Draft version August 3, 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*

ABSTRACT

With 70 binary black hole merger (BBHs) events detected by the LIGO -Virgo Collaboration, it is possible to infer the overall character of the black hole population in the universe. Specifically, the mass distribution of BBHs provides us with valuable information on stellar evolution and binary formation channel. We here aim to test the current estimation of BBHs population mass distribution based on the third Gravitational-wave Transient Catalog (GWTC-3). The project involves: (1) Examine agreement between the observed data and the fitted mode by conducting goodness-of-fit tests; (2) identifying outliers and examining the impact of non-conventional events with leave-one tests on population parameter estimation; (3) adjusting model to better characterize astrophysical phenomena and describe observed result; (4) compare different theoretical models that may describe the BBHs distribution by fitting observed data to multiple models.

1. 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 the merger rate 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 most popular *POWER LAW + PEAK* model (PP) model.

Aiming to evaluate the population inference'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. STRAIN, EVENTS, AND POPULATION

We obtained the estimation of hyper-parameters with two steps. First, the intrinsic parameters such and mass and spin and extrinsic parameters such as redshift and inclination 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.

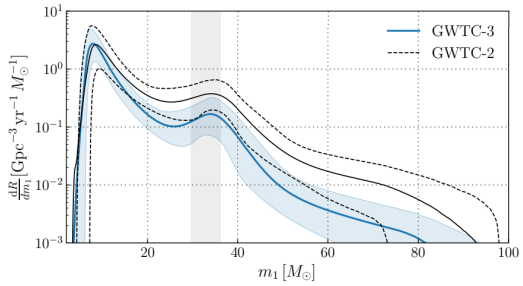

Figure 1. Probability density function of the primary mass distribution for the fiducial Power-Peak (PP) model. (For more on the model, please see Talbot & Thrane (2018)) The solid lines show the posterior population distribution (PPD) and the region shows the 90 percent credible interval of GWTC-3 (blue) 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)

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.

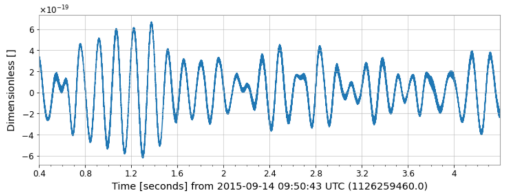

Figure 2. The strain data of BBHs event GW150914 at the LIGO Hanford observatory. The strain contains 4 seconds of data around the event. (Abbott et al. 2021a)

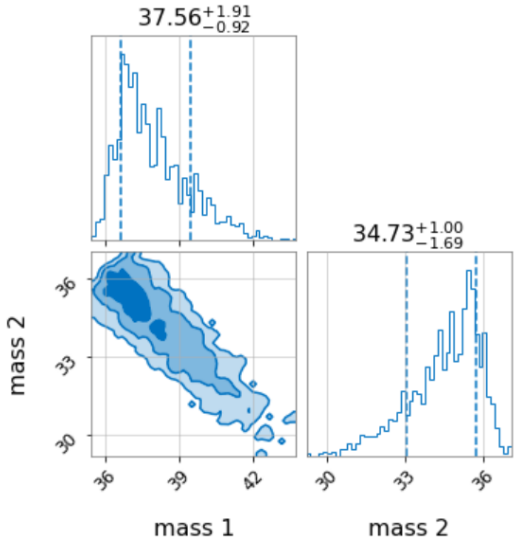

Figure 3. The posterior distribution of primary mass (m1), and secondary mass (m2) of GW150914 and their correlation. In this corner plot, all other parameters are marginalized. The dotted line marked the 90 percent credibility interval——90 percent of the posterior sample is in this mass range.

