Estimating parameters for gravitational waves signal

Thomas Uriot, s1216723@ed.ac.uk

June 28, 2021

Abstract

A Bayesian approach is taken in order to infer on the parameters describing the signal detected by LIGO. A Monte Carlo Markov Chain (MCMC) is used to sample the posterior distributions of the parameters. The 95% credible intervals are computed for each parameter and the posterior distributions are plotted. It is found that the intervals are very precise compared to the ranges used in the prior. Finally, it is argued that the MCMC run is of quality and that the distributions are well sampled.

Keywords: MCMC, gravitational waves; Bayes formula; waveform model; sampling

1 Introduction

Background

Gravitational waves are 'ripples' in spacetime caused by some of the most violent and energetic processes in the Universe. In other words, massive accelerating objects (such as neutron stars or black holes orbiting each other) disrupt space-time in such a way that 'waves' of distorted space radiate from the source (much like the movement of waves created by throwing a stone in the water). The existence of gravitational waves were first predicted by Albert Einstein in 1916 in his general theory of relativity. The first proof of the existence of gravitational waves was made in 1974 through observing a binary pulsar - two extremely dense and heavy stars in orbit around each other [1]. However, one had to wait until September 14, 2015 at 09:50:45 UTC, in order to sense the first 'physical' distortions of spacetime caused by passing gravitational waves. These gravitational waves were generated by two colliding black holes nearly 1.3 billion light years away. The waves were detected using the LIGO interferometers (detectors). The general structure of the LIGO detector is shown in Appendix A. The way LIGO detects gravitational waves is by measuring the time the laser beams take to come back to the source, in both arms. Since the arms are of same length (4 km), it takes the same time for the laser beams to reach the source. As mentioned previously, gravitational waves distort spacetime and would thus cause to alter the arm length. The extra time that the laser beams take to reach the source is then measured. Effectively, the laser beams do not only travel 4 km back and forth, but 1120 km, as the beams are reflected about 280 times. To be sure that the change in arm length is not caused by some noise specific to the detector, two LIGO detectors are operational in the cities of Hanford and Livingston, USA, separated by thousands of miles. The signals (data) detected by the LIGO instruments are used in this paper in order to make inference on the parameters by which the signals are affected.

Relevant Literature

P. Ajith (2008) describes how to numerically compute accurate waveforms from the merger stage between two black holes using a non-spinning phenomenological waveform model. Since

[1]https://www.ligo.caltech.edu/page/what-are-gw

the signal detected by the LIGO instruments which are analyzed in B.P. Abbott et al. (LIGO Scientific Collaboration and Virgo Collaboration, 2016), comes from the collision of two black holes, the waveform model described by P. Ajith is used here. The likelihood function for the data observed by the ground-based laser interferometers is constructed using the former waveform model with its set of parameters. Veitch, J. et al. (2015) is closely related to our paper as it makes use of Monte Carlo Markov Chain (MCMC) to estimate parameters for compact binaries.

Aims and Plan

In this paper, a Bayesian approach is used in order to infer on the parameters describing the signal detected by LIGO. This will enable us to update our knowledge about the parameters of interest as well as gaining a better understanding of black holes merges (by estimating the final merged mass, for instance). A Monte Carlo Markov Chain (MCMC) will be used to sample the posterior distributions in order to infer on the parameters.

Firstly, the data and the theory will be introduced. This will be done by giving an overview of the Bayesian approach applied to our context as well as briefly describing the Metropolis Hastings Algorithm used in the MCMC. Then, several diagnostics will be performed to assess the quality of the results from the MCMC run. Finally, 95% credible intervals for each of the parameters will be computed and the posterior distributions will be plotted.

2 Data and Theory

2.1 Bayes formula

The international BHH (Black Hole Hunter) collaboration has detected a signal in their gravitational wave detectors (Arnor and Gondor). The data collected from the two detectors contain the following: a list of frequencies at which the data is measured, common to both detectors, the real and imaginary parts of the detector response in the Fourier domain, for the two detectors.

The Bayes formula is given by

$$P(\Theta|y, M) = \frac{P(y|\Theta, M)P(\Theta|M)}{P(y, M)} \propto P(y|\Theta, M)P(\Theta|M) \tag{1}$$

where Θ represent the vector of parameters of some model M and y is the data (observations) collected. In the Bayesian setting, $P(\Theta|y, M)$ is called the posterior distribution, while $P(y|\Theta, M)$ is the likelihood and $P(\Theta|M)$ is the prior distribution. The term $P(y, M)$ in (1) which only depends on the data y can be omitted since it does not add any information to the parameters. The aim is to compute the posterior distribution $P(\Theta|y, M)$ and use it to infer on the parameters.

