Model Selection in Gravitational Wave Astronomy

Kate Napier, Dr. James Clark, Dr. Deirdre Shoemaker

Center for Relativistic Astrophysics

Abstract

The several detections of gravitational waves by the Laser Interferometer Gravitational-wave Observatory (LIGO) and the Virgo detector are providing insights about the nature of gravity in our Universe [1]. As the number of detected gravitational waves increases, it will be necessary to employ computationally inexpensive methods to extract the parameters of the gravitational wave sources. This proof of concept study utilizes principal component analysis (PCA) to try to reduce the number of vectors needed to describe binary black hole (BBH) parameter space. The results of this study suggest that performing PCA on face-on BBH systems, those at zero degrees inclination, adds an unnecessary level of complexity. However, PCA is a beneficial method to apply to waveforms that are morphologically complex such as those from edge-on BBH systems, those at 90 degrees inclination. Running PCA on a catalog of specific waveforms in one area of parameter space can inform how to construct waveforms in other regions of parameter space. If these waveforms can be constructed to a high level of accuracy using only a few principal components (PCs), it will significantly reduce the computational cost associated with generating template waveforms. Instead of creating a waveform template for every parameter combination, the PCs can be used to construct waveforms from similar areas of parameter space.

Introduction/Background Information

The detection of six gravitational waves from binary black hole and neutron star systems has brought about the advent of a whole new way of studying the Universe [1]. Now, in addition to using light to study astrophysical phenomena, we can use gravity. When two compact accelerating objects, such as black holes or neutron stars, collide, inspiral, and merge, they produce distortions in the fabric of spacetime, similar to how a stone dropped in water creates ripples. Because gravitational waves interact weakly with matter, they are extremely difficult to detect; nonetheless, they carry information with them about their cataclysmic origins.

LIGO, a National Science Foundation funded project, is the most sensitive scientific instrument in the world and is capable of detecting gravitational waves. LIGO consists of two Michelson-like interferometers, one located in Livingston, Louisiana and the other in Hanford, Washington. At both locations, the interferometers have 4-kilometer long perpendicular arms. A laser beam is split at the intersection of the interferometer’s two arms. The split beam travels down the length of the two arms and then recombines. Because gravitational waves stretch spacetime in one direction and shrink it in another, if a gravitational wave hits the detector, it will alter the path lengths of the arms. This causes the light to recombine constructively and an interference pattern is observed at the photodetector. The interference pattern tells us more about the origin of the gravitational waves and the properties of the system that produced them [1].

Scientists in the LIGO Scientific Collaboration have employed various models to study gravitational waves. When a gravitational wave is detected, we are able to compare them to our models for which we know the parameters, and thus we can determine the source parameters. Different models have various strengths and weaknesses, and improving them will broaden our understanding of the systems that create gravitational waves. The Georgia Tech LIGO group utilizes waveforms that are created by solving Albert Einstein’s equations of General Relativity (GR) numerically in order to study the coalescence of binary black hole systems [2]. With numerical relativity (NR), the equations governing the last few orbits and merging of two black holes can be solved, and thus, remaining mysteries about gravity can be unraveled [3-5].

Creating NR waveforms is computationally expensive [6]. Therefore, this proof of concept study seeks to reduce the computational cost and still capture the most important characteristics of the waveform data set. This study employs a tool derived from linear algebra called PCA to predict where in parameter space we should add more NR simulations. PCA studies the variation in a data set, with the first PC accounting for the maximum variation. The subsequent PCs account for as much variation as possible under the constraint that they are orthogonal to the preceding ones in parameter space. After PCA is performed, waveforms can be constructed by projecting them onto the new PCA basis. Each PC gets assigned a different weight, called a beta value, unique to the waveform that is being constructed. Therefore, a waveform can be constructed as a linear combination of the PCs. It is advantageous to use PCA because it reduces the dimensionality of the data set. For instance, it is common that only a few PCs can construct a waveform to a high level of accuracy. If the PCs of a set of waveforms cannot represent the main features of a gravitational wave nearby in parameter space, it will indicate that a new NR simulation is needed there. Overall, PCA is convenient, mathematically simple, and computationally cheap, and the software for it exists.