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 parameters from Planck 2015, 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. (Talbot et al. 2019)

We first define the hyper-parameters we want to estimate with our model. For instance, the most important ones of the *POWER LAW + PEAK* model are the index of the power law α, the lower mass cutoff m_{min}, the position of the peak μ_m, the width of the peak σ_m, and the contribution of the peak in the entire distribution λ_{peak}. 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 hyperparameters 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, also referred to as the likelihood. 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 relative number of samples per bin represent the probability (similar

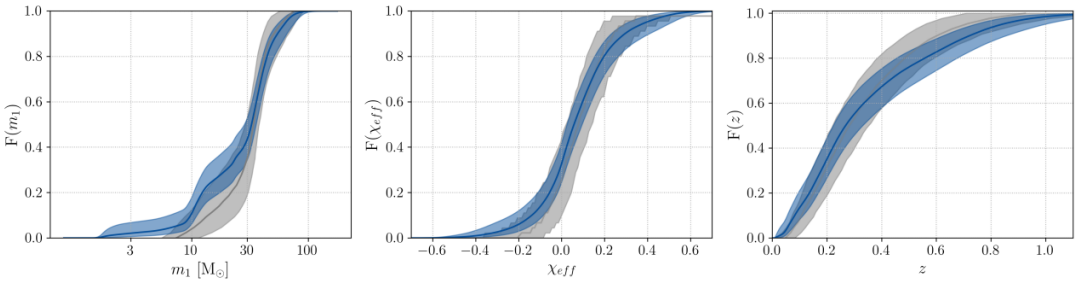

Figure 4. The empirical cumulative density function of primary mass, effective inspiral spin, and redshift. The blue region represent the distribution from GWTC-3, while the grey represent the distribution from GWTC-2. The solid lines show the median while the shaded area are 90 percent credible interval. (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.

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.

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

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

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. As we are interested in the astrophysical properties of black hole population, we correct this bias by including a "detection probability" term.

$$\mathcal{L}(d|\Lambda) \propto \frac{\prod_i^N \int d\theta_i \mathcal{L}(d_i|\theta_i) p(\theta_i|\Lambda)}{p_{\rm det}(\Lambda)^N} \tag{5}$$

$$p_{\rm det}(\Lambda) = \int p(\theta|\Lambda) p_{\rm det}(\theta) d\theta \tag{6}$$

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 do the bayesian inference.

3. TESTING THE MASS DISTRIBUTION ESTIMATION OF BBHS WITH GWTC-3 DATA

3.1. *Outliers and Leave-one-out Tests*

In previous studies, two BBHs events (GW190814 and GW190917) are identified as outliers as they have very low secondary mass. The rest of the population can not predict the existence of these two black holes, as

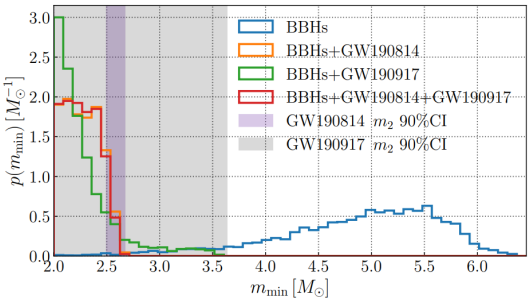

Figure 5. The posterior distribution of the minimum black hole mass inferred with the PP model. Posterior including and excluding GW190817 and GW190917 are shown. Inclusion of either event significantly impacted the distribution. Thus, we conclude that these two events may belong to a separate population that cannot be predicted by the rest of the events. (The LIGO Scientific Collaboration et al. 2021)

shown in Fig 5. The probability with and without these two events have almost no overlaps, and it is very unlikely that there is a value that can predict both scenario. We wonder if any other events have similar constraining power that with or without them in the population inference, the hyper-parameters estimated would not agree.

Here, we conducted leave-one-out test for all 69 events (excluding the previously identified outliers.) By not including one event in each trial and comparing the result with the result obtained from original dataset, we are able to study the impact of each event. Two sample Kolmogorov–Smirnov tests (KS tests) were used to compare the probability distribution of hyper-parameters and quantify the differences. The two sample KS test measure the difference between two population's cumulative distribution function; higher KS test statistic implies that the two sample are more distinct. For all hyper-parameters in the PP model, the events yields the highest statistic value in KS tests are picked out for closer analysis.

