GLOBAL ANALYSIS FOR

SPACE-BASED GRAVITATIONAL WAVE OBSERVATORIES

by

Travis James Robson

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Physics

MONTANA STATE UNIVERSITY
Bozeman, Montana

November, 2018

DEDICATION

For my parents Bill and Karen, and my brother Billy. Without you I would never have made it this far.

ACKNOWLEDGEMENTS

I would like to thank my adviser Neil Cornish. I have learned an incredible amount working with Neil, and he has provided me with a great opportunity. Thanks also goes to Nico Yunes for keeping me on my toes. I also want to thank David Anderson and Blake Moore for their useful conversation and great company.

TABLE OF CONTENTS

TABLE OF CONTENTS – CONTINUED

TABLE OF CONTENTS – CONTINUED

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES – CONTINUED

Figure Page

LIST OF FIGURES – CONTINUED

Figure Page

ABSTRACT

The Laser Interferometer Space Antenna (LISA) is a space-based gravitational wave detector in development under a joint venture between ESA and NASA. LISA will be sensitive to a wealth of signals from a variety of sources—both astrophysical and instrumental. Since many of these signals will be overlapping we must carry out a global analysis where we model everything believed to be present in the data *simultaneously*. To analyze the data this way we must understand what types of signals we expect, develop fast signal generators, and develop data analysis algorithms to handle this problem. We must also be flexible to characterize signals that we do not expect such as instrumental glitches of unknown morphology, or exotic astrophysical sources. We employ the Markov Chain Monte Carlo algorithm to address these multiple facets of the global analysis problem through a Bayesian approach. We have developed fast models for a variety of sources, characterized what we can learn about the sources, and assessed the nature of LISA's global analysis problem.

1

INTRODUCTION

The scientific community has entered the Gravitational Wave Astronomy age with the detections of gravitational waves made by the Laser Interferometer Gravitational wave Observatory (LIGO) [1–6]. Even LIGO's modest number of detections have led to significant contributions to astrophysics [7–13], but LIGO's narrow window of sensitivity leaves us blind to an enormous amount of gravitational wave activity in our universe. Theoretical astrophysics studies hypothesize [14–17], and some observations [18] guarantee, that there are a wealth of gravitational wave sources in the millihertz frequency regime, a regime to which LIGO is insensitive. To detect and study these sources the Laser Interferometer Space Antenna (LISA) has been proposed [19]. The LISA mission has been revitalized as a joint ESA-NASA mission within the past 2 years. The revival of the LISA mission is owed to the discovery of gravitational waves by LIGO and the tremendously successfully technology demonstration: LISA Pathfinder [20,21]. LISA Pathfinder, carried out by the ESA, consisted of a single spacecraft and successfully demonstrated the feasibility of many key technologies that LISA will use. When active, LISA will view the entire cosmic sky in a wide frequency window ranging from 10^{-4} to 10^{-1} Hz. This frequency window is situated above the nanohertz regime to which pulsar timing arrays are sensitive. More importantly, LISA's sensitivity window is situated directly below LIGO's kilohertz sensitivity, allowing for powerful joint detections. LISA will be composed of 3 spacecraft, separated by 2.5 million km, arranged as a quasi-equilateral triangle placed in an Earth-trailing orbit [19]. These spacecraft will house inertial masses and optical benches supporting 6 laser links to measure the masses'

relative separations. Phasemeters are used to measure the phase of the laser links. Gravitational waves perturb this phase allowing for the detection and characterization of a gravitational wave source.

The gravitational waves from the sources in this millihertz frequency regime will generate interesting new science. Most of the gravitational waves to bathe LISA are generated within our own galaxy, produced by tens of millions of compact galactic binaries (GBs) [14,22,23]. These galactic binaries will typically consist of white dwarfs with the occasional white dwarf-neutron star, binary neutron star, and neutron star-black hole pair. It is expected that tens of thousands of these GBs will be resolved by LISA over the course of its nominal 4 year mission lifetime [24]. There will be information extracted through the estimation of the gravitational wave power emitted by the unresolved GBs. Galactic Binaries will provide a wealth of science, giving the best measurements of the Milky Way's shape [25], informing our understanding of binary [26] and triple [27] evolution. The main source of gravitational waves are binary systems. There are many kinds of binary systems other than galactic binaries. Binaries of different masses, mass ratios, separations, and materials will result in different orbital dynamics because different physics will be relevant depending on these parameters. This means that binaries offer a rich variety of physics to study. We will learn more about galactic evolution through the study of binary black holes (BBHs) [16] detected by LISA. LIGO's detections have forced us to pose a new question: what is the origin of stellar mass binary black holes [15]? The LIGO black holes were lower in mass than expected and LISA will greatly aid the effort of explaining that result. Another variety of binary is the extreme mass ratio inspiral (EMRI) which consist of a low mass body zooming and whirling around a massive black hole for many orbits. LISA will test General Relativity with unprecedented precision through the detection of EMRIs [17].

The enormous number of sources detectable by LISA is of great benefit to science but is a challenge to the data analyst. For example, the unresolved GBs will constitute a non-stationary noise source [28, 29] louder than the instrument noise for some range of frequencies. When confronted with the problem of detection and signal characterization a standard technique used is matched filtering [30, 31]. Matched filtering identifies the best-fit parameters for a signal in the data by generating a guess of the signal for various parameters and identifying the guess which best matched the data. The resulting best-fit signal is subtracted from the data and the next signal is hunted for, i.e. a source by source approach. The space of parameters for LISA sources is enormous [32] rendering matched filtering useless. The parameter space is so voluminous because there are so many sources which we expect to be able to resolve and because some of the sources' waveforms are extremely sensitive to the parameters thus demanding a more dense population of parameter space in the matched filtering process. Worse yet, due to the vast number of signals, there will exist many small overlaps which severely bias a source by source approach. For example, small overlaps occur when a BBH signal overlaps thousands of GB signals, leading to a strong bias in the recovered parameters of the BBH. Our trouble does not stop there: instrumental glitches of unknown shape will pollute the data, and in the spirit of science, we must be prepared to characterize the unknown, identifying astrophysical signals for which we have no models. These complications demand a global fit to the data, i.e. simultaneously fitting for all astrophysical, instrumental signals, and incorporating the ability to handle unexpected signals.

The Mock LISA Data Challenges [33–35] gave the LISA data analyst community a chance to develop algorithms to solve the global analysis problem on a blind data set. The full data set was never solved in its entirety, due to the cancellation of the old LISA mission, though much progress was made. In these challenges many

source by source algorithms were created which as discussed before is a standard approach to the detection/characterization but unfortunately falls short for LISA. One such algorithm developed by Cornish and Larson [36] was a modified version of the electromagnetic astronomy CLEAN algorithm [37]. Their code identifies the current brightest source, and then partially subtracts it. The next brightest source (potentially the same source) is identified and then again is partially subtracted. This repeats until a stopping condition is met. An alternative attempt was made by Rubbo et al [38] who implemented a slice and dice method where the current bright signal is *fully* subtracted. However, an improvement upon the explicit source-by-source subtraction is made: at each iteration a simultaneous least squares fit is found to the current set of bright sources' parameters (a stopping condition was also used). A more successful attempt at the global analysis problem was conducted by Crowder & Cornish [39]. They analyzed the data in frequency chunks and utilized the F statistic [40] which maximizes the likelihood over the extrinsic signal parameters. This was combined with a Markov Chain Monte Carlo (MCMC) to calculate the posterior for the remaining parameters by using the maximized likelihood. To determine the number of sources Crowder & Cornish used the Laplace approximation to the Bayesian evidence. Littenberg [41] also utilized the F statistic maximization and broke the data up into frequency chunks, but used a Reversible Jump MCMC to find the appropriate number of sources to explain the data. With the rebirth of LISA we are sure to see a new slew of algorithms designed to address LISA's data analysis needs. In fact, a new mock data challenge [42] is already in motion.

To address the aforementioned complications, and build off past experience gained by the community, we wish to solve LISA's global analysis problem through a fully Bayesian approach. Bayesian probability theory has been utilized many times in gravitational wave astronomy problems [43–46]. Bayesian inference provides the

analyst with the tools to incorporate data into an updated belief about a hypothesis or to calculate a relative belief between competing hypotheses. These two exercises are referred to as parameter estimation and model selection. Parameter estimation provides a best-fit set of parameters—given both the data and a prior belief about the parameters—and describes their associated error characterized by a posterior probability distribution. Model selection provides the probability of competing models to explain data, as measured by the evidence. A classic model selection problem is the standard detection problem: is there a signal in the data or not.

The powerful Markov Chain Monte Carlo (MCMC) [47] algorithm is frequently used to solve these problems of Bayesian inference. The MCMC is initialized at set of parameters which specify the model. A proposal probability distribution is sampled, generating a guess at a new set of parameters which may explain the data. The probability that these parameters are chosen to be the new set of parameters is given by the Metropolis-Hastings ratio. The sequence of parameter sets is referred to as a Markov chain. The original MCMC prescription, while guaranteed to converge given enough time, can converge incredibly slow when the posterior distribution is multi-modal. What makes an MCMC incredibly powerful is that it is a quasi-hill climber in that the chain tends to climb likelihood surfaces to reach the peak. However, to characterize the error, the MCMC can explore downwards, but the probability that the MCMC will traverse down a mode and find itself in another can be amazingly small. Therefore, characterizing more than one mode can take a long time computationally. The MCMC can be augmented by multiple MCMCs running in parallel at various temperatures, i.e. using artificially flattened likelihoods, with occasional communication. This algorithm, called parallel tempering MCMC (PTMCMC) [48], converges quicker than the standard MCMC algorithm in the presence multi-modal posterior distributions. The standard MCMC can also be

generalized to handle a trans-dimensional parameter space. Reversible Jump MCMC (RJMCMC) [49] is a method which tackles the model selection problem of figuring out how many parameters are appropriate to describe the data.

The development of a grand MCMC which can solve the global analysis problem for LISA will be a challenge. Many of the complications that will be apart of the real LISA data set have been simplified away by the Mock LISA data challenges. Difficulty arises when you are unsure of how many sources truly explain the data. While in principle all of the information needed to answer this question will lie within the Markov chains generated by the MCMC, digesting this information into a usable form poses a challenge. We will need an evolving source catalog containing the sources we have detected and stating the degree to which we are confident of that detection. We must also state the best-fit parameters of these systems. Additional complications arise due to the fact that we are unsure as to the precise nature of the physics at play in a given source beforehand. We must be flexible enough to allow for a variety of physics to be used when trying to characterize the LISA data. To date, in these mock data sets, the galactic binaries were assumed to be isolated, and to have circular orbits with mildly chirping frequency evolution [50, 51], due to either mass transfer or gravitational wave emission. There will be a variety of environments in which binaries live [52–54] invalidating these niceties thus complicating their gravitational wave signals. Most stars are not found in binary systems [55, 56]. Most stars reside in triples, quadruples, or more! Companion stars to the gravitational wave emitting binary will severely modulate the inner binary's signal, rendering the isolated binary model useless or biased [27]. Unlike LIGO, LISA will observe many binaries in eccentric orbits created by environmental effects [57, 58]. Eccentricity adds a richness and complexity to the signal. When signals become complicated, such as in the eccentric case, generating their waveform typically becomes computationally

expensive. This can make an analysis impossible to carry out in a reasonable period of time. So an effort must be made to increase the speed at which waveforms can be generated such as creating approximate waveforms in the domain of analysis. Algorithmic and computational speed ups must also be sought. We also need to be concerned with signals from unknown sources. There will also be signals for which we have no model and therefore no way to generate their waveform. To characterize un-modeled signals and to distinguish them as astrophysical or instrumental in origin Bayeswave [43] was developed for LIGO. Glitches are guaranteed to exist in LISA data as LISA Pathfinder has already demonstrated and so we must develop similar capabilities to handle this.

I now outline the rest of this manuscript where references indicate original work associated with this thesis.

To assess the science that can be done with LISA we need to estimate the signals of interesting systems and LISA's sensitivity to said signal. In Chapter 1 we demonstrate how to construct sensitivity curves [59] to allow for such estimations. We also calculate approximate signals—with code provided on Github—so that astrophysicists can readily calculate signal-to-noise ratios for their favorite source.

In Chapter 2 we find that LISA's new proposed design will allow for the detection of tens of thousands of GBs [24] and estimate which of their properties will be characterized and how well. The unresolved binaries constitute confusion noise for which we provide estimates.

In chapter 3 we theoretically study the implications of modeling the GBs in the global analysis [60]. The unresolved GB confusion noise gets amended by subtraction errors in the resolved GBs. We demonstrate that this subtraction error biases our estimate of the noise and inflates the signal parameter errors. We further show that this subtraction error inflates the uncertainty in signals that overlap the resolved GBs,

such as BBHs that overlap thousands of GB signals.

Chapter 4 explores the scenario of a GB accompanied by another star i.e. a triple system [27] in a hierarchical orbit. In an hierarchical orbit the inner binary's center of mass is accelerated, alternately redshifting and blueshifting the gravitational wave signal. We determine how tight the outer orbit must be to detect this effect. We find out which parameters of the outer orbit can be measured due to this mechanism and how well the parameters are determined.

When performing a Bayesian analysis to the data we must generate the gravitational wave signal perhaps millions of times. We must therefore have quick signal generators to make this analysis practical. In chapter 5 we develop a fast frequency-domain model in the stationary phase approximation (SPA) for inspiraling binaries of arbitrary eccentricity [61]. We verify that the SPA accurately approximates the true answer and develop a novel numerical scheme which maximizes the time and phase at coalescence for this model given data.

In chapter 6 we begin investigations into un-modeled signals [62]. We model gravitational wave bursts and instrumental glitches in LISA by a single sine-Gaussian wavelet. We determine how well signals can be reconstructed and whether we can ascertain the signal's origin, i.e. whether it is instrumental or astrophysical.

Lastly, in chapter 7 we summarize this dissertation and discuss future work needed to solve LISA's global analysis problem.

CHAPTER ONE

THE CONSTRUCTION AND USE OF LISA SENSITIVITY CURVES

Contributions of Authors and Co-Authors

Manuscript in Chapter 1

Author: Travis Robson

Contributions: Wrote the sky dependent estimates section. Developed the open source Python code. Advised Chang Liu and co-developed sky dependent code.

Co-Author: Dr. Neil Cornish

Contributions: Conceived the design study. Performed useful calculations for the study. Wrote the manuscript and edited it.

Co-Author: Chang Liu

Contributions: Helped develop the sky dependent code.

Manuscript Information Page

Travis Robson, Neil Cornish, Chang Liu

Classical and Quantum Gravity

Status of Manuscript:

Officially submitted to a peer-reviewed journal

Submitted to IOP Publishing

Abstract

The Laser Interferometer Space Antenna (LISA) will open the mHz band of the gravitational wave spectrum for exploration. Sensitivity curves are a useful tool for surveying the types of sources that can be detected by the LISA mission. Here we describe how the sensitivity curve is constructed, and how it can be used to compute the signal-to-noise ratio for a wide range of binary systems. We adopt the 2018 LISA Phase-0 reference design parameters. We consider both sky-averaged sensitivities, and the sensitivity to sources at particular sky locations. The calculations are included in a publicly available *Python* notebook.

Introduction

We describe the construction and use of LISA sensitivity curves, the computation of signal-to-noise ratios, and how to plot signal strengths against the sensitivity curve. Figure 1 shows an example of a sensitivity/source plot taken from the LISA L3 mission proposal [63]. The idea, in plotting signal and noise curves in this manner, is that the height a signal is above the sensitivity curve indicates how loud it will be.

The literature on this topic can be very confusing, with a profusion of conventions and notation. Unfortunately, some of the choices that have now become standard are misleading, but it is probably too late to change the conventions now. Sensitivity curves are useful for making a quick assessment of what signals may be detectable. While not used for actual data analysis, the sensitivity curve, and signal representations that are shown with them, are designed to represent the quantities that are used in the data analysis. See Ref. [64] for a review of gravitational wave sensitivity curves. While useful, the sky-averaged sensitivity can be misleading as there is often significant variation in the sensitivity with sky location. To this end,

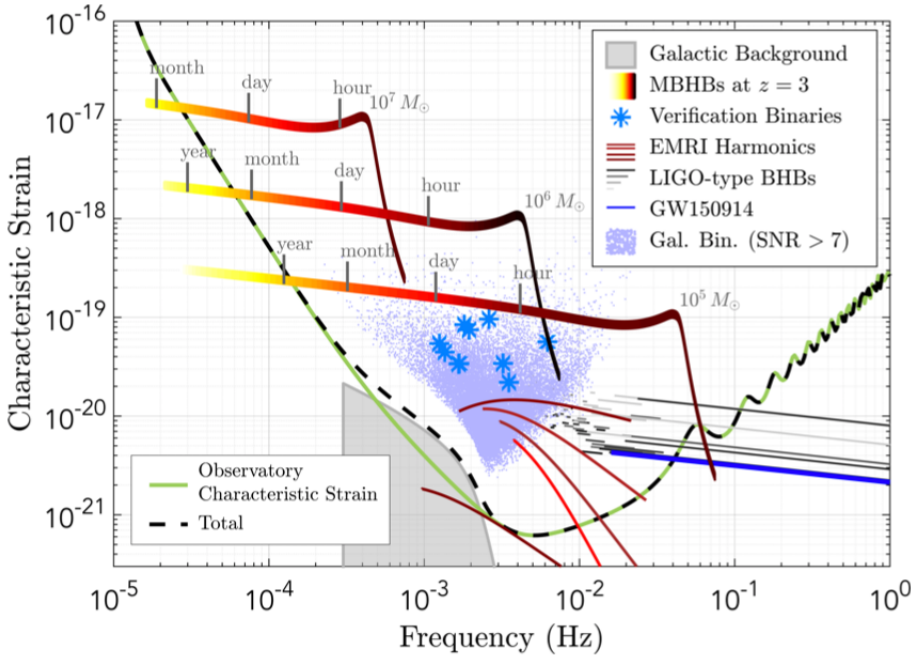

Figure 1.1: A plot taken from the LISA L3 mission proposal showing the expected sensitivity (green line) and a variety of possible sources (various colors) in units of dimensionless characteristic strain.

we also provide expressions and tools for computing the signal-to-noise ratio as a function of sky location, averaged over inclination and polarization angles.

Python code and a Jupyter notebook for generating the results shown in this document can be downloaded from GitHub [65]. The Jupyter notebook can be edited and executed directly in your browser without the need to install any software using the MyBinder version [66].

1.1 Sensitivity Curves

The LISA sensitivity curve can be well approximated by the equation

$$S_n(f) = \frac{10}{3L^2}\left(P_{\rm OMS}(f) + \frac{4P_{\rm acc}(f)}{(2\pi f)^4}\right)\left(1 + \frac{6}{10}\left(\frac{f}{f_*}\right)^2\right) + S_c(f)\,, \qquad (1.1)$$

where $L = 2.5$ Gm, $f_* = 19.09$ mHz, and expressions for $P_{\rm OMS}(f)$, $P_{\rm acc}(f)$ and $S_c(f)$ are given in equations (1.10), (1.11) and (1.14) below. Here we explain how this curve is computed and how it can be used (and sometimes mis-used).

The simplest type of sensitivity curve, and the one used by the ground-based detector community, is the power spectral density of the detector noise $P_n(f)$, or the amplitude spectral density $\sqrt{P_n(f)}$. The mean-squared noise in the frequency band $[f_1, f_2]$ is just the integral of $P_n(f)$ over that band. But for a detector like LISA, where signals may have wavelengths that are shorter than the arms of the detector, it is conventional to include the ensuing arm-length penalty in the sensitivity curve [67]. The strain spectral sensitivity is then defined in terms of the square root of the effective noise power spectral density

$$S_n(f) = \frac{P_n(f)}{\mathcal{R}(f)}, \tag{1.2}$$

where $\mathcal{R}(f)$ is the sky and polarization averaged signal response function of the instrument. The signal response function $\mathcal{R}(f)$ relates the power spectral density of the incident gravitational wave signals to the power spectral density of the signal recorded in the detector. As such, it might have been more logical to include this factor in the expression of the signals, but early on it was decided to apply the inverse of this factor to the noise power to define a sensitivity curve - c'est la vie.

The response function can be computed by working in the frequency domain, where the gravitational wave amplitude in the detector, $\tilde{h}(f)$, is related to the plus and cross gravitational wave amplitudes via

$$\tilde{h}(f) = F^{+}(f)\tilde{h}_{+}(f) + F^{\times}(f)\tilde{h}_{\times}(f)\,, \tag{1.3}$$

where $F^{+}(\theta, \phi, \psi, f)$ and $F^{\times}(\theta, \phi, \psi, f)$ are the (complex) frequency dependent

detector response functions, which depend on the sky location (θ, ϕ) and polarization angle ψ of the source. The sky/polarization averaged spectral power of the signal in the detector, $\langle \tilde{h}(f)\tilde{h}^*(f)\rangle$ is related to the raw spectral signal power $|\tilde{h}_+(f)|^2+|\tilde{h}_\times(f)|^2$ by the response function:

$$\begin{aligned}\langle \tilde{h}(f)\tilde{h}^*(f)\rangle &= \langle F^+(f)F^{+*}(f)\rangle|\tilde{h}_+(f)|^2 + \langle F^\times(f)F^{\times *}(f)\rangle|\tilde{h}_\times(f)|^2 \\ &= \mathcal{R}(f)\left(|\tilde{h}_+(f)|^2 + |\tilde{h}_\times(f)|^2\right) \end{aligned} \tag{1.4}$$

where $\mathcal{R}(f) = \langle F^+(f)F^{+*}(f)\rangle = \langle F^\times(f)F^{\times *}(f)\rangle$, and the angle brackets indicate the sky/polarization average

$$\langle X\rangle \equiv \frac{1}{4\pi^2}\int_0^\pi d\psi \int_0^{2\pi} d\phi \int_0^\pi X\, \sin\theta\, d\theta. \tag{1.5}$$

For a right-angle interferometer operating in the long wavelength limit, such as LIGO/Virgo, the antenna patterns are real and independent of frequency, and are given by

$$\begin{aligned} F^+ &= \frac{1}{2}(1+\cos^2\theta)\cos(2\phi)\cos(2\psi) - \cos\theta \sin 2\phi \sin 2\psi \\ F^\times &= \frac{1}{2}(1+\cos^2\theta)\cos(2\phi)\sin(2\psi) + \cos\theta \sin 2\phi \cos 2\psi\ . \end{aligned} \tag{1.6}$$

For LIGO we have

$$\mathcal{R} = \langle F^{+2}\rangle = \langle F^{\times 2}\rangle = \frac{1}{32}\int_{-1}^{1}(1+6x^2+x^4)dx = \frac{1}{5}\,. \tag{1.7}$$

In the LIGO literature this factor is applied to the signals, leaving the sensitivity curve to be just the power spectral density of the noise. The full expressions for $F^+(f)$ and $F^\times(f)$ for the Michelson-style interferometry signals for LISA are much

more complicated than those for LIGO (they are given in equations (5), (6), (16) and (17) of Ref. [68].) For a 3-arm LISA, there are two independent channels for $f < f_*$ and three for $f > f_*$, where $f_* = c/(2\pi L)$ is the transfer frequency [69]. For the current LISA design, $L = 2.5$ Gm, and $f_* = 19.09$ mHz. The standard convention is to define $\mathcal{R}(f)$ as being summed over the plus and cross channels. For sources that have frequency components $f > f_*$, it is more accurate to consider the 3-channel expressions given in Ref. [69]. The full expression for $\mathcal{R}(f)$ is not known in closed form, but to leading order is given by

$$\mathcal{R}(f) = \frac{3}{10} - \frac{507}{5040}\left(\frac{f}{f_*}\right) + \dots \tag{1.8}$$

The first term, 3/10, is a factor of $2 \times \sin^2(60^\circ) = 3/2$ larger than the corresponding LIGO result due to the LISA having two low-frequency channels, and arms that make an angle of 60°, as opposed to the 90° angle for LIGO. The full expression for $\mathcal{R}(f)$ has to be computed numerically [67], and has the form shown in Figure 2. The transfer function can be well-fit by the curve

$$\mathcal{R}(f) = \frac{3}{10}\frac{1}{(1 + 0.6(f/f_*)^2)} \ . \tag{1.9}$$

Note that many publications quote the number 3/20 for the low frequency limit of $R(f)$. The factor of two larger value quoted in eqn. (1.9) comes from summing over the two independent low-frequency data channels.

The current "official" model for the power spectral density of the LISA noise $P_n(f)$ is based on the Payload Description Document, and is referenced in the "LISA Strain Curves" document LISA-LCST-SGS-TN-001. The single-link optical

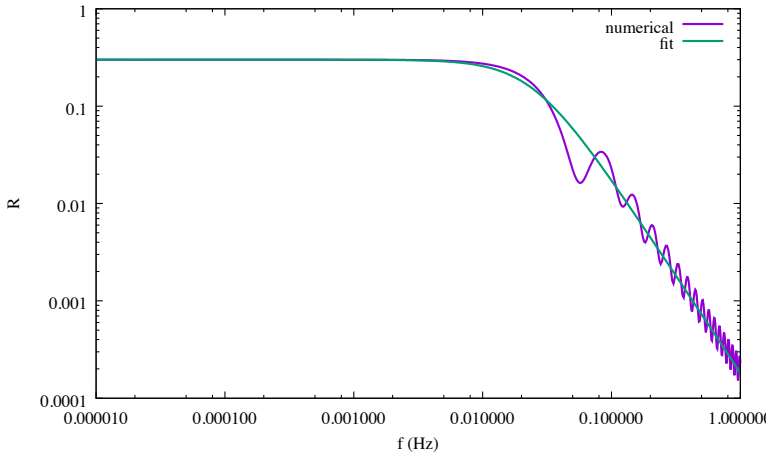

Figure 1.2: The signal transfer function $\mathcal{R}(f)$ for the combination of two Michelson-style LISA data channels, and the analytic fit from equation (1.9).

metrology noise is quoted as

$$P_{\rm OMS} = (1.5 \times 10^{-11}\,{\rm m})^2 \left(1 + \left(\frac{2\,{\rm mHz}}{f}\right)^4\right)\,{\rm Hz}^{-1}\,, \tag{1.10}$$

and the single test mass acceleration noise is quoted as

$$P_{\rm acc} = (3 \times 10^{-15}\,{\rm m\,s^{-2}})^2 \left(1 + \left(\frac{0.4\,{\rm mHz}}{f}\right)^2\right)\left(1 + \left(\frac{f}{8\,{\rm mHz}}\right)^4\right)\,{\rm Hz}^{-1}\,. \tag{1.11}$$

The total noise in a Michelson-style LISA data channel is then [68]

$$P_n(f) = \frac{P_{\rm OMS}}{L^2} + 2(1 + \cos^2(f/f_*))\frac{P_{\rm acc}}{(2\pi f)^4 L^2}\,. \tag{1.12}$$

Note that the Michelson-style response has four contributions from the optical metrology noise and sixteen from the test mass acceleration noise. We convert from displacement to strain by dividing by the round-trip light travel distance $2L$, so, for example, the factor of $4P_{\rm OMS}$ gets divided by $(2L)^2$, leading to the expression seen in (1.12). The same factor of $1/(2L)$ is also applied to the path-length change caused by

the gravitational wave, so it cancels out in the likelihood function and the SNR, and the choice to divide by $2L$ is an unimportant convention. A good analytic model for the sensitivity curve that is sufficient for most purposes is given by combining (1.9) and (1.12):

$$S_n(f) = \frac{10}{3L^2}\left(P_{\rm OMS} + 2(1+\cos^2(f/f_*))\frac{P_{\rm acc}}{(2\pi f)^4}\right)\left(1+\frac{6}{10}\left(\frac{f}{f_*}\right)^2\right). \tag{1.13}$$

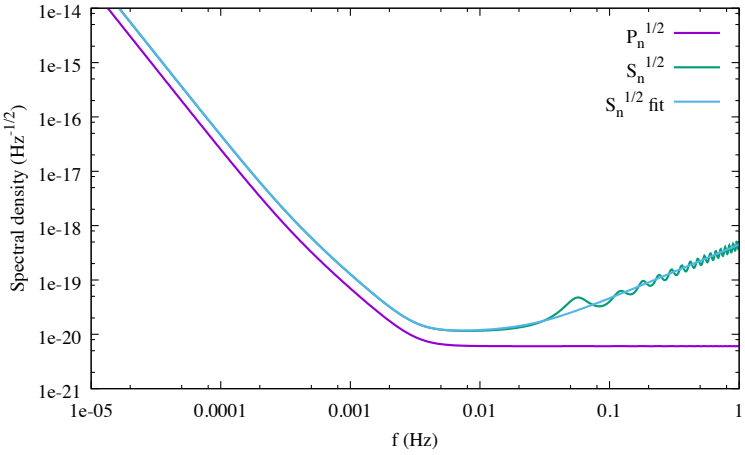

Figure 1.3: The amplitude spectral density of the noise, and the corresponding sensitivity curve, found by dividing $P_n(f)$ by $\mathcal{R}(f)$. The analytic fit to $S_n(f)$ given in equation (1.1) is also shown.

In addition to the instrument noise, unresolved galactic binaries will act as an effective noise source (though one that is not stationary). The galactic confusion noise goes down as the mission progresses and more foreground sources are removed. Estimates for the confusion noise using the new LISA design are given in Ref. [70], and are well fit by the function

$$S_c(f) = A\, f^{-7/3}\, e^{-f^\alpha + \beta f \sin(\kappa f)}\, [1+\tanh(\gamma(f_k - f))]\ \text{Hz}^{-1} \tag{1.14}$$

with fit parameters given in Table 1. Note that the amplitude quoted here is half the

	6 mo	1 yr	2 yr	4 yr
α	0.133	0.171	0.165	0.138
β	243	292	299	-221
κ	482	1020	611	521
γ	917	1680	1340	1680
f_k	0.00258	0.00215	0.00173	0.00113

Table 1.1: Parameters of the analytic fit to the Galactic confusion noise as described by equation (1.14). The amplitude A has been fixed to 9×10^{-45}. Note that the knee frequency f_k decreases with observation time and γ increase with observation time, leading to a steeper drop off in confusion noise.

value quoted in Ref. [70] since here we are using two-channel sensitivity curves. The full sensitivity curve is found by adding $S_c(f)$ to $S_n(f)$.

Figure 4 shows the contribution of the galactic confusion noise assuming a 4-year mission, along with the updated sensitivity curve that includes the confusion noise. Note that the confusion noise shown here is an average value - in practice it will vary over a year as the LISA antenna pattern sweeps across the galaxy.

1.2 Binary Sources

The majority of LISA sources will be binaries of various masses and mass ratios. For simplicity we will focus here on quasi-circular, non-spinning comparable mass binaries, and only consider the dominant quadrupole harmonic. Extreme mass ratio binaries, which may be highly eccentric, require a more involved treatment. While we ignore spin, the signal-to-noise ratios we compute should be good to within a factor

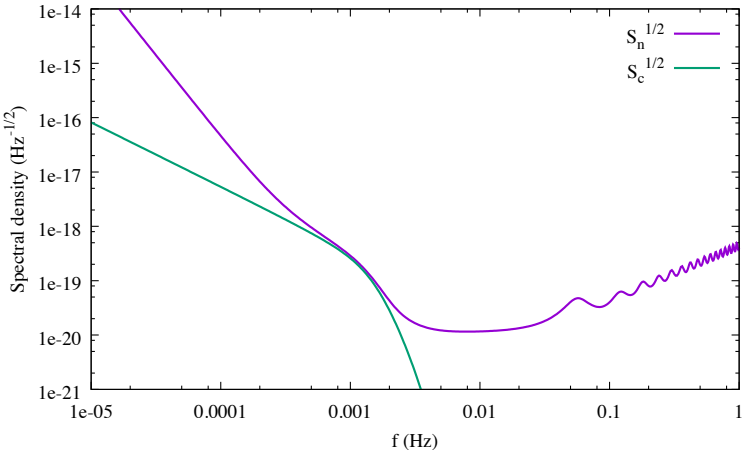

Figure 1.4: The amplitude spectral density of the galactic noise, $S_c^{1/2}$, and the full sensitivity curve combining the instrument noise and the galactic confusion noise, $S_n^{1/2}$, for a 4-year mission lifetime.

of two or so for spinning systems. Our model for the waveforms is then

$$\begin{aligned}\tilde{h}_+(f) &= A(f)\frac{(1+\cos^2\iota)}{2}\,e^{i\Psi(f)}\\ \tilde{h}_\times(f) &= iA(f)\cos\iota\,e^{i\Psi(f)}\,,\end{aligned} \tag{1.15}$$

where ι describes the inclination of the orbit relative to the line of sight, and $A(f)$ and $\Psi(f)$ are the amplitude and phase of the wave. To compute the sky/polarization averaged SNR we only need to know $A(f)$, and in some cases, how the frequency evolves with time, $f(t)$.

Earlier we related the sky and polarization averaged power spectral density of the signal to the power spectral density seen in the detector via equation (1.4). For binary systems it is natural to extend the angle averaging to include the inclination angle:

$$\langle\tilde{h}(f)\tilde{h}^*(f)\rangle = \mathcal{R}(f)A^2(f)\frac{1}{2}\int_{-1}^{1}\left(\frac{(1+x^2)^2}{4}+x^2\right)dx = \frac{4}{5}\mathcal{R}(f)A^2(f)\,. \tag{1.16}$$

Note that for LIGO we recover the well-known pre-factor $\sqrt{(4/5)\mathcal{R}(f)} = 2/5$ that is applied to the GW amplitude to account for averaging over the source location and orientation [71].

The amplitude signal-to-noise ratio ρ for a deterministic signal $\tilde{h}(f)$ is given by

$$\rho^2 = 4 \int \frac{|\tilde{h}(f)|^2}{P_n(f)}\, df = 4 \int_{f=0}^{\infty} \frac{f|\tilde{h}(f)|^2}{P_n(f)}\, d(\ln f)\,. \tag{1.17}$$

Averaging over sky location, inclination and polarization we have

$$\overline{\rho^2} = \frac{16}{5} \int \frac{fA^2(f)}{P_n(f)}\, d(\ln f) = \frac{16}{5} \int \frac{(2fT)S_h(f)}{S_n(f)}\, d(\ln f)\,, \tag{1.18}$$

Where T is the observation time and $S_h(f)$ is the one-sided, angle averaged, power spectral density of the signal,

$$S_h(f) = \frac{A^2(f)}{2T}\,. \tag{1.19}$$

If you took a Fourier transform of the data, $d = h + n$, then ignoring any correlations between the signal and the noise, the power spectral density of the data would equal to $S_d(f) = S_h(f) + P_n(f)$. In other words, $S_h(f)$ is the power spectral density of the signal. The factor of $(2fT)$ that appears in the expression for the optimal signal-to-noise shows that the signal is effectively boosted relative to the noise by using templates to coherently extract the signal. Rather than plotting the signal power directly (which often lies below the sensitivity curve), the convention is to plot $h^2_{\rm eff} = 16f(2fT)S_h(f)/5$, to account for the boost we get from the coherent signal extraction.

