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

Pavani Jairam
Duke University, Durham, NC 27708, USA

Mentor: Ryan Magee
LIGO Laboratory, California Institute of Technology, Pasadena, CA 91125, USA
(Dated: September 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 studied the feasibility of using matched filter searches for CCSNe with a phenomenological GW model that aims to be representative of CCSNe waveforms. I examined the impact of stochasticity on the g-mode dominated emission of CCSNe, investigated if the randomness of waveforms is manageable for generating a parameter space, and designed a template bank of CCSNe gravitational waveforms. I successfully generated a template bank and determined that matched filtering is feasible for CCSN waveforms.

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 underlying processes of 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]. During the lifetime of a high mass star, the gravitational pressure and degeneracy pressure are balanced. 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 degeneracy pressure limit is met and the core collapses increasing the temperature. The core becomes progressively more dense, which 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 shockwave created 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.

Notably, the accretion disks formed during this process can be gravitationally unstable due to the fallback from the collapsing star [5], which could potentially emit GWs. Within the Milky way, these violent and energetic supernovae are rare events that occur once or twice a century [6]. 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 CCSNe would allow further studies into the mechanisms behind them.

One major difficulty with supernova searches is modeling the highly stochastic nature of CCSN from the accretion disk inflows. Current supernova searches, including Principal Component Analysis (PCA), machine learning, and coherent WaveBurst are weakly modeled. Thus, for the approach of this project, we explore the feasibility of matched filtered searches for supernovae. The matched filtered search method correlates a template bank of gravitational waveforms to the detected data to determine if a gravitational wave is present. We want to test how well this method is able to cover parameter space of supernovae to conclude if matched filtering is a plausible method in comparison to a generic search.

The supernova model we base the template bank on follows the phenomenological Astone et al. paper [8] model. This model demonstrates the CCSN GW emission dominated by g-mode oscillations because even with random behavior, they are the dominant feature in numerical relativity simulations. In the methods section, I will describe the approaches I took to test the feasibility of matched filter searches for CCSN. In the results and discussion section, I outline the trends and outcome of my tested methods and highlight the key features of the template banks I have created. Finally, the conclusion wraps up my findings which determine that matched filter searches are in fact feasible and even follow trends in other CCSN GW detection methods.

II. METHODS

For this project, the supernova waveform followed the phenomenological model from Astone et. al [8] which captures the key features of CCSNe waveforms. More specifically, this model simulates the g-mode emission with a damped harmonic oscillator with a random forcing to mimic the random inflows of a accretion disk as shown in Equation 1. I solved this differential equation following the sympletic-Euler method as mentioned in Astone et. al [8].

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

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.

These 7 parameters restrict the behavior of the angular frequency, $\omega(t)$, and are set by the simulation of the waveform. This means that the parameter space of the supernova is 7-dimensional, which is more than the typical gravitational wave search of 4 parameters for phenomena such as black holes and neutron stars. Another important feature of Eq. 1 is the acceleration, $a(t)$, that captures the stochasticity of the inflows of the supernova. This acceleration acts as a pseudo eighth parameter, and is randomized during the simulation.

In order to test the feasibility of a template bank, I needed to be able to generate supernova waveform templates. To generate these templates, in the parameter space, the set of points with high similarity need to be found. This similarity value can be calculated by finding the match. The match quantifies how similar two waveforms are from a range of 0 to 1, with 1 being the same waveform. The match is calculated using Equation **??** which consists of finding the inner product of the waveforms.

The match is used to account for the shifts in time of the waveform. Thus, we want to find the time at which the waveforms are maximally overlapped using Eqn. 2 [9].

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

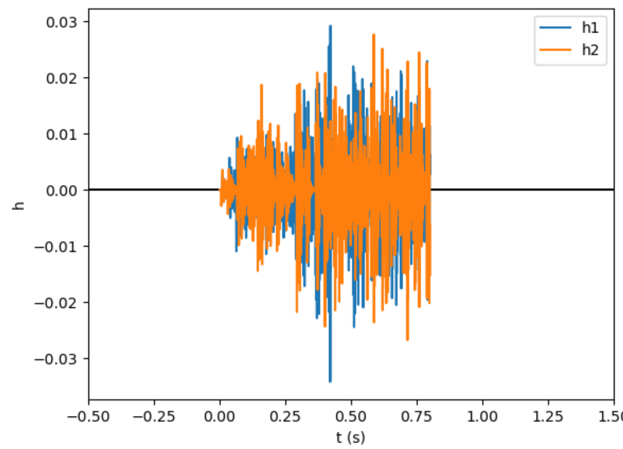


FIG. 1: The waveforms generated in Fig. 2 overlapped. h_1 is the waveform from Fig. 2a and h_2 is the waveform from Fig. 2b. The quantitative overlap to determine how similar these two waveforms are is calculated using Equation 2.

