Developing the Astronomy of Black Hole Mergers with Gravitational Waves

Dustin Leininger

August 11, 2014

Contents

Chapter 1

Introduction

1.1 Future Astronomy

As physicists continues to probe further back into the history of the universe and deeper into the mechanisms of stellar events, there also grows a demand for more observational data. Thus far, the fields of astronomy and observational astrophysics have been dominated primarily by work focusing on physics deductible from observations of the electromagnetic spectrum. However, there are many instances in which these methods simply fail to shed light on phenomenon of interest. Some such examples include the physics within the core of an exploding star, the evolution of black holes, as well as the mechanics of certain binary stellar systems and their mergers. Fortunately, there are other means of deducing physics concerning these and other cosmic events for which traditional methods fall short.

The detection of gravitational waves would not only provide one of the final verifications of Einstein's theory of general relativity, but also a new observational resource for astronomy capable of probing phenomenon previously inaccessible by electromagnetic observations. Gravitational waves are propagating perturbations in spacetime generated by quadrapole moments in matter. These are normally simply accelerating asymmetric masses including such things as orbiting binaries or spinning bodies with significant mountains and valleys. The larger the acceleration or the asymmetry, the more significant the perturbation. For these reasons binary black hole systems are likely candidates for gravitational wave observation particularly while in their rapidly orbiting late stage inspirall, merger, and ring down phases.

Since gravitational waves are generated by the motion and distribution of mass, they can encode and carry physically significant data about the evolving states of their progenitors regardless of their emission of electromagnetic radiation. Therefore, gravitational waves offer an alternative method of observing the cosmos allowing for the possibility of achieving unprecedented insights into physical processes.

1.2 Data Analysis

In order to develop a new astronomy, a complete data analysis machinery must be developed and set in place which can deduce physics from observations. Like observations of electromagnetic radiation, certain information about the progenitors of the gravitational waves can be inferred from various properties of the waves themselves. Among the properties to be determined, the sky location of progenitor events is of central importance. However, a surprising number of insights might be gained from the analysis of detected waves.

Briefly, there are two broad classifications of gravitational waves; continuous waves and bursts. The former are generated by events which occur on long time scales which continuously emit gravitational waves throughout their evolution. Inspiralling binaries are one such example which can emit gravitational waves for many years. The latter classification refers to progenitor events which occur on a much shorter time scale. These include events such as supernova which emit gravitational waves with durations on the order of seconds or less. Some events may transition from one type to the other during their evolution. In the late stage inspirall, merger, and ringdown of a binary system gravitational waves are produced that are distinct from the earlier stages of the inspirall. Firstly, the amplitude of the waves greatly increases during these stages. This means that for binaries with continuous waves of small amplitude, these late stage events may be all that can be detected. Lastly, the merger and ringdown occur very quickly and therefore fall into the domain of burst gravitational waves. It is with these kinds of burst waves that the current report is primarily concerned, and more specifically the determination of the sky location of their sources. In Fig. 1.1, an example waveform of this kind can be seen.

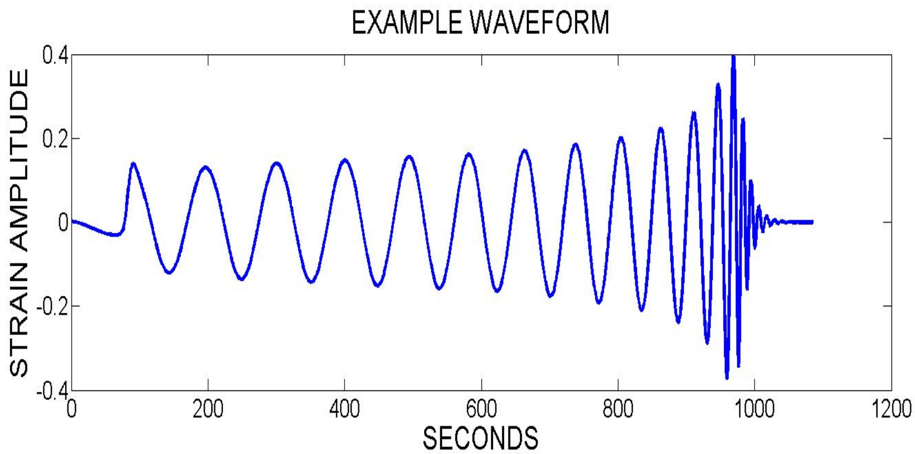


Figure 1.1: This is a plot of a simulated gravitational wave signal generated by a binary black hole system in which the black holes were of equal mass and zero spin. As can been seen, the early segment of this wave is very regular in frequency and amplitude. The time during and before this stage of the systems evolution are a source of continuous gravitational waves. There can also been seen a remarkable and rapid increase in both amplitude and frequency towards the later stages of the systems evolution during which the binary system merges and rings-down. This region falls under the perview of burst gravitational wave analysis. For very distant or small systems this may be the only detectable stage of the entire evolution of the system.

