Determining the Feasibility of Matched Filter Searches for Core-Collapse Supernovae
LIGO SURF 2022 Interim Report #2

Pavani Jairam
Duke University, Durham, NC 27708, USA

Mentor: Ryan Magee
LIGO Laboratory, California Institute of Technology, Pasadena, CA 91125, USA
(Dated: July 2022)

With the efforts of the Laser Interferometer Gravitational-Wave Observatory (LIGO) collaboration, gravitational waves (GWs) have been successfully detected from black hole mergers, neutron stars, and neutron star-black hole binaries. However, there are other violent phenomena, such as core-collapse supernovae (CCSNe), that are potential candidates for gravitational wave studies. CCSNe are of particular interest because they emit other astrophysical messengers such as neutrinos and electromagnetic rays. I will study the feasibility of using matched filter searches for CCSNe with a phenomenological GW model that aims to be representative of CCSNe waveforms. I will examine the impact of stochasticity on the g-mode dominated emission, design a template bank of CCSNe gravitational waveforms, and compare a search benchmarked against numerical relativity simulations.

I. INTRODUCTION

Gravitational waves (GWs) are ripples in the spacetime that are caused by violent processes in the Universe. GWs have been predicted by Einstein in his general theory of relativity in 1915. In 2015, the first gravitational wave signal, GW150914, was detected from a black hole merger by the Laser Interferometer Gravitational-Wave Observatory (LIGO) confirming Einstein's theory of relativity and opening the door to studying other astrophysical phenomena with GWs [1, 2]. In particular, core-collapse supernovae (CCSNe) are promising candidates for GW models. CCSNe are a known multi-messenger astronomy candidate which can be studied using different sources such as GWs, electromagnetic rays, and neutrinos to aid in our understanding of the Universe [3]. Studying the neutrinos and GWs from CCSNe in the local universe and Milky Way provides insight on the mechanisms and processes behind the core collapse and shock wave of these violent explosions.

CCSNe occur specifically with high mass stars, which have a mass greater than approximately $8M_{\odot}$[4]. The death of a high mass star starts when hydrogen is exhausted and helium begins to burn causing heavier elements to be produced. Once iron is produced in the core, the core collapses increasing the temperature. The core becomes progressively more dense, which eventually causes neutrons to be formed from the electron capture reaction and neutrinos to be released. As the collapse of the core accelerates, there is a bounce when the nuclear forces become repulsive, creating a shock wave through the outer layers of the star. Finally, a shock from the neutrinos in the core blasts off the outer layers of the star which leads to accretion. The remnant of this explosion is either a black hole or neutron star. The accretion disks formed during this process can be gravitationally unstable due to the fallback from the collapsing star [5].

These violent and energetic supernovae are rare events that occur once or twice a century in the Milky Way [6] and can potentially be detected by GWs. With the upcoming fourth run, O4 [7], it is possible that a CCSN signal would be strong enough to be observed on Earth. Therefore, being able to analyze the GW signals from CCSN would allow further studies into the mechanisms behind them.

II. OBJECTIVES

The goal of this project is to study supernovae waveform models and matched filter pipelines to design an efficient matched filter search for CCSN and calibrate the distance of progenitor stars with existing simulations. More specifically, this project will explore the feasibility of matched filtered searches for supernovae. The matched filtered search method correlates a bank of template gravitational waveforms to the detected data to determine if a gravitational wave is present, which can be used to detect the gravitational waves of supernovae.

We focus on phenomenological GW emission dominated by g-mode oscillations because because even with random behavior, they are the dominant feature in numerical relativity simulations. According to the Astone et al. paper [8], the strain, or the amplitude, of the GW is similar to a damped harmonic oscillator as the g-mode emissions are characterized by the random bursts from the accretion disk inflows that add acceleration to the collapse. Thus, a random driving force is added to simulate the effects of these random inflows on the g-mode dominated emission the g-modes. For this project, I want to quantify the impact of the randomness on the GW strain feature.

