

Markov Chain Monte Carlo (MCMC) and Bayesian Inference for Gravitational Waves

Christine Andersson

Lund Observatory
Lund University



2021-EXA179

Degree project of 15 higher education credits
June 2021

Supervisor: David Hobbs

Lund Observatory
Box 43
SE-221 00 Lund
Sweden

Abstract

The Laser Interferometer Space Antenna (LISA) is a space borne gravitational wave detector set to launch in 2034, with the objective of detecting and studying the Gravitational Waves (GWs) of our universe. So far, ground-based detectors such as the Laser Interferometer Gravitational-Wave Observatory (LIGO) have been successful in detecting GWs, but the limitations of ground based detectors is what makes LISA so special. With three separate space crafts, each 2.5 million kilometers apart, the detector is expected to measure gravitational radiation within the frequency regime of 0.1 mHz to 1 Hz. With LISA, astronomers will be able to determine the type, mass, as well as the energy released by GW sources. From the data measured with LISA, scientists will be able to challenge Einstein's theory of General Relativity through the lens of gravity, and perhaps uncover vital information about our universe's past, present and future.

As LISA follows an Earth trailing orbit around the Sun, it is expected to detect GWs emitted by numerous sources simultaneously. This results in one single signal that contains the information of all detectable gravitational sources in the universe. In order to study a single source, the development of a mathematical computer model is required to extract the desired information. This thesis implements Bayesian inference with a stochastic sampling algorithm known as Markov Chain Monte Carlo (MCMC) to tackle the multi-dimensional problem and statistically recover the parameters of the GW.

In this thesis, we focus on the ecliptic coordinates of the source, which are just two out of the seven parameters of a GW. We found that MCMC was successful in the localisation of the source from a simulated gravitational wave strain. The ecliptic coordinates were recovered with a standard deviation of less than 1 degree. Expanding the program, we were also able to test the effectiveness of MCMC in the presence of multiple waves, and how their amplitude and frequencies affect the algorithms ability to recover the true position. Lastly, we conclude this paper with an alternative suggestion for extracting the parameters using Multinest, as well as comment on additional research.

Populärvetenskaplig beskrivning

Gravitation Waves (GWs), first predicted by Albert Einstein in the early 20th century, are a warping of spacetime invisible to the naked eye. Back in September of 2016, the most sensitive GW detector on Earth called LIGO - the Laser Interferometer Gravitational Wave Observatory - made the first ever direct observation of a GW. The GW was generated by the merging of two massive black holes about 1.4 billion light years away. More such GWs exist and travel through the universe from all directions and distances. In fact, whenever

an object has mass, it has the potential of warping space-time with its gravitational attractiveness. When an object is massive enough, such as a black hole or neutron star, a change in its velocity can create a large enough ripple in space-time to be detectable. The most common formation of GWs is when two massive gravitationally bound stellar objects in a binary system spiral inward and merge.

LIGO uses laser beams and light sensors stretched four kilometers apart to detect GWs, but even then, the detector is not sensitive enough to detect effectively. The European Space Agency (ESA) decided to build a GW detector sensitive enough to detect hundreds, if not thousands, of GWs with sensors stretched 2.5 million kilometers apart. How can such a detector exist on Earth? Well, it can't. The answer lies in the detector's name. The Laser Interferometer Space Antenna (LISA) is a GW detector stationed in space. It revolves around the Sun, trailing behind Earth, meant to collect GW signals for up to 5 years starting in 2034. There is a lot of work to do before then to make sure all the data collected by LISA is analyzable. One issue is determining where the GW originated and providing a coordinate system to pinpoint the direction of the source in space. Our thesis explores the effectiveness of using a Markov Chain Monte Carlo (MCMC), with the implementation of Baye's Theorem from statistics, in a computer algorithm to locate GW sources.

A good way of understanding how MCMC works is by imagining a robot that makes random jumps around a hill with the goal of finding the peak. Each time the robot lands in a spot higher than its previous record height, it stores the location and continues this pattern until it eventually reaches the peak. In this case, the peak is the coordinates of the GW source. Implementing Baye's Theorem allows us to utilize our prior knowledge of GWs to weed out any physically meaningless data. This allows the algorithm to be more efficient as it restricts the "robot's" area of investigation.

Our project explores the basics of this method. The results obtained showed us now MCMC is able to recover the position of a GW source from it's strain signal to an accuracy of ± 0.6 degrees. The algorithm is also able to 'tune' to different frequencies of GWs to extract only the desired relevant information. Multiple sources of the same frequency were also detectable, albeit more research needs to be done about problems where source positions overlap, causing confusion in the results. Additionally, the algorithm can also be modified to search and recover all 7 parameters of a GW at the same time. By studying these GWs in greater detail, astronomers will be able to probe the universe from an angle we were unable to in the past. Physicists will be able to use LISA to challenge theories such as General Relativity, by studying the gravity of the early universe.

Contents

1 Introduction	5
2 Theory	5
2.1 LISA	5
2.2 Gravitational Waves	7
2.2.1 Binary systems	8
2.2.2 Detector geometry and orientation	9
2.2.3 Parameters	10
2.2.4 Strain	11
2.2.5 Polarisation	12
2.3 Bayesian Analysis	14
2.4 Markov Chain Monte Carlo	16
3 Modelling	17
3.1 GW simulator	17
3.2 MCMC algorithm	17
3.2.1 Priors	18
3.3 Modifications	19
4 Results	19
4.1 Standard Run with single GW source	19
4.2 Adding a second GW source	27
4.3 2 GW sources with varying frequency	29
5 Discussion	30
5.1 Standard run	30
5.2 Two sources	31
5.3 Alternative	33
6 Conclusion	34

7 Appendices

35

8 References

37

List of acronyms

GW - Graviational Waves

LISA - Laser Interferometer Space Antenna

LIGO - Laser Interferometer Gravitational Wave Observatory

ESA - European Space Agency

MCMC - Markov Chain Monte Carlo

EMRIs - Extreme Mass Ratio Inspirals

SgrA* - Sagittarius A*

SMBH - Super Massive Black Hole

BH - Black Hole

WDs - White Dwarfs

NS - Neutron Stars

1 Introduction

Over the course of the last few years, astronomers were able to directly detect and study the gravitational waves produced by orbiting high masses. Ground based gravitational wave detectors such as LIGO (Laser Interferometer Gravitational Wave Observatory) were successful in detecting gravitational waves within a high frequency regime, but more sensitive detectors are required to study a wider range of gravitational wave inducing sources. LISA, the Laser Interferometer Space Antenna, is an ongoing mission plan led by the European Space Agency to place a gravitational wave detector in space, and is set to launch in 2034. With a frequency regime of 0.1 mHz to 1 Hz, LISA will be able to probe a wide variety of gravitational wave sources from all depths of space. By observing the effects of general relativity in our universe with gravity instead of the usual technique of electromagnetic radiation, astronomers hope to gain a better understanding of the formation, evolution, and even future of our Universe through this new lens.

In order to maximise the data we can access with LISA, we need to first investigate the most accurate and efficient way of extracting the parameters of the gravitational waves. Gravitational waves can be characterised by a total of seven parameters, and when taking into account the large number of simultaneous waves LISA will record, the problem becomes high dimensional. Stochastic algorithms such as MCMC has proven to be a powerful tool in the detection of exoplanets for the Gaia telescope (Ranalli et al. 2018). Therefore, using a similar approach, this thesis explores the capacity to which the MCMC can extract the characteristics of hypothetical LISA-like GW signals on the basis of Bayesian Inference. Specifically, this thesis focuses on using statistics to accurately extract the location of the gravitational wave sources in space to an acceptable degree of standard deviation.

This thesis begins with a brief introduction to the LISA mission in section 2.1, the basic theory of gravitational waves and detector geometry in section 2.2, followed by an introduction to statistical Bayesian inference (section 2.3) and how the Markov Chain Monte Carlo is implemented (section 2.4). The main components of the algorithm are described in the modelling section 3 with the experimental results presented in the results section 4. In the discussion section 5 we examine the results and its implications as well as suggest alternative approaches to the problem.

2 Theory

2.1 LISA

Ground based gravitational wave detectors, such as the Laser Interferometer Gravitational Wave Observatory (LIGO) in the United States, have successfully detected the passing of GWs with the first ever direct observation made back in September of 2015 (Abbott et al. 2016). Even with new detections being made on Earth, scientists and space agencies are

pursuing a new space based detector project - known as the Laser Interferometer Space Antenna (LISA) mission - that will increase the sensitivity of detection to give further insight into GWs and our understanding of the subject.

Led by the European Space Agency (ESA), LISA is planned to launch in 2034 where it will follow an Earth trailing heliocentric orbit around the Sun to simultaneously detect and measure multiple GWs from any location in space. Ground based detectors are limited to a frequency regime of 10Hz or larger, whereas LISA aims to detect GWs within a frequency regime of 0.1 mHz to 1 Hz (Danzmann 2000). These frequencies correspond to longer wavelengths that occur in heavier and wider orbiting binary systems, different from the sources that LIGO can detect. Fig. 1 below illustrates the range of detectable sources by LISA, where the frequency range of typical sources are indicated. The sensitivity of a detector has much to do with the length of its laser beams. In the case of LIGO, the detector consists of a two-arm Michelson interferometer stretching 4 kilometers both ways (Aasi et al. 2015).

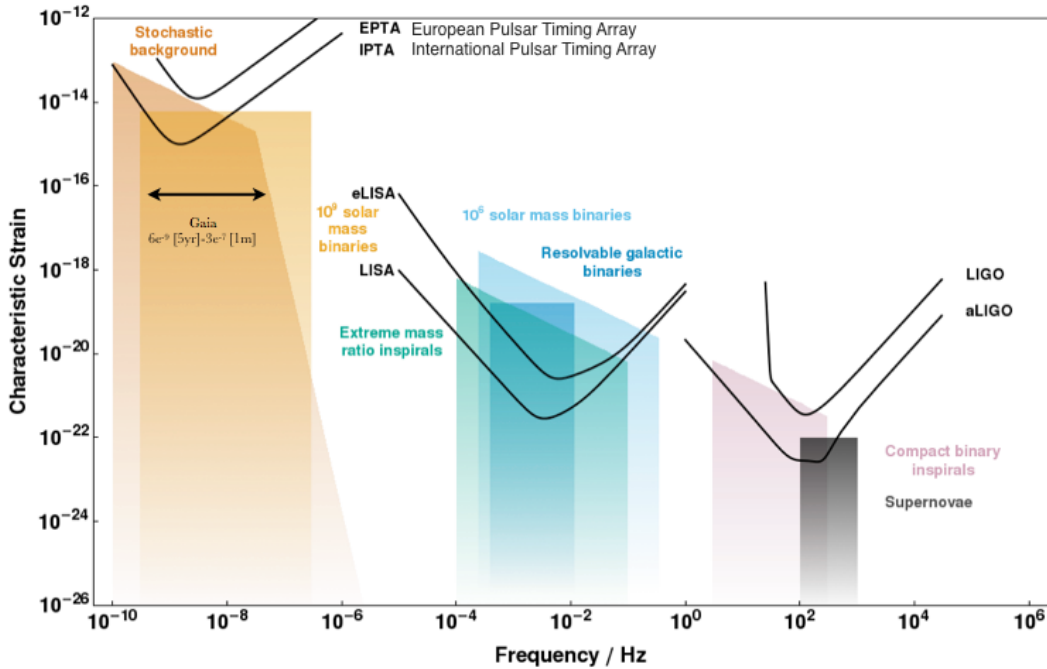


Figure 1: Comparison of space borne detector (LISA) to ground-based detector (LIGO) and Gaia telescope sensitivity. A variety of typical sources are indicated in their corresponding frequency ranges - from supernovae and binary star mergers in the kHz range to EMRIs and resolvable binaries in the mHz range. It can be seen that sources ranging from $10^{-4} - 10$ Hz will only be detectable with a space interferometer. Figure taken from (Robson et al. 2019).