Chapter 2

Methods

2.1 NR-Simulation Catalogs

Given the complexity and nature of Einstein's field equations, many systems are analytically intractable. Therefore, much work has been done to solve these systems numerically. The development of data analysis tools is heavily dependent on these simulations provided by numerical relativity. First, they act as the predictions against which observations must be compared in order to test theory. Second, during the development of the techniques and tools of data analysis, they act as test data on which an algorithm and/or code maybe be tested to check for validity. However, these simulations are difficult to construct and can be computationally expensive. Therefore, the available simulations can be limited in variety.

For analysis purposes, the simulated waveforms are usually grouped together into sets called catalogs. These are matrices each column of which is a different simulated waveform. All of the waveforms in a catalog usually share a certain set of fixed parameters while other parameters are varied so as to created an approximate representative sample of the given set of fixed parameters. In the analysis reported here, three catalogs were implemented. The waveforms from each catalog all came from simulations of binary black hole systems. All of the waveforms in the first catalog, referred to as the Q-series, came from simulations in which neither of the black holes were spinning. There were 33 waveforms in the Q-series over which the mass ratio of the black holes was varied. The 81 waveforms in the second catalog, called the HR-series, were generated in simulations in which both black holes were spinning at the same rate and for which their spin angular momenta were parallel to the orbital angular momentum of the system. In this catalog, the mass ratio was again varied as was the magnitude of the spins. Lastly, the third catalog, referred to as the RO3-series, contained 19 waveforms in which the spin angular momenta of the black holes were allowed to freely precess during inspiral. The mass ratio, spin magnitude, and initial angular displacement between the spin angular momenta of the black holes were varied in each waveform (one spin initially parallel to the orbital angular momentum and the other initially askew varied angles).

Work has already been done to successfully enable the identification of unknown waveforms as belonging to one of these catalogs. This effectively allows for the identification of the type of source of the gravitational wave as well as some of its other parameters. The focus of the work reported here was to broaden the capabilities of the analysis to include the ability to identify the sky location of the progenitor allowing for a more complete astronomy of gravitational waves.

2.2 Principle Component Analysis [1]

The first step of the analysis used here is Principle Component Analysis (PCA) which implements a matrix decomposition technique from linear algebra called Singular Value Decomposition (SVD) to factor the catalog into more computationally useful matrices.

Singular Value Decomposition

$$A_{(mxn)} = U_{(mxm)} \Sigma_{(mxn)} V^T_{(nxn)} \tag{2.1}$$

SVD decomposes a matrix A into three factor matrices. The first is often denoted by U. It's columns form what are called the left singular vectors which obey the equation $c_1 A = c_1 \sigma_1$ where c_1 is the first column of U and σ_1 is its corresponding singular value. The third matrix is denoted by V^T. The columns of V form the right singular vectors of A and obey the same equations as the the columns of U with the exception that the columns of V operate on the right of A as indicated by its name. The middle matrix (Σ) is a diagonal matrix whose elements correspond to the singular values of the columns of U and V.

U and V also have other useful properties on which PCA is based. The columns of U are the eigenvectors of the the covariance matrix of A (AA^T) and their singular values are the corresponding eigenvalues. Similarly, the columns of V and their corresponding singular values form the eignevectors and eigenvalues of A^TA. Ultimately, this means that we can use either the columns of U as a basis for the column space of A or the columns of V as a basis for the row space of A. Since the ultimate goal is to be able to determine certain physics concerning gravitational waves (which form the columns of the catalogs) the matrix U and its eigenvalues are of primary interest.

When the matrix A is a waveform catalog, the column-wise normalized matrix U becomes an orthonormal basis set which spans the column space of the catalog. Each basis vector (column of U) of the column space of A encodes a predominant feature shared among the waveforms in the catalog. In more mathematical terms, the eigenvalue is the scalar resultant of the projection of its corresponding basis vector onto the catalog and therefore a measure of its prevalence in the catalog. This means that the basis vectors may be ordered by their prevalence in the catalog in accordance with their eigenvalues (from greatest to least). Looking back at Eq. 2.1, there are as many basis vectors as there are waveforms in the catalog which means that there is (ideally) no loss of information in the decomposition.

