Inferring Gravitational Wave Source Properties from Intermediate Pipeline Output with Machine Learning

Julianna Levanti, Ryan Magee
LIGO Caltech SURF Final Report

The LIGO-Virgo-KAGRA collaboration provides low-latency (near-real time) localization using the signal-to-noise ratio measured for a single point in the search parameter space. Parameter estimation pipelines subsequently sample the full parameter space to obtain more accurate estimates of the localization. However, this process is computationally expensive. The multi-messenger detection of the binary neutron star merger GW170817 confirmed the need for accurate and fast data products. Some detection pipelines utilize singular value decomposition to reduce the filtering cost. This project uses machine learning to input signal-to-noise ratios from singular value decomposition time series into a simulation-based inference (SBI), a likelihood-free inference algorithm, which outputs a posterior with an accurate parameter estimation, such as a sky map, to localize compact binary coalescences and infer other source properties.

I. INTRODUCTION

Gravitational Waves (GWs), first detected in 2015 by the Laser Interferometer Gravitational-Wave Observatory (LIGO), [1]. GWs are physical ripples in the fabric of space and time, stretching and compressing space. GWs originate from a compact binary coalescence (CBC), the inspiral and merge of two extremely massive objects such as black holes or neutron stars. GWs can also originate from other exotic events in the universe, such as supernovae, but this project will focus solely on CBCs. GWs can be detected by laser interferometers such as LIGO, which uses laser interference to measure the impact of passing GWs [1]. Information encoded in the GW signal detected by LIGO can give scientists valuable information about each source, like distance and location.

The first binary neutron star merger, detected by LIGO in August 2017 [2], was a breakthrough in our understanding of astrophysics. The European Space Agency's INTEGRAL Telescope and NASA's Fermi Gamma-ray Space Telescope observed a brief gamma-ray burst from the source [3]. The Hubble Space Telescope and The Chandra X-ray Telescope also detected electromagnetic (EM) radiation from the same direction [3]. Further evidence shows that the James Webb Space Telescope detected mid-infrared emission of exotic heavy-element tellurium [4]. Analysis of the GW data and the EM counterparts support that the progenitor was most likely a binary neutron star (BNS) merger. The event is important for our understanding of the universe due to its GW and EM counterparts [2, 5–7]. Studies have shown efforts towards low-latency GW detection [8]. This project aims to improve the accuracy of our current low-latency data products.

Pipeline filtering analysis computes a comparison between a large number of modeled CBC waveforms (Figure 1) and the detector output to produce a signal-to-noise ratio (SNR). SNRs can contain GW signals from compact binaries covered by background noise [9]. The data set in Figure 1 contains waveforms accounting for a variety of masses and spins, which encompasses information about source parameters, but seem to look similar in appearance.

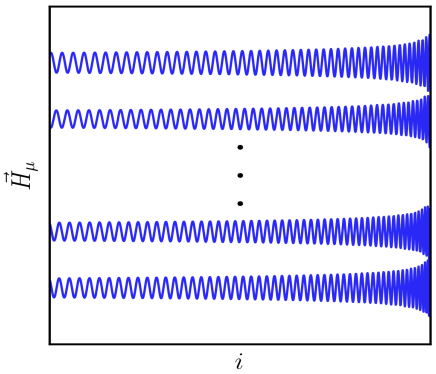

FIG. 1: Simulated waveform candidates. $\vec{H}_\mu$ is a unit less measurement, referred to as strain. Strain is the fractional change in distance inteferometer lasers are stretched or compressed by a passing gravitational wave, relative to the original length.