The future of gravitational wave astronomy will depend heavily on models that are able to capture the physics of some of the most energetic processes in the Universe. Since the number of detected gravitational waves is increasing rapidly, it is important to employ computationally inexpensive methods in order to perform parameter estimation efficiently and accurately so that the results can be shared with the public in a timely manner. PCA may prove to be the desired methodology.

Methods/Procedures/Materials

This research utilizes the bhextractor repository. The bhextractor repository was chosen because it is the only resource that employs the desired methodology. The repository was downloaded from github.com, edited, and run in python.

A BBH system's masses, spins, location, and orientation all affect how gravitational-wave signals appear once they reach Earth. In this study, the two parameters that are varied include the mass ratio of the two black holes, q, and the inclination, *i*, of the binary system. The code allows the user to input the minimum and maximum mass ratios of the waveforms as well as the step size. The inclination is also set prior to running the code. The code creates the waveforms based on the Effective One Body (EOB) formalism, an analytical approach that strives to accurately describe the motion and radiation of coalescing binary black holes. The EOB formalism consists of three main parts: a description of the conservative Hamiltonian part of the dynamics of binary black holes, an expression for the radiation-reaction part, and a description of the emitted gravitational waveform from a coalescing binary system [7]. Since EOB waveforms are solved analytically instead of numerically, it takes significantly less time to generate them. Thus, for the scope of this project, they are the most useful. Previous work has shown a high level of agreement between NR and EOB waveforms, so this work can eventually be extended to NR waveforms [8].

After the waveforms are generated, their amplitudes are extracted. For a wave propagating in the z-direction, the amplitude $A^{\mu\nu}$ is written as

$$A^{\mu\nu} = h_+\epsilon_+^{\mu\nu} + h_x\epsilon_x^{\mu\nu},$$

where $\epsilon_+^{\mu\nu}$ and $\epsilon_x^{\mu\nu}$ are the unit polarization tensors given by

$$\epsilon_{+}^{\mu\nu} \equiv \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\epsilon_{x}^{\mu\nu} \equiv \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

All of the amplitudes are truncated to be the length of the shortest amplitude in the catalog. Then the amplitudes are peak aligned. This was so that PCA, the method described in more mathematical detail below, did not have to account for trivial variance in the catalog. To smooth the amplitudes and minimize spectral ringing, a tukey window was applied to all of the amplitudes. A tukey window, also known as a tapered cosine window, is essentially a cosine lobe of width $\alpha N/2$ that is convolved with a rectangular window of width $(1-\alpha/2)N$. The amplitudes are normalized so that if the inner product of an amplitude with itself is calculated, it returns one.

In order to decrease the number of vectors needed to describe the EOB catalog and to generalize the data set, PCA is performed on the waveform amplitudes. The basis set is constructed by determining the eigenvectors of the dataset's covariance matrix. In this study, the amplitudes in the catalog $\{H_1, H_2, \ldots, H_M\}$ were arranged into a matrix **H**. **H** has dimensions of M x N, where M represents the number of waveforms and N represents the length of the amplitudes. A matrix of the mean-subtracted amplitudes, $\mathbf{\Psi}$, is formed whose columns are given by

$$\Psi_i = H_i - \frac{1}{M}\sum_{i=1}^{M} H_i \ .$$

The covariance matrix, with dimensions N x N is then determined by

$$C = \frac{1}{M}\Psi\Psi^T.$$

A set of basis vectors, $\{e_1, e_2, \ldots, e_M\}$ that span the parameter space is formed by the normalized eigenvectors of **C**. The covariance matrix's eigenvalues, λ_i, provide information about how the eigenvectors span the parameter space. The eigenvectors are ordered by their eigenvalues. The first principal component (PC), with the largest eigenvalue, represents the direction of the largest variance in the catalog. Each subsequent PC accounts for as much variation as possible under the constraint that it is perpendicular to the preceding PCs in parameter space. Any waveform can be constructed using a linear combination of the principal components. The weights assigned to the PCs are called the beta values. This study is concerned with only the first and second PCs as they represent the vast majority of the variance of the catalogs studied.