Recall from above that there are as many basis vectors (columns of U) as there are waveforms in the catalog used in the PCA. This number can often be rather large for computation purposes. However, it is possible to use fewer than the total number of basis vectors so long as enough are retained to sufficiently span the space. This can be determined using an eigenvalue energy method. The eigenvalues of all the basis vectors were collected into a single vector, ordered from greatest to least, which was then normalized. Then the values were summed until the total reached some acceptable threshold (.9 in the analysis reported here). Then only the basis vectors whose eigenvalues contributed to the sum were used in the rest of the analysis. In this way, a much smaller computationally tractable approximate basis was constructed which still sufficiently spanned the catalogs. The basis vectors used in the final analysis are called the Principle Components (PCs). As can be seen in Fig. 2.1, remarkably few PCs (compared to the size of the catalogs) are required to span each catalog. In the analysis reported here, 2 PCs were used on the Q-series analysis, 5 for the RO3-series analysis, and 8 for the HR series.

If the original catalog is large enough, then U can be considered to be the basis for all waveforms sharing the same parameters as those in the catalog. As an example, the columns of the normalized U matrix which results from the application of PCA to the Q-series catalog may be used as an approximate basis for all waveforms generated by non-spinning binary black hole systems and not just those simulations used in the Q-series catalog. In this way, this algorithm gains the ability to effectively analyze unsimulated waveforms. The amazing ability to generalize the available simulations and accurately handle unsimulated waveforms is part of what makes analysis so powerful.

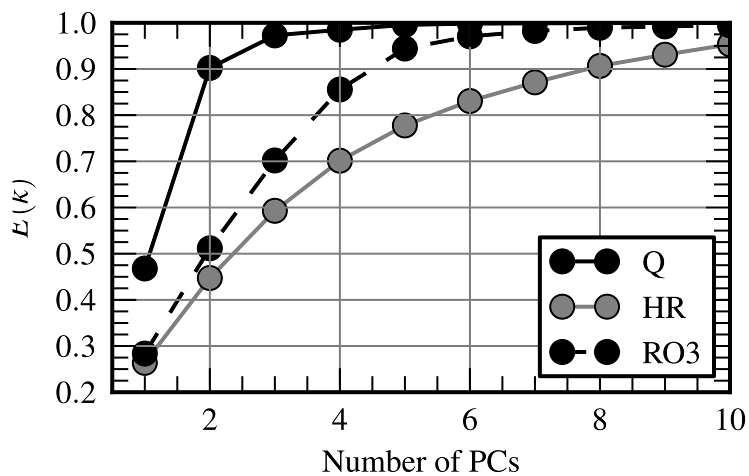


Figure 2.1: In this figure, the eigen-energy is plotted as a function of the number of PCs. Impressively few of the PCs are needed in order to sufficiently span the catalogs. Only 2 PCs are required for the Q-series, 5 for the RO3-series, and 8 for the HR series.

2.3 Bayesian Model Selection and Parameter Estimation [1] [2]

The second crucial element of analysis reported here is based on Bayesian Statistics in which Bayes Theorem pays a central role. Bayes Theorem relates the probability the that a hypothesis is true given the data set $\{D_k\}$ (called the posterior and denoted $P(H|\{D_k\}, I)$) to the probability that the data would have been observed given that the hypothesis were true (called the likelihood and denoted $P(\{D_k\}|H, I)$). This is extremely useful as it is usually much easier to compute the likelihood than the quantity of primary interest, the posterior. In Bayes Thm., the posterior is also related to the probability that the hypothesis is true without support from the data (called the prior and denoted $P(H|I)$) and the probability that the data would have been observed given no hypothesis (called the evidence and denoted $P(\{D_k\}|I)$). However, the evidence is often ignored as a scale factor to be computed during normalization after the fact. Another important note is that at large samples the likelihood almost always dominates the prior. However, both quantities can be of essential importance depending on the problem at hand.

Bayes Theorem

$$P(H|\{D_k\}, I) = \frac{P(\{D_k\}|H, I) \times P(H|I)}{P(\{D_k\}|I)} \tag{2.2}$$

Bayes Thm. is primarily used as a parameter estimation tool in which the resulting posterior is the probability distribution function of some quantity of interest. In the current analysis the sky locations (right ascension and declination) of the source were of primary concern. To ensure an unbiased posterior, a prior is normally chosen which exhibits maximal ignorance. This often results in the assignment of a flat prior distribution in which each possibility is given equal weight. In the analysis reported here, flat priors were implemented for both the right ascension and declination. Although, it is important to note that a flat prior is not the most unbiased assignment for the declination as it will favor the poles. Since the likelihood function so strongly dominates the analysis, this was not seen as a serious error but will be revisited and revised in future work (for a full length discussion of the sky localization analysis, see 2.4).

