



This is the accepted manuscript made available via CHORUS. The article has been published as:

Optimizing the placement of numerical relativity simulations using a mismatch predicting neural network

Deborah Ferguson

Phys. Rev. D **107**, 024034 — Published 24 January 2023

DOI: [10.1103/PhysRevD.107.024034](https://doi.org/10.1103/PhysRevD.107.024034)

Optimizing the Placement of Numerical Relativity Simulations using a Mismatch Predicting Neural Network

Deborah Ferguson¹

¹*Center for Gravitational Physics and Department of Physics,
The University of Texas at Austin, Austin, TX 78712*

Gravitational wave observations from merging compact objects are becoming commonplace, and as detectors improve and gravitational wave sources become more varied, it is increasingly important to have dense and expansive template banks of predicted gravitational waveforms. Since numerical relativity is the only way to fully solve the non-linear merger regime of general relativity for comparably massed systems, numerical relativity simulations are critical for gravitational wave detection and analysis. These simulations are computationally expensive, with each simulation placing one point within the high dimensional parameter space of binary black hole coalescences. This makes it important to have a method of placing new simulations in ways that use our computational resources optimally while ensuring sufficient coverage of the parameter space. Accomplishing this requires predicting the impact of a new set of parameters before performing the simulation. To this effect, this paper introduces a neural network to predict the mismatch between the gravitational waves of two binary systems. Using this network, we then show how we can propose new numerical relativity simulations that will provide the most benefit. We also use the network to identify gaps in existing public catalogs and identify degeneracies in the binary black hole parameter space.

I. INTRODUCTION

As of the end of its third observing run, the LIGO-Virgo-KAGRA (LVK) Collaboration has detected 90 gravitational wave (GW) signals from merging binary systems [1–6]. With the LVK’s fourth observing run, anticipated to begin in 2023, and the fifth observing run a few years later, we expect to observe hundreds of more signals with higher signal-to-noise ratios (SNRs). Beyond current detectors, in the coming decades, we also expect to see next generation GW detectors come online, including the space-based Laser Interferometer Space Antenna (LISA) [7] and the ground-based Einstein Telescope (ET) [8, 9] and Cosmic Explorer (CE) [10]. All of these new detectors will have unprecedented sensitivity and are expected to see orders of magnitude more signals than we have to date [11, 12]. LISA will be sensitive to lower frequencies than the ground-based detectors, giving us a window into entirely different sources, including supermassive black hole (BH) binaries which may remain in the frequency band for months to years.

Detecting and interpreting these events requires understanding the expected GW signatures of astrophysical events [13–15]. While analytic solutions exist for single BHs, binary black holes (BBHs) have no analytic solution for full, nonlinear general relativity (GR). Many methods have been developed to find solutions to the BBH problem. When the BHs are far apart, Post Newtonian (PN) can be used to compute GW waveforms [16–24]. In the highly unequal mass ratio regime, self-force approximations can be used to solve BBH systems and output expected waveforms [25–38].

However, in the highly non-linear regime of the merger of comparably massed systems, numerical relativity (NR) is the most accurate approach [39, 40]. NR computationally solves Einstein’s equations by performing a 3+1

decomposition, such as in the BSSN formulation [41], in order to evolve spacelike surfaces along a timelike path. Many NR codes, including those based on the Einstein Toolkit (ETK), use finite differencing to solve 2nd-order non-linear PDEs [39, 42]. This enables them to solve for the dynamics of the system and the gravitational radiation emitted [39, 42]. These simulations require significant computational resources; each simulation takes an initial BBH configuration and can run for weeks to months, placing one point in the parameter space of BBHs. When considering quasicircular, vacuum BBH systems, the parameter space is 7 dimensional, consisting of the spin vectors for each of the initial BHs as well as the ratio of the masses of the initial BHs. This dimensionality increases when considering eccentric systems or systems with matter.

Detection pipelines and parameter estimation tools require dense, accurate template banks to detect and interpret the events from the observed detector data [43–49]. Due to the high computational cost of NR, in order to produce template banks efficiently and with arbitrary density, analytic models are constructed using information from NR [14, 50–55]. In some cases, NR waveforms are used directly in the analysis [56, 57], to avoid systematic errors from the models and/or to provide parameter coverage not yet available in models. In either situation, NR waveforms are an integral part of GW astronomy; therefore, it is crucial to have a dense bank of NR simulations which cover the desired parameter space. Current public NR catalogs [58–63] have dense coverage for comparable mass systems, even for those with spins misaligned with the orbital angular momentum. They are lacking coverage in the precessing space for more unequal massed systems and have only a trace number of simulations where one BH is more than 10 times the mass of the other. While this coverage has proved acceptable for current GW detector sensitivities, the incredible de-

velopments expected in future GW detectors will require denser, more accurate, and more expansive waveform catalogs [64]. Achieving the accuracy needed for future GW events requires very high resolutions in current NR codes. This becomes even more significant for unequal mass and highly spinning BHs as both of these cases require higher resolution to resolve the spacetime around the BHs; this increases the computational cost significantly.

Given the high computational cost and time consuming nature of NR simulations, particularly at high resolution, it is imperative to construct NR catalogs carefully with their cost and use in mind. In order to be more strategic about NR simulation placement, we need to understand which simulations will be most beneficial and which regions of parameter space will need to be populated more densely. This requires a method of predicting how important a proposed simulation will be before performing it. Analytic models can be used to this effect by efficiently predicting waveforms for given initial parameters [14, 50–55]. Surrogate models can also be used for this purpose as they interpolate between existing simulations to predict waveforms for given parameters [65]. Using these models to predict the importance of a proposed simulation would require creating a waveform and then comparing it to all existing NR waveforms, a very costly endeavor. The use of models also adds in potential systematic biases dependent upon the waveform model used. Work has also been done to use the uncertainty in models to suggest new simulation parameters [66]

We introduce a more efficient and independent method of estimating the significance of a new simulation using a neural network. First emerging in 1944 and resurging in the 1980s, neural networks are a type of machine learning that are able to recognize patterns and make predictions based on input values [67–71]. The networks consist of a series of layers, each of which may contain many nodes. Each node functions as a single logistic regression unit, using weights and biases to input many values and output one. Arranging these nodes in layers allows the network to recognize complex patterns and relationships. Training the network entails optimizing the weights and biases of the nodes in order for the network to output the desired values for labeled training data. This process begins by initializing the weights and biases randomly and then propagating the input parameters of the training data to obtain the network output. This is compared to the desired output for the training data, and backpropagation is used to update the weights and biases accordingly. This process is repeated until the accuracy reaches a plateau or a specified desired value. In recent years, neural networks have been used in the GW and NR fields [72–74].

In this paper, we introduce a neural network to predict the mismatch between two BBH systems given only the initial parameters of the binaries. The mismatch is frequently used in data analysis to measure the difference between two waveforms, so this allows us to gauge how different a proposed simulation would be from those

in existing NR catalogs. Having a neural network that predicts the mismatch avoids the need to generate waveforms and perform costly mismatch calculations. This also allows numerical relativists to use and even train this network on new NR simulations in regions of parameter space that models don’t yet cover. For instance, while this network is currently trained on quasicircular simulations, it can be extended and retrained on eccentric simulations.

This paper begins by introducing the match prediction network. The network, its training data, and its accuracy are outlined in Section II.

Section III goes through the results of using this network to study and fill in the parameter space. The primary motivation for the creation of this network is to optimally place NR simulations to fill in catalogs efficiently. This implementation is described in Section III A. The network also provides us more insight into the behavior of waveforms throughout the parameter space. Using this, in Section III B, we note gaps in the current coverage of public NR waveform catalogs. Section III C uses the network to identify degeneracies in parameter space where significantly different parameters can lead to extremely similar waveforms.

The model described in this paper is available for use at https://github.com/deborahferguson/mismatch_prediction.

II. NEURAL NETWORK TO PREDICT MATCH

In order to identify regions in parameter space that would provide significant scientific benefit if filled, we first need to define a metric to assess how different a resulting waveform would be from existing waveforms. For this paper, we use the mismatch as our metric [75, 76], defined as:

$$\mathcal{MM} = 1 - \max_{t, \phi} \mathcal{O}[h_1, h_2] \equiv 1 - \max_{t, \phi} \frac{\langle h_1 | h_2 \rangle}{\sqrt{\langle h_1 | h_1 \rangle \langle h_2 | h_2 \rangle}}, \quad (1)$$

where

$$\langle h_1 | h_2 \rangle = 2 \int_{f_0}^{\infty} \frac{h_1^* h_2 + h_1 h_2^*}{S_n} df. \quad (2)$$

h_1 and h_2 are the frequency domain strain of the gravitational waves, and S_n is the one-sided power spectral density of the detector. The mismatch is normalized to $0 \leq \mathcal{MM} \leq 1$ with $\mathcal{MM} = 0$ corresponding to two identical waveforms.

With this measure defined, we now need a way to predict it when the waveforms for one or both of the binary systems have not been generated. This section introduces a neural network that inputs the initial parameters of two quasicircular BBH systems and outputs their predicted mismatch.

Quasicircular BBH systems are uniquely described by the masses and spins of their component BHs. All quantities pertaining to the simulations of vacuum BBHs can

scale with the total mass, so in NR, the total mass is typically set to 1. We can then define the symmetric mass ratio, $\eta = m_1 m_2 / (m_1 + m_2)^2$, reducing our parameter space by 1 dimension. η is often used when characterizing the frequency evolution of the inspiral of two BHs and is bounded to $0 \leq \eta \leq 0.25$, where $\eta = 0.25$ corresponds to an equal mass system. We will use the dimensionless spin vectors, $\mathbf{a} = \mathbf{J}/m^2$ where \mathbf{J} is the angular momentum of the BH. The magnitude of the dimensionless spin vector is constrained to $0 < a < 1$.