In order to produce a high sensitivity detector, the laser requires a longer path in order to detect smaller disturbances in spacetime. LISA comprises of three separate spacecrafts

each orbiting around the Sun at different inclinations. Combined together, the detectors form an equilateral triangle with arm lengths 2.5 million kilometers apart (Consortium 2021), illustrated in Fig. 2 below, on a fixed plane that is inclined 60° from the plane of orbit and the Sun, trailing the Earth at 20° (Sweetser 2005). The spacecrafts exchange laser beams in order to detect a GW-subjected change in distance between matter. When no GW is present, the laser beams all interfere destructively at the detectors. Once a GW passes through and there is a significant change in the proper distance between particles, a change in the superposition of the laser beams will cause a signal to be produced and recorded by the detector.

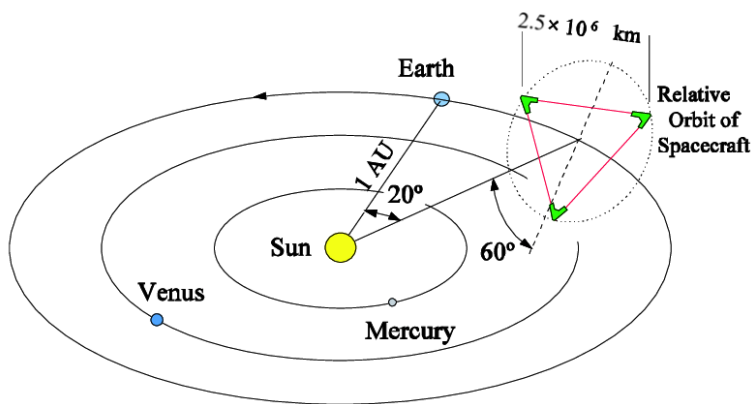


Figure 2: Schematic diagram of LISA's proposed trail behind Earth, orbiting the Sun. Figure modified from (Hough et al. 2005).

2.2 Gravitational Waves

Any object with a mass possesses a gravitational attractiveness that warps spacetime. Some massive objects in space include neutron stars, galaxies, and even black holes. Large stellar objects like the Sun have strong gravitational potentials which results in planets, asteroids and other stellar bodies orbiting the distorted space.

When an object with large enough mass changes its velocity, a wave distortion is created in space time and propagates at the speed of light. An oscillating change in space time under these circumstances is called a Gravitational Wave. Two high mass objects, such as two black holes or neutron stars, can become gravitationally bound to one another and orbit each other in a binary system. As each body follows circular motion, an acceleration is applied to their velocities, and together the system emits ripples in space-time that are detectable as GWs.

This thesis focus' on Compact Binaries and Inspiral GWs, which are binary systems consisting of two bodies orbit around each other, illustrated in Fig. 3, slowly converging.

Inspiral occurs when energy is released during the process of compact orbiting. This release of energy allows the orbit to shrink until eventually the two bodies collide. This was predicted by Einstein’s theory of General Relativity, where two heavy mass objects orbiting around each other release gravitational radiation as a result of the acceleration force.

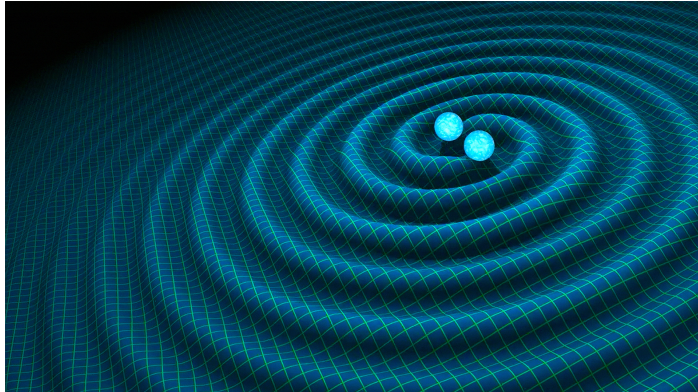


Figure 3: An artist’s interpretation of a compact binary system generating gravitational waves. Figure taken from NASA.

This gravitational radiation emitted carries the energy and momentum of the system, which in turn affects the orbital decay of the binary (Collaboration 2021). There are a couple of standard GW systems: Binary neutron star (BNS); BinaryBlack Hole (BBH); Neutron star-Black hole binary (NSBH) (Caltech 2021). LISA’s frequency regime shifts the focus onto systems that ground-based detectors cannot detect, such as Extreme Mass Ratio Inspirals (EMRIs) or ultra-compact binaries (NASA 2021).

2.2.1 Binary systems

Sagittarius A* (SgrA*) is a high mass radio source at the center of the Milky Way. Measurements of stellar orbits around SgrA* and its X-ray and polarized near-infrared emissions provide strong support for it being the supermassive black hole (SMBH) at the center of our galaxy (Eckart et al. 2017), which means SgrA* will provide a nearby EMRI system to observe with LISA. EMRIs are, in theory, easily distinguishable in measured data due to the strong GW signals created as a result of their high masses. Stellar remnants such as neutron stars (NSs), stellar black holes (BHs) and white dwarfs (WDs) are sometimes locked into an inspiral orbit near the event horizon of SMBHs without being tidally stripped. These interactions result in the production of GWs that are detectable by interferometer detectors such as LISA when the SMBH is of a mass around $(10^4\text{--}10^7) M_{\odot}$ - providing vital information about the SMBH as well as the evolutionary dynamic of our Galaxy (Amaro-Seoane et al. 2012).

On a smaller scale, we have ultra-compact binaries that often consist of pairs of WDs or

NSs, many of which have orbital periods ranging from minutes to hours (Suvorov 2021). These systems produce relatively weak signals due to their low mass, but will dominate the detections made by LISA due to their abundance in nearby galaxies (Amaro-Seoane et al. 2012). Unresolvable GW signals from these compact systems will create background noise in the form of a combination of weak GW signals from simultaneous sources. This is known as ‘Stochastic Gravitational Waves’ (Chakrabarty 1999). They are unavoidable by LISA and should be taken into consideration when simulating detectable data. Difficult to detect with telescopes and electromagnetic systems, LISA will provide a unique perspective on the formation, evolution, merger rate, and various other physical aspects of these GW inducing sources.

2.2.2 Detector geometry and orientation

In order to recover the position of a GW’s source, we need to clarify the orientation of the source in the sky as well as it’s relative orientation with respect to the detector. Figure 4 below illustrates the coordinate system from the reference frame of the Sun. The recovered ecliptic coordinates describing the position of the source in this thesis will be from this reference frame.

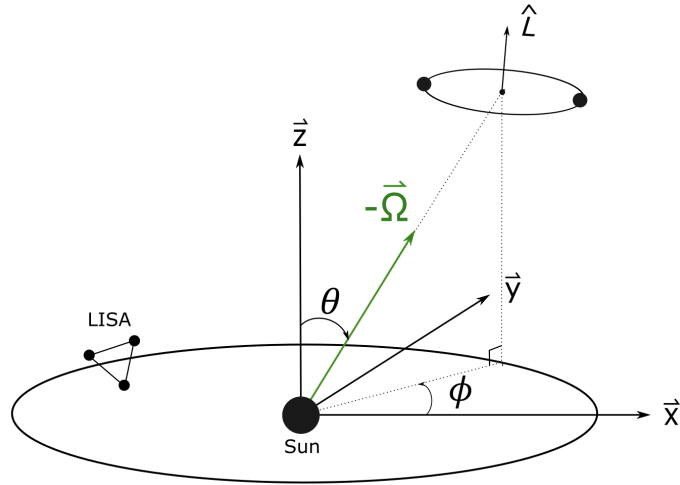


Figure 4: Coordinate orientation in the Sun’s reference frame. The vector $\bar{\Omega}$ is the direction of propagation of the GW, with the location of the source denoted by θ and ϕ in ecliptic coordinates. Here \hat{L} is the angular momentum of the system.

Using a GW simulator in this project, we also need to lay out the orientation of the vector space from the reference frame of the source. This is illustrated in Fig. 5 below.

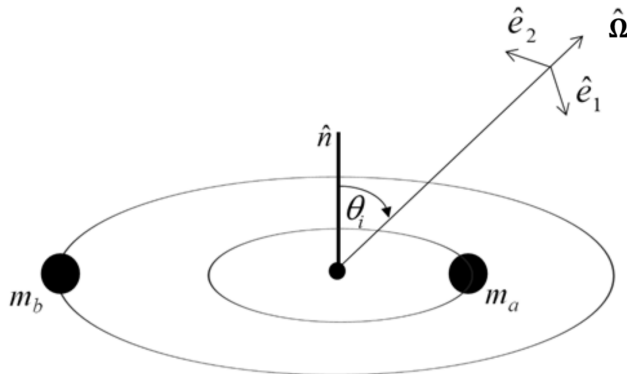


Figure 5: Vector orientation of a binary system with orbiting masses m_a and m_b . \hat{n} represents the normal vector perpendicular to the plane containing the masses, and is parallel to the direction of angular momentum L . The vector $-\hat{\Omega}$ points in the opposite direction of GW propagation, with an inclination angle θ_i from \hat{n} . \hat{e}_1 is in the plane of $(\hat{n}, \hat{\Omega})$, and \hat{e}_2 is perpendicular to it. Figure modified from (Hilborn 2019).

In order to describe a source pointed towards the direction of $-\hat{\Omega}$ using the ecliptic coordinate system (θ, ϕ) , we construct the following orthogonal triad

$$\begin{aligned}
 e_1 &= \sin \phi \hat{x} - \cos \phi \hat{y} \\
 e_2 &= \cos \theta [\cos \phi \hat{x} + \sin \phi \hat{y}] - \sin \theta \hat{z} \\
 \hat{\Omega} &= \sin \theta [\cos \phi \hat{x} + \sin \phi \hat{y}] + \cos \theta \hat{z}
 \end{aligned}
 \tag{2.1}$$

2.2.3 Parameters

In this section, we briefly introduce the 7 main parameters that describe the characteristics of a GW (Cornish & Crowder 2005). The parameters, that describe a Newtonian GW, are given under the state vector λ in the following equation

$$\bar{\lambda} \rightarrow (\theta, \phi, l, \psi, \phi_0, f, \mathcal{A}).
 \tag{2.2}$$

l is the inclination angle and ψ the polarisation angle of the system. ϕ_0 is the initial angle, used to restrict the range to $[0, \pi]$ when working with the vector orientation of the system in the GW simulation.

The frequency of the GW source, f , is given by

$$f = 2f_{\text{orb}} = \frac{1}{\pi} \sqrt{\frac{M_1 + M_2}{r^3}},
 \tag{2.3}$$

where $M_{1,2}$ are the masses of the binary and r the distance between the masses. The factor 2 corrects the frequency of the GW, as it is twice the orbital frequency f (Hilborn 2019).

The amplitude, \mathcal{A} , is given by

$$\mathcal{A} = \frac{2M_1M_2}{rd}, \quad (2.4)$$

with d representing the distance between the source and the detector.