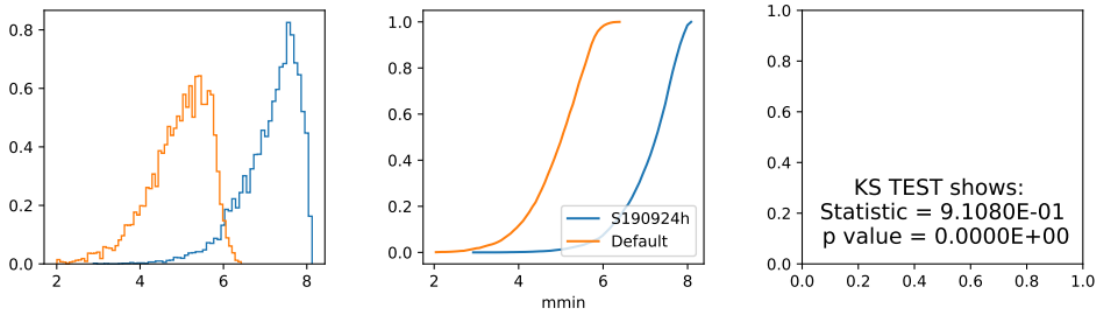

Figure 6. The plot on the left showed the posterior probability distribution of the minimum mass of black holes inferred from the default 69 events population and the population excluding GW190924, respectively. The middle plot shows the cumulative distribution that is used to conduct the KS test, and the KS test result is shown at the right.

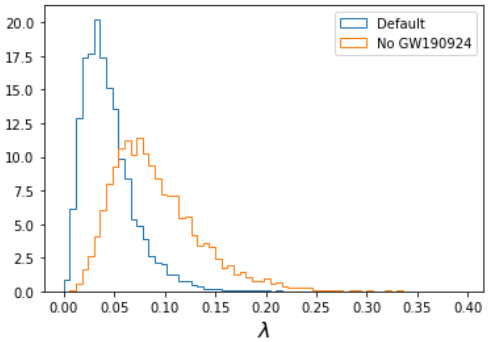

Figure 7. The figure shows posterior probability distribution of λ inferred by the default population and the leave-one-out test without GW190924. The default population predicts a much lower λ where smaller ratio of the total population are predicted to belong to the Gaussian peak.

No new outliers were found as none of the events entirely disagree with the rest of the population. However, several events have strong constraining power on certain hyper-parameters and may be potential source of interests. Future observations may help clarify whether other distinct sub-populations exist. We will discuss these events of interest in the following subsections.

3.1.1. *GW190924: Low Mass*

The secondary black hole of GW190924 has a mass of roughly 5 $M_\odot$, the lowest among all 69 events. Other lowest masses are around 7 $M_\odot$ (including uncertainties.) The posterior probability distribution of the leave-one out test in Figure 6 also shows this potential gap in the minimum mass. The KS test statistic value on minimum mass is over 0.9, much higher than the second highest statistic value of just over 0.1. It is possible that this black hole belongs to a separate population, or be more aligned with the two previously identified outliers. However, though the distributions are quite different, they still have significant overlap and may agree with each other. Thus, we cannot conclude GW190924 s an outlier.

GW190924 also has the highest KS test statistic value on λ, a hyper-parameter in the PP model that characterize the normalized ratio between the black hole population falls into the Gaussian peak and the power law. Without GW190924, the inferred λ centered at around 0.07; with it, the he inferred λ centered at around 0.3. Excluding GW190924, less back holes are following the power law and more black holes are in the Gaussian peak. (Figure 7) This is highly correlated with the minimum mass: As the Gaussian peak is centered around 35 $M_\odot$, the lowest mass population is dominated by the power law. The power law also predicts low mass black holes are much more common than the high mass ones. Together, small decrease on the minimum mass would significantly impact λ as the it allows large portion of the events to be presented at the power law dominated, lowest mass range. Similar to minimum mass, we cannot conclude GW190924 to be an outlier based on λ as the two result overlap significantly.