I will solve the differential equation in Astone et al. [8] to calibrate the emission from the distance-dependence parameter to numerical relativity simulations. I will first generate supernova waveforms to see if they are capable of covering the search space and exhibit similar behavior to numerically generated waveforms. To do this, I will have to sample a discrete number of points to calculate the GW emission. This creates a template bank, which is needed for matched filtering based searches, which finds if there is a high similarity in the signal-to-noise ratio (SNR) in the detector data. I can then compare the SNR to see if there is a high similarity in the data, which may indicate the presence of a gravitational wave. In doing so, I hope to analyze the associated randomness of the g-mode and conclude if the matched filter search can handle the random elements of a supernova.

III. APPROACH

For this project I will be utilizing the methods of Astone et al. [8] to reproduce the results present in the paper, which uses a phenomenological template to capture the key features of CCSNe. The paper focuses on primarily the excitation of g-modes of progenitor neutron stars (PNS) in which its frequency starts at 100 Hz and increases as the mass of the PNS grows. The signal starts right after the bounce or up to a 200 ms delay and ends with an explosion or formation of a black hole. This emission is simulated as a damped harmonic oscillator with a random forcing as frequency varies over time. The following parameters of the supernova waveform is comprised of the post-bounce time and three frequency constants of the frequency evolution: f_0 which describes the starting frequency at the start of the signal, f_{1s} which is the frequency of the signal one second after the bounce, f_{driver} which is the driving frequency, Q which is the Q factor quantifying the ratio of energy stored to energy lost per cycle, t_2 which indicates the time that the frequency polynomial is maximum, t_{ini} which is the start of the signal relative to the bounce, and t_{end} which is the end of the signal relative to the bounce.

The paper models the inflows of the g-modes as accelerations, but does not explore the dependence of the amplitude. I plan on exploring the possible values of this amplitude parameter which can be calibrated to generate distance-dependent templates. To address this, I will first have to develop code to model the CCSN waveform, and once I can reproduce the paper's results, I will compare the waveform similarities by calculating the overlap, determine the impact of stochasticity of the g-mode emission, and look at the randomness of simulated supernova waveform. By comparing a search based on a template bank of CCSNe gravitational waveforms against numerical relativity simulations, I can then determine if matched filter searches are viable to predict the behavior of the g-mode.

IV. PROGRESS UPDATE

The first task of my project was to duplicate the results of the Astone et al. paper [8]. To generate the waveform similar to the Astone paper, I solved the following differential equation 1 cited in the paper using the symplectic Euler method.

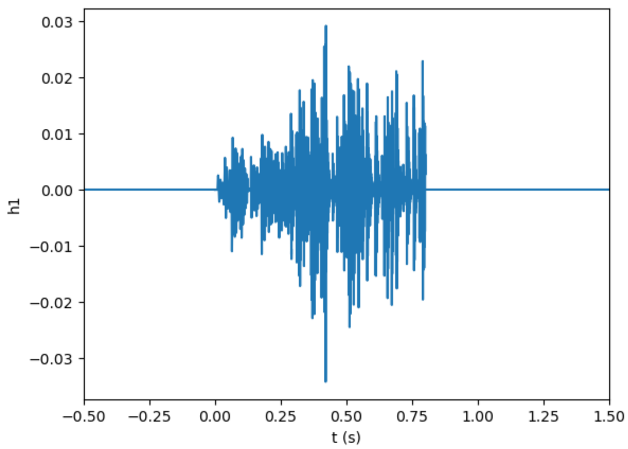
$$\partial_{tt}h + \frac{\omega(t)}{Q}\partial_t h + \omega(t)^2 h = a(t) \tag{1}$$

I utilized the same 7 parameters that describes the gravitational wave emission of a CCSN: f_0, f_{1s}, f_{driver}, Q, t_2, t_{ini}, and t_{end}. From this, I was able to find the strain in the time-domain of a randomized waveform. The waveform is randomized to mimic the stochasticity of accelerations of the CCSNe. When I first generated the waveform with my code, my waveform looked like a diagonal line, and not the expected oscillations. Comparing my results to my mentor's code, I ultimately figured out that the issue was that I was oversampling the times series. For instance, I did not include the 10 when I was calculating the strain using the symplectic Euler method and I was also missing a factor of 10 when generating random impulses. Once I fixed my code, I was then able to generate a plot similar to the Astone et al. paper as shown in Fig. 1a. Fig. 1b is another waveform created with this same code, and since my code randomizes the impulses of the supernova, this waveform has different amplitudes at different times.