Let us define $\lambda = \eta, \mathbf{a}_1, \mathbf{a}_2$ to be the initial parameters for a single BBH system. This network inputs λ_1 and λ_2 and outputs the mismatch between the GWs each system would emit. This structure (and the architecture of the network as described in Section II B) can be seen in Fig. 1.

In Section II A, we describe the training data, including the NR waveforms used and the details of the match computation. Section II B describes the architecture of the neural network, and we discuss the accuracy of the network in Section II C.

A. Training Data

This network is trained on public NR waveforms from the SXS catalog [61]. The uncertainty in the mismatch due to computational limitations within the training data, including finite resolution and extrapolation, is of order 10^{-4} .

The training data consists of BBH systems with $0.0826 \leq \eta \leq 0.25$ with spin magnitudes $0 \leq a_1, a_2 \leq 0.9695$ pointing in various directions. As the symmetric mass ratio decreases, the coverage in the spin parameters becomes less dense. We discuss how this coverage affects the error in the neural network in Section II C.

For BBH systems where one or both of the BH spins are not aligned or antialigned with the angular momentum, the spins and the orbital plane precess throughout the evolution of the system. In order to uniquely identify a precessing simulation, the spins must be specified at a given reference point in the simulation; a reference frequency is generally used for this purpose. To identify the spins at a given reference frequency, we need simulations that reach the specified frequency and contain spin data at that frequency. Gravitational wave frequency scales with the total mass of the binary, so we define our frequency cutoff in terms of a system with total mass $M_{tot} = 1M_\odot$. We select $f_{ref} = 1840$ Hz as our reference frequency for spins and as the frequency at which to begin our mismatch calculation. For the starting frequency of 20 Hz used by many LVK analyses, this would correspond to a total detector frame mass of $92M_\odot$. This was chosen so as to include 95% of waveforms in the SXS public catalog.

Selecting only waveforms that have a starting frequency below f_{ref} and contain spin data at f_{ref} leads to 1885 waveforms. These are then split into training,

development, and test sets with ratios of 0.8/0.1/0.1 respectively. Fig. 2 shows the coverage of the parameter space for the NR waveforms used for this analysis.

Using PyCBC [44] and LALSimulation [77], we compute the mismatch between all pairs of waveforms within each set on a flat noise curve beginning at f_{ref} . This allows us to study the parameter space independently of detector and total mass, though we anticipate training versions of this network on detector noise curves in future work. For this network, we use only the $\ell = 2, m = 2$ mode to simplify the mismatch calculation and allow for a phase shift to be used to maximize the overlap. It also removes the dependence on binary orientation. While considering only the dominant mode will have an impact on the mismatches for systems where higher modes are significant, it is still informative for exploring the parameter space efficiently, so we reserve the consideration of higher modes for future work. We define the spin vectors at f_{ref} , rotated such that the BHs are separated along the x-axis and the orbital frequency lies along the z-axis.

B. Network Architecture

In order to predict the mismatch, we create a fully connected neural network using TensorFlow [78] and Keras [79]. The network consists of 15 hidden layers, each with 56 nodes, as shown in Fig. 1. This structure was chosen after exploratory models were created across a grid consisting of the number of hidden layers and nodes per layer. The development error did not change dramatically for most of the configurations, and the architecture resulting in the lowest development error was chosen. Further tuning of the network architecture will likely be a topic of future work.

The network was trained using an Adam optimizer [80] with a learning rate of 0.001 for 100 epochs using a batch size of 16. In order to add regularization and reduce overfitting, we impose a constrained max norm value of 3. Each hidden layer uses a ReLU activation function [81] and the final layer uses a sigmoid function to ensure the resulting mismatch is constrained between 0 and 1. Predicting mismatches with the network takes very little time since forward propagation through the network can be efficiently parallelized. The network takes about 3 seconds to load, and then each match requires approximately 3×10^{-5} seconds to compute.

C. Network Accuracy

To compute the accuracy of the network on data it has not seen, we take the simulations reserved for the development set and compute the mismatches between all possible simulation pairs. These mismatches are compared to the predicted mismatches and the mean absolute error is computed. This is also done for the test set. After training the network for 100 epochs, it has a

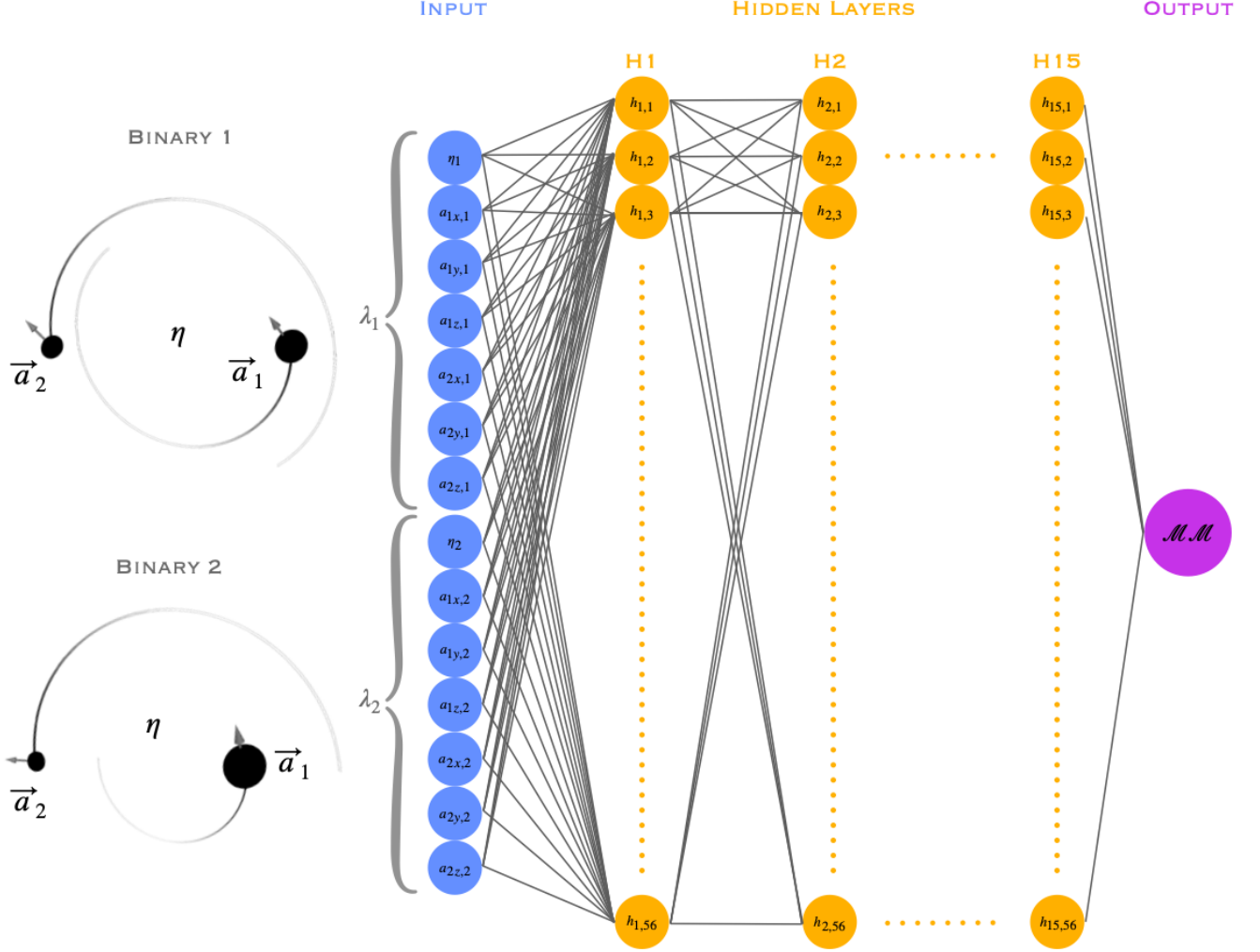


FIG. 1: Diagram showing the structure of the neural network including the input parameters (η , a_{1x} , a_{1y} , a_{1z} , a_{2x} , a_{2y} , a_{2z} for each of the two binary systems) and the output (the mismatch between the gravitational waves emitted from the two systems). The network consists of 15 hidden layers, each with 56 nodes.

training error of 0.00967, a development error of 0.0157, and a test error of 0.018. Therefore, our uncertainty in our mismatch prediction is 0.018. Fig. 3a shows the predicted mismatch versus the computed mismatch for the test set. In the absence of all error, the points would fall on the $y = x$ line. Fig. 3b shows the distribution of the error in the predicted mismatch for the test set.

This network inputs λ_1 and λ_2 resulting in 14 input parameters. The mismatch is order independent so the order of λ_1 and λ_2 in the input layer should not affect the predicted mismatch. After training the network, when swapping the order of λ_1 and λ_2 , we find that the average difference between the predicted mismatches is 0.016, which is less than the network's uncertainty.

Since the training data is not evenly dispersed throughout the entire parameter space, it is expected that the error in the network is not constant in all regions of parameter space. Figures 4 - 8 show the mean absolute

error binned by the various input parameters to the network. Since the network inputs the parameters for both binaries, several of the Figures 4 - 8 overlay the error distributions associated with each of the two systems, with the first system in blue and the second in orange.

Fig. 4 shows the mean absolute error of the model binned by the symmetric mass ratio with bin widths of 0.02. The average error goes up with decreasing η , as is expected due to the scarcity of data points at low η . There is also an interesting spike in the error around $\eta = 0.16$. Referring back to Fig. 2, it is apparent that the density of points drops off by $\eta < 0.16$, so it is expected that the error would be high there, though the reason for the specific peak is not obvious.

