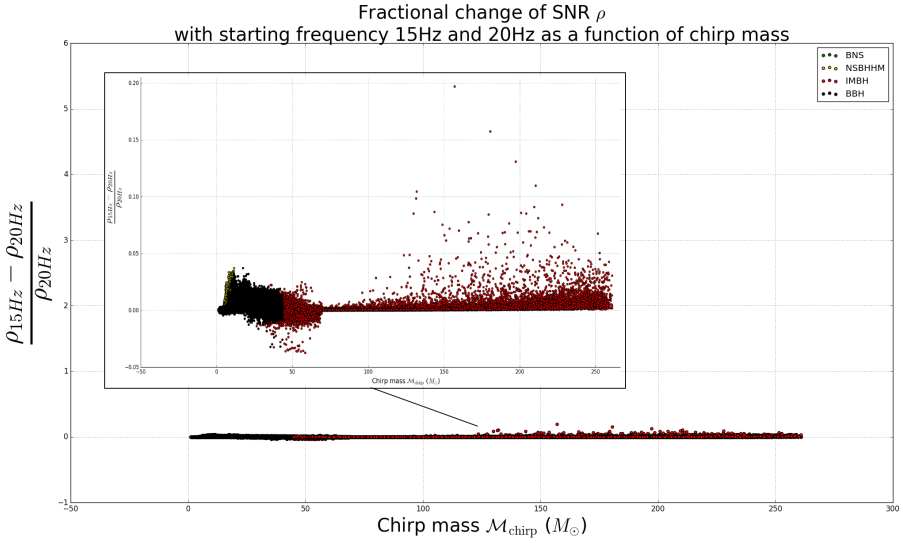
With the presence of glitches, the whitened $h(t)$ would have values higher than the expected values from the coalescences of binary systems that `gstlal` is aiming to detect. $h(t)$ gating removes a segment of signals, centered at the peak of the glitch with a window of 0.25 s on each side, in a sense that it is all replaced by zeros when the whitened $h(t)$ is greater than a threshold value as shown in Figure 17. Currently, the threshold value is taken as 50 [13], which is the lowest value before `gstlal` starts missing injections in the high mass region. However, since the amplitude of signal increases with the chirp mass of the binary system, real signals coming from higher mass binary systems may have amplitude higher than the threshold value that is currently used, causing `gstlal` to be insensitive to these higher mass binary systems. Even worse, the chirp time of a signal decreases with the chirp mass of the binary system. Figure 18b shows an example of GW signals from an IMBH binary system. As a result, glitches resemble to those signals and cause more false alarm triggers. An naive approach to solve this issue is to raise the threshold value such that those louder signals would not be gated out. Yet



(a) Fractional change of SNR ρ with starting frequency 20 Hz and 30 Hz as a function of chirp mass



(b) Fractional change of SNR ρ with starting frequency 15 Hz and 30 Hz as a function of chirp mass



(c) Fractional change of SNR ρ with starting frequency 15 Hz and 20 Hz as a function of chirp mass

