LIGO SURF Final Report: Estimating the parameters of a population of coalescing compact binaries

Naomi Gendler
Mentors: Larry Price and Vivien Raymond
(Dated: September 26, 2014)

The aLIGO network stands to make hundreds of detections over the lifetime of the project. While there is much to be learned from the parameters of single events, the parameter distribution of the population of events is also of great interest for astrophysics. The goal of this project is to develop the tools for estimating such population distributions and accounting for selection bias in such inferences. We will then apply the method to a simulated population of binary systems of neutron stars in order to estimate their mass distribution.

I. INTRODUCTION/BACKGROUND

aLIGO will have a likely detection rate of 40 events per year, with a lower limit of 0.4 per year and upper limit of 400 per year [1]. While the detections themselves will be a hugely significant discovery, we'd like to be able to also use these detections in order to make inferences about the astrophysical sources, paving a path for gravitational wave astronomy. This project will develop a method for inferring parameters of a population of such sources. We use the example of the mass distribution of binary neutron stars in order to develop the method. We would like to infer the parameters of the population's mass distribution (for example, the mean and standard deviation for a gaussian) by using a Markov-Chain Monte Carlo method to sample parameter space and obtain distributions for each parameter governing the universe-set distribution of a population.

There is evidence to suggest that two different types of supernovae could give rise to a bimodal distribution of neutron star masses [2]. Namely, these supernovae are iron-core collapse and electron capture [3]. All stars begin their lives by converting hydrogen into helium at a decreasing rate until gravity overpowers and the core contracts. In an iron-core collapse supernova, this contraction causes the core to burn elements of increasing mass until the core has produced iron. Up until this point, the star has supported itself by fusing elements into heavier elements. Iron, however, is very stable so the core gives in to the force of gravity and begins to collapse [4]. In the case of

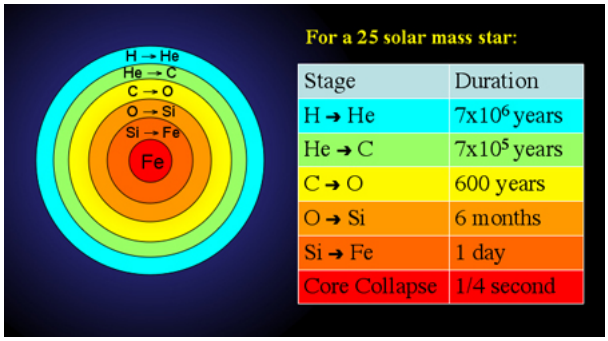


FIG. 1: A visual representation of the stages of element fusion the neutron star goes through during an iron-core collapse supernova. [Image: Swinburne University]

electron-capture supernovae, a white dwarf's core is supported by electron degeneracy pressure caused by the Pauli Exclusion Principle (no two electrons can occupy the same state). However, as the mass of the white dwarf increases, it reaches the Chandrasekhar limit, at which point the core can no longer support itself and collapses [5]. These two supernovae are speculated to give rise to neutron stars of slightly different masses, suggesting that we ought to look for a bimodal distribution of neutron star masses [3], and we would like to be able to develop machinery that is able to construct this. What we would like to do in this project is develop the tools able to identify a possible distribution with peaks at 1.35 $M_\odot$ stars (resulting from iron-core collapse supernovae) and 1.5 $M_\odot$ stars (resulting from electron-capture supernovae) [6].

The detection of gravitational waves will allow us to do this, as gravitational waveforms depend explicitly on the mass (in some form) of the source. We can thus use the waveform of the gravitational wave to figure out what the mass of the binary neutron star is, as demonstrated in Eqs. 3 and 4. The neutron stars coalesce, spinning inward at an increasing rate, much like a figure skater bringing her arms in as she twirls. As the spinning gets faster, the

frequency of the emitted gravitational wave increases as well. This frequency evolution is a quantity determined by the chirp mass, given by:

$$\mathcal{M} = \mu^{3/5} M^{2/5} \tag{1}$$

where $\mu = \frac{m_1 m_2}{m_1 + m_2}$ is the classic reduced mass of the binary system, and M is the total mass of the system. We can

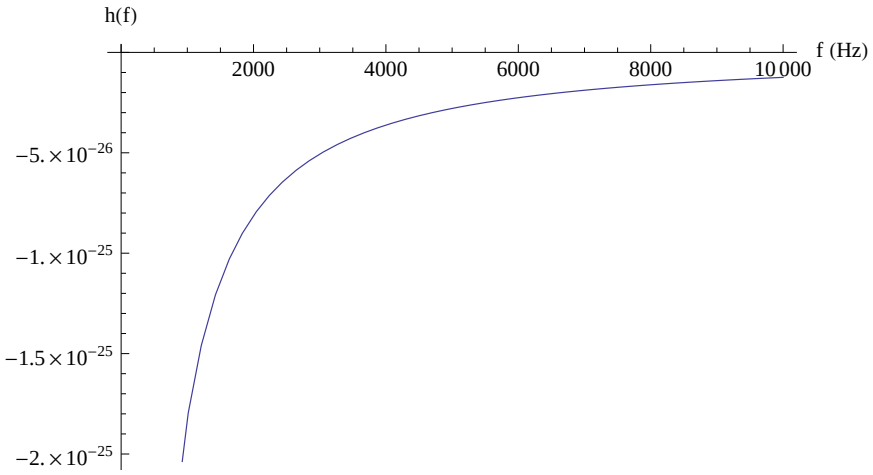


FIG. 2: An inspiral binary neutron star produces a gravitational wave of increasing frequency.

determine the waveform of a gravitational wave caused by an inspiral binary star by computing the metric perturbation from the post-Newtonian expansion of the source [7]. We must take into account the degree to which the gravitational wave strain is induced in a detector. This strain is given by [8]:

$$\tilde{h}(f) = \tilde{h}_+(f) F_+(\theta, \phi, \psi) + \tilde{h}_\times(f) F_\times(\theta, \phi, \psi) \tag{2}$$

where $\tilde{h}_{+,\times}$ are the plus and cross polarizations, and $F_{+,\times}$ are the antenna responses to the incident signal, given in terms of the altitude, θ, the azimuth, ϕ, and the polarization, ψ [9]:

$$F_+ = -\cos(\theta)\sin(2\psi)\sin(2\phi) - \frac{1}{2}\cos^2(\theta)\cos(2\psi)\cos(2\phi) - \frac{1}{2}\cos(2\psi)\cos(2\phi),$$

and

$$F_\times = \frac{1}{2}\cos^2(\theta)\sin(2\psi)\cos(2\phi) + \frac{1}{2}\sin(2\psi)\cos(2\phi) - \cos(\theta)\sin(2\phi)\cos(2\psi)$$