Since the likelihood is meant to be the probability of measuring the data given the hypothesis to be true, it is often modeled after the expected noise of the observed data. For an uncorrelated data set with Gaussian noise, the likelihood is simply the product over the Gaussian noise on each data point in the set. The addition of each data point in the set alters the posterior distribution until it reflects the estimate preferred by the data. The Likelihood function used in this analysis is Eq. 2.3. The "model" term in Eq. 2.3 is a reconstruction from the PCs using parameters which are at first randomly chosen but then begin to converge to higher likelihood values using the Nested Sampling Algorithm described in [2]. For computational purposes, the logarithm of the likelihood was computed which simply and conveniently changes the product to a sum.

Likelihood Function

$$\log(L) = -2\ \Delta F \times \sum_i \frac{(wave(i) - model(i))^2}{noise(i)} \tag{2.3}$$

In earlier work, which preceded this report but with which this report is still concerned, the focus was to determine which among the catalogs an injected signal most likely belongs thereby determining the most likely type of source (among other parameters). This can be done by evaluating the ratio of two posteriors, the first in which the "model" in Eq. 2.3 was generated by the PCs of one catalog, and the second in which the "model" in Eq. 2.3 was generated using the PCs of a second catalog. By assuming *a priori* that the waveform has an equal probability of being from either catalog, the priors for the posteriors of the two catalogs become equal and cancel out in the ratio. Furthermore, the evidence in the denominator of each posterior can be ignored at this stage as a scale factor. This leaves only the ratio of the likelihoods which is often referred to as the Bayes Factor. Also, in the literature, the proposition "that the waveform belongs to the first catalog" is referred to as Hypothesis 1 and is usually denoted H_1. A similar convention is adopted for the proposition "that the waveform belongs to the second catalog". Again, it is common to compute the logarithm of the Bayes Factor such that a positive $B_{H_1H_2}$ in Eq. 2.4 is consistent with a preference in the data for H_1 and a negative $B_{H_1H_2}$ with a preference for H_2.

Bayes Factor

$$B_{H_1H_2} = \frac{P(\{D_k\}|H_1, I)}{P(\{D_k\}|H_2, I))} \tag{2.4}$$

The probability that a waveform belongs to a given catalog i ($P(\{D_k\}|H_i, I)$) is determined by the fit and coefficients of the linear combination of the PCs belonging to that catalog used in the likelihood function. This means that $P(\{D_k\}|H_i, I)$ is a multi-parameter probability. Therefore, what is needed is a method of reducing the multiple posteriors (one for each coefficient in the linear combination) into a single posterior for the entire hypothesis. This is done in two steps. The first is to use Eq. 2.5 known as the Product Rule which states that the probability, given only that the hypothesis is true, of observing data $\{D_k\}$ corresponding to a fixed set of parameters (in this case coefficients) $\{\theta_i\}$ is equal to the product of the probability that the data would have been observed, given the parameters and the hypothesis, and the probability that given that the hypothesis were true given that the set of parameters.

Product Rule

$$P(\{D_k\}, \{\theta_i\}|H) = P(\{D_k\}|\{\theta_i\}, H) \times P(\{\theta_i\}|H) \tag{2.5}$$

The second step to computing ($P(\{D_k\}|H_i)$) is Eq. 2.6 and is called the Marginalization Rule. It allows for the elimination the dependence on $\{\theta_i\}$ in $P(\{D_k\}, \{\theta_i\}|H_i)$ by summing over all possible values of the elements of $\{\theta_i\}$. Since the values of the coefficients vary continuously, the summation becomes an integral. The limits of integration are determined by the upper and lower bounds on the prior for the parameter estimation on the coefficients.

Marginalization Rule

$$P(\{D_k\}|H_i) = \int ... \int_i P(\{D_k\}|\{\theta_i\}, H_i) \times P(\{\theta_i\}|H_i) \; d\{\theta_i\} \tag{2.6}$$

By sampling the likelihood and prior, it is ultimately this integral which is computed by the Nested Sampling Algorithm. As a final note, the Bayes Factor computed in actual analyses is always the ratio of H_i to noise which is simply checking to see if there is any signal present at all. To discern the overall Bayes Factor between two hypotheses, simply take the difference between their respective Bayes Factors.