The variables of interest in this project are (θ, ϕ) . They are known as the elliptic coordinates and reveal the location of the GW in space. The angle θ represents the ecliptic co-latitude, and has the range $[0, \pi]$. ϕ , the ecliptic co-longitude, has the range $[0, 2\pi]$. Simply by determining the location of the GW from the detector's strain signal gives us the information needed to find out other physical properties of the waves and source. The triangulation method for gravitational waves (Fairhurst 2011) incorporate a network of ground-based detectors to reduce the uncertainty in the localization of sources. Comparing the differences in strain signal can aid in determining the polarisation of the waves. Additionally, the relative amplitude of the GW can be determined using a purely geometric approach, uncovering more information about binary systems which we otherwise would not be able to obtain (Hilborn 2019).

2.2.4 Strain

The amplitude of the signal recorded when a passing GW is detected is often referred to as the ‘‘strain’’ and is dimensionless. It is fundamentally a measure of the change in displacement of matter over the total distance, $\Delta L/L$ (Chaudhuri 2016). However, to fully describe a GW, a more complicated strain equation exists so that other physical properties of the wave - such as its amplitude, orbital period and polarisation - are quantifiable. The following equations and derivations were taken and adapted from (Hilborn 2019).

The strain signal for a low frequency GW in the case of LISA can be expressed as the addition of both polarization modes,

$$s(t) = A_+ F^+(t) \cos 2\Phi(t) + A_\times F^\times(t) \sin 2\Phi(t). \quad (2.5)$$

which can be simplified using geometric identities and Eq. (2.8) as

$$s(t) = A(t) \cos \Psi(t). \quad (2.6)$$

The simplified strain in Eq. (2.6) is dependent on the modulation amplitude $A(t)$ and the GW phase $\Psi(t)$ - a combination of the time-dependent orbital phase $\Phi(t)$ and the initial phase angle Φ_d ,

$$\Psi(t) = 2\Phi(t) + \Phi_d. \quad (2.7)$$

Here, Φ_d also describes the Doppler modulation, which occurs due to the relative motion between the source and the detector (Cornish & Larson 2003).

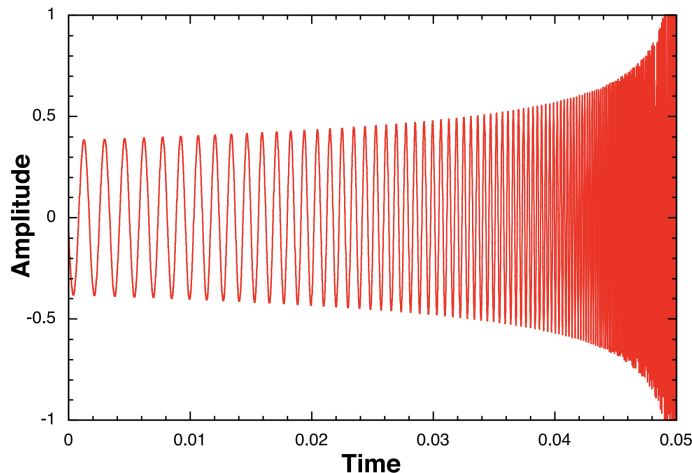


Figure 6: Example of an inspiral compact binary’s GW strain with arbitrary units. Figure taken from Utah State University’s General Relativity Notes: <http://www.physics.usu.edu/Wheeler/GenRel2013/Notes/GravitationalWaves.pdf>.

Fig. 6 above shows a plot of the strain, or amplitude, of a compact binary GW as a function of time. As the binary begins to merge, the period of the system’s orbit decreased and more energy is released in the form of gravitational waves - thus increasing the amplitude and frequency of the signal. The modulation amplitude, $A(t)$, in the strain signal is defined as (Barrellon-Vernay 2018),

$$A(t) = \sqrt{A_+ F^+(t)^2 + A_\times F^\times(t)^2}, \quad (2.8)$$

where A_+ and A_\times are the polarisation tensor amplitudes discussed in section 2.2.5 below.

2.2.5 Polarisation

The main effect of GWs on matter is the compression and stretching of spacetime. As mentioned earlier, the warping of spacetime comes from the energy and angular momentum imparted by the original source. This oscillatory change in relative distance between particles is caused by what is called the “polarisation” of GWs. There are two modes, named plus (+) and cross (×) polarisation (Addison 2014). The effects of either on a ring of test particles is illustrated in Fig. 7 below, in quarter period time frames and can be applied generally to all matter.

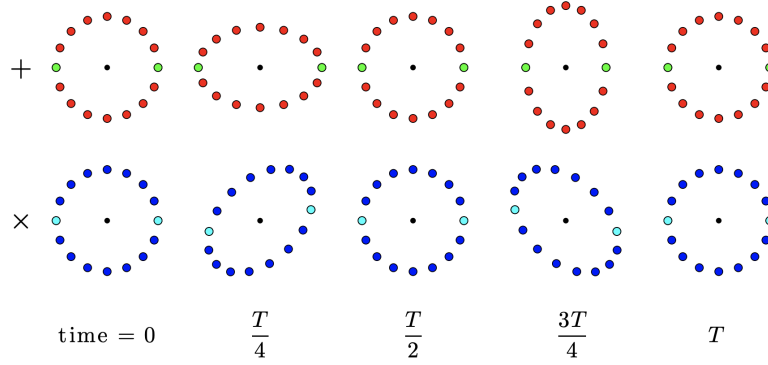


Figure 7: The effects of polarisation on a ring of test particles. Plus polarisation is illustrated in the top row and cross polarisation on the bottom. Figure taken from (Addison 2014).

The amplitude of the polarisation tensor tells us the size, or strength, of the warping of space time due to a specific wave. It is related to the direction of propagation of the wave, as well as the direction of the angular momentum

$$A_+ = \mathcal{A}[1 + (\hat{L} \cdot \hat{\Omega})^2] \quad (2.9)$$

$$A_\times = 2\mathcal{A}\hat{L} \cdot \hat{\Omega} \quad (2.10)$$

with $\hat{\Omega}$ the direction of propagation of the GW and \hat{L} the normal vector of the angular momentum. Additionally, we need to define the polarisation of the detector's beam pattern (Cornish & Crowder 2005), which depend on the antenna's orientation (Giampieri 1998) with respect to the GW's direction of propagation

$$F_+(t) = \frac{1}{2}[\cos 2\psi D^+(t) - \sin 2\psi D^\times(t)] \quad (2.11)$$

$$F_\times(t) = \frac{1}{2}[\sin 2\psi D^+(t) + \cos 2\psi D^\times(t)] \quad (2.12)$$

where the functions $D^+(t)$ and $D^\times(t)$ contain information about the orbital phase and orientation of the LISA spacecrafts (Rubbo et al. 2004),

$$\begin{aligned}
D^+(t) = \frac{\sqrt{3}}{64} [& -36 \sin^2 \theta \sin [2\alpha(t) - 2\lambda] + (3 \\
& + \cos 2\theta) (\cos 2\phi \{9 \sin 2\lambda - \sin [4\alpha(t) - 2\lambda]\} \\
& + \sin 2\phi \{\cos [4\alpha(t) - 2\lambda] - 9 \cos 2\lambda\}) \\
& - 4\sqrt{3} \sin 2\theta \{\sin [3\alpha(t) - 2\lambda - \phi] \\
& - 3 \sin [\alpha(t) - 2\lambda + \phi]\}]
\end{aligned} \tag{2.13}$$

$$\begin{aligned}
D^\times(t) = \frac{1}{16} (& \sqrt{3} \cos \theta \{9 \cos (2\lambda - 2\phi) \\
& - \cos [4\alpha(t) - 2\lambda - 2\phi]\} \\
& - 6 \sin \theta \{\cos [3\alpha(t) - 2\lambda - \phi] \\
& + 3 \cos [\alpha(t) - 2\lambda + \phi]\})
\end{aligned} \tag{2.14}$$

2.3 Bayesian Analysis

The data obtained by LISA will consist of a single signal containing the information of multiple GWs measured simultaneously. In the case of such complex data, the application of statistical methods such as Bayesian Inference can increase the efficiency of algorithms to better solve for the unknown. Bayesian Inference stems from Bayes' Theorem, of which we will give a brief explanation for in this section.

First introduced by Thomas Bayes in the 1770s, Bayes' Theorem tells us what the probability of a scenario is, given some known prior knowledge of a certain conditions that is related to the event. This allows for a more accurate assessment of the probability, as prior knowledge is taken into consideration instead of simply assigning the same probability distribution without any individualization or taking into account background information (Kunin 2018).

Bayes' rule gives us the following posterior density

$$P(\theta|\text{data}) = \frac{P(\text{data}|\theta) \times P(\theta)}{P(\text{data})} . \tag{2.15}$$

$P(\theta)$ is the prior probability. This is the probability assumed before obtaining background knowledge. $P(\text{data})$ is the evidence. This is the probability of data as determined by accumulating a variety of results from all possible values of scenario θ and is affected by preexisting knowledge. $P(\text{data}|\theta)$ is known as the likelihood function. It tells us to what degree the prior knowledge predicts the data (Box & Tiao 2011). Fig. 8 below gives a visual representation of how the likelihood is affected by the prior knowledge of the scenario. If the data is assumed to follow a normal distribution and a sample is taken from the population, the likelihood is an adjusted distribution that takes into account the results of the sample.

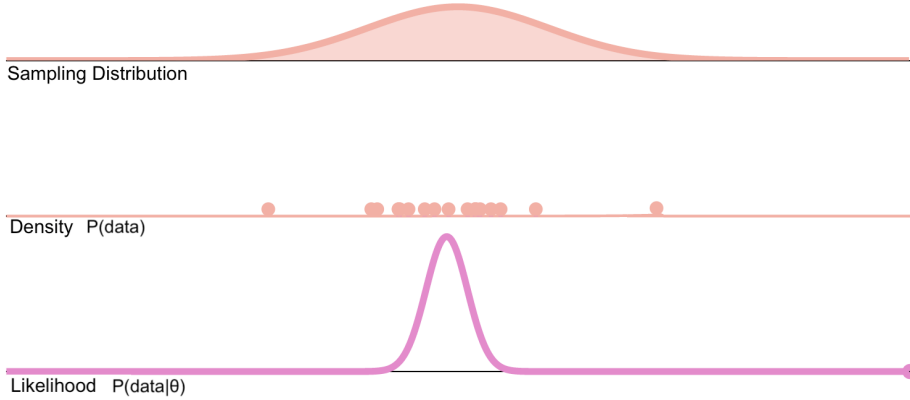


Figure 8: The likelihood function derived from a random sample $n = 19$ taken from a normal distribution. Figure modified from (Kunin 2018).

Statistical inference is when data about a probability distribution is taken to determine the properties of a population. Applying Bayes' Theorem in this process is called Bayesian inference. One of the main concepts of Bayesian statistics is how prior knowledge in an event should be updated when new evidence is obtained. Therefore, the more evidence that is gathered, the more accurate the posterior probability will be.

There are two key parts to Bayesian inference: parameter estimation and hypothesis testing (Smith et al. 2020). In this thesis, the posterior density given in Eq. (2.15) is the parameter estimation and the evidence, $P(data)$, is the hypothesis testing. With the use of a sampling technique such as the Markov Chain Monte Carlo (discussed in further detail in the next section), Bayesian inference can accurately evaluate the posterior and evidence of the problem, given a set of prior beliefs.