For the waveform model we use the original phenomenological inspiral-merger-ringdown (IMR) model, known as PhenomA [72]. While more accurate models now

exist, such as the latest PhenomP model [73, 74], which includes spin-precession, PhenomA is good enough for making graphs and estimating SNRs. The PhenomA amplitude is given by

$$A(f) \equiv \sqrt{\frac{5}{24}}\frac{(G\mathcal{M}/c^3)^{5/6} f_0^{-7/6}}{\pi^{2/3}(D_L/c)} \begin{cases} \left(\frac{f}{f_0}\right)^{-7/6} & \text{if } f < f_0 \\ \left(\frac{f}{f_0}\right)^{-2/3} & \text{if } f_0 \le f < f_1 \\ w\,\mathcal{L}(f,\, f_1,\, f_2) & \text{if } f_1 \le f < f_3\,, \end{cases} \tag{1.20}$$

where

$$f_k \equiv \frac{a_k\eta^2 + b_k\eta + c_k}{\pi(GM/c^3)}\,, \tag{1.21}$$

$$\mathcal{L}(f, f_1, f_2) \equiv \left(\frac{1}{2\pi}\right)\frac{f_2}{(f-f_1)^2 + f_2^2/4}\,, \tag{1.22}$$

and

$$w \equiv \frac{\pi f_2}{2}\left(\frac{f_0}{f_1}\right)^{2/3}\,. \tag{1.23}$$

Here $M = m_1 + m_2$ is the total mass, $\eta = m_1 m_2/M^2$ is the symmetric mass ratio and $\mathcal{M} = (m_1 m_2)^{3/5}/M^{1/5}$ is the chirp mass. The G's and c's have been included for those that are not used to working in natural units. Note that the combinations GM/c^3 and D_L/c both have units of time. A useful number to remember is that the mass of the Sun, $GM_\odot/c^3$, is approximately 5 microseconds in natural units. The coefficients for the transition frequencies f_k are given in Table 2. Roughly speaking, f_0 is the merger frequency, f_1 is the ringdown frequency, f_2 is decay-width of the ringdown and f_3 is the cut-off frequency.

	a_k	b_k	c_k
f_0	2.9740×10^{-1}	4.4810×10^{-2}	9.5560×10^{-2}
f_1	5.9411×10^{-1}	8.9794×10^{-2}	1.9111×10^{-1}
f_2	5.0801×10^{-1}	7.7515×10^{-2}	2.2369×10^{-2}
f_3	8.4845×10^{-1}	1.2848×10^{-1}	2.7299×10^{-1}

Table 1.2: Polynomial coefficients of the transition frequencies.

The final ingredient we need for computing the SNR is the frequency range covered by the signal. For comparable mass black holes, with $M > 10^4 M_\odot$, the signal will sweep across the LISA band and merge in less than the mission lifetime. However, for lower mass systems, such as stellar origin black holes that will merger in the LIGO band a decade or so later, or for white dwarf binaries, which may be millions of years from merger, we need to specify the start and end frequencies for the SNR integration. To leading post-Newtonian order, the frequency as a function of time is given by

$$f(t) = \frac{1}{8\pi(G\mathcal{M}/c^3)}\left(\frac{5(G\mathcal{M}/c^3)}{t-t_c}\right)^{3/8}, \tag{1.24}$$

where t_c is the time of coalescence. For example, an equal mass binary at $z = 3$ with a total source-frame mass of $M = 10^6\,M_\odot$ will have a GW frequency of 2.93×10^{-5} Hz one year prior to merger. Note that it is the detector frame mass, $M_z = M(1+z)$, that should be used in equations (1.20) and (1.24). For these high mass systems it makes sense to plot $h_{\rm eff}$ across the entire LISA band, and not worry about setting limits in the SNR integration. Tracks in $\sqrt{S_h}$ and $h_{\rm eff}$ for the aforementioned source are shown in Figure 5. In contrast, a source similar to GW150914 that is 5 years

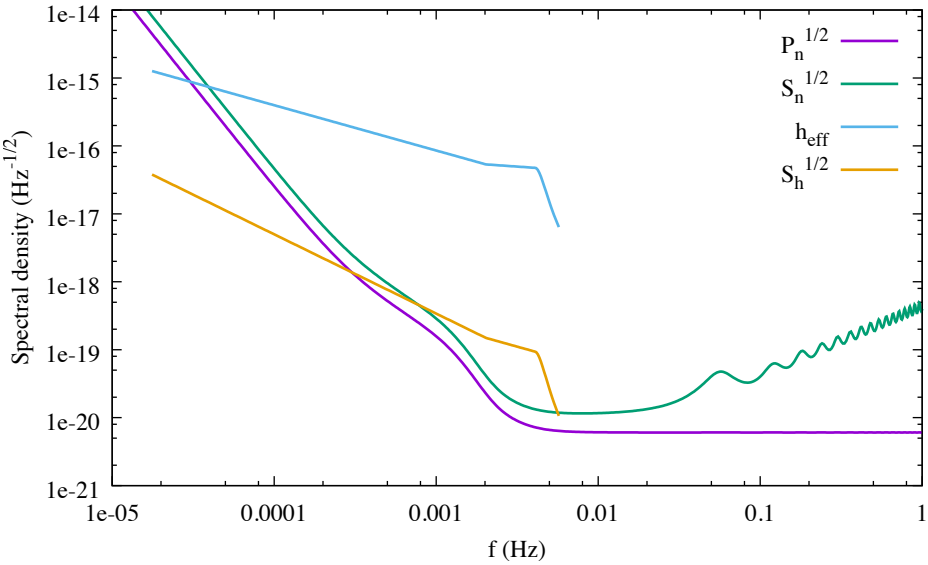

Figure 1.5: The amplitude spectral density of the noise $\sqrt{P_n}$, and the amplitude sensitivity curve $\sqrt{S_n}$ are plotted against the raw strain spectral density $\sqrt{S_h}$ and the effective strain spectral density $h_{\rm eff}$ for an equal mass black hole binary at $z = 3$ with source frame total mass $M = 10^6\, M_\odot$. This system is so bright that even its raw amplitude will be visible in the detector. However, the effective amplitude $h_{\rm eff}$ that appears in the numerator of the SNR calculation better communicates the true brightness of the source. The area between the $h_{\rm eff}$ curve and the $S_n^{1/2}$ curve roughly corresponding to the optimal SNR of 2626. Note that this graph differs slightly from the one shown in Figure 1, which plots dimensionless characteristic strain $h_c(f) = \sqrt{fS(f)}$ rather than strain spectral density $\sqrt{S}$.

from merger when LISA turns on will sweep from $f = 16$ mHz to $f = 29$ mHz over the nominal 4 year mission lifetime.

For galactic binaries the time to merger is typically very large compared to the mission lifetime, and the frequencies will evolve very little over the course of the mission. Taylor expanding (1.24) we find

$$f(t) = f_{\rm in} + \frac{96}{5}\pi^{8/3}(G\mathcal{M}/c^3)^{5/3} f_{\rm in}^{11/3}(t - t_{\rm in}) + \dots\,, \tag{1.25}$$

where $f_{\rm in}$ is the GW frequency at the start of the observation, at time $t_{\rm in}$. For typical

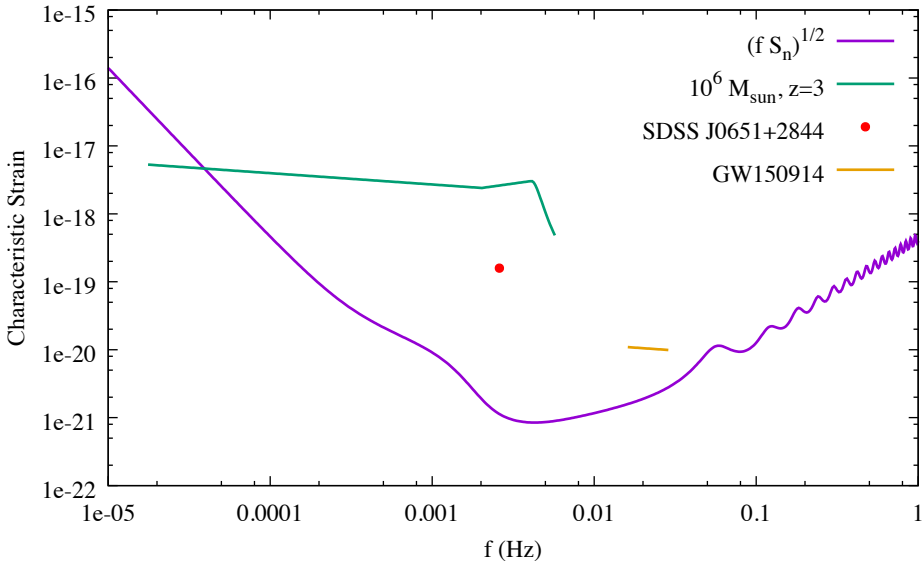

Figure 1.6: The sensitivity curve in terms of characteristic strain, $\sqrt{fS_n}$ is compared to three types of signal: an equal mass black hole binary at $z = 3$ with source-frame total mass $M = 10^6\,M_\odot$; the galactic verification binary SDSS J0651+2844 observed for 4 years; and a signal similar to the first LIGO detection GW150914 if the LISA observation started 5 years prior to merger and continued for 4 years.

galactic binaries the change in frequency Δf during the mission lifetime is so small that it no longer makes sense to the plot the signals as tracks. Rather, the signals are plotted as points with an amplitude $h_{\rm GB}$ that follows from evaluating the SNR integral:

$$\overline{\rho^2} = \frac{16}{5}\int_{f_{\rm in}}^{f_{\rm in}+\Delta f} \frac{A^2(f)}{S_n(f)}\,df \approx \frac{16}{5}\frac{\Delta f A^2(f_{\rm in})}{S_n(f_{\rm in})} \equiv \frac{h^2_{\rm GB}(f_{\rm in})}{S_n(f_{\rm in})} \tag{1.26}$$

where

$$h_{\rm GB} = \frac{8\,T^{1/2}(G\mathcal{M}/c^3)^{5/3}\pi^{2/3}f_{\rm in}^{2/3}}{5^{1/2}(D_L/c)}\,. \tag{1.27}$$

For example, SDSS J0651+2844 which has $D_L \sim 1$ Kpc, $m_1 \sim 0.5\,M_\odot$, $m_2 \sim 0.25\,M_\odot$, and $f_{\rm in} = 2.6$ mHz, will produce a strain spectral density of $h_{\rm GB} = 2.8 \times 10^{18}\,{\rm Hz}^{-1/2}$ and have an SNR of 140 assuming a 4 year mission lifetime. Of course, it not strictly correct to compute angle averaged SNRs for a source with a known sky location and

orientation, nor does it make much sense to plot its amplitude against against an all-sky averaged sensitivity curve, but doing so allows us to put all LISA sources on a single graph.

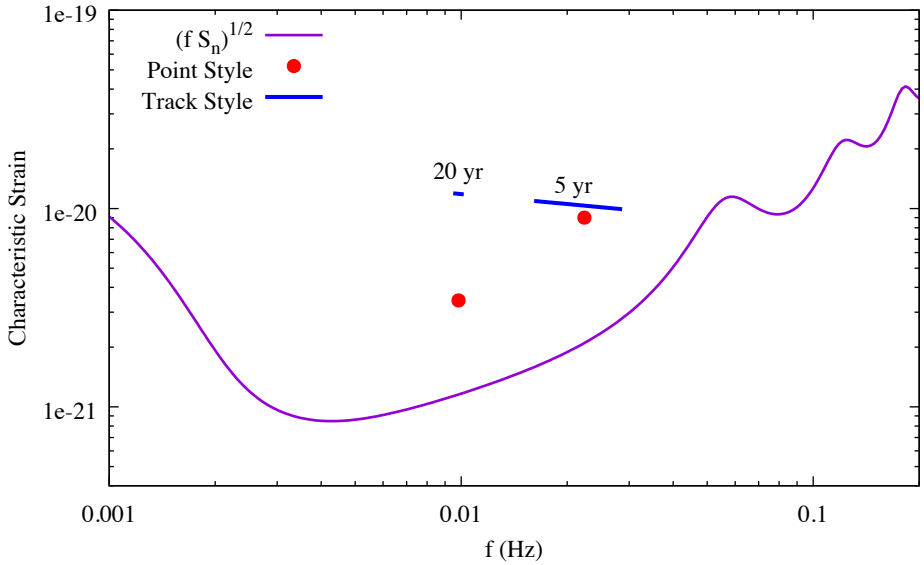

Figure 1.7: The characteristic strain produced by a GW150914 type system that is either 20 years or 5 years from merger at the beginning of the LISA observation. The track-style representation, where the SNR is estimate from the area under the curve, is compared to the point style representation used for slowly evolving white-dwarf binaries, where the SNR is given by the ratio of the height of the sensitivity curve and the hight of the point.

The differing conventions between how slowly evolving and rapidly evolving signals are plotted can be problematic for stellar origin black hole binaries (SOBHBs). For example, if a GW150914 type system was 20 years from merger when LISA started observations, it would be emitting at a gravitational wave frequency of 9.5 mHz, and four years later it would be emitting at 10.4 mHz, producing a track that runs for just $\Delta \ln f = 0.09$. Since the frequency range is so short, the questions becomes do we treat the system as evolving, and plot a track as we do for massive black holes, or do we treat the system as non-evolving, and plot a point as we do for galactic binaries?

Figure 1.7 shows that the two choices paint an inconsistent picture. If the track is longer, such as for a system that is 5 years from merger, the two representations look more consistent. To arrive at consistent representations, where sources appear at almost the same height when shown as evolving tracks or non-evolving points, we recommend switching from tracks to points when $\Delta \ln f < 0.5$.

1.2.1 EMRIs and other complicated signals

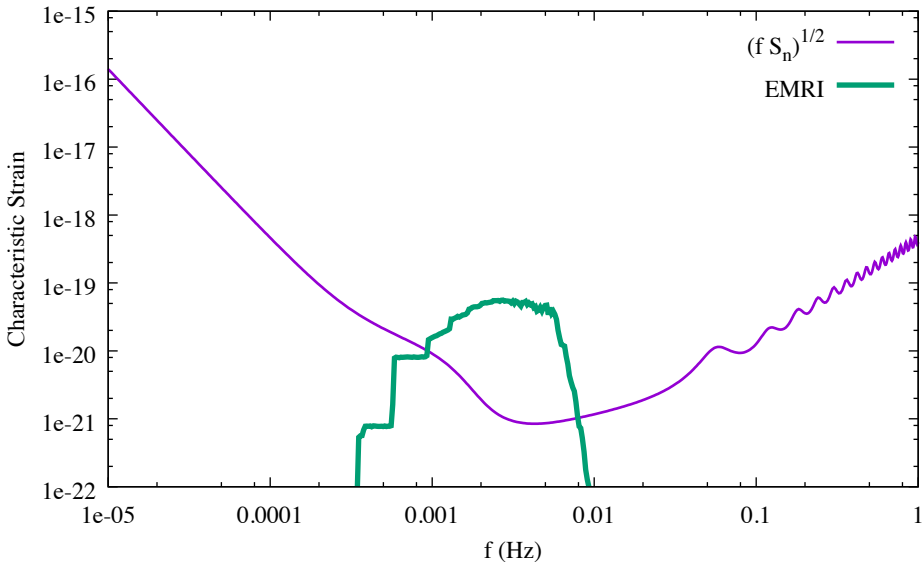

Figure 1.8: The characteristic strain produced by a $20\,M_\odot$ - $10^6 M_\odot, \chi = 0.5$ EMRI at 4 Gpc.

Some sources produce signals that can not be accurately described by simple frequency domain models. Extreme Mass Ratio Inspirals and rapidly precessing spinning black hole binaries fall into that category. The inclination and polarization averaging we used for quasi-circular binaries is not applicable these systems, but for simplicity we still plot these signals against the standard sky and polarization averaged sensitivity curve. Writing the sky and polarization averaged signal-to-noise

squared as

$$\overline{\rho}^2_{\Omega,\psi} = 4\int \frac{|\tilde{h}(f)|^2}{S_n(f)}\,df = \int_{f=0}^{\infty} \frac{4f^2|\tilde{h}(f)|^2}{(fS_n(f))}\,d(\ln f)\,, \tag{1.28}$$

indicates that an appropriate quantity to plot against the characteristic sensitivity $(fS_n(f))^{1/2}$ is the dimensionless characteristic strain

$$h_c(f) = 2f\left(|\tilde{h}_+(f)|^2 + |\tilde{h}_\times(f)|^2\right)^{1/2}. \tag{1.29}$$

We generate the Barycenter signals $h_+(t)$ and $h_\times(t)$, Fourier transform, and form $h_c(f)$. To beautify the plots we smooth the numerically generated $h_c(f)$ using a running average over ~ 100 frequency bins. As an example, we generated augmented analytic kludge (AAK) EMRI waveforms [75] using the code provided at GitHub [76] for a $20\,M_\odot$ stellar remnant black hole falling into a $10^6 M_\odot, \chi = 0.5$ spinning supermassive black hole at distance of 4 Gpc, starting 4 years before merger with an eccentricity of 0.5. This system has SNR = 52, and the effective strain shown in Figure 1.8.

1.2.2 Sky Dependent Estimates

The sky averaged signal-to-noise ratios are useful for a first brush look at what systems might be detectable, but the signal-to-noise ratio can vary significantly across the sky, especially for short duration signals [77]. To incorporate the sky location dependence we must revert to using signals which have not been averaged over the sky location. The sky-location dependent SNRs are particularly useful for sources with known locations, such as the galactic verification binaries. We will continue to average over inclination angle and polarization angle as these are usually not well constrained.

The Michelson-type signal with spacecraft 1 at the vertex is given by

$$s_1(t) = \frac{\delta\ell_{12}(t-2L/c) + \delta\ell_{21}(t-L/c)}{2L} - \frac{\delta\ell_{13}(t-2L/c) + \delta\ell_{31}(t-L/c)}{2L} \ , \quad (1.30)$$

where the GW induced variation in LISA arm lengths between spacecraft i and j are given by

$$\frac{\delta\ell_{ij}(t)}{L} = \ \tfrac{1}{2}(1+\cos^2\iota) d_{ij}^{+}(t) h^{+}(\xi_i) + \cos\iota \, d_{ij}^{\times}(t) h^{\times}(\xi_i) \ . \quad (1.31)$$

The detector terms $d_{ij}^{+,\times} = d_{ij}^{+,\times}(t; \mathbf{x}_i, \mathbf{x}_j, \theta, \phi, \iota, \psi)$ [78] describes LISA's geometry through their dependence on the spacecraft position $\mathbf{x}_i$. The variable ξ_i define surfaces of constant gravitational wave phase at spacecraft i. The frequency domain representation of this signal can be found using the stationary phase approximation [79]

$$\frac{\widetilde{\delta\ell}_{ij}(f)}{L} = \ \left[\tfrac{1}{2}(1+\cos^2\iota) d_{ij}^{+}(t_*) + i\cos\iota \, d_{ij}^{\times}(t_*)\right] A(f) e^{i(\Psi(f)+\delta\Psi_i(f))} \quad (1.32)$$

where $A(f)$ and $\Psi(f)$ are the amplitude and phase. The stationary time is given by the relation $t_*(f) = \Psi'(f)/2\pi$. This is used to map $d_{ij}^{+}(t)$ to $d_{ij}^{+}(f)$ *etc.* The motion of the LISA detector also imparts a phase shift $\delta\Psi_i(f) = 2\pi f \hat{\mathbf{k}} \cdot \mathbf{x}_i(t_*(f))/c$, where $\hat{\mathbf{k}}$ defines the line of sight vector to the source. The PhenomA amplitude is given in equation (1.20). The phase is written in terms of a power series expansion that is motivated the post Newtonian expansion, with additional terms that are found by fitting to numerical relativity simulations:

$$\Psi(f) = 2\pi f t_0 + \phi_0 + \psi_0 f^{-5/3} + \psi_2 f^{-1} + \psi_3 f^{-2/3} + \psi_4 f^{-1/3} + \psi_6 f^{1/3}. \quad (1.33)$$

The expansion coefficients ψ_i depend on the masses and are given in terms of a numerical look-up table [72].

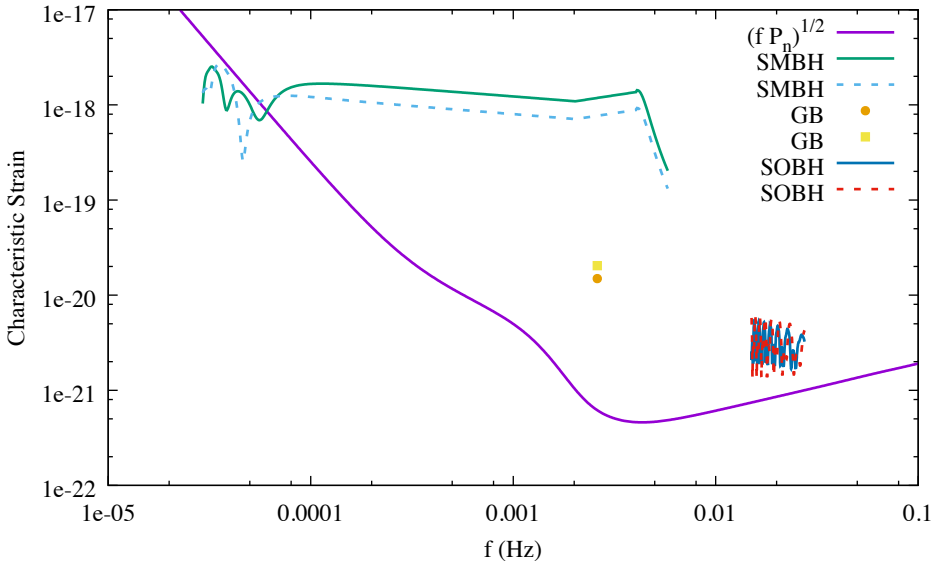

Figure 1.9: The effective amplitudes of a SMBH, GB and SOBH are shown for two sky locations. Sine there is no sky averaging they are plotted against the characteristic noise $\sqrt{fP_n(f)}$. The masses are the same as for the sources shown shown in Figure 1.6 for the sky averaged case.

The average over the source inclination and polarization angles, $\langle|\tilde{s}_1(f)|^2\rangle_{\iota,\psi}$, where the subscripts to the angle brackets denote which variables are averaged over, can be shown to be equivalent to computing the signal at two fixed values of inclination and polarization:

$$\left\langle|\tilde{s}_1(f)|^2\right\rangle_{\iota,\psi} = \frac{8}{5}\left(|\tilde{s}_1(f)|^2|_{\iota=\frac{\pi}{2},\psi=0} + |\tilde{s}_1(f)|^2|_{\iota=\frac{\pi}{2},\psi=\frac{\pi}{4}}\right) . \tag{1.34}$$

This allows us to compute the orientation averaged signal using just two calls to the waveform generator. The sky location dependent signal-to-noise ratio is then

$$\bar{\rho}^2_{\iota,\psi}(\theta,\phi) = 4\int_{f=0}^{\infty} \frac{f^2\left\langle|\tilde{s}_1(f,\theta,\phi)|^2\right\rangle_{\iota,\psi}}{fP_n(f)} d(\log f) . \tag{1.35}$$

To give a visual impression of the signal strength we plot the characteristic strain $\tilde{h}_{\rm eff}(f) = 2f \langle|\tilde{s}_1(f)|\rangle_{\iota,\psi}$ against the characteristic noise amplitude in the Michelson channel $\sqrt{fP_n(f)}$. For a galactic binary we take a similar approach to calculate SNRs as for the sky averaged case:

$$\bar{\rho}^2_{\iota,\psi}(\theta,\phi) = 4\int_{f_{\rm in}}^{f_{\rm in}+\Delta f} \frac{\langle|\tilde{s}_1(f)|^2\rangle_{\iota,\psi}}{P_n(f)} df \approx 4\frac{\langle|\tilde{s}_1(f_{\rm in})|^2\rangle_{\iota,\psi}}{P_n(f_{\rm in})}\Delta f \ . \tag{1.36}$$

An example of a sky-location dependent sensitivity plot is shown in Figure 1.9 for the same sources shown previously in Figure 1.6, but now at two different sky locations $(\theta = 0.5, \phi = 2.3)$ and $(\theta = 1.1, \phi = 1.5)$. The signal-to-noise ratio for these sky-dependent sources are 3106 and 2017 for the super massive black hole binary, 207 and 151 for the galactic binary, and 4.39 and 4.51 for the stellar origin black hole binary.

The oscillations in the tracks seen in Figure 1.9 are due to the time dependent antenna pattern. For the supermassive black hole system most of the modulation is seen at low frequencies where the system spends many months. The evolution of the signal becomes far more rapid as it sweeps to higher frequencies, and the detector is effectively stationary on this timescales, so the amplitude no longer oscillates. The oscillations are very pronounced for the stellar origin black hole binary, which slowly evolves over the entire 4 year observation period.

The sky location of the source plays a very large role in its detectability and our ability to characterize the source parameters. For example, the standard deviation of the SNR across the sky for the equal mass binary black hole system shown in Figure 1.6 is 76% of the sky averaged value. The galactic binary and LIGO binary had deviations of 77% and 39% of the sky averaged value respectively. In Figure 1.10 we show the variation in the SNR across the sky for two binary black hole, one with detector frame total mass $4 \times 10^6 M_\odot$, and another with detector frame total mass

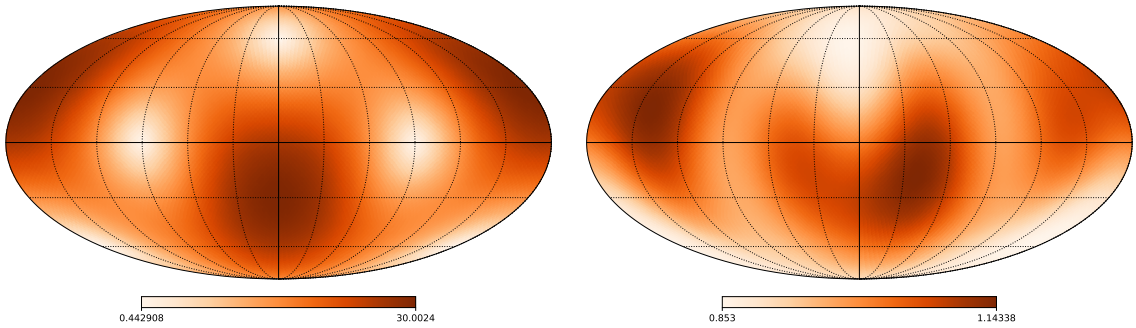

Figure 1.10: The sky location can greatly affect the signal-to-noise for a source. The maps show the SNR scaled by the all-sky average. The sky map on the left is for a $10^6 M_\odot$ black hole binary at $z = 3$, while the map on the right is for a $10^4 M_\odot$ black hole binary at $z = 0.6$. Note the large difference in dynamic range for the two color maps. The standard deviation of the scaled SNR is 76% for the $10^6 M_\odot$ system and 7% for the $10^4 M_\odot$ mass system. The variance is smaller for lower mass systems as the signal accumulates more slowly over time and partially averages out the time variation of the LISA antenna pattern.

$1.6 \times 10^4 M_\odot$. For the more massive system, most of the SNR is accumulated in the two days around merger, and since the detector is effectively stationary during this time, we recover the fixed quadrupolar antenna pattern. For the less massive system, where the merger lasts for a longer time, the sky map is more uniform.

Figure 1.11 shows the signal-to-noise grows in time for the stellar origin black hole binary for various sky locations. We see that the rate of SNR growth depends on the sky location at any given time due to the changing orientation of the LISA antenna pattern. If the sources lies in a sensitive region of the antenna pattern then the SNR will grow quickly compared to when the source lies in an insensitive region.

An additional factor that will impact the SNR as a function of sky location is the time variation of the galactic confusion noise. The confusion noise is loudest when the peak of the antenna pattern sweeps across the galaxy, and quietest when pointed away from the galaxy. The confusion noise will vary adiabatically such that $S_c(f, t)$. Using

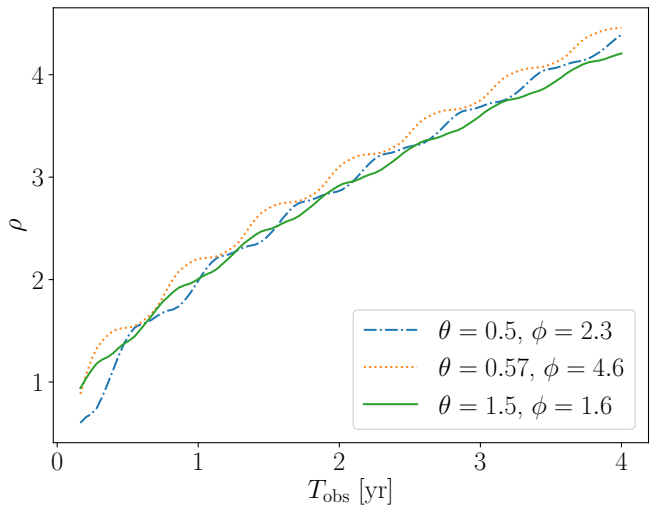

Figure 1.11: For the LIGO binary this figure displays how the signal-to-noise grows with time depending on the sky location of the binary. The rate of SNR gain also depend on the sky location. The green line demonstrates a more uniform growth in SNR over time while the dotted orange line shows periods of quick growth followed by periods of slow growth.

the stationary phase approximation, the time dependence gets mapped to frequency dependence via $t_*(f)$, which allows for the time variation of the confusion noise to be incorporated in the SNR integral (1.35). We defer an analysis of this effect to a future study.

1.3 Acknowledgments

We appreciate the input and feedback from Martin Hewitson, Emanuele Berti, Paul McNamara, and Davide Gerosa. TR and NJC appreciate the support of the NASA grant NNX16AB98G. CL acknowledges the support from the UCAS Joint PhD Training Program.

CHAPTER TWO

GALACTIC BINARY SCIENCE WITH THE NEW LISA DESIGN

Contributions of Authors and Co-Authors

Manuscript in Chapter 2

Author: Dr. Neil Cornish

Contributions: Conceived of the study design and wrote the manuscript.

Co-Author: Travis Robson

Contributions: Utilized code, developed by Neil Cornish in a previous study, to perform the analysis.

Manuscript Information Page

Neil Cornish, Travis Robson

Journal of Physics: Conference Series

Status of Manuscript:

Published in a conference series

Published by IOP Publishing

Published 2017, J. Phys.: Conf. Ser. **840** 012024

Abstract

Building on the great success of the LISA Pathfinder mission, the outlines of a new LISA mission design were laid out at the 11$^{\text{th}}$ International LISA Symposium in Zurich. The revised design calls for three identical spacecraft forming an equilateral triangle with 2.5 million kilometer sides, and two laser links per side delivering full polarization sensitivity. With the demonstrated Pathfinder performance for the disturbance reduction system, and a well studied design for the laser metrology, it is anticipated that the new mission will have a sensitivity very close to the original LISA design. This implies that the mid-band performance, between 0.5 mHz and 3 mHz, will be limited by unresolved signals from compact binaries in our galaxy. Here we use the new LISA design to compute updated estimates for the galactic confusion noise, the number of resolvable galactic binaries, and the accuracy to which key parameters of these systems can be measured.

2.1 Introduction

The first direct detection of gravitational waves by the advanced LIGO [1] and the spectacular success of the LISA Pathfinder mission [21] added a renewed energy to the 11$^{\text{th}}$ International LISA Symposium in Zurich. During the symposium, the LISA Consortium held a half-day meeting to begin the preparation of a formal mission proposal for the L3 launch opportunity in ESA's Cosmic Vision Program. The many decades of experience with LISA technologies and designs meant that the broad outlines of the mission quickly fell into place. The new LISA mission [80] is envisioned to comprise of three identical spacecraft in a triangular formation separated by 2.5 million km in an Earth trailing orbit, with six continuously operating laser links providing heterodyne laser interferometry with pm/$\sqrt{\text{Hz}}$ sensitivity along each

arm. The nominal design calls for 2 W of laser power being delivered to the optical system, with 30 cm telescopes transmitting and receiving the laser light between the spacecraft. The six laser links allow the synthesis of two Michelson-like channels that provide instantaneous measurements of the two gravitational wave polarization states, and a third Sagnac-like channel that is relatively insensitive to gravitational waves that can be used to monitor the average noise level in the detector.

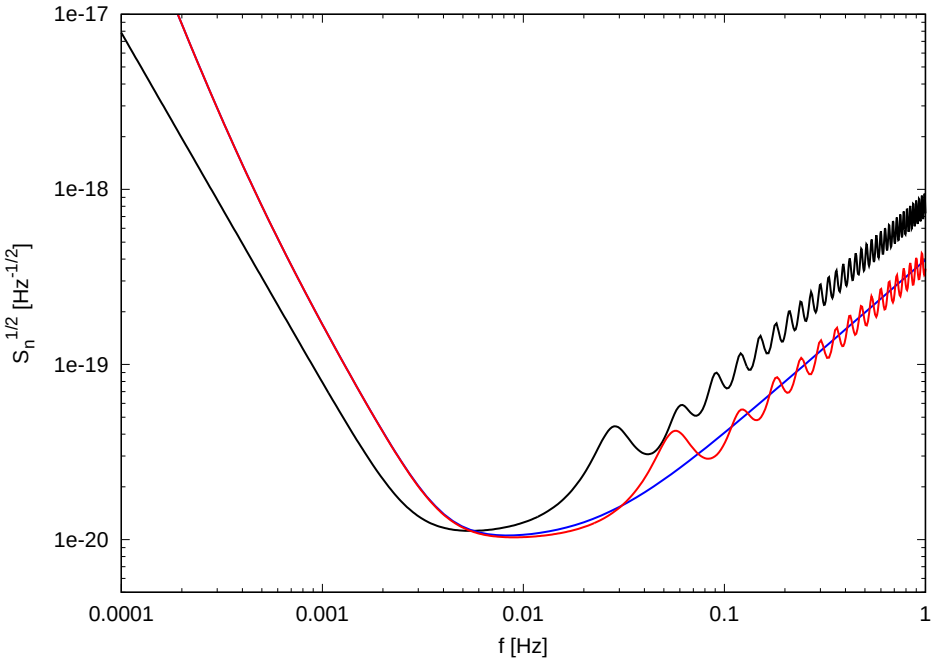

Figure 2.1: Sensitivity curves for the 1998 LISA design (black) and the new LISA design (red). Also shown is a simplified model for the sensitivity (blue), given in equation (1) of the text. The differences in the old and new sensitivity curves are almost entirely due to the reduction in arm length from 5 million km to 2.5 million km, which has the effect of reducing the low frequency sensitivity while enhancing the high frequency sensitivity.