GW data from CBCs are dependent on a high number of physical dimensions, such as mass, spin, distance, and inclination. Previous studies face issues in computing accurate values for all the dimensions due to the high parameter space [10–12]. GW strain observed on Earth depends on an array of 15 parameters. The parameters are as follows: mass of the primary object, mass of the secondary object, luminosity distance, the integration constant, time of coalescence, position of ascension, position of declination, inclination, polarization angle, the spin, angle, and orientation of the primary object, and the spin, angle, and orientation of the secondary object [13]. Although parameter estimation is difficult to compute in high dimensions, LIGO is able to create low-latency sky maps and measure SNRs by imposing constraints on the signal parameter space.

A. Singular Value Decomposition

The project planned to generate low-latency sky maps accurately by using singular value decomposition (SVD)

(Equation 1) vectors instead of original waveforms. Since search pipelines like gstLAL compute SVD in normal detections, we want to expand the utility of the pipeline to infer source properties directly by using the SVD vectors. With the intuition that the project is successful, SVD can be applied to other pipelines. Using SVD waveforms reduces the amount of GW filtering required to analyze a given region of a parameter space [14].

In Equation 1, h are GW time series, a are reconstruction coefficients, and u are SVD abstract basis vectors, all indexed by μ [14]. Equation 2 represents SNR, ρ, which is equal to the noise-weighted inner product of a waveform template and raw strain data, d. Substituting Equation 1 in for h, we can use distributive properties and replace the new inner product with the SVD SNR, Q, shown in Equation 3, and create SVD SNR Equation 4, which can be compared similarly with SNR Equation 1.

$$h = \sum a_\mu u^\mu \tag{1}$$

$$\rho =< h|d > \tag{2}$$

$$Q =< u|d > \tag{3}$$

$$\rho = \sum a_\mu Q^\mu \tag{4}$$

Equation 1 is a breakdown of each waveform template into SVD basis vectors, whereas Equation 4 is a breakdown of each original SNR into SVD SNRs. SVD basis vectors (Figure 2) are abstract and do not contain any concrete evidence regarding the dimensions we hope to reveal. All of the abstract waveforms are orthogonal and do not overlap with one another [10, 11]. LIGO creates sky maps with SNRs as the input data, but we used SVD SNR, Q, and attempted to infer GW source properties from this abstract SNR.

B. Machine Learning

The modeled waveforms represented in Figure 1 can be compiled into a parameter space shown in Figure 3, where each point represents an individual template. Computing abstract waveforms from SVD that will fit in the gaps of the sample space in Figure 3 through a neural network is an example of how to apply machine learning. A specific number of SVDs can be added together in various ways to create each template in the space, which means that SVDs can include information about each template in the space in Figure 3. Normally, low-latency sky maps are created by selecting one template or point in Figure 3. Sky maps generated by full parameter estimation will input every template in the space. Since the empty space (represented by X in Figure 3) compares to the plotted templates with high accuracy, it is possible to cover the entire area in the space by using the SVDs instead of just a single waveform, or every waveform plotted in the space.

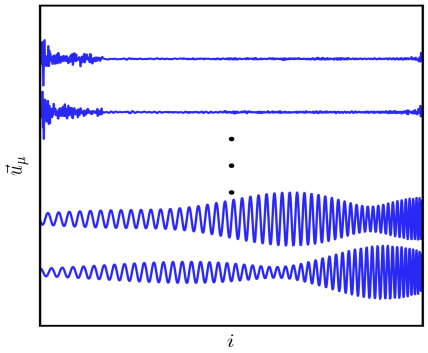

FIG. 2: The result of abstract waveforms computed through SVD using original gravitational waveforms.

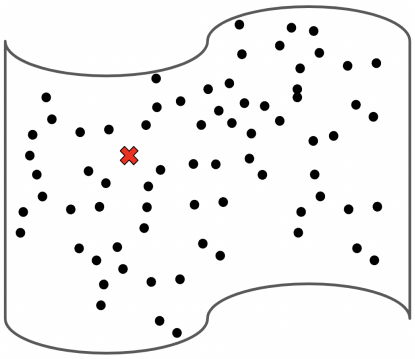

FIG. 3: Search pipelines conduct gravitational wave searches by discretely sampling this continuous signal manifold, which is represented in this shape due to high dimensionality. The manifold represents pre-SVD modeled waveforms under a parameter space. The red X indicates an area where a waveform can be created with the SVD waveforms extracted from Figure 2 with high accuracy.