In reality, prior probabilities are often a distribution of data which can be described with a probability density function instead of a single number value. A standard deviation (σ) is given, in addition to the mean (μ), and a specific sampling distribution is assigned depending of the nature of the problem. For this thesis, a Gaussian distribution is chosen. This is the prior distribution as we presume the probability distribution of GW sources should be an equal number both below and above the mean value. With this distribution, we can generate random data which then forms the likelihood distribution. With each iteration the function is updated to fit the results from the data taken. For a Gaussian probability density function, we have the following equation (Ribeiro 2004),

$$P(x; \mu, \sigma) = \frac{1}{\sigma_{AL}\sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma_{AL}^2}\right) \quad (2.16)$$

where x is the real argument, σ the standard deviation and μ the calculated mean. (Ranalli et al. 2018) uses Bayesian inference concepts with a normal distribution similar to the scope of our project, and the technique used can be applied to this project. In our case, x is the

observed gravitational wave strain, σ_{AL} the systematic measurement uncertainty (or noise) in LISA, and μ the computed strain value. The residual $(x - \mu)$ in the above equation tells us the difference between the computed strain value to that of the observed. Following a normal distribution, the mean value of the residual should be zero with a standard deviation σ_{AL} .

2.4 Markov Chain Monte Carlo

Gathering enough data for some of the probabilities in Bayesian inference, such as the evidence, can take a lot of time. One way to increase efficiency and obtain accurate results is to use a stochastic sampling algorithms such as Markov Chain Monte Carlo (MCMC). MCMC takes samples from probability distributions and uses machine learning to compute the distribution over parameters. It is useful when direct sampling is challenging as it extracts random samples from a probability distribution. Since LISA is a mission plan still under development, taking direct samples from space observations is not possible, meaning we are required to sample randomly generated data from a sampling algorithm. By using the MCMC method, the algorithm can update the evidence used in Bayesian inference to guide the model towards more accurate values. The samples rely heavily on the posterior as MCMC rejects or accepted values based on prior conditioning. The larger the chain, the more data is collected, which statistically leads to a more reliable distribution.

For this project, we choose to apply the Metropolis-Hastings implementation of MCMC as it has proven to be reliable in multi-dimensional cases (Robert 2015). The intent is to investigate whether MCMC can efficiently determine the two coordinate parameters of a GW, and then expand the algorithm to eventually extract all seven parameters of the GW from its signal. For this to work, the algorithm we choose requires to perform similarly under high-dimensional analysis, and thus the Metropolis-Hastings approach appears promising.

Metropolis-Hastings takes samples from a proposal distribution and “walks” along the ecliptic coordinate space. As it does so, it compares the likelihood of the sample to the simulated wave, rejecting or accepting the sample based on the residuals between the two - demonstrated in figure 9 below. The algorithm then jumps to a new random location and repeats the same process as it slowly converges on one specific point in space.

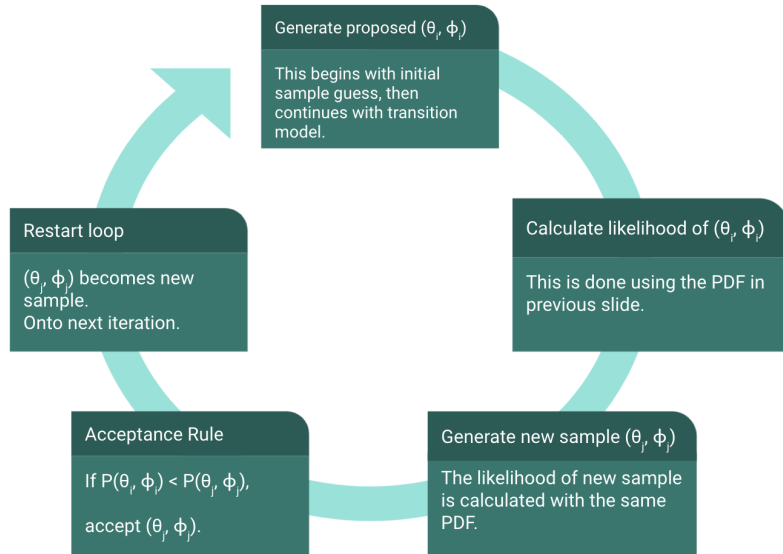


Figure 9: Diagram describing the selection process for a Metropolis-Hastings function in python. A proposed sample of the ecliptic coordinate is used to compute the likelihood, and then compared to a second sample generated from a random normal distribution. The sample with the higher likelihood is accepted, and the process is repeated in the next iteration.

3 Modelling

3.1 GW simulator

In order to investigate the effectiveness of MCMC in solving for GW parameters, we are required to first simulate hypothetical GWs that LISA would measure. We used a GW simulator program (Barrellon-Vernay 2018) developed at Lund University as a foundation for our algorithm. The program takes into account the spacetime geometry, GW parameters and polarisation under section 2.2 in the file `GW.py`, to produce a GW strain as a function of time in `genGW.py`. The program focuses on just two of the seven parameters of a GW: θ and ϕ . We import the files along with `const.py`, a file full of predetermined constants, to then generate an “observed” GW strain which is then used in our program alongside a second “computed” GW strain, using the ecliptic coordinates generated by the MCMC, to compute the residual in Eq. (2.16).

3.2 MCMC algorithm

The MCMC implemented in this project was developed based on a tutorial published and written by Joseph Moukarzel on Github (Moukarzel 2019). The Metropolis-Hastings code

we implement uses the same concept but with different parameters. This way, we were able to utilise the same Bayesian Inference techniques used by Moukarzel to solve for our GWs. The algorithm initiates the chain with a starting sample guess. This initial value is chosen manually to begin the sampling process that eventually leads to convergence. We consider the first half of the samples drawn by the chain to be “burn-in” values that are discarded when studying the posterior distribution. This way, there is no bias contribution from the initial steps taken by the MCMC in possible low-probability regions. Storing the accepted and rejected samples in each iteration into separate arrays, we then plotted histograms of 30 bins to display the frequency of coordinates sampled. The density for the histogram is set to `True` and displays the frequency for each bin in terms of “counts” in a bin i divided by the total counts in each bin multiplied by the width of the bin,

$$\text{Density} = \frac{\text{counts}}{\sum \text{counts} \times (\text{bin width})}. \quad (3.17)$$

Using the `norm.fit` function, we then fitted a gaussian curve to the histogram to find the mean and standard deviation of the final convergence of the location.

3.2.1 Priors

As the algorithm samples random data, a wide range of possible ecliptic coordinate guesses will be returned. The application of prior knowledge in Bayes’ Theorem removes any physically meaningless data from being inferred (Trotta 2008). This works as a tool to guide the MCMC closer to realistic values, removing any unwanted guesses from the pool of data used in the inference process. Values that are deemed acceptable fall under the condition that the subsequent values are of higher likelihood than the previous iteration. In our case, the location of a GW source in the sky is limited to the coordinate range $[0, \pi]$ for θ and $[0, 2\pi]$ for ϕ . The range of θ and ϕ are defined in the prior parameter using the `priorGW` function, which automatically rejects any values outside of the range.

The prior parameter, however, does not account for the possible out-of-range sampling that Bayesian inference uses to derive the likelihood function. The Metropolis-Hastings function uses a transition model to draw samples of $[\theta, \phi]$. Since we assume the distribution of data to be gaussian, the `numpy.random.normal` routine is implemented in the `transition model(x)` function, with an updated mean value each iteration. This allows the algorithm to use existing evidence from previous samples, exactly how the likelihood function in Bayes’ Theorem is applied, to adapt to the new information. In order for the random normal distribution to sample ecliptic coordinates within a valid range, we embed the routine into `transition model(x)` with a wraparound technique in the form of `if/elif` statements. For example, for $(x < 0) :$, we add π or 2π depending on if it is a θ or ϕ coordinate.

3.3 Modifications

Once the program was successful in determining the location of a single GW, we added modifications to see whether the MCMC could effectively work in the same way under the condition of two GW sources being detected simultaneously. To add an additional GW signal, both waves were generated separately, and then the two strains were added together to make one strain.

Thus far we were able to generate a strain signal that contained multiple GWs with different locations. In reality, the GWs detected by LISA will have different frequencies (as well as different values for the other 4 parameters). We expanded our `GenerateGW` (Barrellon-Vernay 2018) function to include a frequency parameter to see how the program differentiates between GWs of different frequencies. This was an easy modification done by adding an extra parameter to `GenerateGW(theta, phi) → GenerateGW(f, theta, phi)` so that the frequency became an input variable. The next section presents the results obtained from various runs with alternative conditions.

4 Results

All simulations were run with constant values for the spiraling masses $M_1 = 0.5 M_\odot$, $M_2 = 0.02 M_\odot$. These were the original mass values used in Barrellon-Vernay (2018)'s simulation. The LISA detector is estimated to determine locations of massive black hole binaries with a precision of about 10 arc minutes, which is around 0.003 radians. The overall accuracy for varying types of sources is up to 1 degree, or 0.017 radians (Jennrich 2011). For this reason, the uncertainty of LISA in our program was set to 0.01 radians, a value in between this range of uncertainty, and is assumed to be the standard deviation when sampling from the transition model. LISA is planned to record data while in orbit for around 5 years. In view of simplicity, we set the time of recording data to 2 years, which is enough to display the measured periodicity of the strain in our simulation.

4.1 Standard Run with single GW source

The generated GW contains a unique strain that varies with time. Fig. 10 below shows an example of a GW strain signal obtained with Eq. (2.5). The strain has a clear period of 1 year, which correlates with the periodicity of LISA's orbit around the Sun. Within the one year, we notice the strain blowing up around the half year mark, which is most likely due to position of LISA being on the opposite side of it's orbit when data was first being recorded.

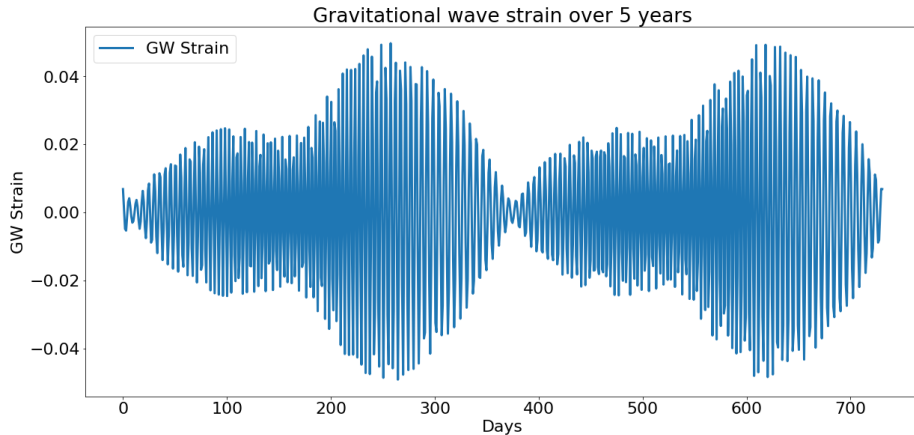


Figure 10: Strain of the generated GW over a period of 2 years.

In the standard run, we used a single GW in the sky and determined the number of iterations we should use to conserve the algorithm’s efficiency whilst allowing the chain to draw enough samples to converge at the desired location. After discarding the burn-in values, the list of sample coordinates should still contain enough values to produce statistically meaningful results. Table [1](#) below displays the mean, μ , and standard deviation, σ , of the histograms displaying the distribution of (θ, ϕ) sampled by the MCMC for different iterations.

Table 1: Simulation results for observed GW source at $(45^\circ, 70^\circ)$. $f = 10$ mHz. Standard deviation displayed in brackets.