Figure 1 compares the predicted Michelson-equivalent sensitivity of the LISA design from the 1998 pre-phase A report [81] and the new LISA design [80]. Also

shown is a simple approximation to the new sensitivity curve [82] given by

$$S_n(f) = \frac{20}{3}\frac{1}{L^2}\left(1+\left(\frac{f}{1.29 f_*}\right)^2\right)\left(S_p(f)+\frac{4S_a(f)}{(2\pi f)^4}\right). \tag{2.1}$$

Here $f_* = c/(2\pi L)$ is the transfer frequency, $L = 2.5$ million km is the arm length, $S_p(f) = 8.9\times10^{-23}\ \mathrm{m^2\,Hz^{-1}}$ is the white position noise, and

$$S_a(f) = 9\times10^{-30}\left(1+16\left(\left(\frac{10^{-4}\,\mathrm{Hz}}{f}\right)^2+\left(\frac{2\times10^{-5}\,\mathrm{Hz}}{f}\right)^{10}\right)\right)\ \mathrm{m^2\,s^{-4}\,Hz^{-1}} \tag{2.2}$$

is the colored acceleration noise. The acceleration noise model is based on a fit to the measured LISA Pathfinder performance.

The new LISA design balances technical and budgetary constraints against a set of science requirements. A key component in this iterative design process is an assessment of the science reach of the instrument. The first step in such an assessment is to consider compact galactic binaries as these are expected to be the most numerous of the sources; so numerous in fact that the unresolved component of the galactic population produces an effectively stochastic "confusion noise". Here we provide estimates for the confusion noise level as it evolves over the lifetime of the mission, as well as estimates for the number of resolvable galactic binaries and how well key parameters, such as the orbital periods and sky location, can be determined. Our work updates the confusion noise estimates of Nissanke *et al* [83].

2.2 Galactic Confusion Noise

The starting point for our analysis is a model for the population of compact binaries in our galaxy. We use an updated version [84] of the Nelemans *et al* [85] model provided to us by Valeriya Korol and Gijs Nelemans. The space density of

interacting white dwarf binaries is reduced by a factor of ten relative to earlier models in response to the findings of recent observational studies [86,87]. The model predicts there are $\sim$ 26 million galactic binaries emitting in the frequency band 0.1-10 mHz. The confusion noise is estimated using the iterative subtraction procedure described in Timpano *et al* [51] using a fast algorithm to generate the signals discussed in Cornish and Littenberg [50]. The first step is to generate a realization of the instrument noise in each data channel, then co-add the signals from the population of galactic binaries. A smooth fit to the the power spectral density of the signal plus noise is found using a variant of the BayesLine algorithm [88]. This fit is used as an initial estimate of the effective noise spectral density, against which the signal-to-noise ratio (SNR) of each binary is computed. All binaries with SNR > 7 are then subtracted from the data, and the estimate for the power spectral density of the remaining signal plus noise is computed. The SNRs of the remaining systems are re-computed using the updated (lowered) noise model, and those above threshold are subtracted. The subtraction procedure is iterated several times until few additional systems are flagged as loud. The end product is an estimate for the confusion noise and a list of resolvable sources. The mission duration has a significant impact on the level of the confusion noise. The longer the mission continues the more sources are resolved and the lower the confusion noise.

Shown in the first panel of Figure 2 are sensitivity curves that include the galactic confusion noise after 6 months, 1 year, 2 years and 4 years of data collection. The second panel in Figure 2 focuses on the confusion noise contribution, and compares our analytic fits to the results of the simulation. The analytic fits to the confusion noise have the functional form

$$S_c(f) = A\, f^{-7/3}\, e^{-f^{\alpha} - \beta f \sin(\kappa f)}\, [1 + \tanh(\gamma(f_k - f))]\ \mathrm{Hz}^{-1} \tag{2.3}$$

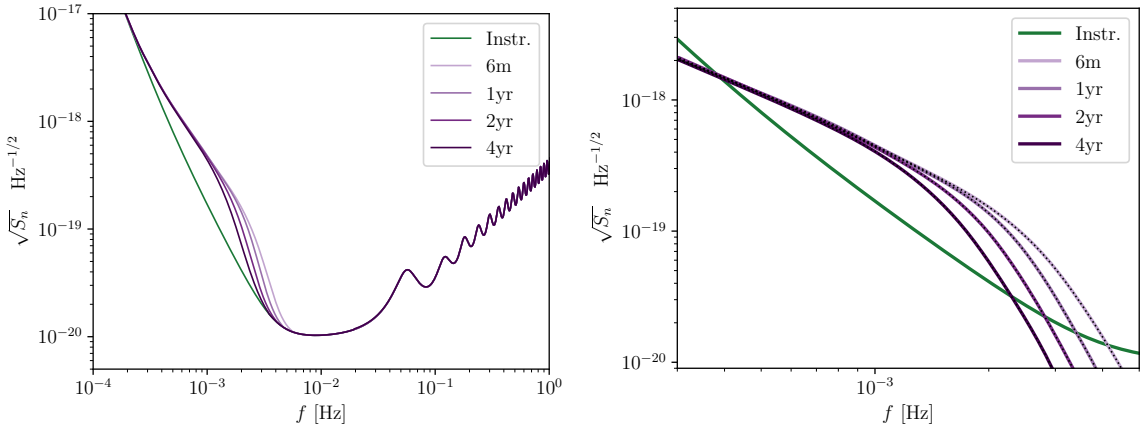

Figure 2.2: The panel on the left shows the sensitivity curve for the new LISA design including the estimated confusion noise level for a range of mission durations. The panel on the right focuses on the confusion noise impact on the sensitivity, and compares our analytic fit (black dashed lines) to the output of the simulation.

The overall $f^{-7/3}$ slope is the theoretical prediction for the power spectrum of a population of quasi-circular binaries evolving due to gravitational wave emission. The other multiplicative factors were chosen to match the departures from the power law spectrum. The form of the fit was originally suggested by Stas Babak. The parameters for the fits are given in Table 1. The overall amplitude and slope of there fit remain roughly constant with increasing observing time. The main change is in the knee frequency f_k, which gets steadily smaller with time.

2.3 Galactic Binary Science

While the unresolved galactic binaries are a source of noise that negatively impacts the science that can be done with other sources, the resolvable galactic binaries tell a rich story about stellar evolution and the distribution of stars in the galaxy. The distribution of resolved signals as a function of orbital period can be used to constrain population synthesis models. The orbital period is very well measured

	6 mo	1 yr	2 yr	4 yr
α	0.133	0.171	0.165	0.138
β	243	292	299	-221
κ	482	1020	611	521
γ	917	1680	1340	1680
f_k	0.00258	0.00215	0.00173	0.00113

Table 2.1: Parameters of the analytic fit the Galactic confusion noise as described by equation (1). The amplitude A has been fixed to $1.80e-44$. Note that the knee frequency f_k decreases with observation time as expected from Figure 2. Similarly, γ increase with observation time to leading to steeper drop off in confusion noise with increasing frequency.

even for systems at the detection threshold. Figure 3 shows the number density of resolved signals per $1/T_{\rm obs}$ frequency bin for a 4 year mission using both Michelson-like data channels.

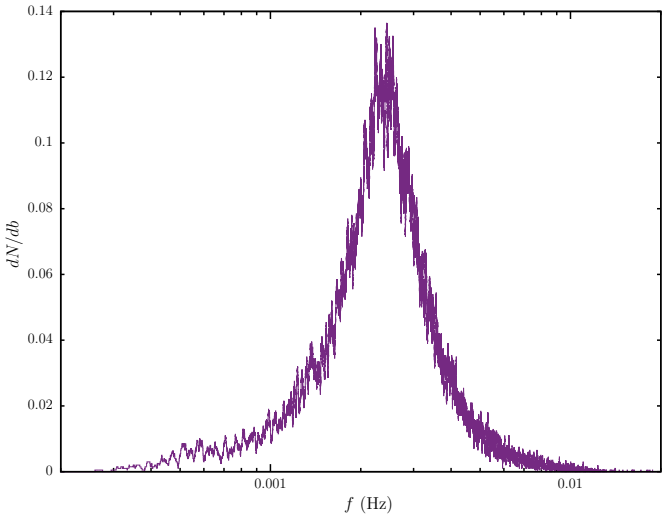

Figure 2.3: The number density of resolved signals per frequency bin for a 4 year mission using both Michelson-like data channels. Note that the density of resolved signals is largest between 2 - 3 mHz, which is where the Galactic confusion noise changes the most over time as shown in Figure 2.

CHAPTER THREE

IMPACT OF GALACTIC FOREGROUND CHARACTERIZATION ON A GLOBAL ANALYSIS FOR THE LISA GRAVITATIONAL WAVE OBSERVATORY

Contributions of Authors and Co-Authors

Manuscript in Chapter 3

Author: Travis Robson

Contributions: Performed the analytic calculations. Modified and further developed code provided by Neil Cornish. Wrote first draft of manuscript.

Co-Author: Dr. Neil Cornish

Contributions: Conceived the study design. Edited the manuscript. Wrote MCMC code for Bayesian interpretation of the results.

Manuscript Information Page

Travis Robson, Neil Cornish

Classical and Quantum Gravity

Status of Manuscript:

Published in a peer-reviewed journal

Published by IOP Publishing

Published November, 2017, Class. Quantum Grav. **34** 244002

Abstract

The Laser Interferometer Space Antenna (LISA) will explore the source-rich milli-Hertz band of the gravitational wave spectrum. In contrast to ground based detectors, where typical signals are short-lived and discrete, LISA signals are typically long-lived and over-lapping, thus requiring a global data analysis solution that is very different to the source-by-source analysis that has been developed for ground based gravitational wave astronomy. Across the LISA band, gravitational waves are both signals *and* noise. The dominant contribution to this so-called confusion noise (better termed unresolved signal noise) is expected to come from short period galactic white dwarf binaries, but all sources, including massive black hole binaries and extreme mass ratio captures will also contribute. Previous estimates for the galactic confusion noise have assumed perfect signal subtraction. Here we provide analytic estimates for the signal subtraction residuals and the impact they have on parameter estimation while for the first time incorporating the effects of noise modeling. The analytic estimates are found using a maximum likelihood approximation to the full global Bayesian analysis. We find that while the confusion noise is *lowered* in the global analysis, the waveform errors for individual sources are *increased* relative to estimates for isolated signals. We provide estimates for how parameter estimation errors are inflated from various parts of a global analysis.

3.1 Introduction

The recent discovery of gravitational waves [1] and the outstanding success of the LISA pathfinder mission [21] have given new life to the LISA mission. Building on decades of study, an updated LISA mission concept [80] was recently submitted to address the European Space Agency's "Gravitational Universe" science theme with

a launch scheduled for the early 2030's. The plan is to fly three identical spacecraft connected by six laser links forming a triangular detector with 2.5 million km long arms.

It has long been recognized that the LISA mission will suffer from "an embarrassment of riches", delivering data sets so packed with signals that extracting information about individual sources will require the development of unique data analysis techniques. Significant attention was given to this problem through the 2000's, culminating in a series of Mock LISA Data Challenges (MLDCs) [89–94] that produced some promising proof-of-principle solutions. The demise of the original LISA project in 2010 halted this effort, but work is now resuming following the re-birth of the mission. In addition to finding an implementable solution to the global analysis problem, there is also interest in producing reliable estimates for the science that can be achieved, including the number of sources of each type that can be resolved, and how well they can be characterized. A key input to these studies are estimates for the confusion noise from unresolved sources, as this adds to the instrument noise, and reduces the signal-to-noise ratio (SNR) of the resolved systems. These confusion noise estimates use variants of the idealized iterative source subtraction scheme introduced by Timpano *et al* [51]. We recently applied this technique to produce confusion noise estimates [70] that were used in the design study for the new LISA mission concept [80]. There are, however, several deficiencies with the simple confusion noise estimates: they assume that the confusion noise is stationary when in fact it oscillates with a 6 month period; they neglect the parameter estimation errors for the subtracted signals and the waveform residuals; and they also neglect the impact the removal has on other resolvable signals such as massive black holes. Here we generalize the Timpano *et al* [51] method to account for the waveform residuals and the impact they have on the detection of other signals. Some of our results were derived previously

by Cutler and Harms [95] in studies of foreground subtraction for the Big Bang Observer, but many results are new, including analytic estimates for power spectrum of the waveform residuals, incorporating the process of noise modeling, and the impact on parameter estimation for other sources. We find that the parameter estimation errors caused by other resolved signals are typically small compared to those due to instrument noise and unresolved signals. The exception to this rule is when two signals have very high overlap, such as sometimes occurs for galactic binaries with near identical orbital periods and sky locations [96].

Electromagnetic observations have identified ~ 50 galactic binaries with orbital periods that put their predicted gravitational signals in the LISA band [80]. Those that rise above the noise are refer to as "verification binaries". Population synthesis models predict that there are far more detectable systems waiting to be discovered, though the estimates have been lowered in the past decade as on-going surveys have been used to re-calibrate the models [14,86,87]. It is estimated that there are hundreds of millions of galactic binaries GBs emitting gravitational waves in our galaxy. In the mid-band of the LISA sensitivity, between $\sim 0.5 - 3$ mHz, gravitational waves from these systems are expected to dominate over instrument noise, with the unresolved component producing what is termed "confusion noise".

There have been several previous attempts at estimating the galactic confusion noise [14, 51, 70, 97, 98]. To characterize the confusion noise one must first determine how many galactic binaries are resolvable, but in order to figure out which binaries are resolvable, one must already have an estimate for the noise. The ideal solution is to perform a global fit, *e.g.* a full Bayesian analysis that fits both resolvable sources and noise at the same time as done by Littenberg [41]. Unfortunately, this procedure is extremely computationally intensive, and more efficient techniques are needed if we want to consider a range of population models and detector configurations for design

studies. To this end, Timpano *et al* [51] developed an iterative subtraction scheme which starts with a simulated data set comprised of an instrument noise realization and the superposition of all gravitational waves produced by synthetic population of galactic binaries. The signal-to-noise ratio (SNR) of the GBs is calculated, and those above a specified threshold SNR are subtracted perfectly *i.e.* the true waveform is removed from the data stream using the simulated signal parameters. The noise estimate is updated after the bright signals are removed, and the SNRs of the remaining sources are re-computed. Those above the detection threshold are removed, and the whole process is repeated. It typically takes 5-6 iterations for the solution to converge.

It is this assumption of perfect signal recovery we wish to address in this paper. In reality the instrument plus confusion noise realization will randomly perturb the estimated parameters for the resolvable systems, resulting in an inaccurate signal recovery. Here we use the Maximum-Likelihood approximation and Fisher information matrix to estimate the parameter errors and waveform residuals.

The outline of this paper is as follows: In Section 3.2 we briefly review the galactic population model used to produce realizations of the LISA data used in the analysis. Next, in Section 3.3 we provide a review of relevant Maximum-Likelihood methods and how they can be used to estimation the noise-induced errors in signal recovery and parameter estimation of the resolved GBs. In Section 3.4 we extend the usual ML analysis to include noise spectral estimation, and in Section 3.5 we illustrate the relevance of the ML to a full Bayesian analysis using a simple model of a sinusoid in stationary Gaussian noise. In Section 3.6 we apply the Maximum-Likelihood approach to the global fitting of multiple signals and drive expressions for how the interaction between the signals impacts waveform and parameter estimation errors. We conclude in Section 3.7 by computing an improved estimate for the galactic confusion noise

that takes into account parameter estimation errors in the bright source removal.

3.2 Instrument and Galactic Population Models

Our galaxy simulations use realizations of the the Toonen *et al* [84] population model provided by Valeriya Korol and Gijs Nelemans. The space density of interacting white dwarf binaries is reduced by a factor of ten relative to earlier models in response to the findings of recent observational studies [86,87]. The population has ~ 26 million systems with gravitational wave frequencies above 0.1 mHz. The signals from these systems are simulated using an improved version of the fast waveform generation algorithm of Cornish and Littenberg [50]. The improved algorithm removes the need to sum over terms in the Fourier convolution by referencing the carrier frequency f_0 to the nearest integer multiple of the sample frequency, such that $f_0 = m/T_{\rm obs} + \delta f$, and absorbing the factor of $e^{2\pi i \delta f t}$ into the slowly varying part of the signal. This removes the need for the sum in equation (A24) of Ref. [50], and significantly speeds up the waveform generation. In our analysis we use the full set of first-generation time-delay interferometry (TDI) [99] variables $\tilde{X}(f), \tilde{Y}(f), \tilde{Z}(f)$ given in Ref. [50], but when displaying results we show the more familiar Michelson-equivalent signals. To obtain the Michelson equivalent sensitivity we make use of the relation $S_X = 4 \sin^2 (f/f_*) S_M$ where S_X is the noise as seen in the TDI X data channel and S_M is the equivalent Michelson noise and $f_* = c/(2\pi L)$ is the transfer frequency. Our instrument noise model assumes white position noise $S_p(f)$ and colored acceleration noise $S_a(f)$ with spectral densities

$$\begin{aligned} S_p(f) &= 8.9 \times 10^{-23} \mathrm{m}^2 \mathrm{Hz}^{-1} \\ S_a(f) &= 9 \times 10^{-30} \left[1 + 16 \left(\left(\frac{10^{-4}\mathrm{Hz}}{f} \right)^2 + \left(\frac{2 \times 10^{-5}\mathrm{Hz}}{f} \right)^{10} \right) \right] \mathrm{m}^2 \mathrm{s}^{-4} \mathrm{Hz}^{-1} \end{aligned} \quad (3.1)$$

Under the assumption that the noise levels are the same in each link, we can form the noise-orthogonal $\{A, E, T\}$ channels [69]

$$\begin{aligned} A &= \tfrac{1}{3}\,(2X - Y - Z) \\ E &= \tfrac{1}{\sqrt{3}}(Z - Y) \\ T &= \tfrac{1}{3}\,(X + Y + Z)\ . \end{aligned} \tag{3.2}$$

Below the transfer frequency f_*, where most signals are found, the T channel is far less sensitive to gravitational waves, and does not contribute to our analysis.

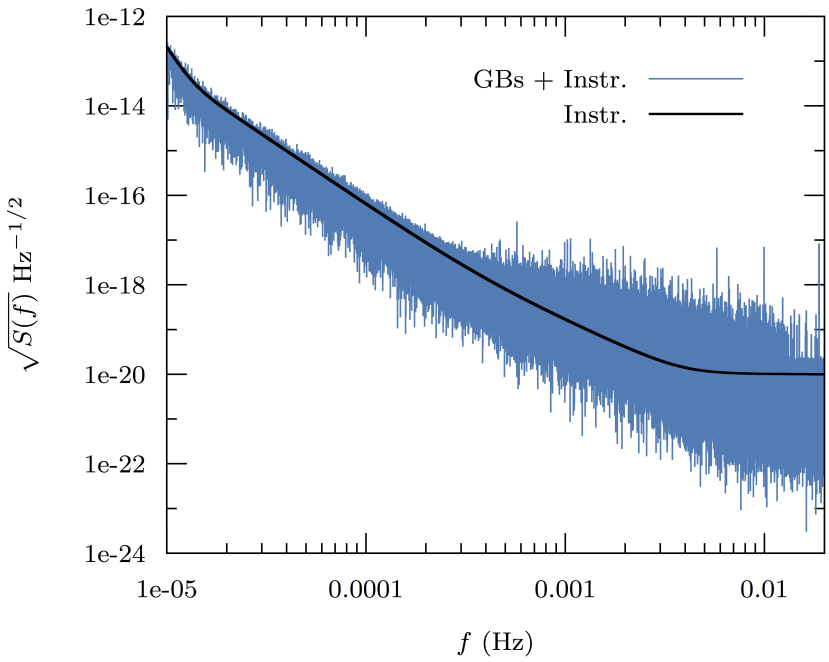

Figure 3.1: A realization of the Michelson-equivalent strain spectral density in the X channel, comprised of instrument noise and signals from our galactic binary population using a four year observation period. Note that signal from the galactic binaries dominates the instrument noise across the band $0.3 - 20$ mHz.

Figure 3.1 shows a realization of instrument noise (assumed to be Gaussian and stationary), combined with the signal from all of the relevant GBs in the population

synthesis realization plotted as the Michelson-equivalent sensitivity. We see that the galactic foreground rises above the instrument noise across the frequency range $0.3 - 20$ mHz.

3.3 Parameter Estimation and Waveform Errors

In the high SNR regime, the likelihood is strongly peaked about the true model parameters, which allows for a Maximum-Likelihood analysis. Many useful results can be derived from a Taylor expansion of the likelihood about the true parameters. Here we provide a brief review of the Maximum-Likelihood (ML) approximation, deriving results for the parameter estimation errors and waveform residuals. We follow with a discussion of how the ML analysis relates to a global Bayesian analysis.

3.3.1 Maximum-Likelihood Review

Consider the simple case of data $\mathbf{s}$ comprised of a gravitational wave signal $\mathbf{h}_T = \mathbf{h}(\vec{\lambda}_T)$ and stationary, Gaussian noise $\mathbf{n}$. The likelihood of observing $\mathbf{s}$ given the presence of a gravitational wave signal $\mathbf{h}_T$ is then

$$p(\mathbf{s}|\mathbf{h}) = e^{-G/2} e^{-(\mathbf{s}-\mathbf{h}(\vec{\lambda})|\mathbf{s}-\mathbf{h}(\vec{\lambda}))/2} = e^{-G/2} e^{-\chi^2/2} \ , \tag{3.3}$$

where

$$(\mathbf{g}|\mathbf{k}) \coloneqq 2 \sum_{I=\{A,E\}} \int_0^\infty \frac{\tilde{g}_I(f)\tilde{k}_I^*(f) + \tilde{g}_I^*(f)\tilde{k}_I(f)}{S_{n,I}(f)} df \ , \tag{3.4}$$

defines the noise-weighted inner product taken across all independent data channels, and the one-sided noise spectral density in channel I is given by the expectation value

$$\mathbb{E}[\tilde{n}_I(f)\tilde{n}_J^*(f')] = \frac{1}{2}\delta(f - f')S_{n,I}(f)\delta_{IJ} \ . \tag{3.5}$$

The noise $S_{n,I}(f)$ will include instrument noise and unresolved gravitational wave signals. The normalization factor G is given by

$$G = \sum_{I=\{A,E\}} \int_0^\infty T \log\left[\pi T S_{n,I}(f)\right] df \,. \tag{3.6}$$

The traditional derivation of the maximum likelihood solution assumes that the noise model $S_{n,I}(f)$ is known, and that the signal model, $\mathbf{h}(\vec{\lambda})$ is close to the true signal $\mathbf{h}_T = \mathbf{h}(\vec{\lambda}_T)$. The signal model is then Taylor expanded about the true parameters:

$$\mathbf{h}(\vec{\lambda}) = \mathbf{h}_T + \partial_i \mathbf{h}_T \Delta\lambda^i + \mathcal{O}\left(\Delta\lambda^2\right), \tag{3.7}$$

where $\Delta\vec{\lambda} = \vec{\lambda} - \vec{\lambda}_T$. The chi-squared in the likelihood can then be expanded as

$$\begin{aligned} \chi^2 &= (\mathbf{s}-\mathbf{h}|\mathbf{s}-\mathbf{h}) \\ &= (\mathbf{n}|\mathbf{n}) - 2(\mathbf{n}|\partial_i \mathbf{h}_T)\Delta\lambda^i + (\partial_i \mathbf{h}_T|\partial_j \mathbf{h}_T)\Delta\lambda^i\Delta\lambda^j + \mathcal{O}(\Delta\lambda^3)\,. \end{aligned} \tag{3.8}$$

The maximum likelihood solution is found by setting $\partial_i \chi^2 = 0$, which yields

$$\Delta\lambda^j = (\mathbf{n}|\partial_i \mathbf{h}_T)\left(\Gamma^{-1}\right)^{ij} + \ldots \tag{3.9}$$

where

$$\Gamma_{ij} = (\partial_i \mathbf{h}_T|\partial_j \mathbf{h}_T) \tag{3.10}$$

is the Fisher information matrix. Using the identity $\mathbb{E}[(\mathbf{n}|\mathbf{g})(\mathbf{n}|\mathbf{k})] = (\mathbf{g}|\mathbf{k})$ we find that the error covariance matrix is given to leading order in the signal-to-noise ratio by the inverse of the Fisher information matrix:

$$C^{ij} = \mathbb{E}[\Delta\lambda^i \Delta\lambda^j] = \left(\Gamma^{-1}\right)^{ij} + \mathcal{O}\left(\mathrm{SNR}\right)^{-1} \,, \tag{3.11}$$

where the SNR ρ is given by $\rho^2(\mathbf{h}) = (\mathbf{h}|\mathbf{h})$. See Vallisneri [100] for a more in depth presentation that discusses some of the potential pitfalls in using the Fisher Information matrix approximation to perform parameter error estimation. It is important to note that there are higher order corrections to the signal parameters and covariance matrix which appear in Cutler and Flanagan [101] as equations (A31) and (A35).

3.3.2 Signal Residuals

We can use the maximum likelihood approximation to study noise induced errors in the parameter recovery and signal subtraction for galactic binaries. A closely related analysis was performed by Cutler and Harms [95] in the context of subtracting the signals from neutron stars to allow for the detection of a primordial stochastic background for the Big Bang Observer mission concept. We extend their analysis to include noise modeling, and derive new expressions for the impact the foreground removal has on parameter estimation for other sources such as massive black hole mergers and extreme mass ratio insprials (EMRIs).

The noise-induced parameter estimation errors (3.9) result in waveform errors

$$\Delta\mathbf{h} = \mathbf{h}_T - \mathbf{h} \simeq -\partial_i \mathbf{h}_T \Delta\lambda^i + \dots \tag{3.12}$$

An example of the observed signal from a galactic binary and the noise-induced subtraction residual is shown in Figure 3.2. Note that the residual is below the reference noise level as the waveform error is down-weighted by the level of overlap between noise and parameter derivatives of the signal.

The waveform error has zero mean, $\mathbb{E}[\Delta\mathbf{h}] = 0$, and variance

$$\mathbb{E}[\rho^2_{\Delta h}] \coloneqq \mathbb{E}[(\Delta\mathbf{h}|\Delta\mathbf{h})] = (\partial_i \mathbf{h}_T|\partial_j \mathbf{h}_T)\mathbb{E}[\Delta\lambda^i \Delta\lambda^j] \approx \Gamma_{ij}\left(\Gamma^{-1}\right)^{ij} = D\,. \tag{3.13}$$

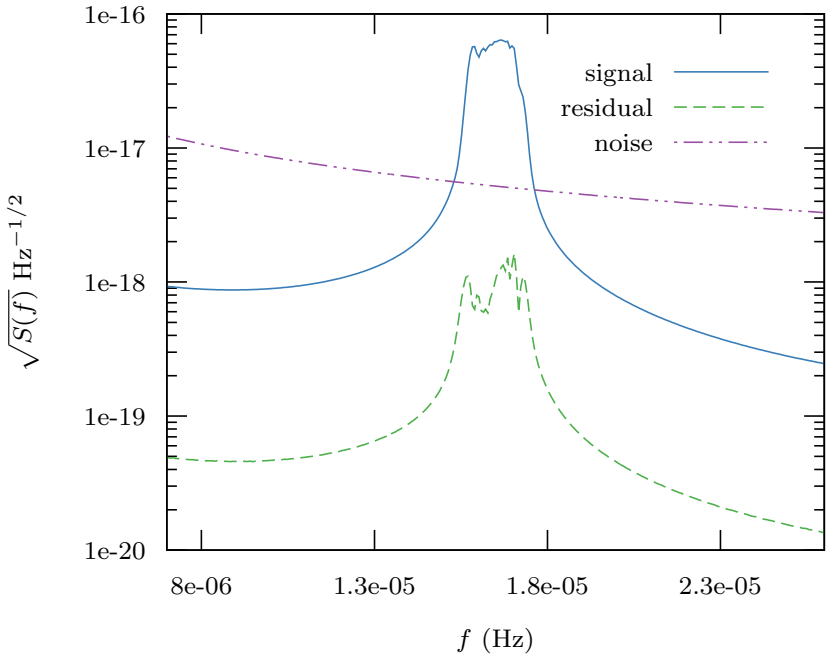

Figure 3.2: This figure presents an example of the Michelson-equivalent strain spectral density for a signal (solid, blue), the corresponding waveform error (dashed, green), and the noise level (dot-dash, purple). The SNR of the signal was ~ 110 and the residual SNR was ~ 5. This signal results from one of the sources identified as bright after a 6 month observation run. Note the broadening that results from Doppler shift and frequency evolution.

In the final step we have assumed that the error covariance matrix is approximated by the inverse of the Fisher matrix. The SNR of the residual depends only on the parameter dimension in the signal model and not upon the strength of the signal. Each term in the sum for Δh is random walk induced by the noise realization, i.e. $\Delta h \sim n\sqrt{D}$ as there are D terms. This means that $|\Delta\tilde{h}|^2 \sim (n^* n)D$ and the inner product is weighted by the RMS noise resulting in a dependence only on model dimension. It can be shown that the variance of $\rho^2_{\Delta h}$ is $2D$ and the skew is $1/D^2$.

The expectation value of the chi-squared can be written as

$$\begin{aligned}\mathbb{E}[\chi^2] &= \mathbb{E}[(\mathbf{n}+\Delta\mathbf{h}|\mathbf{n}+\Delta\mathbf{h})] \\ &= N + \mathbb{E}[\rho^2_{\Delta h}] + 2\mathbb{E}\left[(\Delta\mathbf{h}|\mathbf{n})\right] = N - D\,, \end{aligned} \tag{3.14}$$

where the last step follows from

$$\mathbb{E}[(\Delta\mathbf{h}|\mathbf{n})] = -\mathbb{E}[(\mathbf{h}_{,i}|\mathbf{n})(\Gamma^{-1})^{ij}(\mathbf{n}|\mathbf{h}_{,j})] = -D\ , \tag{3.15}$$

and N is the number of data samples. We see that the signal residuals are anti-correlated with the noise, which results in a reduction in the chi-squared. Part of the noise gets absorbed by the signal model, which will ultimately result in a lowering of the confusion noise estimate relative to that found assuming perfect signal subtraction. Note that power spectrum of the residual $\mathbf{s}-\mathbf{h}$ has expectation value

$$\begin{aligned}\mathbb{E}[(\mathbf{s}(f)-\mathbf{h}(f))(\mathbf{s}(f')-\mathbf{h}(f'))^*] &= \frac{1}{2}S_{n,0}(f)\delta(f-f') \\ &\quad -\partial_i\tilde{\mathbf{h}}_T(f)\partial_j\tilde{\mathbf{h}}^*_T(f')\left(\Gamma_0^{-1}\right)^{ij}\delta(f-f') \\ &= \frac{1}{2}\left(S_{n,0}(f) - S_{\Delta\mathbf{h}}(f)\right)\delta(f-f')\,, \end{aligned} \tag{3.16}$$

where

$$\frac{1}{2}\delta(f-f')S_{\Delta\mathbf{h}}(f) = \partial_i\tilde{\mathbf{h}}_T(f)\partial_j\tilde{\mathbf{h}}^*_T(f')\left(\Gamma_0^{-1}\right)^{ij}\delta(f-f') \tag{3.17}$$

$$= \mathbb{E}[\Delta\tilde{\mathbf{h}}(f)\Delta\tilde{\mathbf{h}}^*(f')] \tag{3.18}$$

is the power spectral density of $\Delta\mathbf{h}$. Note that we made use of $\mathbb{E}\left[\tilde{\mathbf{n}}(f)(\mathbf{n}|\partial_i\mathbf{h}_T)\right] = \partial_i\tilde{\mathbf{h}}_T(f)$.