The waveform is:

$$\tilde{h}(f) = \left(\frac{1\text{Mpc}}{D_{\text{eff}}}\right) \mathcal{A}_{1\text{Mpc}}(M, \mu) f^{-7/6} \mathrm{e}^{-i\Psi(f; M, \mu)} \tag{3}$$

where

$$\mathcal{A}_{1\text{Mpc}} = -\left(\frac{5}{24\pi}\right)^{1/2} \left(\frac{GM_\odot/c^2}{1\text{Mpc}}\right) \left(\frac{\pi G M_\odot}{c^3}\right)^{-1/6} \left(\frac{\mathcal{M}}{M_\odot}\right)^{-5/6}, \tag{4}$$

$$\Psi(f; M, \mu) = 2\pi f t_0 - 2\phi_0 - \pi/4 + \frac{3}{128\eta}\left[v^{-5} + \left(\frac{3715}{756} + \frac{55}{9}\eta\right)v^{-3} - 16\pi v^{-2} + \left(\frac{15293365}{508032} + \frac{27145}{504}\eta + \frac{3085}{72}\eta^2\right)v^{-1}\right], \tag{5}$$

$$v = \left(\frac{GM}{c^3}\pi f\right)^{1/3}, \tag{6}$$

and

$$D_{\text{eff}} = D\left[F_+^2\left(\frac{1+\cos^2\iota}{2}\right)^2 + F_\times^2\cos^2(\iota)\right]^{-1/2} \tag{7}$$

Here, D is the distance to the source, ι is the angle between the line connecting the source and the observer and the angular momentum vector, $\mathcal{M}$ is the chirp mass, η is the symmetric mass ratio, given by $\eta = \frac{\mu}{M}$, and the phase at coalescence is ϕ_c. ϕ_0 is the termination phase, given by:

$$\phi_0 = \phi_c - \frac{1}{2}\arctan\left(\frac{F_\times}{F_+}\frac{2\cos(\iota)}{1+\cos^2(\iota)}\right) \tag{8}$$

The parameters named above are classified as extrinsic or intrinsic. An extrinsic parameter is one for which we need not resort to a template bank for comparison [10]. In other words, an intrinsic parameter is one that is the same regardless of the parameters of the detector, whereas an extrinsic parameter is one that changes depending on the observer. The extrinsic parameters here are:

- distance (D)
- sky location (θ, ϕ)
- polarization (ψ)
- coalescence phase (ϕ_0)
- inclination (ι)

and the intrinsic parameters are:

- chirp mass ($\mathcal{M}$)
- symmetric mass ratio (η)

We should note that given a binary system, we may only determine the *chirp mass*, since only the chirp mass and the total mass show up in the signal-to-noise ratio. This is important for actual experiments, since although it would be nice to measure the individual masses, inverting the chirp mass and mass ratio distributions leads to component mass distributions with broad peaks. Finding the chirp mass, on the other hand, results in a sharply peaked distribution. Thus, the chirp mass is more easily found–there is no way to separate out the masses of the individual neutron stars without incurring large error.

II. WORK DONE

A. Theory: Hierarchical Modeling

We would like to obtain the mass distribution of the binary neutron star population given the noisy mass distributions of the individual events. To do this, we follow the technique developed in [11] and [12]. The first goal in this project is to obtain the probability distribution for the masses of binary neutron stars on an event by event basis. We obtain this distribution by applying Bayes theorem for an individual event, say event n out of N events [13]:

$$p(\boldsymbol{\theta_n}|\mathbf{D_n}) = \frac{p(\boldsymbol{\theta}_n)p(\mathbf{D_n}|\boldsymbol{\theta}_n)}{p(\mathbf{D_n})} \tag{9}$$

where $p(\mathbf{D_n}|\boldsymbol{\theta})$ is the likelihood of a set of parameter values given the data set (as opposed to the *probability* of obtaining a data set given a model), given by [11]:

$$p(\mathbf{D_n}|\boldsymbol{\theta}) \propto \exp\left(\langle \mathbf{D_n}|h(\boldsymbol{\theta})\rangle - \frac{1}{2}\langle h(\boldsymbol{\theta})|h(\boldsymbol{\theta})\rangle\right) \tag{10}$$

where h is the waveform model exhibited by the data set, $\mathbf{D_n}$, and $\langle a|b\rangle$ is the inner product of the Fourier transform of the signal which picks out the waveform model from the noise in the data, given by [11]:

$$\langle a|b\rangle = 4\int_0^\infty \frac{\tilde{a}(f)\tilde{b}^*(f)}{|S_n(f)|}df \tag{11}$$

where $S_n(f)$ is the one-sided power spectral density of the noise.

$\boldsymbol{\theta_n}$ is the set of intrinsic and extrinsic parameters spelled out in the Introduction, $p(\mathbf{D_n})$ is the probability of obtaining this particular data set, given by $p(\mathbf{D_n}) = \int d\boldsymbol{\theta_n} p(\mathbf{D_n}|\boldsymbol{\theta_n})$ and $p(\boldsymbol{\theta})$ is our prior–in this case, the information we already have about the probability of measuring a a set of parameter values. Using this data stream, we can construct the waveform of the gravitational wave (accounting for noise). We compare this data to our model with $p(\mathbf{D_n}|\boldsymbol{\theta})$. We take the prior here to be flat, a general, uninformative prior. Finally, we can marginalize our PDF over all but one parameter, mass, to produce a one-dimensional PDF [14].

