

Gravitational Wave Signal Extraction from LISA

by

Matthew Smith

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

BACHELOR OF SCIENCE

in

The Faculty of Science

(Physics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2023

© Matthew Smith 2023

Abstract

Ever since their existence was confirmed in late 2015, gravitational waves have been an invaluable tool for humanity to probe the universe. In order to examine gravitational wave sources never before detected, a new gravitational wave observatory is being situated in outer space. Known as the Laser Interferometer Space Antenna (LISA), this detector will host data streams different from any current generation detector. In preparation for LISA's launch, gravitational wave extraction has been done on mock data sets. These data sets include a variety of artefacts including large, transient signals from technical malfunctions. Parameter estimation of the gravitational wave sources within these data sets has been performed. Because previous methods have shown biases in the amplitude of the gravitational wave source, the artefact removal methods are scrutinized. Specifically, the effect of the roll-off and window length of the window functions used to remove artefacts are studied. This is done using two different gravitational wave signal models: one accounting for the applied window function and one without. It was found that the optimal window parameters for amplitude bias mitigation differed between the two signal models. For the unmodified model the optimal parameters were found to be 1000 and 0 for window length and padding, respectively. For the modified model, the optimal parameters were found to be 3000 and 0 for window length and padding, respectively. In any case, parameter estimation using the modified signal templates generally improved amplitude bias mitigation, but did not remove it entirely. Future work including methods for optimization and further bias mitigation have been proposed.

Preface

Chapter 3.1 is based on work conducted by Kye Emond under Dr. Scott Oser's supervision. This chapter details his contributions on creating detection methods for glitches, designing glitch removal techniques with Tukey window functions, and utilizing the F-statistic and Monte Carlo Markov Chain sampling for parameter estimation of the galactic binary gravitational wave sources in Spritz. He also implemented an analysis package in Python necessary to carry out many of the above tasks.

Table of Contents

Abstract ii

Preface iii

Table of Contents iv

List of Tables vi

List of Figures vii

Acknowledgements xii

1 Introduction 1

 1.1 Gravitational Waves 1

 1.2 Detecting Gravitational Waves 3

 1.3 Laser Interferometer Space Antenna 5

 1.4 Glitches 7

2 Theory 11

 2.1 Simplified Parameter Estimation 11

 2.2 Data Specifics 14

3 Methods 16

 3.1 Glitch Detection 16

Table of Contents

3.2	Artefact Removal	18
3.2.1	Tukey Window Creation	18
3.2.2	Window Function Study	22
4	Results	24
4.1	3D F-Statistic Landscapes	24
4.1.1	Unmodified F-Statistic	24
4.1.2	Modified F-Statistic	26
4.2	Source Detection Criteria	28
4.3	Modified F-Statistic Efficacy	29
5	Discussion and Analysis	33
	Bibliography	35
	Appendix	
A	First Appendix	39

List of Tables

2.1	Summary of the eight parameters relevant for sinusoidal GW signals. Note that A is the Pythagorean addition of A_+ and A_\times , and physically encodes various properties of the system, including mass.	12
A.1	Summary of the best-fit amplitude values as depicted in 4.5. The window parameters used for this calculation is 2000 and 0 for window length and padding, respectively. .	40

List of Figures

1.1	A visual depiction of the strain of a GW. As the phase changes, the GW stretches or squeezes different directions of space. The plus component h_+ modulates in a plus shape, whereas the cross component h_\times traces an x-shape over time. Here the wave-vector is orthogonal to the plane of the page. Figure reproduced from [10].	2
1.2	Sample schematic of laser interferometry apparatus for GW detection. Laser light is split, sent down both arms, recombined, and measured by the photodetector. In the case of LIGO, the arms are 4 kilometers long. Figure adapted from [1].	3
1.3	A schematic of the LISA detector. LISA is depicted by the red triangular formation of three spacecraft (black dots and circles). Trailing the Earth by around 50 million km, LISA's centre of mass sweeps out an orbit of radius 1 AU around the Sun (solid line). The LISA constellation tumbles as it rotates around the Sun, so the path of one of the spacecraft is also shown (dotted line). The LISA constellation's 2.5 million km arms are magnified by 15 times here for illustrative purposes. This figure is from [3].	5

1.4	The bandwidth and detection sensitivity of LISA, Advanced LIGO, and LIGO. Advanced LIGO is a second-generation improvement upon the original LIGO design. Various characteristic GW sources are present for each of the frequency bands. Any sources with characteristic GW strain amplitude above the LISA or LIGO curves are detectable by that respective detector. The black hole binaries, extreme mass ratio inspirals (EMRIs), and neutron star binaries are represented by lines depicting their frequency and strain amplitude evolution over time. Sources of relatively constant GW emission, like verification galactic binaries (GBs) and supernovae, are depicted by dots. Note that LISA's bandwidth does not compete with that of the LIGO family of detectors, and both parties are able to observe GWs from different phenomena. Figure adapted from [6].	6
1.5	Four examples of time domain signals in sample LISA data. (a) depicts a pure GW signal from several continuous wave sources. (b) depicts a reasonably sharp and loud glitch. (c) also encodes a glitch, but of a smaller amplitude than that of (b). (d) is simply stationary Gaussian noise. Note the qualitative similarities between panels (b), (c), and (d). Despite carrying juxtaposing information, the time series of (b), (c), and (d) are almost indistinguishable. While glitches can be as obvious as in (b), they can also be hard to tell from other, desirable signals.	9
1.6	Amplitude spectral density (ASD) of example data from LISA. The red line depicts LISA's response to continuous wave GBs in the absence of noise and other artefacts. The large concentration of peaks in the 10^{-4} to 10^{-2} Hz range correspond to GW signals from GBs in LISA's bandwidth. The blue line depicts LISA's response with no artefact mitigation. Note that these artefacts completely drown out the GW information. The sharp dips in power in frequencies beyond 10^{-2} Hz are attributed to LISA's deteriorating sensitivity at larger frequencies.	10

3.1	A sample glitch in the time series data. The black dashed line represents the threshold for amplitude statistic violation. The region in which the amplitude and quadratic statistic are violated are denoted by the red and blue dotted lines, respectively. Because the amplitude and quadratic statistic must be simultaneously violated for a valid glitch trigger event, the leftmost blue and rightmost red line define the glitch detection window. The data outside the glitch detection window consists of Gaussian noise and GW signals.	19
3.2	The time series of 3.1 after zeroing of the points within the glitch detection window. The purple dotted lines represent the intersection of the violation of the amplitude and quadratic statistics from 3.1. The solid blue line depicts one of the two hyper parameters in Tukey window function creation: padding. This is applied symmetrically around the glitch window. The data outside the glitch detection window consists of Gaussian noise and GW signals.	20
3.3	A sample window function created as a result of the process depicted in 3.1 and 3.2. The purple dotted lines represent the intersection of the violation of the amplitude and quadratic statistics from 3.1. The solid blue line depicts the padding parameter. Here, the distance between the leftmost blue and purple lines represents the half of the window length parameter. Note that the window length here is chosen to make the entire glitch detection region zeroed, but this is not generally true.	21
3.4	An example window function to be applied to artefact-ridden data. In this example, there would be two artefacts in the data localized at 1500 and 6500 seconds respectively. Note the tapered cosine around the zeroed regions, characteristic of the Tukey window.	22

4.1	Δ Amplitude as a function of window length and padding, for the lowest frequency source (1.81 mHz). Here Δ Amplitude is defined as the difference between the $\log(\text{true amplitude})$ and $\log(\text{best-fit amplitude})$ for this source, as determined by the unmodified F-statistic. There is a linear trend with respect to increasing window length, with a smaller dependence on padding. Minimal Δ Amplitude is desirable, so the ideal window parameters are situated at the bottom of the linear trend. . . .	25
4.2	Δ Amplitude as a function of window length and padding, for the lowest frequency source (1.81 mHz). Here Δ Amplitude is defined as the difference between the $\log(\text{true amplitude})$ and $\log(\text{best-fit amplitude})$ for this source, as determined by the modified F-statistic. There is a linear trend with respect to increasing window length, with a smaller and opposite dependence on padding. Minimal Δ Amplitude is desirable, so the ideal window parameters are situated at the bottom of the linear trend.	27
4.3	The number of sources detected as a function of window length for the unmodified F-statistic. Padding is held constant at 0. The number of detected sources is relatively constant in the mid 30s for small window lengths. There is a significant drop in sources detected beyond a window length of 2500.	28
4.4	The number of sources detected as a function of window length for the modified F-statistic. Padding is held constant at 0. The number of detected sources is relatively constant in the mid 30s for large window lengths. There is a significant drop in sources detected before a window length of around 1800.	29

4.5	<p>ΔAmplitude for the 36 sources in the data set. Here ΔAmplitude is defined as the difference between the $\log(\text{true amplitude})$ and $\log(\text{best-fit amplitude})$. The sources are ordered in terms of increasing frequency. The best-fit amplitude was determined by maximizing the F-statistic when using a padding of 0 and window length of 2000. This was done for both the unmodified (blue dots) and modified (red dots) cases. Points closer to the horizontal 0 line indicate a smaller amplitude bias. Values are tabulated in the appendix A.</p>	30
4.6	<p>The difference in $\Delta\text{Amplitude}$ between the unmodified and modified F-statistic cases for the 36 sources in the data set. Here ΔAmplitude is defined as the difference between the $\log(\text{true amplitude})$ and $\log(\text{best-fit amplitude})$. The best-fit amplitude was determined by maximizing the F-statistic when using a padding of 0 and window length of 2000. The sources are ordered in terms of increasing frequency. Points above the the horizontal 0 line indicate where the modified case mitigates amplitude bias better than the unmodified case, whereas points below this axis indicate the opposite.</p>	31

Acknowledgements

I would like to thank my supervisor Dr. Scott Oser for his patience and invaluable guidance for both the theory and practical implementation of this project. I am grateful for the opportunity to work on a collaboration as impressive and influential as the LISA mission. I would also like to thank Dr. Jess McIver for helping me gain access to various LISA resources and a Compute Canada allocation, and to Kye Emond for laying a significant groundwork for this project. Special thanks to my dear friend Ryan Kaufmann for proofreading, and to Willow Benitz for providing moral support throughout this project.

Chapter 1

Introduction

1.1 Gravitational Waves

Long before the Laser Interferometer Gravitational-Wave Observatory (LIGO) first recorded gravitational waves (GWs) in late 2015, the predicted existence of GWs from Einstein's theory of general relativity excited physicists and astronomers across the globe. However, ever since LIGO's pivotal work in the field, the true gravity of GW astronomy was unearthed. There are a few advantages GW astronomy has over its electromagnetic counterparts. One of the main advantages of GW astronomy is the vast cosmological reach it permits [11]. Most electromagnetic telescopes measure the power of a signal, which drops off as $1/r^2$. GW detectors instead measure the squeezing and stretching of space-time related to the amplitude of a GW, which drops off as $1/r$. Moreover, GWs travel through space-time at the speed of light, effectively undisturbed by any obstacles [11]. This allows GW observatories to probe parts of the universe in which light is otherwise obstructed. Hence, a variety of interesting astronomical systems including pulsars, supernovae, black hole and neutron star mergers, and galactic binary (GB) systems can be scrutinized via their emission of GWs [6].