Our waveform model for a galactic binary has $D = 9$ parameters: the sky

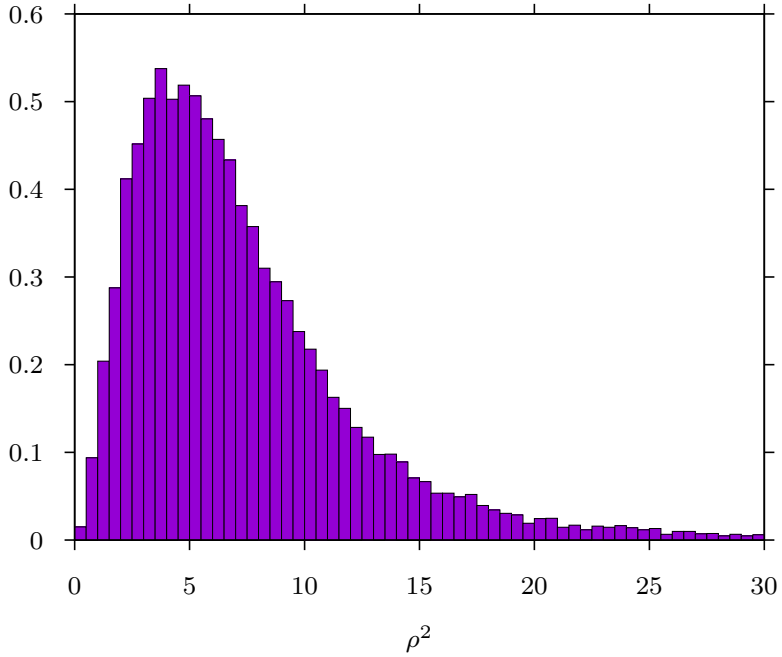

Figure 3.3: Histogram of the SNRs of the waveform errors resulting for the iterative subtraction scheme. The average value was 7.6.

location (θ, ϕ); inclination and polarization angles (ι, ψ); amplitude A; reference phase ϕ_0 and reference gravitational wave frequency f_0; and first and second frequency derivatives $\dot{f}_0$ and $\ddot{f}_0$. For most galactic binaries the frequency derivatives are poorly constrained and the effective model dimension is closer to $D = 7$. The relevant quantity for estimating which systems have detectable frequency evolution are the number of frequency bins of evolution, $\alpha = \dot{f}_0 T_{\rm obs}^2$ and $\beta = \ddot{f}_0 T_{\rm obs}^3$, and the SNR. Roughly speaking, frequency evolution through $\sim (7/\mathrm{SNR})$ bins is detectable [102], and to leading post-Newtonian order we have

$$\begin{aligned} \alpha &= 1.5 \left(\frac{f_0}{4\,\mathrm{mHz}}\right)^{11/3} \left(\frac{\mathcal{M}}{0.25 M_\odot}\right)^{5/3} \left(\frac{T_{\rm obs}}{4\,\mathrm{yrs}}\right)^2 \\ \beta &= 1.8 \left(\frac{f_0}{25\,\mathrm{mHz}}\right)^{19/3} \left(\frac{\mathcal{M}}{0.25 M_\odot}\right)^{10/3} \left(\frac{T_{\rm obs}}{4\,\mathrm{yrs}}\right)^3 . \end{aligned} \tag{3.19}$$

The chirp mass $\mathcal{M}$ has been scaled to the mode of the population distribution [103]. From these expressions we see that only the loudest, most massive and highest-frequency systems will have a measurable second frequency derivative, and that most systems below 3 mHz will show no measurable frequency evolution at all. Including poorly constrained parameters in the model can lead to ill-conditioned Fisher matrices with badly behaved inverses. One solution is to reduce the model dimension by eliminating parameter λ_i whenever the inner product $(\partial_i \mathbf{h}_T | \partial_i \mathbf{h}_T)$ drops below some threshold. An alternative solution is to replace the Fisher matrix $\mathbf{\Gamma}$ with a matrix formed from the augmented Fisher matrix, $\mathbf{K}$, which includes derivatives of the priors (see Section 3.5 for details). We adopt the latter approach and include Gaussian priors on α, β centered on zero with width $\sigma = 10$. The priors condition the matrix and when the parameters are poorly constrained by the data, have the effect of reducing the model dimension. Using the more stable approximation $\mathbb{E}[\Delta\lambda^i \Delta\lambda^j] \approx (K^{-1})^{ij}$, yields $\mathbb{E}[\rho^2_{\Delta h}] \approx D_{\rm eff}$, where $D_{\rm eff}$ is the effective dimension of the model, defined by the number of parameters that have posterior distributions that differ measurably from their priors (a notion that can be made precise using the Kullback-Leibler divergence).

Figure 3.3 shows a histogram of square SNRs of the waveform residuals for the galactic binaries that are deemed detectable by the iterative subtraction scheme discussed in Section 3.7. The average value for there residual SNR^2 of 7.6 is less than the full model dimension $D = 9$, and consistent with our estimate for the effective dimension.

3.4 Maximum-Likelihood Approximation with Noise Estimation

The standard treatment of the maximum likelihood expansion assumes that the noise spectrum is known. If the detectable gravitational wave signals are infrequent and short-lived, as is currently the case for compact binary mergers in LIGO, then

"off-source" data from times where no detectable signals are present can be used to estimate the noise spectrum. These estimates will include instrument noise and unresolved gravitational wave signals. The option of making off-source estimates will not be available for LISA, and the noise spectrum will have to be inferred along with the signal model. Our derivation will assume that we have a parameterized model for the noise, such as the cubic spline model used by the BayesLine algorithm [104].

To get an understanding for how noise modeling impacts the maximum likelihood calculation, consider a simple example with zero mean, additive, white Gaussian noise and N data samples with likelihood

$$p(\mathbf{s}|\mathbf{h}) = \prod_{k=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(s_k-h_k)^2/(2\sigma^2)}\,. \tag{3.20}$$

The un-perturbed ($\mathbf{h} = 0$) noise level is given by sample variance

$$\sigma_0^2 = \frac{1}{N}\sum_{k=1}^{N} n_k^2\,. \tag{3.21}$$

We could expand σ^2 about the theoretical variance, σ_*^2, but it is simpler to expand σ^2 about the sample variance: $\sigma^2 = \sigma_0^2 + \Delta\kappa$ so that $\Delta\kappa = 0$ when $\mathbf{h} = 0$. The typical difference between the sample variance and the theoretical variance will be by an amount that scales as the standard deviation of the sample variance, $\Delta\sigma_0^2 = \sqrt{2}\sigma_*^2/\sqrt{N}$. The log likelihood can be expanded:

$$\begin{aligned}\log p(\mathbf{s}|\mathbf{h}) &= \text{const} + \left(1 - \frac{\Delta\kappa}{\sigma_0^2}\right)\left((\mathbf{n}|\partial_i\mathbf{h}_T)_0\Delta\lambda^i - \frac{1}{2}(\partial_i\mathbf{h}_T|\partial_j\mathbf{h}_T)_0\Delta\lambda^i\Delta\lambda^j\right) \\ &\quad -\frac{N}{2}\left(1 + \frac{\Delta\kappa^2}{2\sigma_0^4}\right) + \dots\,,\end{aligned} \tag{3.22}$$

where the notation $(a|b)_0$ denotes that the noise weighted inner product is taken

with respect to the un-perturbed noise level σ_0^2. Setting $\partial_{\Delta\kappa} \log p(\mathbf{s}|\mathbf{h}) = 0$ and $\partial_{\Delta\lambda^k} \log p(\mathbf{s}|\mathbf{h}) = 0$, yields the ML solution

$$\begin{aligned} \Delta\lambda^j &= (\mathbf{n}|\partial_i \mathbf{h}_T)_0 \left(\Gamma_0^{-1}\right)^{ij} \\ \Delta\kappa &= -\frac{\sigma_0^2}{N}(\mathbf{n}|\partial_i \mathbf{h}_T)_0 (\mathbf{n}|\partial_j \mathbf{h}_T)_0 \left(\Gamma_0^{-1}\right)^{ij} . \end{aligned} \tag{3.23}$$

We see that the leading order ML solution for the signal parameters is unchanged from the fixed noise case. The updated noise estimate $\sigma^2_{\rm ML} = \sigma_0^2 + \Delta\kappa$ is *lowered* relative to its true value, as can be seen by taking the expectation value

$$\mathbb{E}[\sigma^2] = \sigma_0^2 \left(1 - \frac{D}{N}\right) . \tag{3.24}$$

While the ML waveform removes some of the noise, this is now accounted for in the ML estimate for the noise, such that the expected value of the chi-squared is again just the dimension of the data: $\mathbb{E}[\chi^2] = N$.

From expanding the likelihood around the signal and noise parameters $\vec{\eta} = \{\vec{\lambda}, \kappa\}$ and maximizing the likelihood obtains the form

$$p(\mathbf{s}|\Delta\vec{\eta}) = \frac{1}{\sqrt{2\pi \ \det \mathbf{\Gamma}^{-1}}} e^{-\frac{1}{2}\Gamma_{\mu\nu}\Delta\eta^\mu \Delta\eta^\nu} \tag{3.25}$$

where $\Gamma_{\mu\nu} = -\partial_\mu \partial_\nu \log p(\mathbf{s}|\mathbf{h})|_{\rm max}$. Note that we are using Greek indicies to denote the entire collection of parameters. One can read off the Fisher matrix from the log-likelihood:

$$\mathbf{\Gamma} = \begin{pmatrix} \Gamma_{0,ij} & (\mathbf{n}|\partial_j \mathbf{h}_T)_0/\sigma_0^2 \\ (\mathbf{n}|\partial_i \mathbf{h}_T)_0/\sigma_0^2 & 2\sigma_0^4/N \end{pmatrix} , \tag{3.26}$$

where $\Gamma_{0,ij}$ is the Fisher matrix obtained from the signal-only ML analysis discussed

in the previous section. The full Fisher matrix can be inverted by recognizing that the off-diagonal terms $(\mathbf{n}|\partial_j \mathbf{h}_T)_0/\sigma_0^2$ are small compared to the block diagonal terms. We find that the variances for the signal parameters are inflated:

$$\left(\Gamma^{-1}\right)^{ii} \approx \left(\Gamma_0^{-1}\right)^{ii} + \frac{2}{N}\left(\Gamma_0^{-1}\right)^{il}\left(\Gamma_0^{-1}\right)^{ki}(\mathbf{n}|\partial_l \mathbf{h}_T)_0(\mathbf{n}|\partial_k \mathbf{h}_T)_0\,. \tag{3.27}$$

On average $(\Gamma^{-1})^{ij} \rightarrow \left(\Gamma_0^{-1}\right)^{ij}(1+2/N)$ where in the limit of large N we obtain the original Fisher matrix. The signal model parameter variances are inflated by covariances with the noise model parameters as they both attempt to capture pieces of the signal. Note that covariance of the parameter shifts $\Delta\lambda^j$ and $\Delta\kappa$ from their true values, as computed in (3.23), does not equal the inverse of the Fisher matrix, $\mathbb{E}\left[\Delta\eta^\nu \Delta\eta^\mu\right] \neq (\Gamma^{-1})^{\mu\nu}$. This is because the noise modeling changes the shape of the likelihood, and not just the location of the peak. However, we see from (3.25) that Γ^{-1} does indeed describe the parameter covariances.

The ML expansion for a colored noise model is considerably more involved, and we relegate the details to Appendix A. To keep the notation simple we suppress the sum over data channels. Introducing the parameterized noise model

$$S_n(f;\vec{\theta}) = S_{n,0}(f) + \Delta\theta^a \partial_a S_n(f) + \frac{1}{2}\Delta\theta^a \Delta\theta^b \partial_a \partial_b S_n(f) + \dots\,, \tag{3.28}$$

(noise model derivatives are evaluated at the ML values after differentiation) where $S_{n,0}(f)$ is some smooth estimate of the instrument noise and unresolved signals (assuming perfect subtraction of the resolvable signals), we find the leading-order solution for the signal parameters has the same form as in (3.23), while the noise model parameters are expressed in their full form in appendix A. There exists factors $(TS_{n,0} - 2\tilde{n}^*\tilde{n})/S_{n,0}$ which accounts for the difference between the theoretical noise model and fluctuation seen in a particular noise realization, i.e. the difference between

σ_*^2 and σ_0^2 in the white noise toy model, as evidenced by its expectation value vanishing. Neglecting this difference and considering the expectation value of $\Delta\theta^j$ we obtain the simplification

$$\mathbb{E}\left[\Delta\theta^j\right] \approx \; -\left[\int T\frac{S_{n,i}S_{n,j}}{S_{n,0}^2}df\right]^{-1}(\partial_a\mathbf{h}_T|\partial_b\mathbf{h}_T)_i\left(\Gamma_0^{-1}\right)^{ab} \;, \tag{3.29}$$

where the notation $(x|y)_a$ defines the inner product

$$(x|y)_a = 4\int\frac{(\tilde{x}\tilde{y}^* + \tilde{x}^*\tilde{y})}{S_{n,0}}\frac{\partial_a S_n}{S_{n,0}}df\,. \tag{3.30}$$

The white-noise case (3.23) can be recovered by setting $S_{n,0} = \frac{2}{T}\sigma_0^2$ and $\partial_a S_n(f) = \frac{2}{T}$, so that $\left[\int T\frac{S_{n,i}S_{n,j}}{S_{n,0}^2}df\right]^{-1} = \sigma_0^4/N$, and $(x|y)_a = (x|y)_0/\sigma_0^2$.

We can now compute the expectation value of the noise perturbation $\Delta S_n(f) = \Delta\theta^a\partial_a S_n(f) + \ldots$

$$\mathbb{E}[\Delta S_n(f)] = -\left[\int T\frac{S_{n,a}S_{n,b}}{S_{n,0}^2}df\right]^{-1}\left(\Gamma_0^{-1}\right)^{ij}(\partial_i\mathbf{h}_T|\partial_j\mathbf{h}_T)_a\partial_b S_n(f)\,. \tag{3.31}$$

Note that the perturbation to the noise model is negative, as it must be given that the signal model absorbs some of the noise. One would expect that $\Delta S_n(f)$ should be a smoothed representation of $-S_{\Delta h}(f)$ for an effective noise model, mopping up errors introduced by the signal ML.

Similar to the white noise model above we may obtain the signal model variances for a general noise model by making the same appeals to neglecting differences between the theoretical and sample variance and averaging over many noise realizations

$$\left(\Gamma^{-1}\right)^{ii} \approx \left(\Gamma_0^{-1}\right)^{ii} + 2\left(\Gamma_0^{-1}\right)^{im}\left(\Gamma_0^{-1}\right)^{ni}\left(\partial_m \mathbf{h}_T | \partial_n \mathbf{h}_T\right)_{ab} \times \left(\int_0^\infty T \frac{S_{n,a} S_{n,b}}{S_{n,0}^2} df\right)^{-1} . \tag{3.32}$$

An effective noise model would minimize the factors $\left(\partial_m \mathbf{h}_T | \partial_n \mathbf{h}_T\right)_{ab}$ such that $\left(\Gamma^{-1}\right)^{ii} \to \left(\Gamma_0^{-1}\right)^{ii}$. We can turn this into a more useful expression by taking advantage of the compact (in the frequency domain) nature of the GB signal and assume the noise PSD is roughly constant

$$\left(\Gamma^{-1}\right)^{ii} \approx \left(\Gamma_0^{-1}\right)^{ii}\left(1 + \frac{2}{T\Delta f}\right) , \tag{3.33}$$

where Δf is the bandwidth of the signal such that $T\Delta f$ is the number of frequency bins the GB spans. For sources that span many frequency bins such that the noise PSD cannot be assumed to be constant this serves as an upper limit for the increase in the parameter errors. Note that other terms exist when considering covariances of the signal model. Again, we see that when the source occupies a large bandwidth we recover the variances for when the noise is known. For a 3 mHz source that experiences a $\Delta f \approx 0.6 \times 10^{-6}$ Hz Doppler shift spreading due to LISA's orbital motion, the parameter variances will be inflated by 10% for a one year observation span, dropping to 3% after four years.

3.5 Relating Bayesian Inference and Frequentist Maximum Likelihood Estimation

The LISA data will include overlapping signals from an unknown number of sources of different types. Bayesian inference provides a powerful and flexible

framework for inferring the number and properties of the resolvable sources. In addition to the ontological differences between the Bayesian and Frequentist approach to statistical inference - Bayesian inference considers the data to be known and the signal parameters to be uncertain while Frequentist inference considers the signal parameters to be fixed and the data to be uncertain - Bayesian inference typically integrates over uncertainty (marginalization), while Frequentist analysis employs maximization. Despite these differences, the maximum likelihood analysis we have described can be used to estimate results from Bayesian inference by way of a Taylor expansion of the posterior distribution $p(\vec{\lambda}|\mathbf{s}) = p(\mathbf{s}|\vec{\lambda})p(\vec{\lambda})/p(\mathbf{s})$. Expanding about the mode of the posterior distribution (also known as the maximum a posteriori probability (MAP) estimate, $\vec{\lambda}_{\rm MAP}$), we have the quadratic approximation

$$p(\vec{\lambda}|\mathbf{s}) \simeq (2\pi)^{-D/2}(\det\boldsymbol{K})^{1/2}e^{-\frac{1}{2}K_{ij}\Delta\lambda^i\Delta\lambda^j}\,, \tag{3.34}$$

where $\Delta\vec{\lambda} = \vec{\lambda} - \vec{\lambda}_{\rm MAP}$ and

$$K_{ij} = -\partial_i\partial_j \log\left(p(\mathbf{s}|\vec{\lambda})p(\vec{\lambda})\right)\Big|_{\rm MAP}\,. \tag{3.35}$$

When the likelihood is more informative than the prior, $\vec{\lambda}_{\rm MAP} \approx \vec{\lambda}_{\rm ML}$ and $\mathbf{K}$ is well approximated by the Fisher information matrix $\mathbf{\Gamma}$, though even small contribution from the derivatives of the log prior can have a an important stabilizing effect on $\mathbf{K}$.

To illustrate the relationship between the maximum likelihood analysis and Bayesian inference we produced simulated data consisting of stationary, Gaussian white noise with variance σ^2 and a sinusoidal signal $h(A, f_0, t_0, \phi_0) = A\cos(2\pi f_0(t - t_0) + \phi_0)$. We held the phase parameter $\phi_0 = \pi$ fixed in the analysis as otherwise there is a near perfect degeneracy between the time offset t_0 and the phase offset ϕ_0, which significantly complicates the analysis. The simulated data consisted of

$N = 10^4$ evenly spaced samples spanning $T = 100$ seconds, with $A = \sqrt{2}$, $f_0 = 0.25$ Hz, $t_0 = 1$ second and $\sigma^2 = 100$. The noise level was set to yield a signal-to-noise ratio of SNR $= 10$. A plot of the simulated data and signal are shown in the upper panel of Figure 3.4.

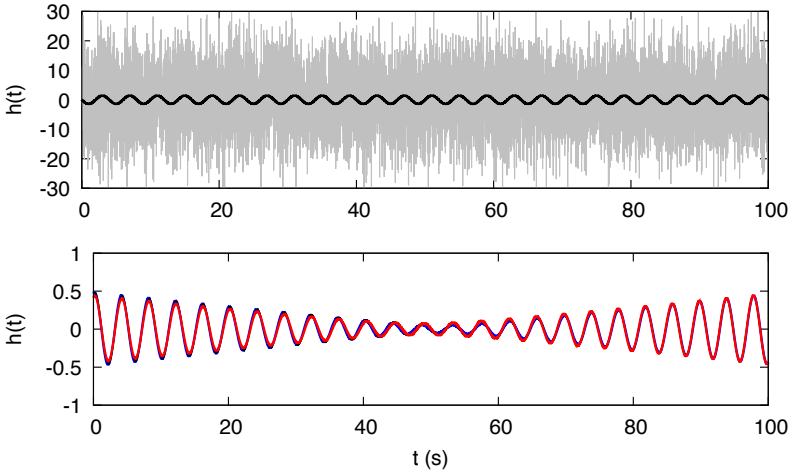

Figure 3.4: The upper panel shows simulated data (grey) and signal (black) for the simple sinusoid model. The lower panel shows the waveform residuals computed using the analytic maximum likelihood (red), and the numerical maximum a posteriori (blue) and mean (black) values from a MCMC analysis.

Table 3.1: Parameter Error Estimates

Parameter	σ_Γ	$\sigma_{\rm MCMC}$	$\Delta\lambda_{\rm ML}$	$\Delta\lambda_{\rm MAP}$	$\Delta\lambda_{\rm mean}$
A	0.141	0.135	-0.078	-0.048	-0.067
f_0	0.00055	0.00054	0.0010	0.0011	0.0010
t_0	0.125	0.123	0.195	0.214	0.217

For the Bayesian analysis we assumed uniform priors on the signal and noise model parameters (A, f_0, t_0, σ^2) across a range that was much wider than the predicted statistical errors so that posterior distribution and the likelihood were effectively

identical. A Markov Chain Monte Carlo (MCMC) simulation was used to compute the mean and variance of the signal parameters and the waveform error, while (3.23) and (3.25) were used to estimate the parameter shifts at maximum likelihood and the variances. The MCMC and ML derived values for the parameter shifts and standard deviations are listed in Table 3.1 for a particular noise realization. The marginalized posterior distributions for the parameters are compared to the predictions of the Gaussian approximation (3.34) in Figure 3.5. The agreement between the ML and MCMC seen in this example was typical of what we found when repeating the analysis for different noise realizations. That is not to say that the Gaussian approximation will be this accurate in more realistic settings where the noise is more complicated and the parameters are highly correlated [100], but it does provide useful order-of-magnitude estimates in most situations.

The displacement of the parameters from their true values is related to the waveform error Δh shown in the lower panel of Figure 3.4 which displays the ML, MAP and mean waveform errors. Note that the mean waveform error is found by averaging the waveform errors, and *not* by using the mean parameter values to compute a waveform. The particular shape of the residual, which appears to have an amplitude envelope that is narrow at the center and wide at the edges, is just an accident of the particular noise realization. Other noise realizations give different looking residuals, some of which peak at the center and are narrow at the edges of the time window.

The good agreement between the maximum likelihood and mean waveform residual hides a key difference between the frequentist and Bayesian analyses: in the Bayesian global fit the waveform uncertainties are marginalized over, while the frequentist analysis uses point estimates. Rather than subtracting a particular point estimate for each signal from the data, the Bayesian approach subtracts a range of

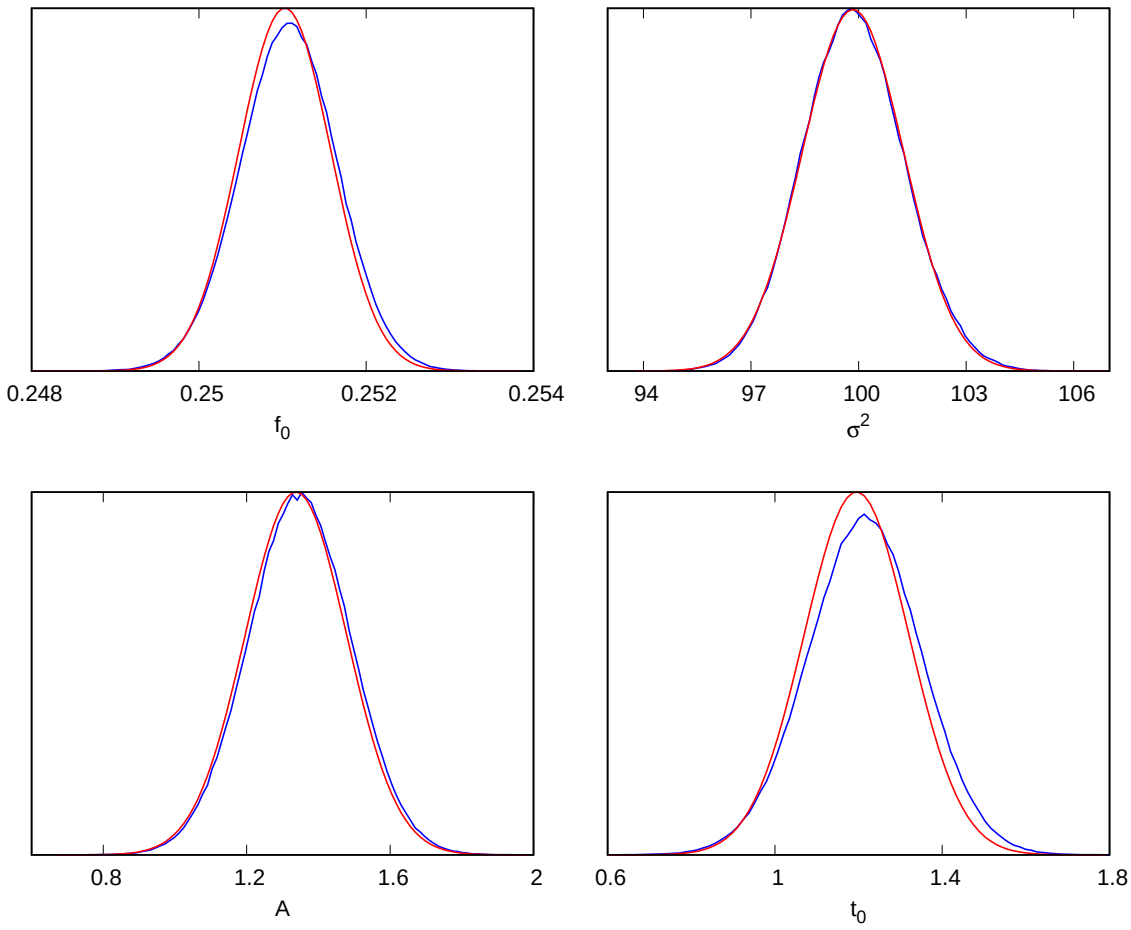

Figure 3.5: Marginalized posterior distributions for the signal and noise parameters from a MCMC simulation (blue) and the Gaussian ML approximation (red).

estimates for each signal such that the residual is consistent with the noise model. This procedure is illustrated in Figure 3.6 for the sinusoid signal model, where the waveform residuals from each iteration of the MCMC analysis are used to produce a histogram of the residual at each time sample. Also shown in Figure 3.6 is the MAP point estimate for the waveform residual. Notice that the full posterior distribution for the waveform residuals has significant spread about the point estimate.

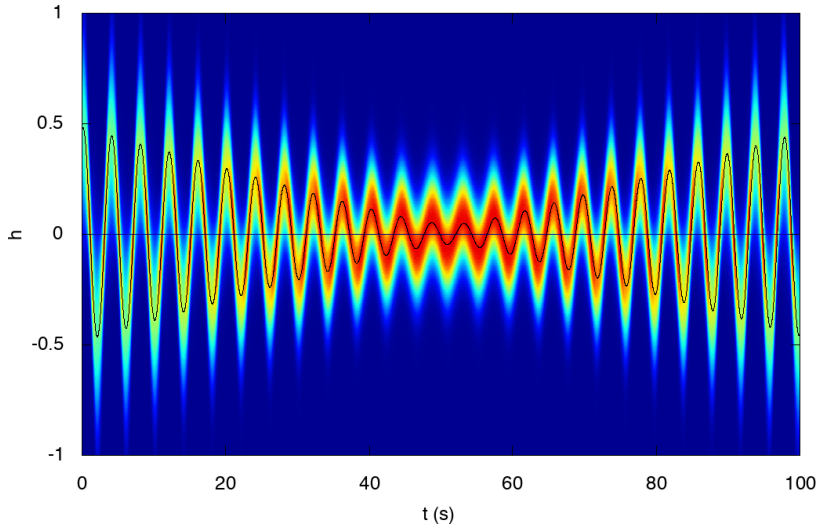

Figure 3.6: The posterior distribution for the waveform error computed from the MCMC analysis using a heat map to indicate the posterior weight, compared to the MAP point estimate (thin black line).

3.6 Multiple Sources

The LISA data will contain many signals that partially overlap in both time and frequency. Extracting information about these signals necessitates finding a global solution for all signals that can be resolved - that is, signals that are both sufficiently loud and sufficiently distinct to be individually identified. The full LISA data stream can be written as

$$\mathbf{s} = \mathbf{H} + \mathbf{n}\,, \tag{3.36}$$

where $\mathbf{H} = \sum \mathbf{h}_T$ denotes the sum total of all gravitational wave signals, which can be further separated into a resolved, $\mathbf{H}_R$, and un-resolved $\mathbf{H}_U$, component. The unresolved component is often referred to as "confusion noise", the largest component of which is expected to come from white dwarf binaries in our galaxy. The ML analysis for single sources can be applied to multiple sources by replacing $\mathbf{h}$ with $\mathbf{H}_R$, and by replacing $\mathbf{n}$ by $\mathbf{n} + \mathbf{H}_U$. The resolved signals will include a large number of bright

galactic binaries, along with multiple supermassive black holes and EMRIs.

The parameter estimation for the resolved systems will be impacted by the unresolved signals, which add to the effective noise level, and by the other resolved signals due to signal overlap. Some of the details of this calculation have been delegated to appendix B. To simplify the discussion imagine that the resolved signals consist of bright galactic binaries $\mathbf{H}_G$ and a single massive black hole binary $\bar{\mathbf{h}}$. The parameter vector $\vec{\lambda}$ runs over the full set of galactic binary parameters (denoted by indices early in the alphabet, $\lambda^a, \lambda^b \ldots$ *etc*) and the black hole parameters (denoted by indices later in the alphabet $\lambda^i, \lambda^j \ldots$ *etc*). The full set of signal parameters for the galactic sources and the black hole are indicated by Greek indices. The Fisher information matrix for the combined solution, $\Gamma_{\alpha\beta}$ can be broken into a block diagonal part $B_{\alpha\beta}$ formed from a galactic-binary block, $G_{ab} = (\partial_a \mathbf{H}_G | \partial_b \mathbf{H}_G)$, and a black hole block $\bar{B}_{ij} = (\partial_i \bar{\mathbf{h}} | \partial_j \bar{\mathbf{h}})$, and a mixed block $M_{aj} = (\partial_a \mathbf{H}_G | \partial_j \bar{\mathbf{h}})$. The waveform error for the black hole signal is then

$$\Delta\bar{\mathbf{h}} = -\partial_i \bar{\mathbf{h}} \left(\mathbf{n} | \partial_\alpha \mathbf{H}_R\right) \left(\Gamma^{-1}\right)^{i\alpha} + \ldots \ . \tag{3.37}$$

The expectation value for the squared SNR of the black hole waveform residual is then

$$\mathbb{E}\left[\left(\Delta\bar{\mathbf{h}} | \Delta\bar{\mathbf{h}}\right)\right] = \bar{B}_{ij} \left(\Gamma^{-1}\right)^{ij} + \ldots \ , \tag{3.38}$$

Assuming the mixture terms M_{aj} are small compared to the terms on the diagonal, the inverse of the full Fisher matrix can be expanded as

$$\begin{aligned} \left(\Gamma^{-1}\right)^{\alpha\beta} \quad &= \quad \left(B^{-1}\right)^{\alpha\beta} - \left(B^{-1}\right)^{\alpha\mu} M_{\mu\nu} \left(B^{-1}\right)^{\nu\beta} \\ &+ \quad \left(B^{-1}\right)^{\alpha\mu} M_{\mu\nu} \left(B^{-1}\right)^{\nu\gamma} M_{\gamma\eta} \left(B^{-1}\right)^{\eta\beta} + \ldots \ . \end{aligned} \tag{3.39}$$

The black-hole block of the inverse, $(\Gamma^{-1})^{ij}$, lacks the second term since $(B^{-1})^{i\mu} M_{\mu\nu} (B^{-1})^{\nu j} = \left(\bar{B}^{-1}\right)^{ik} M_{kl} \left(\bar{B}^{-1}\right)^{lj} = 0$. This calculation is shown more explicitly show in appendix B we therefore have

$$\mathbb{E}\left[\left(\Delta\bar{\mathbf{h}}|\Delta\bar{\mathbf{h}}\right)\right] = \bar{D} + \mathbb{E}\left[\left(\Delta\bar{\mathbf{h}}|\Delta\mathbf{H}_G\right)^2\right] + \dots \,. \tag{3.40}$$

where $\bar{D}$ is the dimension of the black hole model. We see that the black hole waveform residuals are inflated from the isolated source result by the mixture term. The second line is obtained using Isserilis's theorem and dropping the cross terms $\mathbb{E}[\Delta\lambda^i \Delta\lambda^a]$ which would produce higher order corrections. Using $\mathbb{E}[\Delta\tilde{H}_G(f)\Delta\tilde{H}^*_G(f')] = \frac{1}{2} S_{\Delta H_G}(f)\delta(f - f')$ we have

$$\mathbb{E}\left[\left(\Delta\bar{\mathbf{h}}|\Delta\mathbf{H}_G\right)^2\right] = 4\int_0^\infty \frac{|\Delta\bar{h}(f)|^2 S_{\Delta H_G}(f)}{S^2_{n,0}(f)} df \,. \tag{3.41}$$

If we were to switch the roles of the black hole and the resolved galactic binaries we would find the the squared SNR of the galactic residuals were inflated by exactly the same amount:

$$\mathbb{E}\left[(\Delta\mathbf{H}_G|\Delta\mathbf{H}_G)\right] = D_G + \mathbb{E}\left[\left(\Delta\bar{\mathbf{h}}|\Delta\mathbf{H}_G\right)^2\right] \,. \tag{3.42}$$

On the other hand, the squared SNR of the full residual is equal to the total parameter dimension:

$$\mathbb{E}\left[(\Delta\mathbf{H}_R|\Delta\mathbf{H}_R)\right] = \Gamma_{\alpha\beta}\left(\Gamma^{-1}\right)^{\alpha\beta} = \bar{D} + D_G \,. \tag{3.43}$$

These results can be reconciled by noting that

$$\mathbb{E}\left[(\Delta\mathbf{H}_R|\Delta\mathbf{H}_R)\right] = \mathbb{E}\left[(\Delta\mathbf{H}_G|\Delta\mathbf{H}_G)\right] + \mathbb{E}\left[\left(\Delta\bar{\mathbf{h}}|\Delta\bar{\mathbf{h}}\right)\right] + 2\mathbb{E}\left[\left(\Delta\bar{\mathbf{h}}|\Delta\mathbf{H}_G\right)\right] \,, \tag{3.44}$$

and using

$$\begin{aligned}\mathbb{E}\left[\left(\Delta\bar{\mathbf{h}}|\Delta\mathbf{H}_G\right)\right] &= M_{ia}\Gamma_{\alpha\beta}\left(\Gamma^{-1}\right)^{i\alpha}\left(\Gamma^{-1}\right)^{a\beta} \\ &= -M_{ia}M_{jb}\left(\bar{B}^{-1}\right)^{ij}\left(G^{-1}\right)^{ab} \approx -\mathbb{E}\left[\left(\Delta\bar{\mathbf{h}}|\Delta\mathbf{H}_G\right)^2\right] . \quad (3.45)\end{aligned}$$

Thus we see that the extra residual for each source class is canceled by the cross-correlation of the residuals between the source classes. Note that the results for the SNR of the signal residuals are unchanged to the order we are considering when using the full noise model or the unperturbed noise model.

Next we consider the impact on the black hole parameter estimation errors caused by fitting the bright galactic binaries. The variance in the parameter estimation errors can be estimated from the diagonal entries of inverse of the full Fisher information matrix

$$\left(\Gamma^{-1}\right)^{ii} = \left(\bar{B}^{-1}\right)^{ii} + \left(\bar{B}^{-1}\right)^{ik} M_{ka}\left(G^{-1}\right)^{ab} M_{bn}\left(\bar{B}^{-1}\right)^{ni} + \dots . \quad (3.46)$$

The second term in the expansion comes from correlations between the black hole signal and the resolved galactic binaries as is positive definite since $\mathbf{x}^T\mathbb{G}\mathbf{x} \geq 0$ for a positive-definite matrix. Thus, the simultaneous fitting of the galactic binary signals and the black hole signal tends to inflate the parameter estimation errors. Expanding to leading order the second term is given by

$$\bar{B}^{ik}M_{ka}G^{ab}M_{bn}\bar{B}^{ni} = 4\left(\bar{B}_0^{-1}\right)^{ik}\left(\bar{B}_0^{-1}\right)^{ni}\left(\mathbb{R}\int_0^\infty \frac{\partial_k\bar{\mathbf{h}}^*\partial_n\bar{\mathbf{h}}\,S_{\Delta H_G}}{S_{n,0}^2(f)}df\right) . \quad (3.47)$$

Using an estimate to the waveform errors of the resolved sources we can express this result in a more useful form:

$$4\left(\mathbb{R}\int_0^\infty \frac{\partial_k\bar{\mathbf{h}}^*\partial_n\bar{\mathbf{h}}\,S_{\Delta H_G}}{S^2_{n,0}(f)}df\right) \approx 2D_{\rm eff}\left(\mathbb{R}\int_0^\infty \frac{\partial_k\bar{\mathbf{h}}^*\partial_n\bar{\mathbf{h}}}{S_{n,0}(f)}\frac{dN}{db}df\right)$$
$$\leq 2D_{\rm eff}\left(\mathbb{R}\int_0^\infty \frac{\partial_k\bar{\mathbf{h}}^*\partial_n\bar{\mathbf{h}}}{S_{n,0}(f)}df\right)\left(\frac{dN}{db}\right)_{\rm max}. \quad (3.48)$$

This implies that the covariance matrix inflates with the following upper bound:

$$\left(\Gamma^{-1}\right)^{ii} \leq \left(\bar{B}_0^{-1}\right)^{ii}\left(1+2D_{\rm eff}\left(\frac{dN}{db}\right)_{\rm max}\right), \quad (3.49)$$

where dN/db of sources resolved per frequency bin. In the next section we will obtain an estimate for the noise due to GB waveform errors which will allow us to obtain a more useful expression for this overlap term. A similar inflation of GB parameter variances results from the overlap with the BH signal.