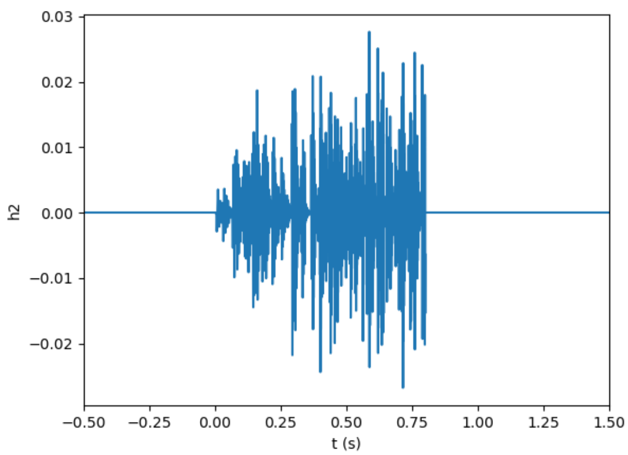
This randomly generated CCSN waveform has a duration of 1 second, which is the expected duration of this phenomenon, and demonstrates the behavior of a damped harmonic oscillator. I then wrote a Python script that would take in any values for the 7 parameters as the input and generate a waveform as the output. I also fixed the seed using Numpy's random.seed so that the perturbations of the waveform would not be randomized with each run and it would be easier to compare the behavior of the waveform with changes to the parameters. To get a sense of how the waveform changes with the f_0, f_{1s}, f_{driver}, and Q parameters, I used the CCSN waveform script to compare the minimum and maximum parameters listed in the Astone et al. paper [8] as shown in Fig. 2. From this, I noticed that the general shape of the waveform was retained with changes to the amplitude. Moving forward, I needed a quantitative metric to compare the waveforms rather than eyeballing.

Hence, I compared the figures by calculating the overlap of the two waveforms I generated in Fig. 3 [9, 10]. The overlap quantifies how similar two waveforms are from a range of 0 to 1, with 1 being the same waveform. The overlap is calculated using Equation 2 which consists of finding the inner product of the waveforms.

$$overlap = \frac{\langle h_1 | h_2 \rangle}{\sqrt{\langle h_1 | h_1 \rangle \langle h_2 | h_2 \rangle}} \tag{2}$$



(a) Waveform 1



(b) Waveform 2

FIG. 1: Supernova waveforms generated following the methods and the parameters $f_0 = 100\ Hz$, $f_{1s} = 700\ Hz$, $f_{driver} = 200\ Hz$, $Q = 10$, $t_2 = 1.25\ s$, $t_{ini} = 0\ s$, and $t_{end} = 0.8\ s$ used in Astone et al. [8]. The impulses of the supernovae were randomly generated using Numpy.

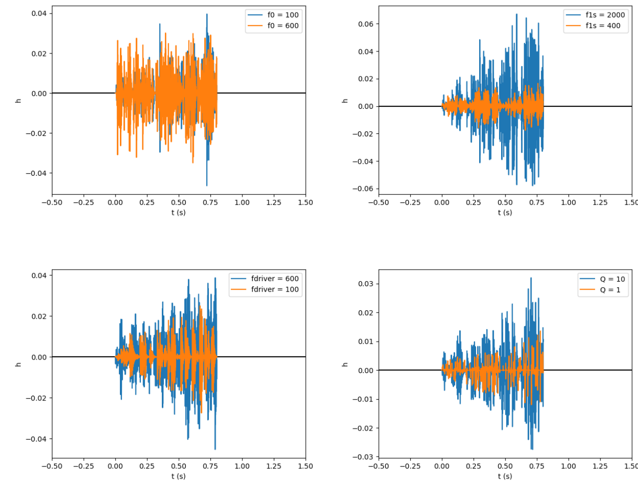


FIG. 2: The comparison between the minimum and maximum values of the f_0, f_{1s}, f_{driver}, and Q parameters of CCSNe as listed in Astone et al. [8]. The seed was fixed when generating these plots. In general, tweaking the parameters demonstrates that there are some similarities in the overall behavior of the waveform.