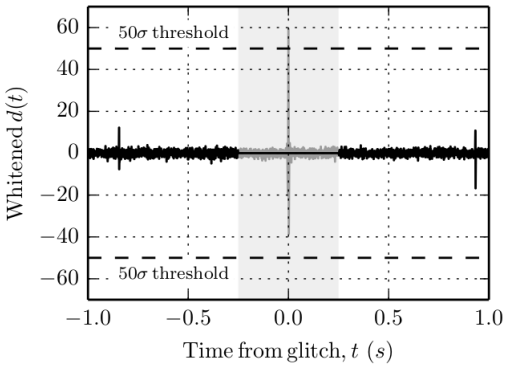
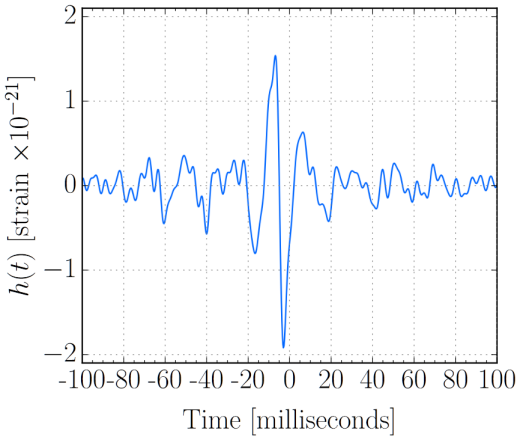
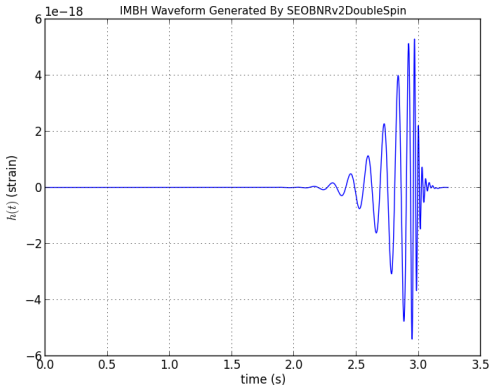


FIG. 17: When the whitened $h(t)$ exceeds a pre-defined threshold (50σ in this figure), the data centered at the peak with a window of 0.25 s on each side (shaded area) is regarded as a glitch and is replaced by zeros. (Figure taken from [6])



(a) An example of (blip) glitches.



(b) An example of GW signals from IMBH binary system.

this solution is not ideal because `gstlal` would have a higher tolerance of glitchy signals in lower mass region and hence allowing more noise in the search and causing more false alarm triggers.

Here we propose using a linear scale of $h(t)$ gate threshold based on chirp mass instead of using a single value for all masses. This simple solution addresses the issue mentioned above and now `gstlal` has different $h(t)$ gate threshold values in different mass regions. Figure 18 shows a linear scale for $h(t)$ gate threshold. User will specify the threshold value for template with lowest chirp mass and also that for template with the highest chirp mass. The program then computes an appropriate threshold value according to this linear scale and the highest chirp mass in a sub-bank. Of course, more tests can be performed to see how we should scale the threshold value for different chirp masses to optimize sensitivity but a linear scale which is proposed here is better than using a single value for the whole region.

This idea was implemented in the program `gstlal_inspiral` and `gstlal_inspiral_pipe` where user can either specify a single value of $h(t)$ gate threshold value through the option `--ht-gate-threshold` or specify a linear scale of threshold value through the option `--ht-gate-threshold-linear` with the format `mchirp_low:threshold_low-mchirp_high:threshold_high`.

Template bank	Uberbank
Starting frequency	30 Hz
Search Duration	3 days
Threshold value (default gating)	50
Linear scale (linear gating)	$\mathcal{M}_{\text{chirp, low}} = 0.8M_\odot$ $\text{Threshold}_{\text{low}} = 12.0$ $\mathcal{M}_{\text{chirp, high}} = 45.0M_\odot$ $\text{Threshold}_{\text{high}} = 100.0$

TABLE IV: The specifications of the two `gstlal` runs to test the new linear gating scheme.

B. Results of linear $h(t)$ gating

Two `gstlal` off-line runs were performed, one with default gating scheme and without vetoes in order to simulate the situation during on-line searches; one with our new linear gating scheme and without vetoes. The specifications of the two runs are tabulated in Table IV.

Figure 19 shows two χ^2 vs. ρ in H1 diagrams for two `gstlal` runs. The diagram on the left is the default gating run whereas that on the right is the linear gating run. We can see that the crosses corresponding to noise [14] with high values of χ^2 (enclosed inside the blue box) were removed and all injections recovered in default gating run were also recovered in linear gating run.