Parameter estimation relies on Bayes' Theorem and an extensive computation for a likelihood function. In pipelines, waveforms are computed from parameters, which then is transformed into an SNR. All of these calculations produce the proper information that allows for the likelihood computation. Instead of a lengthy calculation, our project looks to input SVD SNRs that are previously computed from detection pipelines, through a neural network to compute a likelihood of the physical properties of a GW source. Using simulation-based inference, machine learning can help us "front-load" the expensive computa-

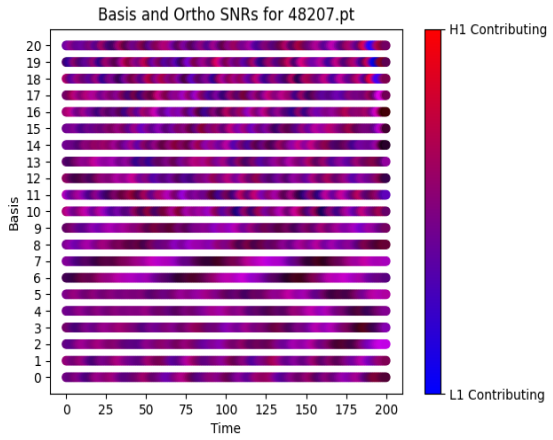

FIG. 4: SVD SNRs from H1 and L1 data, represented as an input "image" for the convolutional neural network.

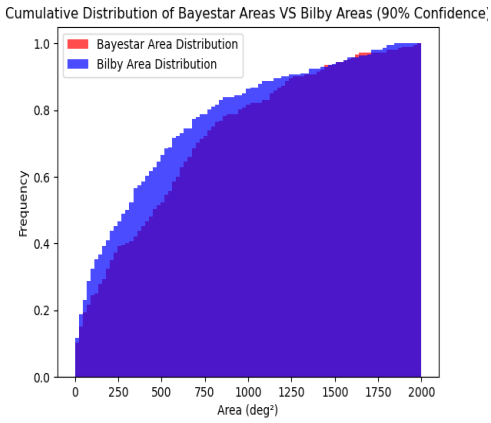

FIG. 5: A cumulative area distribution comparison of BILBY and BAYESTAR results. BILBY surpasses BAYESTAR with higher frequencies at smaller areas of degrees^2.

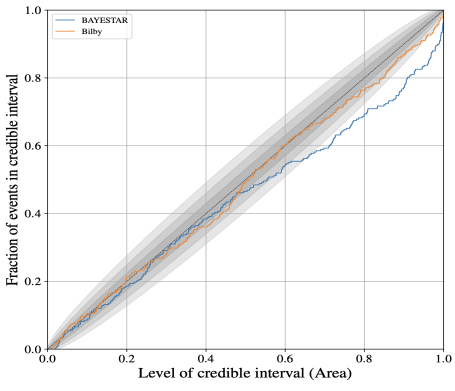

FIG. 6: The Probability-Probability (PP) plot visually compares the two algorithms and how well they perform based on the theoretical model, comparing the confidence level on the x-axis, and the probability that the true value falls within the confidence level on the y-axis [18].

tions to build a posterior and calculate the likelihood of parameters, given simulated parameters and a large set of injection data. The project uses a convolutional neural network to connect abstract SNRs that are produced by SVD to the physical properties of GW candidates. Convolutional neural networks (CNNs) are common in image classification and are used to categorize patterns and shapes in data sets. The convolution layer takes a kernel, slides it across our input data, and computes a dot product with the weights represented in the kernel. The output is a feature map, which represents specific and common patterns in the data. The pooling layer performs max pooling, where a kernel slides across the feature map, and selects the largest value within that frame. Pooling reduces dimensionality, computational load, and over-fitting. We have adjusted 8 factors included in the CNN such as the kernel size and learning rate, to achieve accurate posterior and sky map results. The use of a neural network has been confirmed and tested as a reliable structure for a machine learning algorithm [15].