3.1.2. *GW190521: High Mass*

The primary mass of GW190521 is roughly 85 $M_\odot$, and has been flagged as a possible high-mass outlier in GWTC-2. As shown in Figure 8, the maximum mass predicted with or without GW190521 is centered around 85 $M_\odot$ and 70 $M_\odot$, respectively. The uncertainty of black hole mass on high mass events are often large, as the gravitational wave signal we detected are shorter and have lower frequencies. This is then reflected on the large spread on the probability distribution of maximum mass. The two distribution have significant overlap and we cannot conclude that GW190521 is an outlier. GW190521 also do not have significant impact on other hyper-parameters that defines the population as high mass events are, by nature, very rare in the total population.

It is worth noting that though we do not see GW190521 as an outlier, high mass events still lack an astrophysical explanation. GW190521, along

with GW190602_175927 and GW190519_153544 (Abbott et al. 2021b), have primary mass above 45 $M_\odot$ at a $> 99\%$ confidence interval, above the mass range predicted by the PPSN model. They may involve with hierarchical mergers or other unknown astrophysical process.

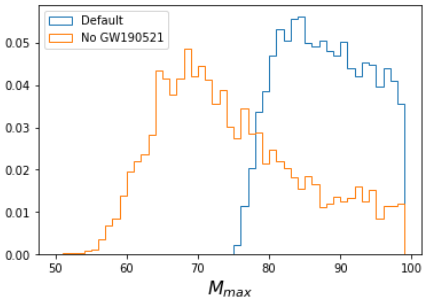

Figure 8. The figure shows posterior probability distribution of maximum mass inferred by the default population and the leave-one-out test without GW190521. The distribution stopped at 100 $M_\odot$ as the it is the highest allowed range of the prior.

3.1.3. *GW190412: Low mass ratio and well-measured tilt angle*

GW190412 is an outstanding events for both its unusually low mass ratio and its well-measured, positive aligned spin.

The two black holes of GW190412 have masses of around 30 $M_\odot$ and 8 $M_\odot$, making the mass ratio below 0.3, the lowest among all events. However, due to the high uncertainty on mass ratio and the scarcity of such unbalanced binaries, GW190412 does not significantly change the estimation of population mass ratio distribution, as shown in Fig 9.

GW190412 also has a relatively high constraining power on the component spin tilts ($cos\theta$) and effective inspiral spin distribution, mostly because its uncertainty is low compare to other events. Here, we used *DEFAULT* model to describe cos_t. The model contains two sub-population: One normal distribution centered at $\theta = 0$ (the spin of individual black hole is in the same direction of the binary's angular momentum), which is motivated by black hole binaries that are formed together; the other one with isotropically oriented tilts that captures the binaries that are formed separately and with no preferred orientation. σ_{spin} is the width of the aligned distribution; a lower σ_{spin} indicate that the population is more concentrated and more binaries are formed together. As shown in Figure 10, the inclusion of GW190412 shift the peak of σ_{spin} distribution to an lower value, and the uncertainty also dropped significantly. With future observation with lower uncertainty, we expect to have better estimation of hyper-parameters related with black hole spin and tilt.

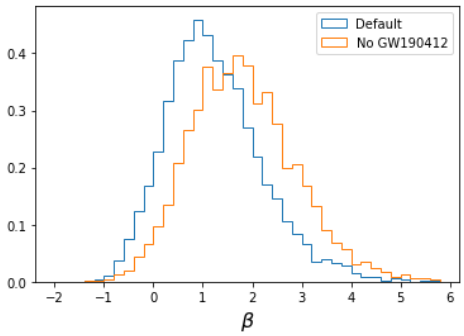

Figure 9. The mass ratio of BBHs primary and secondary mass is modeled with a power law where equal mass binaries are more common. The figure shows posterior probability distribution of mass ratio power law index inferred by the default population and the leave-one-out test without GW190412. The power law index is sightly higher when excluding the unbalanced binary GW190412, but the difference is small.

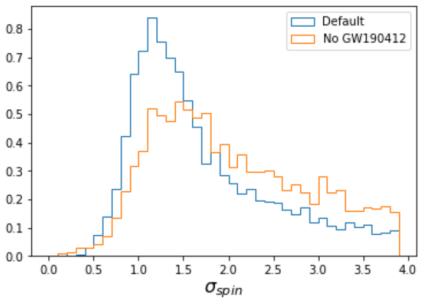

Figure 10. The figure shows posterior probability distribution of σ_{spin} inferred by the default population and the leave-one-out test without GW190412. Without GW190412, the estimation is more widespread and have higher uncertainties.