In our context, the log of the prior distributions for the parameters are chosen to be uniformely distributed. The ranges of the priors are taken to represent realistic assumptions about the signal, from previous knowledge or observational constraints. The vector of parameters Θ contains 6 parameters: (θ, ϕ) - the sky position of the source relative to axes centred at the Arnor detector, t_0 - the time (in s) at which the signal was observed in the Arnor's detector; (M, η) - the total mass (in solar masses) and symmetric mass ratio of the binary system, $M = m_1 + m_2$, $\eta = m_1 m_2/M^2$, where m_1 and m_2 are the component masses;

and d, the distance to the source in Gpc.

As mentioned in the introduction, the likelihood function for the data collected, is constructed using the waveform model described by P. Ajith (2008) with its set of parameters. The log-likelihood function for the two detectors is proportional to

$$(d_1|h_1) - \frac{1}{2}(h_1|h_1) + (d_2|h_2) - \frac{1}{2}(h_2|h_2), \tag{2}$$

where d_1 and d_2 are the data collected from the Arnor and Gondor detectors, respectively. Similarly, the waveform model corresponding to the Arnor detector is given by h_1 and to the Gondor detector by h_2. The waveform model returns a real and imaginary part for each of the frequencies at which the data is measured. In addition, one has

$$(d_1|h_1) \approx 4 \cdot df \cdot \sum_{i=1}^{N_f} \frac{\mathrm{Re}(d_1[i])\mathrm{Re}(h_1[i]) + \mathrm{Im}(d_1[i])\mathrm{Im}(h_1[i])}{S_n(f_i)}, \tag{3}$$

where N_f is the total number of measurements (frequencies), $\mathrm{Re}(d_1[i])$ is the real part of the Arnor detector response measured at the i_{th} frequency and $\mathrm{Re}(h_1[i])$ is the real part of the waveform model evaluated at the i_{th} frequency, for the Arnor detector. Similarly, for the Gondor detector. Furthermore, $S_n(f_i)$ is the noise spectral density, which is the same for both detectors and given by

$$S_n(f_i) = 10^{-49}\left[\left(\frac{f_i}{215}\right)^{-4.14} - 5\left(\frac{f_i}{215}\right)^2 + 111\left(\frac{2 - 2\left(\frac{f_i}{215}\right)^2 + \left(\frac{f_i}{215}\right)^4}{2 + \left(\frac{f_i}{215}\right)^2}\right)\right], \tag{4}$$

where f_i is the i_{th} frequency. As mentioned previously, the aim is to integrate over the log-posterior distribution (which is simply the log-likelihood added to the log-prior) in order to obtain the marginal distributions for each of the parameters. In order to compute the integral, one can sample the posterior distribution a large number of times in order to approximate the integral by a sum over the sample. Here, the sampling of the parameter space is based on the MCMC technique which is briefly described next.

2.2 MCMC

To effectively generate a sequence of samples, one can use a Markov chain with stationary distribution equal to the target distribution (i.e the posterior distribution in this case). The Markov chain must satisfy the detailed balance equation

$$P(\Theta)P(\Theta, \Theta') = P(\Theta')P(\Theta', \Theta), \tag{5}$$

where

$$P(\Theta_i = \Theta'|\Theta_{i-1} = \Theta). \tag{6}$$

Here, $P(\Theta)$ is the posterior distribution, and $\Theta = \{\theta, \phi, t_0, M, \eta, d\}$.

In this paper, The Metropolis-Hastings algorithm is used to compute a Markov chain with these properties. The algorithm consists of choosing a starting point, then, at step i:

1. Propose a new point, Θ', by drawing from a proposal distribution $q(\Theta', \Theta_i)$. Here, the proposal distribution is taken to be Gaussian.