This overlap of the waveforms finds the maximum overlap by iterating over the frequency steps of the strain's frequency domain. The match calculation shown in Eqn. 3 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. The result of this can be Fourier transformed to find the maximum time overlap in the frequency domain which results in Eqn. 2. Because of this, the time and the phase of the waveform is taken into account to quantify the similarity between two waveforms.

$$match = [\langle h_1|h_2\rangle]_{t_0,\phi_0=0} \tag{3}$$

The inner product of the two waveforms is calculated using Equation 4 which integrates the waveforms in the frequency-domain.

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

To determine the feasibility of using matched filtering for supernova waveforms, I took two approaches: change one parameter at a time, and compare waveforms with smaller time windows. For changing one parameter at a time, I compared a range of values of one parameter with itself while fixing the randomization of the inflows and time windows of the waveform. Thus, 5 parameter spaces were be examined: f_0, f_{1s}, f_{driver}, Q, and t_2. The reasoning behind changing one parameter at a time is to analyze the waveform match if there are small differences between the waveforms being compared. If these small steps indicate a high performing match between the supernova waveforms, then we could scale up to a higher dimensional space which would include changing all 7 parameters at once and include the supernova randomization. From these tests, we would be able to determine if matched filtering is feasible, which means that creating a template bank is feasible.

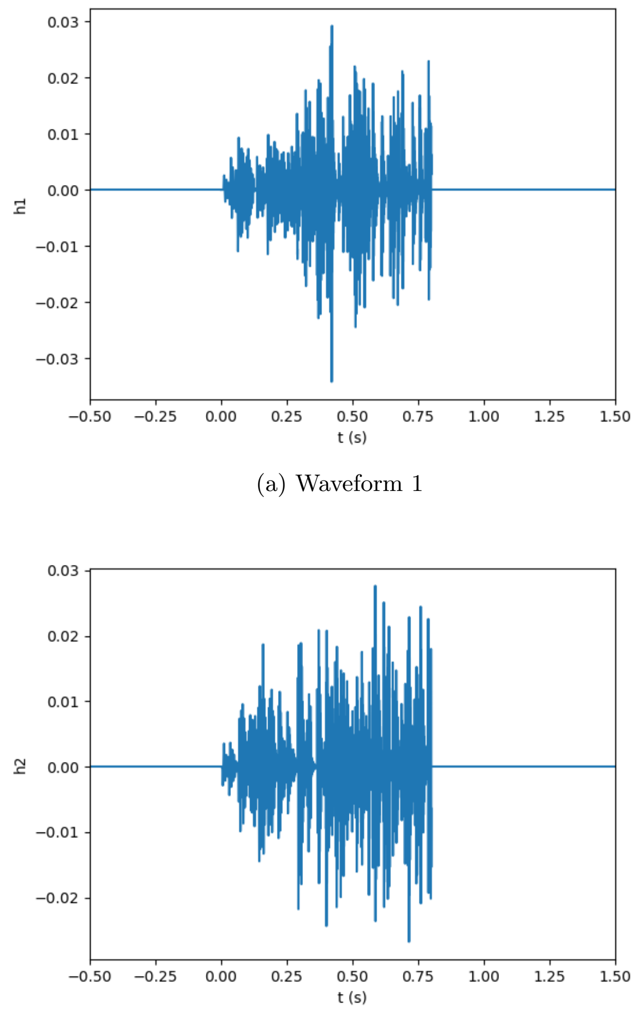
III. RESULTS AND DISCUSSION

The first task was to duplicate the results of the Astone et al. paper [8] to generate a supernova waveform. I utilized the same simulated 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}. The sampling rate I utilized was 16384 Hz. 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.

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. This tapers the end of the waveform at the end so that it matches the waveform of the Astone et. al paper. I verified that the other calculations I had done to generate the waveform were the same. I generated a waveform plot similar to the Astone et al. paper as shown in Fig. 2a. Fig. 2b 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 0.8 second, which is the expected duration of this phenomenon based on the parameters set by the simulation, 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. The quantitative metric I used to compare the waveforms was the match as mentioned previously, so when two waveforms are overlapped as shown in Fig. 1, the match would determine how similar these two waveforms are [10, 11].