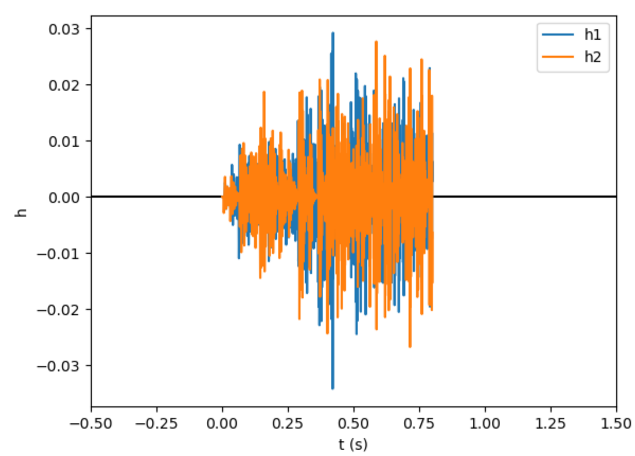
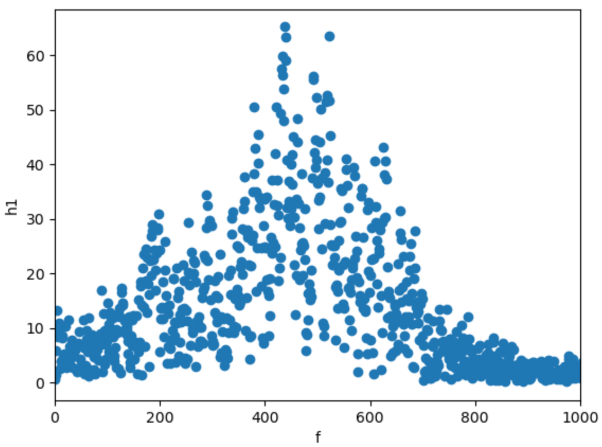


FIG. 3: The waveforms generated in Fig. 1 overlapped. h_1 is the waveform from Fig. 1a and h_2 is the waveform from Fig. 1b. The quantitative overlap to determine how similar these two waveforms are is calculated using Equation 2.

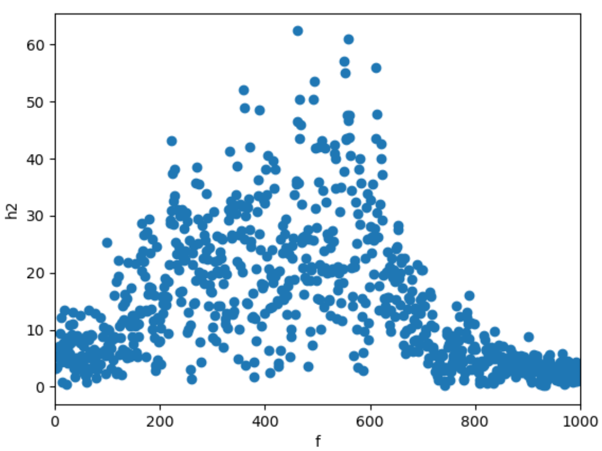
The inner product of the two waveforms is calculated using Equation 3 which integrates the waveforms in the frequency-domain.

$$\langle h_1|h_2\rangle = 4\int_0^\infty \frac{\tilde{h}_1^*(f)\tilde{h}_2(f)}{S_n(f)}df \qquad (3)$$