2. Evaluate the target distribution at the new point. Compute the Metropolis-Hastings ratio

$$H = \frac{P(\Theta')q(\Theta_i, \Theta')}{P(\Theta_i)q(\Theta', \Theta_i)}.$$

3. Draw a random sample, α, from a uniform U[0,1] distribution. If $\alpha < H$, set $\Theta_{i+1} = \Theta'$, otherwise set $\Theta_{i+1} = \Theta_i$. In words, if $H < \alpha$, the move to the new point Θ' is not accepted and the next point Θ_{i+1} is the same as the current point Θ_i. In which case, one has to repeat step 1.

The way new points are proposed is important since if the new proposed point is too extreme, the numerator will be very small compared to the denominator, in the ratio H, leading to rejection. Generally, the optimal acceptance rate is taken to be around 25%. In addition to this rate, there are other criterion used to assess the performance (quality) of a MCMC run, that will be used in section 3. The MCMC run and the various criterion will be computed using R.

MCMC diagnostics

Graphical criterion include whether the points sampled are actually sampling the region of interest of the posterior distribution. This can be examined by looking at the trace plots and checking whether the 'burn-in' phase is finished and whether the chain is moving unidirectionally. Related to the previous point, one must check whether the posterior distributions are well sampled, for all the parameters. If it is the case, then the graphs of the distributions should be smooth. In other words, the sampled points must represent the distribution as a whole. This may not always be the case as the samples are autocorrelated. A large number of iterations (samples) is necessary in order to get independent samples. The measure used to assess the autocorrelation is the effective sample size, in general, it is good to aim to have an effective sample size of 1000. Lastly, one can run multiple MCMC chains starting at different points and checking whether they converge to the same value. This is checked using the Gelman-Rubin convergence diagnostic, where one is looking for the result of the test to be smaller than 1.1.

3 Results

3.1 Posterior Distributions and Inference

Figure 1 below shows the posterior distributions (histograms) for all of the 6 parameters, with their associated trace plots. The one-dimensional posteriors are smooth and thus well sampled. The trace plots do not contain the burn-in phase and fluctuate unidirectionally around the mean. Furthermore, the acceptance rate for the chain is 33% which is fairly close to 25%. Therefore, it is sensible to use the values contained in the chain to give a 95% credible intervals.

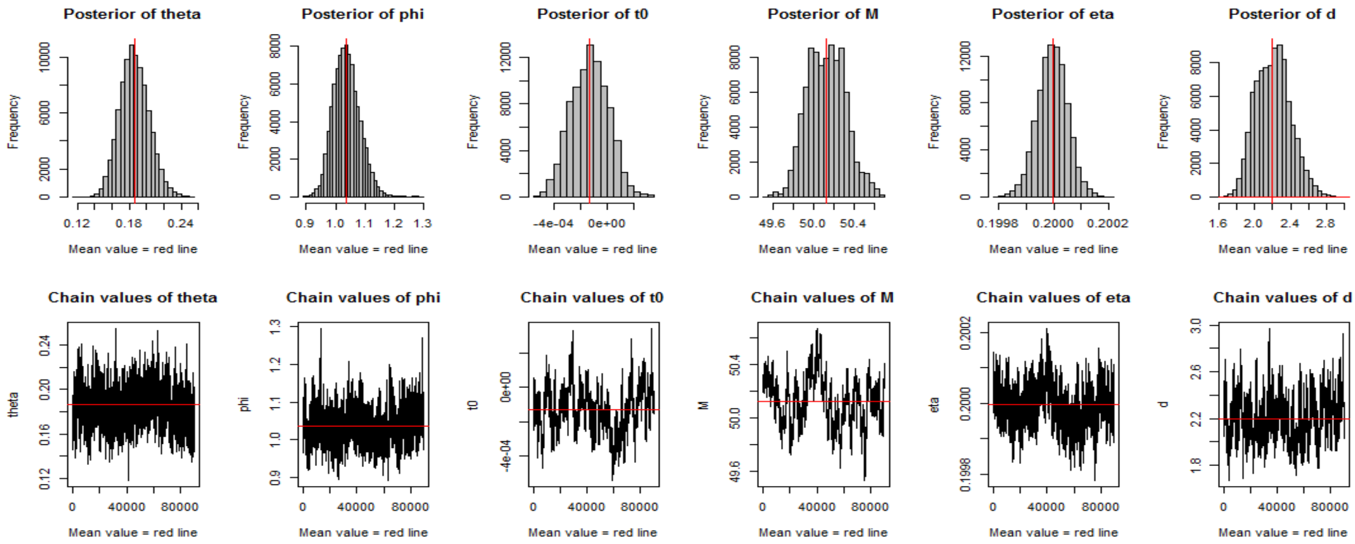

Figure 1: One-dimensional posterior distributions and trace plots for the 6 parameters, for 100,000 iterations.