2.4 Sky Localization

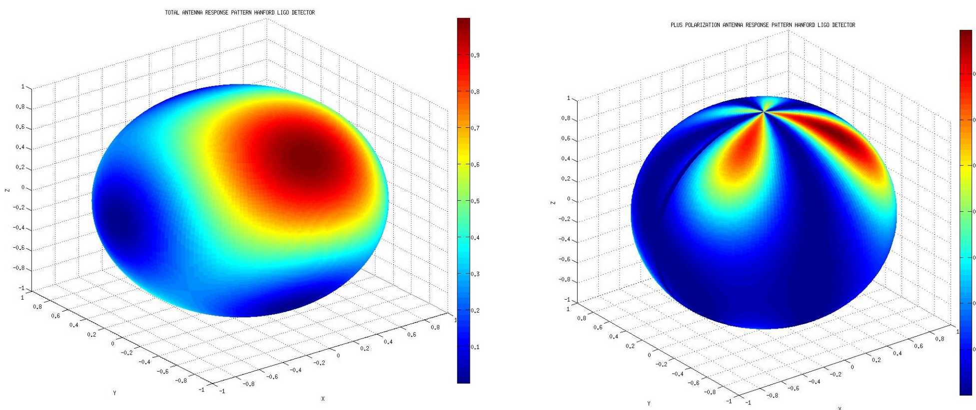
As mentioned earlier, the primary focus of the work done this summer was to broaden the capabilities of the existing analysis to include estimates of the sky locations (including right ascension and declination) of the sources of gravitational waves generated by binary black hole systems. In order to determine the sky location, multiple detectors must be used in the analysis. This introduces new complications to the analysis including the time delay between various detectors and the different antenna response patterns of the detectors which were the primary focus in this report.

First, the time shift between the different detectors physically originates in their different locations on the surface of the earth. A single wave which passes through multiple detectors will reach each detector at a slightly different time. The detected signals therefore need to be time shifted and aligned such that they are able to compared. In the analysis, the center of the earth was taken to be the origin. Since each detector remains stationary on the surface of the earth with respect to the Earth's center, the time shift from the center to a given detector remains constant. Once the simulated signal is injected at each detector, a short segment of code simply applies the calculated time shifts of each detector to their respective injections effectively time shifting them all to the center of the earth.

Next, some time must be taken to briefly describe the antenna response pattern of the detectors and how they effect the analysis. The gravitational wave detectors are extremely large interferometers meaning that they have an intrinsically planar geometry. This implies that the sensitivity of the detector has directional bias. They are optimally sensitive to waves propagating normal to their plane and minimally responsive to those traveling parallel. In Fig. 2.2 (A), the computed sensitivity of the Washington LIGO detector as it varies over the entire sky is projected onto the celestial sphere. It can bee seen that the detector is optimally sensitive to waves which come from directly above and below the detector and diminishes to zero and the peripheries.

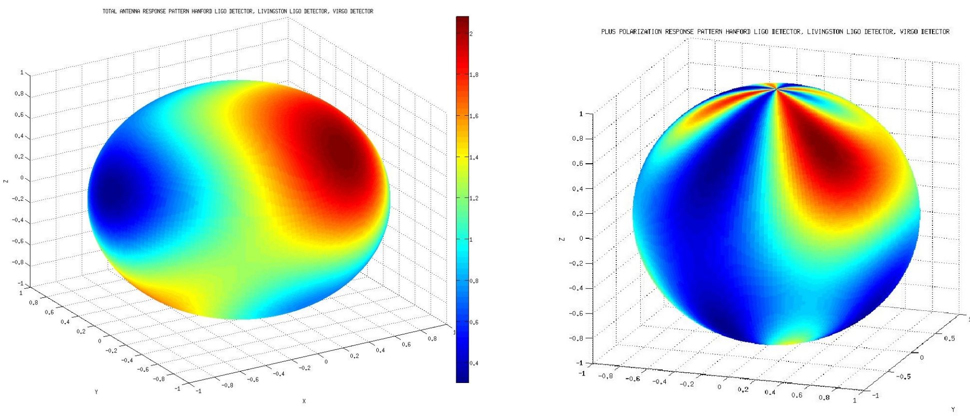
Two things are important to note at this point. The first is that the antenna response is dependent on both the plus and cross polarization of the incident gravitational wave. For computational expediency and testing purposes, only the plus polarization was applied to injected waves in the current analysis. Plots of the antenna response pattern resulting only from the plus polarization can be seen below in Fig. 2.2 (B) and Fig. 2.3 (B). Future work will implement both the plus and cross polarization for a more complete and real world analysis. Second, is the significant effect of the antenna response pattern on the Bayesian parameter estimation.