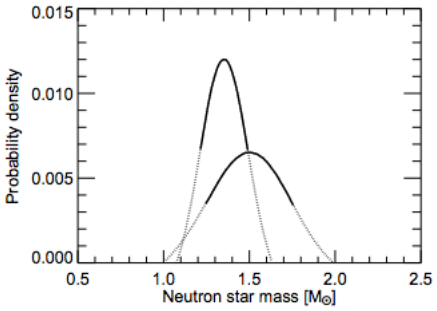


FIG. 3: A possible mass distribution for binary neutron stars found from radio observations of pulsars. [Image: B. Kiziltan, *et. al.*]

Next, we want to use the individual probability distributions in order to construct a so-called distribution of distributions. In order to do this, we apply Bayes theorem once more to a set of observed distributions. Assuming that the events are independent, we arrive at the following equality:

$$p(\{\boldsymbol{\theta_n}\}_{n=1}^N|\{\mathbf{D_n}\}_{n=1}^N) = \prod_{n=1}^N p(\boldsymbol{\theta_n}|\mathbf{D_n}) \tag{12}$$

We would like to find a distribution of some parameter, t, governed by a set of global parameters, $\boldsymbol{\alpha}$. Ultimately, we want an expression for $p(\boldsymbol{\alpha}|\{\mathbf{D_n}\}_{n=1}^N)$. We derive this by first applying Bayes' Theorem:

$$p(\boldsymbol{\alpha}|\{\mathbf{D_n}\}_{n=1}^N) = \frac{p(\boldsymbol{\alpha})p(\{\mathbf{D_n}\}_{n=1}^N|\boldsymbol{\alpha})}{p(\{\mathbf{D_n}\}_{n=1}^N)} \tag{13}$$

Because the events are independent, we can write this expression as a product of individual events and then marginalize over the set of local parameters:

$$p(\boldsymbol{\alpha}|\{\mathbf{D_n}\}_{n=1}^N) = \prod_{n=1}^N \frac{p(\boldsymbol{\alpha})}{p(\mathbf{D_n})}\int p(\mathbf{D_n}|\boldsymbol{\theta_n})p(\boldsymbol{\theta_n}|\boldsymbol{\alpha})d\boldsymbol{\theta_n} \tag{14}$$

Using the assumption that our parameter of interest, t, separates from the other parameters, we arrive at a final analytic expression:

$$p(\boldsymbol{\alpha}|\{\mathbf{D_n}\}_{n=1}^N) = p(\boldsymbol{\alpha})\prod_{n=1}^N \int p(\boldsymbol{\theta}|\mathbf{D_n})\frac{f_\alpha(t_n)}{p(t_n)}\mathrm{d}\boldsymbol{\theta} \tag{15}$$

where $f_\alpha(t) \equiv p(t|\boldsymbol{\alpha})$.

Finally, we approximate this integral by sampling the distribution at K spots, evaluating and summing the function at those points, and dividing by the number of points. This gives us our final, computable expression:

$$p(\boldsymbol{\alpha}|\{\mathbf{D_n}\}_{n=1}^{N}) = p(\boldsymbol{\alpha}) \prod_{n=1}^{N} \frac{1}{K_n} \sum_{k=1}^{K_n} \frac{f_\alpha(t_n^{(k)})}{p(t_n^{(k)})} \tag{16}$$

B. Theory: Selection Bias

A big problem when inferring population parameters from experimental data is selection bias. Selection bias is a classic problem in astrophysics in that observed objects may look dimmer if they are farther away. In the case of LIGO, weak signals are not taken into account. If the signal-to-noise ratio is below some threshold, the data is thrown out. This method does not account for weak or low-amplitude signals that really were there, and so making inferences from such data will be biased. This needs to be added into our above derivation.

We consider the following situation, laid out in [15]: we have a set of independent detection time periods. When a signal above a set threshold is detected the data (a trigger) is recorded as $\mathbf{D_n}$ along with a positive indicator, I^+. When no signal reaches threshold, the data is thrown out but a negative indicator, I^-, is recorded.

When a trigger is recorded, we can write the posterior distribution on $\boldsymbol{\alpha}$ using Bayes Theorem:

$$p(\boldsymbol{\alpha}|\mathbf{D_n}, I^+) = \frac{p(\boldsymbol{\alpha})}{p(\mathbf{D_n})} \sum_{\{\mathcal{H}\}} p(\mathbf{D_n}|\boldsymbol{\alpha}, \mathcal{H}) p(\mathcal{H}|\boldsymbol{\alpha}) \tag{17}$$

where we have marginalized over the set of local models, $\{\mathcal{H}\}$, of which there must be at least one that describes the presence of a signal and one that describes the absence of a signal.

In order to compute the first factor in the sum, we marginalize over local model parameter space, $\Theta_{\mathcal{H}}$:

$$p(\mathbf{D_n}|\boldsymbol{\alpha}, \mathcal{H}) = \int_{\Theta_{\mathcal{H}}} p(\mathbf{D_n}|\boldsymbol{\theta}_{\mathcal{H}}, \boldsymbol{\alpha}, \mathcal{H}) p(\boldsymbol{\theta}_{\mathcal{H}}|\boldsymbol{\alpha}, \mathcal{H}) \mathrm{d}\boldsymbol{\theta}_{\mathcal{H}} \tag{18}$$

In the complementary case, we start out again by applying Bayes Theorem:

$$p(\boldsymbol{\alpha}|I^-) = \frac{p(\boldsymbol{\alpha})}{p(I^-)} \sum_{\{\mathcal{H}\}} p(I^-|\boldsymbol{\alpha}, \mathcal{H}) p(\mathcal{H}|\boldsymbol{\alpha}) \tag{19}$$

Since we have now don't have the actual data, when calculating the first factor in the sum, we must marginalize over not just local parameter space, but also over the missing data up to the threshold:

$$p(I^-|\boldsymbol{\alpha}, \mathcal{H}) = \int_{\Theta_{\mathcal{H}}} \int_{\mathbf{D} \leq \mathbf{D}_{th}} p(\mathbf{D_n}|\boldsymbol{\theta}_{\mathcal{H}}, \boldsymbol{\alpha}, \mathcal{H}) p(\boldsymbol{\theta}_{\mathcal{H}}|\boldsymbol{\alpha}, \mathcal{H}) \mathrm{d}\mathbf{D_n} \mathrm{d}\boldsymbol{\theta}_{\mathcal{H}} \tag{20}$$

The goal now is to determine the probability of our set of global parameters, $\boldsymbol{\alpha}$, given this set of data and indicators. As usual, we apply Bayes Theorem:

$$p(\boldsymbol{\alpha}|\{\mathbf{D_n}\}_{n=1}^{N}, I) = \frac{p(\boldsymbol{\alpha}) p(\{\mathbf{D_n}\}_{n=1}^{N}, I|\boldsymbol{\alpha})}{p(\{\mathbf{D_n}\}_{n=1}^{N}, I)} \tag{21}$$