In Fig. 5, we see the mean absolute error binned by the difference in the symmetric mass ratio between the two binary systems. A peak is apparent around $\Delta\eta = 0.09$. This is consistent with the peak at $\eta = 0.16$ since most

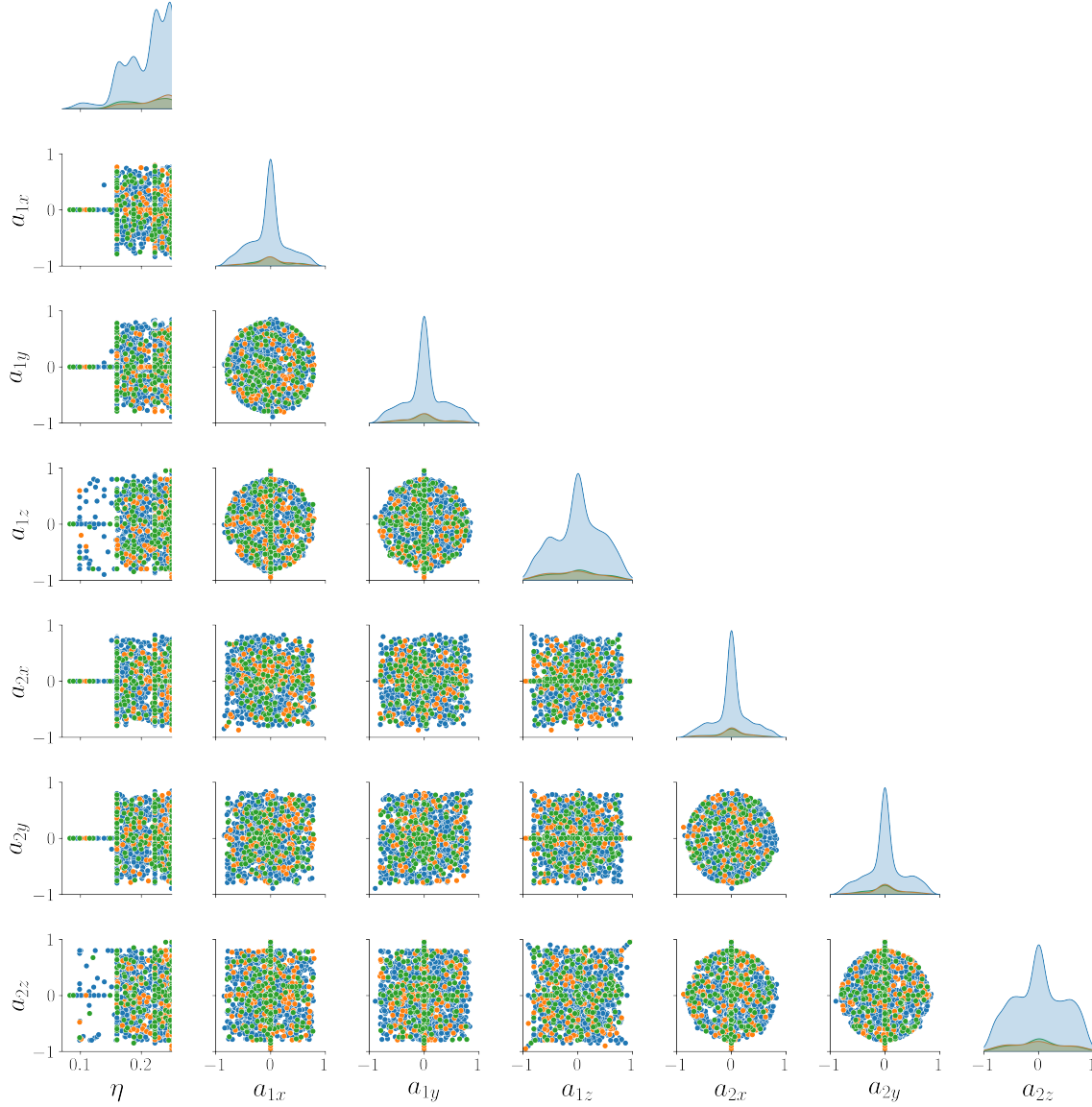


FIG. 2: Coverage of the parameter space by NR simulations in training (blue), development (orange), and test sets (green).

simulations fall at $\eta = 0.25$. For cases that have a binary with $\eta = 0.16$, we will have many points with high error at $\Delta\eta = 0.09$.

In Fig. 6, the data is binned by the spin magnitudes for both the primary and secondary black holes with bins of width 0.1. For the primary black hole, the error grows with increasing spin magnitude, peaking at $a_1 \approx 0.85$. The error binned by the secondary spin magnitude shows a wide peak around $a_1 \approx 0.7$. In both cases, there is a minimum error around $a_{1,2} \approx 0.25$.

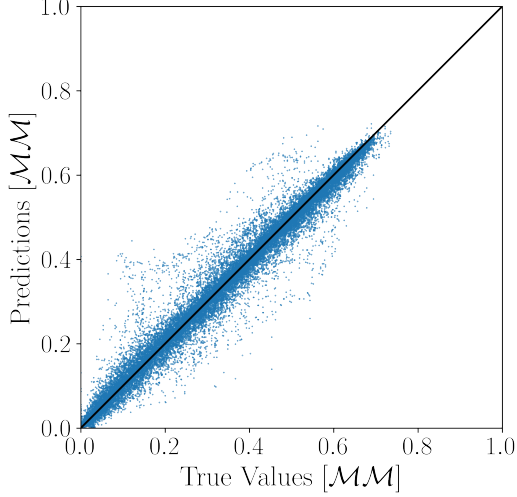
The first panel of Fig. 7 bins the data by χ_{eff} , a quantity used to represent the effective spin of the binary

aligned with the orbital angular momentum [82]:

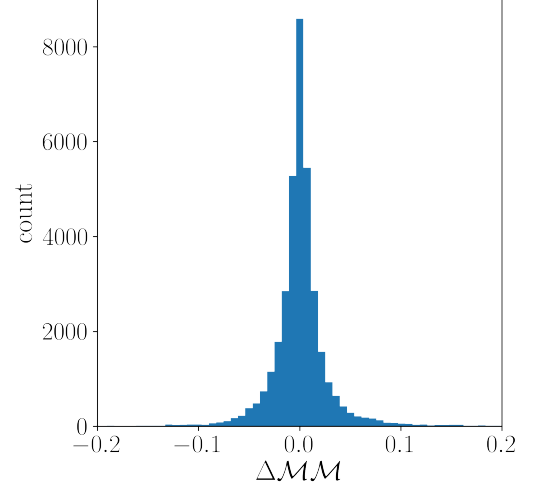
$$\chi_{eff} = \frac{m_1 a_1 \cos(\theta_1) + m_2 a_2 \cos(\theta_2)}{m_1 + m_2} \quad (3)$$

where θ_1 and θ_2 are the angles of the BH spin vectors with the orbital angular momentum. χ_{eff} ranges from -1 to 1, with $\chi_{eff} = 1$ occurring when both BHs are maximally spinning and aligned with the orbital angular momentum and $\chi_{eff} = -1$ occurring when both BHs are maximally spinning and anti-aligned with the orbital angular momentum.

The second panel of Fig. 7 bins the data by χ_p , a parameter used to describe the precession of a bi-



(a) The mismatches predicted by the network versus the mismatches computed from the waveforms in the test set.



(b) The distribution of the error in the predicted mismatches for waveforms in the test set.

FIG. 3: The error in the predicted mismatch for the waveforms in the test set.

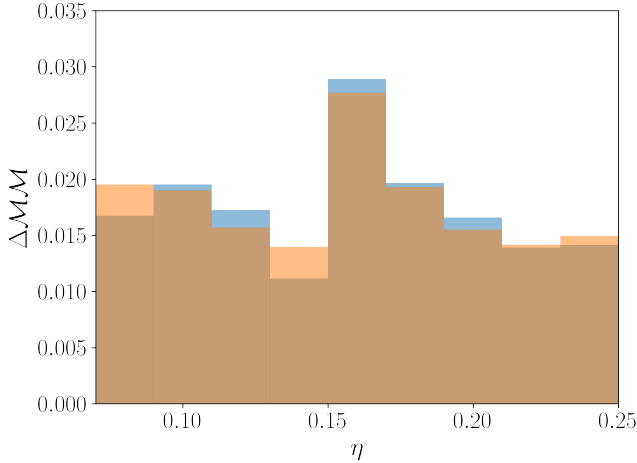


FIG. 4: The mean absolute error in the mismatch for the test set binned by the symmetric mass ratio. The blue represents the first system and the orange represents the second system.

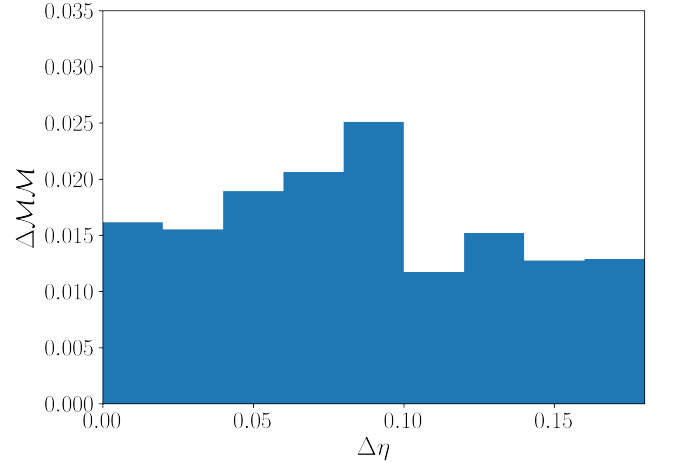


FIG. 5: The mean absolute error in the mismatch for the test set binned by the difference in the symmetric mass ratio between the two systems.

nary [50, 83–85]:

$$\chi_p = \max \left(a_1 \sin(\theta_1), \frac{4m_2^2 + 3m_1m_2}{4m_1^2 + 3m_1m_2} a_2 \theta_2 \right). \quad (4)$$

When the spins of the component BHs are not aligned with the orbital angular momentum, they drive a precession of the orbital plane; χ_p quantifies this effect. Ranging from 0 to 1, a value of $\chi_p = 0$ signifies no precession, i.e. any spin is aligned with the orbital angular momentum and the orbital plane will not precess, and a value

of $\chi_p = 1$ denotes maximal precession.