After PCA is performed on the catalog, the first beta value corresponding to the first PC is plotted as a function of mass ratio. A Gaussian interpolant is fit to the points and the 95% confidence interval is also plotted. The purpose of the interpolant is to inform how to construct an amplitude that is not contained in the original catalog by providing an estimate of the beta value to use with the first PC. Using the prediction and +/- 95% confidence level interval, the mass ratio corresponding to the largest range in beta value is determined. This mass ratio is denoted by m_1. Likewise, m_2, the mass ratio corresponding to the second largest range in beta value is also determined. At m_2, a grid of ten equally spaced beta values between the minimum and maximum of the confidence interval is determined. At each of these points, an EOB amplitude is constructed by multiplying the beta value corresponding to the given mass ratio by the first principal component and then adding this to the mean.

Then, in the same way the original catalogs were created, an EOB amplitude is created with mass ratio m_2. The mean squared error is calculated using each of the constructed EOB amplitudes and the true EOB amplitude. The distribution of the mean squared error values is plotted as a function of the first beta value.

Next, in order to determine how the interpolation changes with the addition of an extra amplitude, a new catalog is created that is identical to the original catalog but with the addition of an amplitude with mass ratio m_1. PCA is re-performed on the catalog. Once again, a range of beta values are selected at mass ratio m_2. The EOB amplitudes are constructed using these new beta values, the new first PC, and the new mean. The distribution of the mean squared error values is plotted again and compared to the first distribution. If the values are significantly minimized with the addition of the new amplitude, it proves that it is beneficial to add a new simulation in this area of parameter space.

Finally, all of the above steps are repeated for the second PC.

Results

The bhextractor code was initially run for two BBH system configurations to generate two catalogs. In the first catalog, the mass ratios varied from one to ten with a step size of one. The inclination of the system was set to be zero, corresponding to a face-on orientation. The second catalog was identical to the first except the orientation was changed to 90 degrees, or an edge-on orientation. Figures 1 and 2 show aligned, tukey-windowed amplitudes for the face-on and edge-on systems, respectively. The amplitudes for the edge-on system are more morphologically complex than the ones from the face-on system because as the inclination increases, higher-order modes from spherical harmonics contribute more to the resulting signal [9].

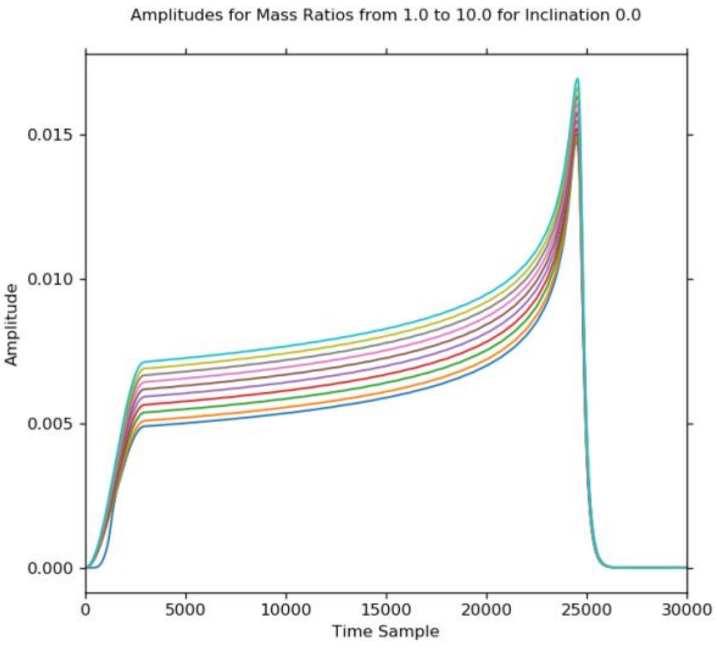

Figure 1. A catalog of amplitudes with *q*=1-10 and *i*=0.0.

