Rates and populations of compact binary mergers

Caltech LIGO SURF 2019

Phoebe McClincy

Pennsylvania State University

Mentor: Alan Weinstein

California Institute of Technology

Version 3.0

9 May 2019

Rates and populations of compact binary mergers

Caltech LIGO SURF 2019

Phoebe McClincy

Pennsylvania State University

Mentor: Alan Weinstein

California Institute of Technology

Abstract

The Advanced LIGO and Virgo detectors have been observing the cosmos in search of gravitational waves (GW) since 2000. Both detectors were upgraded to Advanced versions that began observing in 2015. In their first and second observing runs (O1 and O2, respectively), they detected 10 GW signals from binary black hole (BBH) mergers, and one from a binary neutron star (BNS) merger, all with high significance (low probability of being due to instrumental noise fluctuations). Already in the first month O3, which began in April 2019, two BNS signals, four BBH signals, and one NSBH candidate have been seen with such high significance. These three categories are collectively known as compact binary coalescence (CBC). In the coming years, as the detectors' sensitivity is improved, we expect to accumulate tens, hundreds, or thousands of CBC events. From such large samples, we expect to be able to infer the underlying population of CBC systems as a function of their masses, component black hole spins, and redshift. This, in turn, will allow us to better understand the astrophysical processes governing the formation, evolution, and final fate of such systems, as tracers of the most massive stars. In this project, we aim to develop tools and techniques to accomplish this through detailed simulation and Bayesian inference.

I. INTRODUCTION

Since Advanced LIGO and Virgo began collecting data in 2015, the collaboration has made 14 detections of GW signals from BBH mergers, three from BNS, and has identified one potential NSBH candidate[1]. As our detectors improve, so will the number of events that we are able to recover. In the near future, we expect to detect significantly more events of this nature. With a population size akin to the one we expect to see, we are able to identify and measure redshift dependency of component mass and spin, and use this information to infer the mechanisms of formation and evolution that govern such binaries.

LIGO intends to measure the binary population, and compare it with parametrized models; as such, we can analyze both the primary and secondary component masses. This is parametrized by analyzing the primary component mass, as well as the mass ratio (q) of the primary and secondary component masses.

There are different assumptions that may be made about such parametrization, as well as other parameters. In [2], three models were presented for the BBH primary mass distribution, denoted Models A, B, and C. In regards to mass, Model A fixes m_{min} to be 5 $M_\odot$, and allows m_{max} to vary. Model B allows both mass limits to vary. Model C allows for a second component of Gaussian nature to appear in the distribution due to the pair instability in massive progenitor stars; the power law distribution fits at lower masses, and the Gaussian distribution fits at higher masses. In this project, we aim to produce a best-fit Model C and incorporate redshift dependence for future, larger populations of BBH [2] .

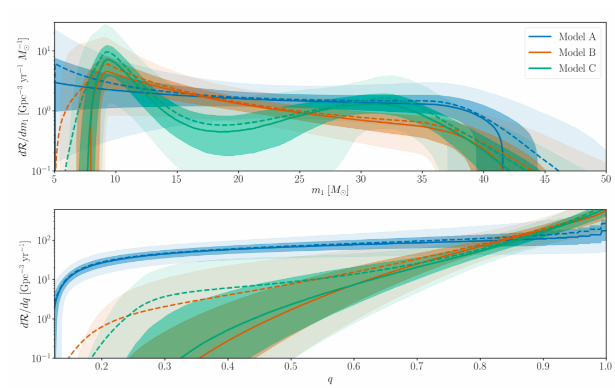

FIG. 1. Differential merger rate distribution for BBH as a function of primary mass and mass ratio (q); Models A, B, and C. At lower masses, Model C follows a power law distribution, and at higher masses, it follows a Gaussian distribution. This distribution is based on data from O1 and O2 only (10 BBH mergers); we aim to revise this model for future, larger populations of BBH[2].

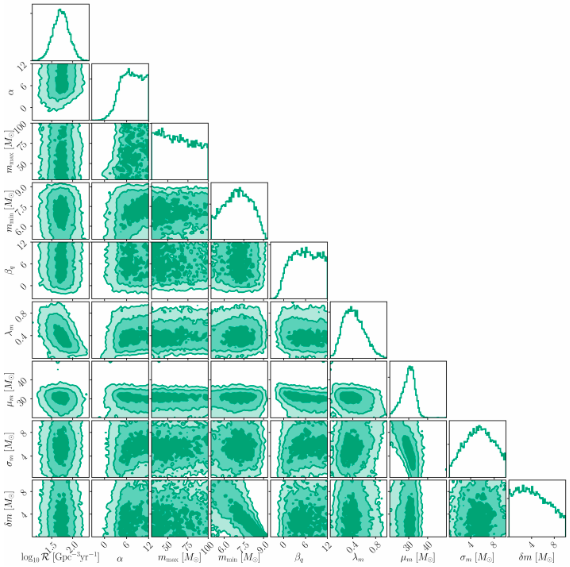

FIG. 2. Posterior distributions for parameters for Model C. The low-mass power law component is described by α, β, $\mathrm{m}_{\mathrm{max}}$, and $\mathrm{m}_{\mathrm{min}}$. The high-mass Gaussian component is described by mean mass μ_m and standard deviation σ_m, as well as the fraction of BH in the Gaussian component, λ_m[2].