One of the most outstanding effects of the antenna pattern is that it scales the Bayes factors. This is expected as it follows directly from the fact that it is harder to distinguish waves that cannot be as well detected. For the same reasons, it is reasonable to expect that the sky location estimates in the "darker" regions will be less reliable than those in more "visible" regions. These expectation full be explored more thoroughly in Ch. 3.



(A) **(B)**

Figure 2.2: In (A) the total response pattern of the Hanford LIGO detector including both plus and cross polarization contributions as projected onto the celestial sphere. It is clear that detector is most sensitive to the areas on the sphere which lie directly above and beneath it and attenuates to zero along it periphery. In (B) the antenna response pattern of the Hanford LIGO detector resulting from only the plus polarization contribution. This is a much less physical result as it depicts only the detector's response to the plus polarization components of prospective gravitational waves as their sources vary over the sky.



(A) **(B)**

Figure 2.3: In (A) the combined total response pattern of the Hanford LIGO detector, Livingston LIGO detector, and the Virgo detector including both plus and cross polarization contributions is projected onto the celestial sphere. The combination of three detectors has significantly increased the visibility. Particularly, there is no longer a connected bind of blind space which wraps around he entire sphere. In (B) the antenna response pattern of the aforementioned trio of detectors resulting from only the plus polarization contribution is projected onto the celestial sphere. Again, this is a much less physical result as it depicts only the detectors' response to the plus polarization components of prospective gravitational waves as their sources vary over the sky. However, this is the only component used in the analysis presented later in this report and this is therefore the antenna response pattern implemented in the analysis.

Chapter 3

Results

In this analysis, 48 sky locations were selected for injection sites using the program MEALPix (more information about MEALPix can be found at http://www.gwastro.org/for20scientists/mealpix-matlab-healpix-interface-1). Each waveform from each catalog was injected at each of the 48 sky locations and then compared against each of the three catalogs in the analysis. This ultimately resulted in 6500+ jobs which are currently running on the computing clusters at Caltech. As of the time of writing of this report, only roughly one sixth of the jobs had finished. Fortunately, the analysis output data as it progressed. At least partial data was produced concerning all sky locations of all waveforms of the RO3 catalog. Since the rest of the data was simply too incomplete, all of the preliminary results presented here will are restricted to the analysis of the RO3 catalog.

3.1 Sky Localization

As mentioned earlier the aim of this project was to implement sky location estimates into the analysis. In Ch. 2.4 above, the method for constructing the estimates for the right ascension and declination is described. Each iteration of the nested sampling generates an estimate of the sky location. A histogram of the estimates generates a distribution of likely values examples of which can be seen below in Fig. 3.1 and Fig. 3.2. The waveforms were injected with an SNR of 50 which is particularly high for real world simulations but appropriate to the early stage testing reported here which simply demonstrates the ability to determine sky location.

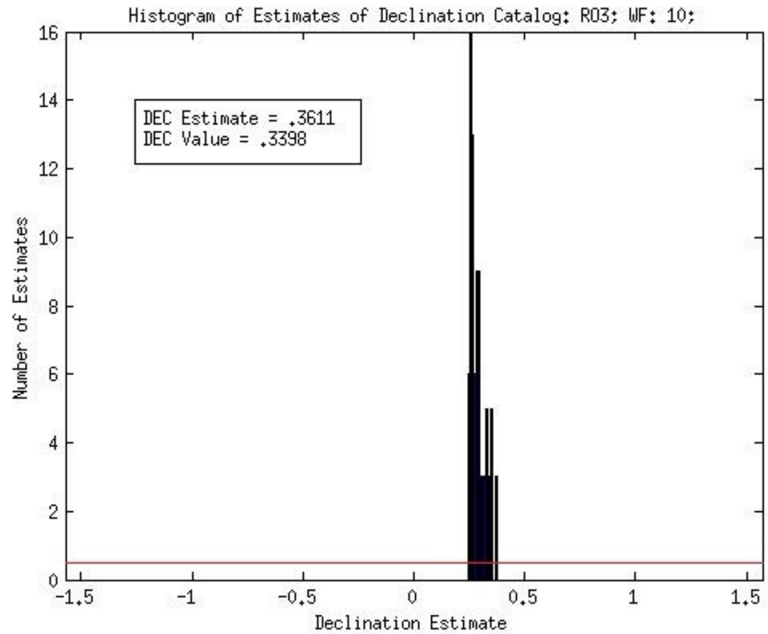


Figure 3.1: Above is plotted a histogram of all the estimates of the declination of the 10th waveform in the RO3 catalog on its entire range of possible values from -$\frac{\pi}{2}$ to $\frac{\pi}{2}$. The flat red line in the diagram is the prior distribution. As can been seen, the distribution is sharply peaked about the best estimate of .3611 which is remarkably close to the true value for the injection which was .3398