II. MOTIVATION

We researched two sky map algorithms. BAYESTAR, which computes low-latency sky maps by assuming a singular template [16], as represented in Figure 3, and BILBY, which takes many templates into account for a better localized sky map [17]. However, BILBY has a computing time on the order of hours to days. We extracted data in square degrees and compared the two systems in a 90% confidence area. To filter data and study true results, criteria were set at SNR signal greater than 9, and a false alarm ratio (FAR) less than 1 yr^{-1} or less than 2 yr^{-1}. BAYESTAR and BILBY data is pulled from mock data challenges (MDCs) that were run in the lead-up to LIGO's fourth observing run. MDCs look at all uploads done to a low-latency database and compares the two sets of sky maps.

In Figures 5 and 6, evidence shows that BILBY's complex computations can produce a more accurate sky map than BAYESTAR's low-latency algorithm. Figure 5 shows that BILBY computes smaller areas to the square degree than BAYESTAR, and Figure 6 shows that BILBY performs closer to the theoretical model, which tells us there is room for improvement in low-latency. The project strives to achieve each system's strengths of low-latency and more accurate localization with the use of SVDs and machine learning.

III. METHODS

The project goal is to generate low-latency sky maps more accurately by 1) using abstract waveforms formulated by singular value decomposition already present in data pipelines and 2) utilizing machine learning techniques such as simulation-based inference (SBI) that would reduce computational costs.

SVD SNRs were ingested into an SBI-motivated framework that included the CNN to compile a parameter estimation and demonstrate parameter probability. We ran simulations to train the algorithm, allowed it to learn the shape of our current data, and built a posterior. From this, we compared simulated source parameters and tested the validity of the posterior. Our final goal is to construct an accurate sky map showing the 50% and 90% confidence areas of the sky where the source is located. Sky maps were generated using the inferred parameters, right-ascension, and declination.

IV. RESULTS

By utilizing LALSuite, a public collection of gravitational wave software, we set boundaries for each parameter and simulated injections with varying parameters to binary coalescence gravitational wave sources. Initially, a fully connected neural network was used for processing data in the machine learning framework.

Due to the high level of dimensions applicable to binary coalescences, we began testing by setting all parameters in the LALApps configuration to fixed or 0, which allowed distance to be a free parameter, with boundaries of 100 to 1000 megaparsecs (mpc). Varying sample sizes from 10 000 injections (5000 from each interferometer) were processed to each include a varying distance and an SVD time series, and added into a training network, which resulted in a posterior. Plotting posteriors demonstrated about 800 samples were needed to train the network to achieve accurate results. To test significance, we used measured SVD responses for known sources and compared how well our network predicted the true value. Randomly selected true distance values were displayed against posteriors. 27% of true values lay outside the predicted distribution (73% accuracy rate). Half of the true values that fell outside the posterior were underestimated due to the parameter boundary in place, for there are signals that have been detected from distances further than 1000 mpc. If a larger distribution of distances were included in the training, posteriors would be underestimating true values at a smaller percentage. We increased this prior boundary to 5000 mpc in the following tests.

After one-dimensional training and plotting success, a second free dimension, inclination, was added to the injections. With a total of 10 000 injections, 3250 samples were used to draw accurate posteriors. Approximately 15% of true values fell outside the inclination distribution while zero true values fell outside the distance distribution, giving us a high parameter accuracy rate of 85% accuracy rate. To check for possible bad data, more posteriors were created using data from only one of LIGO's observatories, H1 (Handford). Between posteriors drawn from H1 solely compared with H1 and L1 (Livingston), distributions had an increase in variance for H1 and similar accuracy when predicting true values. For the following tests, we have used data from both H1 and L1.