One of the simplest models for a GW is simply a ripple in the fabric of space-time. Because these waves reflect deviations from the otherwise flat fabric, they have a profound effect on the local region of space-time that they interact with. Specifically, the waves simultaneously squeeze and stretch physical space perpendicular to their direction of propagation [11]. This phenomenon is generally denoted as a strain h . Analogous to their electromagnetic counterparts, GWs also have associated polarizations. However, because a GW simultaneously squeezes and stretches space-time

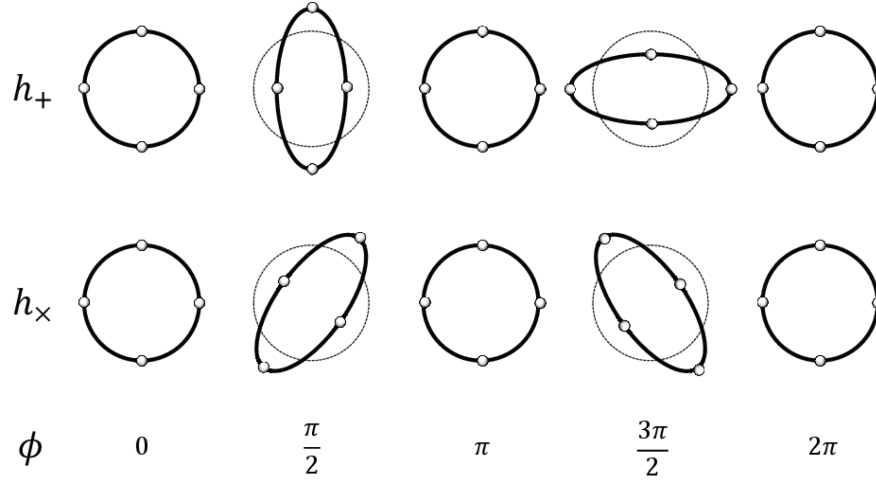


Figure 1.1: A visual depiction of the strain of a GW. As the phase changes, the GW stretches or squeezes different directions of space. The plus component h_+ modulates in a plus shape, whereas the cross component h_\times traces an x-shape over time. Here the wave-vector is orthogonal to the plane of the page. Figure reproduced from [10].

in orthogonal directions, a conventional rotation by $\frac{\pi}{2}$ is insufficient to define a new polarization. Instead, the two polarizations differ by a rotation of $\frac{\pi}{4}$, and are suitably named plus and cross, respectively [12]. This is depicted in 1.1.

It then follows that any GW can be mathematically described as a linear combination of the components of the strain corresponding to the plus and cross polarizations [8],

$$\mathbf{h}(t) = h_+(t)\boldsymbol{\epsilon}_+ + h_\times(t)\boldsymbol{\epsilon}_\times \quad (1.1)$$

$$\boldsymbol{\epsilon}_+ = \cos(2\psi)\hat{e}_+ - \sin(2\psi)\hat{e}_\times \quad (1.2)$$

$$\boldsymbol{\epsilon}_\times = \sin(2\psi)\hat{e}_+ + \cos(2\psi)\hat{e}_\times. \quad (1.3)$$

Here ψ is a polarization angle parameter, and $\hat{e}_{+/\times}$ denote unit tensors in the plus or cross directions, respectively. Of course, the functional forms of $h_+(t)$ and $h_\times(t)$ are dependent on the specifics of the astronomical system that created the GW.

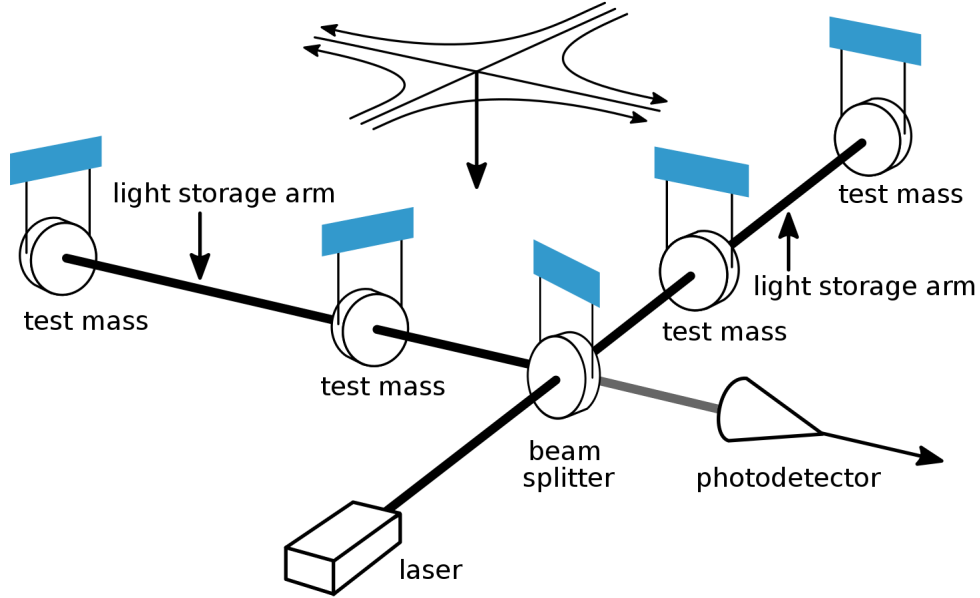


Figure 1.2: Sample schematic of laser interferometry apparatus for GW detection. Laser light is split, sent down both arms, recombined, and measured by the photodetector. In the case of LIGO, the arms are 4 kilometers long. Figure adapted from [1].

1.2 Detecting Gravitational Waves

In order to extract the valuable information that GWs carry, they must first be measured. In order to do this, GW detectors measure the strain of a GW via its squeezing and stretching of physical space or distances [11]. Specifically, the detectors implemented today use a method known as laser interferometry. As shown in 1.2, the process involves firing a high-powered laser towards a beam splitter. This optical device splits the laser and sends it down two equal length paths in an L shape. The two components of the beam are then reflected at the end of their respective arms. The beams are eventually recombined, and the reconstructed beam is measured by a photo-detector. Now, if a GW was incident on the detector during this process, a difference in length would be introduced between the two arms. This would cause a difference in phase between the two components of the beam, subsequently resulting in a detectable interference pattern [11].

The complexity of this apparatus is largely due to the vanishingly small impact GWs have

on their surroundings. Even from the extremely high energy systems that are of interest, the impact their GWs have on the detector's length is approximately 1 part in 10^{20} . This means that the detectors are required to detect length differences fractions the size of an atom. Put into more digestible terms, this is comparable to being able to measure the fluctuations in the distance between Earth and Alpha Centauri— which is approximately 4.3 light years away— on the order of the width of a human hair. It then follows that the calibration of these detectors must be extremely precise in order to collect meaningful data.

One of the most physically important aspects of the detector is the arm-length itself. Because the arm-length of a detector impacts the wavelength of the GWs detectable, its frequency bandwidth is largely dependent on this length [6]. In the case of LIGO, the detector arms are 4 kilometers long, which equates to an effective frequency range of around 10Hz to 10KHz [11]. Because different sources are characterized by different frequency regimes, this allows current detectors like LIGO to detect GWs from sources like black hole and neutron star binaries, among others [12].

While the most famous example might be the United States' LIGO Hanford and Livingston, GW detectors now exist across the globe. Two notable examples include the European Gravitational Observatory (EGO/VIRGO) in Italy [7], and the Kamioka Gravitational Wave Detector (KAGRA) in Japan [4]. Moreover, LIGO-India and the Einstein Telescope are two proposed future GW observatories [6]. Creation of new GW observatories helps to create the global network of detectors which compliment each other to improve the resolution of the location of various GW sources [12].

While ground-based detectors have already made excellent strides in the field of GW astronomy, they face a number of issues that limit their detection. First of all, because they are necessarily extremely sensitive, detectors are prone to noise originating from human, atmospheric, and seismic activity [11]. Attempts have been made to mitigate these effects with KAGRA and the proposed Einstein Observatory by moving underground, but creating detectors significantly deeper is financially prohibitive [6]. Moreover, ground-based detectors are very limited in their physical size. The lasers in these apparatuses require a straight line-of-sight, so the Earth's curvature makes significant terraforming necessary for detector arms much longer than a few kilometers [12]. In order to

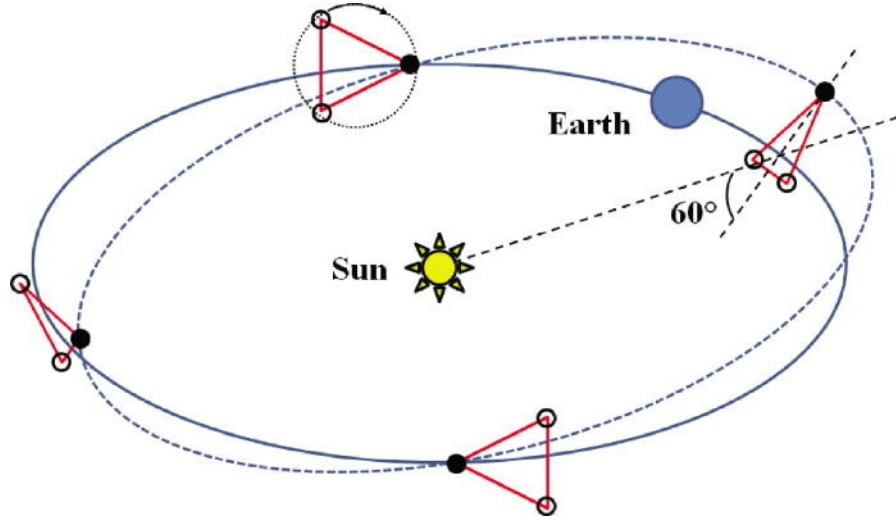


Figure 1.3: A schematic of the LISA detector. LISA is depicted by the red triangular formation of three spacecraft (black dots and circles). Trailing the Earth by around 50 million km, LISA’s centre of mass sweeps out an orbit of radius 1 AU around the Sun (solid line). The LISA constellation tumbles as it rotates around the Sun, so the path of one of the spacecraft is also shown (dotted line). The LISA constellation’s 2.5 million km arms are magnified by 15 times here for illustrative purposes. This figure is from [3].

probe a lower frequency regime, and hence a variety of different sources, detectors with enormous arm-lengths are required.

1.3 Laser Interferometer Space Antenna

Given the issues faced with ground-based detectors, space is a perfect candidate for expansion. Space is devoid of the seismic and atmospheric fluctuations that impair ground-based detectors, and the physical space required for the arm-lengths of a detector is practically limitless.

For this reason, the Laser Interferometer Space Antenna (LISA) project was born (1.3). LISA is a joint ESA and NASA project to create a space-based GW detector consisting of three identical spacecrafts that emit and measure laser responses in a triangular formation [2]. Planned for launch in the mid 2030s, LISA’s arms will be 2.5 million km long, allowing for an effective frequency detection range of 0.1mHz to 1Hz. This will allow LISA to observe massive black hole binaries

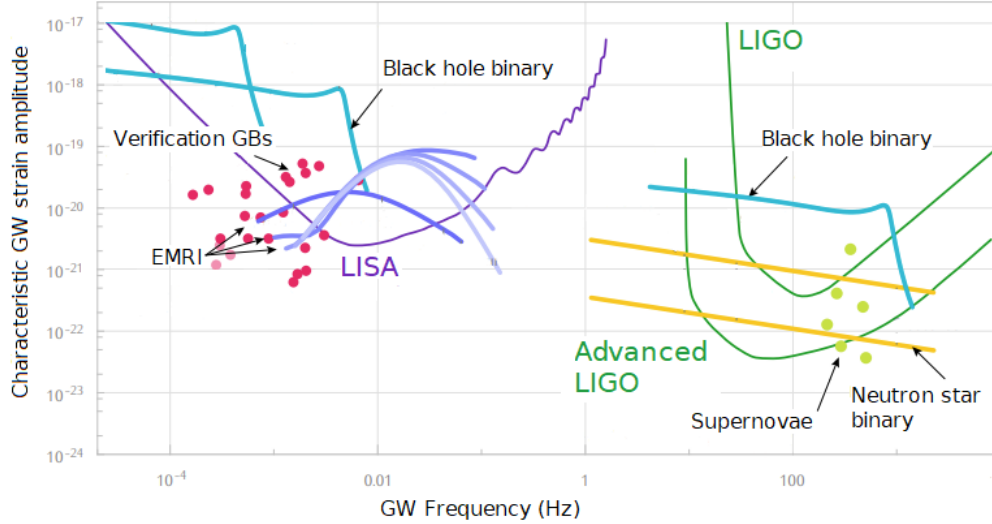


Figure 1.4: The bandwidth and detection sensitivity of LISA, Advanced LIGO, and LIGO. Advanced LIGO is a second-generation improvement upon the original LIGO design. Various characteristic GW sources are present for each of the frequency bands. Any sources with characteristic GW strain amplitude above the LISA or LIGO curves are detectable by that respective detector. The black hole binaries, extreme mass ratio inspirals (EMRIs), and neutron star binaries are represented by lines depicting their frequency and strain amplitude evolution over time. Sources of relatively constant GW emission, like verification galactic binaries (GBs) and supernovae, are depicted by dots. Note that LISA’s bandwidth does not compete with that of the LIGO family of detectors, and both parties are able to observe GWs from different phenomena. Figure adapted from [6].