These preliminary results seem to indicate that the analysis is able to determine the sky location. The impressive accuracy of the estimates is of course influenced by the high SNR used in the injections. However, it was only the scope of this report to demonstrate the ability to deduce sky location at all.

While the variation in the sky location estimates do seem to correlate as expected with the antenna pattern, the precision of the estimates and the sparse sampling of the sky made it difficult to graphically represent. However, the difference in the actual and estimated sky locations can be

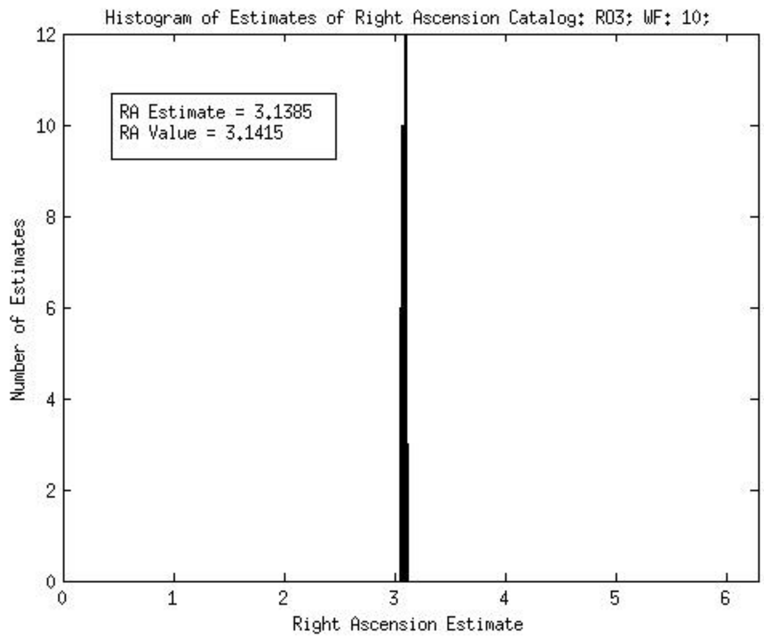
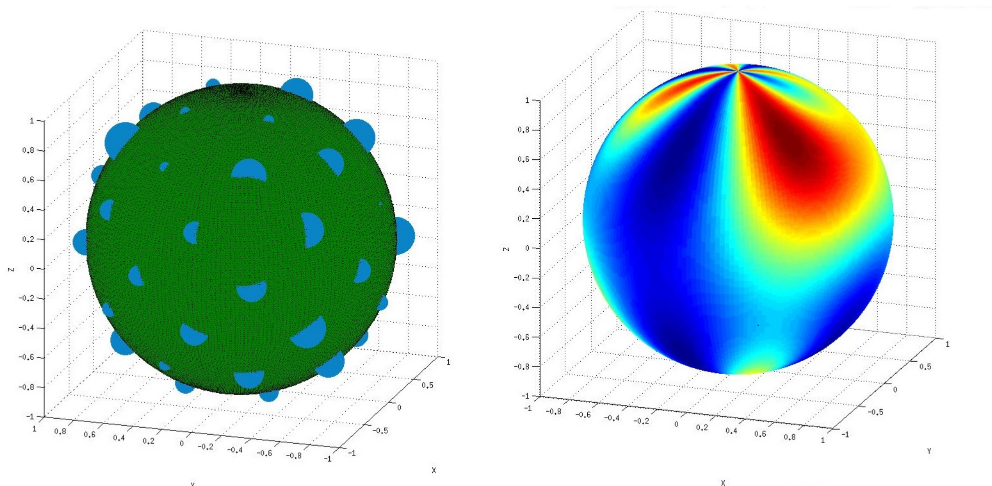


Figure 3.2: Above is plotted a histogram of all the estimates of the declination of the 10th waveform in the RO3 catalog on its entire range of possible values from 0 to 2π. Although not depicted here, the prior distribution for this estimate was similarly flat and very small in amplitude compared to the histogram. Similar to Fig. 3.1, the distribution is sharply peaked about the best estimate of 3.1385 which is again very close to the true value for the injection which was 3.1415

seen as projected onto the celestial sphere in Fig. 3.3 alongside the antenna response pattern used in the analysis and described above in Ch. 2.4.