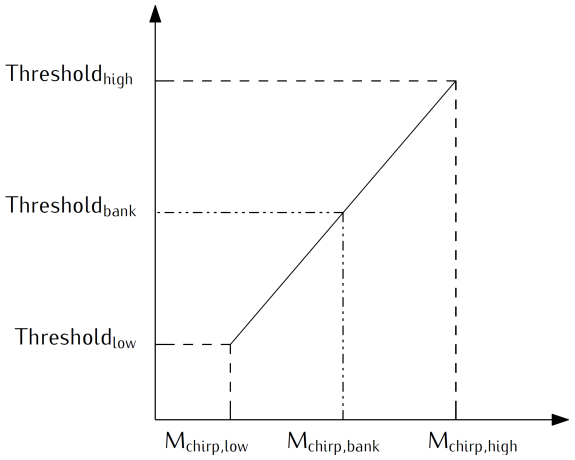


FIG. 18: A linear scale for $h(t)$ gate threshold.

C. Conclusions and Future Work for Linear $h(t)$ Gating

We can conclude from those plots (Figure 19) that the linear gating scheme has managed to reduce noise and clean the data in absence of vetoes and also recover the same set of injections as compared with the default gating scheme, meaning that it is at least doing as good as the default gating scheme.

It should be noted that we are not claiming the linear scale we used in the test is optimal, meaning that we can explore different linear scales to get better results for the linear $h(t)$ gating scheme. It is also worthwhile to run even more tests with a longer search duration so as to be exposed to more noise; or with more software injections and a combined (uberbank + IMBH) bank to see the performance of linear $h(t)$ gating scheme for binary systems with different masses. The modification on `gstlal` for allowing this linear $h(t)$ gating scheme has been pushed.

V. BANK SPLITTING

A. Motivation of Bank Splitting

Before performing singular value decomposition to the template bank, it is split into many similar sub-banks in order to optimize the compression effect and performance of singular value decomposition, which is computationally expensive. Also, the way we put templates together in a sub-bank is crucial because this will also affect the sensitivity of the search as background estimation is done per sub-bank. By default, the template bank is first binned by the effective spin parameter χ. Within each bin, the templates are then sorted by the chirp mass $\mathcal{M}$ in ascending order and N_T of the templates are grouped to form a split bank that should contain similar templates, as shown in Figure 15. Finally, Split banks are sorted in ascending order by the chirp mass of the first template and some number of split banks are put together, forming a sub-bank. We proposal several more schemes to split the template bank, which will be discussed in the coming subsections.

B. Proposal I: Sorting by Chirp Time τ

This proposal of bank splitter sorts templates by chirp time τ in descending order. As the chirp time of a template decreases with the chirp mass of the binary system as we previously discussed, it is thought to be a good proxy for similar templates, so basically we are trying to group templates that are "similar" together. Figure 20 shows how this bank splitter would split a template bank into many split-banks and hence sub-banks. Intuitively, the compression should be better when we group similar templates together, therefore it has the potential to outperform the default bank splitter.

In our implementation of `gstlal_bank_splitter`, the routine `lalsim_chirptime` from `SimIMRSEOBNRv2ChirpTimeSingleSpin` in `lalsimulation` library is used to compute the chirp time of a template where the routine uses thin-spline interpolation for 'higher mass' binary systems when the product of lower cut-off frequency and total mass in solar mass $f_{\rm low} \times m_{\rm total} \geq 120$ or 3.5-PN expansion to estimate the chirp time τ for 'lower mass' binary systems when the condition is not satisfied.

C. Proposal II: Sub-banks Sorted By Chirp Time of the First Template

This proposal of bank splitter is basically the same as the default bank splitter except that the output sub-banks are sorted by the chirp time of their first template in descending order.