MCMC Iteration	Residual mean	All θ mean [deg]	All ϕ mean [deg]	Accepted θ mean [deg]	Accepted ϕ mean [deg]
500	0 (0.0004)	45.028 (0.588)	69.985 (0.597)	46.349 (1.882)	73.196 (2.858)
1000	0 (0.0007)	44.939 (0.583)	70.008 (0.594)	45.853 (1.727)	71.688 (2.058)
5000	0 (0.0003)	45.019 (0.583)	70.009 (0.590)	45.306 (0.944)	70.799 (2.141)

In the table above, “All θ ” or “All ϕ ” stand for the combination of accepted and rejected θ and ϕ values. “Accepted θ ” or “Accepted ϕ ” stand for only the values MCMC accepted. All 3 different iteration lengths of the program successfully deduced the location of the GW to an accuracy of up to 3 degrees. At 500 iterations, we got slightly higher standard deviations for the accepted θ and ϕ values when compared to results from 5000 iterations. Overall, the residual mean was closest to 0 at 5000 iterations (see Fig. [11](#) for the histogram of residuals), and a general trend could be spotted where the standard deviation decreased with increasing iteration.

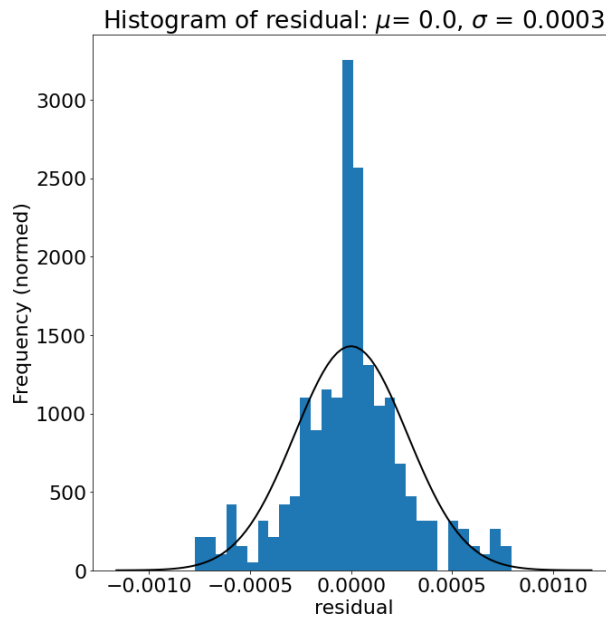


Figure 11: Histogram of the last 50% of residuals computed for 5000 iterations. The gaussian fit is displayed with a black line.

The Metropolis-Hastings function is conditioned to accepted new values that have a residual smaller than the previously accepted sample. This comes from Eq. (2.16), where the difference of the computed and observed strain is used to compare the precision of the computed strain to the true value.

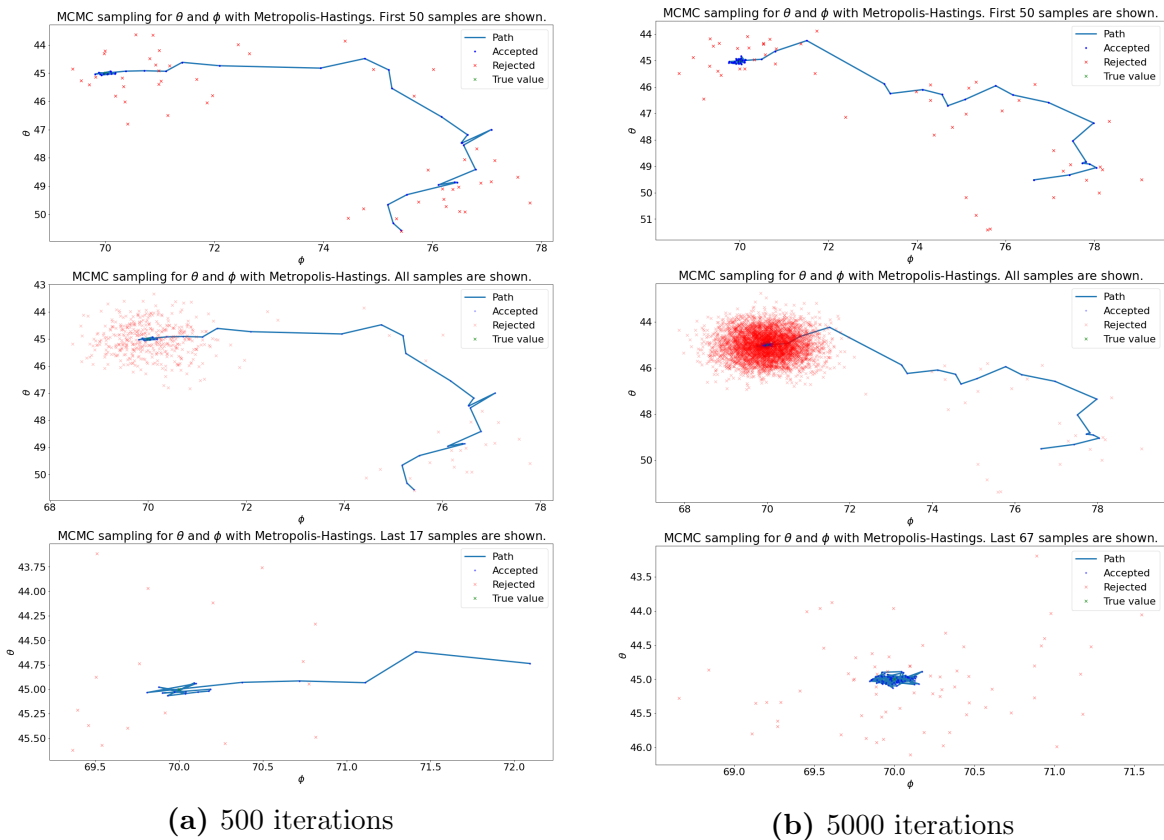


Figure 12: Graphical representation of the path taken by the MCMC to converge to the final value of the source location. From top to bottom: First 50 samples, All samples, last 50% of samples. Red marker represents a rejected sample, while the blue path traces the accepted samples.

Fig. 12 above traces the path the chain took to converging on the location of the GW source. In both 12a and 12b, the rejected samples are seen forming a gaussian distributed red halo around the true value. The last cell in each figure shows the last 50 percent of the samples and by comparing the two different iterations, we could see that while not all the burn-in values were properly discarded for 500 iterations, they were in 5000 iterations. Therefore, we could assume that 5000 iterations were a good sampling size for the MCMC to converge on the right location without taking too much time to compute.

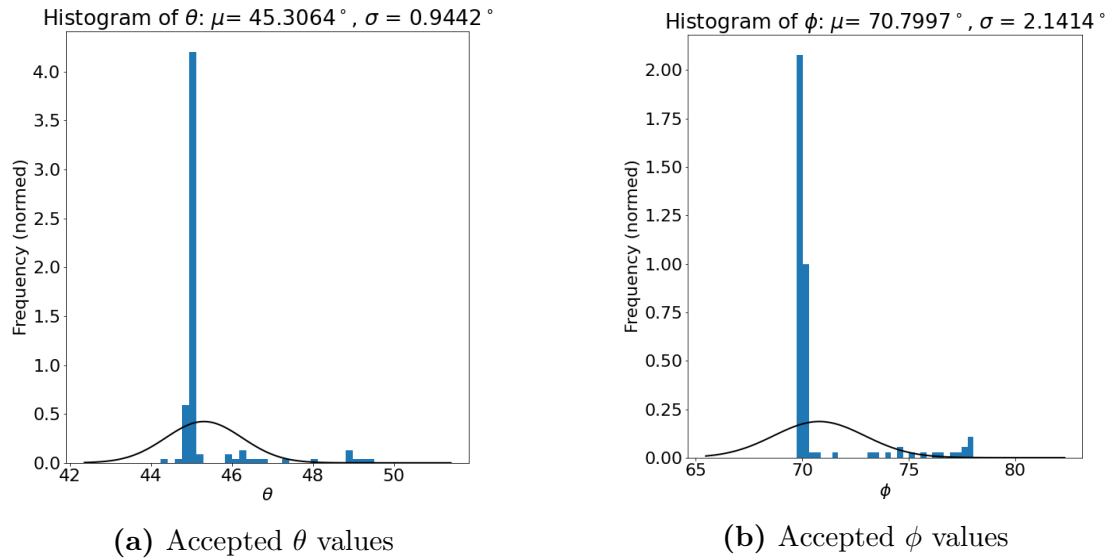


Figure 13: Last 50% of accepted jumps from 5000 iterations for θ and ϕ presented in a histogram and fitted with a gaussian approximation (black line).

The histograms in Fig. 13 above contain only values from the accepted array. Although the most frequently sampled coordinates are around 45 and 70 degrees, the standard deviation of these histograms are large. This is due to the accepted array containing much fewer values than the total samples taken, and thus the chain had not converged even after removing the first 50% of values. Compared to the histograms of the last 50% of all sampled jumps for 5000 iterations, Fig. 14 below shows much more evenly distributed frequencies of θ and ϕ . For the purpose of meaningful statistical analysis, we focus our attention to the histograms for all jumps and not just the chain's accepted values as they include valuable statistical information. Additionally, comparing the mean values and standard deviation of results from the two approaches, results from the histogram that incorporated both accepted and rejected values provided a more accurate and precise results.

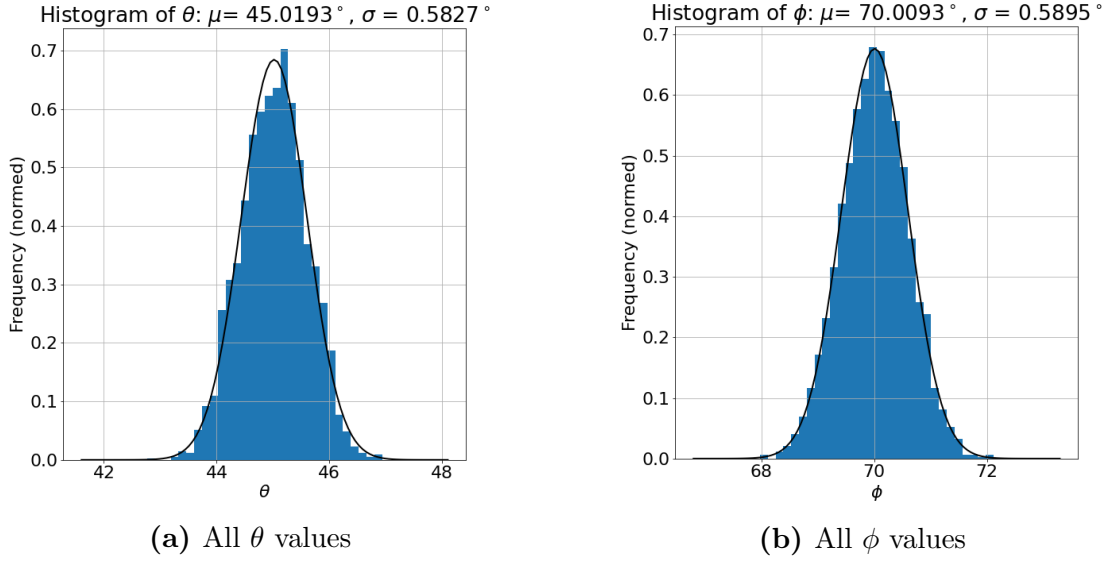


Figure 14: Histogram showing the posterior distribution of the coordinate parameters of a GW. The last 50% of all jumps for each ecliptic coordinate is included. Black line represents the gaussian approximation.

Using 5000 iterations as the standard amount of samples the algorithm will draw, the next step is to investigate how the systematic uncertainty, σ_{AL} in Eq. (2.16), in LISA's detector affects the standard deviation in the histograms for θ and ϕ . The results are plotted in Fig. 15 below, showcasing how the relation between systematic uncertainty and the standard deviation is linear for small values of uncertainty.