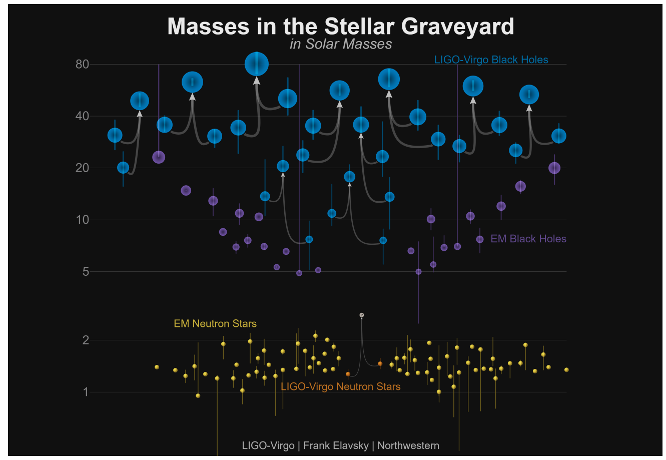

FIG. 3. Distribution of NS and BH masses, as detected via electromagnetic radiation and Advanced LIGO. There are regions of scarcity below 1 $M_{\odot}$ and between 2-5 $M_{\odot}$, as well as a possible mass gap above 50 $M_{\odot}$.

II. MOTIVATIONS

At present, we recognize two regions of scarcity within the mass distribution of NS and BH; one below 1 $M_{\odot}$, and one between approximately 2 and 5 $M_{\odot}$. There is evidence that another, beginning at approximately 50 $M_{\odot}$, exists. We have not yet made a detection of black holes formed via astrophysical processes in this region. With a larger distribution such as those we have discussed, detections in this alleged mass gap may be evident. A large number of detections in this region would disallow a mass gap's existence there; however, a small number of detections in this region could indicate the existence of new physics to form black holes of such a mass. Having the tools to analyze larger populations of BBH will enable us to also analyze possible detections in this region.

Because we are considering future, larger populations of BBH, it is also interesting to examine population properties. In this case, we plan to examine the mass distribution, as well as spin and redshift. By examining such distributions, we are able to deduce information about the nature of BBH. We stand to learn much about mechanisms of formation, including formation scenarios and formation rates.

III. METHODS

We will ultimately use the method of Bayesian inference for this project. The first step in this method is to construct a posterior distribution, given by:

$$p(\theta|d) = \frac{\mathcal{L}(d|\theta)\pi(\theta)}{\mathcal{Z}}. \tag{1}$$

The posterior distribution is the probability density of θ, which describes the parameters of the model, given the strain data from the detectors, d. $\mathcal{L}(d|\theta)$ represents the likelihood function of the strain data given the parameters of the model, $\pi(\theta)$ represents the prior distribution of the model parameters, and $\mathcal{Z}$ is the normalization factor, also called the evidence:

$$\mathcal{Z} \equiv \int d\theta \mathcal{L}(d|\theta)\pi(\theta) \tag{2}$$

We can thus define evidences for both the signal and the noise:

$$\mathcal{Z}_{signal} \equiv \int d\theta \mathcal{L}(d|\theta)\pi(\theta), \tag{3}$$

$$\mathcal{Z}_{noise} \equiv \mathcal{L}(d|n). \tag{4}$$

We can then define the Bayes factor, or the ratio of evidence, for signal and noise to be:

$$BF^S_N = \frac{\mathcal{Z}_{signal}}{\mathcal{Z}_{noise}}. \tag{5}$$

We can use this technique to determine preferred models, as well. For example, given two models 1 and 2, we can produce a Bayes factor:

$$BF^1_2 = \frac{\mathcal{Z}_1}{\mathcal{Z}_2}. \tag{6}$$

It is important to note that this Bayes factor is a generalized approximation. We can thus compare different models and determine which produces the best fit for our actual data[3].

Another key quantity to determine in this process is the volume-time ⟨VT⟩, which describes the area of space-time in which we are observing[4]:

$$\langle VT \rangle = \frac{4}{3}\pi D_{obs}^3 t. \tag{7}$$

We can analyze a synthetic population of many BBH, whose underlying astrophysical distribution is known, by estimating ⟨VT⟩. During our analysis of such synthetic data, we first ask if we are able to recover the known, simulated GW signals. Signals are generally detected via a technique known as matched filtering, in which raw data is correlated against model waveforms calculated based on certain expected parameters[5]. We will use a somewhat similar approach to detect the simulated signals. We then compare our theoretical ⟨VT⟩ to the observed ⟨VT⟩ recovered from our simulations. Given confirmation by this analysis that we will be able to identify large populations of BBH using our currently implemented methods, we may then proceed to verifying what tools must be developed to analyze such populations.