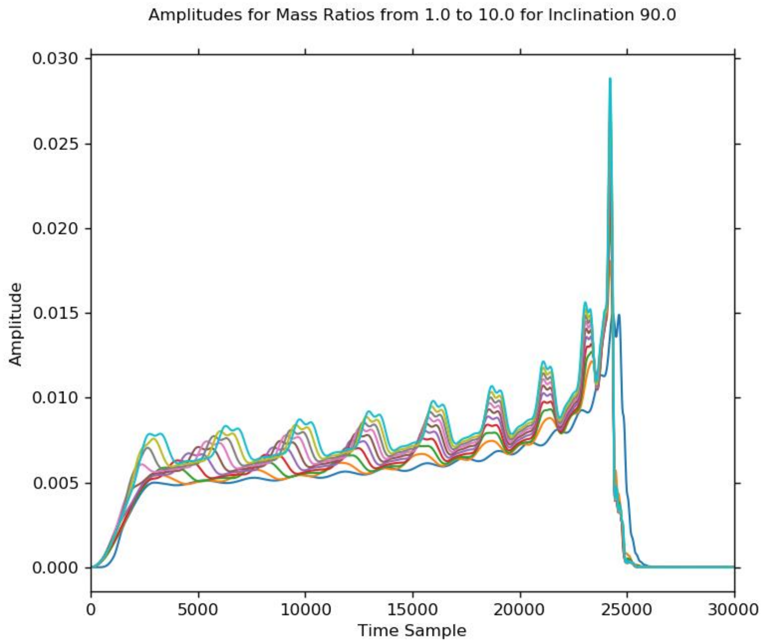

Figure 2. A catalog of amplitudes with *q*=1-10 and *i*=90.0.

After PCA was performed on the two catalogs, the explained variance ratio was plotted. As seen in Figure 3, for the system with face-on inclination, nearly all of the variance in the catalog is captured by the first PC. This suggests that PCA is unnecessary. One amplitude can simply be scaled by a factor to obtain another amplitude in the catalog. However, for more complex waveforms, such as those from edge-on systems, PCA is a beneficial method. The first PC accounts for less than 80 percent of the total variance. Thus, the second PC still contributes significantly to the total variance. Therefore, only edge-on systems were examined for the duration of this study.

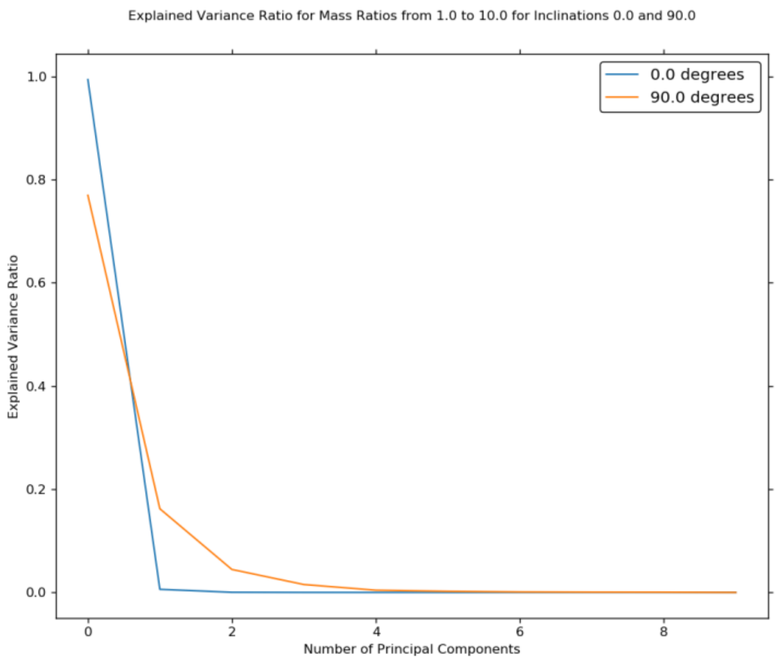

Figure 3. Comparison of the explained variance ratios for the catalogs created from BBH systems oriented at 0.0 and 90.0 degrees.