The strain calculated using the Astone et al. method was in the time-domain. Therefore, I Fourier transformed the waveforms in Fig. 1 from the time-domain to the frequency-domain as seen in Fig. 4. I also used the power spectral density (PSD) for O4 by using GSTLAL's Python script that generates the PSD given the frequency and amplitude spectral density (ASD) [11]. For this part of my project, I faced some challenges getting the Fourier transform to be representative of the waveforms I generated. I used the wrong Numpy function that calculates the Fourier transformation. In the end, I was able to correctly Fourier transform the strain by using Numpy's fast Fourier transform function. After I implemented this code, I plotted the Fourier transform but it was of no use because the frequency bounds on the plot were in the tens of thousands–much larger than what LIGO uses. Thus, I bounded the figures from 0 to 1000 Hz because that range is the frequency observed by LIGO. Both plots peak between 400 to 600 Hz and might be related to the accelerations and the driving frequency which is indicative of the stochastic nature of CCSNe. Then, using Equation 2, the calculated overlap of the waveforms was only 10%.



(a) Fourier Transformation of Waveform 1



(b) Fourier Transformation of Waveform 2

FIG. 4: Fourier transforms of the waveforms generated in 1. The absolute value of the amplitude was plotted. The peaks are between 400 to 600 Hz and highlights the stochasticity of CCSNe.

Moving forward, I also compared my code to the CCSN code from Astone et al. to see if there were possible discrepancies in modeling the waveform. I found that the paper accounted for an extra factor of 10 after the end of the relative signal bounce time to mimic the extra damping of the supernova explosion or black hole formation. I verified that the other calculations I had done to generate the waveform were the same.

Once I was sure that my waveform was representative of the paper's results, I wanted to accurately measure the overlap by also accounting for all the possible times of the waveform. The overlap calculated in Eqn. 2 only accounts for the amplitude. Thus, we want to find the time at which the waveforms are maximally overlapped using Eqn. 4 [12].

$$\langle h_1|h_2\rangle = 4\int_0^\infty \frac{\tilde{h}_1^*(f)\tilde{h}_2(f)}{S_n(f)} e^{2\pi i f t} df \tag{4}$$

This version of the overlap finds the maximum overlap by iterating over the frequency steps of the strain's frequency domain. In Eqn. 5, the amplitude of the time domain encodes $h_+$and $h_\times$.

$$h(t) = h_+(t) + i h_\times(t) \tag{5}$$

The match calculation shown in Eqn. 6 searches the waveform for the peak signal and accounts for the shifts of the waveform for the maximum time overlap. The highest possible value would be 1. Using this calculation, h is now based on Eqn. 5, and not just the amplitude. The result of this can be Fourier transformed to find the maximum time overlap in the frequency domain which results in Eqn. 4.

$$match = [\langle h_1|h_2\rangle]_{t_0,\phi_0=0} \tag{6}$$

Because I want to compare multiple waveforms to find the maximum overlap, I wrote a script that generalizes the code for calculating the overlap between two waveforms so that I can compare waveforms of varying parameters or changing the random seed. To presample the parameter space, I generated multiple waveforms by fixing the supernova parameters and randomizing the seed. I did this by writing a script in which one waveform can be compared to a set number of realizations. First, I generated one waveform and then in a for loop, I would compare a randomized waveform with the same parameters. As the script loops, if the overlap is higher than other values in the dictionary, the waveform gets added. After the script runs, the final entry in the dictionary stored contains the maximum overlap between the first waveform and another realization that was generated.

I found that running many waveforms is time consuming, so we optimized the code that calculated the generated the waveform and calculated the overlap. We downsampled the sampling rate that I had used previously of 16384 Hz to 2048 Hz. I also divided out the factor of 10 for the time step scaling to reduce the time to calculate a waveform. We also restructured the code that avoided unnecessary duplications of calculations such as the overlap of the first wave. These changes reduce the original runtime of around 4 seconds to around 0.5 seconds. For this version of the script, the highest overlap was 40.3%. From this still low overlap, we determined that the stochasticity of the supernova waveforms was likely too high to be able to use matched filtering.