D. Results: Performance of Different Bank Splitters

We can measure the compression efficiency and performance of the bank splitters in terms of latency or throughput. The following benchmarks are used to measure the latency:

1. using a program called `gstlal_inspiral_flopulator` that estimates the floating-point operations per seconds (FLOPS) needed to perform singular value decomposition to the sub-bank. Smaller the FLOPS needed, higher the compression efficiency;

2. monitoring the time required to complete `gstlal_inspiral` jobs, where matched filtering is actually done and they are the most

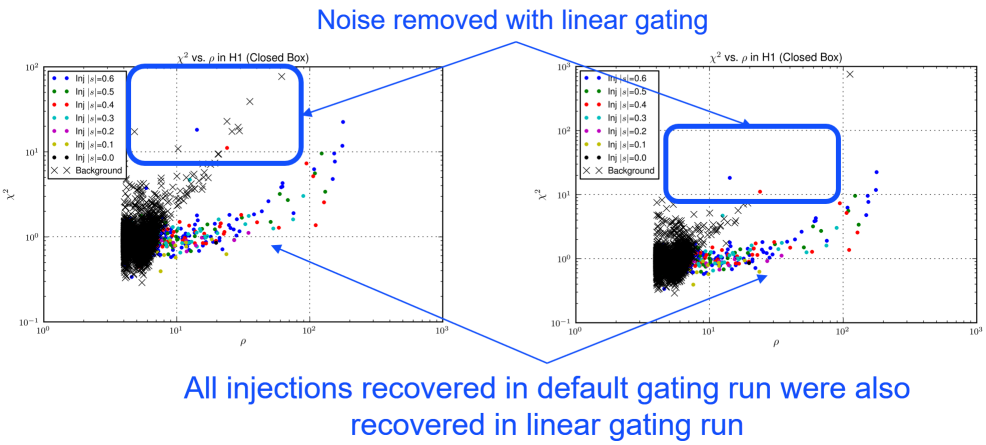


FIG. 19: Comparison of χ^2 vs. ρ in H1 for two runs (Left: Default gating; Right: Linear gating). Colored dots on the diagrams represent injections with different spins and crosses represent background.

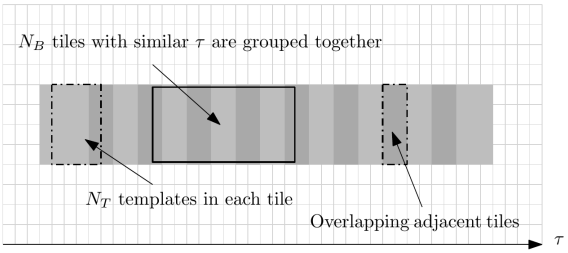


FIG. 20: Sorting by chirp time bank splitter

expensive jobs in `gstlal`. Shorter the time required, higher the performance.

Figure 21a, 21b and 21c show the 2D histograms of number of floating-point operations per second needed to perform singular value decomposition to sub-banks split by default bank splitter, proposal I bank splitter and proposal II bank splitter respectively. We can see that the default bank splitter and proposal II bank splitter perform better than proposal I bank splitter because the total MFLOPS required by default bank splitter and proposal II bank splitter are smaller than that of proposal I bank splitter by 37.9%. Also, more sub-banks split by proposal I bank splitter require more operations sampling at the frequency of 2048 Hz.

Figure 22a, 22b and 22c show the histograms of time required to complete `gstlal_inspiral` jobs on template bank split by default, proposal I and proposal II bank splitter respectively. We can see that the number of `gstlal_inspiral` jobs that take a long time ($>$ 600 minutes) for template bank split by proposal I bank splitter is smaller than that split by the default and proposal II bank splitter. Also, the maximum time needed for a `gstlal_inspiral` job to complete is smaller for template bank split by proposal I bank splitter (around 600 minutes)than the default and proposal II bank splitter (around 700 minutes). The reason we think why the chirp time bank splitter did not outperform the default bank splitter is that the chirp time function we were using might not be accurate enough.

Figure 23a, 23b and 23c show the plots of range of `gstlal` search for different ranges of chirp mass as a function of combined false-alarm rate (FAR) with template bank split by default ,proposal I and proposal II bank splitter respectively. This is one of the measures of the sensitivity of a search, longer the range, more sensitive of the search. Although the estimated range of different ranges of chirp masses with template bank split by various bank splitter differ from each other, statistically speaking they are all the same within the error.