Using the assumption that each event is independent, we write this as a product. We also note that $p(\{\mathbf{D_n}\}_{n=1}^{N}, I) = p(\{\mathbf{D_n}\}_{n=1}^{N})$:

$$p(\boldsymbol{\alpha}|\{\mathbf{D_n}\}_{n=1}^{N}, I) = \frac{p(\boldsymbol{\alpha})}{\{\mathbf{D_n}\}_{n=1}^{N}} \prod_{i=1}^{N} p(\mathbf{D_i}, I|\boldsymbol{\alpha}) \tag{22}$$

We now can split up this product into events with a trigger (say there are n of them) and ones without:

$$p(\boldsymbol{\alpha}|\{\mathbf{D_n}\}_{n=1}^{N}, I) = \frac{p(\boldsymbol{\alpha})}{\{\mathbf{D_n}\}_{n=1}^{N}} \prod_{j=1}^{n} p(\mathbf{D_j}, I^{+}|\boldsymbol{\alpha}) \prod_{k=1}^{N-n} p(I^{-}|\boldsymbol{\alpha}) \tag{23}$$

Now suppose that there are two models–one that predicts a trigger ($\mathcal{H}^{+}$) and one that predicts no trigger ($\mathcal{H}^{-}$). We can then split up each product into the sums laid out in Eqs. 17 and 19:

$$p(\boldsymbol{\alpha}|\{\mathbf{D_n}\}_{n=1}^{N}, I) = \frac{p(\boldsymbol{\alpha})}{p(\{\mathbf{D_n}\}_{n=1}^{N})} \prod_{j=1}^{n} \left[p(\mathbf{D_j}, I^{+}|\boldsymbol{\alpha}, \mathcal{H}^{+})p(\mathcal{H}^{+}|\boldsymbol{\alpha}) + p(\mathbf{D_j}, I^{+}|\boldsymbol{\alpha}, \mathcal{H}^{-})p(\mathcal{H}^{-}|\boldsymbol{\alpha})\right] \times \prod_{k=1}^{N-n} \left[p(I^{-}|\boldsymbol{\alpha}, \mathcal{H}^{+})p(\mathcal{H}^{+}|\boldsymbol{\alpha}) + p(I^{-}|\boldsymbol{\alpha}, \mathcal{H}^{-})p(\mathcal{H}^{-}|\boldsymbol{\alpha})\right] \tag{24}$$

We now note that the first term in this equality is the posterior distribution derived in Section II A. Thus, we have our final result:

$$p(\boldsymbol{\alpha}|\{\mathbf{D_n}\}_{n=1}^{N}, I) = \frac{p(\boldsymbol{\alpha})}{p(\{\mathbf{D_n}\}_{n=1}^{N})} \prod_{j=1}^{n} \left[\frac{1}{K_j}\sum_{i=1}^{K_j} \frac{f_{\alpha}(t_j^{(k)})}{p(t_j^{(k)})} + p(\mathbf{D_j}, I^{+}|\boldsymbol{\alpha}, \mathcal{H}^{-})p(\mathcal{H}^{-}|\boldsymbol{\alpha})\right] \times \prod_{k=1}^{N-n} \left[p(I^{-}|\boldsymbol{\alpha}, \mathcal{H}^{+})p(\mathcal{H}^{+}|\boldsymbol{\alpha}) + p(I^{-}|\boldsymbol{\alpha}, \mathcal{H}^{-})p(\mathcal{H}^{-}|\boldsymbol{\alpha})\right] \tag{25}$$

C. Toy Model: 1D Population Distributions

My first completed task was using a NumPy module called emcee [16] in order to randomly sample a Gaussian distribution, which simulates collecting data from aLIGO. Each sample was given a random standard deviation, which simulates the noise that would occur in actual measurements. I was then able to reconstruct the distribution, as shown in Fig. 4 using the formalism laid out in this section to compute the likelihoods. We constructed a true population distribution that was a Gaussian with parameters $\mu = 0.4$ and $\sigma = 0.1$. In our calculations, we assumed that the samples contained Gaussian noise and were being drawn from a normal distribution. The module uses a number of walkers in order to sample the distribution in steps and reconstruct the original distribution from an array of samples taken from the "measured" distributions. The results of this are shown in Fig. 4. The implemented method worked exactly as expected, with the true values of μ and σ centered at the estimated distributions.

In order to complexity our toy model to be more realistic, we had to sample from a chirp mass distribution instead of a unimodal Gaussian. Because there is a suggestion that there are two mass-classes of binary neutron stars, we started with a bimodal distribution of the following form:

$$p(m) = \frac{h}{\sqrt{2\pi}\sigma_1} \mathrm{e}^{\frac{(m-\mu_1)^2}{2\sigma_1^2}} + \frac{(1-h)}{\sqrt{2\pi}\sigma_2} \mathrm{e}^{\frac{(m-\mu_2)^2}{2\sigma_2^2}} \tag{26}$$

The graph of this distribution is displayed in Fig. 5

We would like to make a change of variable from this distribution in the individual masses to a probability distribution of the chirp mass. To do this, we first make a change of variable from $p(m_1, m_2)$, where m_1 and m_2 are the individual masses of the binary system, to the distribution $p(\mathcal{M}, \eta)$, where $\mathcal{M}$ is the chirp mass of the system and η is the symmetric mass ratio. In particular, $\mathcal{M} = \frac{(m_1 m_2)^{3/5}}{(m_1+m_2)^{1/5}}$ and $\eta = \frac{m_1 m_2}{(m_1+m_2)^2}$. The necessary Jacobian for this transformation is:

$$\mathcal{J}(\mathcal{M}, \eta; m_1, m_2) = \frac{\mathcal{M}}{\eta^{6/5}\sqrt{1-4\eta}} \tag{27}$$

