Draft version July 13, 2022
Typeset using LaTeX **twocolumn** style in AASTeX631

Modeling the Mass Distribution of Binary Black Hole Mergers with GWTC-3
LIGO Surf 2022 Interim Report 1

Jingyi Zhang,[1] Supervised by: Alan J. Weinstein,[2] and Jacob, Golomb[2]

[1] *Smith College*
[2] *LIGO Laboratory, California Institute of Technology*

1. MOTIVATION: BLACK HOLE POPULATION IN THE UNIVERSE

With roughly 70 binary black hole merger (BBHs) events detected by the Advanced LIGO and Advanced Virgo Scientific Collaboration, it is possible to the infer the overall character of black hole population in the universe. Based on the third Gravitational-wave Transient Catalog (GWTC-3), The LIGO Scientific Collaboration et al. (2021) looked into the mass and spin distribution, overall merger rate, as well as the cosmological evolution of binary black holes. These information provides valuable information on high mass star formation and history of stellar evolution, as well as how compact objects fit in the evolution of the universe.

The BBHs population has been fitted with models and described with the associated population hyper-parameters. For example, Figure 1 shows the estimated primary mass distribution of BBHs using the currently most commonly accepted *POWER LAW + PEAK* model (PP) model.

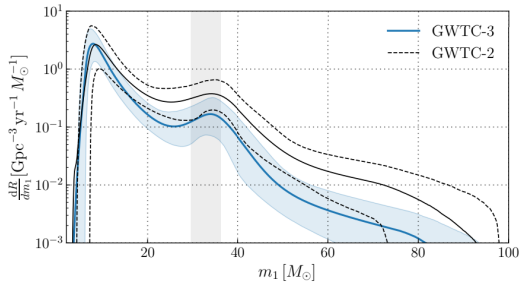

Figure 1. Probability density function of the primary mass distribution for the fiducial Power-Peak (PP) model. (For more on the model, please see Talbot & Thrane (2018)) The solid lines show the posterior population distribution (PPD) and the region shows the 90 percent credible interval of GWTC-3 (blue) and GWTC-2 (black), respectively. As shown in the figure, GWTC-3 suggests that the primary mass has more prominent peaks than we previously observed.(The LIGO Scientific Collaboration et al. 2021)

Aiming to evaluate the parameter estimation's performance, compare different black hole population models, discuss effects of outliers, and improve the general modeling process; we want to conduct goodness-of-fit tests on each step of the hyper-parameter estimation. In addition to supplement the GWTC-3 estimations, the test can also help to better understand the expected polulation from the upcoming forth observation run (O4.) For this project, I focus on the mass distribution of BBHs.

2. METHODS: STRAIN, EVENTS, AND POPULATION

We obtained the estimation of hyper-parameters with two steps. First, mass, spin, redshift, inclination, etc. of the individual BBHs are estimated from the strain data from the interferometer. The individual BBHs parameters are then used to estimate the astrophysical distribution of the population. The two-step approach, also known as the hierarchical Bayesian inference (Talbot et al. 2019), is described below.

2.1. *Single event parameters*

15 parameters describes an individual BBHs events: 8 intrinsic parameters that characterized the spin and mass of the two black holes, 7 extrinsic parameters that describe the binary's position and orbit as seen by us. Packages such as BILBY (Ashton et al. 2019) and LALInference (Veitch et al. 2015) use bayesian inference to estimate the parameters from the strain data. (Figure 2.) The result is given in the form of posterior samples. Each sample contain a possible combination of the parameters that may create the obseved signal. The distribution of parameters in the posterior sample represent the probability distribution of each parameters including how they may be correlated, as shown in Figure 3.