E. Conclusion and Future Work for Bank Splitting

We can conclude from the plots that the default, proposal I and proposal II bank splitters have the same sensitivity (within the error). No conclusive statement can be drawn for the bank splitters regarding their performance because although proposal I takes a lot more floating-point operations per second than the other two bank splitters, the number of filtering (i.e. `gstlal_inspiral`) jobs greater than 600 minutes as well as the the maximum time required by a filtering job is also lower.

Another observation regarding proposal I bank splitter is that more sub-banks require more operations sampling at the frequency of 2048 Hz. This inspires us to design another bank splitter taking advantage of this: grouping templates that have similar chirp times as calculated with different starting frequencies. Hopefully this scheme of bank splitting will lower the number of sub-banks that require a lot of operations at high sampling rate and hence we can change the distribution of time of `gstlal_inspiral` jobs with more sub-banks taking less time to finish and by arranging the jobs appropriately to reduce the time to finish the whole filtering process. This

idea has been implemented and tests will be performed on this bank splitter to see its performance in terms of the sensitivity of the search as well as the latency of the filtering jobs.

VI. ACKNOWLEDGEMENT

The author would like to express his sincere gratitude to his mentors Surabhi Sachdev, Kent Blackburn, Alan Weinstein in LIGO laboratory at California Institute of Technology and Tjonnie Li at The Chinese University of Hong Kong. The author is also grateful to LIGO Laboratory, LIGO SURF program and Caltech SFP office for providing resources and organizing the SURF program; as well as the computational resources provided by the Leonard E Parker Center for Gravitation, Cosmology and Astrophysics at University of Wisconsin-Milwaukee.

This summer research would not be possible without National Science Foundation (NSF), Charles K. Kao Research Exchange Scholarship, CN YANG Research Exchange Scholarship for the financial support. The author would also like to thank department of physics and Professor MC Chu at CUHK for organizing the CUHK SURE program and the support to the author.

[1] 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. (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. **116**, 061102 (2016).

[2] B. S. Sathyaprakash and B. F. Schutz, Living Reviews in Relativity **12** (2009), 10.12942/lrr-2009-2, arXiv:0903.0338.

[3] LIGO Science Collobration, "Introduction to LIGO and Gravitational Waves: Continuous Gravitational Waves," `http://www.ligo.org/science/GW-Continuous.php` (), accessed: 14 May 2016.

[4] LIGO Science Collobration, "Introduction to LIGO and Gravitational Waves: Stochastic Gravitational Waves," `http://www.ligo.org/science/GW-Stochastic.php` (), accessed: 14 May 2016.

[5] LIGO Science Collobration, "Introduction to LIGO and Gravitational Waves: Burst Gravitational Waves," `http://www.ligo.org/science/GW-Burst.php` (), accessed: 14 May 2016.

[6] C. Messick, K. Blackburn, P. Brady, P. Brockill, K. Cannon, R. Cariou, S. Caudill, S. J. Chamberlin, J. D. E. Creighton, R. Everett, C. Hanna, D. Keppel, R. N. Lang, T. G. F. Li, D. Meacher, A. Nielsen, C. Pankow, S. Privitera, H. Qi, S. Sachdev, L. Sadeghian, L. Singer, E. G. Thomas, L. Wade, M. Wade, A. Weinstein, and K. Wiesner, ArXiv e-prints (2016), arXiv:1604.04324 [astro-ph.IM].

[7] 1 dimension for each component mass (hence 1 + 1 since there are two bodies) and 3 dimensions for each component spin (hence 3 + 3 since there are 2 bodies).

[8] 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. (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. D **93**, 122003 (2016).

[9] 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., ArXiv e-prints (2016), arXiv:1606.04856 [gr-qc].