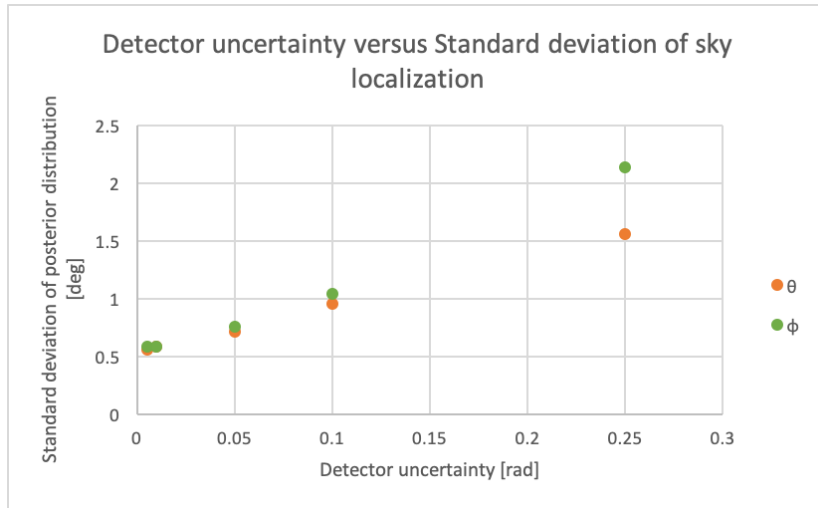


Figure 15: Scatter plot showing how the systematic error in LISA's detector affects the standard deviation of the gaussian fitted to the histogram for all samples obtained by MCMC algorithm. Green represents the standard deviation for ϕ and orange for θ .

Keeping σ_{AL} at 0.01 radians for the rest of this project, we then took a look at how the amplitude of a GW affects the accuracy of the recovered ecliptic coordinates obtained by the algorithm. The amplitude of the GW is defined by Eq.(2.4), where increasing the distance r between two stellar bodies will decrease the amplitude of the GW. Table 6 in the appendix displays the recovered ecliptic longitude ϕ and latitude θ for varying distances r , and Fig. 16 below is the results plotted with respect to the wave’s amplitude.

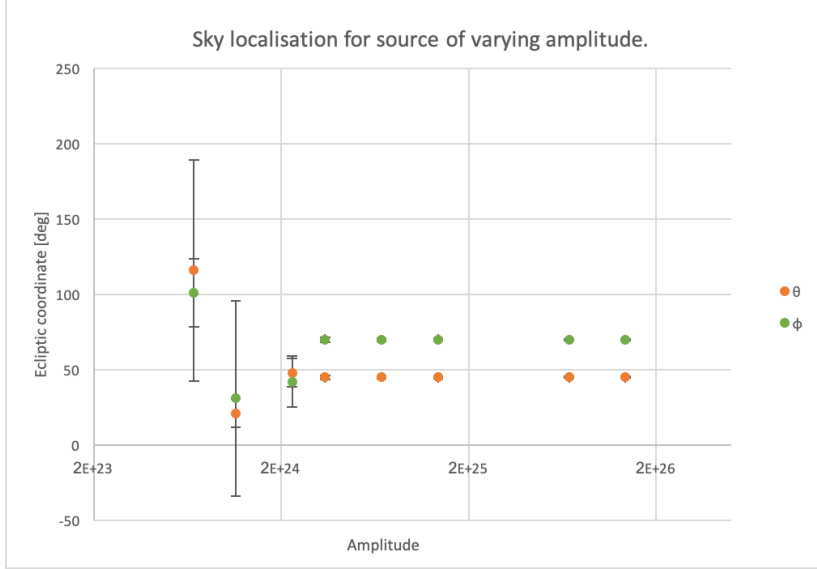


Figure 16: Semi-log plot for a single GW source at $(45^\circ, 70^\circ)$ showing how the amplitude of the GW affects the accuracy of the recovered position. The ecliptic longitude is denoted with green dots and ecliptic latitude in orange.

Keeping everything else constant, including the chain initialisation guess, we found that distances larger than 5.00×10^{-4} pc, which corresponds to an amplitude of 3.43×10^{24} (Eq. 2.4), is where the algorithm starts to confuse the position of the source. At a distance of 7.50×10^{-4} pc, the algorithm completely fails to accurately detect the source due to the amplitude being too small for proper resolving. To confirm that the resolution of the detector really is dependent on the system noise σ_{AL} , we repeated the run for 7.50×10^{-4} pc with a smaller σ_{AL} to see if the algorithm could recover the position. Table 2 below displays the results.

Table 2: The recovered positions of a source with small amplitude at $(45^\circ, 70^\circ)$, for two different degrees of systematic detector noise. MCMC Iteration = 5000, $f = 10\text{mHZ}$. Standard deviation displayed in brackets.

System noise σ_{AL} [rad]	Separation r [pc]	All θ mean [deg]	All ϕ mean [deg]
0.01	7.50×10^{-4}	48.106 (9.600)	42.107 (16.76)
0.005	7.50×10^{-4}	45.004 (1.059)	70.302 (1.430)

The figures below show how the convergence of MCMC changes when the distance between the two masses increase and the amplitude decreases.

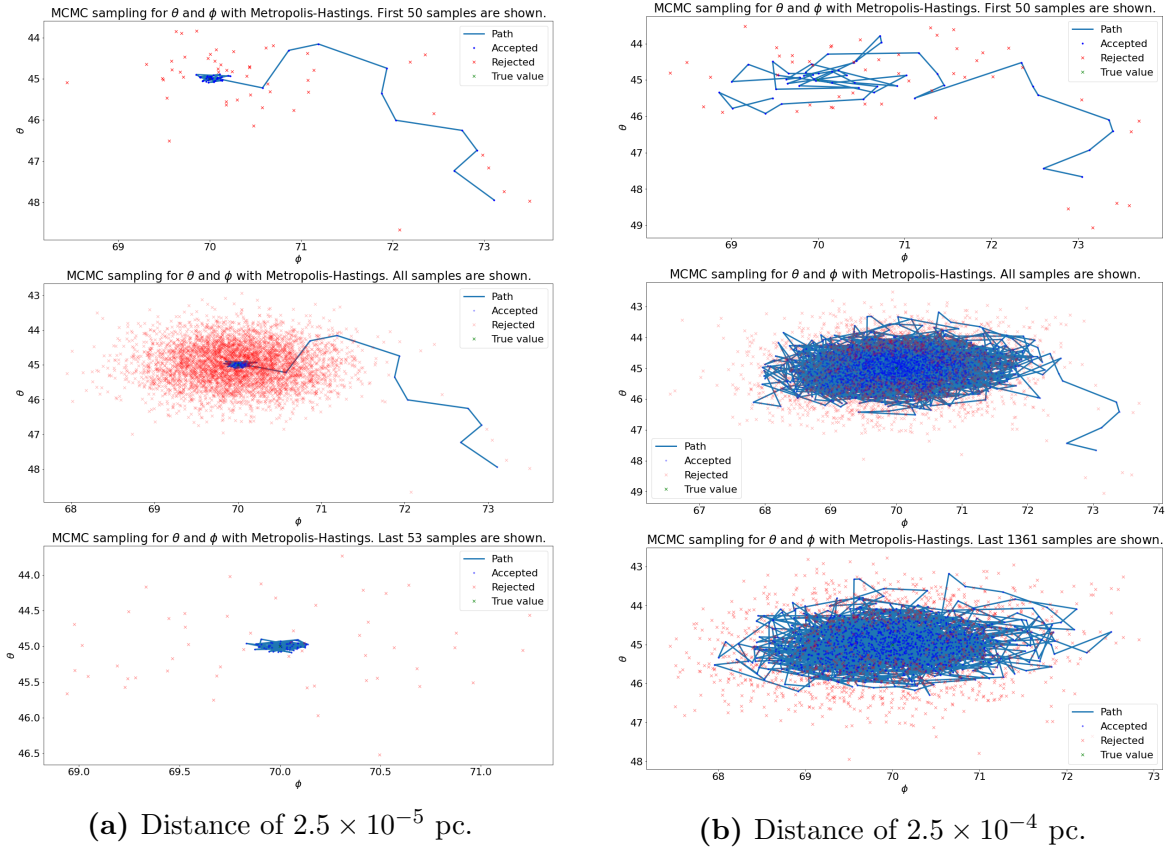


Figure 17: Comparison of the paths taken by MCMC for two separate sources that vary in separation r between the two masses of the source. Left figure seen with precise convergence and right seen with imprecise convergence. From top to bottom: First 50 samples, All samples, last 50% of samples. Red marker represents a rejected sample while the blue path traces the accepted samples.

4.2 Adding a second GW source

In reality, LISA will be sensitive to multiple gravitational wave sources simultaneously, so it is imperative that our algorithm can distinguish between different sources. Table 3 below shows the mean convergence value of the ecliptic coordinates when an additional wave is added to the strain signal. The first wave was set at $(45^\circ, 70^\circ)$, and the location of the second wave is displayed in the first column. Apart from θ and ϕ , the two waves have identical parameters.

Table 3: Simulation results for two identical GWs. MCMC Iteration = 5000, $f = 10\text{mHZ}$. Standard deviation displayed in brackets.

GW 2 coordinates [deg]	Angular distance R [deg]	Residual mean	All θ mean [deg]	All ϕ mean [deg]
(50, 65)	4.2	0 (0.0163)	47.596 (0.585)	67.293 (0.586)
(50, 65)	28.3	0 (0.0289)	118.488 (1.526)	96.867 (0.984)
(55, 60)	4.2	0 (0.0225)	44.738 (0.589)	69.997 (0.569)
(55, 60)	11.3	0 (0.0219)	55.255 (0.565)	59.968 (0.574)

In this experimental run, we varied the starting point of MCMC to see how the chain initialisation process affects the results of our data. In Table 3 above, R represents the angular distance between the first GW source at $(45^\circ, 70^\circ)$ and the initial starting point for the MCMC. The angular distance is calculated with

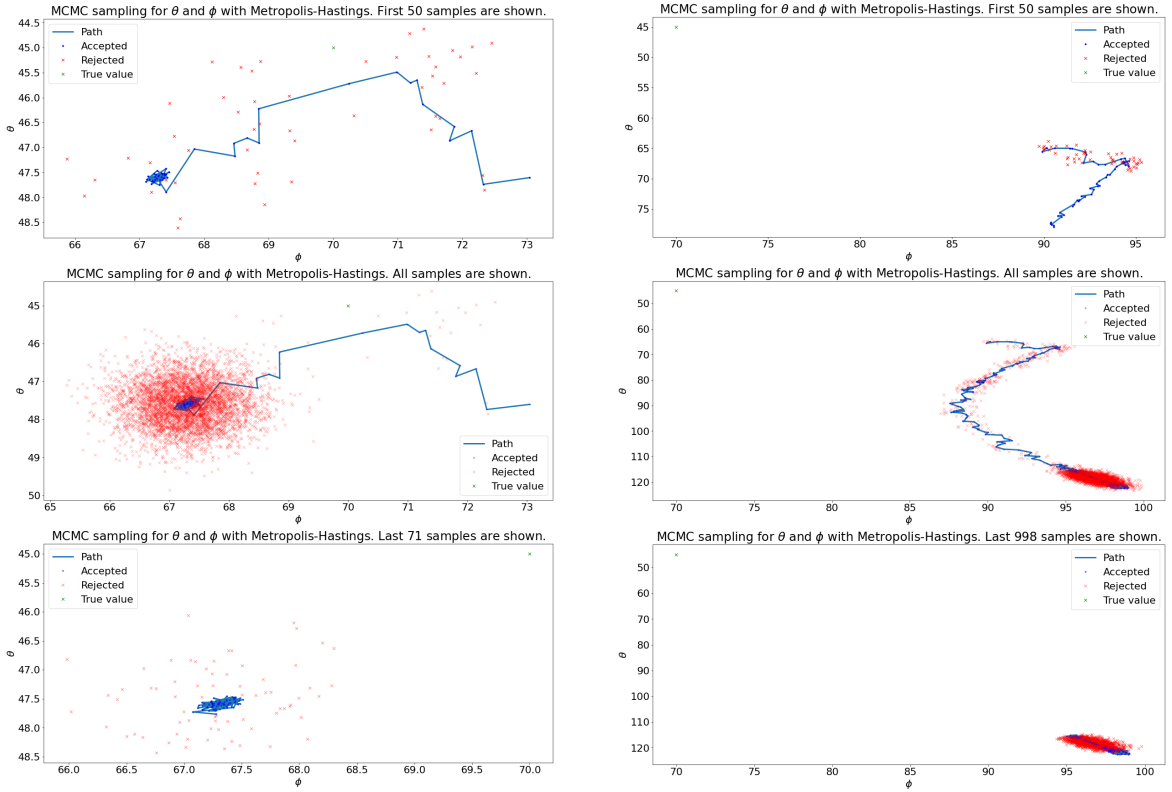
$$R = \sqrt{\theta_i^2 + \phi_i^2}, \quad (4.18)$$

where θ_i is the difference between 45° and the initial guess for θ , and ϕ_i the difference between 70° and the initial guess for ϕ .