Figure 4 shows the first two PCs for the edge-on catalog with mass ratios from 1-10 in step size of one. This catalog is called the “original” catalog.

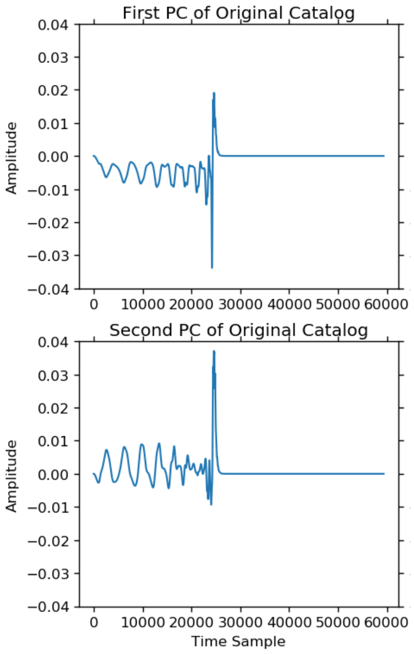

Figure 4. The first two PCs for the original catalog with *q*=1-10 and *i*=90.0.

Figure 5 shows the first beta value plotted as a function of mass ratio for each amplitude in the original catalog, the interpolation, and the 95% confidence interval. Although the eye may deceive you, the widest range in the first beta value corresponds to a mass ratio of 1.47 (m_1). The second widest range in the first beta value corresponds to a mass ratio of 9.53 (m_2).

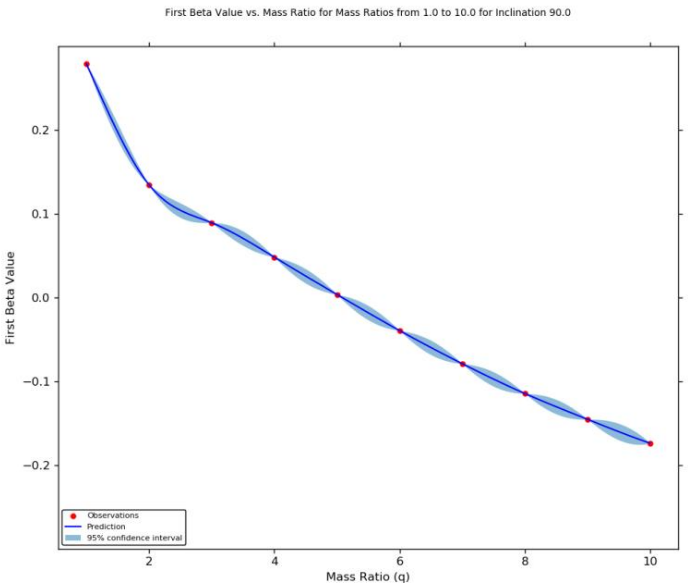

Figure 5. The interpolation of the first beta value as a function of mass ratio for the original catalog with *q*=1-10 and *i*=90.0.

For the “new” catalog with the additional amplitude at mass ratio m_1, the first two PCs look very similar to those for the original catalog (Figure 6).