(MBHBs), extreme mass ratio inspirals (EMRIs), and ultra-compact binaries [9]. Note that LISA’s frequency range does not compete with the frequency ranges of ground-based detectors; LISA will complement, not replace, ground-based detectors. Given that some GW sources increase in frequency over time, LISA can also be an invaluable tool to prognosticate important high-frequency merger events for its ground-based cousins to detect [12]. A better comparison of LISA and LIGO’s detection ranges can be found in 1.4.

Unfortunately LISA also faces some significant difficulties. Given that LISA will orbit the Sun, the differences in the positions of the spacecraft will necessarily cause a relative motion of up to 15 m/s between them [5]. While this introduces technical engineering challenges with

laser calibration, the larger issue with LISA's unequal arm lengths is the laser frequency noise that it introduces. The frequency of the lasers in GW detectors naturally fluctuate by several MHz. These fluctuations naturally cancel in equal-length ground detectors, but instead couple and completely drown out signals in LISA's detection chain [6]. In order to reconcile this otherwise catastrophic noise contribution, a more complicated scheme, known as time-delay interferometry (TDI), combines the signal output from the three spacecraft in post-processing to create data free from frequency noise. Specifically, this relies on properly time-shifting and linearly combining the independent laser measurements [15]. After this TDI treatment we are left with three data streams corresponding to the three side lengths of LISA's constellation: $X(t)$, $Y(t)$, and $Z(t)$. Additionally, when dealing with sinusoidal signals, it is convention to create so-called orthogonal modes because they simplify data analysis computations. Denoted $A(t)$, $E(t)$, and $T(t)$, these channels are simply linear combinations of $X(t)$, $Y(t)$ and $Z(t)$ [15],

$$A(t) = \frac{1}{\sqrt{2}}[Z(t) - X(t)] \quad (1.4)$$

$$E(t) = \frac{1}{\sqrt{6}}[X(t) - 2Y(t) + Z(t)] \quad (1.5)$$

$$T(t) = \frac{1}{\sqrt{3}}[X(t) + Y(t) + Z(t)]. \quad (1.6)$$

With these three channels, GW signal data from LISA can be extracted and analysis can be performed.

1.4 Glitches

As a test run for LISA's feasibility, a single spacecraft probe known as the LISA Pathfinder mission was launched in 2015 [6]. Because LISA Pathfinder is a single satellite that does not exchange laser links within an array, it is unable to detect GWs. However, the mission did answer vital questions about LISA's noise model and its subsequent sensitivity within its frequency bandwidth

[12]. Moreover, the data streams from Pathfinder exhibited glitches. Glitches are abnormalities in the data stream that are created by some, often unknown, technical issue with the apparatus. While their morphology is reliant on their origin, glitches are usually characterized as short burst signals with immense power. It is important to note that glitches are not a new discovery for the GW community. Ground-based detectors already experience their own host of glitches and research is actively being done to explore and scrutinize methods for their removal [11].

In order to better understand why glitches are so detrimental to the data streams, it is first important to understand typical glitch morphology. 1.5 depicts the differences in the time-domain of different signals, including glitches.

Moreover, because the noise model is simple in frequency space, all data analysis is done in this domain. As such, loud and sharp glitches in the time domain spread out their high-power over a wide range of frequencies upon Fourier transformation. As shown in 1.6, glitches drown out any signal from GWs, so their detection and removal is paramount for the LISA project.

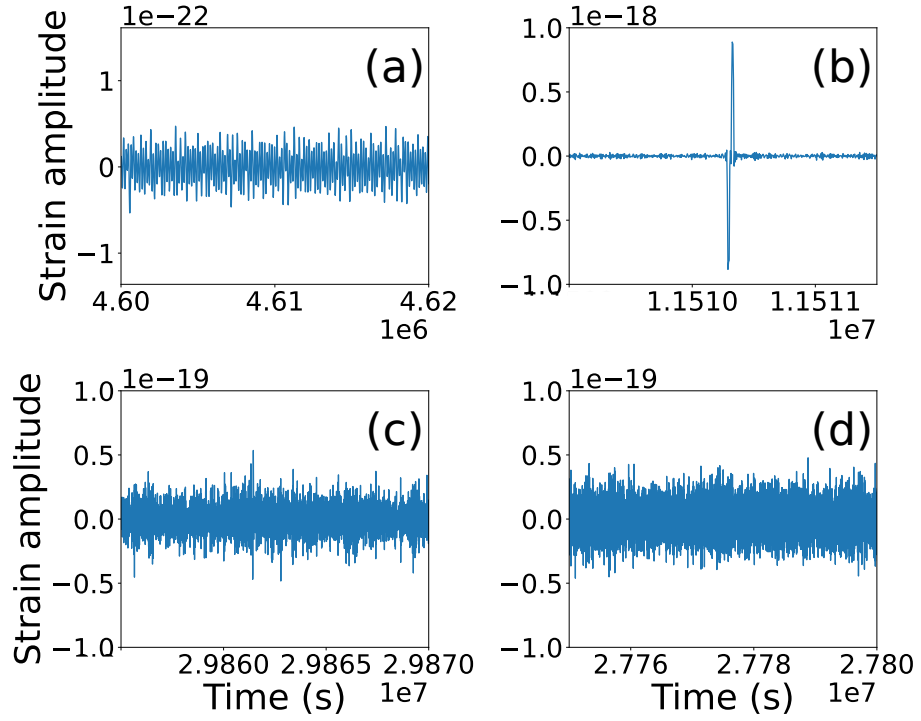


Figure 1.5: Four examples of time domain signals in sample LISA data. (a) depicts a pure GW signal from several continuous wave sources. (b) depicts a reasonably sharp and loud glitch. (c) also encodes a glitch, but of a smaller amplitude than that of (b). (d) is simply stationary Gaussian noise. Note the qualitative similarities between panels (b), (c), and (d). Despite carrying juxtaposing information, the time series of (b), (c), and (d) are almost indistinguishable. While glitches can be as obvious as in (b), they can also be hard to tell from other, desirable signals.

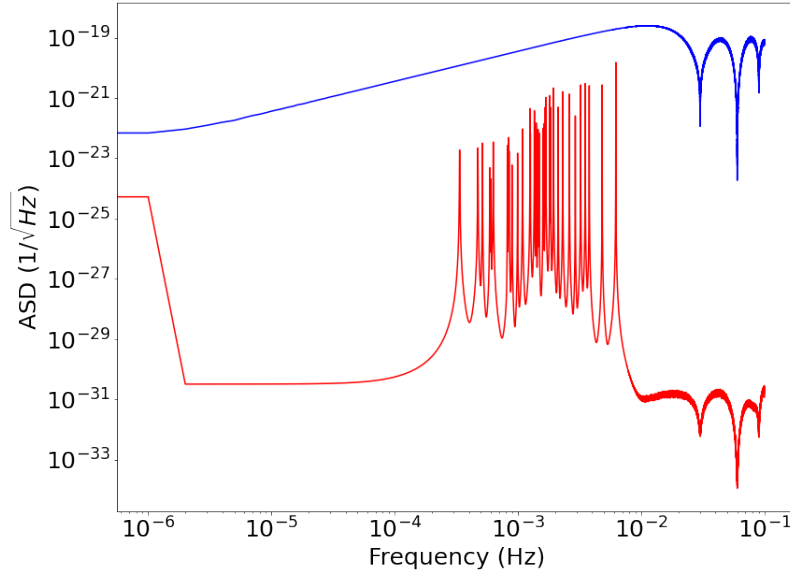


Figure 1.6: Amplitude spectral density (ASD) of example data from LISA. The red line depicts LISA’s response to continuous wave GBs in the absence of noise and other artefacts. The large concentration of peaks in the 10^{-4} to 10^{-2} Hz range correspond to GW signals from GBs in LISA’s bandwidth. The blue line depicts LISA’s response with no artefact mitigation. Note that these artefacts completely drown out the GW information. The sharp dips in power in frequencies beyond 10^{-2} Hz are attributed to LISA’s deteriorating sensitivity at larger frequencies.

Chapter 2

Theory

2.1 Simplified Parameter Estimation

As is the case with any GW signal analysis, a significant portion of this project is reliant on extracting the parameters of the sources encoded in our data streams. As such, considering a toy-model of a continuous wave sinusoidal model with a Gaussian noise background, this section will describe the mathematical and statistical framework necessary for parameter estimation from time series data.

First, understanding the relevant source parameters for GW strain is crucial. Building on the general form in equation 1.1,

$$\mathbf{h}(t) = \text{Re}[(A_+ \boldsymbol{\epsilon}_+ + e^{i\frac{3\pi}{2}} A_\times \boldsymbol{\epsilon}_\times) e^{i\Psi(t)}] \quad (2.1)$$

$$\Psi(t) = 2\pi f t + \pi \dot{f} t^2 + \phi_0 \quad (2.2)$$

$$A_+ \propto f^{\frac{2}{3}} (1 + \cos^2 \iota) \quad (2.3)$$

$$A_\times \propto f^{\frac{2}{3}} \cos \iota, \quad (2.4)$$

an equation for the strain of a relatively continuous wave source ($\dot{f} \ll 1$) is given [8]. Here, the parameters of interest are the frequency of the source f , its temporal derivative \dot{f} , its initial phase ϕ_0 , its inclination ι , and its amplitudes from the respective polarization components $A_{+/\times}$. Note from equations 1.2 and 1.3, $\boldsymbol{\epsilon}_{+/\times}$ encodes a polarization angle ψ in the strain as well. There are two additional source parameters that are also of interest. While not explicitly present in 2.1,

2.1. Simplified Parameter Estimation

Symbol	Parameter	Description
λ	Ecliptic latitude	With β , describes sky position of the source
β	Ecliptic longitude	With λ , describes sky position of the source
f	Frequency	The rate of revolution of the system
\dot{f}	Frequency derivative	The rate at which the frequency changes
A	Amplitude	Property of the wave
ι	Inclination	Angle of the tilt between the source and LISA
ψ	Polarization	Angle determining components of each polarization
ϕ_0	Initial phase	Property of the wave; dependent on t_0

Table 2.1: Summary of the eight parameters relevant for sinusoidal GW signals. Note that A is the Pythagorean addition of A_+ and A_\times , and physically encodes various properties of the system, including mass.

the ecliptic latitude and longitude of the source with respect to LISA’s position will also naturally have an impact on the measured GW signal. A summary of the relevant parameters are presented in 2.1.

Defining the parameters as θ , data as d , and using Bayes’ Theorem [14],

$$p(\theta|d) = \frac{p(d|\theta)p(\theta)}{p(d)},$$

an expression relating the parameters to the data is derived. Defining $p(\theta)$ as the prior $\pi(\theta)$, $p(d|\theta)$ as the likelihood function $\mathcal{L}(d|\theta)$, and $p(d)$ as the evidence \mathcal{Z} , the above equation can be rewritten,

$$p(\theta|d) = \frac{\mathcal{L}(d|\theta)\pi(\theta)}{\mathcal{Z}}. \quad (2.5)$$

Above, \mathcal{L} encodes the data and model information, π encodes previous knowledge of the parameters, and the evidence \mathcal{Z} acts as a normalization factor. $p(\theta|d)$, or the posterior, describes a probability distribution of the output parameters given the data. Here, the peak of the posterior is defined by the parameters that maximize \mathcal{L} [13].