Other than the dip at $\chi_{eff} \approx -1$, the error seems to be higher at strongly negative χ_{eff} and lower at positive χ_{eff} . This may be explained by the fewer data points at negative χ_{eff} than positive. The error seems to generally increase as χ_p increases, though it seems to peak at around $\chi_p \approx 0.6$. This is expected due to the relative scarcity of precessing training points at high χ_p , particularly as η decreases.

Fig. 8 shows the error when the data is binned by each of the components of the primary (a) and secondary (b)

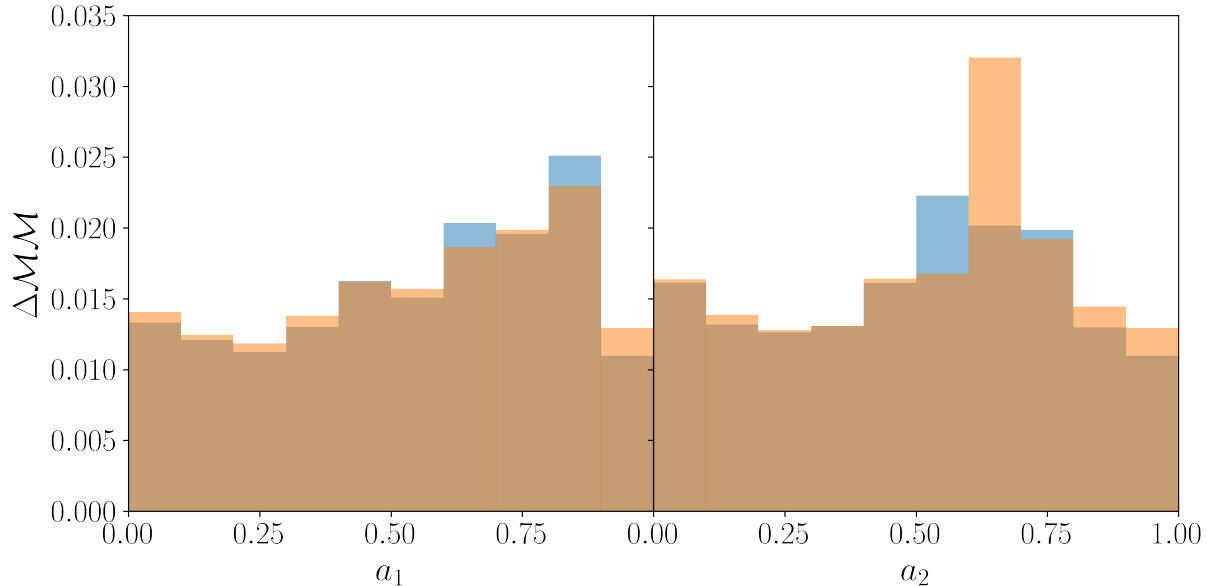


FIG. 6: The mean absolute error in the mismatch for the test set binned by the magnitudes of the spins of the primary (left) and secondary (right) black holes. The blue represents the first system and the orange represents the second system.

spin. For both the primary and secondary black hole, the error seems to grow as the x- or y-component of the spin moves away from 0. These results are anticipated and consistent with the errors when binned by χ_p . For the primary black hole, the error seems to grow at negative a_z , consistent with the results for χ_{eff} . The trends seem more pronounced for the primary black hole than the secondary black hole, which is expected given that the spin of the primary black hole is known to play a larger role in the shape of the waveform.

Looking specifically at outlier points with error greater than 0.2, we find that the average χ_{eff} is strongly negative and χ_p is higher than in the dataset as a whole. The trends in the error appear consistent with expectations due to the density of training points throughout the parameter space. As we perform new simulations and add them to the training data, we expect these errors to decrease.

III. RESULTS

This network enables us to explore the parameter space more efficiently and fill it in more effectively. With the network in hand to predict how significant a simulation will be, we can create an automated tool to suggest the next simulation that should be performed. This helps us optimally use the computational resources and time available to us. With the high computational cost of NR simulations, this will be particularly important as

we prepare dense catalogs for future GW observations. Section III A introduces such a tool.

Similarly, by predicting the mismatch between points covered by existing catalogs and points not yet populated, we can note gaps in the parameter space that haven't been filled to a certain threshold in mismatch. There are regions of parameter space known to be insufficiently filled, but there are other regions whose statuses are less clear. Identifying gaps in coverage will be important for assessing and improving the accuracy of models. Section III B shows such gaps in the current public waveform catalogs.

Both of the above scenarios have involved using the network to predict the mismatch between unpopulated and populated points in parameter space. We can also use the network simply to explore the mismatch behavior across the parameter space. By doing this, we can note degeneracies, points in parameter space that are significantly separated in terms of spins or mass ratio and yet have low mismatch with one another. Section III C utilizes the network to explore such degeneracies.

A. Template Placement

There are several approaches one could take to identify the next simulation to perform depending on one's needs. In this section, we introduce a tool to use this new network to propose parameters for a simulation that would maximize the minimum mismatch between the new sim-

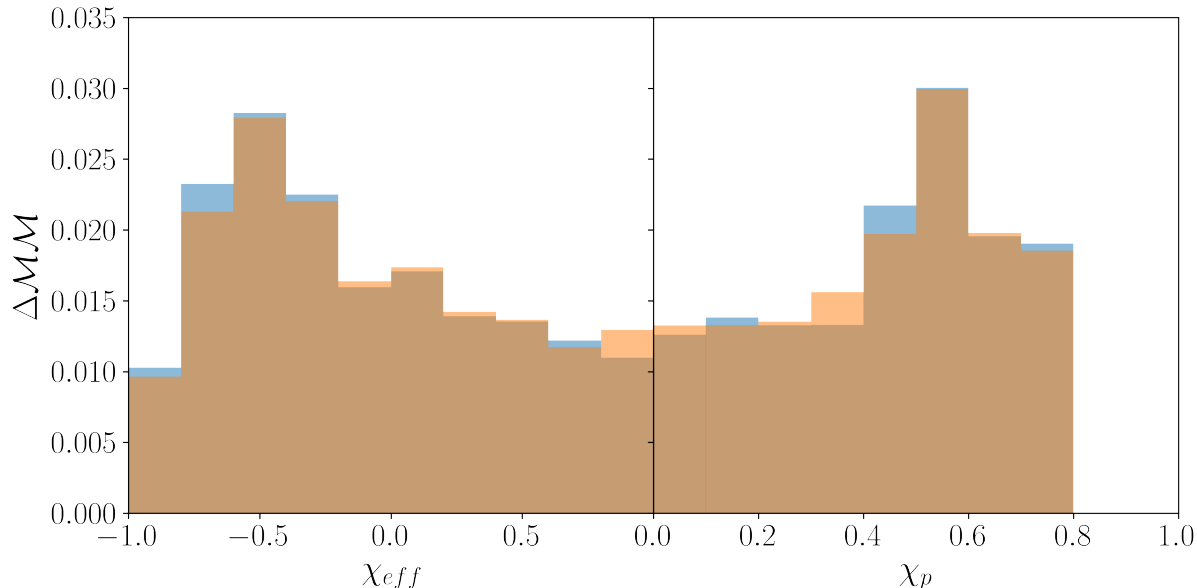


FIG. 7: The mean absolute error in the mismatch for the test set binned by χ_{eff} (left) and χ_p (right). The blue represents the first system and the orange represents the second system.

ulation and existing simulations. In other words, it maximizes the distance between the new simulation and its nearest neighbor (with distance characterized by the mismatch). This will efficiently cover the edges of the desired parameter space and populate gaps within it.

The current parameter space coverage does not lead to a very smooth function to maximize; we have a seven dimensional parameter space with many local maxima, so the suggestion for the next simulation depends greatly upon the initial guess fed to the maximizing function. To find a global maximum in the presence of local maxima, we utilize the scipy implementation of basin hopping [86]. Developed to minimize the energy of a cluster of atoms, basin hopping is a method of identifying the global extremum when many local extrema may be present. The method cycles through two steps, a local minimization, and then a perturbation. Given a starting point, it performs a local minimization, finding the nearest local extremum. It then perturbs the coordinates and performs another local minimization beginning at the new coordinates. Depending on the height of the newly found extremum, the new coordinates are either accepted or rejected. Another perturbation is then performed from either the new or old coordinates. This cycle continues for a given number of iterations.

Without doing a full brute force analysis, we can't be certain if a global maximum has been found. Instead, we can apply the same process from multiple different starting points and test whether they converge to the same maximum. We find that while they don't approach the exact same parameters, they do all approach

maxima of nearly the same height with mismatches of $\approx 0.623 \pm 0.016$. The difference in the peaks is within the uncertainty of the network and implies the presence of multiple, very comparable local maxima, any of which could be deemed the global maximum.

Now that we're confident that the basin hopping optimization method is finding the parameters that lead to a maximal mismatch, we can apply it multiple times to suggest future simulations. In this paper, we use a greedy approach, proposing simulations one at a time. After a simulation is suggested, we add it as an existing data point and then perform another optimization to search for the next best point. Fig. 9 shows the results of doing this 20 times, restricting $\eta \geq 0.0826$.

The parameters output by the optimization algorithm are generally in regions where we would anticipate requiring more points. All the suggested points have a symmetric mass ratio of $\eta \lesssim 0.16$ and many have higher x and y spin components to fill in gaps in the precessing space. The mismatch between the first suggested simulation and its nearest neighbor was ~ 0.6 while the mismatch between the 20th proposed simulation and its nearest neighbor was ~ 0.3 . With only 20 simulations, we have reduced the largest mismatch resulting from a gap in parameter space by a factor of 2.