The strain calculated using the Astone et al. method was in the time-domain. Therefore, I Fourier transformed the waveforms in Fig. 2 from the time-domain to the frequency-domain as seen in Fig. 3. I also used the power spectral density (PSD) for the O4 run by using GSTLAL's Python script that generates the PSD given the frequency and amplitude spectral density (ASD) [12]. The range that is the frequency observed by LIGO is from 0 to 1000 Hz, thus I bounded the Fourier transformations in the frequency domain from 0 to 1000 Hz as well. Both



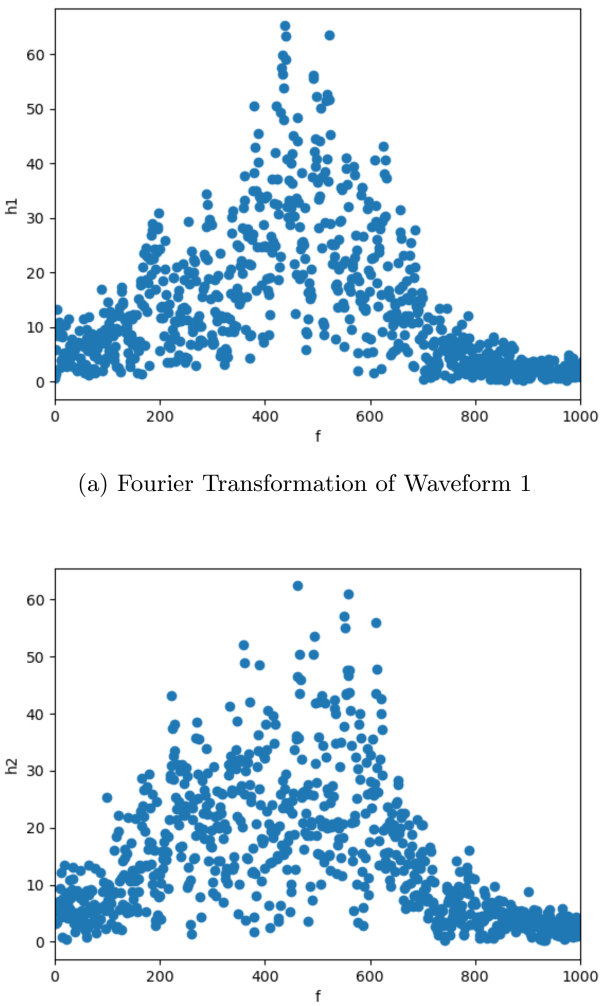
(a) Waveform 1

(b) Waveform 2

FIG. 2: 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.

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.

Because I wanted 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 param-



(a) Fourier Transformation of Waveform 1

(b) Fourier Transformation of Waveform 2

FIG. 3: Fourier transforms of the waveforms generated in 2. The absolute value of the amplitude was plotted. The peaks are between 400 to 600 Hz and highlights the stochasticity of CCSNe.

eters. 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 was 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. This change in the time step scaling decreased the number of time shifts. We also restructured the code that avoided unnecessary duplications of calculations. These changes reduce the original runtime of around 4 seconds to around 0.5 seconds. Initially, to test how waveform randomizations were affecting the match of the waveforms, I generated 10,000 waveform realizations following the 7 parameters set by the simulation like in the Astone et. al paper, and randomized the waveform. With this method having a time window of 0.8 seconds for the CCSN waveform, the highest overlap between these waveforms was 40.3%.

The other possible method for presampling the parameter space that I experimented with to examine how the match changes 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. 4. 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. 4 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. 4 where there are high overlaps that deviate one point away from the diagonal.

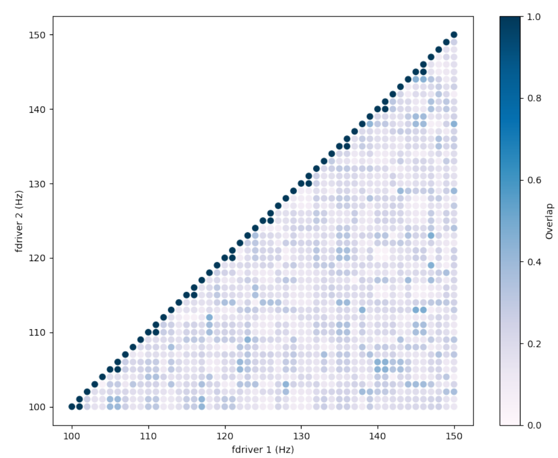


FIG. 4: The parameter space of driving frequencies of 100 to 130 Hz. The dark blue data points indicate that there is a high overlap close to the diagonal, which marks the same waveform.