To set forth our goal of an accurate sky map, the next trial included two localizing parameters, right ascension and declination. 60 000 total injections were created and 9000 samples were drawn to plot accurate posteriors. Out of 60 parameters in 4 dimensions, 23% of true values fell outside the posterior (77% accuracy rate). Initial observations from the 4-dimensional posteriors include the lack of stability of right ascension and inclination distribution. Even with an increased number of samples, there is a high variance in the distribution.

The posteriors drawn in the first three trials have an average accuracy of about 78.3%, however, the final goal is to reach a 90% accuracy rate. Furthermore precision of the right ascension and inclination distributions still struggled. Due to these observations, we planned to increase both precision and accuracy by adjusting the neural network format to a CNN. Due to the initial formatting required by SBI, SVD data matrices were compressed into a one-dimensional list. This formatting can forego some vital information about our data, where the framework can struggle with separating each time series from the next. Although the neural network is performing well for accuracy, adjustments were made to preserve matrix structure, in hopes of better results in accuracy and precision.

After adjustments to the neural network, initial observations show an accuracy rate for various sample sizes of 83% or greater. A sample size of 15 000 out of 50 000 injections represented the greatest accuracy and precision. Each distribution is mid-narrow in precision, with a total accuracy rate of 90%. As expected, distance and declination are precise and accurate. We notice a bimodal variance for right ascension distributions and uniform distributions for inclination. However, a bimodal distribution is excepted for inclination. Through each adjustment to the neural network, sample sizes, and reparameterizing, our accuracy rate generally increased since the initial trial. At this point in the process, we trusted that the neural network was performing at its best.

Using the previous posterior distributions discussed, we were able to use the right ascension and declination and directly plot a sky map for each injection's inferred posterior. A generated sky map is shown in Figure 7. The model created can identify the true localization but has a hard time calculating more precise confidence areas.

The wide confidence areas calculated can be due to several reasons, possibly a poor neural network architecture, noise realization, or maybe ridding of important information when using SVDs that are unaccounted for. The project's future direction includes using data of the

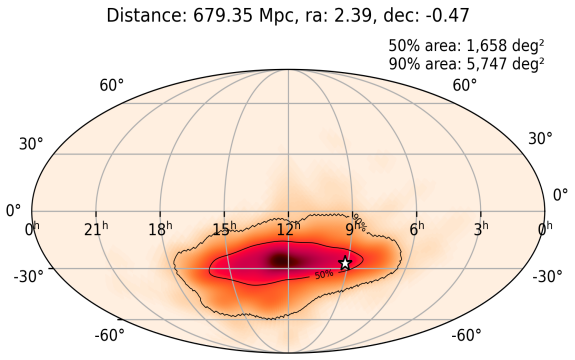

FIG. 7: A sky map is produced directly using posterior distributions from an observed injection in the sample set. The title represents the true values of the injection. The 50% and 90% confidence areas are also computed.

time delay detected by LIGO as a tool for inferring the parameter estimations we wish to produce in our neural network. Due to the geographical separation distance of the detectors, as a GW passes through Earth, there is a difference in time from when each detector experiences the GW. This is defined as the time of the detection at the instrument, which equals the time the wave passes through Earth's center, plus a delay that depends on the right ascension and declination of the source. The right ascension and declination of a GW source can affect the time delays between detections at different LIGO sites, and these time delays are crucial pieces of information used to determine the properties and location of the source.

V. IMPLICATIONS

Data attainment and detection times have been minimized in recent research. Data interpretation still requires extreme expenses and effort. It is important to localize sources quickly and allow for EM counterpart follow-up to test and study the connections between GWs and EM radiation. The project offers an open door into a new interpretation methodology to infer gravitational wave source properties fast with accurate precision and has made great strides in creating accurate, low-latency sky maps.