One thing to note is that the estimated parameters are in the detector frame, the properties we see on earth. However, many BBHs events happened far away enough that we must take cosmology into consideration. We can directly estimate the luminosity distance from the strain data. With the assumption of current date cosmological

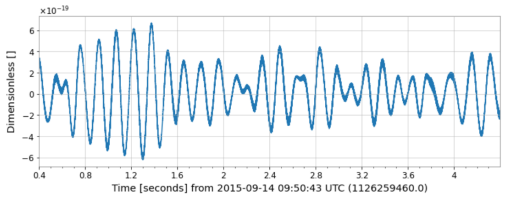

Figure 2. The strain data of BBHs event GW150914 at the LIGO Hanford observatory. The strain contains 4 seconds of data around the event. (Abbott et al. 2021)

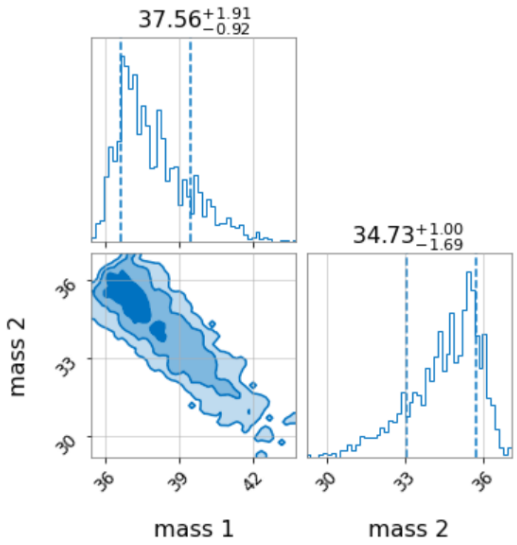

Figure 3. The posterior distribution of primary mass (m1), and secondary mass (m2) of GW150914 and there correlation. In this corner plot, all other parameters are marginalized. The dotted line marked the 90 percent credibility interval——90 percent of the posterior sample is in this mass range.

parameters, we can calculate the redshift of the event. (Hogg 1999) The redshift can then be used to calculate the source frame parameters of the events. The source frame parameter is independent of the observer and reflect the astrophysical information that we would like to use.

2.2. *Population Hyper-Parameters*

Similar to single event parameter estimation where we take in observed strain data as the likelihood to obtain the posterior distribution, (In this case, as we don't know much about the event, we use uninformative prior in the bayesian inference.) we use "data" of the individual events to obtain posterior distribution of the population hyper-parameter. Here, the "data" is the parameter estimation of all events.

We first define the hyper-parameters we want to estimate with out model. For instance, the most important ones of the *POWER LAW + PEAK* model are the index of the power law, the lower mass cutoff, the position of the peak, the width of the peak, and the contribution of the peak in the entire distribution. Several other hyper-parameters describe the detailed shape of the distribution, but they may not be as important.

With the hyper-parameters defined, we can construct the likelihood for the bayesian inference. For each BBHs, we can calculate the probability of getting such events under a defined population model (the model is characterized with hyper-parameters). Given that all events are independent, the probability of getting all the events—-the likelihood–is the product of all individual probability.

$$\mathcal{L}(d|\Lambda) = \prod_i^N P(\theta_i|\Lambda) \tag{1}$$

Where Λ is the hyper-parameter and θ_i is the parameter of the i^{th} event.

However, we do not have the exact value of the parameter describing each single events. We can only infer their probability distribution from the strain data.

$$\mathcal{L}(d|\Lambda) = \prod_i^N \int d\theta_i \mathcal{L}(d_i|\theta_i) p(\theta_i|\Lambda) \tag{2}$$

Here the $\mathcal{L}(d_i|\theta_i)$ term is the probability of getting θ_i for the i^{th} event given the strain data. We integrate over all θ_i that we can get from the strain data to get the total probability, hence the $\int d\theta_i \mathcal{L}(d_i|\theta_i)$ term. In reality, we do not have a smooth probability distribution, but rather discrete samples that represent the events. The number of samples represent the probability (similar to how histogram is used to describe probability density.) Instead of reintegration, we add up the probability of each sample to construct our hyper-parameter likelihood.

$$\mathcal{L}(d|\Lambda) = \prod_i^N \sum_{\theta_i \sim \mathcal{L}(d_i|\theta_i)} p(\theta_i|\Lambda) \tag{3}$$

Finally, like all other astronomical observation, gravitational detection is subjected to selection effect: more massive black holes that create stronger signals are easier to detect. We correct this bias by including a "detection probability" term.

$$\mathcal{L}(d|\Lambda) \propto \frac{\prod_i^N \int d\theta_i \mathcal{L}(d_i|\theta_i) p(\theta_i|\Lambda)}{p_{\rm det}(\Lambda)^N} \tag{4}$$

$$p_{\rm det}(\Lambda) = \int p(\theta|\Lambda) p_{\rm det}(\theta) d\theta \tag{5}$$

The "detection probability" term include the probability of detecting each parameter and the probability of detection such parameter under certain model.