3.7 Galactic Confusion Noise

To obtain an estimate of the Galactic confusion foreground we employ an iterative subtraction scheme. Previously, this scheme was performed with perfect removal of source waveform [51, 70], but clearly noise will lead to errors in the parameters estimation and signal subtraction, and a reduction of the confusion noise estimate.

The revised procedure is as follows: A simulated data set is produced that includes a realization of the instrument noise and the sum of the strains due to the galactic binaries H from the galactic population model. A smooth fit to the power spectral density of the instrument noise and the signals is used as an initial estimate for the noise in each data channel. Next we identify sources which are loud (SNR > 7) relative to this noise estimate and subtract the best-fit waveform

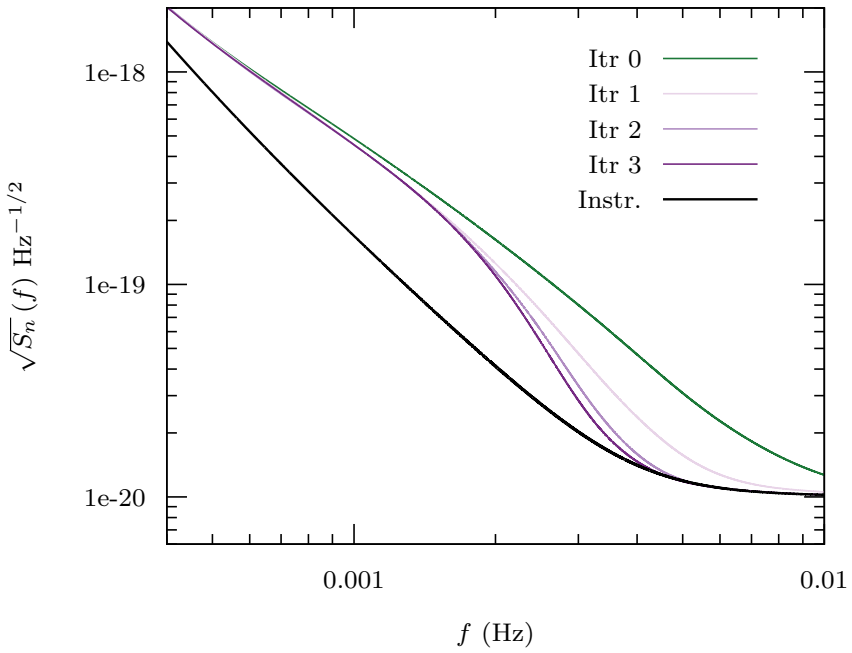

Figure 3.7: Estimates of the combined power spectral density of the combined instrument noise and galactic signals for the first 3 iterations on the removal process. These simulations are for a 4-year mission lifetime. For reference, the instrument noise contribution is shown as a heavy black line.

$\mathbf{h}(\vec{\lambda})_{\text{bestfit}} = \mathbf{h}_T - \partial_i \mathbf{h}_T \Delta\lambda^i$ from the data (where equation (9) defines $\Delta\lambda^i$). A smooth fit to the power spectral density of the remaining signals and noise is computed, and signals above the SNR threshold for the updated noise estimate are identified and subtracted. As can be seen in Figure 3.7, the subtraction procedure quickly converges.

The number of sources which can be resolved converges after just 5 or 6 iterations. The simulations included the signals from 29,624,349 galactic binaries, 3,185,491 of which are AM CVn. Table 2 displays how many sources were identified and removed at each iteration. A total of 30,442 sources were considered bright using both the A and E channels while only 19,112 using the X channel. The number density of sources, measured in terms of the number per a frequency bin, dN/db, is shown in Figure 3.8 for a single channel and dual channel analysis. We see that

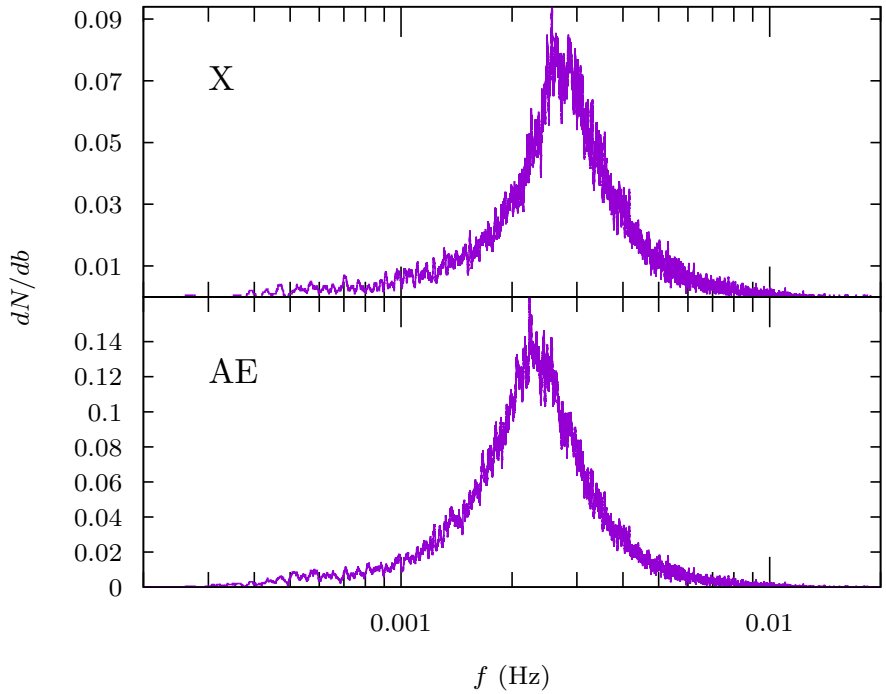

Figure 3.8: The number of sources resolved per frequency bin, dN/db, for a single channel analysis X, and a dual channel analysis AE, for a 4-year mission duration. Note the difference in the vertical scales; the combined information in the AE data stream allows for more sources to be resolved. The density of resolvable systems peaks at around $2-3$ mHz.

more sources can be resolved when multiple data channels are used in the analysis. Attempts have been made to deal with the identification and subtraction of signals which overlap [36, 105] and of how many sources per frequency band [29] can be resolved. Here we see that with two data channels and a 4-year mission the peak density is roughly one source resolved per ten frequency bins. In our simulations we made the simplifying assumption that the augment Fisher matrix for the galactic population is block diagonal. That is, we ignored correlations between galactic signals. This approximation is reasonable when dN/db is small, but may be questionable in the highest density regions and for the occasional systems that happen to have high overlap. We will re-visit this complication in a future study, as the parameter

Table 3.2: Number of sources identified as bright at each iteration

Iteration	X	A,E
1	8,910	13,697
2	14,415	22,658
3	16,970	27,027
4	18,170	28,999
5	18,765	29,913
6	19,005	30,284
7	19,085	30,445

estimation errors grow significantly for highly overlapping systems [96].

Figure 3.9 compares the Michelson-equivalent strain power spectral densities for the imperfect and perfect subtraction scheme. The dashed lines show galactic confusion noise and the solid lines show the combined instrument and confusion noise. Note that the differences between the PSD arise where the most sources are resolved as one would expect (see Figure 3.8). In the dual A, E channel the PSD is lower as indicated by the noise levels specified by the reference frequency 2 mHz.

We can estimate the power spectral density of the combined waveform residual, $S_{\Delta H_R}(f) = \frac{T}{2}\mathbb{E}\left[|\Delta\tilde{H}_R(f)|^2\right]$ by applying (3.13) to the full compliment of N resolved binaries:

$$\mathbb{E}\left[\rho^2_{\Delta H_R}\right] = N D_{\text{eff}} = 4\int_0^\infty \frac{\mathbb{E}\left[|\Delta\tilde{H}_R(f)|^2\right]}{S_{n,0}(f)} df \ ,$$

Considering the contribution in a small frequency range Δf centered at f we find

$$S_{\Delta H_R}(f) = \frac{D_{\text{eff}}}{2}\frac{dN}{db} S_{n,0}(f) \ . \tag{3.50}$$

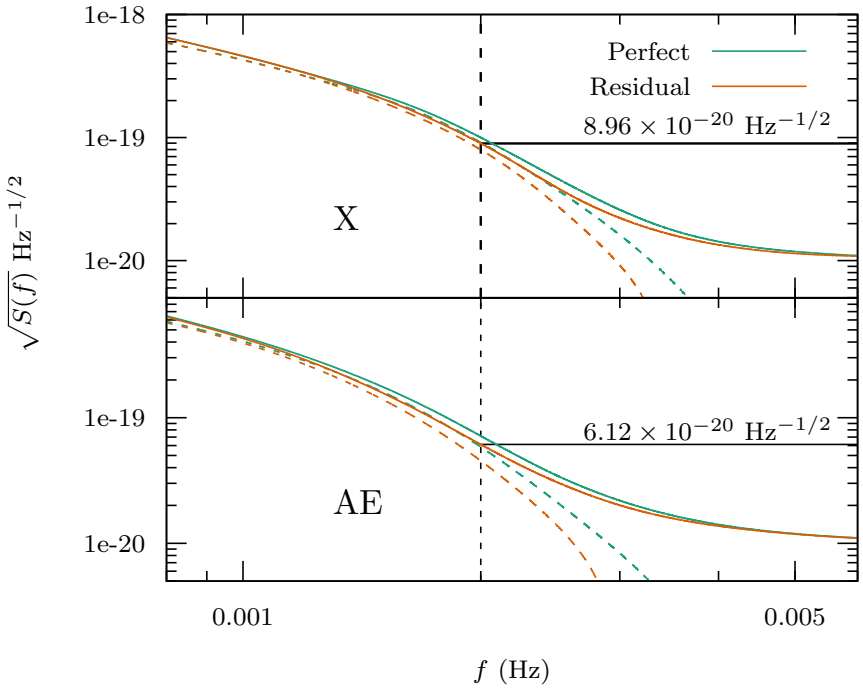

Figure 3.9: The Michelson-equivalent stain PSD in both the X and combined A, E data channels. The dashed lines represent the the GB confusion noise. A reference frequency of 2 mHz has been denoted by a vertical dashed line for noise level comparison on the corresponding horizontal solid black lines.

Figure 3.10 compares this analytic estimate to the numerical value found from the iterative subtraction scheme. The prediction lines up quite well with a small deviation at the frequencies where the number of resolved sources per bin peaks.

3.8 Discussion

We have used the maximum likelihood approximation to derive a number of analytic results pertaining to the LISA global analysis problem. A simple toy model was used to demonstrate the relevance of these estimates to a full Bayesian analysis, though we cannot guarantee that the approximations will be as reliable when applied to LISA data analysis. We extended the standard maximum likelihood analysis to include noise modeling, and found that the estimated noise level is lowered

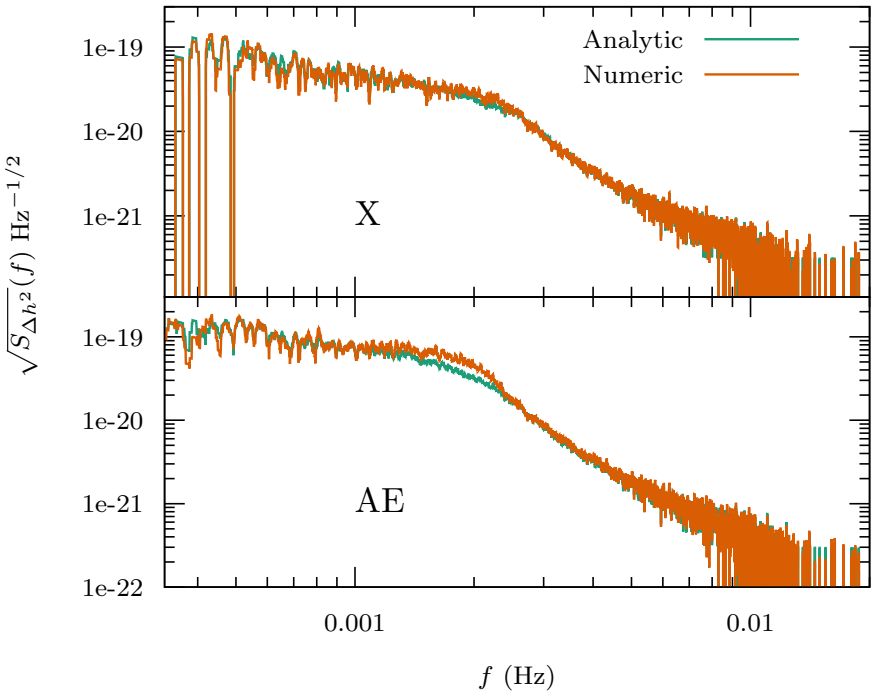

Figure 3.10: Comparison between the predicted (teal) strain power spectral density and that obtained through the imperfect subtraction scheme (orange) of the GB confusion strictly plotted as their Michelson-equivalent strain.

whenever signals are subtracted from the data. We applied our general results to the simultaneous fitting of a black hole binary and a collection of galactic binaries and found that the errors in the black hole waveform recovery are increased, as wells as the variances of the source parameters. It is important to note that it is not the overlap of the signals which causes the inflation of variances, but rather, it is due to the waveform subtraction errors. We concluded by incorporating parameter estimation errors in the estimation of the galactic confusion noise, and derived a useful expression that can be used to predict the reduction in the confusion noise in terms of the number density of resolved signals. Equations (3.33) and (3.49) provide quick estimates for how parameter estimation errors are inflated by noise fitting and source confusion.

Acknowledgments

We are grateful for the support provided by NASA grant NNX16AB98G.

Appendix A

Here we derive the more general signal plus noise model ML analysis results presented in section 4. Beginning with the log-likelihood obtained from equation (3.3)

$$\begin{aligned}\log p\left(\mathbf{s}\middle|\mathbf{h}(\vec{\lambda}), S_n(f;\vec{\theta})\right) = \; & -\tfrac{1}{2}\int_0^\infty T\log\left[\pi T S_n(f;\vec{\theta})\right] df \\ & -\tfrac{1}{2}\left(\mathbf{s}-\mathbf{h}(\vec{\lambda})\middle|\mathbf{s}-\mathbf{h}(\vec{\lambda})\right)_{\vec{\theta}} \; , \end{aligned} \tag{3.51}$$

where the noise-weighted inner product is now parameterized by $\vec{\theta}$ as denoted by $(\cdot|\cdot)_{\vec{\theta}}$. We may expand about the true noise model $S_{n,0}(f)$ and the true signal model $\mathbf{h}_T$ and maximized to obtain estimates of $\Delta\vec{\theta}$ and $\Delta\vec{\lambda}$. Expanding out the normalization constant and dropping terms which are constant with respect the maximization gives

$$\begin{aligned}-\frac{1}{2}\int_0^\infty \log\left[2\pi T S_n(f;\vec{\theta})\right] df \approx \; & -\tfrac{1}{2}\Delta\theta^i\int_0^\infty \tfrac{T S_{n,i}}{S_{n,0}} df - \tfrac{1}{4}\Delta\theta^i\Delta\theta^j\int_0^\infty \tfrac{T S_{n,ij}}{S_{n,0}} df \\ & +\tfrac{1}{4}\Delta\theta^i\Delta\theta^j\int_0^\infty \tfrac{T S_{n,i} S_{n,j}}{S_{n,0}^2} df \; , \end{aligned} \tag{3.52}$$

where T is the observation period. The noise-weighted inner product when expanded takes the form

$$(\mathbf{a}|\mathbf{b})_{\vec{\theta}} \approx (\mathbf{a}|\mathbf{b})_0 - \Delta\theta^i(\mathbf{a}|\mathbf{b})_i - \frac{1}{2}\Delta\theta^i\Delta\theta^j(\mathbf{a}|\mathbf{b})_{ij} + \Delta\theta^i\Delta\theta^j(\mathbf{a}|\mathbf{b})_{i;j} \; , \tag{3.53}$$

where

$$
\begin{aligned}
(\mathbf{a}|\mathbf{b})_i &\coloneqq 4\mathbb{R}\int_0^\infty \frac{\tilde{a}^*\tilde{b}}{S_{n,0}}\frac{S_{n,i}}{S_{n,0}}df \ , \\
(\mathbf{a}|\mathbf{b})_{ij} &\coloneqq 4\mathbb{R}\int_0^\infty \frac{\tilde{a}^*\tilde{b}}{S_{n,0}}\frac{S_{n,ij}}{S_{n,0}}df \ , \\
(\mathbf{a}|\mathbf{b})_{i;j} &\coloneqq 4\mathbb{R}\int_0^\infty \frac{\tilde{a}^*\tilde{b}}{S_{n,0}}\frac{S_{n,i}S_{n,j}}{S_{n,0}^2}df \ .
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
-\frac{1}{2}\left(\mathbf{s}-\mathbf{h}(\vec{\lambda})\middle|\mathbf{s}-\mathbf{h}(\vec{\lambda})\right)_{\vec{\theta}} \approx \quad & \tfrac{1}{2}\Delta\theta^i(\mathbf{n}|\mathbf{n})_i + \tfrac{1}{4}\Delta\theta^i\Delta\theta^j(\mathbf{n}|\mathbf{n})_{ij} \\
& -\tfrac{1}{2}\Delta\theta^i\Delta\theta^j(\mathbf{n}|\mathbf{n})_{i;j} + \Delta\theta^i(\mathbf{n}|\Delta\mathbf{h})_i \\
& +\tfrac{1}{2}(\mathbf{n}|\Delta\mathbf{h})_0 + \tfrac{1}{2}\Delta\theta^i(\Delta\mathbf{h}|\Delta\mathbf{h})_i \\
& -\tfrac{1}{2}(\Delta\mathbf{h}|\Delta\mathbf{h})_0 \qquad (3.54)
\end{aligned}
$$

Collecting terms, and maximizing the likelihood yields the solution

$$
\begin{aligned}
\Delta\lambda^j &= (\mathbf{n}|\partial_j\mathbf{h}_T)_{\vec{\theta}}\left(\Gamma^{-1}_{\vec{\theta}}\right)^{ij} \qquad (3.55) \\
\Delta\theta^j &\approx \left[\int \frac{S_{n,i}S_{n,j}}{S_{n,0}^2}\left(\frac{TS_{n,0}-4\tilde{n}^*\tilde{n}}{S_{n,0}}\right)df - \int \frac{S_{n,ij}}{S_{n,0}}\left(\frac{TS_{n,0}-2\tilde{n}^*\tilde{n}}{S_{n,0}}\right)df\right]^{-1} \\
&\quad \times \left[\int \frac{S_{n,i}}{S_{n,0}}\left(\frac{TS_{n,0}-2\tilde{n}^*\tilde{n}}{S_{n,0}}\right)df - 2(\mathbf{n}|\Delta\mathbf{h})_i - (\Delta\mathbf{h}|\Delta\mathbf{h})_i\right] \ , \qquad (3.56)
\end{aligned}
$$

where $\Delta\mathbf{h}$ need only be kept to leading order:

$$
\begin{aligned}
\Delta\theta^j \approx \quad & \left[\int \tfrac{S_{n,i}S_{n,j}}{S_{n,0}^2}\left(\tfrac{TS_{n,0}-4\tilde{n}^*\tilde{n}}{S_{n,0}}\right)df - \int \tfrac{S_{n,ij}}{S_{n,0}}\left(\tfrac{TS_{n,0}-2\tilde{n}^*\tilde{n}}{S_{n,0}}\right)df\right]^{-1} \\
& \times\Bigg[\int \tfrac{S_{n,i}}{S_{n,0}}\left(\tfrac{TS_{n,0}-2\tilde{n}^*\tilde{n}}{S_{n,0}}\right)df + 2(\mathbf{n}|\partial_a\mathbf{h}_T)_i(\mathbf{n}|\partial_b\mathbf{h}_T)_0\left(\Gamma_0^{-1}\right)^{ab} \\
& \qquad - (\partial_a\mathbf{h}_T|\partial_b\mathbf{h}_T)_i(\mathbf{n}|\partial_c\mathbf{h}_T)_0(\mathbf{n}|\partial_d\mathbf{h}_T)_0\left(\Gamma_0^{-1}\right)^{ac}\left(\Gamma_0^{-1}\right)^{bd}\Bigg] . \qquad (3.57)
\end{aligned}
$$

The Fisher matrix can be obtained similarly to the toy white noise problem presented in Section 3.4:

$$\mathbf{\Gamma} = \begin{pmatrix} (\partial_i \mathbf{h}_T|\partial_j \mathbf{h}_T)_0 + \frac{1}{2}(\mathbf{n}|\partial_{ij}\mathbf{h}_T)_0 & (\mathbf{n}|\partial_j \mathbf{h}_T)_i \\ (\mathbf{n}|\partial_i \mathbf{h}_T)_j & \frac{1}{2}\int \frac{S_{n,ij}}{S_{n,0}}\left(\frac{2\tilde{n}^*\tilde{n}-TS_{n,0}}{S_{n,0}}\right)df \\ & +\frac{1}{2}\int \frac{S_{n,i}S_{n,j}}{S_{n,0}^2}\left(\frac{4\tilde{n}^*\tilde{n}-TS_{n,0}}{S_{n,0}}\right)df \end{pmatrix}. \tag{3.58}$$

Inverting the Fisher matrix and focusing on the diagonal elements we find

$$\begin{aligned} \left(\Gamma^{-1}\right)^{ii} \approx & \left(\Gamma_0^{-1}\right)^{ii} - \frac{1}{2}\left(\Gamma_0^{-1}\right)^{im}(\mathbf{n}|\partial_{mn}\mathbf{h}_T)_0\left(\Gamma_0^{-1}\right)^{ni} \\ & +2\left(\Gamma_0^{-1}\right)^{im}(\mathbf{n}|\partial_m \mathbf{h}_T)_a\left[\int_0^\infty \frac{S_{n,ab}}{S_{n,0}}\left(\frac{2\tilde{n}^*\tilde{n}-TS_{n,0}}{S_{n,0}}\right)df \right. \\ & \left. + \int_0^\infty \frac{S_{n,a}S_{n,b}}{S_{n,0}^2}\left(\frac{4\tilde{n}^*\tilde{n}-TS_{n,0}}{S_{n,0}}\right)df\right]^{-1}(\mathbf{n}|\partial_n \mathbf{h}_T)_b\left(\Gamma_0^{-1}\right)^{ni}. \end{aligned} \tag{3.59}$$

Appendix B

In this appendix we demonstrate that the SNR of the BH residuals are inflated by overlaps with galactic binary residuals, and that these overlaps also inflate parameter estimation errors. The expectation value for squared SNR of the the BH residuals is given by

$$\begin{aligned} \mathbb{E}\left[\left(\Delta\bar{\mathbf{h}}|\Delta\bar{\mathbf{h}}\right)\right] &= \bar{B}_{ij}\left(\left(\bar{B}^{-1}\right)^{ij} + \left(\bar{B}^{-1}\right)^{ik} M_{ka}\left(G^{-1}\right)^{ab} M_{al}\left(\bar{B}^{-1}\right)^{lj}\right) + \dots \\ &= \bar{D} + \left(\bar{B}^{-1}\right)^{ij}\left(G^{-1}\right)^{ab} M_{ja}M_{bi} + \dots . \end{aligned} \tag{3.60}$$

where $\bar{D}$ is the dimension of the black hole model. The second term containing a mixture between BH and GW signals can be expressed as

$$\begin{aligned}\left(\bar{B}^{-1}\right)^{ij}\left(G^{-1}\right)^{ab}M_{ja}M_{bi} &= \mathbb{E}\left[\Delta\lambda^i\Delta\lambda^j\right]\mathbb{E}\left[\Delta\lambda^a\Delta\lambda^b\right]\left(\partial_j\bar{\mathbf{h}}|\partial_a\mathbf{H}_G\right)\left(\partial_b\mathbf{H}_G|\partial_i\bar{\mathbf{h}}\right)\\ &\approx \mathbb{E}\left[\Delta\lambda^i\Delta\lambda^j\Delta\lambda^a\Delta\lambda^b\right]\left(\partial_j\bar{\mathbf{h}}|\partial_a\mathbf{H}_G\right)\left(\partial_b\mathbf{H}_G|\partial_i\bar{\mathbf{h}}\right)\\ &= \mathbb{E}\left[\left(\Delta\bar{\mathbf{h}}|\Delta\mathbf{H}_G\right)^2\right]\ . \end{aligned} \tag{3.61}$$

As for the approximate covariance matrix

$$\left(\Gamma^{-1}\right)^{ii} = \left(\bar{B}^{-1}\right)^{ii} + \left(\bar{B}^{-1}\right)^{ik}M_{ka}\left(G^{-1}\right)^{ab}M_{bn}\left(\bar{B}^{-1}\right)^{ni} + \dots\ , \tag{3.62}$$

the second term, again containing mixture factors can be simplified as follows

$$\begin{aligned}\bar{B}^{ik}M_{ka}G^{ab}M_{bm}\bar{B}^{ni} &= \mathbb{E}\left[\Delta\lambda^a\Delta\lambda^b\left(\partial_k\bar{\mathbf{h}}|\partial_a\mathbf{H}_R\right)_0\left(\partial_b\mathbf{H}_R|\partial_n\bar{\mathbf{h}}\right)_0\right]\left(\bar{B}_0^{-1}\right)^{ik}\left(\bar{B}_0^{-1}\right)^{ni},\\ &= \mathbb{E}\left[(\mathbf{h}_{,k}|\Delta\mathbf{H}_G)_0(\Delta\mathbf{H}_G|\mathbf{h}_{,n})_0\right]\left(\bar{B}_0^{-1}\right)^{ik}\left(\bar{B}_0^{-1}\right)^{ni}\\ &= 4\left(\bar{B}_0^{-1}\right)^{ik}\left(\bar{B}_0^{-1}\right)^{ni}\left(\mathbb{R}\int_0^\infty\frac{\partial_k\bar{\mathbf{h}}^*\partial_n\bar{\mathbf{h}}\,S_{\Delta H_G}}{S_{n,0}^2(f)}df\right)\ . \end{aligned} \tag{3.63}$$

where $\partial_a H_R = \partial_a H_G$ as the derivatives are with respect to GB parameters.

CHAPTER FOUR

DETECTING HIERARCHICAL STELLAR SYSTEMS WITH LISA

Contributions of Authors and Co-Authors

Manuscript in Chapter 4

Author: Travis Robson

Contributions: Developed the model and code needed for the study. Performed the analysis. Wrote the first draft of the manuscript.

Co-Author: Dr. Neil Cornish

Contributions: Conceived the study design. Provided feedback/guidance for the analysis. Outlined the first draft and edited the manuscript.

Co-Author: Dr. Nicola Tamanini

Contributions: Edited the manuscript. Performed EKL calculation.

Co-Author: Dr. Silvia Toonen

Contributions: Wrote the astrophysics section of the manuscript.

Manuscript Information Page

Travis Robson, Neil Cornish, Nicola Tamanini, Silvia Toonen

Physical Review D

Status of Manuscript:

Published in a peer-reviewed journal

Published by American Physical Society

Published September, 2018, Phys. Rev. D **98**, 064012

Abstract

A significant fraction of stars are members of gravitationally bound hierarchies containing three or more components. Almost all low mass stars in binaries with periods shorter than three days are part of a hierarchical system. We therefore anticipate that a large fraction of compact galactic binaries detected by the Laser Interferometer Space Antenna (LISA) will be members of hierarchical triple or quadruple system. The acceleration imparted by the hierarchical companions can be detected in the gravitational wave signal for outer periods as large as 100 years. For systems with periods that are shorter than, or comparable to, the mission lifetime, it will be possible to measure the period and eccentricity of the outer orbit. LISA observations of hierarchical stellar systems will provide insight into stellar evolution, including the role that Lidov-Kozai oscillations play in driving systems towards merger.

4.1 Introduction

The Laser Interferometer Space Antenna (LISA) [63] is expected to individually resolve the signals from tens of thousands of compact galactic binaries during its nominal four year mission lifetime [70]. Roughly 13% of low mass stellar systems contain three or more stars [106–108], and roughly 96% of low mass binaries with periods shorter than 3 days are part of a larger hierarchy [109, 110]. While the multiplicity distribution for ultra-compact binaries is currently unknown, it is reasonable to expect that a significant fraction of compact galactic binary systems detected by LISA will be members of a hierarchical system. Indeed, dynamical effects in hierarchical stellar systems such as oscillations in the eccentricity due to the (regular & eccentric) Lidov-Kozai mechanism [111–113] can cause the inner binary

orbit to harden, potentially enhancing the fraction of compact binary systems with companions [114–120].

The hierarchal companions to the ultra-compact binaries detected by LISA will impart accelerations on the center of mass of the binary that can lead to observable doppler shifts in the signals. This effect has previously been considered in the context of LISA observations of extreme mass ratio binaries [121], and merging black hole binaries detected by LIGO and LISA [122–125]. The mathematical description is essentially identical to that in pulsar timing, where the orbital parameters of pulsars found in binary systems can be inferred from modulations of the radio pulses [126]. One difference between the radio and gravitational wave analyses is that wavelengths of the gravitational waves are significantly larger than the gravitational radii of the stars, which modifies the calculation of the Shapiro time-delay.

Here we consider LISA observations of compact galactic binaries in hierarchical systems and identify three main regimes that are governed by the ratio of the outer orbital period to the observation time: (1) When the outer period is much larger than the observation time the hierarchical orbit imparts an overall unobservable doppler shift; (2) When the outer period is up to a factor ten or so larger than the observation time the influence of the companion can be detected; (3) When the outer period is shorter than or comparable to the observation time, the eccentricity and period of the hierarchical orbit can be inferred. In rare cases a fourth regime can occur: (4) The acceleration due to the hierarchal perturber can be mistaken for frequency changes due to gravitational wave emission or mass transfer. This regime occurs when the outer period is larger than the observation time, and the chirp mass and gravitational wave frequency of the compact binary lie in a narrow range of values. The precise location of the boundaries between the four cases depends on several factors, including the signal-to-noise ratio, the gravitational wave frequency, the mass ratio between

the inner binary and the perturber, and the eccentricity of the outer orbit. Using a simple Fisher matrix based estimate for when the frequency change of a nearly monochromatic signal can be detected, we arrive at the condition that, on average, the outer binary can be detected when the period of the outer orbit, P_2 obeys the inequality

$$P_2 \lesssim 43.2 \text{ yrs} \left(\frac{\rho}{10} \cdot \frac{m_c}{1.0 M_\odot} \cdot \frac{f}{5 \text{ mHz}}\right)^{3/4} \left(\frac{m_2}{2 M_\odot}\right)^{-1/2} \times \left(\frac{T_{\rm obs}}{4 \text{ yr}}\right)^{3/8} \left(\frac{1 + \frac{1}{2} e_2^2}{(1 - e_2^2)^{5/2}}\right)^{3/8} . \quad (4.1)$$

This expression is valid for $P_2 > T_{\rm obs}$, where $T_{\rm obs}$ is the observation time. For shorter periods higher derivatives of the frequency change with respect to time need to be accounted for. Other quantities that appear in the expression are the signal-to-noise ratio ρ, the mass of the perturber m_c, the gravitational wave frequency f, the total mass of the system m_2 and the eccentricity of the hierarchical orbit e_2, and we work in geometrical units with $G = c = 1$. To derive this expression we computed the root-mean-square (RMS) line-of-sight acceleration of inner binary due to the distant companion, averaging over the orbital period and orientation. Note that some systems will be detectable with longer periods if the orientation and phase of the orbit is more favorable. Also note that LISA is expected to detect hundreds of galactic binaries with signal-to-noise ratios (SNR) $\rho > 100$ [70], and for these systems it will be possible to detect systems with $P_2 > 100$ years.

The outline of the paper is as follows: In §4.2 we review what is known about compact binaries in hierarchical systems. In §4.3 we summarize the models and methods used in our study. The orbital model is described in more detail in §4.4 and the gravitational wave modeling is outlined in §4.5. The detectability of hierarchical companions is considered in §4.6, and the characterization of the orbits is investigated

in §4.7. The possibility of confusing the acceleration caused by a distant perturbed with orbital evolution due to radiation reaction or mass transfer is discussed in §4.8. We conclude with a summary and discussion of future studies in §4.9.

4.2 Compact Binaries In Hierarchical Systems

The majority of stars are members of multiple systems, including binaries, triples and higher-order hierarchies. The triple fraction is best known for stellar systems with main-sequence components, in particular for lower mass stars of F- and G-type where a triple fraction of 11-20% are found [106, 108, 127]. There are indications that the triple fraction increases for higher mass stars as for binaries [128, 129]. The period distribution of the inner and outer orbits of triples with F- and G-type primaries are distributed similarly as those of binaries, however, with the additional constraint that the triple is dynamically stable [130]. As a result the inner orbits tend to be more compact, leading more often to mass transfer episodes and compact binaries (Toonen et al. in prep.). Besides the initial structure of the triple, three-body dynamics can provide additional means to harden the inner binary. The classical low-order quadrupole approximation of the three-body leads to Lidov-Kozai cycles, in which the mutual inclination between the two orbits and the eccentricity of the inner binary vary periodically [111–113]. The octupole terms give rise to the eccentric Lidov-Kozai mechanism, in which even higher eccentricities can be reached [113]. As a result the (regular and eccentric) Lidov-Kozai mechanism is linked to a variety of exotic astrophysical phenomena, such as stellar mergers in isolated triples [131–135], stellar mergers in binaries near massive black holes [136–140], low-mass X-ray binaries [141–143], blue stragglers [118, 144], as well as enhanced dissipation through gravitational wave emission and tides [145, 146]. Due to the latter mechanism, also known as high-eccentricity migration or Lidov-Kozai cycles with tidal friction (LKCTF), the inner

binary tightens forming hot Jupiters, e.g. [147–152], and compact binaries [117, 118, 145,146]; Observationally roughly 96% of low-mass binaries with periods shorter than three days have outer companions [109, 110].