Unfortunately, searching over the 8-dimensional parameter space for maxima of this likelihood function is necessarily prohibitive from a computational standpoint. As such, a clever method

utilizing a coarse F-statistic gridsearch followed by Monte Carlo Markov Chain (MCMC) sampling is utilized to determine the parameters.

This maximized likelihood function, called the F-statistic, can be maximized via a coarse grid search over parameter space. First, the dimensionality of the problem can be reduced from 8 to 3 dimensions. In order to do this, β , λ , f , and \dot{f} are held constant. For a given choice of these parameters, the complexity of the problem is reduced significantly. Maximizing the F-statistic becomes analogous to fitting the data to four different signal templates: a real and complex part for each polarization. As such, the remaining four parameters that maximize the F-statistic can be analytically computed [12]. Moreover, as mentioned previously, these sources have a relatively constant frequency, so it is safe to hold $\dot{f} \approx 0$ for the grid search.

With this, the entire sky can be parametrized by β and λ , and the frequency range can be chosen reasonably given LISA's bandwidth. Varying these 3 parameters and comparing their F-statistics gives an estimate for the parameters of the source with accuracy proportional to the grid spacing chosen. Any point on this grid that has an F-statistic value beyond the 70th percentile is considered a detected source.

There is an issue with solely relying on a grid search for computing the source parameters. Because the fineness of the grid is limited by computational resources and time, the grid search is prone to falling into local maxima of the F-statistic. Further, the F-statistic calculations only provide the best-fit parameters, but provide no information about their uncertainties. In order to characterize a confidence interval for the found parameters, a mapping of the posterior distribution is necessary.

As such, we utilize the estimates of the parameters from the grid search as a starting point for a Monte Carlo Markov Chain (MCMC) sampling routine. In this MCMC method, so-called walkers undergo a random walk through the entire 8-dimensional parameter space to map out the posterior distribution. The probability of moving to different points is characterized by the transition probability of the Markov Chain [14]. Due to the properties of a Markov Chain, it is necessary to include a burn-in phase before the walkers reach their stationary distributions. After

a sufficiently long run, the walkers will have created a mapping of the likelihood function, and the maxima can be extracted. Note that these walkers can also get trapped within local maxima, so it is important to run the MCMC with various different initial positions for the walkers [14].

Once the posterior distributions are extracted from the MCMC sampling, all 8 source parameters of the sinusoidal signal with associated uncertainties are now successfully extracted from the data.

2.2 Data Specifics

In preparation for LISA’s launch, the LISA consortium has created several mock datasets. One of which is known as the Spritz dataset. This dataset contains what LISA data *should* look like over the timescale of a year. Naturally, the dataset contains time and frequency series data streams from each of the $X(t)$, $Y(t)$, and $Z(t)$ channels. In terms of the contents of these streams, there are a few key characteristics to note.

Noise Model

First of all, the data contains general background noise. This noise is motivated by LISA’s proposed noise model. This model has been constructed via intricacies in its hardware, and supported by observations made by LISA Pathfinder. In future real data sets, unresolvable low-power background GW sources are also contained in this noise.

Gaps

The data also contains gaps. Like all GW detectors, LISA will likely face technical issues that cause periods of no data collection in the middle of observation runs. As such, these periods of time contain no information about incoming GW signals, and should be treated accordingly for data analysis. In terms of the actual data stream, gaps are data points in the time series represented by NaNs.

Glitches

As foreshadowed in Chapter 1.4, the Spritz data set also contains glitches. Modelled after those seen by LISA Pathfinder, these glitches are randomly distributed throughout the time series. The detailed discussion supplied in Chapter 1.4 applies precisely to the glitches found in Spritz.

GW Signals

Finally, the GW sources present are from well-documented verification GBs. As previously alluded to, unlike merger events, GB signals are characterized by relatively continuous sinusoidal waves in the time domain. Transferring to the frequency domain, these signals essentially resemble delta functions. Moreover, the parameters that describe the signals from the GBs are already well documented via electromagnetic astronomy. As such, if the task of removing artefacts is successful, the parameters from the extracted signal can be compared to their assumed true values. Holistically, using these GB sources as controls, LISA data analysis techniques can be perfected in preparation for eventual observation of unknown sources. There are 36 sources of varying signal-to-noise in the 10^{-4} to 10^{-2} Hz range.

For ease of analysis, three separate data sets are included: one with just the GW sources, one with the GW sources and the noise background, and one with everything listed above. Given the full dataset, methods to identify and remove the gaps and glitches can be used to extract solely the GW signal. Using the methods outlined in Chapter 2.1, parameters can be computed and compared to their known true values. Under this lens, gap and glitch removal methods can be scrutinized and adjusted.

Chapter 3

Methods

3.1 Glitch Detection

One of the most common types of glitches can be seen in panel (b) of 1.5. This type of glitch appears almost as a delta function in the time domain. This sharp spike will spread out over a wide range of frequencies upon Fourier transformation, entirely drowning out any comparatively small GW signal peaks. Fortunately, it is only these large glitches that cause the biggest issues. Keeping this in mind, a relatively crude method utilizing their amplitude and slope can be devised for their detection.¹

In order to differentiate between noise, glitches, and GWs, the autocovariance $K_{XX}(\tau = t_2 - t_1)$ of the noise is a useful property. An estimate for this value can be computed from the data after crudely removing any signal that strays too far from the expected 0 mean of the amplitude noise. This mitigates effects of the loudest glitches, allowing for an approximate estimate of K_{XX} . With this, assuming Gaussian noise with 0 mean, the amplitude A can be described by the PDF,

$$p(A) = \frac{1}{\sigma\sqrt{2\pi}}e^{-A^2/2\sigma^2}. \quad (3.1)$$

where $\sigma^2 = K_{XX}(0)$ by definition. Furthermore, the so-called quadratic statistic, $Q = A^2 + \frac{d}{dt}(A)^2$, is an invented parameter that encodes information about the slope of the amplitude. Given the previous assumptions about the Gaussian background noise, the A^2 random variable can be described by the gamma distribution,

¹This sub-chapter is based on prior work conducted by Kye Emond under Dr. Scott Oser's supervision.

$$A^2 \sim \Gamma(\frac{1}{2}, 2K_{XX}(0)) \triangleq \Gamma_{A^2} \quad (3.2)$$

Likewise, approximating $\frac{d}{dt}(A)^2$ with a finite difference scheme from the discrete data as $(A_{i+1} - A_{i-1})^2$, $\frac{d}{dt}(A)^2$ can be described by the gamma distribution,

$$\frac{d}{dt}(A)^2 \sim \Gamma(\frac{1}{2}, 4(K_{XX}(0) - K_{XX}(2))) \triangleq \Gamma_{\frac{d}{dt}(A)^2} \quad (3.3)$$

Given that the amplitude and slope should be independent random variables, their joint PDF is given by their product. With this, the PDF describing the quadratic statistic is derived,

$$p(Q) = \int_0^Q \Gamma_{A^2}(x) \cdot \Gamma_{\frac{d}{dt}(A)^2}(Q - x) dx. \quad (3.4)$$

In order to avoid removing GW signals, a glitch is deemed a glitch if and only if $|A| > a$ and $|Q| > q$, where a and q are hyper-parameters to be tweaked for a reasonable balance between data preservation and glitch removal. Hence, putting equations 3.1 and 3.4 together, the probability of finding a trigger event can be described,

$$P(\text{Trigger}) = P(|A| > a \wedge |Q| > q). \quad (3.5)$$

Finally, we can use the fact that LISA has three data streams to our advantage. First of all, the TDI T channel has been found to be less sensitive to GWs. As such, if a trigger event happens in the T channel, there is a higher probability that a glitch has been identified. Furthermore, a glitch affecting a single laser link should only impact two of the three TDI channels X , Y , and Z , whereas a GW should impact all three. As such, if a trigger occurs in exactly two-thirds of these channels there is also a higher confidence of that flag being a glitch. As such, the final probability of a glitch flag follows,

$$P(\text{Glitch flag}) = P(XYZ \text{ Trigger} \vee T \text{ Trigger}). \quad (3.6)$$

3.2 Artefact Removal

Once the parameters of the glitch detection chain are satisfactory, both the gaps and glitches need to be removed from the data. As previously mentioned, artefacts are relatively localized in the time domain, so it is ideal to remove them in this space. However, the removal of artefacts is a non-trivial task. A natural idea is to simply zero out the data where the artefact is localized to. However, this abrupt cutting violently alters the GW signal encoded in the data and can cause a smearing of the signal's power across the frequency spectrum, a phenomenon known as spectral leakage. In order to combat this issue, the artefact is zeroed out and then the neighboring time segments are multiplied by a smooth function known as a Tukey window. In order to get the desired result of minimizing spectral leakage while still removing the power of the glitches, window length and padding are two hyper-parameters that must be adjusted accordingly.

3.2.1 Tukey Window Creation

To start, a glitch is identified using the methods outlined in chapter 3.1. As shown in 3.1, a large glitch is located in the time series via simultaneous violations of the amplitude and quadratic statistic. Here, the glitch detection window is defined as the region between the inner blue vertical line and the inner red vertical line, as per the detection criteria outlined in the previous chapter.

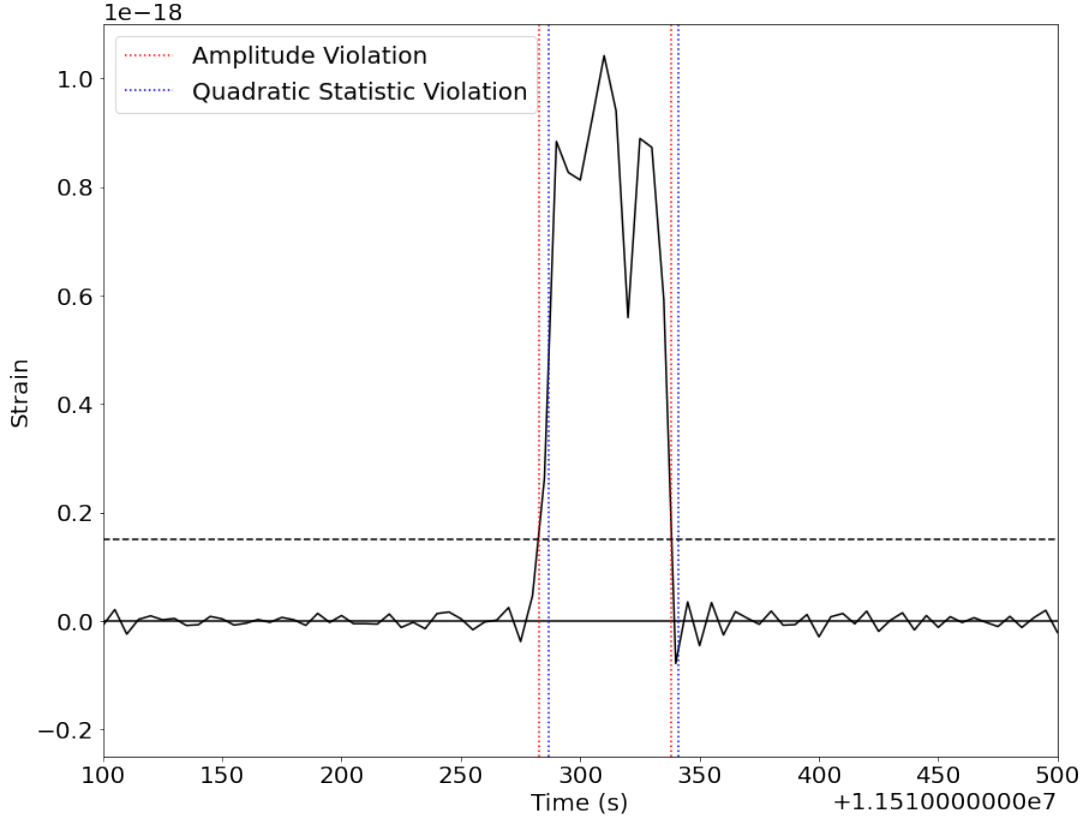


Figure 3.1: A sample glitch in the time series data. The black dashed line represents the threshold for amplitude statistic violation. The region in which the amplitude and quadratic statistic are violated are denoted by the red and blue dotted lines, respectively. Because the amplitude and quadratic statistic must be simultaneously violated for a valid glitch trigger event, the leftmost blue and rightmost red line define the glitch detection window. The data outside the glitch detection window consists of Gaussian noise and GW signals.