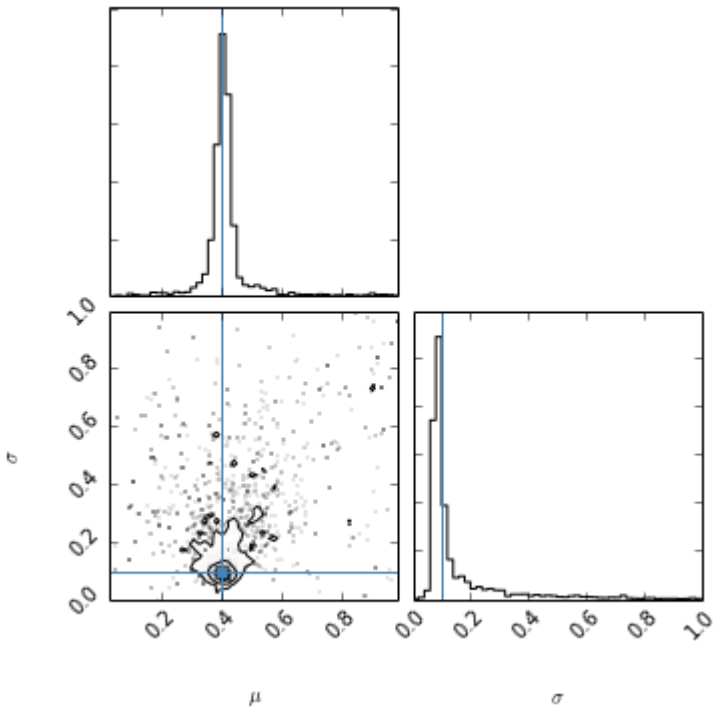


FIG. 4: Reconstructed posterior probability distribution. The true values of μ and σ are shown with blue lines. For $\mu = 0.4$ and $\sigma = 0.1$, the recovered distribution had mean $\bar{\mu} = 0.39$ and standard deviation $\bar{\sigma} = 0.056$. 100 walkers sampled the distribution 1000 times.

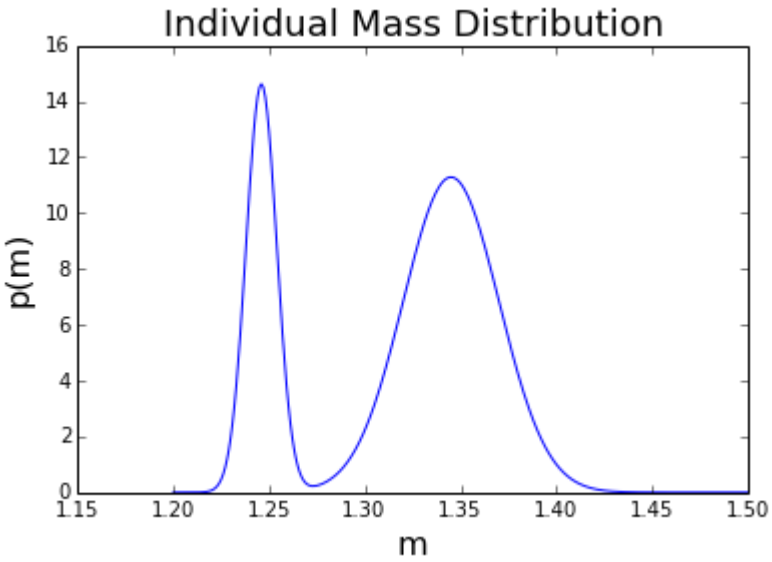


FIG. 5: A possible distribution for the individual masses in the population of binary neutron star systems.

To get an expression for $p(\mathcal{M})$, we integrate over η:

$$p(\mathcal{M}) = \int_0^{0.25} \mathrm{p}(m_1(\mathcal{M},\eta))\mathrm{p}(m_2(\mathcal{M},\eta))\mathcal{J}(\mathcal{M},\eta;m_1,m_2)\mathrm{d}\eta \tag{28}$$

where the limits of integration go from 0 to 0.25 because η is at a maximum when $m_1 = m_2$, which translates into

a value of 0.25 for η. This integral cannot be done analytically, but by iterating the integral over many values of $\mathcal{M}$, ranging from 0 to 3 $M_\odot$ and plotting the values we arrive at the distribution showcased in Fig. 6.

It is worth it to point out how the three peaks of the chirp mass distribution arise–the leftmost peak occurs by combining the left peak of the individual mass distribution with itself in the chirp mass equation. Similarly, the rightmost peak in the chirp mass occurs by combining the right peak of the individual mass distribution with itself. Finally, the middle peak arises from a combination of the right and left peaks of the individual mass distribution. The red lines in the figure represent the values of the chirp mass with the three combinations of μ_1 and μ_2 as inputs.

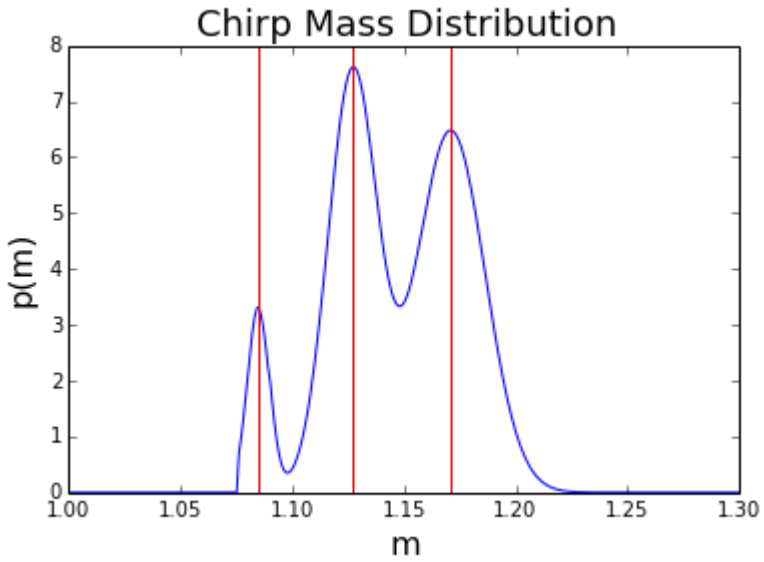


FIG. 6: The chirp mass distribution calculated via a change of variable on individual masses drawn from two bimodal Gaussian distributions.