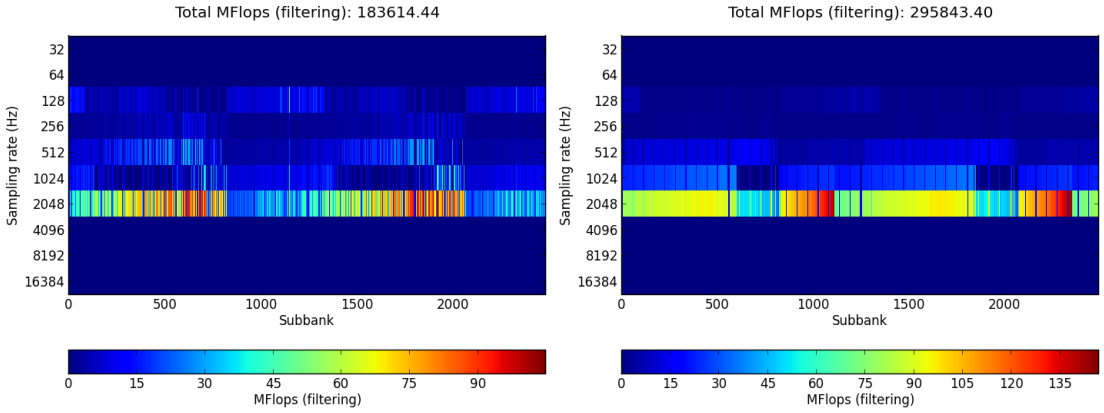
[10] The LIGO Scientific Collaboration, the Virgo Collaboration, B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, and et al. (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. **116**, 241103 (2016).

[11] K. Cannon, A. Chapman, C. Hanna, D. Keppel, A. C. Searle, and A. J. Weinstein, Phys. Rev. D **82**, 044025 (2010).

[12] This SNR is different from the usual one in a sense that the usual SNR compares the signals with a set of templates (which may not be identical) whereas this SNR compares the signals with themselves (which are injections we put to the pipeline).

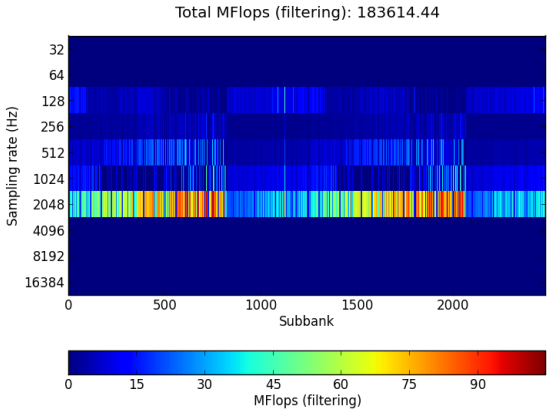
[13] After whitening, the data will have a variance of 1, therefore the threshold 50σ is converted to 50 here.

[14] The background is estimated by the coincident triggers in time-slid H1 and L1 data, which are not going to be real gravitational wave signals.