Now that the glitch has been identified, a padding parameter is and applied symmetrically as depicted in 3.2. This interval defines the region that will be affected by the Tukey window function.

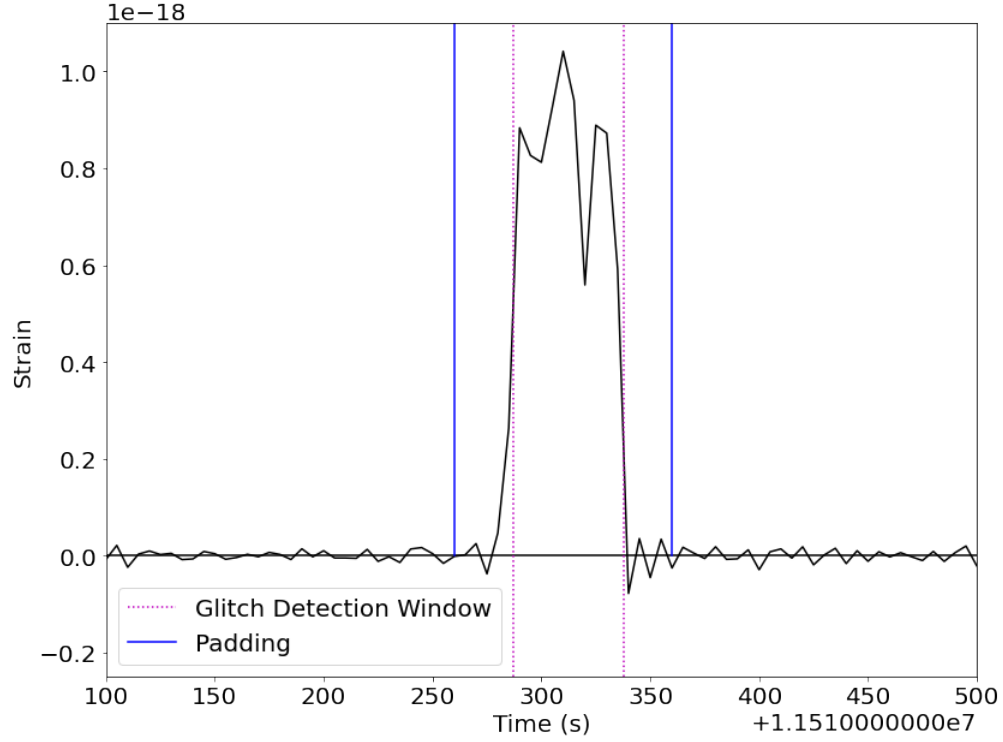


Figure 3.2: The time series of 3.1 after zeroing of the points within the glitch detection window. The purple dotted lines represent the intersection of the violation of the amplitude and quadratic statistics from 3.1. The solid blue line depicts one of the two hyper parameters in Tukey window function creation: padding. This is applied symmetrically around the glitch window. The data outside the glitch detection window consists of Gaussian noise and GW signals.

Now that the glitch detection window and padding have been defined, a framework of the Tukey window function can be created 3.3. However, a final parameter, denoted the window length, is required in order to define what proportion of the window is occupied the cosine. In the case of 3.3, half of the window length is the distance between the leftmost blue and purple lines.

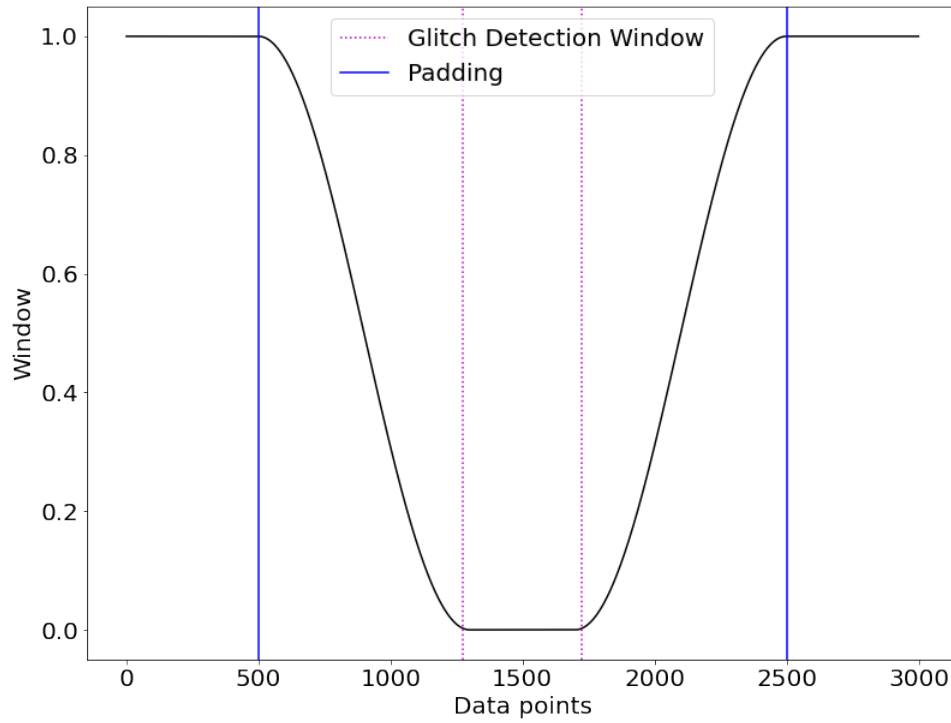


Figure 3.3: A sample window function created as a result of the process depicted in 3.1 and 3.2. The purple dotted lines represent the intersection of the violation of the amplitude and quadratic statistics from 3.1. The solid blue line depicts the padding parameter. Here, the distance between the leftmost blue and purple lines represents the half of the window length parameter. Note that the window length here is chosen to make the entire glitch detection region zeroed, but this is not generally true.

Now that the glitch has been zeroed and the Tukey window function has been created, this function can be multiplied to the time series in order to smooth out the effects of the otherwise rough cutting. Note that 3.3 only depicts a single window, and the entire time series will require a superposition of many window functions in order to remove all of the glitches. An example depicting a larger window function for a time series with two glitches is depicted in 3.4.

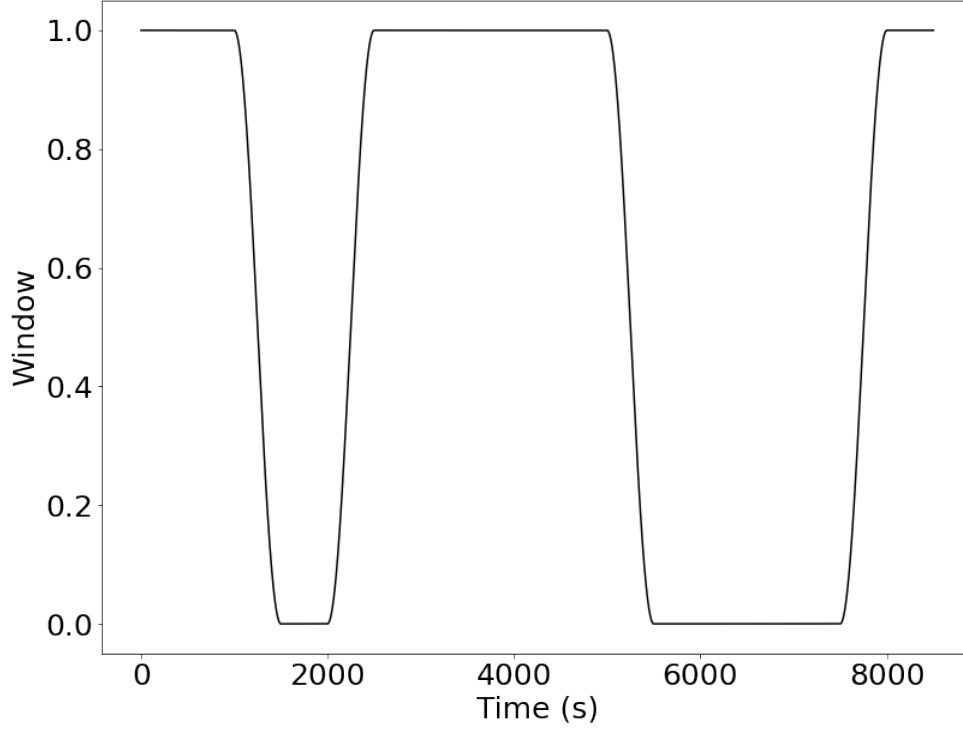


Figure 3.4: An example window function to be applied to artefact-ridden data. In this example, there would be two artefacts in the data localized at 1500 and 6500 seconds respectively. Note the tapered cosine around the zeroed regions, characteristic of the Tukey window.

3.2.2 Window Function Study

Previous work on Spritz by Kye Emond has shown interesting results. Utilizing the parameter estimation methods outlined in chapter 3.1 after removing glitches with Tukey window functions as per chapter 3.2.1, significant biases were seen between the calculated and true values. In particular, the most troublesome source parameter was amplitude. For all of the 36 GB sources in the data set, the true amplitude was consistently larger than the calculated value by over an order of magnitude.

The most likely explanation for this discrepancy is due to the excessive cutting of glitches. Note that, after zeroing out artefacts and applying a window function, the GW signal carried in the data is fundamentally changed. Because there is an underlying GW signal beneath the glitches,

applying the window function to the dataset also removes some of the GW power. The model used for parameter estimation assumes perfect sinusoidal signals, so the zeroed out regions of the time series could translate to the calculated best fit amplitude being lower than desired.

The natural solution to attempt to fix this discrepancy would be to incorporate the information about the window function into the model. However, this method has a crucial flaw. Each grid point in the F-statistic grid search necessitates a slightly different model function because of the varying frequency and sky position. While this task in of itself is computationally expensive, the effect is exacerbated by accounting for the window function in each permutation of the model. Because analysis is done in the frequency domain, and the window function must be multiplied to the model in the time domain, a convolution operation between the model and window function in frequency space must be performed. When repeated over several thousand grid points, this method becomes computationally prohibitive.

An alternate method would be to systematically adjust the window parameters. Because the original work used an arbitrary set of nominal window function parameters, it is possible that varying the window length and padding for the window functions could resolve the amplitude biases. In any case, it is physically interesting to determine the optimal set of window parameters needed to minimize the amount of GW signal cut, and maximize the amount of glitch power cut.

In order to determine these optimal parameters, a two-dimensional grid is initialized. Based on the nominal values of 50 and 3000 for padding and window length, the grid varies the padding from 0 to 100 data points in increments of 10, and the window length from 1000 to 3000 data points in increments of 100. In order to forgo an entire F-statistic gridsearch, the true values for λ , β , f , and \dot{f} for each source are utilized and the F-statistic is analytically maximized for each grid point. Thus, the landscape of the maximal F-statistic with respect to window length and padding is mapped for each source. The same process is also repeated utilizing instead a model that incorporates the respective window function. Coined the modified F-statistic, this treatment is used to explore the efficacy of the first proposed method for amplitude bias mitigation without extensive computational resources.

Chapter 4

Results

4.1 3D F-Statistic Landscapes

4.1.1 Unmodified F-Statistic

The landscape of the maximal F-statistic with respect to window length and padding was computed for all 36 sources. Taking the best-fit amplitude that corresponded to the maximal F-statistic for each point, a mapping of the difference between the true amplitude and calculated amplitude for each grid point was trivially derived. While each source had a unique distribution, the general trend was relatively consistent among all 36 sources. An example of the amplitude mapping for the lowest frequency source is shown in 4.1.