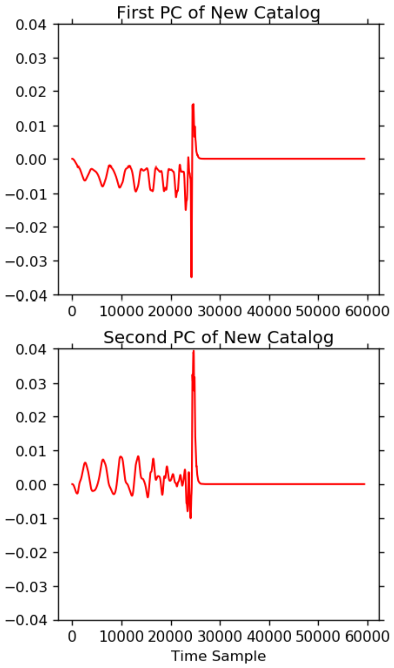

Figure 6. The first two PCs for the new catalog with *q*=1-10 and *q*=1.47 and *i*=90.0.

The interpolation plot changes with the addition of a waveform at mass ratio 1.47 (Figure 7). The 95% confidence interval range for the beta values is widened for some of the mass ratios. The mean squared error distributions for each catalog were overlaid on each other in order to aid in the comparison (Figure 8).

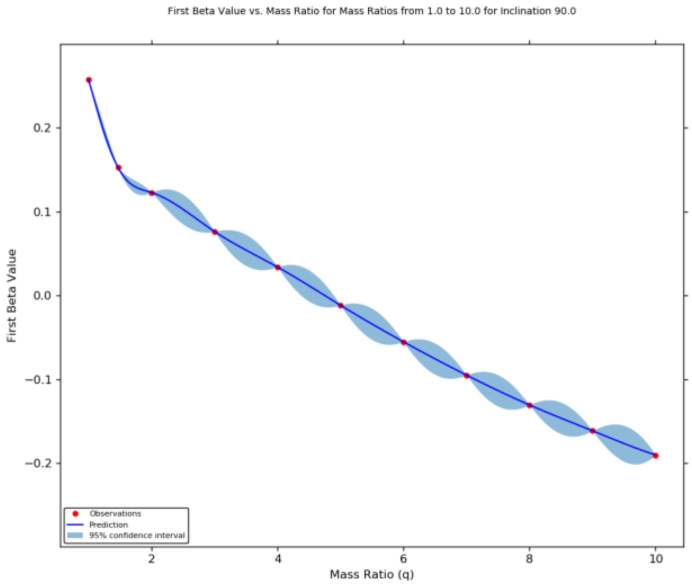

Figure 7. The interpolation of the first beta value as a function of mass ratio for the new catalog with q=1-10 and q=1.47 and i=90.0.

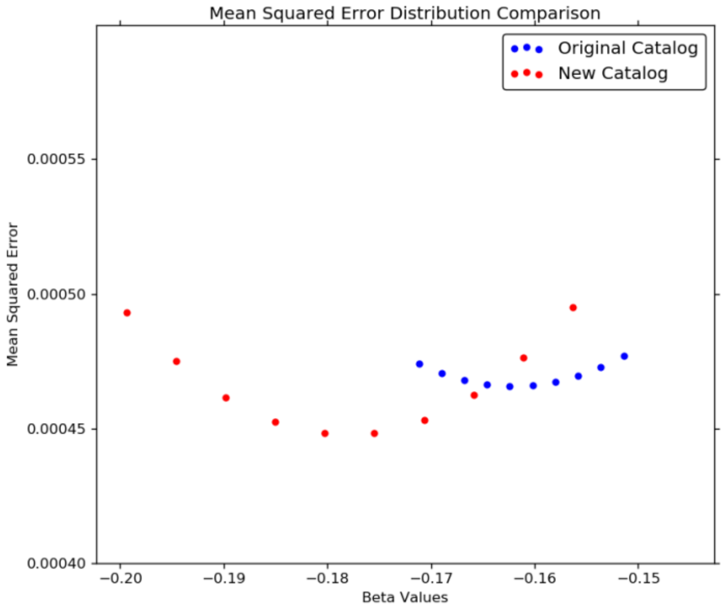

Figure 8. Comparison of the mean squared error distribution for the first beta value for the original and new catalogs.