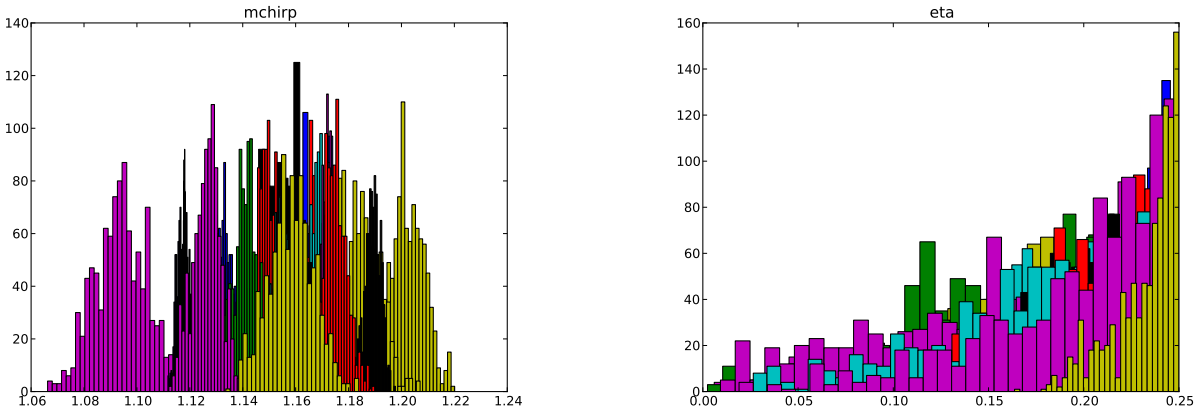
The next step is to use the technique outlined in Sec. II A in order to recreate this distribution. This time, we generated samples from the 2-dimensional distribution in chirp mass and symmetric mass ratio and gave them random standard deviations to simulate noise in the data. The samples we obtained are shown in Fig. 7. We then used emcee's sampler in order to recover the five parameters that govern the original 2-dimensional distribution (μ_1, μ_2, σ_1, σ_2, and h). Because this distribution has multiple peaks and is not well-known, the sampler used previously could not sample the function adequately. Thus, we used the Parallel Tempering sampler. This sampler initializes the walkers at different "temperatures." Each temperature of walkers explores an altered likelihood that allows them to explore multi-modal distributions more easily. The results of this sampling are displayed in Fig. 8

There are several things to notice about these results. First of all, though the distributions of μ_1, μ_2, σ_1, and σ_2 are doubly peaked, the higher peak is centered roughly at the true values. This means that if we were to obtain results like this from a real set of detections, our "best guess" would be fairly accurate. Another feature to see is that the distribution of h seems to be pushed up against the upper limit of our prior, which is unusual. Again, though, the other peak of the distribution is centered roughly at the true value of h.

D. Toy Model: Selection Bias

The next piece of the puzzle was to implement the technique laid out in Sec. II B. To estimate the parameters of a population accounting for selection bias, we use the emcee module once again, this time altering our likelihood with the added factors from Eq. 25. We consider a scenario in which we are trying to estimate the parameters of a gaussian mass distribution, as well as the rate of actual signals, and we measure the signal-to-noise ratio given to us by a detector. Below a threshold SNR, $\rho_{\rm th}$, the mass measurements and SNR are not recorded. We call detections with SNRs greater than the threshold "triggers." We assume that under the signal model, the SNR is drawn from a non-central chi-squared distribution with a non-centrality parameter given by the optimal SNR [15]:

$$\rho^2_{\rm opt} = \left(\frac{1Mpc}{M_\odot}\right)^2 \mathcal{A}^2_{\rm 1Mpc} \int_{20}^{1500} \frac{f^{-7/3}}{S_h(f)} {\rm d}f \tag{29}$$



(a) Samples drawn in $\mathcal{M}$

(b) Samples drawn in η

FIG. 7: Samples drawn from the 2-dimensional original distribution

where $S_h(f)$ is the power-spectral noise distribution, shown in Fig. 9.

The necessary factors for the likelihood are given by:

$$p(\mathbf{D_j}, I^+|\boldsymbol{\alpha}, \mathcal{H}^+) = \int_0^{d_L^{\max}} \mathrm{d}d_L \frac{3d_L^2}{2\Delta m^2 (d_L^{\max})^3} \exp\left\{-\frac{1}{2}\left(\rho_j^2 + \rho_{opt}^2(\boldsymbol{m_j}, d_L)\right)\right\} \times I_0(\sqrt{\rho_j}\rho_{\mathrm{opt}}(\boldsymbol{m_j}, d_L)) \times \exp\left(-\frac{(\mu - m_1)^2 + (\mu - m_2)^2}{2\sigma^2}\right) \quad (30)$$

where I_0 is a modified Bessel function of the first kind. A derivation of this expression can be found in [17]. This expression represents the probability that a trigger occurs given that there was an actual signal.

$$p(\mathbf{D_j}, I^+|\boldsymbol{\alpha}, \mathcal{H}^-) = \frac{1}{2\Delta m^2} \exp(-\rho_j^2/2) \quad (31)$$

This represents the likelihood that a trigger occurs given that there was no real symbol.

$$p(I^-|\boldsymbol{\alpha}, \mathcal{H}^+) = \int_0^{\rho_{th}^2} \mathrm{d}x \int_{\mathcal{S}_m^2} \mathrm{d}\boldsymbol{m} \int_0^{d_L^{\max}} \mathrm{d}d_L \frac{3d_L^2}{2\Delta m^2 (d_L^{\max})^3} \exp\left(-\frac{1}{2}\left[x + \rho_{\mathrm{opt}}^2(\boldsymbol{m}, d_L)\right]\right) \times I_0\left(\sqrt{x}\rho_{\mathrm{opt}}(\boldsymbol{m}, d_L)\right) \times \exp\left(-\frac{(\mu_m - m_1)^2 + (\mu_m - m_2)^2}{2\sigma_m^2}\right) \quad (32)$$

where x represents the unknown true SNR-squared which we marginalize out. This equation represents the probability that a trigger does not occur given that in fact there was a signal.

$$p(I^-|\boldsymbol{\alpha}, \mathcal{H}^-) = \int_0^{\rho_{th}} \exp\left(-x/2\right) \mathrm{d}x \quad (33)$$

Finally, this expression represents the probability that there was no trigger given that there was not a signal.