As shown in 4.1, the optimal window parameters when using unmodified GW signal templates simultaneously minimize window length and padding, although there is a steeper linear trend in the window length.

How this relates physically to the shape of the window function is interesting. Because the window length essentially parametrizes the roll-off of the window function, a smaller window length relates to a steep descent to the bottom of the well, and hence a small amount of the window function is occupied by a tapered cosine. Ultimately, a small window length results in a larger amount of the glitch detection window being zeroed. Because a steep descent suffers from a larger degree of spectral leakage than one that is more gradual, this result suggests that the degree in which spectral leakage affects parameter estimation is vastly outweighed by the impact of residual glitch power.

The fact that the maximal F-statistic also corresponds to a minimized padding is not an entirely

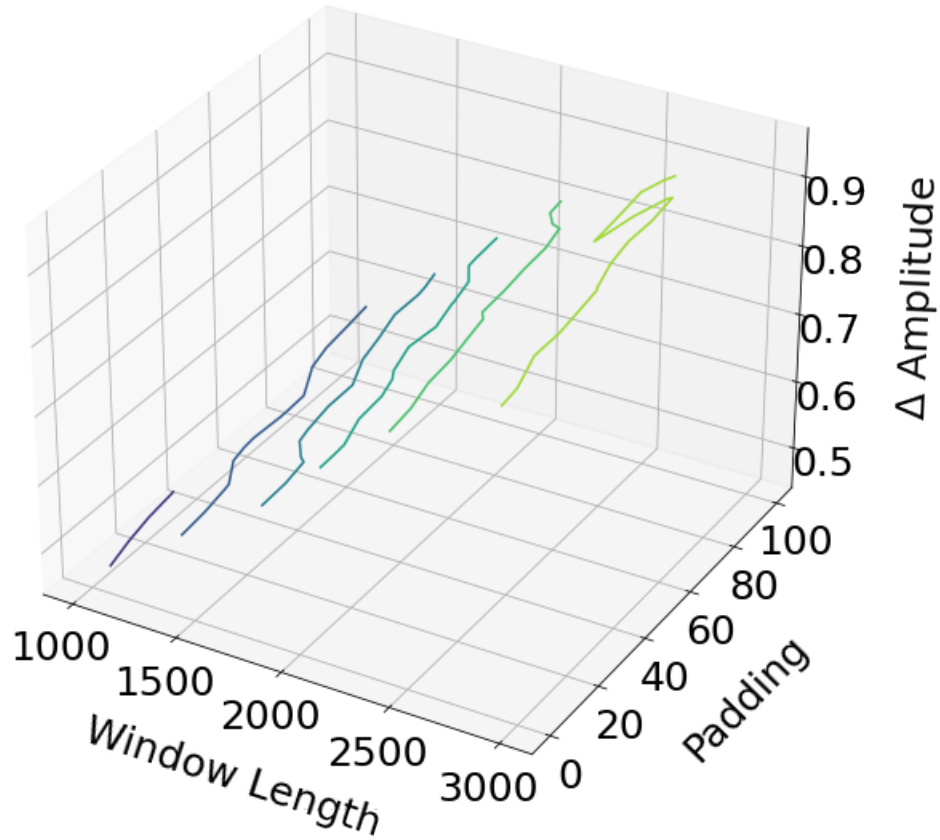


Figure 4.1: Δ Amplitude as a function of window length and padding, for the lowest frequency source (1.81 mHz). Here Δ Amplitude is defined as the difference between the $\log(\text{true amplitude})$ and $\log(\text{best-fit amplitude})$ for this source, as determined by the unmodified F-statistic. There is a linear trend with respect to increasing window length, with a smaller dependence on padding. Minimal Δ Amplitude is desirable, so the ideal window parameters are situated at the bottom of the linear trend.

surprising result. Because the length of the entire window is defined by twice the padding length plus the glitch detection window, a preference for smaller padding indicates that the glitch detection statistics already sufficiently capture most of the glitch. As a result, adding additional padding increases the degree in which actual signal is cut, without any benefit of reducing additional glitch power.

4.1.2 Modified F-Statistic

A modified F-statistic has been created that takes into account the windowing function that was applied to the data in order to remove glitches. Specifically, the F-statistic no longer naively uses an unmodified template for the GW signal, and alters it according to the characteristics of the windows throughout the data. Repeating the above process with this new F-statistic provides interesting results. The maximal F-statistic landscape was still relatively consistent among sources, but differed significantly from its unmodified counterparts. A sample distribution using the lowest frequency source is depicted in 4.2.

There are a few key differences between 4.1 and 4.2. First of all, as expected, the modified F-statistic has a lower $\Delta\text{Amplitude}$ value across the entire grid because the modified F-statistic signal templates more closely resemble the data. The more peculiar discrepancy comes from the dependence on the window length. If the previous amplitude bias was truly a result of the data not matching the model, we would expect the distribution in 4.2 to be relatively flat, centered around a $\Delta\text{Amplitude}$ of 0. However, it is clear that this change, while it did have a positive impact in mitigating amplitude bias, was insufficient to remove the bias entirely. Moreover, the linear trend for the modified F-statistic with respect to window length is in the opposite direction as its unmodified counterpart. In contrast to the unmodified analysis, a large amount of the window length should be occupied by a tapered cosine in this domain. This observation suggests that, when the model accounts for the degree of glitch cutting, the effects of spectral leakage now dominate in hindering parameter estimation. Of course, the fact that minimal padding is preferred aligns with the previous analysis in the unmodified case.

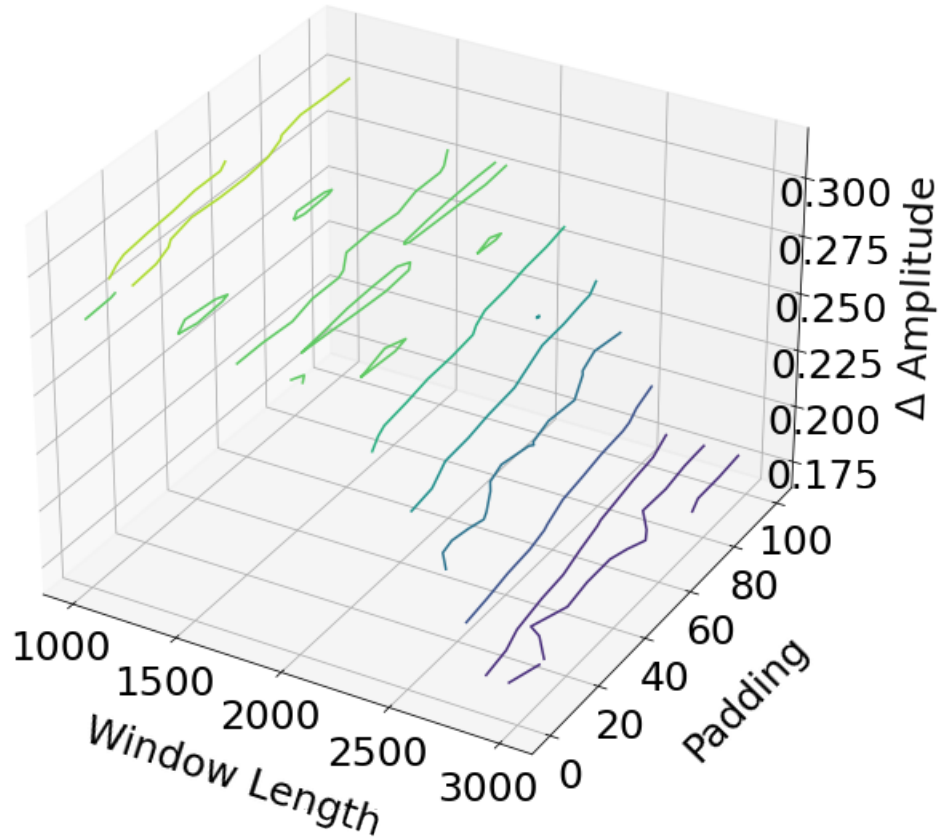


Figure 4.2: Δ Amplitude as a function of window length and padding, for the lowest frequency source (1.81 mHz). Here Δ Amplitude is defined as the difference between the $\log(\text{true amplitude})$ and $\log(\text{best-fit amplitude})$ for this source, as determined by the modified F-statistic. There is a linear trend with respect to increasing window length, with a smaller and opposite dependence on padding. Minimal Δ Amplitude is desirable, so the ideal window parameters are situated at the bottom of the linear trend.

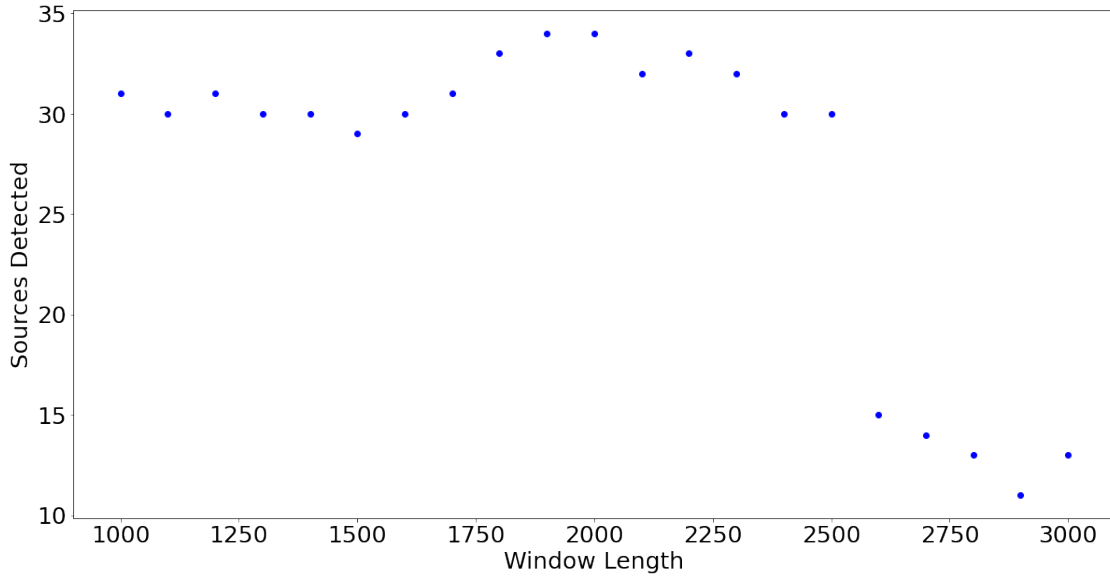


Figure 4.3: The number of sources detected as a function of window length for the unmodified F-statistic. Padding is held constant at 0. The number of detected sources is relatively constant in the mid 30s for small window lengths. There is a significant drop in sources detected beyond a window length of 2500.

4.2 Source Detection Criteria

While the landscape of the F-statistic with respect to the window parameters is a useful diagnostic tool, it alone cannot fully describe the optimal source parameters. This is due to the fact that this plot is completely oblivious to the number of sources detected. Utilizing the source detection criterion as outlined in chapter 2.1, the number of sources detected as a function of window length can be produced. For this analysis, the padding is minimized.

Fortunately, as shown in 4.3, the number of sources detected is relatively constant from a window length between 1000 and 2500 for the the unmodified case. Beyond a window length of 2500, there is a large drop in the number of sources detected. This is a similar to the trend to 4.1, as a large window length is undesirable in both plots. As such, for the unmodified case it appears that simultaneously minimizing both window length and padding is ideal.