While the parameters are in areas that aren't entirely surprising, they improve the parameter space coverage much more efficiently than maximizing the separation of the parameters would. To assess this, we used the same optimization technique to optimize for the maximum parameter separation rather than mismatch. As a metric

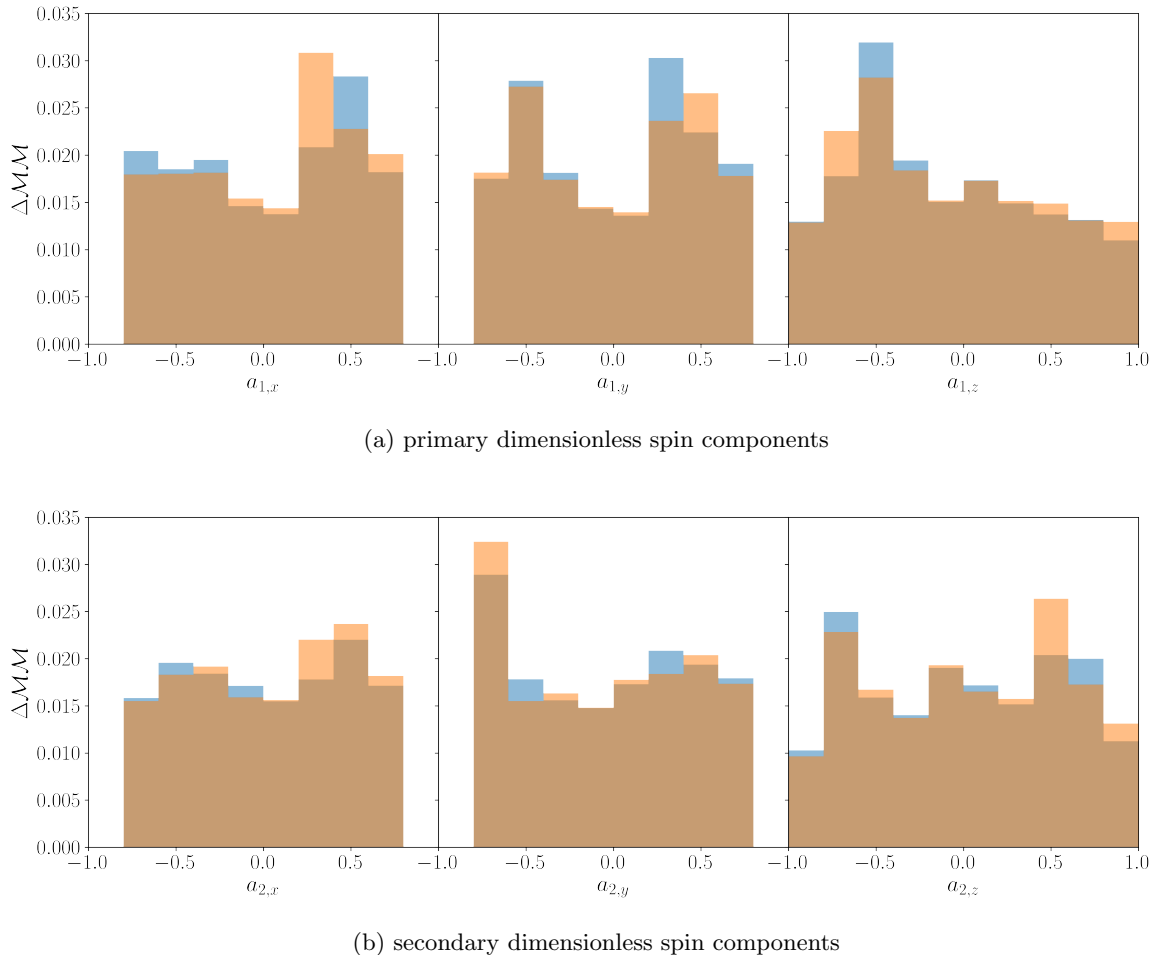


FIG. 8: The mean absolute error in the mismatch for the test set binned by the components of the spin. The blue represents the first system and the orange represents the second system.

for the separation in the parameter space, we compute the magnitude of the separation along each of the seven axes normalized by the considered range. It does begin to fill in the parameter space, but after 20 simulations, there are still regions of parameter space with a mismatch of 0.6 with their nearest neighbor.

B. Identifying Gaps

There are several public NR catalogs including those from SXS [61], RIT [60, 62], and Maya [59]. Combining these three catalogs, there are some regions of parameter space that appear to have dense coverage and others that appear rather sparse. This section reports on a brute force analysis to identify gaps within the coverage of these three public catalogs. We randomly sample 100000 points across the seven dimensional parameter space and, for each point, find their lowest mismatch with existing simulations. Those points whose mismatch

with their nearest neighbor is greater than 0.1 are shown in Fig. 10. The uncertainty of 0.018 in the mismatch prediction contributes some uncertainty to these points, but it is clear that any point included in this figure still represents a significant gap.

We only show simulations with $\eta \geq 0.16$ in this figure since it is well known that the parameter space below $\eta = 0.16$ is sparsely populated. By including only those points with $\eta \geq 0.16$, we can identify gaps in what appears to be a well populated region of parameter space. Of the 53,019 points in this region, 57 have a mismatch with their nearest neighbor greater than 0.1. The NR catalogs have dense coverage for $\eta \geq 0.19$, but in the region $0.16 \leq \eta \leq 0.19$, more simulations are needed. In particular, in terms of the primary BH, more simulations are needed with anti-aligned or precessing spins. In terms of the secondary BH, all spin configurations are still needed.

To understand the impact of these gaps on the training of analytic models, we assess the performance of IMRPhenomPv3HM [87] at two points in the precessing

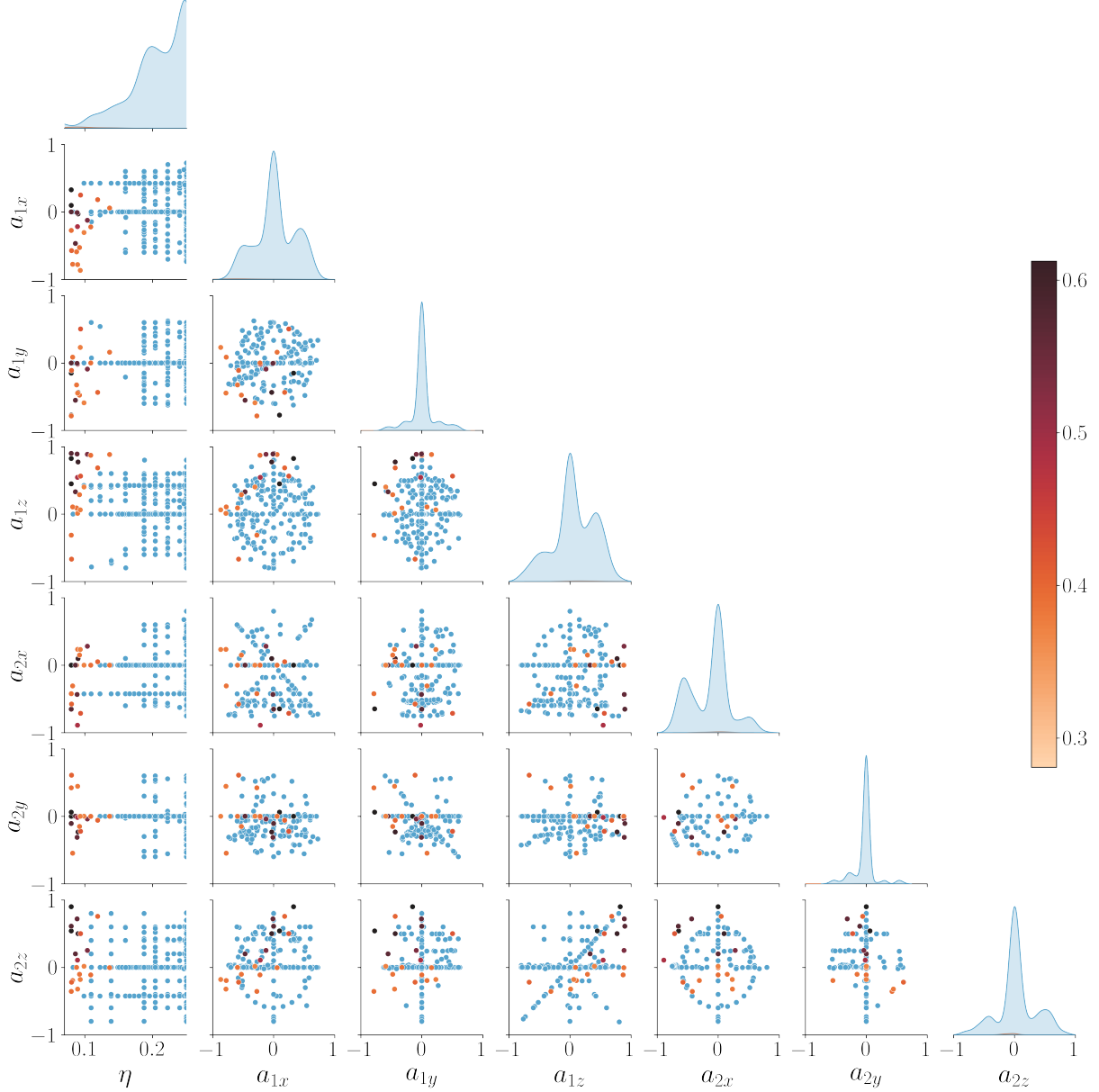


FIG. 9: Suggested simulations, selected so as to maximize their mismatch with existing simulations. Blue points are existing simulations from the MAYA catalog, and red points are suggested new simulations. The shade denotes a point's mismatch with its nearest neighbor at the time it is suggested, with darker being a higher mismatch.

space, one in a region densely populated by NR points (mismatch of 0.997 with the nearest catalog point) and one in a gap (mismatch of 0.839 with nearest catalog point). We find that the mismatch between the model waveform and the NR waveform is about an order of magnitude higher for the system placed within a gap.