We assume a uniform distribution of binary neutron stars in volume out to a maximum distance of 600 Mpc. We generate mass samples from a gaussian of mean $1.2M_\odot$ and standard deviation $0.1M_\odot$, along with concurrently drawing SNRs and luminosity distances. Of these events, we use a rate of $4.18 \times 10^{-7}\mathrm{Mpc}^{-3}\mathrm{yr}^{-1}$ to assign the correct number of these events to be signals. For each set of masses and distances, the drawn SNR tells us whether a trigger was recorded by the detector. We can then use the appropriate masses in our likelihood and run emcee. The outcome of this sampling is shown in Fig. 10.

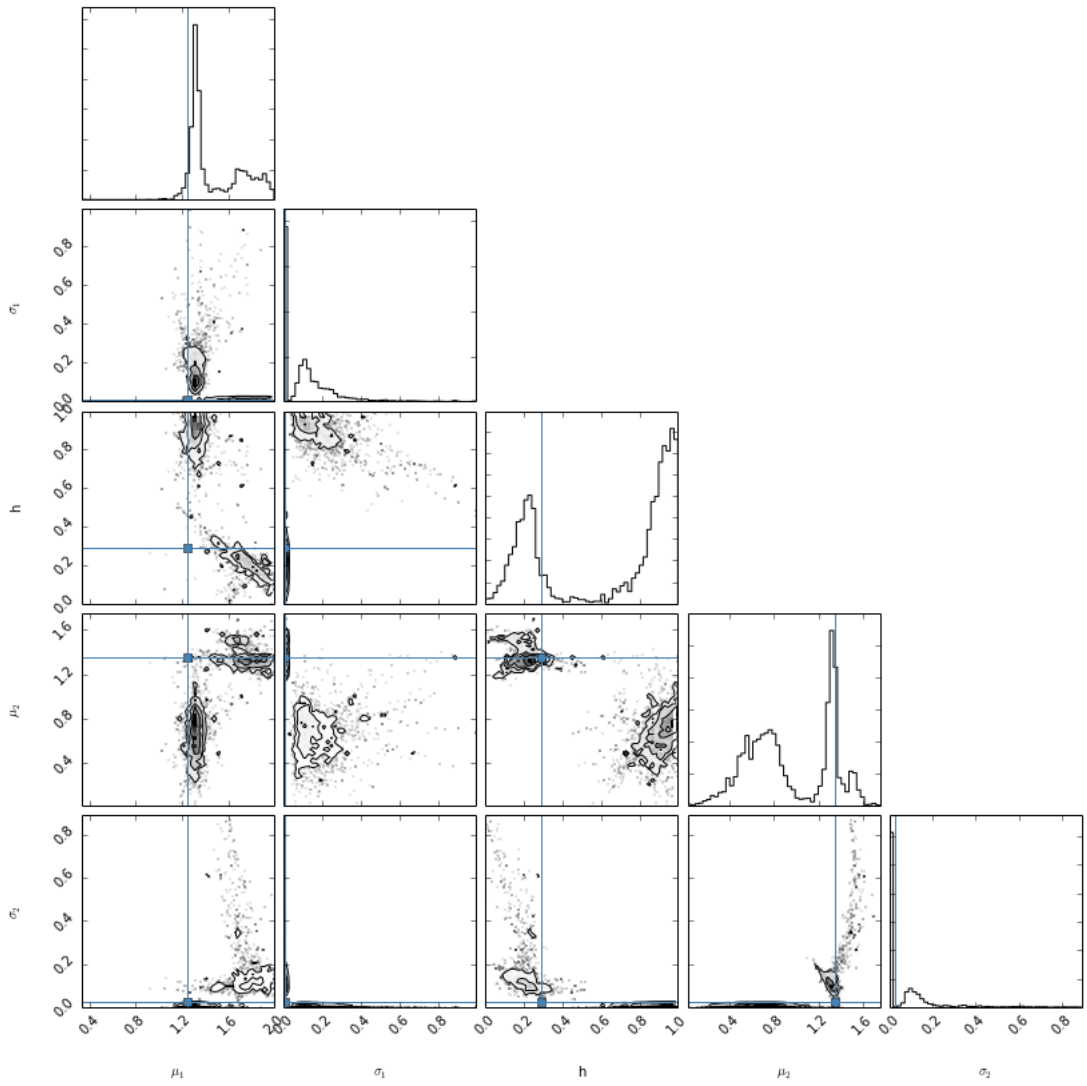


FIG. 8: This parameter estimation was calculated using 100 walkers at 20 different temperatures. The true parameters for the original chirp mass distribution were $\mu_1 = 1.246, \mu_2 = 1.345, \sigma_1 = 0.008, \sigma_2 = 0.025$, and $h = 0.293$.

E. Full Chirp Mass Distribution Estimation

Finally, we are ready to put together all previous parts in order to create a technique that takes a set of data from aLIGO and runs it through a parameter estimation pipeline, taking into account selection bias effects. We start with a set of mass measurements, each measurement its own distribution due to noise in the detector. We draw these samples from a 2-dimensional distribution in chirp mass and symmetric mass ratio, just as we did in Section II C. We then use the same technique outlined in Section II D to estimate the five parameters of the original distribution, as well as the rate of events. Again, we use the Parallel Tempering method, as our distribution is once again multi-modal.

Unfortunately, at this time there are a few technical problems in the code we developed by the end of the summer. The complete program has been written up in Python–however, as it stands there are errors concerning array matching due to the "noise" in the measurements. However, we have the full framework laid out for a full pipeline, enabling us to do accurate parameter estimations on realistic populations of aLIGO sources.

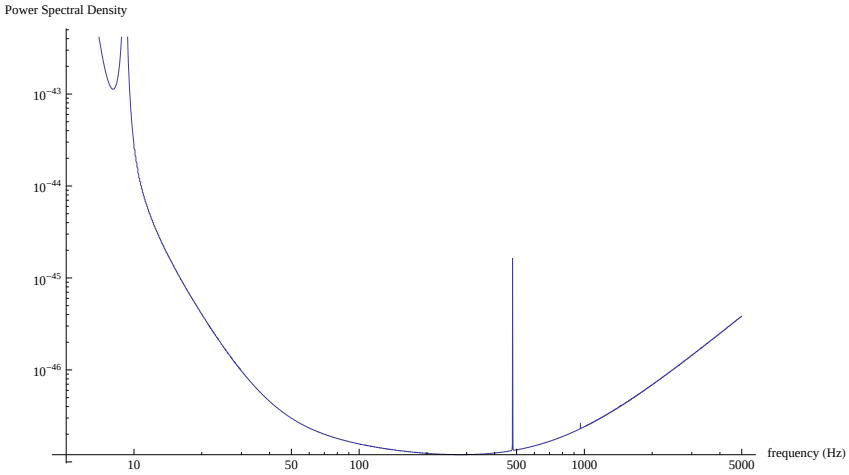


FIG. 9: The power spectral density noise curve for LIGO.