In the context of GW sources, Lidov-Kozai cycles are relevant, as the gravitational wave inspiral time of a close (inner) binary with compact objects can be significantly reduced, if an outer star is present. Whereas isolated compact binaries need to be formed at periods $\lesssim$ 0.3 days to merge within a Hubble time, the presence of an outer companion extends the inner period range to hundreds of days if LKCTF is efficient. Even wider inner orbits can be brought to merge or collide if the triple system is weakly hierarchical for which the secular perturbation theory breaks down [120,136,153,154]. Such mergers of compact objects occur in orbits with higher residual eccentricities e.g. [155–157].

On the observational side, our knowledge of the triple fraction and orbital structures of triples with compact objects is limited. The highly complete sample of WDs within 20pc from the Sun, contains one to two triples with an inner compact double WD, showing that indeed its possible to form such objects e.g. [154]. Moreover, out of about 130 objects in total, there is one confirmed isolated compact double WD and four candidates, indicating that triple sources are relatively abundant.

When shifting our attention from compact double white dwarfs to wide systems, there are only two such binaries within 20pc. This is in contradiction to theory, from which one would expect 15-30 such systems within 20 pc [154]. As destruction mechanisms (e.g. dynamical interactions or stellar winds) are not efficient enough to explain the discrepancy, it has been claimed that the progenitor systems are not formed as efficiently as expected e.g. [154] or that the wide double white dwarfs have been missed observationally [158], however in the state-of-the-art sample of Gaia no new wide double white dwarfs were found within 20 pc [159].

Interesting to mention is PSR J0337+1715, the millisecond pulsar in a hierarchical triple with two white dwarfs [160] with periods of 1.6d and 327d. As both white dwarfs are low-mass helium dwarfs, the system demonstrates that it is possible in nature for a triple to survive several phases of mass transfer (see, e.g. [161, 162] for possible formation scenarios), and have outer periods in the range of the LISA mission lifetime.

4.3 Summary of Models and Methods

The natural separation of scales found in hierarchical systems allows us to make a number of simplifying assumptions. The few-body Hamiltonian for a hierarchical system can be expanded in the ratio of the semi-major axes yielding terms at monopole, quadrupole, octupole and higher orders [163–165]. Here we are mostly interested in 2:1 and 2:2 component hierarchies where the semi-major axis of the binary components are much smaller than semi-major axis of the overall system. Because the hierarchical periods we are considering will be comparable to or larger than the observation time, we can restrict our analysis to the leading order, Newtonian monopole interactions. In this approximation, the motion of the binaries separates from that of the hierarchical system, and each can be treated as a separate Keplerian system. The center of mass of the inner binary follows a Keplerian orbit around the distant perturber. We are justified in doing this since the Lidov-Kozai [166, 167] and eccentric Lidov-Kozai [163, 168] oscillations induced by the quadrupole and octapole terms occur on timescales that are long compared to the period of the hierarchical orbit, and very much longer then the observation time. The same is true for the high-order post-Newtonian effects such as periastron precession.

To check that the monopole description is sufficient for our analysis, we can derive qualitative estimate of the contribution of the Lidov-Kozai and eccentric

Lidov-Kozai mechanisms on the triple systems under consideration. We compute the variation over the LISA mission lifetime of the inner (e_1) and outer (e_2) eccentricities and the inner (ι_1) and outer (ι_2) inclination angles, due to the combined effect of those two mechanisms. We consider the expressions for the time variation of these physical quantities given in [113] and in order to obtain rough conservative estimates we set all geometrical factors (sines and cosines) to unity and then add the absolute value of any additive term in those expressions. From this analysis we find that the variation over a 4 year LISA mission duration of the outer inclination angle (ι_2) and the outer eccentricity (e_2) are always negligible, irrespectively of the triple's parameters. Similarly the change on the values of the inner inclination angle (ι_1) and inner eccentricity (e_1) is always below a few percent, unless we have very small WD masses ($\lesssim 0.2 M_\odot$), large inner eccentricities ($e_1 \gtrsim 0.3$) and/or low GW frequencies ($f < 1$ mHz). We will leave the investigation of these particular systems to future studies and in what follow we will assume the quantities ι_1, e_1, ι_2, e_2 to be constant. This approximation should work well for the majority of the systems that LISA will be able to observe.

In light of this, we allow for the outer hierarchical orbit to be eccentric, but make the further simplifying assumption that the inner orbit responsible for the gravitational wave emission is circular. We can justify this choice in two ways. First, gravitational radiation acts to quickly circularize orbits, and second, even if effects such as Lidov-Kozai oscillations have managed to maintain the eccentricity of the inner binary, our results will be little changed, at least for moderate eccentricities. The reasoning is as follows: for slowly evolving, moderately eccentric systems the gravitational wave signal can be expressed as a sum of circular binaries with periods at harmonics of the orbital period. The separation of these harmonics in frequency is very much larger than the sidebands imparted by the hierarchical orbit, so there

is zero confusion between the two effects. The sum of circular binary signals for an eccentric system contains almost identical information to that of a single circular binary for the purposes of the current analysis, so in the interests of computational efficiency we neglect the eccentricity of the inner binary. To verify this argument we model a triple system with non-zero inner eccentricity and perform parameter estimation by obtaining the marginalized posteriors for the outer orbit parameters. We compare the posteriors to those for the exact same system but with zero inner eccentricity and find that the posteriors are identical, hence demonstrating that the presence of sidebands due to eccentric of the inner orbit will not affect our results - see the end of section §4.7 for details.

To assess the detectability of the distant companion and the accuracy with which the parameters of the orbits can be inferred we use a mixture of methods. To make quick estimates and derive analytic scalings we compute Fisher information matrices, and to spot check these estimates and provide more detailed results we employ Bayesian inference via the Markov Chain Monte Carlo algorithm.

4.4 Hierarchical Orbit Model

In this section we derive how the perturbing companion affects the center-of-mass motion of the inner binary in the hierarchical orbit which will impart perturbations to the gravitational waveform. We desire to extract the line-of-sight component of the inner binary's center-of-mass velocity. For an isolated binary, its center-of-mass is stationary with respect to the solar system barycenter (ignoring unobservable constant peculiar velocities), but this line-of-sight component of the induced center-of-mass will create a time-dependent red-shift as seen in the barycenter frame. We will use "1" subscripts to denote orbital parameters of the inner gravitational wave emitting binary comprised of masses m_a and m_b for a total mass of m_1. The subscript "2"

will denote the Keplerian outer orbit describing the motion of the perturber m_c and the monopole mass of the inner binary. In our hierarchical approximation, in which we essentially have a circular Keplerian orbit emitting gravitational wave visible to LISA inside of a larger outer Keplerian orbit that is governed by

$$\mathbf{a}_2 = -\frac{Gm_2}{r_2^2}\hat{\mathbf{r}}_2 \ , \tag{4.2}$$

where $\mathbf{a}_2$ is the relative acceleration and $\hat{\mathbf{r}}_2 = \hat{\mathbf{r}}_c - \hat{\mathbf{r}}_1$ is the unit separation vector as defined in an inertial coordinate system of the triple, and $m_2 = m_a + m_b + m_c$.

The solution for the orbital motion is then

$$\mathbf{r}_2(t) = r_2(t)\left(\cos\varphi_2, \sin\varphi_2, 0\right) \ , \tag{4.3}$$

where

$$r_2(t) = \frac{a_2(1-e_2^2)}{1 + e_2\cos\varphi_2} \ , \tag{4.4}$$

defining the standard Keplerian ellipse. The quantities introduced are defined as follows: φ_2 is the orbital phase of the outer orbit, e_2 and a_2 are its eccentricity and semi-major axis respectively. In order to relate the orbital phase to time, its convention to introduce the eccentric anomaly which is an angle related to the orbital phase by the geometric relation

$$\varphi = \varphi_0 + 2\tan^{-1}\left(\sqrt{\frac{1+e}{1-e}}\tan\frac{u}{2}\right) \ , \tag{4.5}$$

where u is the eccentric anomaly. The eccentric anomaly is then related to time

through Kepler's equation

$$n_2(t - T_2) = u_2 - e_2 \sin u_2 \ , \tag{4.6}$$

where $n_2 = \sqrt{m_2/a_2^3}$ defines the mean motion, or mean angular frequency associated with an orbit. The mean motion is related to P_2, the outer period, $n_2 = 2\pi/P_2$, and lastly, the parameter T_2 is the time of pericenter passage, a constant of integration.

The desired velocity of the inner binary's center-of-mass is simply obtained by $\mathbf{v}_1 = (m_c/m_2)\mathbf{v}_2$, and differentiating equation (4.3)

$$\mathbf{v}_1 = \frac{m_c}{m_2}\sqrt{\frac{Gm_2}{p_2}}\,(-\sin\varphi_2, \cos\varphi_2 + e_2, 0) \ , \tag{4.7}$$

where $p_2 = a_2(1 - e_2^2)$ is the semi-latus rectum. Up to this moment we have been working in a coordinate system where the outer orbit defines the xy-plane. We must rotate our system to properly orient it into the coordinate system used by our detector model: the Solar System Barycenter frame. This may be accomplished through a series of Euler rotations: a rotation of $-\omega_2$, around the barycenter's z-axis, then by $-\iota_2$ around the new x-axis, and finally $-\Omega_2$ around the new z-axis, which are given by the following matrices:

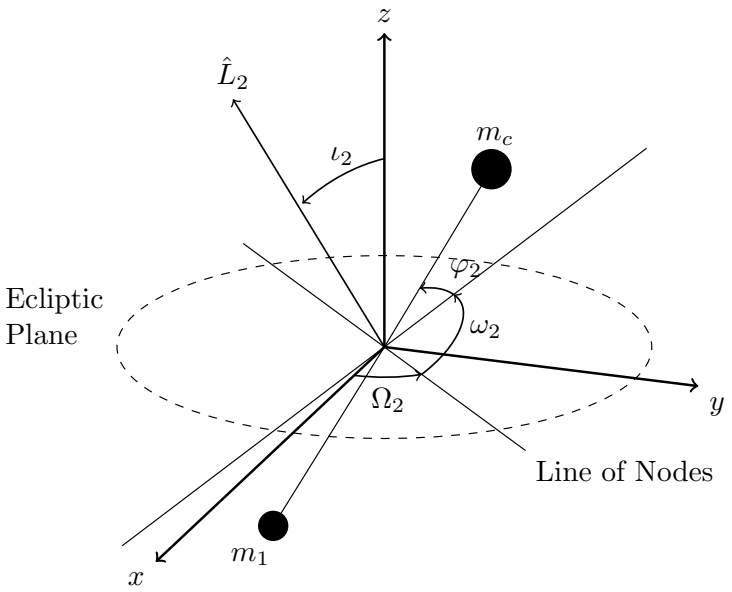

Figure 4.1: The geometry of the outer orbit consisting of m_c and the inner binary's monopole moment m_1 as displayed above, where the orientation angles are with respect to the solar system barycenter frame with the z-axis normal to the plane of the Ecliptic.

$$\mathbb{R}_1 = \begin{pmatrix} \cos\omega_2 & -\sin\omega_2 & 0 \\ \sin\omega_2 & \cos\omega_2 & 0 \\ 0 & 0 & 1 \end{pmatrix} , \tag{4.8}$$

$$\mathbb{R}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\iota_2 & -\sin\iota_2 \\ 0 & \sin\iota_2 & \cos\iota_2 \end{pmatrix} , \tag{4.9}$$

$$\mathbb{R}_3 = \begin{pmatrix} \cos\Omega_2 & -\sin\Omega_2 & 0 \\ \sin\Omega_2 & \cos\Omega_2 & 0 \\ 0 & 0 & 1 \end{pmatrix} , \tag{4.10}$$

operated in the order $\mathbb{R} = \mathbb{R}_3 \cdot \mathbb{R}_2 \cdot \mathbb{R}_1$. As shown in Figure 4.1 the line of ascending

nodes (labelled in the figure) is defined by a rotation of angle Ω_2 from the barycenter x-axis to where the outer orbital plane intersects the Ecliptic. The angle ω_2 defines the rotation angle from the line of nodes to the argument of periapsis (whose position is given by the solid line passing through the semi-major axis of the orbit), and ι_2 is the inclination angle, i.e. the angle between the outer orbit's angular momentum $\hat{L}_2$ and the z-axis of the barycenter coordinates (here neglecting any contribution to angular momentum due to the fact that the inner binary is extended and has an orbit of its own). Note that here that $\cos\iota_1 = \hat{n}\cdot\hat{L}_1$ defines the inner binary's inclination with respect to the line of sight.

Finally, we may construct the desired quantity: the line-of-sight velocity $v_\parallel$. We can use the line-of-sight vector $\hat{\mathbf{n}}$ pointing from the origin of the barycenter coordinates to the triple center-of-mass. Due to the large distances involved we will assume that the sky location (θ,ϕ) of the inner binary and of the triple's center-of-mass are located at the same point on the sky. This vector is given by $\hat{\mathbf{n}} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. At last we obtain the expression

$$
\begin{aligned}
v_\parallel(t) =& \hat{\mathbf{n}}\cdot\mathbb{R}\cdot\mathbf{v}_1 \ , && (4.11)\\
=& \hat{\mathbf{n}}\cdot\mathbb{R}\cdot\left(\frac{m_c}{m_2}\mathbf{v}_2\right) \ , \\
=& \frac{m_c}{m_2}\sqrt{\frac{Gm_2}{p_2}} \\
& \times\Big\{C(\theta,\iota_2,\phi-\Omega_2)\left[\cos(\varphi_2+\omega_2)+e_2\cos\omega_2\right] \\
& - S(\theta,\phi-\Omega_2)\left[\sin(\varphi_2+\omega_2)+e_2\sin\omega_2\right]\Big\} \ , && (4.12)
\end{aligned}
$$

where $C(\theta,\iota_2,\phi-\Omega_2) = \cos\theta\sin\iota_2 + \sin\theta\cos\iota_2\sin(\phi-\Omega_2)$ and $S(\theta,\phi-\Omega_2) = \sin\theta\cos(\phi-\Omega_2)$. In the above form it is unclear how many extra parameters are

truly involved in the modeling of the triple system, so we re-write the line-of-sight velocity in the simpler form

$$v_{\parallel}(t) = \mathcal{A}_2 \left[\sin(\varphi_2 + \varpi) + e_2 \sin(\varpi)\right] \ , \tag{4.13}$$

where $\mathcal{A}_2 = \frac{m_c}{m_2}\sqrt{\frac{m_2}{p_2}}\bar{A}$ and $\bar{A}^2 = C^2 + S^2$ and finally $\varpi = \omega + \bar{\phi}$ where $\tan\bar{\phi} = \frac{C}{-S}$. With the line of sight velocity written this way we can see what combination of parameters can be measured. To specify $v_{\parallel}(t)$ we needed the parameters $n_2, e_2, T_2, \iota_2, \omega_2, \Omega_2, m_c, m_2$ (note that the sky location angles are part of the binary model), but unfortunately we do not have access to all of these parameters due to degeneracies in the model which can be seen from equation (4.13). The parameters ω_2, Ω_2, and ι_2 get lumped into $\mathcal{A}_2$, and ϖ, leaving us in no position to parse the dynamically interesting ι_2 from other orientation angles. This amplitude has an average value of roughly 0.77 which will be used in the analysis contained in later sections. Also hidden in $\mathcal{A}_2$ are m_2, and m_c which we will not have access to individually. The orbital phase φ_2 contains the uninteresting parameter T_2 and is also controlled strongly by the mean motion n_2 and eccentricity e_2 parameters. We are now in a position to incorporate the line-of-sight velocity into the gravitational waveform. For eclipsing systems the Shapiro time delay can break some of the degeneracies and allow us to measure ι_2. We will leave the analysis of the gravitational wave Shapiro time delay to future work.

4.5 Gravitational Wave and Instrument Model

We will first briefly review the gravitational wave model for an isolated galactic binary as seen by LISA, and then incorporate the affects due to the companion body. These are low-mass binaries which are millions of years away from merging

and therefore we will only be capturing them in the inspiral phase. The plus and cross gravitational wave polarizations in the compact binary's barycenter frame are given by

$$h_+ = \frac{2\mathcal{M}}{D_{\rm L}} \left(\pi f_{\rm gw}(t)\right)^{2/3} \left(1 + \cos^2 \iota_1\right) \cos \Psi_{\rm gw} \ , \tag{4.14}$$

$$h_\times = -\frac{4\mathcal{M}}{D_{\rm L}} \left(\pi f_{\rm gw}(t)\right)^{2/3} \cos \iota_1 \sin \Psi_{\rm gw} \ , \tag{4.15}$$

where $D_{\rm L}$ is the luminosity distance, $\mathcal{M} = (m_a m_b)^{3/5}/m_1^{1/5}$ is the chirp mass, $f_{\rm gw}$ is the instantaneous gravitational wave frequency (as measured in the compact binary's barycenter frame), $\Psi_{\rm gw}$ the corresponding gravitational wave phase, and lastly ι_1 is the inclination of the inner binary i.e. $\cos \iota_1 = \hat{L}_1 \cdot \hat{n}$. One may obtain the gravitational wave phase from the frequency through $\Psi_{\rm gw} = 2\pi \int^t f_{\rm gw}(t')dt' + \phi_0$ where ϕ_0 is an arbitrary phase shift.

4.5.1 Numerical Implementation

For galactic binaries whose orbital evolution is dominated by gravitational wave radiation reaction, the frequency evolution is given by

$$f_{\rm gw}(t) = \frac{1}{8\pi\mathcal{M}} \left(\frac{5\mathcal{M}}{t_c - t}\right)^{3/8} \ , \tag{4.16}$$

where t_c is the time of coalescence for the binary; a 3 mHz, 0.265 $M_\odot$ galactic binary will merge in 1 million years. The number of $1/T_{\rm obs}$ frequency bins a fiduciary source

evolves through over the LISA mission lifetime is given by [169]

$$\dot{f}T_{\rm obs}^2 = 5.1 \left(\frac{\mathcal{M}}{0.32M_\odot}\right)^{5/3} \left(\frac{f}{5{\rm mHz}}\right)^{11/3} \left(\frac{T_{\rm obs}}{4{\rm yrs}}\right)^2 \tag{4.17}$$

$$\ddot{f}T_{\rm obs}^3 = 1.5 \times 10^{-4} \times \left(\frac{\mathcal{M}}{0.32M_\odot}\right)^{10/3} \left(\frac{f}{5{\rm mHz}}\right)^{19/3} \left(\frac{T_{\rm obs}}{4{\rm yrs}}\right)^3 . \tag{4.18}$$

The strong frequency dependence in these expressions implies that the higher frequency sources will have more measurable chirps. It is this frequency dependence that will allow us to determine the physics responsible for the evolution of a population of binaries. A similar order of magnitude frequency evolution is experienced by galactic binaries which involve stable mass transfer [170]. A key difference is that mass transfer tends to widen orbits leading to a frequency decrease over time. The mild evolution in gravitational wave frequency lends itself to a Taylor expansion:

$$f_{\rm gw} = f + \dot{f}t + \frac{1}{2}\ddot{f}t^2 , \tag{4.19}$$

whose coefficients are determined by the dynamics at play in the binary. We shall refer to f (and the equivalently red-shifted version during the triples discussion) as the carrier frequency.

Cornish & Littenberg [171] present a frequency domain model for galactic binaries measured by LISA. Under the rigid adiabatic approximation to the LISA motion one is able to perform a fast-slow decomposition, due to the slowly evolving amplitude (varying on timescales of a year mostly due to LISA's motion) and the fast varying phase due the larger carrier frequency (corresponding to orbital periods of minutes to hours for galactic binaries) of the waveform allowing a rapid evaluation of the waveform, specifically the Time-Delay Interferometry (TDI) variables X, Y, and

Z (where tildes over these variables will denote their Fourier transform). Through a linear superposition these data channels form noise-orthogonal variables A, E, and T [172] (of which only A and E will be sensitive to the triple systems of interest).

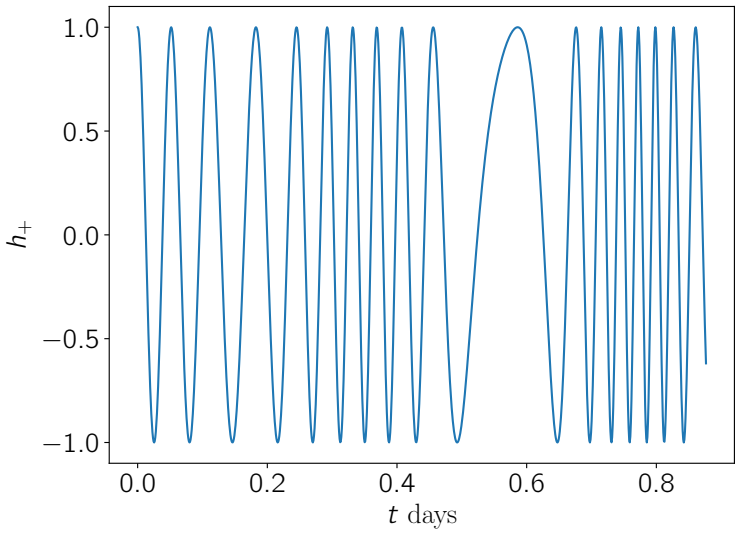

Figure 4.2: The gravitational waveform seen at the solar system barycenter for a system with outer period $P_2 = 1.1$ hours and the line-of-sight velocity amplitude 0.1. The carrier frequency f of the gravitational wave in this example is 5 mHz. The strain amplitude has been normalized to 1 such that the only time dependence crops up in the gravitational wave phase. The frequency of the oscillations is clearly changing.

The presence of a perturbing companion star m_c leads to an acceleration of the center-of-mass with respect to the barycenter frame, hence red-shifting the signal such that the gravitational wave phase gets modified

$$\Psi_{\rm gw} = 2\pi \int^t \left[1 + v_\parallel(t')\right] f_{\rm gw}(t') dt' + \phi_0 \ , \tag{4.20}$$

where $v_\parallel$ is the line-of-sight velocity obtained in the previous section[1]. In Figure 4.2

[1]Note that there should be an additional correction to Eq. (4.20) due to the fact that time in the two frames is related by $t \mapsto t + r_\parallel/c$. This correction is however negligible for the triple systems considered in this work, but it might be relevant for systems closer to coalescence, e.g. hierarchical triple black hole systems observable with LISA

the quantity h_+ (normalized to 1) is displayed for a circular triple system whose outer period was chosen to be very short to exaggerate the effects. The frequency oscillates around the carrier frequency f modulating the gravitational wave phase.

To accommodate these changes to the fast-slow waveform code [171] must be modified. To properly calculate the gravitational wave transfer function one must evaluate it at the gravitational wave frequency observed by the LISA detectors. The Taylor expanded frequency evolution (as in equation (4.19)) is redshifted with respect to the solar system barycenter i.e. $f_{\rm gw} \to (1 + v_\parallel)\, f_{\rm gw}$. The line-of-sight velocity is numerically obtained through equation (4.13) and the inversion of Kepler's equation. For the isolated galactic binaries the gravitational phase may be easily integrated. When this binary resides in a triple system an extra term in the gravitational wave phase integral crops up $2\pi \int v_\parallel f_{\rm gw} dt$ which is numerically integrated, interpolated at the detector sampling intervals, and then appended to the isolated galactic binary gravitational wave phase. These modifications to the gravitational wave frequency get applied to the slow portion of the waveform model, which is sampled at cadence much longer than the orbital period.

The log likelihood function used in our analysis involves noise-weighted inner products of the form

$$(g|k) = 4\mathcal{R} \int_0^\infty \frac{\tilde{g}^*(f)\tilde{k}(f)}{S_n(f)} df \ , \tag{4.21}$$

where g and k are arbitrary waveforms as seen by LISA, and $S_n(f)$ is the one-sided noise power spectral density. Further discussion of this quantity and the noise model for LISA, including both instrumental noise and unresolved galactic binary confusion noise, can be found in references [70, 173, 174]. The signal-to-noise ratio (SNR) ρ is defined as $\rho^2 = (h|h)$ for a given waveform h.

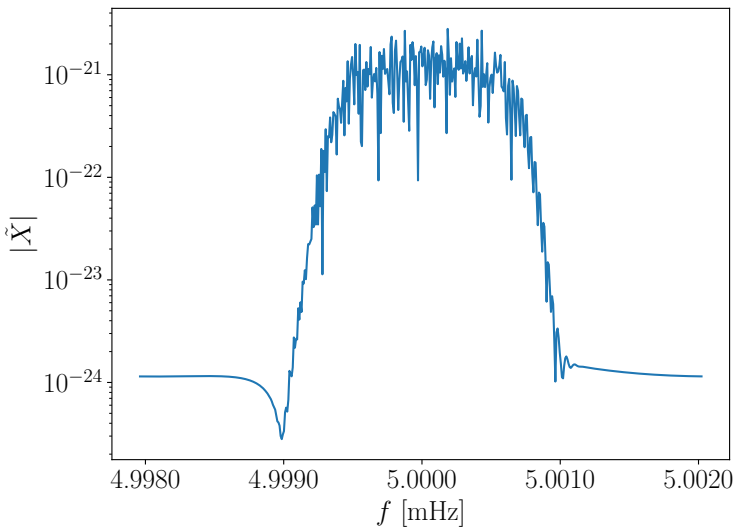

Figure 4.3: The X TDI channel is displayed for a triple system with an outer orbital period $P_2 = 1.5$ years and $e_2 = 0.3$. The presence of the perturbing companion induces harmonics of the carrier frequency and of the harmonics present due to LISA's modulations. The amplitude (therefore distance) of the system was modified to give an SNR of 20. The other parameters of this system were chosen as follows: $\dot{f} = 1.11 \times 10^{-15}$, $\theta = 1.52$, $\phi = 4.577$, $\phi_0 = 0.346$, and $\psi = 1.58$. The masses were chosen as $m_a = 0.6\ M_\odot$, $m_c = 1.0\ M_\odot$, and $\mathcal{M} = 0.32\ M_\odot$. The observation period for this figure was 4 years.

Examples of the frequency domain strain amplitude can be found in Figures 4.3 and 4.4. Both of these waveforms were generated for inner binaries with $f = 5$ mHz, and a chirp mass of $0.32 M_{\odot}$ (which fixes the source frame frequency evolution as determined by General Relativity) for a 4 year observation period at a 15 second cadence. In Figure 4.3 the outer orbit revolves every 1.5 years and has an eccentricity of 0.3. An isolated binary is nearly monochromatic, resulting in a near delta function in the frequency domain, but due to the modulations caused by the cartwheel motion of the LISA observatory around the Sun, picks up side-bands whose phase and relative amplitude are determined by the sky location and gravitational wave polarization of the binary. The introduction of a perturbing third body generates more harmonics of the frequencies already present, and tends to increase the bandwidth of the signal. Increasing the eccentricity of the outer orbit shifts the distribution of power into higher modes of the triple harmonics.

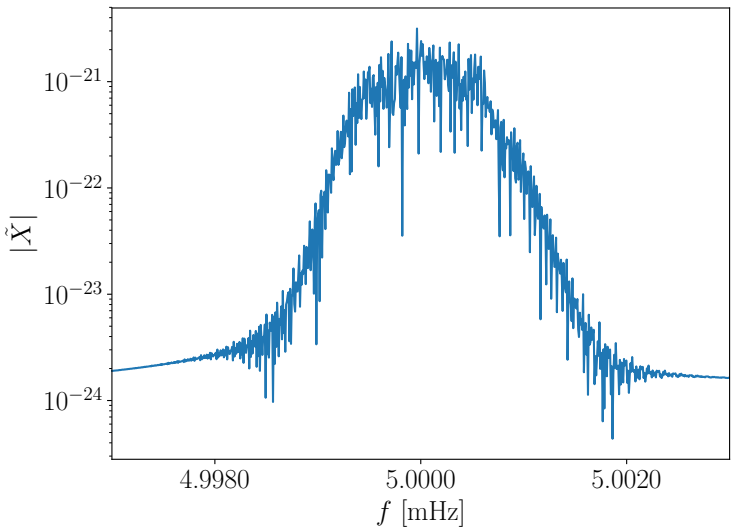

Figure 4.4: The X TDI channel is displayed for triple system with an outer orbital period $P_2 = 0.6$ years and $e_2 = 0.7$. In this example the harmonics induced by the companion star and LISA are interfering. Eccentricity in the outer orbit changes the distribution of power in the triple induced harmonics. The other parameters were chosen to be the same as in Fig. 4.3.

In Figure 4.4 the system has a tighter outer period of 0.6 years, and a larger eccentricity of 0.7. Here the orbital period P_2 is comparable to the orbital timescale of LISA, leading to a strong interference of harmonics. The side-bands of the carrier frequency induce by the triple are now more widely spaced than those imparted by LISA orbit, leading to a much broader signal, which is amplified by the larger eccentricity. If one were to consider a system of even shorter period then the triple induced harmonics separate out into isolated side-bands.

One can gain insight visually into our ability to distinguish a triple system from a binary by considering Fig. 4.5. In both panels the blue lines represent the power in the X TDI channel for an isolated binary signal and orange represents the same data channel for a triple system where the parameters characterizing the inner binary are the same for all signals displayed in this figure. The orange lines are therefore signals of the same system in both panels. The blue lines differ from each other purely in their outer orbital period. These signals are the result of a 4 year observation period. In the regime where the outer period is much larger than the observation period we are still able to measure the frequency evolution of the signal. This is seen in the upper panel of Fig. 4.5 where the outer orbital period was set to 100 years. Note how the red-shifting slightly shifts the frequency of the triple signal to the right. For this very large outer period one might in fact mistake this signal for an isolated binary with slightly different parameters than those in the actual inner binary. In the lower panel we see a more sever difference since the outer orbital period was set to 3.6 years, comparable to the observation period. There is a severe broadening of the signal and the power structure is vastly different than that seen in the upper panel. This difference allows us to detect the presence of the triple and characterize the parameters of the outer orbit which will be demonstrated later.

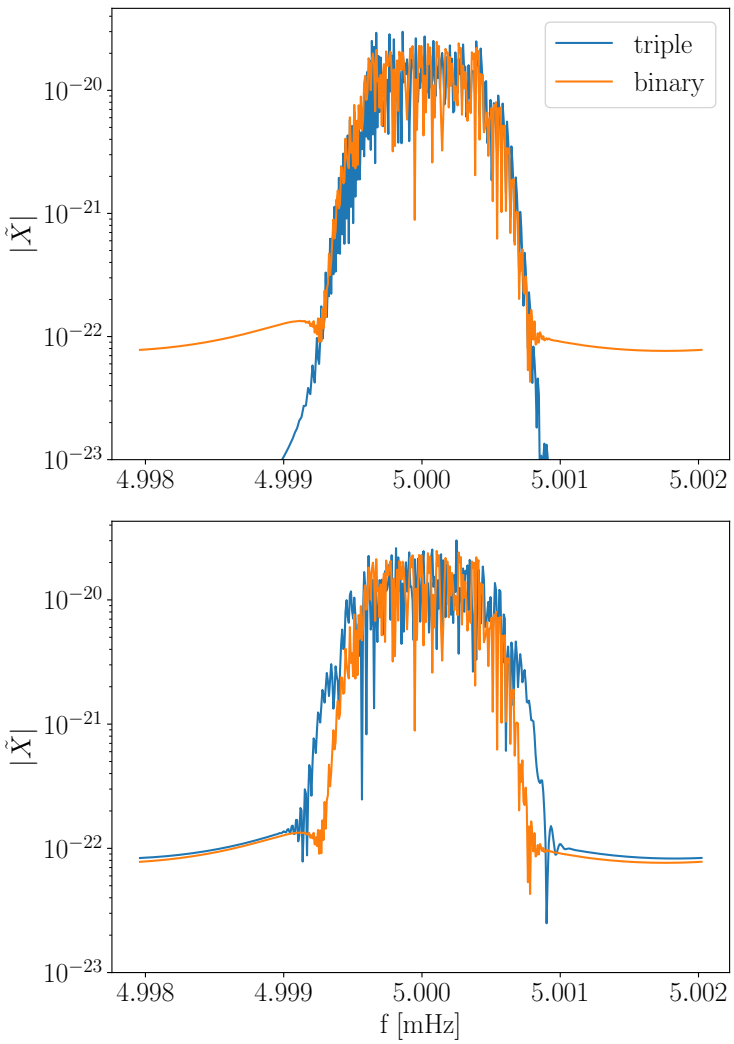

Figure 4.5: The TDI X-channel response for isolated binary signals—denoted in orange—and for triple signals—colored blue. The SNR of these signals were chosen to be 20 and were observed for 4 years. The outer eccentricity of the triple signals was set to 0.3. The outer period in the upper figure is 100 years (much larger than observation period) and the lower figure 3.6 years (comparable to observation period). The other parameters are identical to those chosen for Fig. 3.

4.5.2 Eccentric Inner Binary Signal

In order to demonstrate that presence of gravitational wave harmonics due to an eccentric inner binary, $e_1 \neq 0$, do not affect the characterization of the outer orbit, we extend the waveform model to cover inner binaries with moderate eccentricity.

The gravitational wave plus and cross polarizations at Newtonian order for eccentric binaries are given by

$$h_{+,\times} = -\frac{m_1\eta_1}{D_L}(2\pi m_1 F_1)^{2/3}\sum_{j=1}^{\infty}\Big[C^{(j)}_{+,\times}(e_1,\beta_1,\iota_1)\cos jl_1 + S^{(j)}_{+,\times}(e_1,\beta_1,\iota_1)\sin jl_1\Big] \ , \tag{4.22}$$

where $\eta_1 = m_a m_b/m_1^2$ is the symmetric mass ratio, F_1 is the mean orbital frequency, β_1 defines the pericenter angle, and ι_1 is the inclination angle of the inner binary. The sum is over harmonics of the gravitational wave signal, and the coefficients $C^{(j)}_{+,\times}$ and $S^{(j)}_{+,\times}$ are the amplitudes of these harmonics. Their functional form is given in Moore et al [175].