C. Identifying Degeneracies

It is clear that changing the parameters of the initial BHs affects the morphology of the gravitational waves their coalescence emits. Many of the patterns of these effects are well known, and in some cases, changing different parameters can lead to similar effects on the waveform. It may even be possible that two binary systems with very different initial parameters may lead to extremely similar gravitational waves. In such a situation, analysis pipelines may not be able to distinguish between

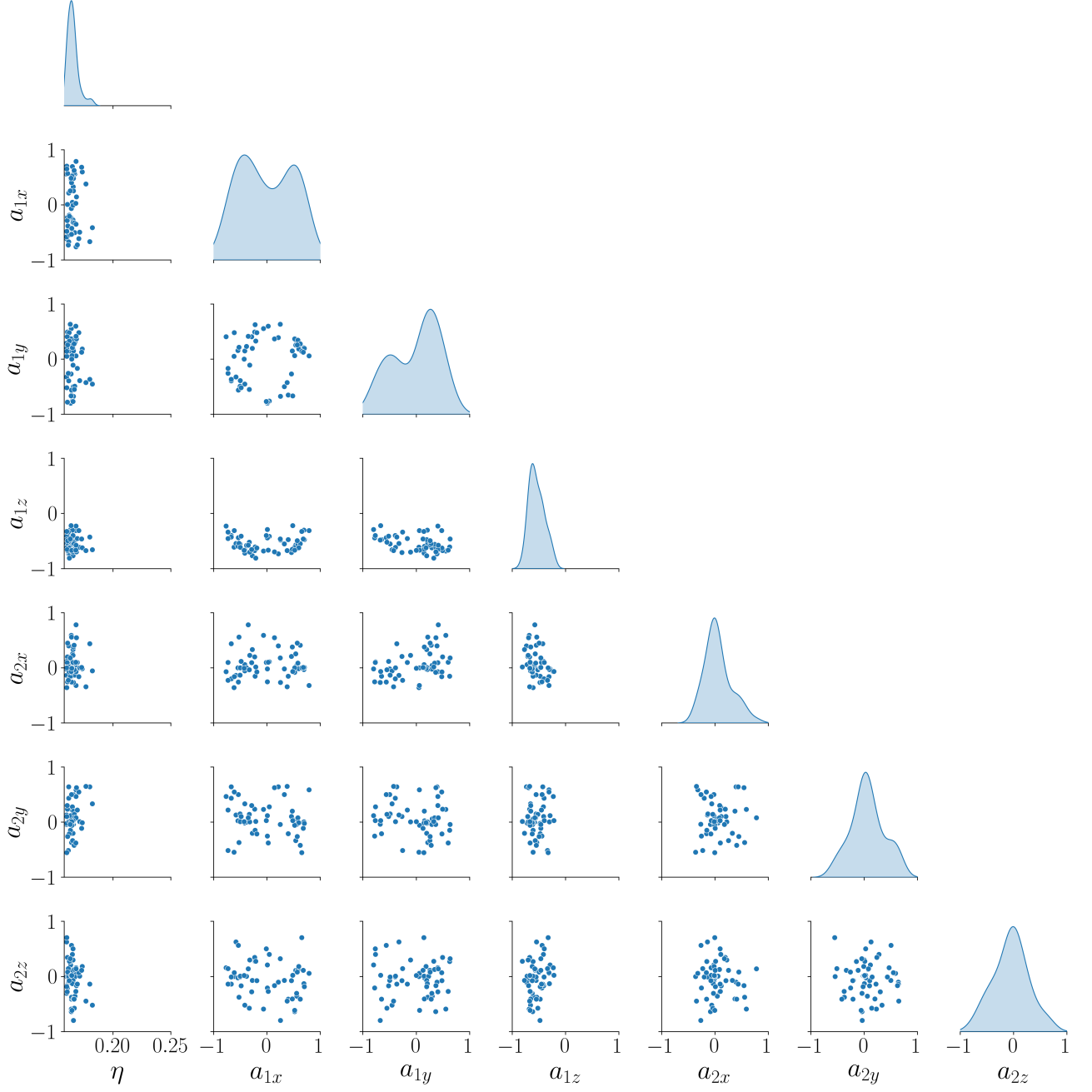


FIG. 10: Randomly sampled points whose mismatch with their nearest neighbor in existing public NR waveform catalogs is greater than 0.1.

the potential sources. This can complicate the population studies performed on GW catalogs [88–90]. In some cases, degeneracies may only appear in certain modes and can be broken when analyzing other modes. If we identify where such degeneracies occur, we can study the systems further to identify any degeneracy breaking properties.

The network presented in this paper can be used to identify such degenerate regions of parameter space. Since a seven dimensional parameter space is challenging

to show visually, we consider the aligned space here as an example, meaning we are only considering systems with spin vectors aligned with the orbital angular momentum of the binary ($a_{1x} = a_{1y} = a_{2x} = a_{2y} = 0$). This also equates to systems where $\chi_p = 0$. Figure 11 shows the mismatch between a reference binary and other binaries throughout the parameter space, plotted as χ_{eff} vs η . This shows the distribution of the predicted mismatch across the aligned spin parameter space when one of the

binaries is held fixed. Each panel uses a different reference simulation, denoted by the black dot.

Both low η and high χ_{eff} cause the binary system to take longer to inspiral when compared to equal mass, nonspinning binaries; this similarity can be seen in Figure 11 where low η , low spinning simulations have a very low mismatch with high η , highly spinning simulations. We can see that even binaries with symmetric mass ratios as low as $\eta \sim 0.08$ can be confused with a highly spinning, equal mass binary. This could have significant implications for population studies, as highly spinning events and highly unequal mass events could be confused. Since this network is trained only on the $\ell = 2, m = 2$ mode, it is possible that this degeneracy will be broken in higher modes; this motivates further study of these systems.

IV. CONCLUSIONS

NR simulations form an important basis for the detection and characterization of GW signals. Whether they are used directly with data analysis or to train analytic models, it is crucial to have a sufficiently dense template bank of NR simulations. Given the high computational cost of NR simulations, it's important to select initial parameters with care. In this paper, we have introduced a way to place new NR simulations using a neural network trained to predict the mismatch between the waveforms of two binary systems.

The network has an average uncertainty of 0.018 in the predicted mismatch. Due to the high density of training points at high η , the error is less at high η and increases as η decreases, particularly for precessing binaries. Even for the most challenging regions of parameter space, the uncertainty stays below ~ 0.03 .

With this network, we were able to optimally select initial parameters so as to maximize the mismatch of a new simulation with its nearest neighbor. With only 20

simulations, we decreased the maximal mismatch in unpopulated regions by a factor of 2. As we prepare for the next generation of GW detectors, it becomes crucial to have a densely populated template bank of highly accurate waveforms. This method will aid us in creating such a template bank, ensuring every simulation performed is chosen to maximize our coverage.

We also used the mismatch predicting network to identify gaps in the coverage of current public NR waveform catalogs. It is clear that more simulations are needed at $\eta < 0.16$, but our analysis also shows that in the region of $0.16 \leq \eta \leq 0.19$, there are significant coverage gaps. This is particularly true for binaries with anti-aligned or precessing spins.

The network also provided us a method to explore the parameter space, seeing how the mismatch changes with each parameter and identifying degeneracies. We demonstrated how highly unequal mass, low spinning simulations can be confused with comparable mass, highly spinning simulations, encouraging more study into how these degeneracies may be broken.

As we prepare for the next generation of GW detectors, we need to have a strong understanding of BBH systems within GR. The network presented in this paper to predict the mismatch between binary systems provides a unique and powerful approach to studying the BBH parameter space. The template placement tool utilizes this network to optimally propose simulations that will efficiently expand our NR catalogs, greatly improving our readiness for the coming improvements to GW detectors.

Acknowledgements This work is supported by grants NASA 80NSSC21K0900 and NSF PHY 2207780. Computing resources were provided by XSEDE PHY120016 and TACC PHY20039. The author thanks Deirdre Shoemaker for support and helpful discussions. The author also thanks Harald Pfeiffer for useful feedback. This work was done as part of the Weinberg Institute, with the identifier UTWI-20-2022.

-
- [1] LIGO Scientific Collaboration and Virgo Collaboration, Phys. Rev. Lett. **116**, 061102 (2016), arXiv:1602.03837 [gr-qc].
 - [2] LIGO Scientific Collaboration and Virgo Collaboration, Phys. Rev. Lett. **119**, 141101 (2017).
 - [3] LIGO Scientific Collaboration and Virgo Collaboration, The Astrophysical Journal **851**, L35 (2017).
 - [4] LIGO Scientific Collaboration and Virgo Collaboration (LIGO Scientific and Virgo Collaboration), Phys. Rev. Lett. **118**, 221101 (2017).
 - [5] LIGO Scientific Collaboration and Virgo Collaboration (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. **116**, 241103 (2016).
 - [6] V. LIGO Scientific Collaboration and KAGRA Collaboration (LIGO Scientific, VIRGO, KAGRA), (2021), arXiv:2111.03606 [gr-qc].
 - [7] T. Robson, N. J. Cornish, and C. Liu, Classical and Quantum Gravity **36**, 105011 (2019).
 - [8] M. Punturo et al., Class. Quant. Grav. **27**, 194002 (2010).
 - [9] M. Abernathy et al., EGO (2011).
 - [10] M. Evans et al., (2021), arXiv:2109.09882 [astro-ph.IM].
 - [11] J. Baker, J. Bellovary, P. L. Bender, E. Berti, R. Caldwell, J. Camp, J. W. Conklin, N. Cornish, C. Cutler, R. DeRosa, M. Eracleous, E. C. Ferrara, S. Francis, M. Hewitson, K. Holley-Bockelmann, A. Hornschemeier, C. Hogan, B. Kamai, B. J. Kelly, J. Shapiro Key, S. L. Larson, J. Livas, S. Manthripragada, K. McKenzie, S. T. McWilliams, G. Mueller, P. Natarajan, K. Numata, N. Rioux, S. R. Sankar, J. Schnittman, D. Shoemaker, D. Shoemaker, J. Slutsky, R. Spero, R. Stebbins, I. Thorpe, M. Vallisneri, B. Ware, P. Wass, A. Yu, and J. Ziemer, arXiv e-prints, arXiv:1907.06482 (2019), arXiv:1907.06482 [astro-ph.IM].
 - [12] P. Amaro-Seoane, H. Audley, S. Babak, J. Baker, E. Barausse, P. Bender, E. Berti, P. Binetruy, M. Born, D. Borzulli, J. Camp, C. Caprini, V. Cardoso, M. Colpi,