We will first consider an ideal situation in which statistical error within our measurements due to detector noise (for individual events) and finite Poissonian statistics (for ensembles of many events) does not exist. Although this is not at all the reality, we will assume this in an initial attempt to uncover and eliminate any systematic errors; rather, we will debug the program we are creating. When all system error has been sufficiently eliminated, we will then consider a realistic situation in which statistical error within our measurements does, in fact, exist.

This error is most apparent in the observed component masses of the binaries that we detect. When a GW is detected, it is assumed that there is an inverse relationship between the signal frequency and the component masses; rather, if frequency is lower, the component masses are greater. The universe is consistently expanding, and thus the events that we observe experience cosmological redshift (z). As such, our observed mass for any given binary would measure higher than that same binary's actual mass:

$$(1+z) = \frac{M_{obs}}{M_{actual}}. \tag{8}$$

How are we able to take this into account and make the subsequent corrections? We can estimate the distance of a binary by its "loudness," or how significant its signal is. We assume that the loudness is inversely proportional to distance (closer events result in higher loudness). By comparing the loudness of any given event to predictions of the waveform amplitude from general relativity, we are able to estimate its distance. While binaries appear in many orientations, there are two extremes: face-on or edge-on. If the binary is face-on, it is circularly polarized, and if it is edge-on, it is linearly polarized[6]; in between, they are elliptically polarized. When we have determined the orientation, we can then compare the binary to the standard siren and determine its distance. From its distance, we are able to determine its cosmological redshift, and thus can correct the observed values for the component masses to their actual values.

Our ultimate objective, however, is to also consider that formation rates and the mass distribution itself are functions of cosmological redshift. At different points in the history of the universe, formation rates and the mass distribution was different due to a difference in redshift. Thus, we would expect a current population's mass distribution to be different than the population from any earlier point in history.

IV. PROJECT SCHEDULE

Throughout the ten weeks this project will be ongoing at Caltech, we aim to accomplish the goals outlined above. We must first confirm that our current matched filter analysis method is able to recover the aforementioned populations of BBH. This should take no longer than a week. We then must begin construction of the programs that will analyze the population data. Testing the programs will be the main goal of the remaining nine weeks. During weeks two through three, developing the software to analyze the population data will be our focus. Throughout weeks four through six, we can eliminate all system and method error from our program. Thus, during weeks seven through nine, we may begin to observe what statistical error is present, and determine methods to correct this. During the final week, we will produce a final product of our program, as well as a final report and presentation detailing our results.

V. CONCLUSION

The ultimate goal of this project is to infer the underlying mass, spin, and redshift distribution of BBH systems using a large number of such detections. From these inferences, we may learn about the formation, evolution, and final fate of massive stars in our universe.

As we are not yet equipped to make hundreds or thousands of BBH detections, we must consider whether we will have tools developed, debugged, and ready to use when such a time comes. Overall, we aim to build such tools via detailed Bayesian inference to simulate a future, larger population of BBH. We aim to first identify and eliminate systematic errors in our methodology using high-statistics simulations, and further our analysis with more realistic statistical errors. We finally will analyze the fit of our model to the simulated data and determine how well we can constrain our population model parameters, such as m_{min} and m_{max}.

[1] B. Abbott, R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. Adhikari, V. Adya, C. Affeldt, *et al.*, "Gwtc-1: a gravitational-wave transient catalog of compact binary mergers observed by ligo and virgo during the first and second observing runs," *arXiv preprint arXiv:1811.12907*, 2018.

[2] B. Abbott, R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. Adhikari, V. Adya, C. Affeldt, *et al.*, "Binary black hole population properties inferred from the first and second observing runs of advanced ligo and advanced virgo," *arXiv preprint arXiv:1811.12940*, 2019.

[3] E. Thrane and C. Talbot, "An introduction to bayesian inference in gravitational-wave astronomy: Parameter estimation, model selection, and hierarchical models," *Publications of the Astronomical Society of Australia*, vol. 36, 2019.

[4] R. Magee, A.-S. Deutsch, P. McClincy, C. Hanna, C. Horst, D. Meacher, C. Messick, S. Shandera, and M. Wade, "Methods for the detection of gravitational waves from subsolar mass ultracompact binaries," *Physical Review D*, vol. 98, no. 10, p. 103024, 2018.

[5] C. Messick, K. Blackburn, P. Brady, P. Brockill, K. Cannon, R. Cariou, S. Caudill, S. J. Chamberlin, J. D. Creighton, R. Everett, *et al.*, "Analysis framework for the prompt discovery of compact binary mergers in gravitational-wave data," *Physical Review D*, vol. 95, no. 4, p. 042001, 2017.

[6] B. P. Abbott, R. Abbott, T. Abbott, M. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, *et al.*, "Properties of the binary black hole merger gw150914," *Physical Review Letters*, vol. 116, no. 24, p. 241102, 2016.