For mild eccentricities the second harmonic is dominant and the frequency associated with it is Taylor expanded as in Equ (4.19)

$$f^{j=2}_{\rm gw} = f + \dot{f}t + \frac{1}{2}\ddot{f}t^2 \ . \tag{4.23}$$

In order to calculate the other harmonics we make use of the Cornish & Littenberg galactic binary waveform generator but with the following harmonic dependent conditions: 1) the gravitational wave frequency seen gets modified by $f^{j}_{\rm gw} \to \frac{j}{2}f^{j=2}_{\rm gw}$, and 2) similarly, the phase must be modified $\Psi^{j}_{\rm gw} \to \frac{j}{2}\Psi^{j=2}_{\rm gw}$. The amplitudes of the TDI variables were then modulated by the harmonic coefficients. For these eccentric waveform we keep the first 4 harmonics in our analysis which is enough to maintain 99% of the gravitational wave power up to $e_1 = 0.2$ (as determined by the harmonic coefficients). There is an important feature to keep in mind about the structure of the gravitational wave signal. The harmonics are separated by $\Delta f = f_1$ Hz which

for for a 1.5 mHz orbital frequency corresponds to 47,000 frequency bins for a 1 year observation period! The bandwidth for these harmonics is a most a couple thousand bins, which tells us that the harmonics are well separated and do not interfere with each other. Nor do the harmonics due to the eccentricity overlap in any way with the harmonics imparted by a hierarchical companion. This is why we are justified in neglecting any eccentricity of the inner binary and using the simpler circular orbit model. This argument is confirmed by simulations in §4.7.

4.6 Detecting Hierarchical Companions

When LISA first detects a triple only the intrinsic gravitational wave frequency of the inner binary will be measurable. As more cycles are accumulated, and the center-of-mass of the inner binary has moved through a significant portion of the outer orbit, the data will support the inclusion of orbitally induced redshifts. We will now estimate when we expect the frequency evolution to be measurable, i.e. for a given source and observation period, and an average oriented source, what P_2's will we be able to detect with the effect of this center-of-mass motion?

From the gravitational wave phase quoted in equation (4.20) it is straightforward to obtain the time derivative of the frequency in the barycenter frame for a binary in a triple system which has negligible source frame frequency evolution

$$\dot{f} = a_{1,\parallel} f \ , \tag{4.24}$$

where $a_{\parallel}$ is the line-of-sight acceleration of the inner binary's center-of-mass. This can obtained by differentiating equation (4.7)

$$\mathbf{a}_1 = -\frac{m_c}{p_2^2} \left(1 + e_2 \cos\varphi_2\right)^2 \left(\cos\varphi_2, \sin\varphi_2, 0\right) \ , \tag{4.25}$$

applying the rotation matrices once more and projecting along the line-of-sight gives

$$a_{1,\parallel} = -\frac{m_c}{p_2^2}\left(1+e_2\cos\varphi_2\right)^2 \times \left[S\cos(\omega_2+\varphi_2)+C\sin(\omega_2+\varphi_2)\right] \ , \tag{4.26}$$

where S and C are defined as before; Cf. equation (4.12). We would like to consider the average magnitude of this acceleration for a given orientation and sky location. We may square this quantity and then average it over the angles ϕ, θ, ω_2, and ι_2

$$\begin{aligned}\left\langle a_{1,\parallel}^2 \right\rangle =& \frac{1}{(4\pi)^2}\int_0^{2\pi} d\phi \int_{-1}^{1} d(\cos\theta) \\ & \times \int_0^{2\pi} d\omega_2 \int_{-1}^{1} d(\cos\iota_2) a_{1,\parallel}^2 \qquad (4.27)\\ =& \frac{m_c^2}{3p_2^4}\left(1+e_2\cos\varphi_2\right)^4 \ . \qquad (4.28)\end{aligned}$$

To calculate the RMS acceleration we average the previous result over the course of an orbit

$$\begin{aligned} a_{\parallel\,\mathrm{RMS}}^2 &= \frac{1}{P_2}\int_0^{P_2} \left\langle a_{1,\parallel}^2 \right\rangle dt & (4.29)\\ &= \frac{1}{P_2}\int_0^{P_2} \left\langle a_{1,\parallel}^2 \right\rangle \dot{\varphi}_2^{-1} d\varphi_2 & (4.30)\\ &= \frac{m_c^2}{3m_2^{4/3}}\left(\frac{2\pi}{P_2}\right)^{8/3} \frac{1+\frac{1}{2}e_2^2}{\left(1-e_2^2\right)^{5/2}} \ . & (4.31)\end{aligned}$$

In the regime that $P_2 > T_{\mathrm{obs}}$ we may Taylor expand the gravitational wave frequency. Equation (4.24), when averaged over angles and over an orbit, provides

us with a rough estimate of the size of $\dot{f}$ for an average outer orbit orientation which started at an average spot in its orbit when LISA began to collect data. With this we can ascertain how many frequency bins this $\dot{f}$ estimate will evolve the carrier frequency through

$$\dot{f}T_{\rm obs}^2 = 573 \left(\frac{P_2}{1\ {\rm yr}}\right)^{-4/3} \left(\frac{m_c}{1M_\odot}\right) \left(\frac{m_2}{2M_\odot}\right)^{-2/3} \times \left(\frac{T_{\rm obs}}{4\ {\rm yr}}\right)^2 \left(\frac{f}{5\ {\rm mHz}}\right) \sqrt{\frac{1+\frac{1}{2}e_2^2}{(1-e_2^2)^{5/2}}}\ . \tag{4.32}$$

In order to ascertain when this effect is measurable we utilize Fisher matrix estimates for the error in measurement of $\dot{f}$. The Fisher matrix, by the Cramer-Rao bound, provides an estimate of the covariance matrix (upon inversion of the Fisher matrix), thereby providing error estimates. The Fisher matrix is defined as

$$\mathbf{\Gamma}_{ij} = (h_{,i}|h_{,j})\ , \tag{4.33}$$

where $h_{,i}$ are derivatives of the waveform with respect to parameter λ^i and then evaluated at the true parameters. For a triple signal whose outer period is larger than LISA's observation period we may readily approximate the frequency evolution of the system by a Taylor expansion as we would for a mildly chirping isolated binary. This allows us to utilize the fast galactic binary waveform to calculate the Fisher matrix.

Seto [169] used a simple toy model for a Fisher matrix analysis to estimate the measurement errors in some of the galactic binary parameters. In appendix A we expand upon these results and investigate how the errors get inflated by including more parameters through the use of the full galactic binary model. We find that the $\dot{f}$ and $\ddot{f}$ errors become inflated through the inclusion of the full set of galactic binary

parameters. The criterion which we use to determine whether $\dot{f}$ is a measurable parameter is that $\dot{f}$ must be larger than 3σ (as estimated by the Fisher matrix) compared to no frequency evolution at all. This yields the expression quoted in equation (4.1) in the introduction, which we repeat here for completeness:

$$P_2 \lesssim 43.2 \text{ yrs} \left(\frac{\rho}{10} \cdot \frac{m_c}{1.0 M_\odot} \cdot \frac{f}{5 \text{ mHz}}\right)^{3/4} \left(\frac{m_2}{2M_\odot}\right)^{-1/2} \times \left(\frac{T_{\text{obs}}}{4 \text{ yr}}\right)^{3/8} \left(\frac{1 + \frac{1}{2}e_2^2}{(1-e_2^2)^{5/2}}\right)^{3/8} . \tag{4.34}$$

A fiducial source with an outer orbital period of 40 years would have a measurable frequency evolution by the time the nominal LISA mission concluded. When the outer eccentricity reaches $e_2 = 0.7$,the outer periods is measurable out to $P_2 = 110$ years for typical systems. We see that for larger companion masses the (assuming fixed total mass m_2), the larger the outer period we can measure. Thus we see an increase in the detectability of triples with large companion masses. The opposite is true as we increase the total mass.

We may make similar applications of the Fisher analysis to ascertain when the gravitational wave carrier frequency becomes biased (i.e. differs from the source frame value in a measurable way) for a given P_2. The RMS line-of-sight velocity is given by

$$v^2_{\parallel \text{ RMS}} = \frac{m_c^2}{3m_2^{4/3}} \left(\frac{2\pi}{P_2}\right)^{2/3} , \tag{4.35}$$

such that when

$$P_2 \lesssim 71.8 \text{ yrs} \left(\frac{\rho}{10} \cdot \frac{m_c}{1.0 M_\odot} \cdot \frac{f}{5 \text{ mHz}}\right)^3 \times \left(\frac{m_2}{2M_\odot}\right)^{-2} \left(\frac{T_{\rm obs}}{4 \text{ yrs}}\right)^3 , \tag{4.36}$$

our measurements of the carrier frequency f will be biased. This is potentially the most concerning result if one is interested in the orbital period distribution of the galactic binaries, as for sources which only have f measured, this yields a quite large range of outer orbital periods which could bias the frequency measurement.

Another question of interest is when the parameter $\ddot{f}$ is measurable (recall that here we are only considering the frequency evolution coming from the center-of-mass motion). Upon measuring f, $\dot{f}$, and $\ddot{f}$ we have the best chance of determining the underlying physics for mildly evolving sources. The RMS jerk is given by

$$\dot{a}^2_{\parallel \text{ RMS}} = \frac{m_c^2}{3 m_2^{4/3}} \left(\frac{2\pi}{P_2}\right)^{14/3} \frac{1 + \frac{19}{2} e_2^2 + \frac{69}{8} e_2^4 + \frac{9}{16} e_2^6}{(1 - e_2^2)^{11/2}} . \tag{4.37}$$

which for fiduciary values becomes measurable when

$$P_2 \lesssim 16.7 \text{ yrs} \left(\frac{\rho}{10} \cdot \frac{m_c}{1.0 M_\odot} \cdot \frac{f}{5 \text{ mHz}}\right)^{3/7} \times \left(\frac{m_2}{2M_\odot}\right)^{-2/7} \left(\frac{T_{\rm obs}}{4 \text{ yrs}}\right)^{3/7} . \tag{4.38}$$

To verify the validity of the preceding results, based off a Fisher matrix analysis, we now spot check the measurability of frequency evolution of a triple system using Markov Chain Monte Carlo (MCMC) simulations. Simulated data was produced for a triple system, and analyzed using the Taylor expanded frequency evolution model.

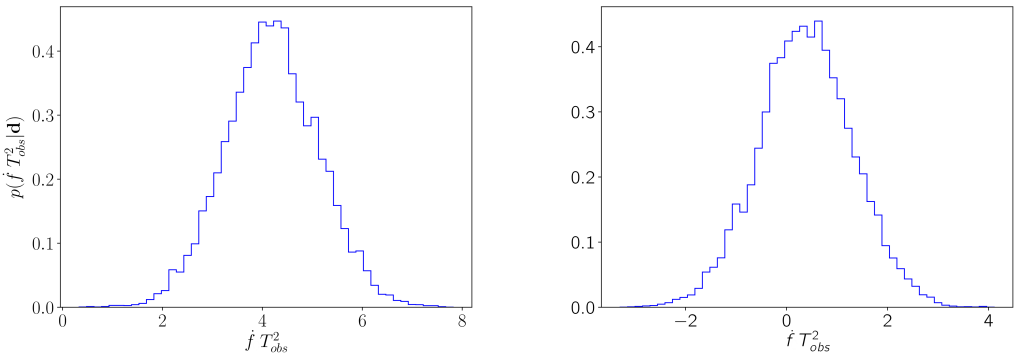

Figure 4.6: Here we display the posteriors for the parameter $\dot{f}T^2_{\rm obs}$ marginalized over all other system parameters. The simulated systems had a total mass of $m_2 = 1.77M_\odot$, with $\mathcal{M} = 0.32M_\odot$, $\rho = 20$, and an observation period of 1 year. The time of pericenter passage for the left hand figure was set to 0 while for the right hand side T_2 was set to $-P_2/8$ i.e. an eighth of an orbit. The remaining parameters $\dot{f}$, θ, ϕ, ψ, and ϕ_0 have been set to those chosen in Fig. 3.

The MCMC consisted of a burn-in phase such that the galactic binary model could search through parameter space to identify a regime in which the triple signal was well described by the binary model. A mixture of Fisher matrix proposal, differential evolution proposals and draws from the prior distribution were utilized to explore the posterior distribution [176,177]. Since we have developed a quick numerical model to generate the signal from these triple systems the proposal distributions may choose any parameters randomly and we can generate a model for those parameters on the spot. Parallel tempering was also used to ensure a wide exploration of parameter space and to move between secondary modes of the posterior.

In Figure 4.6 the posteriors for the parameter $\dot{f}$ (marginalized over all other parameters) are displayed for two triple systems. The outer period was chosen to be 46 years i.e. the value obtained from the relation (4.1) using the modified triple parameters. The errors predicted by the Fisher matrix for $\dot{f}$ are a bit smaller compared to the error measured by the MCMC, suggesting that we might

be marginally overestimating the outer periods we can confidently measure. The difference between these two posteriors is the time of pericenter passage T_2 which differed by an eighth of an orbit between the two systems. This demonstrates that it is very important where we catch the triple in its orbit when LISA turns on, as the measurability of $\dot{f}$ is quite sensitive to T_2. This is especially important point to consider for larger outer period sources. Here we have seen that the Fisher analysis has roughly identified the regime in which we may hope to identify the presence of a triple system depending on where in the orbit we are measuring the gravitational wave signal.

4.7 Characterizing the Hierarchical Orbit

Now that we have ascertained when the effects of a triple system are detectable we would like to know when the parameters of the triple orbit can be measured. To determine this we utilized the Fisher information matrix for the triple signal. The criterion that we use to determine if a parameter is measurable is as follows: if the error in a parameter, as estimated by the Fisher matrix, is less than 50% of its true value then we claim this parameter can be measured. For triple systems the best measured parameter pertaining to the outer orbit is the outer orbital period, and if this quantity can be measured we say that the triple can be characterized (at least to some level).

In Figure 4.7 we display the results of the Fisher matrix based analysis. Systems with carrier frequencies and outer periods in the shaded region have orbits whose parameters *cannot* be measured. To determine the separating line we construct a system with a given carrier frequency f and a very short outer period P_2 and estimate its error with a Fisher analysis. The outer orbital is gradually made larger until its effects on the gravitational signal are marginal such that its error breaches 50%. The

P_2 at which this happens defines the border in Figures 4.7 and 4.8. We see that as the carrier frequency gets larger the outer period can be measured. This is due to this being a redshift phenomenon where the deviations in the frequency observed by LISA are proportional to the frequency itself, coupled with the fact that the error in the frequency is independent (to leading order, see appendix A) of the frequency itself.

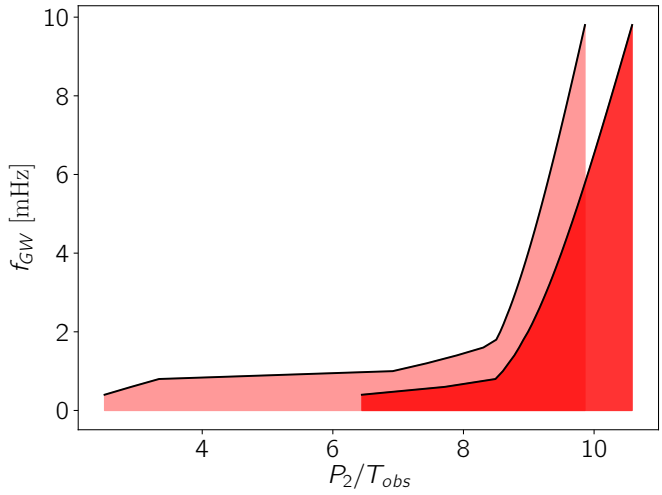

Figure 4.7: The shaded regions denote triple systems which cannot have their outer orbital parameters measured. In other words, systems to the left of the shaded regions can have at least the outer orbital period determined. Note the the orbits period P_2 is given in units of the observation time $T_{\rm obs}$, which could be as large as 10 years. The left black line is for a SNR 20 system and the right line is for SNR 100. This system had the parameters $m_2 = 2.0M_\odot$, $m_c = 1.0M_\odot$, $m_a = 0.5M_\odot$, and $\mathcal{M} = 0.32M_\odot$. The triple systems here had ciruclar outer orbits. The remaining parameters have been set to the same values used to produce Fig. 3.

Figure 4.8 reveals the effect that eccentricity of the outer orbit has on the characterization of the triple parameters. Typically, for larger f, increasing the eccentricity allows one measure orbital periods that are larger than for the circular case. It is important to note that with such large orbits (in fact, any time when

$P_2 > T_{\rm obs}$) these results will depend on where we captured the triple in its orbits. For the systems considered here we chose $\varpi = 0$, and $T_2 = 0$. This Fisher analysis demonstrates that we will be able to characterize the parameters for triple systems whose orbital period is up to 10 times that of the LISA mission lifetime, though the details get slightly modified by the other parameters and SNR.

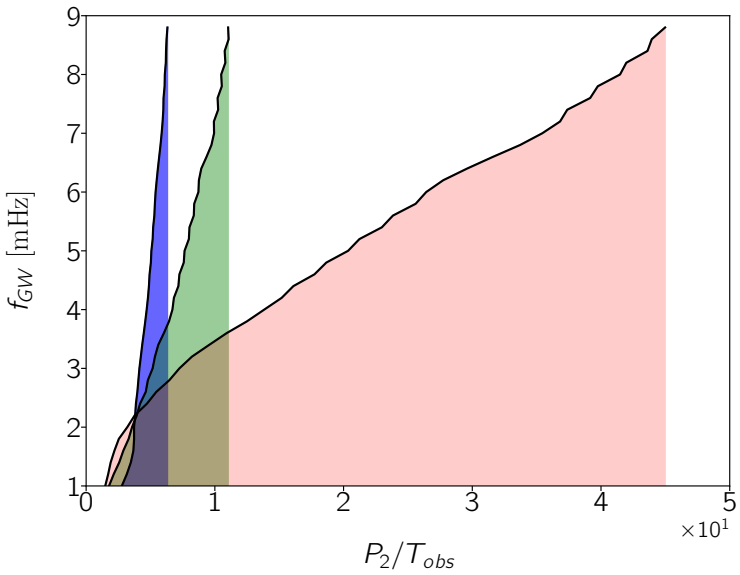

Figure 4.8: Just as in the previous figure the shaded regions denote triple system which cannot be characterized. Note the the orbits period P_2 is given in units of the observation time $T_{\rm obs}$,. These figures are different in that the outer orbits are now eccentric. The red region denotes $e_2 = 0.1$, blue $e_2 = 0.4$, and green $e_2 = 0.9$ These systems had an SNR of 10. We see that more eccentric outer orbits may allow for the characterization of triple systems with much larger outer orbital periods. The mass parameters were chosen to be these same as the previous figure.

Let us now address how well the parameters of the triple system can be measured. The period and eccentricity of the outer orbit has the largest effect on the gravitational wave signal, and are therefore the most readily measured quantities. It is instructive to consider the strong parallels with the pulsar timing case. The analogy is clear; pulsars in a binary emit pulses at a very regular rate, with mild frequency evolution, and the arrival of these pulses gets modulated by Earth's motion and the presence of a

companion. However, for pulsar timing the source is well localized on the sky, whereas the sky localization is generally poor for galactic binaries detected by LISA [178]. Another parameter that is well measured in pulsar timing is $(m_{\rm b}\sin\iota_2)^3/m_{\rm total}^2$, but it is only with the measurement of a Shapiro time delay for eclipsing binaries which allows the masses and inclination to be untangled. An additional effect, which will be negligible for the triples we are considering, is the variation in path length of light which allows the longitude of the ascending node Ω_2 to be measured.

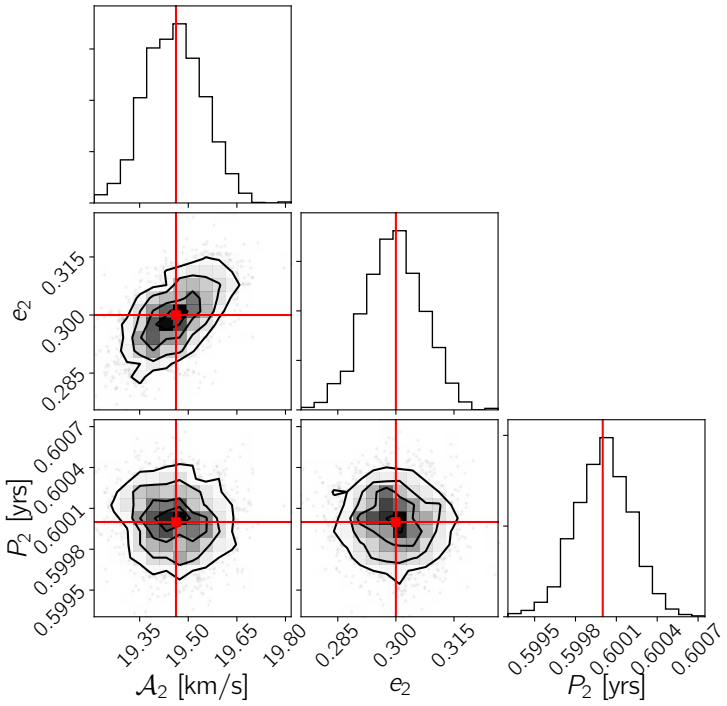

Figure 4.9: This is a corner plot of the posteriors for the parameters $\mathcal{A}_2$, e_2, and P_2. The one dimensional histograms are posteriors marginalized over all other parameters, and the other histograms are joint posteriors between pairs of the aforementioned parameters, marginalized over the rest of the triple system parameters. The line-of-sight velocity amplitude is 19.5 km/sec, its eccentricity 0.3, and outer period 0.6 years. The remaining parameters were chosen to be the same as those chosen in Fig. 3.

Figures 4.9 and 4.10 are the results of MCMC of triple systems with an SNR

of 50 and a range of outer orbital periods and eccentricities. The injected values are marked by red lines and dots in these figures. As expected the outer period and eccentricity are well measured for both systems. For both of these systems the line-of-sight amplitude $\mathcal{A}_2$ is also well measured, but as discussed earlier, on its own not terribly informative which means that the companion mass cannot be determined. One sees that $\mathcal{A}_2$ and e_2 are correlated, which gets amplified in the more eccentric case. The fact that both of these parameters influence the amplitude of the harmonics induced by the triple is responsible for this correlation. In Figure 4.11 marginalized posteriors for the parameters ϖ and T_2 are displayed for the more eccentric system. We see that these quantities are well measured, but they are of little physical interest.

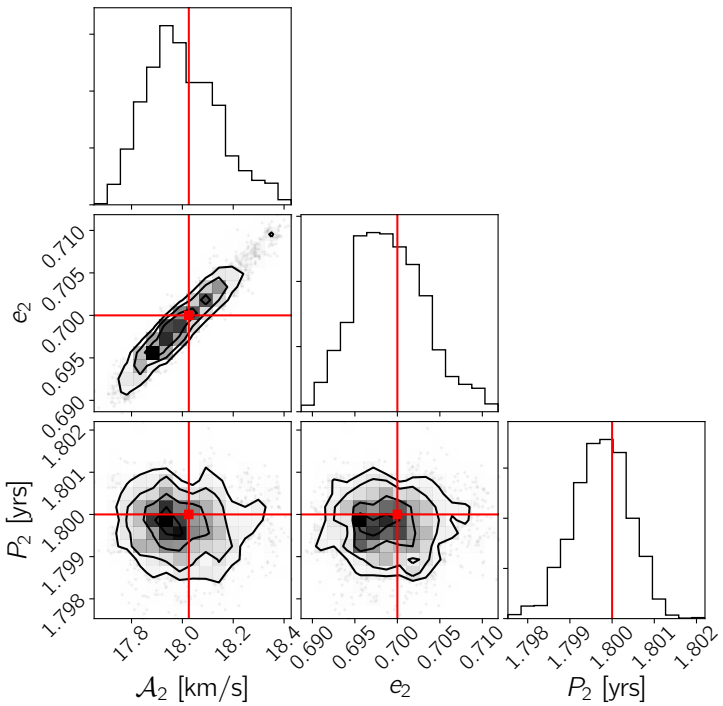

Figure 4.10: This corner plot displays some marginalized posteriors for a more eccentric system with a larger outer period. The line-of-sight velocity amplitude is 18.0 km/sec, its eccentricity 0.7, and outer period 1.8 years. The other parameters were chosen to be the same as the previous corner plot figure.

Typically, as P_2 increases, such that less orbits are captured by LISA, the worse the parameters are characterized. In figure 4.10 the eccentricity has a standard deviation of 0.6% relative to e_2 and the outer period an standard deviation of 0.037% or 14.6 hours. The tighter system with $P_2 = 0.6$ years had its outer orbital period determine to a standard deviation of 4.5 hours. However, this does not seem to hold steadfast for the measurement of eccentricity. The tighter system had a relative standard deviation of 2.5% i.e. larger than the system with a wider orbit. This exception occurs as the outer orbital period starts to encroach upon the LISA modulation frequency (1 year). The distribution of power in the higher modes of the carrier frequency, induced by the triple, get shifted as e_2 changes. These harmonics, when their fundamental frequency $1/P_2$ is comparable to the LISA modulation frequency begin to interfere strongly making it harder to accurately extract the eccentricity.

By considering the posteriors in presented in this section we see that the outer orbital period and outer eccentricity are well measured parameters.

Lastly, we would like to address the effect of eccentricity in the inner orbit on the characterization of the outer orbital parameters. To do this we injected a signal with $e_1 = 0$ and SNR of 20, $e_2 = 0.3$, and $P_2 = 0.6$ years. Again, all other parameters were set to those for Fig. 3. An MCMC was used to calculate the marginalized posterior for the outer orbital parameters denoted by the solid-blue line in the upper panels of Fig 4.12. Next a signal was simulated where the inner binary was eccentric; the parameters were set exactly the same as before but $e_1 = 0.1$ and $\beta_1 = 0$. The resulting marginalized posteriors are displayed by the dashed-red line in the upper panels of Fig 4.12. As an added bonus we obtain the marginalized posterior for the inner eccentricity displayed in the lower left panel of the same figure. This eccentricity is measurable since for this system as evidenced by there being zero posterior weight

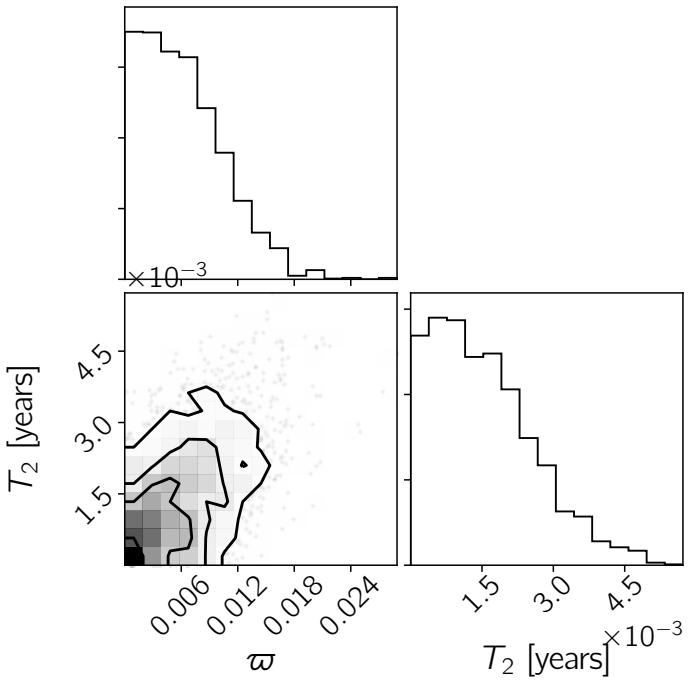

Figure 4.11: These are marginalized posteriors for the variables ϖ and T_2 for the system represented by Fig. 9. Both ϖ and T_2 were set to 0 for this system.

at $e_1 = 0$.

Comparing the posteriors for the important and measurable outer binary parameters e_2 and P_2 we see that they are essentially the same distribution. More specifically, the posteriors for the inner eccentric case lie well within the error associated with the posteriors for the inner circular case. Thus we can conclude that the presence of harmonics in the gravitational waveform due to the existence of eccentricity in the inner binary does not affect the characterization of the outer orbital parameters. The evidence for this is also seen by the lack of any correlation between the outer orbital parameters and e_1. A scatter plot of e_1 and e_2 is displayed in the lower right panel of Fig 4.12, and no correlation is evident. This results from the clean separation of harmonics—i.e. the fact that the bandwidth of each harmonic

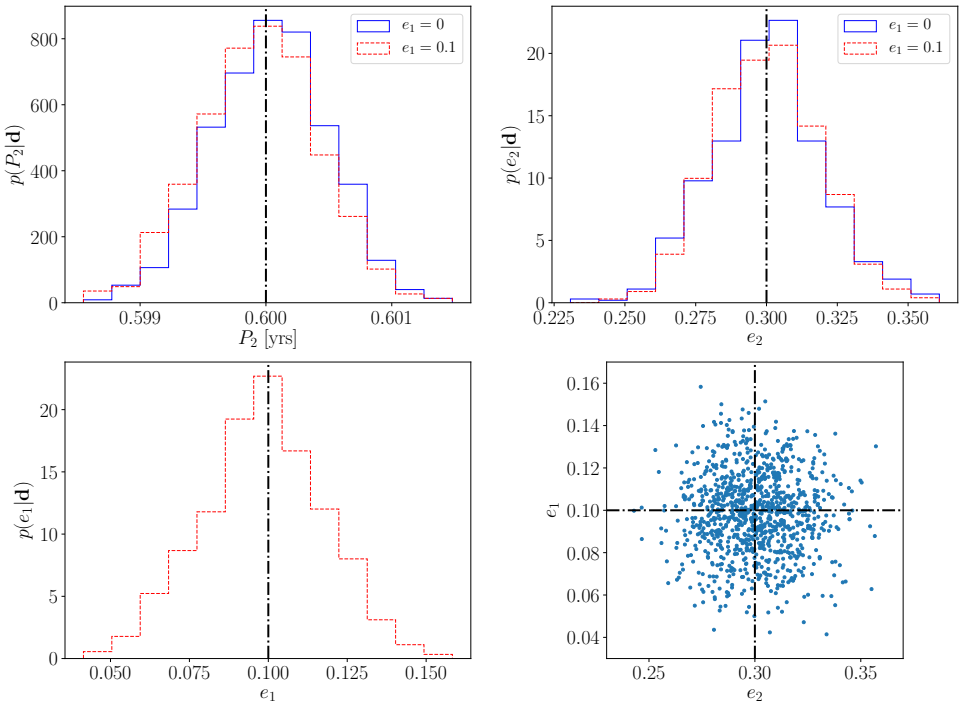

Figure 4.12: The solid-blue lines denote marginalized posteriors for outer orbital parameters for a circular inner binary injection in the upper panels. The red-dashed lines represent the marginalized posteriors for the same parameter for an eccentric inner binary injections. The vertical black lines denote the injected parameter values. The marginalized posteriors in the upper panel are statistically equivalent for each parameter. The lower left panel displays the marginalized posterior for the inner eccentricity. The lower right panel is a scatter plot of the e_1 and e_2 samples from the MCMC and show no correlation.

in frequency is much less than their spacing in frequency. Thus the results we have derived for binaries with circular inner orbits apply unchanged to systems where the inner orbit is eccentric.

4.8 Ambiguous Systems

Assuming a nominal 4 year mission lifetime, it has been estimated that frequency evolution due to gravitational wave emission or mass transfer will be measurable for roughly 9000 isolated galactic binaries [70]. It is interesting to consider if a regime exists where the orbital acceleration due to hierarchical companions may be confused with these effects. The chance of confusion is greatest when only f and $\dot{f}$ are measurable. In most cases, a measurement of $\ddot{f}$ will break the degeneracy. To determine the risk of confusion consider Figure 4.13, which compares the frequency evolution for an isolated binary and a binary in a hierarchical system. The frequency range over which the effects might be confused is very small since the frequency evolution scales very differently: $\dot{f} \propto f$ from the hierarchical orbit (see equation (4.24)) and $\dot{f} \propto f^{11/3}$ for mass transfer and gravitational wave emission. We see that for an outer period of 1 year there is no chance of confusion for this system. Even up to outer periods of 10 years the amount of overlap is small. The system with an outer period of 30 years, which is approaching the largest period for which there is a measurable $\dot{f}$, has the greatest potential for confusion. The larger the gravitational wave frequency the less likely it is that the effects will be confused.

We now directly test how well a binary signal can reproduce a triple signal. To do so we inject a triple system into the LISA data stream and perform an MCMC with simulated annealing utilizing a galactic binary waveform model. The simulated annealing cools down the MCMC such that the chain settles into the peak of the posterior, thus allowing us to find the best values for the parameters as suggested by the data. The maximum posterior signal allows us to calculate the fitting factor

$$FF = \max_{\boldsymbol{\lambda}} \frac{(h_{\mathrm{T}}|h(\boldsymbol{\lambda}))}{\sqrt{(h_{\mathrm{T}}|h_{\mathrm{T}})\,(h(\boldsymbol{\lambda})|h(\boldsymbol{\lambda}))}} \ , \tag{4.39}$$

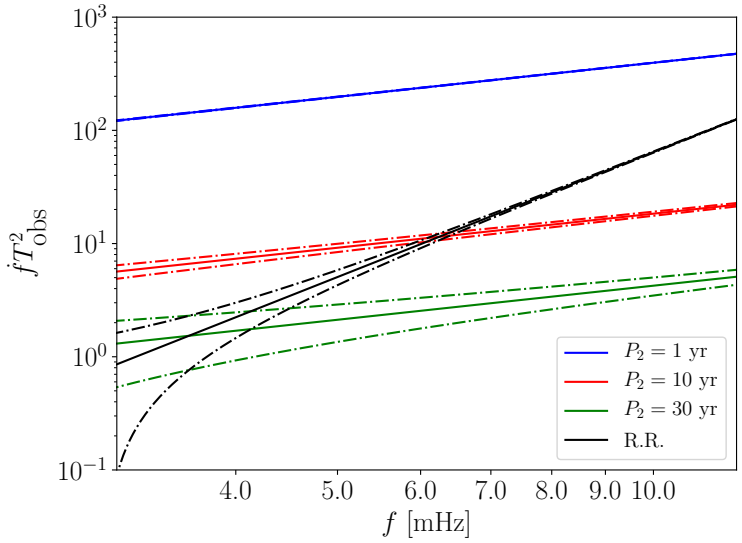

Figure 4.13: The solid lines denote the gravitational wave frequency for an isolated binary and for several binaries in hierarchical orbits with outer periods of 1,10 and 30 years. In each case the outer orbit is circular, and the SNR of the gravitational wave signal is $\rho = 50$. The dotted lines indicate the Fisher matrix error estimate for the frequency derivatives. Note the difference in power laws for the frequency derivatives and the small region of overlap between the curves. The chirp mass was chosen to be 0.32 $M_\odot$, $m_2 = 2$ $M_\odot$, $m_c = 1$ $M_\odot$.

where $\boldsymbol{\lambda}$ are the parameters which maximize the galactic binary model. The fitting factor is a measure of how well the maximum posterior galactic binary waveform $h(\boldsymbol{\lambda}_{\rm max})$ resembles the true triple waveform $h_{\rm T}$, which returns 1 when the signals are equivalent and 0 when they are perfectly orthogonal.