The same analysis can be applied to the modified case. This is depicted in 4.4. Here the number

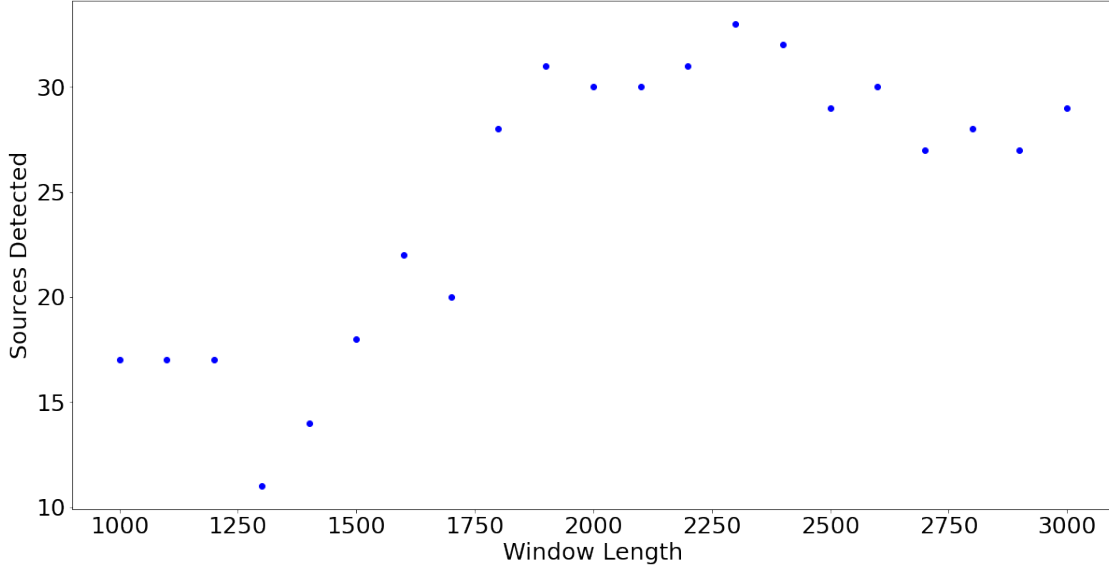


Figure 4.4: The number of sources detected as a function of window length for the modified F-statistic. Padding is held constant at 0. The number of detected sources is relatively constant in the mid 30s for large window lengths. There is a significant drop in sources detected before a window length of around 1800.

of sources detected is maximized and relatively constant beyond a window length of 1800. However, below a window length of 1800, comparatively few sources are detected. This is similar to the trend in 4.2, as here a small window length is undesirable in both plots. Here the discrepancy between the unmodified and modified case persists as the modified case prefers to instead maximize window length for the ideal window parameters.

4.3 Modified F-Statistic Efficacy

Having determined the optimal window parameters for both the unmodified and modified F-statistic cases, it is important to determine the degree of improvement in amplitude bias mitigation the modified F-statistic has over the unmodified F-statistic. In order for a fair comparison, the same set of window parameters needs to be chosen for both cases. Conditioning on number of sources detected, there is an overlap in the maxima in the mid-range of the window lengths. Moreover, the

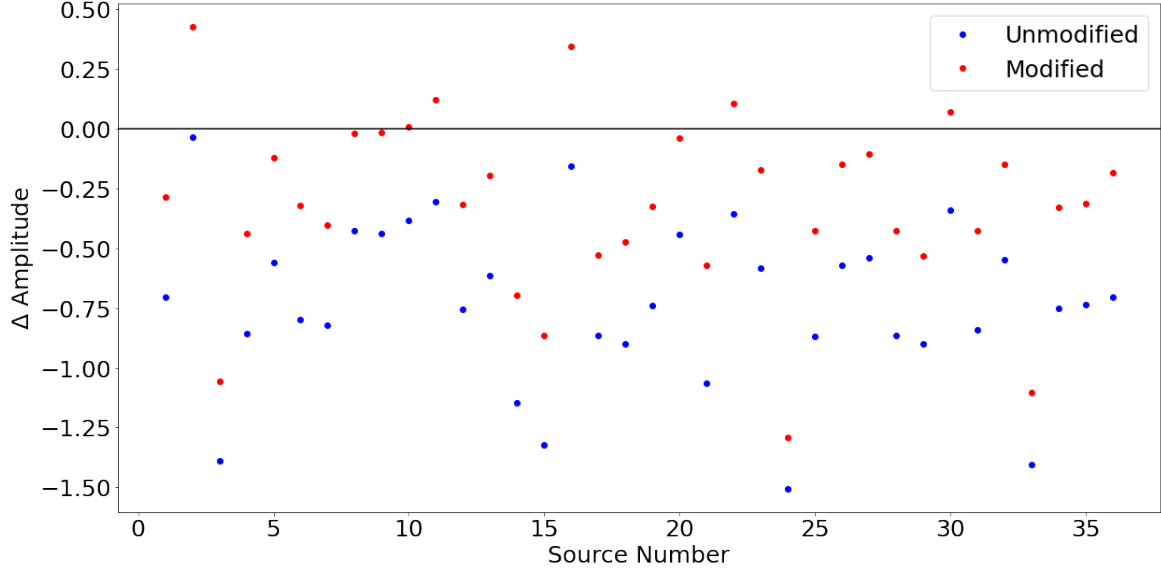


Figure 4.5: $\Delta\text{Amplitude}$ for the 36 sources in the data set. Here $\Delta\text{Amplitude}$ is defined as the difference between the $\log(\text{true amplitude})$ and $\log(\text{best-fit amplitude})$. The sources are ordered in terms of increasing frequency. The best-fit amplitude was determined by maximizing the F-statistic when using a padding of 0 and window length of 2000. This was done for both the unmodified (blue dots) and modified (red dots) cases. Points closer to the horizontal 0 line indicate a smaller amplitude bias. Values are tabulated in the appendix A.

middle window length of 2000 is not preferential to either case's F-statistic landscapes. As such, a padding length of 0 and a window length of 2000 was chosen to undertake this analysis.

Using these window parameters, the best-fit amplitude value was extracted for each of the 36 sources, for both cases. With this, the deviation from the true amplitude for each source can be compared. This is depicted exactly in 4.5.

As previously hypothesized, inclusion of the window function in the model increases the value of the calculated best-fit amplitude. This is evident in the fact that the modified (red) dots are consistently above the unmodified (blue) dots in 4.5. Because the absolute distance of these points to the horizontal 0 line is the physically interesting quantity, the information in 4.5 can be revisualized. Taking the difference of the absolute value of $\Delta\text{Amplitude}$ between the unmodified and modified case, the degree of amplitude mitigation improvement is shown in 4.6.

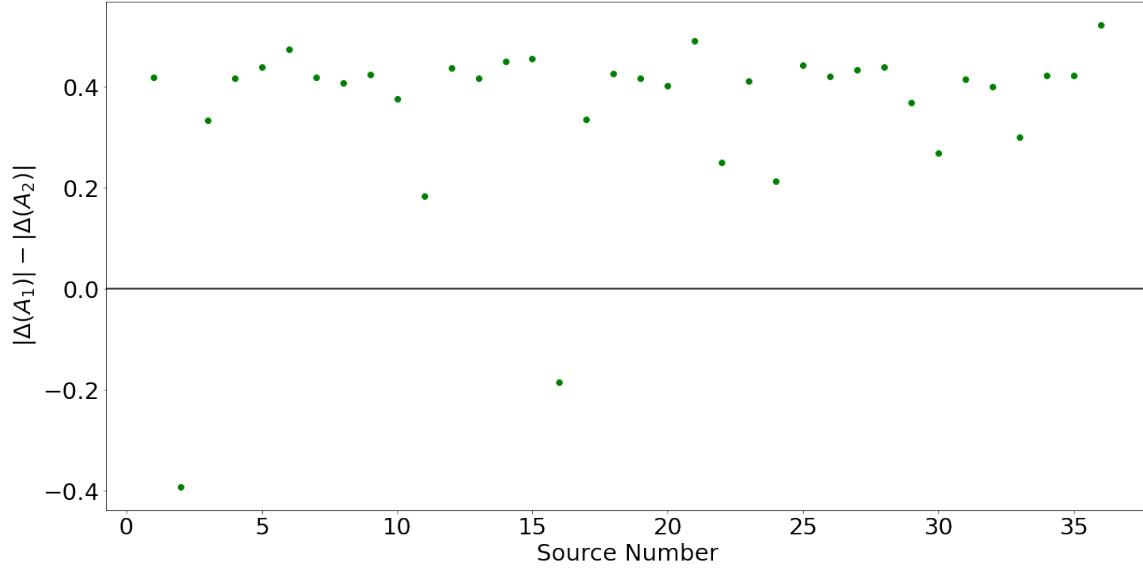


Figure 4.6: The difference in $|\Delta\text{Amplitude}|$ between the unmodified and modified F-statistic cases for the 36 sources in the data set. Here $\Delta\text{Amplitude}$ is defined as the difference between the $\log(\text{true amplitude})$ and $\log(\text{best-fit amplitude})$. The best-fit amplitude was determined by maximizing the F-statistic when using a padding of 0 and window length of 2000. The sources are ordered in terms of increasing frequency. Points above the horizontal 0 line indicate where the modified case mitigates amplitude bias better than the unmodified case, whereas points below this axis indicate the opposite.

As expected, 4.6 depicts the modified F-statistic providing an improvement in amplitude bias over the unmodified case. For most sources, this improvement is relatively constant with an improvement of $0.4 \cdot 10$. However, the second and sixteenth source deviated from this trend as the modified case actually did worse at mitigating the amplitude than the unmodified case. Referring back to 4.5, source 2 and 16 already had a relatively small amount of amplitude bias in the unmodified case. Because all best-fit amplitudes increased when accounting for the window function, the modified case ended up significantly overshooting the horizontal 0 line for these two sources. As to what causes this discrepancy is unclear. Because the sources are ordered in terms of increasing frequency, and these two sources are at different points of the frequency range, it is unlikely that the intrinsic frequency of these sources is causing this effect.

Chapter 5

Discussion and Analysis

Altogether, this project has two main findings. First of all, the optimal parameters for the window length and padding of the Tukey window function have been determined. This has been done when utilizing both the unmodified and modified F-statistic. In the unmodified case, it was found that the amplitude bias mitigation followed a linear trend with respect to window length. Here the optimal window length and padding was found to be their respective minimal values. Fortunately, the values that minimized amplitude bias simultaneously maximized the number of sources detected. As such, the optimal padding and window length for the unmodified case were found to be 0 and 1000, respectively. On the other hand, in the modified case, the opposite was found. The amplitude bias mitigation followed a linear trend with respect to window length, but in the opposite direction. Here the optimal window length was found to be its maximal value, and the padding was again found to be its minimal value. Like the unmodified case, these found parameters also maximized the number of sources detected. As such, the optimal padding and window length for the unmodified case were found to be 0 and 3000, respectively.

The second main finding of this project is the effect of modifying the GW signal templates on amplitude bias mitigation. Because the window function removes power from the underlying GW signal, calculated best-fit amplitudes are consistently underestimated when the model function is purely sinusoidal. As expected, all best-fit amplitudes increased when utilizing the modified signal templates. In most cases, this provided a relatively constant degree of improvement over the unmodified case of $0.4 \cdot 10$. However, for two outlier sources the increase in amplitude that the modified templates provided caused amplitude bias mitigation to decrease. Nonetheless, even after applying the modified window function to the data, very few source's calculated best-fit amplitude

converged to a significant degree to their true values.

The reason that modifying the signal templates to account for the window function does not completely mitigate amplitude biases is unclear. However, given the fact that the two outlier sources exist in different regions of the frequency spectra, it is unlikely that frequency of a source has a significant impact. In any case, this study has revealed that, while in most cases it does provide an improvement, the modified F-statistic is insufficient in resolving the amplitude biases in parameter estimation.

There are several suggestions for future research stemming from this project. First of all, it would be interesting to investigate if any of the other intrinsic source parameters, other than frequency, have a significant impact on amplitude bias mitigation. Further, using a finer grid to vary window length and padding might provide even further optimization of the window parameters. Moreover, given the fact that the effects of the modified F-statistic are generally positive, it would be beneficial to explore methods that permit a full F-statistic gridsearch and MCMC using this model. This might be able to be achieved via code optimization and possible minimization of necessary convolutions and Fourier transformations.