The two sources were located at $(45^\circ, 70^\circ)$ and $(50^\circ, 65^\circ)$. The first run resulted in the chain converging at a point right between the two sources at approximately $[47.6(0.6)^\circ, 67.3(0.6)^\circ]$. During the second run the initial starting point for sampling was further away from both sources, and surprisingly the chain converged on the opposite side of the sky at approximately $[118.5(1.5)^\circ, 97(1)^\circ]$. The path of convergence for these two runs are displayed in Fig. 18 below.

When the second source was placed further away from the first, at $(55^\circ, 60^\circ)$, repeating the run with 4.2° angular distance from source 1 yielded results in which the chain converged

on the location of the first source with a standard deviation of no more than 0.6 degrees. Then the same process was repeated but with the initial starting point placed closer to the second source - at an angular distance 11.3° from source 1. The result was that the chain managed to converge on the location of the second wave with similar accuracy to the previous run.



(a) Second source at $(50, 65)$ with initial guess at angular distance 4.2° .

(b) Second source at $(50^\circ, 65^\circ)$ with initial guess at angular distance 28.3° .

Figure 18: Path taken to converge to the final solution from MCMC sampling. From top to bottom: First 50 samples, All samples, last 50% of samples. Red marker represents a rejected sample while the blue path traces the accepted samples. Green marker represents location of first GW source.

We discovered that, under the condition that the two waves shared the same characteristics, such as frequency, the algorithm became confused and caused the chain to converge right in the middle - if the sources are close together. The MCMC's chain initialising sample guess also plays a role in where the algorithm converges, and in some cases it may not be on either true locations.

We then ran an experiment where we kept the location of the two sources constant at $(45^\circ, 70^\circ)$ and $(55^\circ, 80^\circ)$, changing the angular distance from source 1 by a step of $\sqrt{2}$

degrees closer to source 2 each time, to see at what distance the chain begins to favour source 2 instead of source 1. Table 7 in the appendix lists the recovered positions. We started the initial guess at 1.4° angular distance from the first source, ending at 12.7° , which is equivalent to a 1.4° angular distance from source two. The results were then plotted in the scatter plot below in Fig. 19.

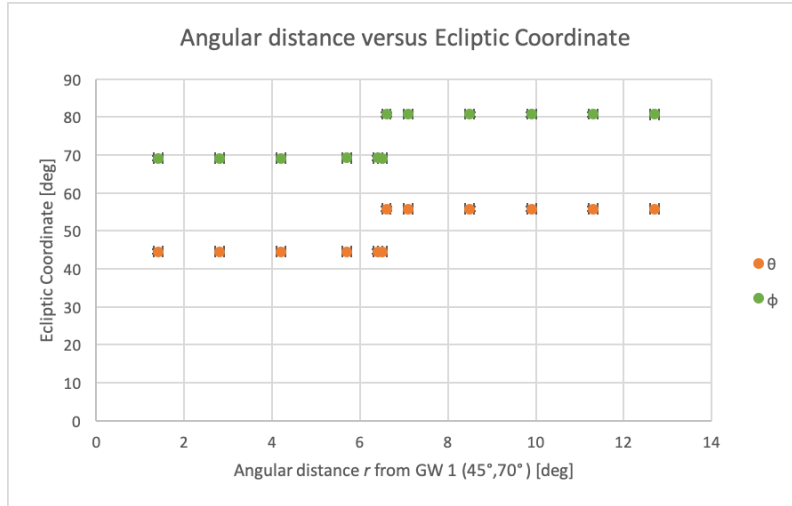


Figure 19: Scatter plot displaying the convergence of MCMC with varying angular distance from source one. Both θ (orange) and ϕ (green) converge towards the second source when the chain initiation guess is located half way to second source.

The results showed that at an angular distance of 6.6° , the MCMC begins to converge towards the second wave. This is just a little over half way distance between the two sources, with 6.6° translating to $(+4.7^\circ, +4.7^\circ)$ away from $(45^\circ, 70^\circ)$.

4.3 2 GW sources with varying frequency

With the modifications made to the program, we now run an experiment with two waves of different frequencies, to see if the algorithm is able to distinguish between the two without confusion. For simplicity, we focused on tuning the program to solve for 10 mHz GWs in the presence of another GW with different frequency. Table 4 below shows how the algorithm was able to recover the location of the first source with a standard deviation of no more than 0.6 degrees.

Table 4: Results of solving for GW at (45,70) with $f = 10$ mhz in the presence of a second GW at varying distance with varying frequency. Iteration = 5000. Standard deviation displayed in brackets

GW 2 coordinates [deg]	Angular Distance R [deg]	GW 2 Frequency [Hz]	All θ mean [deg]	All ϕ mean [deg]
(50, 65)	4.2	1	45.014 (0.580)	70.019 (0.575)
(45, 72)	1.4	1	45.014 (0.585)	69.989 (0.587)
(50, 65)	4.2	0.245	45.001 (0.580)	69.984 (0.576)
(50, 65)	4.2	1.005	45.003 (0.573)	69.995 (0.576)
(50, 65)	4.2	0.013	45.023 (0.583)	69.992 (0.578)

5 Discussion

5.1 Standard run

From the first standard run, we found that running the algorithm with 5000 iterations effectively allowed the MCMC to converge on the location of the GW source with an accuracy of approximately 0.6 degrees. From the trend in Table 1, it can be inferred that increasing the iteration will decrease the standard deviation of the Gaussian fitted to the samples taken by MCMC. However, because of the acceptance condition of the metropolis-hastings function defined by the residual of Eq. (2.16), the more the likelihood function updates, the more rejections MCMC should make as the acceptable steps taken become more restrained in area. Additionally, when comparing the results from all three different iterations, there was less than a 0.01 degree decrease in standard deviation for the posterior probability obtained from the histogram of all θ and ϕ samples. LISA is estimated to typically obtain the location of GW sources to an accuracy of one degree (Jennrich 2011). For EMRIs or black hole binaries with smaller frequencies, the localisation of these sources can improve to an accuracy of approximately 0.2 degrees. Following these guidelines for accuracy of position, the range in which our uncertainty falls into suggests compatible results. In order to preserve the efficiency of the program, we deem iterations over 5000 to be unnecessary in producing results with a much higher statistical significance.

The reason we chose to focus of data from the posterior probability obtained from all samples, and not just accepted samples, is because of the larger standard deviation in the histograms of the accepted samples. In Fig. 13, it can be seen that even after having

removed the first 50% of the accepted values taken, a significant amount of burn-in values were still affecting the standard deviation. Figure 12 shows that both the rejected and accepted values of the last 50% of samples taken have already converged on the position, with samples taken all within a small range of a ± 3 degrees. The histograms of these values in Fig. 14 form a much more noticeable gaussian fit, which provides more statistically meaningful and accurate results with an acceptable standard deviation. Furthermore, each run of the algorithm prints a final value taken from the last elements of the accepted array. This value, however, cannot be used as the most precise location of the source as it provides no statistical significance. The mean value of the posterior density is statistically more reliant and accurate to the true location of the simulated source, which is exactly why we chose to utilise Bayes theorem in our analysis.

The systematic uncertainty of LISA is translated into a background noise in the system. Since the exact error is unknown, we stuck to an expected value of 0.01 radians. In order to demonstrate that our choice does not produce invalid or biased results, we plotted the standard deviation of the results for ecliptic latitude and longitudes for varying systematic uncertainties. The results displayed in Fig. 15 show how the relation between standard deviation and detector uncertainty is linear for small values below 0.25 radians. In the case that LISA's systematic error slightly deviates from the value we chose for this project, the results can be simply be linearly re-scaled.

With a single wave, we could also vary the amplitude by changing the distance between the two spiraling bodies. The path taken by the algorithm in Fig. 17 presents a visual understanding of how the standard deviation of θ and ϕ increases with increasing distance r . The path of convergence in Fig. 17a was concentrated within a smaller range than that of Fig. 17b, suggesting that the precision of the recovered ecliptic coordinates reduced with decreasing amplitude. This was also reflected in the increase in standard deviation in Table 6's results. Fig. 16 shows how the recovered position of the wave deviates from the true value for amplitudes smaller than 3.43×10^{24} . The reason for this is because a weaker amplitude corresponds to a weaker signal, and once the signal becomes smaller than the noise of the system ($S/N < 1$), the wave becomes impossible for the algorithm to distinguish from the noise. Table 2 exhibits how a source, which was indistinguishable by the algorithm for a systematic noise of 0.01 radians, became distinguishable once the noise of the detector was decreased to 0.005 radians. It should also be noted that the standard deviation of the positions decreased as well. This tells us, not only is the algorithm used going to determine how effectively we can recover information about GWs with LISA, but the technical design of LISA and its detector noise level also will determine which waves can be successfully studied.

5.2 Two sources

When a second GW of the same frequency was added to the strain of the detector, the main conclusion made was that the MCMC confused two sources, which resulted in the

chain converging between the two sources. This is most likely the case for two identical sources with overlapping Gaussians when the sources are relatively close together in ecliptic coordinates. The rejected samples in Figure 18a form a Gaussian with a range of approximately 4 degrees (and this appears to also be the case for other runs). For the two sources to be 5 degrees apart in both θ and ϕ suggest a large likelihood of this overlapping scenario. When the second source was placed further from the first, we noticed there was no longer a confusion in the location of the sources, with the chain initialisation guess being the determiner for which source the chain converged on. In this case, the Gaussians of the two sources were assumed to not have overlapped, and once a sample taken by the Markov chain fell within the range of either Gaussians, the chain would converge to the center of said gaussian distribution.