The same procedure was performed for the second PC. Once again, the interpolation plots for the second beta value as a function of mass ratio change with the addition of the new amplitude at mass ratio m_1 (Figures 9 and 10). Finally, the mean squared error distributions for each catalog were once again overlaid on each other in order to aid in the comparison (Figure 11).

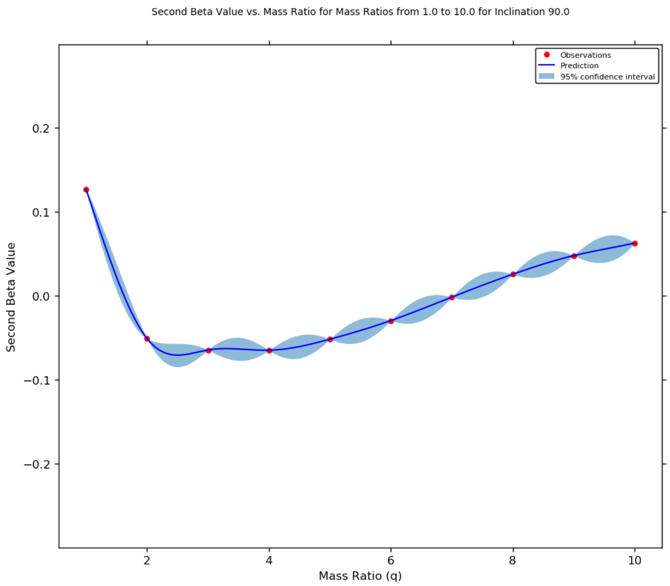

Figure 9. The interpolation of the second beta value as a function of mass ratio for the original catalog with q=1-10 and i=90.0.

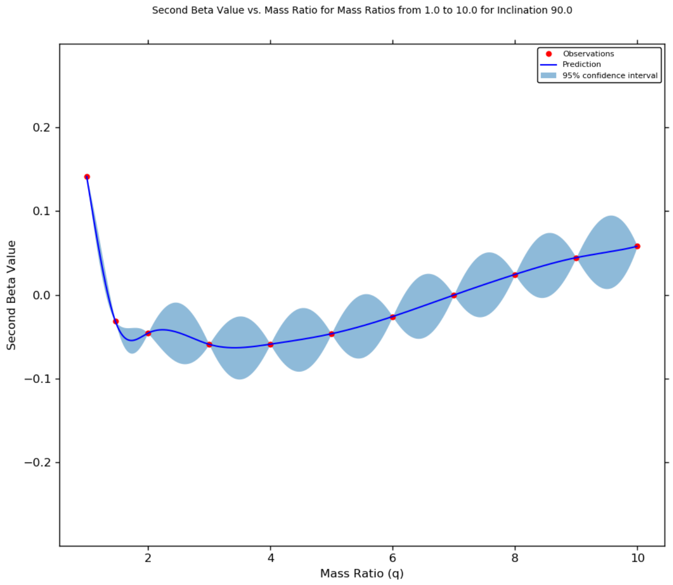

Figure 10. The interpolation of the second beta value as a function of mass ratio for the new catalog with *q*=1-10 and *q*=1.47 and *i*=90.0.