Bibliography

- [1] B Abbott, F. Abbott, Ranav Adhikari, A. Ageev, Bruce Allen, Ria Amin, S. Anderson, Warren Anderson, Melody Araya, H. Armandula, Fedah Asiri, Peter Aufmuth, Carsten Aulbert, Stanislav Babak, R. Balasubramanian, Stefan Ballmer, B. Barish, D. Barker, C. Barker-Patton, and M. Fyffe. Detector description and performance for the first coincidence observations between ligo and geo. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v.517, 154-179 (2004), 05 2016.

- [2] Pau Amaro-Seoane, Heather Audley, Stanislav Babak, John Baker, Enrico Barausse, Peter Bender, Emanuele Berti, Pierre Binetruy, Michael Born, Daniele Bortoluzzi, Jordan Camp, Chiara Caprini, Vitor Cardoso, Monica Colpi, John Conklin, Neil Cornish, Curt Cutler, Karsten Danzmann, Rita Dolesi, Luigi Ferraioli, Valerio Ferroni, Ewan Fitzsimons, Jonathan Gair, Lluís Gesa Bote, Domenico Giardini, Ferran Gibert, Catia Grigani, Hubert Halloin, Gerhard Heinzl, Thomas Hertog, Martin Hewitson, Kelly Holley-Bockelmann, Daniel Hollington, Mauro Hueller, Henri Inchauspe, Philippe Jetzer, Nikos Karnesis, Christian Killow, Antoine Klein, Bill Klipstein, Natalia Korsakova, Shane L Larson, Jeffrey Livas, Ivan Lloro, Nary Man, Davor Mance, Joseph Martino, Ignacio Mateos, Kirk McKenzie, Sean T McWilliams, Cole Miller, Guido Mueller, Germano Nardini, Gijs Nelemans, Miquel Nofrarias, Antoine Petiteau, Paolo Pivato, Eric Plagnol, Ed Porter, Jens Reiche, David Robertson, Norna Robertson, Elena Rossi, Giuliana Russano, Bernard Schutz, Alberto Sesana, David Shoemaker, Jacob Slutsky, Carlos F. Sopuerta, Tim Sumner, Nicola Tamanini, Ira Thorpe, Michael Troebs, Michele Vallisneri, Alberto Vecchio, Daniele Vetrugno, Stefano Vitale, Marta Volonteri, Gudrun Wanner,

- Harry Ward, Peter Wass, William Weber, John Ziemer, and Peter Zweifel. Laser interferometer space antenna, 2017.
- [3] Joseph Amato. Flying in formation: The orbital dynamics of lisa’s three spacecraft. *American Journal of Physics*, 87:18–23, 01 2019.
- [4] and T. Akutsu, M. Ando, K. Arai, Y. Arai, S. Araki, A. Araya, N. Aritomi, H. Asada, Y. Aso, S. Atsuta, K. Awai, S. Bae, L. Baiotti, M. A. Barton, K. Cannon, E. Capocasa, C-S. Chen, T-W. Chiu, K. Cho, Y-K. Chu, K. Craig, W. Creus, K. Doi, K. Eda, Y. Enomoto, R. Flaminio, Y. Fujii, M.-K. Fujimoto, M. Fukunaga, M. Fukushima, T. Furuhata, S. Haino, K. Hasegawa, K. Hashino, K. Hayama, S. Hirobayashi, E. Hirose, B. H. Hsieh, C-Z. Huang, B. Ikenoue, Y. Inoue, K. Ioka, Y. Itoh, K. Izumi, T. Kaji, T. Kajita, M. Kakizaki, M. Kamiizumi, S. Kanbara, N. Kanda, S. Kanemura, M. Kaneyama, G. Kang, J. Kasuya, Y. Kataoka, N. Kawai, S. Kawamura, T. Kawasaki, C. Kim, J. Kim, J. C. Kim, W. S. Kim, Y.-M. Kim, N. Kimura, T. Kinugawa, S. Kirii, Y. Kitaoka, H. Kitazawa, Y. Kojima, K. Kokeyama, K. Komori, A. K. H. Kong, K. Kotake, R. Kozu, R. Kumar, H-S. Kuo, S. Kuroyanagi, H. K. Lee, H. M. Lee, H. W. Lee, M. Leonardi, C-Y. Lin, F-L. Lin, G. C. Liu, Y. Liu, E. Majorana, S. Mano, M. Marchio, T. Matsui, F. Matsushima, Y. Michimura, N. Mio, O. Miyakawa, A. Miyamoto, T. Miyamoto, K. Miyo, S. Miyoki, W. Morii, S. Morisaki, Y. Moriwaki, T. Morozumi, M. Musha, K. Nagano, S. Nagano, K. Nakamura, T. Nakamura, H. Nakano, M. Nakano, K. Nakao, T. Narikawa, L. Naticchioni, L. Nguyen Quynh, W.-T. Ni, A. Nishizawa, Y. Obuchi, T. Ochi, J. J. Oh, S. H. Oh, M. Ohashi, N. Ohishi, M. Ohkawa, K. Okutomi, K. Ono, K. Oohara, C. P. Ooi, S-S. Pan, J. Park, F. E. Peña Arellano, I. Pinto, N. Sago, M. Saijo, S. Saitou, Y. Saito, K. Sakai, Y. Sakai, Y. Sakai, M. Sasai, M. Sasaki, Y. Sasaki, S. Sato, N. Sato, T. Sato, Y. Sekiguchi, N. Seto, M. Shibata, T. Shimoda, H. Shinkai, T. Shishido, A. Shoda, K. Somiya, E. J. Son, A. Sue-masa, T. Suzuki, T. Suzuki, H. Tagoshi, H. Tahara, H. Takahashi, R. Takahashi, A. Takamori, H. Takeda, H. Tanaka, K. Tanaka, T. Tanaka, S. Tanioka, E. N. Tapia San Martin, D. Tatsumi, T. Tomaru, T. Tomura, F. Travasso, K. Tsubono, S. Tsuchida, N. Uchikata, T. Uchiyama, T. Uehara, S. Ueki, K. Ueno, F. Uraguchi, T. Ushiba, M. H. P. M. van Putten, H. Vocca,

- S. Wada, T. Wakamatsu, Y. Watanabe, W-R. Xu, T. Yamada, A. Yamamoto, K. Yamamoto, K. Yamamoto, S. Yamamoto, T. Yamamoto, K. Yokogawa, J. Yokoyama, T. Yokozaawa, T. H. Yoon, T. Yoshioka, H. Yuzurihara, S. Zeidler, and Z.-H. Zhu. KAGRA: 2.5 generation interferometric gravitational wave detector. *Nature Astronomy*, 3(1):35–40, jan 2019.
- [5] Michele Vallisneri Andrzej Krolak, Massimo Tinto. Optimal filtering of the lisa data.
- [6] Simon Barke. *Inter-spacecraft frequency distribution for future gravitational wave observatories*. PhD thesis, Leibniz U., Hannover, 2015.
- [7] B. Caron, A. Dominjon, C. Drezen, R. Flaminio, X. Grave, F. Marion, L. Massonnet, C. Mehmél, R. Morand, B. Mours, V. Sannibale, M. Yvert, D. Babusci, S. Bellucci, S. Candusso, G. Giordano, G. Matone, J.-M. Mackowski, L. Pinard, F. Barone, E. Calloni, L. Di Fiore, M. Flagiello, F. Garuti, A. Grado, M. Longo, M. Lops, S. Marano, L. Milano, S. Solimeno, V. Brisson, F. Cavalier, M. Davier, P. Hello, P. Heusse, P. Mann, Y. Acker, M. Barsuglia, B. Bhawal, F. Bondu, A. Brillet, H. Heitmann, J.-M. Innocent, L. Latrach, C.N. Man, M. Pham-Tu, E. Tournier, M. Taubmann, J.-Y. Vinet, C. Boccara, Ph. Gleyzes, V. Loriette, J.-P. Roger, G. Cagnoli, L. Gammaitoni, J. Kovalik, F. Marchesoni, M. Punturo, M. Beccaria, M. Bernardini, E. Bougleux, S. Braccini, C. Bradaschia, G. Cella, A. Ciampa, E. Cuoco, G. Curci, R. Del Fabbro, R. De Salvo, A. Di Virgilio, D. Enard, I. Ferrante, F. Fidecaro, A. Giamassi, A. Giazotto, L. Holloway, P. La Penna, G. Losurdo, S. Mancini, M. Mazzoni, F. Palla, H.-B. Pan, D. Passuello, P. Pelfer, R. Poggiani, R. Stanga, A. Vicere’, Z. Zhang, V. Ferrari, E. Majorana, P. Puppo, P. Rapagnani, and F. Ricci. The virgo interferometer for gravitational wave detection. *Nuclear Physics B - Proceedings Supplements*, 54(3):167–175, 1997.
- [8] Neil J. Cornish and Tyson B. Littenberg. Tests of bayesian model selection techniques for gravitational wave astronomy. *Physical Review D*, 76(8), oct 2007.
- [9] Lee S. Finn. Detection, measurement, and gravitational radiation. *Physical Review D*, 46(12):5236–5249, dec 1992.

- [10] Andrzej Królak and Mandar Patil. The first detection of gravitational waves. *Universe*, 3(3), 2017.
- [11] Jess McIver and D. H. Shoemaker. Discovering gravitational waves with advanced ligo. *Contemporary Physics*, 61(4):229–255, 2020.
- [12] Markus Otto. Time-delay interferometry simulations for the laser interferometer space antenna., 2015.
- [13] Reinhard Prix and John T Whelan. \mathcal{F} -statistic search for white-dwarf binaries in the first mock LISA data challenge. *Classical and Quantum Gravity*, 24(19):S565–S574, sep 2007.
- [14] Eric Thrane and Colm Talbot. An introduction to bayesian inference in gravitational-wave astronomy: Parameter estimation, model selection, and hierarchical models. *Publications of the Astronomical Society of Australia*, 36, 2019.
- [15] Massimo Tinto and Sanjeev V. Dhurandhar. Time-delay interferometry. *Living Reviews in Relativity*, 8(1), jul 2005.

Appendix A

First Appendix

The following table A.1 summarizes the best-fit amplitude values as depicted in 4.5.

Source Number	Unmodified log(Amplitude)	Modified log(Amplitude)	True log(Amplitude)
1	-22.5078	-22.0882	-21.8035
2	-22.6497	-22.1884	-22.6152
3	-22.9692	-22.6358	-21.5786
4	-22.4047	-21.9878	-21.5483
5	-22.9354	-22.4963	-22.3743
6	-23.1411	-22.6660	-22.3443
7	-22.3456	-21.9268	-21.5242
8	-22.4307	-22.0234	-22.0035
9	-22.4891	-22.0652	-22.0502
10	-23.2074	-22.8190	-22.8253
11	-22.4991	-22.0743	-22.1953
12	-22.6446	-22.2077	-21.8910
13	-22.8734	-22.4560	-22.2608
14	-22.6117	-22.1606	-21.4636
15	-22.7036	-22.2477	-21.3819
16	-23.0355	-22.5341	-22.8776
17	-23.1094	-22.7740	-22.2441
18	-22.8148	-22.3895	-21.9149
19	-22.6793	-22.2630	-21.9388
20	-22.7969	-22.3951	-22.3551
21	-23.1053	-22.6149	-22.0425
22	-22.9607	-22.4971	-22.6037
23	-22.5558	-22.1442	-21.9711
24	-22.8625	-22.6484	-21.3558
25	-22.9427	-22.4993	-22.0742
26	-22.5299	-22.1095	-21.9597
27	-22.5327	-22.0992	-21.9946
28	-22.6283	-22.1894	-21.7631
29	-22.8467	-22.4781	-21.9470
30	-22.8359	-22.4227	-22.4946
31	-22.6383	-22.2232	-21.7961
32	-22.7578	-22.3579	-22.2098
33	-23.3676	-23.0671	-21.9643
34	-22.3728	-21.9497	-21.6227
35	-22.5257	-22.1039	-21.7907
36	-22.9225	-22.3999	-22.2176

Table A.1: Summary of the best-fit amplitude values as depicted in 4.5. The window parameters used for this calculation is 2000 and 0 for window length and padding, respectively.