3.2. *Redshift Evolution of BBHs Merger Rate*

In previous study of BBHs population, a simple power law is used to describe the redshift evolution of BBHs event rate, with more events at higher redshift. With more observation being made and the possibility of future high redshift events, we may be able to incorporate the theory of cosmic star formation rate evolution into our estimation of BBHs event rate.

Madau & Dickinson (2014) suggests that star formation rate does not grows indefinitely with redshift, but rather peaked at $z = 2$ 3. The formation of stellar black holes should obey a similar trends. Fishbach & Kalogera (2021) suggests that for stellar black hole mergers, which LIGO observes, the peak should be at lower redshift as

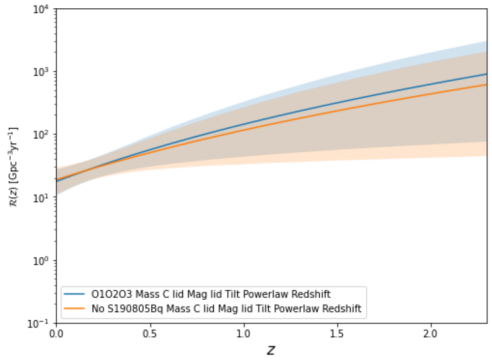

Figure 11. The plot shows how the BBHs event rate change with the redshift. The shaded region marks the 90% confidence interval. The blue region shows the estimation including GW190805, and the orange region shows the estimation with it. The estimation has the lowest uncertainty at z 0.2 where most events happened.

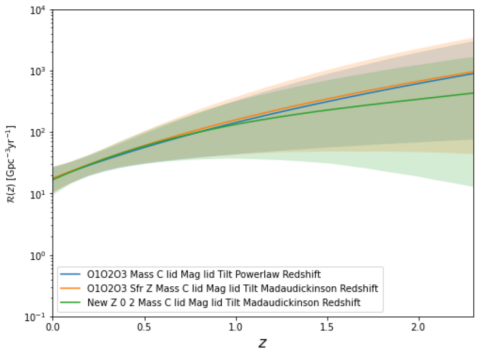

Figure 12. The plot shows BBHs rate evolution with redshift estimated with both simple power law and a broken power law motivated by cosmic stellar formation rate. The shaded area marks the 90% confidence interval. The blue region is predicted with a simple power law. The orange region is predicted by a broken power law with almost no constrain on peak position (at z between 0 to 10.) The uncertainty is slightly higher compare to the blue one due to more parameters presents in the broken power law model. The green region is also predicted by a broken power law, but with a more informative peak position (at z between 0 and 2.) We can see under this scenario, the rate may began to fall at higher redshift where observation is still technically possible.

the delay time between stellar black hole formation and merger is very long.

We cannot estimate the time delay between the stellar formation peak and BBHs event peak with current data, but might be able to better understand the power law index and peak position with future high redshift event. In particular, GW190805 showed us how a single high redshift event may help shape the entire distribution.The events is estimated to be happened at $z > 1$, much further away than most other events. As shown in Figure 11, GW190805 lift up the merger rate at high redshift.

More detailed model and prior selection may also yields better estimation. The current analysis use uninformative prior with large range on the inference of merger rate. In Figure 12, we show that even when keeping a uniform prior, a more informative range can change the merger rate - redshift relation. By only allowing the merger rate peak to happen after the cosmic noon (stellar formation peak,) we can see the slope of rate-redshift relationship become flatter.

REFERENCES

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

—. 2021b, ApJL, 913, L7, doi: 10.3847/2041-8213/abe949

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

Fishbach, M., & Kalogera, V. 2021, ApJL, 914, L30, doi: 10.3847/2041-8213/ac05c4

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

Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415, doi: 10.1146/annurev-astro-081811-125615

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