Figure 2 below shows 3 two-dimensional plots: η with M, ϕ with θ, and t_0 with d. These parameters are chosen to be paired together as they represent either angles, masses, or time and distance, which may be correlated.

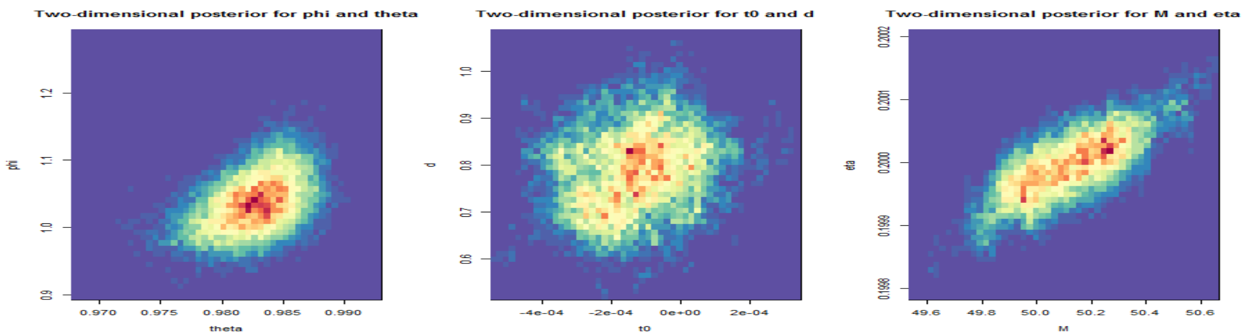

Figure 2: Two-dimensional posterior distributions for chosen parameters.

From Figure 2, one can see that there does not seem to be any obvious correlation between ϕ and θ and between t_0 and d. However, it is clear that M and η are correlated since the distribution is diagonally aligned. In other words, as M increases (decreases), η increases (decreases). This is expected, as the parameters are both determined by the component masses m_1 and m_2. As a reminder

$$M = m_1 + m_2, \qquad \eta = \frac{m_1 m_2}{M^2}. \tag{7}$$

It is also of interest to compute the posterior distributions for the individual component masses, as well as a two-dimensional distribution. To do so, one must express m_1 and m_2 in terms of M and η. My derivation is done in Appendix B. Again, the distributions are smooth the trace plots fluctuate unidirectionally around the mean. As expected, m_1 and m_2 are very correlated.

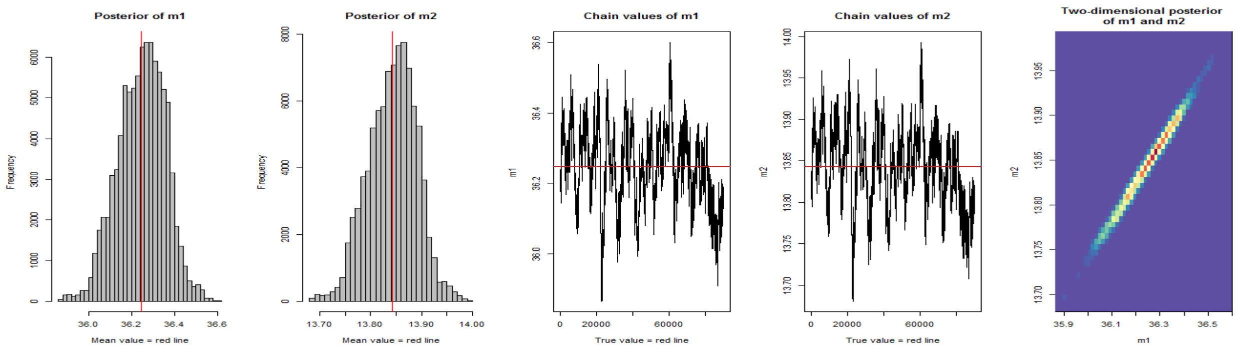

Figure 3: One-dimensional posterior distributions, trace plots and two-dimensional posterior distribution for the component masses.