With the likelihood defined, we pass it to python package GWPOPULATION to define do the bayesian inference.

3. PROJECT: GOODNESS-OF-FIT

For my specific project, we want to conduct multiple goodness-of-fit test to test the accountability of the analysis process as well as better understand the effect of data and models in the analysis.

1. Single and Multiple Parameter Goodness-of-Fit Test We want to test how well the GWTC-3 (The LIGO Scientific Collaboration et al. 2021) population estimation agrees with the predicted distribution. We may use the model to generate events arbitrarily, and see how well we can recover the model through the analysis process, similar to the conparision we conduct between the GWTC-2 and GWTC-3 estimation in Figure 4. (The LIGO Scientific Collaboration et al. 2021) We may conduct χ^2 test or Kolmogorov–Smirnov test (ks-test) to compare the resulting posterior distribution with the model.

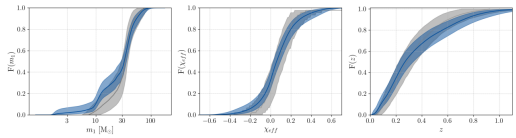

Figure 4. The empirical cumulative density function of primary mass, effective inspiral spin, and redshift. The blue region represent the distribution from GWTC-3, while the grey represent the distribution from GWTC-2. The solid lines show the median while the shaded area are 90 percent credible interval. (The LIGO Scientific Collaboration et al. 2021) For this project, we hope to conduct similar comparison between the observed GWTC-3 population and the theoretical modeled population.

2. Leave-One-Out Test In the analysis of the GWTC-3 population, outliers with extreme mass values were left out because they significantly changed the posterior result and does not agree with the prediction generated by the others. This leave us to wonder whether other events that are included in the analysis would also significantly change the result, and what are the impact of individual events. We may test this by run the estimation many times with one events excluded for each run. Noted that though some events are marked as "outliers", they still bear astrophysical significance.

3. Model Comparison The currently most commonly used *POWER LAW + PEAK* (Talbot & Thrane 2018) model consist of two parts: a power law motivated by the initial mass function (IMF) of stellar formation, where higher mass stars are much rarer than the low mass ones, and a peak motivated by the pulsational pair-instability supernovae (PPSN) graveyard, where stars over a certain threshold would experience explosions and leave remnants of similar masses after the supernovae. However, this is not the only model that describe the binary black hole population. Other parametric model describe the lower and higher mass cutoff differently, and hierarchical mergers may create additional peaks in the mass distribution. Other models attempt to describe the distribution with non-parametric models, where the distribution is almost solely rely on the observed data. We want to see how the posterior distribution change when assuming different models. By comparing the Bayes factor, We want to investigate that whether other models can also yield nicely-fitted result and whether we should consider them for later analysis.

REFERENCES

Abbott, R., Abbott, T. D., Abraham, S., et al. 2021, SoftwareX, 13, 100658, doi: 10.1016/j.softx.2021.100658

Ashton, G., Hübner, M., Lasky, P. D., et al. 2019, ApJS, 241, 27, doi: 10.3847/1538-4365/ab06fc

Hogg, D. W. 1999, arXiv e-prints, astro. https://arxiv.org/abs/astro-ph/9905116

Talbot, C., Smith, R., Thrane, E., & Poole, G. B. 2019, PhRvD, 100, 043030, doi: 10.1103/PhysRevD.100.043030

Talbot, C., & Thrane, E. 2018, ApJ, 856, 173, doi: 10.3847/1538-4357/aab34c

The LIGO Scientific Collaboration, the Virgo Collaboration, the KAGRA Collaboration, et al. 2021, arXiv e-prints, arXiv:2111.03634. https://arxiv.org/abs/2111.03634

Veitch, J., Raymond, V., Farr, B., et al. 2015, PhRvD, 91, 042003, doi: 10.1103/PhysRevD.91.042003