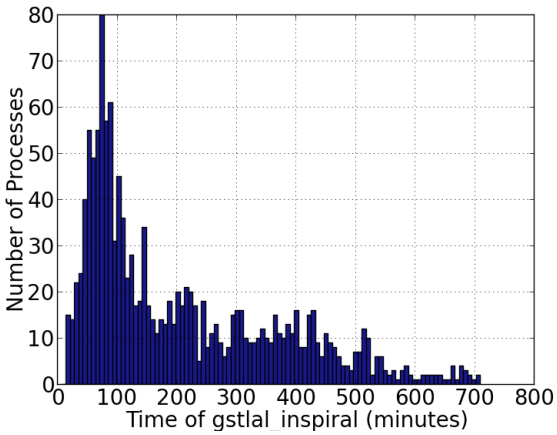
(a) Default bank splitter

(b) Sorting by chirp time bank splitter

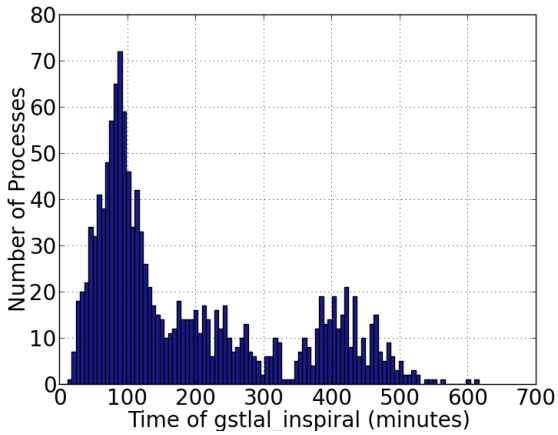


(c) Sub-banks sorted by chirp time of the first template

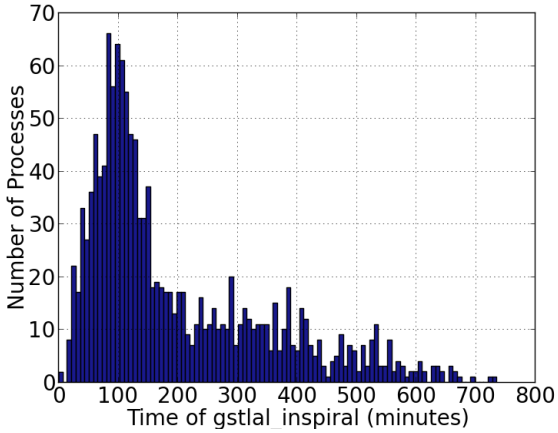
FIG. 21: 2D histograms of MFLOPS needed to perform singular value decomposition to sub-banks split by different bank splitters as estimated by `gstlal_inspiral_flopulator`.



(a) Default bank splitter

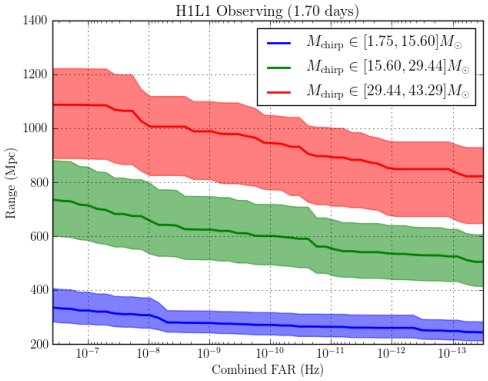


(b) Sorting by chirp time bank splitter

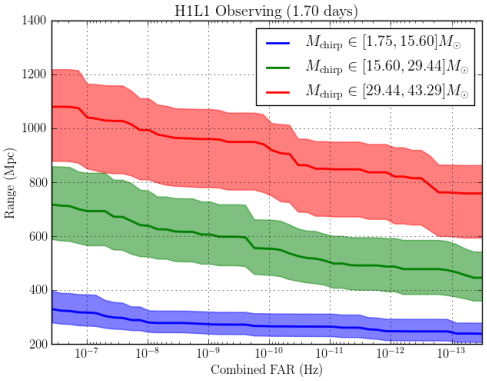


(c) Sub-banks sorted by chirp time of the first template bank splitter

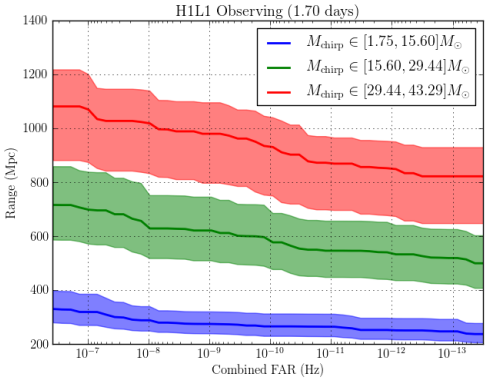
FIG. 22: Histograms of time required to complete `gstlal_inspiral` jobs on template bank split by different bank splitters.



(a) Default bank splitter



(b) Sorting by chirp time bank splitter



(c) Sub-banks sorted by chirp time of the first template bank splitter

FIG. 23: Range of `gstlal` search for different ranges of chirp mass as a function of combined FAR with template bank split by different bank splitters. The solid lines represent the range of a search for a particular range of chirp masses whereas the colored patches represent the errors of the estimation of range.