My findings for the point estimates (here, the mean is used), and the 95% credible intervals of the parameters are summarized in Table 1 below. As interesting feature to note is that although the prior ranges for M, θ and ϕ are large, the resulting 95% credible intervals for the parameters are very precise. In other words, the new knowledge about the parameters is very much improved after incorporating the likelihood function.

Parameters	θ	ϕ	t_0	M
PE	0.18	1.04	-0.00013	50.09
95% CI	[0.15 0.22]	[0.966 1.132]	[-0.00041 0.00015]	[49.78 50.37]
Parameters	η	d	m_1	m_2
PE	0.1999	2.83	36.25	13.84
95% CI	[0.1998 0.2001]	[2.603 3.086]	[36.03 36.45]	[13.75 13.93]

Table 1: Point estimates (PE) and 95% credible intervals (CI) of the parameters.

3.2 MCMC performance

The acceptance rate and the trace plots indicate that the MCMC run performed well and indeed converged. The estimated effective sample size for the thin chain is 4141, which is more than 1000. This means that there is no autocorrelation for 4141 sampled points. The Gelman-Rubin convergence test gives a multivariate potential scale-reduction factor of 1.11. It is thus right above the generally optimal bound of 1.1 (as explained in MCMC diagnostics section). Therefore, one may want to run the chain for a bit longer (150 000 iterations should give the sought after improvement).

4 Conclusion

The Bayesian framework and the Monte Carlo Markov Chain method to sample the posterior distributions for the parameters were introduced. The 6 parameters characterizing the signal detected by the LIGO detectors and the component masses m_1 and m_2 have been estimated. I have found that, the 95% credible intervals for the parameters are very precise in the sense that they are much more narrow than the ranges that were assumed in the prior distributions. This is due to the information added by the likelihood function built from the waveform model described by P. Ajith. It was shown that the MCMC run gave good results, and that the posterior distributions were well sampled, although the scale-reduction factor was 1.11. In further work, it would be of interest to repeat the analysis using different waveform models to construct the likelihood function. One could also check that frequentist confidence intervals under correct prior assumptions agree with the credible intervals found.

Appendix A

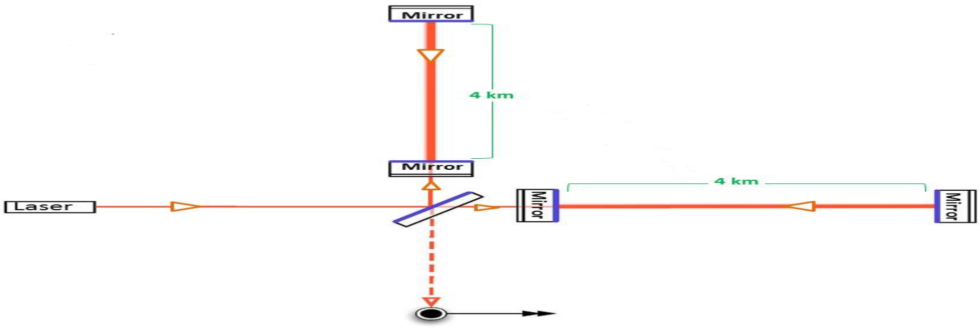

Figure 4: LIGO detector.

Appendix B

The basic relations are

$$M = m_1 + m_2, \quad \eta = \frac{m_1 m_2}{M^2}. \tag{8}$$

Solving for m_2 in the η equation and then substituting m_2 in the M equation, one gets

$$M = m_1 + \frac{\eta M^2}{m_1}$$

which is equivalent to

$$m_1^2 - M m_1 + \eta M^2 = 0. \tag{9}$$

Remembering the relationships in (8), one can use Vieta's formula for quadratic equations and concludes that m_1 and m_2 are roots to (9), (taking m_1 as x). Therefore, we get that

$$m_1 = \frac{M + M(1-4\eta)^{1/2}}{2}, \quad m_2 = \frac{M - M(1-4\eta)^{1/2}}{2}. \tag{10}$$

References

[1] P. Ajith (2008) Gravitational-wave data analysis using binary black-hole waveforms. *Class.Quant.Grav.* **25**, *114033*.

[2] B. P. Abbott et al.(2016) (LIGO Scientific Collaboration and Virgo Collaboration) Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys. Rev. Lett.* **116**, *061102*.

[3] https://www.ligo.caltech.edu/ (Accessed 28th March 2016).

[4] Veitch, J. et al. (2015) Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library. *Physical Review D*, **91(4)**, *042003*.