Another possible method for presampling the parameter space that I experimented with was fixing the random seed and changing one parameter to see how the overlap between the waveforms changes. Essentially, if the parameter changes by 1, I want to quantify how different the supernova waveforms get from one another. Out of the 7 parameters, I first changed the driving frequency parameter, f_{driver}, fixed the other parameters, and fixed the seed to generate the parameter space shown in Fig. 5. I did this by creating a waveform at each f_{driver} value set within a range based on the minimum and maximum parameters. I chose 100 to 130 Hz, which took 118 seconds to run, to avoid a long runtime. Fig. 5 has the maximum overlaps along the diagonal of the plot because the waveforms are the same and should have an overlap of 100%. Looking at waveforms that are different, we found that the overlap can be close to 100% when the driving frequencies are 1 Hz apart (i.e. 100 Hz and 101 Hz). This can be seen in Fig. 5 where there are high overlaps that deviate one point away from the diagonal.

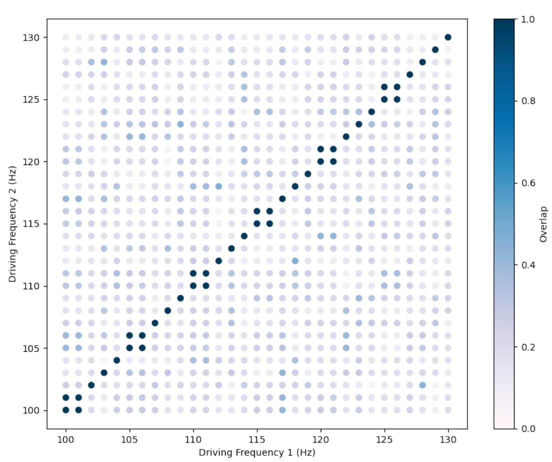


FIG. 5: The parameter space of driving frequencies of 100 to 130 Hz. The dark blue data points indicate that there is a high overlap close to the diagonal, which marks the same waveform.

In Fig. 7, I generated a parameter space for the Q factor value of 1 to 10 which follows the minimum and maximum parameters of Astone et al. For this plot, I compared the varying Q factor values and once again fixed the random seed and the other parameters. Once again, the diagonal indicates the overlap of the same waveform which means that the highest overlap is along the diagonal. All of the overlaps that are not along the diagonal in the Q factor parameter space are above 65%, with several data points close to the diagonal above 90% in Fig. 7.

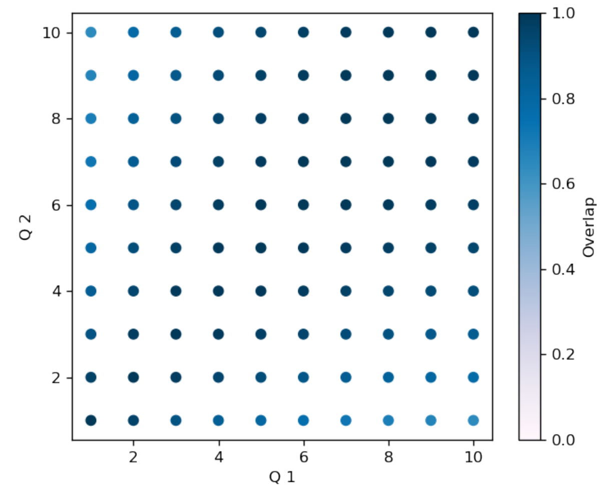


FIG. 6: The parameter space of the Q factors ranging from 1 to 10. The highest overlaps are close to the diagonal line.

I then generated a parameter space for f_{1s}. Much like the f_{driver} parameter, the waveforms that deviated with an increment of 1 from the diagonal had the highest overlap. These results are promising like the f_{driver} because it indicates that it may be possible to construct a template bank using this slight deviation between parameters.