In Fig. 5, 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. 5.

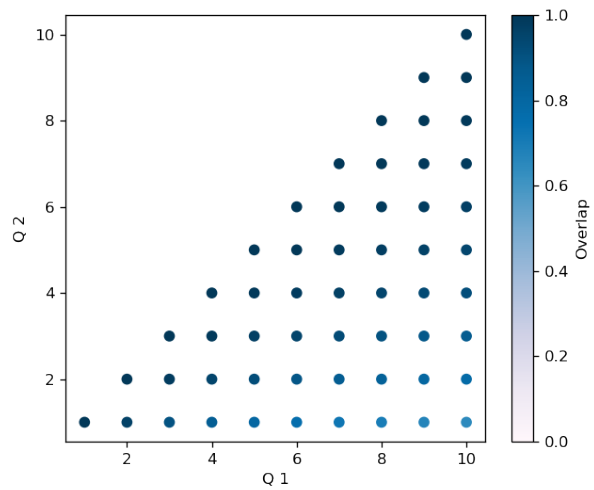


FIG. 5: The parameter space of the Q factors ranging from 1 to 10. The highest overlaps are close to the diagonal line.

I then generated a parameter space for f_{1s} 6. 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. The high match values indicates that there can be a high similarity between the waveforms when one parameter is tweaked. Testing the other parameters t_2 and f_0 as shown in Figs. 7 and 8 respectively. Once again, using the diagonal as a reference of the match being a value of 1, the rest of the parameter spaces of these parameters were close to 1. This indicates that creating a template bank is indeed likely to be feasible when these parameter spaces are able to generate high match values between the waveforms generated. In fact, for example, most of the match values in the f_0 parameter space in Fig. 8 were above 0.8. Thus, the next step was creating a template bank utilizing the whole 7-dimensional parameter space of the CCSN and incorporating the randomization of the impulses.

After this exploration into how the supernova waveform matches hold up with these comparisons, I then moved on to generate a template bank which covers the 7-dimensional space of the supernova waveform. To generate a template bank, I wrote the Python package `SNoW CaPS`: SuperNova Waveforms for Calculating the Parameter Space which follows a sampling algorithm.

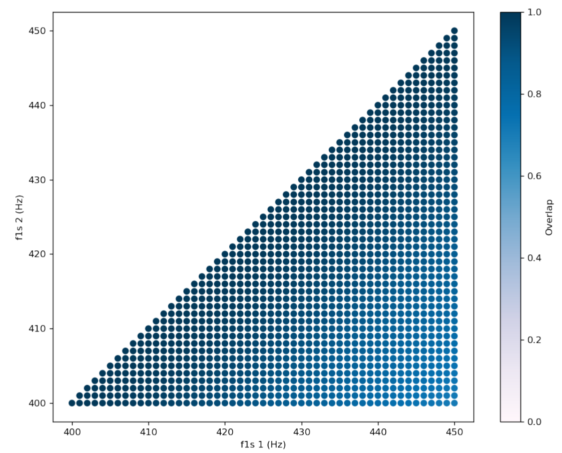


FIG. 6: The parameter space of f_{1s} ranging from 400 to 450 Hz. The diagonal line highlights the overlap of the same waveforms.

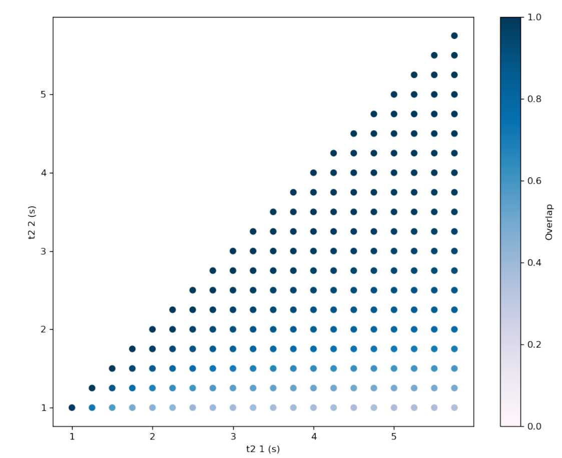


FIG. 7: The parameter space of t_2 ranging from 1 to 10 s. The t_2 has to be bigger than the end time of the bounce. The diagonal line highlights the overlap of the same waveforms.