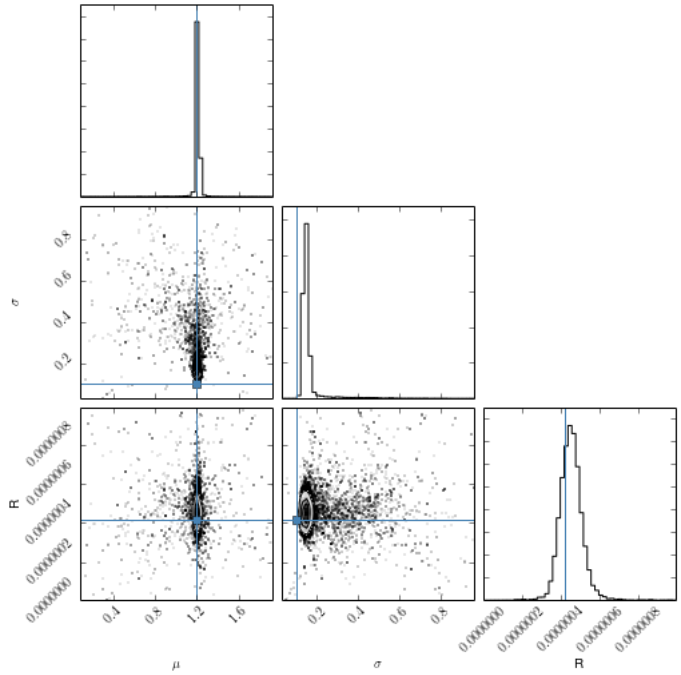


FIG. 10: This parameter estimation was done using 525600 total events, 90 signals, 63 false dismissals, and 27 false triggers. We used 500 walkers with a burn-in period of 100 steps. For true values of $\mu = 1.2M_{\odot}$, $\sigma = 0.1M_{\odot}$, and a rate of $R = 4.18 \times 10^{-7}$ we recovered estimated values of $\bar{\mu} = 1.2M_{\odot}$, $\bar{\sigma} = 0.14M_{\odot}$, and $\bar{R} = 4.5 \times 10^{-7}$

III. FUTURE WORK

This project provides numerous opportunities for future work. First of all, we would like to calculate the number of events we would need to see in order to make accurate estimations of the population parameters. We would also like to do a full 10-parameter estimation for not only chirp mass, but also symmetric mass ratio, sky location, and the other parameters listed in section I. Ultimately, we would like to generate some simulated data and implement the entire method from start to finish. Another important piece of future work is to compare the evidences, which here we've ignored, in order to perform a Bayesian model comparison to select the most accurate model for the chirp mass distribution. Finally, once aLIGO starts making detections we can use this method to do gravitational-wave astronomy. The technique outlined here can be used to explore other population parameters, leading to greater

knowledge about our universe.

[1] LIGO Scientific Collaboration. Predictions for the rates of compact binary coalescences observable by ground-based graviational-wave detectors. *Class. Quantum Grav.*, 27(173001), 2010.
[2] J. Schwab, Ph. Podsiadlowski, and S. Rappaport. Further evidence for the bimodal distribution of neutron-star masses. *ApJ*, 719(1), 2010.
[3] C. Knigge, M.J. Coe, and Ph. Podsiadlowski. Two populations of x-ray pulsars produced by two types of supernovae. *Nature*, 2011.
[4] NASA. Introduction to supernova remnants.
[5] C.L. Fryer and K.C.B. New. Gravitational waves from gravitational collapse. *Living Reviews in Relativity*, 14(1), 2011.
[6] B. Kiziltan, A. Kottas, and S. Thorsett. The neutron star mass distribution. *ApJ*, 2010.
[7] J.D.E. Creighton and W.G. Anderson. *Gravitational-Wave Physics and Astronomy.* Wiley-VCH, 2011.
[8] B. Allen, W.G. Anderson, P.R. Brady, D.A. Brown, and J.D.E. Creighton. Findchirp: an algorithm for detection of gravitational waves from inspiraling compact binaries. *Phys. Rev. D.*, 85(122006), 2012.
[9] W.G. Anderson, P.R. Brady, J.D.E. Creighton, and E.E. Flanagan. Beam pattern response functions and times of arrival for earthbound interferometer. *Phys. Rev. D.*, 63(042003), 2001.
[10] B.J. Owen and B.S. Sathyaprakash. Matched filtering of gravitational waves from inspiraling compact binaries: Computational cost and template placement. *Phys. Rev. D.*, 60(022002), 1998.
[11] I. Mandel. Parameter estimation on gravitational waves from multiple coalescing binaries. *Phys. Rev. D.*, 81(084029), April 2010.
[12] D.W. Hogg, A.D. Myers, and J. Bovy. Inferring the eccentricity distribution. *ApJ*, 725, 2010.
[13] A. Gelman, J.B. Carlin, H.S. Stern, and D.B. Rubin. *Bayesian Data Analysis.* Chapman, 2 edition, 2004.
[14] D.W. Hogg. Data analysis recipes: Probability calculus for inference. 2012.
[15] C. Messenger and J. Veitch. Avoiding selection bias in gravitational wave astronomy. *New J. Phys*, 15(053027), may 2013.
[16] D. Foreman-Mackey, D.W. Hogg, D. Lang, and J. Goodman. emcee: The mcmc hammer. *pasp*, 125:306–312, mar 2013.
[17] J. Veitch and W. Del Pozzo. Analytic marginalisation of phase parameter. *LIGO DCC*, 2013.

IV. ACKNOWLEDGMENTS

I'd like to extend thanks to my wonderful mentors, Larry Price and Vivien Raymond for all their patience and brilliance. I'd also like to thank Alan Weinstein and everyone at LIGO for giving me the opportunity to do this research this summer. Finally, I'd like to thank the National Science Foundation for the financial support.