While these early results seem to bode well for the success of the implementation of sky localization to the analysis much work remains to be done to fully verify the success or failure. When the data is complete, a more thorough and rigorous analysis of the data will be done. Afterwords, the number of sky locations on the analysis will be increased to verify the successful implementation of the antenna response pattern once both plus and cross polarizations are implemented.



(A) **(B)**

Figure 3.3: In Fig. 3.3 (A), the absolute value of the average difference of the actual to estimated right ascension was used as the radius of the circles each of which was plotted at its respective sky location. In Fig. 3.3 (B), the network plus polarization antenna pattern is plotted on the surface of the sphere. While not certainly not conclusive in this preliminary analysis, there can be seen a visual correlation between the radii of the circle (the difference in the estimated and actual values) on the right, and the antenna pattern on the left. Due to the high SNR used in the injections, the differences tended to be very small. Therefore, the radii have been linealy scaled for visualization purposes.

3.2 Bayes Factors

While sky localization is of central importance to the development of a full gravitational wave based astronomy, it is essentially the Bayes factors which are of concern in this algorithm as it determines whether or not a wave has indeed been detected and if so the nature of its progenitor. It is therefore worth the time to quickly look at some of the preliminary data concerning the Bayes factors.

Below can be seen a similar plot to that in Fig. 3.3. However, in this plot, the Bayes factors of signal to noise have been averaged and used as the radii of circles which are again plotted at their respective sky locations on the sphere.

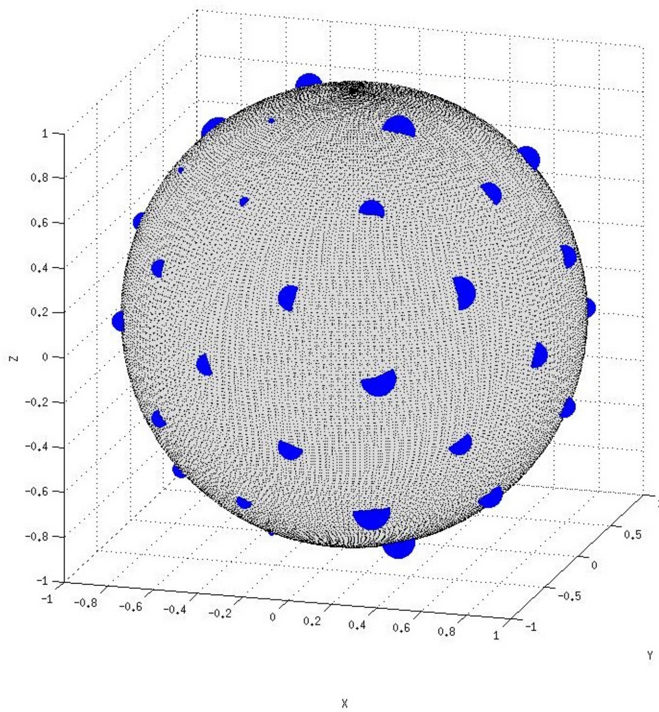


Figure 3.4: I this plot, the average Bayes factor across all wave forms and catalogs has been used as the radii of the circles plotted at the respective sky locations. The size of the circle is therefore a rough indicator of the analsis' ability to distinguish signal from noise at the given sky location. Again, this analysis is influenced by the high SNR. However, the purpose of this plot is again to show the effect of the antenna response pattern on the analysis.

Chapter 4

Conclusions

Given the limited nature of the data presented here, it would be inappropriate to try and draw any hard conclusions. However, It does appear in general that the analysis is capable of deducing sky locations and correctly implementing the combined antenna response patterns as well as other technical problems inherent in the implementation of multiple detectors in the analysis.

Chapter 5

Acknowledgements

I would like to thank Dr. Ik Siong Heng how oversaw my work at Glasgow University as well as James Clark and Nick Mangini at the University of Massachusetts as Amherst for their help and guidance during this assignment.

References

[1] I.S. Heng P. Kalmus J. Louge, C.D. Ott and J.H.C Scargill. Inferring Core-Collapse Supernova Physics with Gravitational Waves. *PHYSICAL REVIEW D*, 86.

[2] D.S. Sivia. *Data Analysis: A Baysian Tutorial*. Oxford University Press, second edition, 2006.