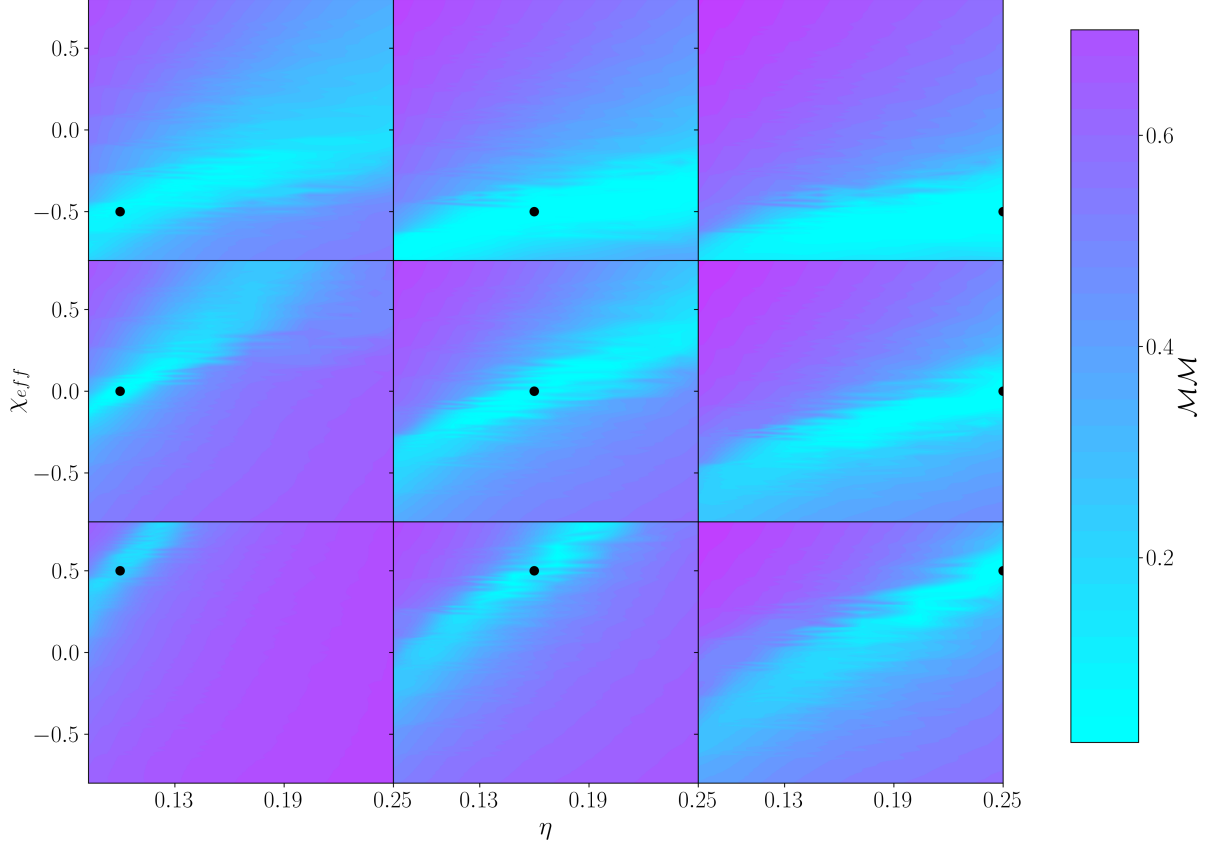


FIG. 11: Mismatch between a reference simulation and other simulations throughout the parameter space. All simulations considered here have BH spins aligned with the orbital angular momentum. The black dot in each panel denotes the reference simulation. From left to right, the reference simulations have symmetric mass ratios of $\eta = 0.1, 0.16, 0.25$, and from top to bottom, the reference simulations have $\chi_{eff} = -0.5, 0, 0.5$.

- J. Conklin, N. Cornish, C. Cutler, K. Danzmann, R. Dolesi, L. Ferraioli, V. Ferroni, E. Fitzsimons, J. Gair, L. Gesa Bote, D. Giardini, F. Gibert, C. Grimaldi, H. Halluin, G. Heinzl, T. Hertog, M. Hewitson, K. Holley-Bockelmann, D. Hollington, M. Hueller, H. Inchauspe, P. Jetzer, N. Karnesis, C. Killow, A. Klein, B. Klipstein, N. Korsakova, S. L. Larson, J. Livas, I. Lloro, N. Man, D. Mance, J. Martino, I. Mateos, K. McKenzie, S. T. McWilliams, C. Miller, G. Mueller, G. Nardini, G. Nelemans, M. Nofrarias, A. Petiteau, P. Pivato, E. Plagnol, E. Porter, J. Reiche, D. Robertson, N. Robertson, E. Rossi, G. Russano, B. Schutz, A. Sesana, D. Shoemaker, J. Slutsky, C. F. Sopuerta, T. Sumner, N. Tamanini, I. Thorpe, M. Troebels, M. Vallisneri, A. Vecchio, D. Vetrugno, S. Vitale, M. Volonteri, G. Wanner, H. Ward, P. Wass, W. Weber, J. Ziemer, and P. Zweifel, arXiv e-prints, arXiv:1702.00786 (2017), arXiv:1702.00786 [astro-ph.IM].
- [13] B. Abbott et al. (LIGO Scientific, Virgo), Phys. Rev. Lett. **119**, 161101 (2017), arXiv:1710.05832 [gr-qc].
- [14] B. Abbott et al. (LIGO Scientific, Virgo), Astrophys. J. Lett. **892**, L3 (2020), arXiv:2001.01761 [astro-ph.HE].
- [15] M. Shibata, S. Fujibayashi, K. Hotokezaka, K. Kiuchi, K. Kyutoku, Y. Sekiguchi, and M. Tanaka, Phys. Rev. D **96**, 123012 (2017), arXiv:1710.07579 [astro-ph.HE].
- [16] S. Isoyama, R. Sturani, and H. Nakano, (2020), 10.1007/978-981-15-4702-7_31-1, arXiv:2012.01350 [gr-qc].
- [17] L. Blanchet, Living Rev. Rel. **17**, 2 (2014), arXiv:1310.1528 [gr-qc].
- [18] G. Schäfer and P. Jaranowski, Living Rev. Rel. **21**, 7 (2018), arXiv:1805.07240 [gr-qc].
- [19] T. Futamase and Y. Itoh, Living Rev. Rel. **10**, 2 (2007).
- [20] L. Blanchet, G. Faye, B. R. Iyer, and S. Sinha, Class. Quant. Grav. **25**, 165003 (2008), [Erratum: Class.Quant.Grav. 29, 239501 (2012)], arXiv:0802.1249 [gr-qc].
- [21] L. Blanchet, G. Faye, B. R. Iyer, and B. Joguet, Phys. Rev. D **65**, 061501 (2002), [Erratum: Phys.Rev.D 71, 129902 (2005)], arXiv:gr-qc/0105099.