An additional experiment was conducted where we placed two sources at an angular distance of 14.1 degrees apart and moved the chain initialisation guess to see at what distance does the chain convergence change from the first source to the second. The results indicated that the shift in convergence occurred when the initial guess was placed ($4.7^\circ, 4.7^\circ$) away from the first source. This angular distance is just far enough from the normally distributed range of 4 degrees mentioned earlier so the initialisation of the chain is not systematically biased to converge on the first wave. This information can later be implemented to further modify the program to solve for multiple identical sources in the sky where a grid of chain initialisation guesses, spaced strategically apart for efficiency, can be executed one after the other to identify positions of the sources in the sky under the condition that their individual gaussian distributions don't overlap. The possibility of the chain converging on an anti-source needs to also be taken into account in order to not misidentify the true position of the source in the sky. Such a problem occurred in the third row of Table 3, where the chain converged far away from both sources, seemingly on the other side of the sky.

Anti-sources are a degeneracy in non-local parameter space and show up as a “false positive” GW source on the opposite side of the sky due to interdependence of parameter quantities (Barrellon-Vernay 2018). The metropolis-hastings function can mistake it for the true source when calculating the difference between computed and observed strain. The degeneracy occurs due to confusion between the sources on either side of the ecliptic equator from the Doppler modulation (Cornish & Larson 2003). A clue that can help identify convergence onto anti-sources is the standard deviation of the posterior probability. For a chain converging on the anti-source, the standard deviation is much higher than for a chain converging on the true value. We see in Table 3 that the standard deviation of the ecliptic latitude in the case of an anti-source is approximately three times larger, and two times larger for the ecliptic longitude. This, however, cannot be used as a standard indicator of false convergence. Anti-sources are a common problem in low resolution calculations where the difference between the true and anti source cannot be distinguished. However, even for high resolution calculations, anti-sources can become an issue when there are multiple sources in the sky simultaneously. One possible solution is to use multi-messenger astronomy, whereby one source is studied using disparate signals - whether that be looking

at the electromagnetic radiation or cosmic rays.

For waves of the same frequency, it is apparent that the algorithm confuses two closely located sources as one single source. This was already predicted by the ESA (Jenrich 2011) as a possible problem, but according to the space agency, LISA will not have an issue in deciphering between sources of different frequencies. We put this to the test in the second modification to the program, where an extra parameter for frequency was added to the GW simulator. For a strain containing the information of two sources of varying frequency, the algorithm was able to “tune” itself to search for sources of a specific frequency, enabling the second wave to be completely invisible. In the second row of Table 4, we created a condition where the Gaussian of the two sources were overlapping. The angular distance between the sources was 1.4 degrees. Even then, the algorithm was successful in recovering the true value of the first source.

In the future, it is possible to expand the GW simulator to choose values for each parameter, and then using MCMC to solve for all 7 of them. It is also crucial to investigate how the algorithm reacts when more GW sources of different parameters are added to the strain. By creating a grid of points for chain initialisation guesses, the program can be run to sample regions of the sky individually to examine how many source positions can be accurately recovered.

5.3 Alternative

The Markov chain works especially well for locating the position of a GW source, however, there are other methods of stochastic sampling with computer algorithms that produce the same - if not better - results. An alternative method to MCMC is ‘nested sampling’. Nested sampling algorithms, like `Multinest`, works similarly to MCMC by sampling for the evidence $P(\text{data})$, in Eq. (2.15) of Baye’s theorem. Nested sampling is known for its flexibility when working in high-dimensional parameter space and is easy to expand upon for equations like the likelihood function (Smith et al. 2020). From this, we can assume that expanding our GW simulation to all 7 parameter spaces and solving for each will not pose a challenge when using nested sampling techniques. An additional advantage of nested sampling is its ability to undergo simultaneous estimate parameters and model selection due to posterior inference being a byproduct of its calculations (Feroz et al. 2019). In a study investigating the efficiency of parallel nested sampling, it was concluded that the method not only produced accurate results, but also reduced the running time of GW solving algorithms significantly (Feroz et al. 2019). Based on the results of such papers, it is worth looking into whether nested sampling will significantly improve the efficiency of recovering GW parameters for LISA compared to the traditional MCMC approach.

6 Conclusion

Gravitational Waves and the information they contain will be able to challenge the validity of Einstein's theory of General Relativity in a way that we were unable to study in the past. LISA will give astrophysicists the tools and data required to challenge our understanding of the Universe and its massive stellar objects. Unlike photons from early stages of the universe that were absorbed and re-emitted by plasmas of hot ions, GWs can travel without interference from matter. For this reason, GWs that originated from a time of high redshift may also provide us with information otherwise unobtainable about the early formation of our Universe. Furthermore, efficient use of stochastic sampling such as MCMC and Bayesian inference can also be broadly implemented in other aspects of astronomy.

We were able to effectively implement an MCMC algorithm to the GW simulator for this project, which allowed us to solve for the position of GW sources in the sky based on their strain signal. Implementing techniques of Bayesian inference within the Metropolis-Hastings function - such as setting up priors and conditions of acceptance - we were able to extract the mean ecliptic longitude and latitude of the source to an accuracy of approximately ± 0.6 degrees. Changing the systematic error demonstrated the effect the detector's accuracy has on the uncertainty of the source position. Adding a second wave was the first step in setting up a realistic simulation where the strain contains information about more than one GW, concluding that a second wave can confuse the algorithm to converge on false positions in the sky. Initial conditions of the MCMC, such as the chain starting point, were also explored to demonstrate how the results can depend on manual choices made when using the program. We then expanded the parameters of the algorithm, increasing the dimensions of the problem. By incorporating frequency, we were able to confirm sources of different frequencies do not interfere with each other.

This thesis has explored how the Markov Chain Monte Carlo stochastic sampling algorithm is effective in recovering the position of GW sources. Further development of the program combined with the use of a high-performance computer will undoubtedly result in a more sophisticated algorithm for recovering information about the GWs detected by LISA. In this paper, we have not considered the effects of stochastic GWs on the detector's accuracy for source localization, nor have we explored all parameter spaces of a GW signal. Further studies made about this project should contain an exploration of all 7 parameter spaces in a GW. We briefly investigated how a variation in amplitude affects the effectiveness of the program, but more variations with all parameters need to be studied. It is not uncommon to explore alternative stochastic sampling methods such as MultiNest, to see how an alternative program compares to the MCMC.

7 Appendices

Table 5: The variation of σ in All θ and All ϕ with changing uncertainty in the detector

Detector uncertainty [rad]	σ for all θ [deg]	σ for all ϕ [deg]
0.005	0.564	0.586
0.01	0.582	0.589
0.05	0.715	0.757
0.1	0.958	1.046
0.25	1.559	2.136

Table 6: For a source at (45,70) with $f = 10$ mhz, the distance r between the two stellar bodies is increased to see its affect on the convergence of MCMC. Iteration = 5000. Standard deviation displayed in brackets

Distance r [pc]	Residual mean	All θ mean [deg]	All ϕ mean [deg]
2.50×10^{-5}	0 (0.0002)	44.981 (0.579)	69.985 (0.598)
1.25×10^{-5}	0 (0.0005)	44.986 (0.578)	69.991 (0.584)
1.25×10^{-4}	0 (0.0002)	45.002 (0.725)	70.014 (0.747)
2.50×10^{-4}	0 (0.0004)	45.029 (0.924)	69.988 (1.088)
5.00×10^{-4}	0 (0.0005)	44.927 (1.263)	70.105 (1.649)
7.50×10^{-4}	0 (0.0010)	48.106 (9.600)	42.107 (16.76)
1.50×10^{-3}	0 (0.0005)	21.155 (9.229)	31.1019 (64.823)
2.5×10^{-3}	0 (0.0003)	115.87 (73.193)	100.859 (22.652)

Table 7: Results of two GW sources at $(45^\circ, 70^\circ)$ and $(55^\circ, 80^\circ)$ respectively, with $f = 10$ mhz. Angular distance between the first source and chain initiation guess is varied. Iteration = 5000. Standard deviation displayed in brackets.

Angular Distance R [deg]	All θ mean [deg]	All ϕ mean [deg]
1.4	44.549 (0.579)	69.166 (0.589)
2.8	44.566 (0.582)	69.148 (0.580)
4.2	44.561 (0.568)	69.138 (0.595)
5.7	44.567 (0.578)	69.176 (0.588)
6.4	44.554 (0.575)	69.179 (0.573)
6.5	44.571 (0.580)	69.164 (0.575)
6.6	55.801 (0.599)	80.733 (0.580)
7.1	55.794 (0.581)	80.715 (0.574)
8.5	55.792 (0.588)	80.711 (0.580)
9.9	55.774 (0.581)	80.731 (0.596)
11.3	55.772 (0.595)	80.739 (0.572)
12.7	55.784 (0.569)	80.706 (0.580)

8 References

References

- Aasi, J., Abadie, J., Abbott, B. P., et al. 2015, *Classical and Quantum Gravity*, 32, 115012
- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, , 116, 061102
- Addison, E. 2014, *All Graduate Theses and Dissertations*
- Amaro-Seoane, P., Aoudia, S., Babak, S., et al. 2012, *Classical and Quantum Gravity*, 29, 124016
- Barrellon-Vernay, R. 2018, *Simulating Gravitational Waves For LISA*, Report 1-20, Lund's University
- Box, G. E. & Tiao, G. C. 2011, *Bayesian inference in statistical analysis*, Vol. 40 (John Wiley & Sons)
- Caltech, L. 2021, *Sources and Types of Gravitational Waves*
- Chakrabarty, I. 1999, *Gravitational Waves: An Introduction*
- Chaudhuri, A. K. 2016, *Gravitational Wave for a pedestrian*
- Collaboration, L. S. 2021, *Introduction to LIGO Gravitational Waves Consortium*, L. 2021, Lisamission.org
- Cornish, N. J. & Crowder, J. 2005, , 72, 043005
- Cornish, N. J. & Larson, S. L. 2003, , 67, 103001
- Danzmann, K. 2000, *Advances in Space Research*, 25, 1129
- Eckart, A., Hüttemann, A., Kiefer, C., et al. 2017, *Foundations of Physics*, 47, 553
- Fairhurst, S. 2011, *New Journal of Physics*, 13, 069602
- Feroz, F., Hobson, M. P., Cameron, E., & Pettitt, A. N. 2019, *The Open Journal of Astrophysics*, 2, 10
- Giampieri, G. 1998, in *Second Edoardo Amaldi Conference on Gravitational Wave Experiments*, ed. E. Coccia, G. Veneziano, & G. Pizzella, 286
- Hilborn, R. C. 2019, *GW170814: Gravitational wave polarization analysis*
- Hough, J., Rowan, S., & Sathyaprakash, B. S. 2005, *Journal of Physics B Atomic Molecular Physics*, 38, S497

- Jennrich, O. 2011, Frequently asked questions
- Kunin, D. 2018, Bayesian Inference
- Moukarzel, J. 2019, From Scratch: Bayesian Inference, Markov Chain Monte Carlo and Metropolis Hastings, in python
- NASA. 2021, LISA - Laser Interferometer Space Antenna - NASA Home Page
- Ranalli, P., Hobbs, D., & Lindegren, L. 2018, , 614, A30
- Ribeiro, M. I. 2004, Institute for Systems and Robotics, Lisboa, Portugal
- Robert, C. P. 2015, The Metropolis-Hastings algorithm
- Robson, T., Cornish, N. J., & Liu, C. 2019, Classical and Quantum Gravity, 36, 105011
- Rubbo, L. J., Cornish, N. J., & Poujade, O. 2004, , 69, 082003
- Smith, R. J. E., Ashton, G., Vajpeyi, A., & Talbot, C. 2020, , 498, 4492
- Suvorov, A. G. 2021, , 503, 5495
- Sweetser, T. H. 2005, Classical and Quantum Gravity, 22, S429
- Trotta, R. 2008, Contemporary Physics, 49, 71