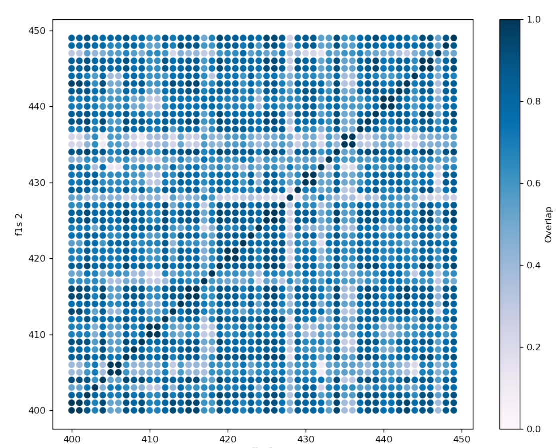


FIG. 7: The parameter space of f_{1s} ranging from 400 to 449 Hz. The diagonal line highlights the overlap of the same waveforms.

V. FUTURE WORK

After finding that the f_{driver}, Q, and f_{1s} factor parameter spaces have promising results for generating a high overlap value, the next step is to examine the overlaps of the f_0 parameter. I can then determine if a search can be constructed on the simulated supernova parameters with a fixed randomization seed. Once I get a sense of the overlap values of changing individual parameters, I want to vary the f_0, f_{1s}, f_{driver}, and Q parameters to qualitatively deduce if those differences between the supernova waveforms affect the overlap. So far, I have been utilizing multiple scripts to test the overlap of different parameters and randomizing the seed which is inefficient. Therefore, moving forward, I will need to clean up my scripts, upload it to a Git repository, and package my code to streamline my overlap calculations.

VI. ACKNOWLEDGEMENTS

This work was supported in part by the Research Experiences for Undergraduates (REU) Program of the National Science Foundation. I would also acknowledge the California Institute of Technology and the LIGO Summer Undergraduate Research Program. I would like to extend my thanks to Alan Weinstein for running this summer program and his continuous support. Finally, I would also like to thank my mentor Ryan Magee for his guidance and mentorship this summer.

VII. REFERENCES

[1] T. L. S. Collaboration and the Virgo Collaboration, (2016), 10.1103/PhysRevLett.116.061102.

[2] C. Messick, K. Blackburn, P. Brady, P. Brockill, K. Cannon, R. Cariou, S. Caudill, S. J. Chamberlin, J. D. Creighton, R. Everett, C. Hanna, D. Keppel, R. N. Lang, T. G. 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, Physical Review D **95** (2017), 10.1103/physrevd.95.042001.

[3] K. Nakamura, S. Horiuchi, M. Tanaka, K. Hayama, T. Takiwaki, and K. Kotake, Monthly Notices of the Royal Astronomical Society **461**, 3296 (2016).

[4] A. Burrows and D. Vartanyan, Nature **589**, 29 (2021).

[5] D. M. Siegel, A. Agarwal, J. Barnes, B. D. Metzger, M. Renzo, and V. A. Villar, “”super-kilonovae” from massive collapsars as signatures of black-hole birth in the pair-instability mass gap,” (2021), arXiv:2111.03094 [astro-ph.HE].

[6] K. Rozwadowska, F. Vissani, and E. Cappellaro, New Astronomy **83**, 101498 (2021).

[7] “LIGO, Virgo and KAGRA Observing Run Plans,” `https://observing.docs.ligo.org/plan` (2022), [Online; accessed 15-May-2022].

[8] P. Astone, P. Cerdá -Durán, I. D. Palma, M. Drago, F. Muciaccia, C. Palomba, and F. Ricci, Physical Review D **98** (2018), 10.1103/physrevd.98.122002.

[9] S. K. Sahay, “Studies in gravitational wave data analysis,” (2002).

[10] T. A. Apostolatos, Phys. Rev. D **52**, 605 (1995).

[11] “GSTLAL psd.py,” `https://git.ligo.org/lscsoft/gstlal/-/blob/master/gstlal/python/psd.py`.

[12] B. Allen, W. G. Anderson, P. R. Brady, D. A. Brown, and J. D. E. Creighton, Phys. Rev. D **85**, 122006 (2012).