- [22] L. Blanchet, T. Damour, G. Esposito-Farese, and B. R. Iyer, *Phys. Rev. D* **71**, 124004 (2005), arXiv:gr-qc/0503044.
- [23] G. Faye, S. Marsat, L. Blanchet, and B. R. Iyer, *Class. Quant. Grav.* **29**, 175004 (2012), arXiv:1204.1043 [gr-qc].
- [24] G. Faye, L. Blanchet, and B. R. Iyer, *Class. Quant. Grav.* **32**, 045016 (2015), arXiv:1409.3546 [gr-qc].
- [25] L. M. Burko and G. Khanna, *Phys. Rev. D* **91**, 104017 (2015), arXiv:1503.05097 [gr-qc].
- [26] K. A. Lackeos and L. M. Burko, *Phys. Rev. D* **86**, 084055 (2012), arXiv:1206.1452 [gr-qc].
- [27] T. Hinderer and E. E. Flanagan, *Phys. Rev. D* **78**, 064028 (2008), arXiv:0805.3337 [gr-qc].
- [28] L. Barack and A. Pound, *Rept. Prog. Phys.* **82**, 016904 (2019), arXiv:1805.10385 [gr-qc].
- [29] E. Poisson, A. Pound, and I. Vega, *Living Rev. Rel.* **14**, 7 (2011), arXiv:1102.0529 [gr-qc].
- [30] A. I. Harte, *Fund. Theor. Phys.* **179**, 327 (2015), arXiv:1405.5077 [gr-qc].
- [31] A. Pound, *Fund. Theor. Phys.* **179**, 399 (2015), arXiv:1506.06245 [gr-qc].
- [32] A. Pound and B. Wardell, (2021), 10.1007/978-981-15-4702-7_38-1, arXiv:2101.04592 [gr-qc].
- [33] M. van de Meent and H. P. Pfeiffer, *Phys. Rev. Lett.* **125**, 181101 (2020), arXiv:2006.12036 [gr-qc].
- [34] U. Sperhake, V. Cardoso, C. D. Ott, E. Schnetter, and H. Witek, *Phys. Rev. D* **84**, 084038 (2011), arXiv:1105.5391 [gr-qc].
- [35] A. Le Tiec et al., *Phys. Rev. D* **88**, 124027 (2013), arXiv:1309.0541 [gr-qc].
- [36] A. Le Tiec, E. Barausse, and A. Buonanno, *Phys. Rev. Lett.* **108**, 131103 (2012), arXiv:1111.5609 [gr-qc].
- [37] M. van de Meent, *Phys. Rev. Lett.* **118**, 011101 (2017), arXiv:1610.03497 [gr-qc].
- [38] A. Zimmerman, A. G. M. Lewis, and H. P. Pfeiffer, *Phys. Rev. Lett.* **117**, 191101 (2016), arXiv:1606.08056 [gr-qc].
- [39] F. Löffler et al., *Class. Quant. Grav.* **29**, 115001 (2012), arXiv:1111.3344 [gr-qc].
- [40] L. Pekowsky, R. O’Shaughnessy, J. Healy, and D. Shoemaker, *Phys. Rev. D* **88**, 024040 (2013).
- [41] T. W. Baumgarte and S. L. Shapiro, *Phys. Rev. D* **59**, 024007 (1998), arXiv:gr-qc/9810065.
- [42] E. Schnetter, S. H. Hawley, and I. Hawke, *Class. Quant. Grav.* **21**, 1465 (2004), arXiv:gr-qc/0310042 [gr-qc].
- [43] C. Capano, I. Harry, S. Privitera, and A. Buonanno, *Phys. Rev. D* **93**, 124007 (2016), arXiv:1602.03509 [gr-qc].
- [44] A. Nitz, I. Harry, D. Brown, C. M. Biwer, J. Willis, T. D. Canton, C. Capano, L. Pekowsky, T. Dent, A. R. Williamson, S. De, G. Davies, M. Cabero, D. Macleod, B. Machenschalk, S. Reyes, P. Kumar, T. Massinger, F. Pannarale, dfinstad, M. Tápai, S. Fairhurst, S. Khan, L. Singer, S. Kumar, A. Nielsen, shasvath, idorrington92, A. Lenon, and H. Gabbard, “gwastro/pycbc: Pycbc release v1.15.4,” (2020).
- [45] T. Dal Canton, A. H. Nitz, A. P. Lundgren, A. B. Nielsen, D. A. Brown, T. Dent, I. W. Harry, B. Krishnan, A. J. Miller, K. Wette, K. Wiesner, and J. L. Willis, *Phys. Rev. D* **90**, 082004 (2014).
- [46] S. A. Usman et al., *Class. Quant. Grav.* **33**, 215004 (2016), arXiv:1508.02357 [gr-qc].
- [47] K. Cannon et al., *Astrophys. J.* **748**, 136 (2012), arXiv:1107.2665 [astro-ph.IM].
- [48] S. Privitera, S. R. Mohapatra, P. Ajith, K. Cannon, N. Fotopoulos, M. A. Frei, C. Hanna, A. J. Weinstein, and J. T. Whelan, *Phys. Rev. D* **89**, 024003 (2014), arXiv:1310.5633 [gr-qc].
- [49] C. Messick, K. Blackburn, P. Brady, P. Brockill, K. Cannon, R. Cariou, S. Caudill, S. J. Chamberlin, J. D. Creighton, R. Everett, and et al., *Physical Review D* **95** (2017), 10.1103/physrevd.95.042001.
- [50] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer, *Phys. Rev. Lett.* **113**, 151101 (2014), arXiv:1308.3271 [gr-qc].
- [51] A. Bohé et al., *Phys. Rev. D* **95**, 044028 (2017), arXiv:1611.03703 [gr-qc].
- [52] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. Jiménez Forteza, and A. Bohé, *Phys. Rev. D* **93**, 044007 (2016), arXiv:1508.07253 [gr-qc].
- [53] J. Blackman, S. E. Field, M. A. Scheel, C. R. Galley, C. D. Ott, M. Boyle, L. E. Kidder, H. P. Pfeiffer, and B. Szilágyi, *Phys. Rev. D* **96**, 024058 (2017), arXiv:1705.07089 [gr-qc].
- [54] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. Jiménez Forteza, and A. Bohé, *Phys. Rev. D* **93**, 044006 (2016), arXiv:1508.07250 [gr-qc].
- [55] A. Taracchini et al., *Phys. Rev. D* **89**, 061502 (2014), arXiv:1311.2544 [gr-qc].
- [56] LIGO Scientific Collaboration and Virgo Collaboration (LIGO Scientific, Virgo), *Phys. Rev. D* **93**, 122004 (2016), [Addendum: *Phys. Rev. D* **94**, 069903 (2016)], arXiv:1602.03843 [gr-qc].
- [57] J. Lange et al., *Phys. Rev. D* **96**, 104041 (2017), arXiv:1705.09833 [gr-qc].
- [58] A. H. Mroue et al., *Phys. Rev. Lett.* **111**, 241104 (2013), arXiv:1304.6077 [gr-qc].
- [59] K. Jani, J. Healy, J. A. Clark, L. London, P. Laguna, and D. Shoemaker, *Class. Quant. Grav.* **33**, 204001 (2016), arXiv:1605.03204 [gr-qc].
- [60] J. Healy, C. O. Lousto, Y. Zlochower, and M. Campanelli, *Class. Quant. Grav.* **34**, 224001 (2017), arXiv:1703.03423 [gr-qc].
- [61] M. Boyle et al., *Class. Quant. Grav.* **36**, 195006 (2019), arXiv:1904.04831 [gr-qc].
- [62] J. Healy, C. O. Lousto, J. Lange, R. O’Shaughnessy, Y. Zlochower, and M. Campanelli, *Phys. Rev. D* **100**, 024021 (2019), arXiv:1901.02553 [gr-qc].
- [63] J. Healy and C. O. Lousto, (2020), arXiv:2007.07910 [gr-qc].
- [64] D. Ferguson, K. Jani, P. Laguna, and D. Shoemaker, *Phys. Rev. D* **104**, 044037 (2021), arXiv:2006.04272 [gr-qc].
- [65] S. E. Field, C. R. Galley, J. S. Hesthaven, J. Kaye, and M. Tiglio, *Phys. Rev. X* **4**, 031006 (2014), arXiv:1308.3565 [gr-qc].
- [66] Z. Doctor, B. Farr, D. E. Holz, and M. Pürrer, *Phys. Rev. D* **96**, 123011 (2017), arXiv:1706.05408 [astro-ph.HE].
- [67] W. S. McCulloch and W. Pitts, *The bulletin of mathematical biophysics* **5**, 115 (1943).
- [68] J. J. Hopfield, *Proceedings of the National Academy of Sciences* **79**, 2554 (1982).
- [69] F. Rosenblatt, *Psychological review* **65** **6**, 386 (1958).
- [70] P. Werbos and P. John, (1974).
- [71] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, *Neural Computation* **1**, 541 (1989).

- [72] L. Haegel and S. Husa, *Class. Quant. Grav.* **37**, 135005 (2020), arXiv:1911.01496 [gr-qc].
- [73] Z. Que, E. Wang, U. Marikar, E. Moreno, J. Ngadiuba, H. Javed, B. Borzyszkowski, T. Aarrestad, V. Loncar, S. Summers, M. Pierini, P. Y Cheung, and W. Luk, arXiv e-prints , arXiv:2106.14089 (2021), arXiv:2106.14089 [cs.LG].
- [74] T. D. Gebhard, N. Kilbertus, I. Harry, and B. Schölkopf, *Phys. Rev. D* **100**, 063015 (2019), arXiv:1904.08693 [astro-ph.IM].
- [75] B. J. Owen, *Phys. Rev. D* **53**, 6749 (1996).
- [76] S. D. Mohanty, *Phys. Rev. D* **57**, 630 (1998).
- [77] LIGO Scientific Collaboration, “LIGO Algorithm Library - LALSuite,” free software (GPL) (2018).
- [78] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, “TensorFlow: Large-scale machine learning on heterogeneous systems,” (2015), software available from tensorflow.org.
- [79] F. Chollet et al., “Keras,” <https://keras.io> (2015).
- [80] D. P. Kingma and J. Ba, arXiv e-prints , arXiv:1412.6980 (2014), arXiv:1412.6980 [cs.LG].
- [81] A. F. Agarap, arXiv e-prints , arXiv:1803.08375 (2018), arXiv:1803.08375 [cs.NE].
- [82] T. Damour, *Phys. Rev. D* **64**, 124013 (2001).
- [83] E. Racine, *Phys. Rev. D* **78**, 044021 (2008), arXiv:0803.1820 [gr-qc].
- [84] P. Schmidt, F. Ohme, and M. Hannam, *Phys. Rev. D* **91**, 024043 (2015), arXiv:1408.1810 [gr-qc].
- [85] P. Schmidt, M. Hannam, and S. Husa, *Phys. Rev. D* **86**, 104063 (2012), arXiv:1207.3088 [gr-qc].
- [86] D. J. Wales and J. P. K. Doye, *The Journal of Physical Chemistry A* **101**, 5111 (1997), <https://doi.org/10.1021/jp970984n>.
- [87] S. Khan, F. Ohme, K. Chatziioannou, and M. Hannam, *Phys. Rev. D* **101**, 024056 (2020), arXiv:1911.06050 [gr-qc].
- [88] B. P. Abbott et al. (LIGO Scientific, Virgo), *Astrophys. J. Lett.* **818**, L22 (2016), arXiv:1602.03846 [astro-ph.HE].
- [89] R. Abbott et al. (LIGO Scientific, Virgo), *Astrophys. J. Lett.* **913**, L7 (2021), arXiv:2010.14533 [astro-ph.HE].
- [90] M. Zevin, S. S. Bavera, C. P. L. Berry, V. Kalogera, T. Fragos, P. Marchant, C. L. Rodriguez, F. Antonini, D. E. Holz, and C. Pankow, *Astrophys. J.* **910**, 152 (2021), arXiv:2011.10057 [astro-ph.HE].