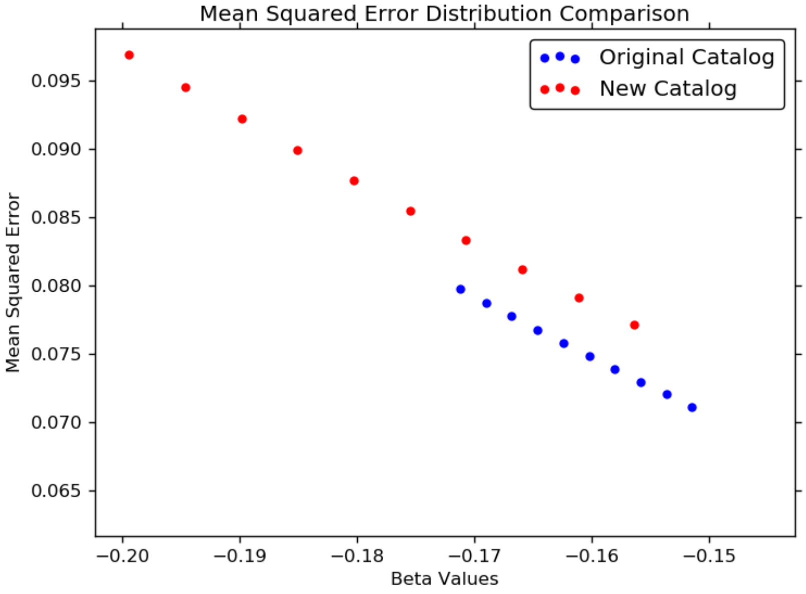

Figure 11. Comparison of the mean squared error distribution for the second beta value for the original and new catalogs.

Discussion/Analysis

This work has shown that it is only advantageous to use PCA to reduce the dimensionality of morphologically complex waveforms. PCA is an unnecessary methodology to apply to simpler waveforms such as those from face-on systems.

One of the main objectives of this study was to answer the following questions: Is it beneficial to add the new amplitude at mass ratio m_1? Does the interpolation significantly improve?

For the first PC, the 95% confidence interval for the beta value range becomes wider for many points in the new interpolation. For example, the range of beta values for mass ratio m_2 changes from approximately -0.17 to -0.15 to -0.20 to -0.165. This means that if an EOB amplitude is constructed with a beta value at the lower end of the new range of beta values, the fluctuation from the mean is weighted too much. On the contrary, if an EOB amplitude is constructed using a beta value at the higher end of the new range of beta values, the fluctuation from the mean is weighted too little. Looking at the mean squared error distribution plot is further confirmation that the additional amplitude at mass ratio m_1 is unnecessary. Although some of the mean squared error values are decreased with the addition of the amplitude at mass ratio m_1, they decrease by a small amount.

It is also beneficial to study the second PC. Once again, the 95% confidence interval for the beta value range is widened for many points in the new interpolation. The mean squared error values are actually higher for the new catalog than they are for the original catalog. This indicates that the original interpolation does a better job representing the true EOB amplitude than the new interpolation. The work completed thus far shows that the PCs can be used to construct EOB amplitudes in nearby areas of parameter space.

Future Work

Further insights into the utility of applying the PCA methodology could be achieved by applying the PCA methodology to NR waveforms, since they are the waveforms that best represent our understanding of GR. In addition to studying BBH systems, it would be instructive to also study binary neutron star systems, especially given the recent detection of a gravitational wave from one such system [10].

Hopefully, future research will be able to generate PCA templates that can be given to the LIGO Scientific Collaboration numerical modelling community in order to refine existing waveform models. As LIGO and VIRGO detect more gravitational waves, techniques will need to be employed to reduce computational cost. PCA might prove to be one such method that maintains high levels of accuracy, but provides a less computationally expensive alternative to methods currently used.

References

[1] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. **116**, 061102 (2016).

[2] K. Jani *et al.*, Class. Quantum Grav. **33**, 204001 (2016).

[3] F. Pretorius, Phys. Rev. Lett. **95**, 121101 (2005).

[4] M. Campanelli, C. O. Lousto, P. Marronetti, and Y. Zlochower, Phys. Rev. Lett. **96**, 111101 (2006).

[5] J. G. Baker, J. Centrella, D. Choi, M. Koppitz, and J. van Meter, Phys. Rev. Lett. **96**, 111102 (2006).

[6] J. Blackman *et al.*, Phys. Rev. D. **96**, 024058 (2017).

[7] T. Damour, Int. J. Mod. Phys. A **23**, 1130 (2008).

[8] T. Damour, A. Nagar and L. Villain, Phys. Rev. D **89**, 024031 (2014).

[9] I. Harry, J.C. Bustillo, and A. Nitz, LIGO Document P1700262, (2017).

[10] B.P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. **119**, 161101 (2017).