This package utilizes three modules: `waveform.py`, `snoverlap.py`, and `snowbank.py`. The `waveform.py` file generates one waveform from the Astone et. al [8] paper. The `snoverlap.py` file generates two waveforms using `waveform.py` and calculates the match. Finally, `snowbank.py` calls `snoverlap.py` to generate the template bank. These scripts are distinct but connected parts of `SNoW CaPS` which ultimately generates a template bank based on the parameters the user sets using the command line.

The template bank algorithm in `snowbank.py` starts by

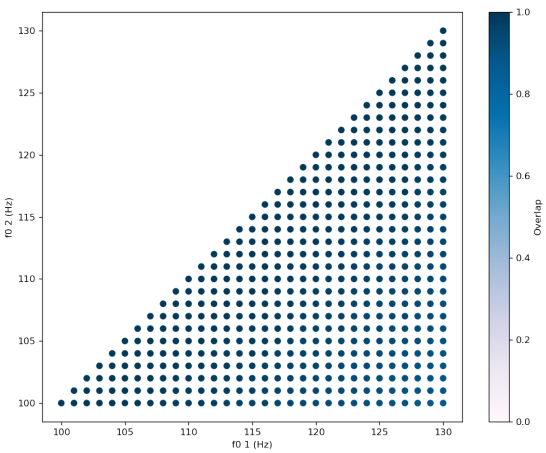


FIG. 8: The parameter space of f_0 ranging from 100 to 130 Hz. The diagonal line highlights the overlap of the same waveforms.

setting a match threshold. Then, it generates an initial waveform using `waveform.py` and stores the parameters. Then, `waveform.py` is called again to generate a second waveform. The match of these two waveforms are then calculated. If the match is less than the threshold, then the parameters are stored. Then, another waveform is generated and the match calculation is repeated with the newly generated waveforms and the stored parameters. However, if the match is greater than the threshold, then the number of successive times the match is greater than the threshold is tracked. If the match is greater than some threshold for some value length of time, then all the stored parameters are saved into a hdf5 file. These stored parameters are what make up the template bank.

When I generated around 10,000 realizations within a time window of 0.05 seconds, I found that I was able to achieve banks with a match greater than 0.97. Thus, generating a template bank with waveforms that have a shorter signal are able to generate high overlaps. Even with randomization, high overlaps are achievable. These high overlaps highlight that the template banks created do not have much signal loss, which means that more of the supernova waveform would be retained when utilizing matched filtering. Comparing these results to current supernova GW detection methods in Szczepańczyk et. al [13], this follows a similar trend of seeing signals that are shorter being more efficient. The shorter signals are more efficient because they are less complex. When the time window is smaller, the randomized impulses has a smaller role on the behavior of the waveform which results in a higher match between the waveforms in the template bank. Thus, matched filtering is feasible and clearly performs similar to other detection methods.

IV. CONCLUSION

I determined that matched-filter searches are feasible for supernova waveforms. High overlaps are achievable for template banks even with randomized inflows and randomized time windows. When comparing the results to other supernova detection methods to the template bank I generated, the shorter signals also perform well. The next step would be to compare the matched filter method to numerical relativity simulations. This way, I could have a deeper analysis of the detection methods and compare the results with the signal-to-noise (SNR) ratio. Future work would also entail generalizing `SNoW CaPS` further to allow the template bank to be completely generated via command line. As of right now, `SNoW CaPS` requires the user to edit the match threshold within the `snowbank.py` file. Finally, the stopping criteria for generating new parameters for the template bank in `SNoW CaPS` can be generalized further.

V. ACKNOWLEDGEMENTS

This work was supported in part by the Research Experiences for Undergraduates (REU) Program of the National Science Foundation. I would also like to 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. Finally, I would also like to thank my mentor Ryan Magee for his guidance and mentorship this summer.

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

[10] S. K. Sahay, "Studies in gravitational wave data analysis," (2002).

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

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

[13] M. J. Szczepań czyk, J. M. Antelis, M. Benjamin, M. Cavaglià, D. Gondek-Rosińska, T. Hansen, S. Klimenko, M. D. Morales, C. Moreno, S. Mukherjee, G. Nurbek, J. Powell, N. Singh, S. Sitmukhambetov, P. Szewczyk, O. Valdez, G. Vedovato, J. Westhouse, M. Zanolin, and Y. Zheng, Physical Review D **104** (2021), 10.1103/physrevd.104.102002.