As the next generation of GW observatories rises, low-latency data interpretation will be imperative for the possible detections scientists foresee with newer technologies. The Einstein Telescope (ET) will attempt to detect the GW and short gamma-ray bursts of BNS mergers at high redshift [19]. The Laser Interferometer Space Antenna (LISA) will attempt to detect GW inspirals from distant CBCs including BNS [20]. Both projects, aiming to locate GWs and EM can succeed even further when accurately localizing sources in low-latency. The goal to calculate the physical properties of the CBC from the abstract waveforms efficiently will allow instantaneous review of possible overlaps in GW with EM data.

This work was supported by the National Science Foundation Research Experience for Undergraduates (NSF REU) program, the LIGO Laboratory Summer Undergraduate Research Fellowship program (NSF LIGO), and the California Institute of Technology Student-Faculty Programs.

[1] B. P. Abbott et al. (LIGO Scientific), Phys. Rev. D **95**, 062003 (2017), arXiv:1602.03845 [gr-qc].
[2] B. P. Abbott et al. (LIGO Scientific, Virgo), Phys. Rev. Lett. **119**, 161101 (2017), arXiv:1710.05832 [gr-qc].
[3] NASA ESA (2017).
[4] A. Levan et al., (2023), arXiv:2307.02098 [astro-ph.HE].
[5] R. Magee et al., Astrophys. J. Lett. **910**, L21 (2021), arXiv:2102.04555 [astro-ph.HE].
[6] B. P. Abbott, R. Abbott, T. D. Abbott, Abernathy, and others., .
[7] B. P. Abbott et al. (LIGO Scientific, Virgo, Fermi GBM, and others), Astrophys. J. Lett. **848**, L12 (2017), arXiv:1710.05833 [astro-ph.HE].
[8] K. Cannon et al., Astrophys. J. **748**, 136 (2012), arXiv:1107.2665 [astro-ph.IM].
[9] A. Reza, A. Dasgupta, and A. S. Sengupta, (2021), arXiv:2101.03226 [gr-qc].
[10] K. Cannon, C. Hanna, and D. Keppel, Phys. Rev. D **84**, 084003 (2011), arXiv:1101.4939 [gr-qc].
[11] K. Cannon, C. Hanna, and D. Keppel, Phys. Rev. D **85**, 081504 (2012), arXiv:1108.5618 [gr-qc].
[12] J. Veitch et al., Phys. Rev. D **91**, 042003 (2015), arXiv:1409.7215 [gr-qc].
[13] V. Raymond, M. V. van der Sluys, I. Mandel, V. Kalogera, C. Rover, and N. Christensen, Class. Quant. Grav. **27**, 114009 (2010), arXiv:0912.3746 [gr-qc].
[14] K. Cannon, A. Chapman, C. Hanna, D. Keppel, A. C. Searle, and A. J. Weinstein, Phys. Rev. D **82**, 044025 (2010), arXiv:1005.0012 [gr-qc].
[15] R. Qiu, P. G. Krastev, K. Gill, and E. Berger, Phys. Lett. B **840**, 137850 (2023), arXiv:2210.15888 [astro-ph.IM].
[16] L. P. Singer and L. R. Price, Phys. Rev. D **93**, 024013 (2016), arXiv:1508.03634 [gr-qc].
[17] P. Baral, S. Morisaki, I. Magaña Hernandez, and J. Creighton, (2023), arXiv:2304.09889 [astro-ph.HE].
[18] S. S. Chaudhary et al., (2023), arXiv:2308.04545 [astro-ph.HE].
[19] Ronchini, S., Branchesi, M., Oganesyan, G., Banerjee, B., Dupletsa, U., Ghirlanda, G., Harms, J., Mapelli, M., and Santoliquido, F., A&A **665**, A97 (2022).

[20] LISA: NASA, ESA.