In Figure 4.14 we show an example where the observation period was 1 year and the carrier frequency f was 3 mHz for a circular outer orbit. We are now strictly concerned with how well we can fit a triple signal with an isolated binary model; not with how well parameters can be measured. The relevant masses for the triple were as follows: $m_a = 0.5M_\odot$, $\mathcal{M} = 0.32M_\odot$, $m_c = 1.0M_\odot$. The parameters T_2 and ω_2 were set to 0. There are three different models under consideration which will be used to fit the injected signal from a triple system. The symbol $\hat{T}$ indicates models

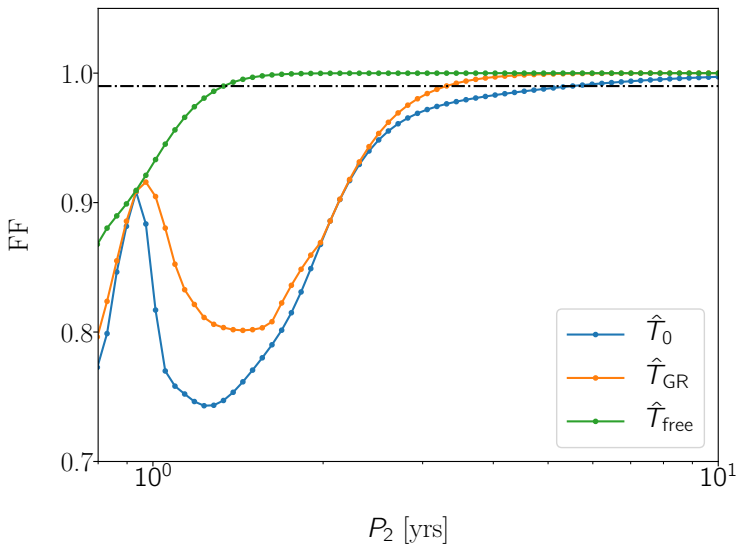

Figure 4.14: The blue line represents the fitting factor (FF) for a purely monochromatic galactic binary mode $\hat{T}_0$), the orange includes $\dot{f}$ in the frequency evolution ($\hat{T}_{\rm GR}$), and the red includes $\dot{f}$ and $\ddot{f}$ ($\hat{T}_{\rm free}$). These fitting factors are for circular outer orbits. As the outer orbital period increase the ability of the binary model to match the triple signal is improved. The parameters θ, ϕ, ψ, and ϕ_0 were chosen to be the same as Fig. 3 for the injected triple signal.

that uses a Taylor expanded frequency evolution The $\hat{T}_0$ model assumes the signal is monochromatic i.e. it is characterized by only f. The $\hat{T}_{\rm GR}$ model utilizes a three term Taylor expansion (i.e. f, $\dot{f}$, and $\ddot{f}$) in which the coefficients are related by the radiation reaction equations. Lastly, we consider the model $\hat{T}_{\rm free}$ which also utilized a three term Taylor expansion, but one in which there is no relation between the coefficients.

The $\hat{T}_0$ model is able to fit the signal from the hierarchical system for outer orbital periods that exceed ~ 4 times the observation period, while the $\hat{T}_{\rm GR}$ model does a little better, and is able to fit the signal for outer orbital periods that exceed ~ 3 times the observation period. The $\hat{T}_{\rm free}$ mode provides a good fit for outer orbital periods that exceed ~ 1.2 times the observation period. When the the outer period is comparable to, or short than the observation time the Taylor expansion representation

of the frequency evolution will begin to fail, and we need to use the full orbital model. Note that in a time-evolving analysis of the LISA data, where the analysis is updated as the data arrives on Earth, the simple Taylor expansion model will initially work well for all systems, but as time goes on it will begin to break down for systems in hierarchical orbits. Long before that happens it will be obvious that these systems are part of a hierarchical system as the frequency derivatives will be far in excess of what we expect from mass transfer or gravitational wave emission (or equivalently, the chirp masses needed to explain the frequency evolution in terms of gravitational wave emission will be much larger than is expected for stellar remnants).

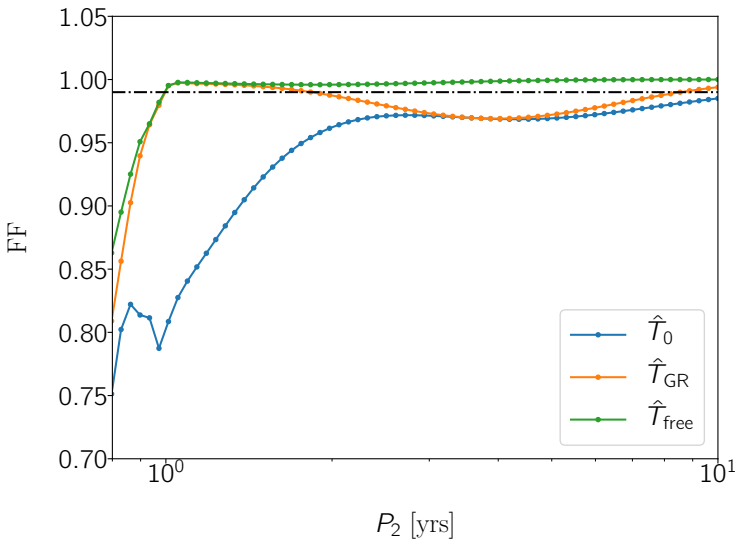

Figure 4.15: The fitting factors between an isolated binary model model and a highly eccentric triple system are displayed above. These systems have a large eccentricity, $e_2 = 0.7$. Fitting factors are larger compared to circular outer orbit case in the previous figure.

The dash horizontal black line on Figure 4.14 denotes a fitting factor of 99%, which is what we expect for a perfectly modeled signal with SNR 20. For a given SNR and model dimension D (for which the galactic binary models we are considering vary from 7 to 9) the presence of noise will cause the fitting factor to deviate from unity

even with a perfect model for the signal. The expectation value for the fitting facto in the presence of noise is [179]

$$\mathrm{FF} = 1 - \frac{D-1}{2\rho^2} \ . \tag{4.40}$$

Above the dashed line it may not be possible to distinguish the Taylor expanded models from the full hierarchical model, though it will still be possible measure some parameters of the hierarchical orbit past where the dashed black line and fitting factor lines cross. Figure 4.13 allows one to see what outer periods could reproduce $\dot{f}$'s which resemble radiation reaction i.e. when the tracks overlap.

In Figure 4.15 the eccentricity of the outer orbit is set to 0.7. We see that the same general description holds. The fitting factors for the various Taylor expansion models decrease as P_2 approaches the LISA orbital timescale. The details of the interference's affects on the fitting factor change, and the fitting factors on the left side of the plot are generally a little higher. This is due to the shift of power to higher modes in the side-bands due to the larger eccentricity such that the most visible fundamental mode has less power. There is again no danger here of mis-modeling, as even larger frequency derivatives will be needed to accurately model these signals. Lastly, in figure 4.16 we see how a 5 mHz source compares. We see that again the broad picture is intact. The shortest outer period at which the Taylor expanded models provides an "acceptable" fitting factor grows, leaving even less room for confusion between the models.

4.9 Discussion

Motivated by possibility that many of the galactic binaries observed by LISA may belong to hierarchical systems, we sought to answer three main questions (1)

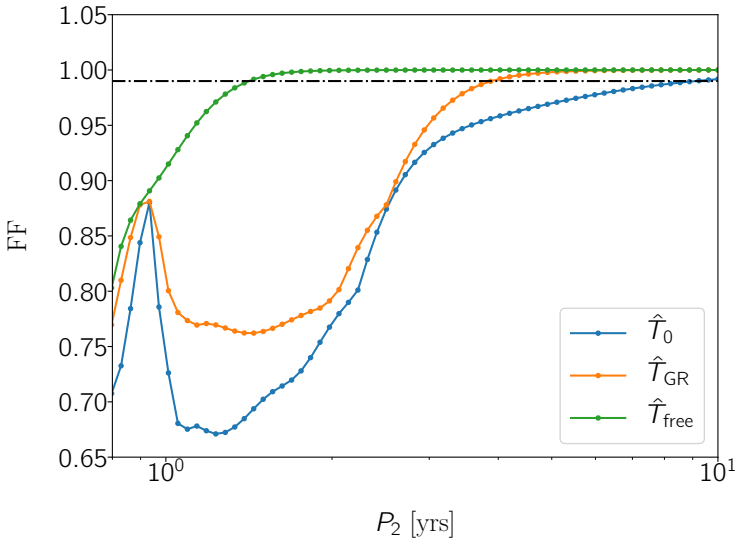

Figure 4.16: These triple systems have a circular outer orbit, but the carrier frequency is 5 mHz. This is more challenging for the isolated binary model to replicate the triple data resulting in worse fitting factors.

Under what circumstances can we detect the affects on a binary in a triple system (2) How well can we characterize the outer orbit of this system and (3) In what regime of parameter space where we may confuse a triple system with an isolated binary. The frequency evolution incurred by a center-of-mass acceleration of the inner binary due to the presence of a perturbing companion will be measurable for outer periods as many as 10 times larger than the LISA mission lifetime. The outer orbital period and eccentricity will be measurable for systems whose outer periods are no larger than a few times the LISA mission lifetime. LISA will likely detect many triple systems and characterize their orbits, and in doing so provide unique insights into the role of that hierarchical companions have on binary evolution.

There will only be a small regime of parameter space in which we would expect to confuse the frequency evolution of an isolated binary with that imparted by a hierarchical companion. Analysis of the LISA data will require a global fit, simultaneously considering all detectable sources to account for covariances between

the signals. One might be concerned about how the presence of binaries in hierarchical orbits will complicate the analysis, but it is only a mild complication. The simple Taylor expansion model will pick up the signals accurately at first, and once it becomes clear that the systems are undergoing large accelerations due to a distant companion, the signal model can be switch to the full orbital model.

There are many avenues for future research. For sufficiently tight systems, Lidov-Kozai oscillation or finite size effects may be measurable. Another interesting scenario is that of eclipsing systems. In pulsar timing, an eclipsing system allows one to disentangle the mass and inclination of a binary through the measurement of time delays in the light. For a triple system, the eclipsing companion might induce a measurable Shapiro time-delay type effect into the gravitational wave, allowing us to learn more about the system.

Acknowledgments

TR and NJC appreciate the support of the NASA grant NNX16AB98G. ST acknowledges support from the Netherlands Research Council NWO (grant VENI [nr. 639.041.645]). We benefited from useful discussions with Valeria Korol and we would like to thank the organizers of the School on Gravitational Waves for Cosmology and Astrophysics held in Benasque (2017) where initial ideas for this work were developed.

4.A Basic Binary Fisher Analysis

In this appendix we generalize the toy model introduced by Seto [169] which rather well approximates the errors in parameter estimation that one faces with a galactic binary signal in LISA. We model the signal as $h = A\cos\left(2\pi f_{\rm gw} t + \phi\right)$, where A is a constant amplitude, ϕ an arbitrary phase shift, and $f_{gw} = f + \dot{f}t + \frac{1}{2}\ddot{f}t^2$ (Note

the difference of $\frac{1}{2}$ for the definition between our $\dot{f}$ and Seto's!). In this section we will investigate how the error analysis changes as we include more or less parameters in the model.

Under the assumption that the gravitational wave frequency is mildly chirping (such that the Taylor expansion is valid) the noise-weighted inner product in the time domain can be approximated as

$$(g|k) = \frac{2}{S_n(f)} \int_0^T g(t)k(t)dt \ . \tag{4.41}$$

The Fisher matrix, in the approximation that many cycles are measured, i.e. $fT_{\rm obs} \gg 1$, can be approximated as

$$\Gamma \approx \rho^2 \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{4}{3}\pi^2 T^2 & \pi^2 T^3 & \frac{2}{5}\pi^2 T^4 & \pi T \\ 0 & \pi^2 T^3 & \frac{4}{5}\pi^2 T^4 & \frac{1}{3}\pi^2 T^5 & \frac{2}{3}\pi T^2 \\ 0 & \frac{2}{5}\pi^2 T^4 & \frac{1}{3}\pi^2 T^5 & \frac{1}{7}\pi^2 T^6 & \frac{1}{4}\pi T^3 \\ 0 & \pi T & \frac{2}{3}\pi T^2 & \frac{1}{4}\pi T^3 & 1 \end{pmatrix} \ , \tag{4.42}$$

where the matrix is ordered as $\log A, f, \dot{f}, \ddot{f}, \phi$. Upon inversion we may obtain estimates of the errors in the parameters of interest by inverting the full Fisher matrix (or in versions where the $\ddot{f}$, and/or $\dot{f}$ dimensions are dropped). When only a monochromatic signal is used the RMS errors are

$$\Delta f T_{\rm obs} = \frac{\sqrt{3}}{\pi}\rho^{-1} \approx 0.06 \left(\frac{10}{\rho}\right) \tag{4.43}$$

$$\Delta\phi = 2\rho^{-1} \approx 0.20 \left(\frac{10}{\rho}\right) \ . \tag{4.44}$$

Including $\dot{f}$ inflates the errors to the following

$$\Delta f T_{\rm obs} = \frac{4\sqrt{3}}{\pi}\rho^{-1} \approx 0.22\left(\frac{10}{\rho}\right) \tag{4.45}$$

$$\Delta \dot{f} T_{\rm obs}^2 = \frac{3\sqrt{5}}{\pi}\rho^{-1} \approx 0.21\left(\frac{10}{\rho}\right) \tag{4.46}$$

$$\Delta\phi = 3\rho^{-1} \approx 0.30\left(\frac{10}{\rho}\right) \ . \tag{4.47}$$

Lastly, if one also includes the $\ddot{f}$ term

$$\Delta f T_{\rm obs} = \frac{10\sqrt{3}}{\pi}\rho^{-1} \approx 0.55\left(\frac{10}{\rho}\right) \tag{4.48}$$

$$\Delta \dot{f} T_{\rm obs}^2 = \frac{18\sqrt{5}}{\pi}\rho^{-1} \approx 1.28\left(\frac{10}{\rho}\right) \tag{4.49}$$

$$\Delta \ddot{f} T_{\rm obs}^3 = \frac{20\sqrt{7}}{\pi}\rho^{-1} \approx 1.68\left(\frac{10}{\rho}\right) \tag{4.50}$$

$$\Delta\phi = 4\rho^{-1} \approx 0.40\left(\frac{10}{\rho}\right) \ . \tag{4.51}$$

Now we will consider a numerically calculated Fisher matrix for a galactic binary seen by LISA which includes only f, and $\dot{f}$ in its frequency evolution. The following matrix is ordered as f, $\cos\theta$, ϕ, $\log A$, $\cos\iota_1$, ψ, ϕ_0, and $\dot{f}$:

$$\Gamma = \begin{pmatrix} 5.05 \times 10^{3} & 2.77 \times 10^{2} & -1.77 \times 10^{2} & 9.85 \times 10^{-4} & 2.08 & 2.49 \times 10^{3} & 1.24 \times 10^{3} & 1.82 \times 10^{3} \\ 2.77 \times 10^{2} & 5.25 \times 10^{2} & -2.01 \times 10^{2} & -3.47 & -3.56 & 2.09 \times 10^{1} & 1.04 \times 10^{1} & 1.26 \times 10^{2} \\ -1.77 \times 10^{2} & -2.01 \times 10^{2} & 3.15 \times 10^{4} & 5.87 \times 10^{-1} & 1.37 \times 10^{1} & -5.87 \times 10^{1} & -2.88 \times 10^{1} & 5.26 \times 10^{2} \\ 9.85 \times 10^{-4} & -3.47 & 5.87 \times 10^{-1} & 4.00 \times 10^{2} & 4.13 \times 10^{2} & -6.67 \times 10^{-1} & 5.52 \times 10^{-9} & 1.32 \times 10^{-3} \\ 2.08 & -3.56 & 1.37 \times 10^{1} & 4.13 \times 10^{2} & 4.27 \times 10^{2} & 1.09 & 5.81e-01 & 1.04 \\ 2.49 \times 10^{3} & 2.09 \times 10^{1} & -5.87 \times 10^{1} & -6.67 \times 10^{-2} & 1.09 & 1.60 \times 10^{3} & 8.00 \times 10^{2} & 8.04 \times 10^{2} \\ 1.24 \times 10^{3} & 1.04 \times 10^{1} & -2.88 \times 10^{1} & 5.52 \times 10^{-9} & 5.81 \times 10^{-1} & 8.00 \times 10^{2} & 4.00 \times 10^{2} & 4.02 \times 10^{2} \\ 1.82 \times 10^{3} & 1.26 \times 10^{2} & 5.26 \times 10^{2} & 1.32 \times 10^{-3} & 1.04 & 8.04 \times 10^{2} & 4.02 \times 10^{2} & 6.95 \times 10^{2} \end{pmatrix} . \tag{4.52}$$

This system had a carrier frequency of 5 mHz, a chirp mass of $0.32 M_{\odot}$, and SNR of 20. The resulting Fisher matrix inverted, gives the following error estimates for the frequency

$$\Delta f T_{\rm obs} = 0.31 \ , \tag{4.53}$$

$$\Delta \dot{f} T_{\rm obs}^{2} = 0.61 \ . \tag{4.54}$$

These errors are rather robust to choices in the parameters of the model. Comparing these results to the toy model considered above we see that the error in f is roughly 3 times larger when using the full galactic binary model and about 6 times larger for $\dot{f}$. This results from the very strong covariance between ϕ_0 and ψ tied with the covariance of both of these parameters with f, and $\dot{f}$. If one considers galactic binaries modeled with $\ddot{f}$ as well one finds that the error in $\ddot{f}$ is about 4 times as great compared to the toy model estimate. These extra inflations are included in the analysis through the body of this paper in which we consider how tight the outer orbit must be for certain features to be measurable.

CHAPTER FIVE

A FOURIER DOMAIN WAVEFORM FOR NON-SPINNING BINARIES WITH ARBITRARY ECCENTRICITY

Contributions of Authors and Co-Authors

Manuscript in Chapter 5

Author: Blake Moore

Contributions: Analytically developed the waveform model. Developed the waveform and analysis code. Wrote the initial draft of the manuscript.

Author: Travis Robson

Contributions: Analytically developed the waveform model, and the code.

Author: Dr. Nicholas Loutrel

Contributions: Advised the project.

Author: Dr. Nicolas Yunes

Contributions: Conceived of design study. Edited the manuscript.

Manuscript Information Page

Blake Moore, Travis Robson, Nicholas Loutrel, Nicolas Yunes

Classical and Quantum Gravity

Status of Manuscript:

Officially published in a peer-reviewed journal

Published by IOP Publishing

Published November, 2018, Class. Quantum Grav. Volume 35, Number 23

Abstract

Although the gravitational waves observed by advanced LIGO and Virgo are consistent with compact binaries in a quasi-circular inspiral prior to coalescence, eccentric inspirals are also expected to occur in Nature. Due to their complexity, we currently lack ready-to-use, analytic waveforms in the Fourier domain valid for sufficiently high eccentricity, and such models are crucial to coherently extract weak signals from the noise. We here take the first steps to derive and properly validate an analytic waveform model in the Fourier domain that is valid for inspirals of arbitrary orbital eccentricity. As a proof-of-concept, we build this model to leading post-Newtonian order by combining the stationary phase approximation, a truncated sum of harmonics, and an analytic representation of hypergeometric functions. Through comparisons with numerical post-Newtonian waveforms, we determine how many harmonics are required for a faithful (matches above 99%) representation of the signal up to orbital eccentricities as large as 0.9. As a first byproduct of this analysis, we present a novel technique to maximize the match of eccentric signals over time of coalescence and phase at coalescence. As a second byproduct, we determine which of the different approximations we employ leads to the largest loss in match, which could be used to systematically improve the model because of our analytic control. The future extension of this model to higher post-Newtonian order will allow for an accurate and fast phenomenological hybrid that can account for arbitrary eccentricity inspirals and mergers.

5.1 Introduction

Eccentric binaries circularize rapidly as their orbital separation shrinks due to the emission of gravitational waves (GWs). Since the target sources of ground-

based detectors, such as the Advanced Laser Interferometer Gravitational-Wave Observatory (aLIGO) [180], advanced Virgo (aVirgo) [181], LIGO-India [182], and KAGRA [183], are thought to form at large initial separations, one expects their orbital separation and eccentricity will have decreased considerably by the time they are detectable. As such, the modeling of GWs has focused on quasi-circular binaries, and indeed all current aLIGO/aVirgo detections can be captured well with quasi-circular GW models [184–189].

Several astrophysical scenarios, however, suggest that some small number of binaries could have moderate eccentricities while emitting GWs at frequencies in the sensitivity band of ground-based detectors, and these different formation scenarios can be constrained through detection of eccentric signals [190, 191]. A very small number of weakly eccentric sources emitting detectable GWs are expected to be formed through isolated stellar evolution (field binaries). Kowalska et al. [58] simulated field binary evolution and found typical eccentricities of $\sim 10^{-4}$, with roughly 1% of binaries having eccentricities greater than 0.01 when emitting GWs detectable by ground based networks.

In contrast to binaries formed through isolated stellar evolution, binaries formed in dense stellar regions, such as globular clusters, are significantly more likely to be eccentric due to many-body interactions, such as the Kozai-Lidov mechanism [192–195]. The latter is a form of orbital resonance where oscillations in inclination and eccentricity are induced in a hierarchical triple [166]. Recently, Rodriguez et al. [196] (see also Samsing [197]) incorporated post-Newtonian (PN) effects[1] in orbital dynamics and found that 10% of binaries in globular clusters emitting GWs in the sensitivity band of ground based detectors will have eccentricities

[1]The post-Newtonian approximation is one in which the field equations are solved assuming small velocities and weak gravitational fields in an expansion in powers of $(\frac{v}{c})$, where v is the orbital velocity and c is the speed of light [198]. By nPN order we mean an expansion to order $(v/c)^{2n}$.

greater than 0.1. For more studies which focus on the eccentricity distribution of sources for ground-based detectors see [199–201].

Since there may be some small number of detectable binaries with non-negligible eccentricity, it is natural to consider what error is incurred by neglecting eccentricity in the modeling. Martel and Poisson [202] computed the fitting factor (FF), i.e. the overlap maximized over all parameters of the model, between quasi-circular templates and eccentric signals for a variety of sources at leading PN order. This study showed that as the eccentricity of the signal increases and the total mass decreases, the FF decreases. Since the percent loss in detection rate scales like $1 - FF^3$ [203], neglecting a moderate eccentricity in source modeling can lead to a significant loss in detection rate. Loss in detection rate has also been the focus of several other studies [204–206], which varied the range of masses considered and the PN order, all leading to similar conclusions: for low mass systems, the loss in match due to sub-optimal templates is significant when $e \geq 0.05$, and for higher mass systems when $e \geq 0.1$.

But even if an eccentric signal is detected with quasi-circular templates, the lack of eccentricity modeling will nevertheless lead to an associated parameter bias. Favata [207] showed that the systematic error in the symmetric mass ratio incurred by neglecting eccentricity in the model exceeds the statistical error of aLIGO for initial eccentricities as small as $e \sim 2 \times 10^{-3}$ for a binary neutron star system. For the third-generation Einstein Telescope [208], there is significant parameter error for even smaller orbital eccentricities. Schematically, this is because parameter biases become important when the match (M), i.e. the overlap without maximizing over intrinsic parameters, between quasi-circular templates and eccentric signals drops below $1 - D/(2\rho^2)$ [209], where D is the effective dimensionality of the model and ρ is the signal-to-noise ratio of the detection. Thus, even though the FF may be high, the M may still not be high enough for high signal-to-noise ratio events.

The need for eccentric waveforms has therefore encouraged some research in eccentric modeling. The Post-Circular formalism (PC), introduced at Newtonian order in [210], was one of the first attempts to provide an analytic frequency domain waveform that incorporated eccentricity. The philosophy of this formalism is to expand all relevant quantities in the small eccentricity limit. Tanay, et al. [211] extended this work to 2PN order, and Ref. [212] extended it to 3PN order, but keeping only leading-order in eccentricity corrections. While the PC models are computationally fast, they are not able to handle moderate eccentricities, and so other modeling efforts have combined analytic and numerical methods to arrive at a more accurate model. Pierro et al. [213] solved the equations of motion necessary to build the Fourier domain model of [214] without making a small eccentricity approximation by combining special functions (hypergeometric functions and Bessel functions) with certain numerical inversions. While exact in some regards within the PN approximation, this model is computationally expensive and it has only been extended to 1PN order so far [215].

We here take the first steps toward the construction and validation of a ready-to-use and computationally efficient waveform model in the Fourier domain that is valid to arbitrary eccentricity. The new model combines the accuracy of [213] with the efficiency of the PC models, without requiring the evaluation of costly special functions in the Fourier phase of the frequency response. Instead, the model is constructed by combining elements of the stationary-phase approximation, a truncated sum over harmonics and an analytic representation of hypergeometric functions. Schematically, the Fourier transform of the plus- and cross-polarizations

in the new model is

$$\tilde{h}_{+,\times}(f) = \frac{m\eta}{R} \sum_{j=1}^{N} \tilde{\mathcal{A}}^{(j)}_{+,\times}(f)\, e^{i\psi_j(f)} \ . \tag{5.1}$$

where $\tilde{\mathcal{A}}^{(j)}_{+,\times}$ is an analytic, slowly-varying, complex Fourier amplitude and ψ_j is an analytic, rapidly-varying, real Fourier phase, where j is the harmonic index and N is the truncation index, with m, η and R the total mass, symmetric mass ratio and luminosity distance to the source respectively.

The key of the new model is an analytic prescription for the Fourier amplitude and for the Fourier phase, with the truncation index determined from a match analysis. Although the Fourier phase can be solved for exactly in terms of hypergeometric functions, this representation is not computationally efficient. Instead, we explored different analytic representations of hypergeometric functions, and found that Taylor expansions about small eccentricity do an exceptional job at capturing the exact result. The Fourier amplitude, on the other hand, is not expanded in small eccentricity, and it is instead kept in its exact PN form. The truncation index is determined by requiring that the match between the truncated series and an infinite series be at least 99%. We find in practice that for most initial eccentricity cases the sum need only be taken to the tenth term or less. We then verify that the resulting waveforms are faithful (with matches $\sim 99\%$ to numerical PN waveforms) for aLIGO sources with initial orbital eccentricities as high as 0.9, as shown in Fig. 5.1 for a black hole–neutron star binary (BH-NS). All throughout, we work to leading PN order, focusing mostly on faithfulness measures for the aLIGO detector, but the approach can easily be extended to higher PN order and to other detectors.

Two main byproducts are also generated from this analysis. First, we develop a new method to efficiently maximize the overlap over the time and phase of coalescence

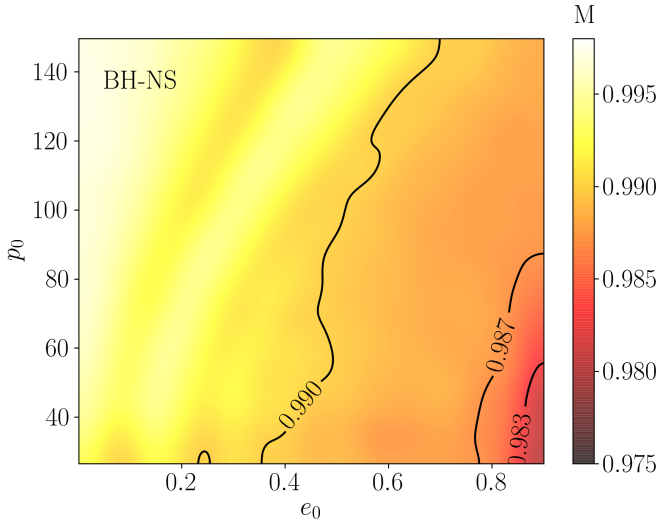

Figure 5.1: Match between our Fourier domain model and a fully numerical PN waveform as a function of initial orbital eccentricity, e_0, and initial dimensionless semilatus rectum, p_0, for a $(10, 1.4)M_\odot$ black hole - neutron star binary (BH-NS). The match is greater than 99% for more than half of the explored parameter space. The decay in match at high initial eccentricity and small initial semilatus rectum is due to finite time effects.

of the new eccentric model. Maximization over these extrinsic parameters is a solved problem for quasi-circular binaries, but the latter must be generalized non-trivially when including several harmonics with comparable power. Second, we investigate which elements of the approximations that make up our new model lead to the largest loss in accuracy. This error analysis therefore allows us to identify which elements should be taken to higher order if a higher match is desired. Due to the analytic control of the waveform model, such extensions to higher order are straightforward.

The remainder of this paper presents the details of the results described above. Section 5.2 reviews the fundamentals of eccentric GW emission, including the parameterization of the orbit and its time evolution, as well as the decomposition of the signal into a sum of harmonics of the mean orbital frequency. Section 5.3 reviews some basic data-analysis measures to compare waveform models, while presenting the

new method to efficiently maximize over the time and phase of coalescence of eccentric templates. Section 5.4 discusses the details of previous models and it introduces the new models we develop in this paper. Section 5.5 studies which of the different approximations used in the new model leads to the largest loss in match. Section 5.6 carries out a faithfulness study of the new analytic model. Section 5.7 concludes and points to future research. Throughout this work we use geometric units ($c = 1 = G$).

5.2 Fundamentals of Eccentric Binary GW Emission

In this section we begin by reviewing the Newtonian parameterization of the Kepler problem in the absence of radiation reaction. We show that the associated time domain GW waveform can be decomposed into a sum of harmonics of the mean orbital frequency. We then review how radiation reaction affects the orbital dynamics and how the application of the stationary phase approximation (SPA) leads to a Fourier response with similar structure to the time domain harmonic decomposition.

5.2.1 Newtonian Emission in the Absence of Radiation Reaction

In the Newtonian treatment of the two-body problem, the dynamics of an elliptical orbit restricted to a plane are described by:

$$r = a(1 - e\cos u), \tag{5.2a}$$

$$\phi - \phi_0 = 2\arctan\left[\left(\frac{1+e}{1-e}\right)^{1/2}\tan\frac{u}{2}\right], \tag{5.2b}$$

$$l = 2\pi F(t - t_0) = u - e\sin u. \tag{5.2c}$$

Here r is the magnitude of the relative separation vector, which we choose to be on the x–y plane $\vec{r} = (r\cos\phi, r\sin\phi, 0)$, ϕ is the orbital phase, e is the orbital eccentricity, m

is the total mass, and F is the mean orbital frequency defined by $F = 1/P$, where P is the orbital period. The semi-major axis, a, is related to the mean orbital frequency via Kepler's third law: $(2\pi F)^2 a^3 = m$. The angles l and u are the mean anomaly and eccentric anomaly, respectively, while the constants t_0 and ϕ_0 arise from integration and specify the initial orientation at some time t_0. From the above equations, one can also easily derive the following differential equations

$$\dot{r} = (2\pi m F)^{1/3} \frac{e \sin u}{(1 - e \cos u)}, \tag{5.3a}$$

$$\dot{\phi} = \frac{2\pi F (1 + e\cos\phi)^2}{(1 - e^2)^{3/2}}. \tag{5.3b}$$

Even in the Newtonian treatment, one is unable to analytically solve for the orbital separation and phase as explicit functions of time. Instead, one is forced to numerically invert Kepler's equation, Eq. (5.2c), in order to obtain the eccentric anomaly as a function of time, and thus the orbital separation and phase as functions of time. Alternatively, one can solve the differential equations presented above to obtain the orbital phase and the separation distance as a function of time.

In General Relativity, the accelerated motion of massive bodies leads to the emission of GWs. Following Martel and Poisson [202], the GW polarizations are given by

$$\begin{aligned} h_+ = -\frac{m\eta}{pR}\Bigg\{ &\left[2\cos(2\phi - 2\beta) + \frac{5}{2} e \cos(\phi - 2\beta) \right. \\ &\left. + \frac{1}{2} e \cos(3\phi - 2\beta) + e^2 \cos 2\beta \right] (1 + \cos^2\iota) \\ &+ \left[e\cos\phi + e^2 \right] \sin^2\iota \Bigg\}, \end{aligned} \tag{5.4}$$

and

$$h_\times = -\frac{m\eta}{pR}\big[4\sin(2\phi - 2\beta) + 5e\sin(\phi - 2\beta) + e\sin(3\phi - 2\beta) - 2e^2\sin 2\beta\big]\cos\iota \ , \tag{5.5}$$

where R is the luminosity distance and $\eta = m_1 m_2/m^2$ is the symmetric mass ratio, with $m_{1,2}$ the component masses. The angles β and ι are the polar angles describing the polarization axes. The *dimensionless* semilatus rectum, p, is related to the eccentricity and semi-major axis by $a = pm/(1-e^2)$. Following Moreno-Garrido, et al. [216] we decompose the time domain signal into harmonics of the mean orbital frequency such that the signal takes the form

$$h_{+,\times}(t) = -\frac{m\eta}{R}\left(2\pi m F\right)^{2/3}\sum_{j=1}^{\infty}\Big[C^{(j)}_{+,\times}\cos jl + S^{(j)}_{+,\times}\sin jl\Big] \ . \tag{5.6}$$

The harmonic coefficients $C^{(j)}_{+,\times}$ and $S^{(j)}_{+,\times}$ are functions of the orbital eccentricity e and the polarization angles (i, β).

We now briefly review how these coefficients are obtained. Making use of the relation $\cos\phi = e^{-1}\left[a(1-e^2)/r - 1\right]$, we express the strain polarizations in Eqs. (5.4) and (5.5) in the form

$$h_{+,\times} = B_1\cos\phi + B_2\left(\frac{a}{r}\right)^2\cos\phi + B_3\sin\phi + B_4\left(\frac{a}{r}\right)^2\sin\phi \ . \tag{5.7}$$

Here the B_i are functions of the orbital eccentricity, mean orbital frequency, and the angles ι and β, which can be found in Appendix A of [216]. Neglecting radiation-

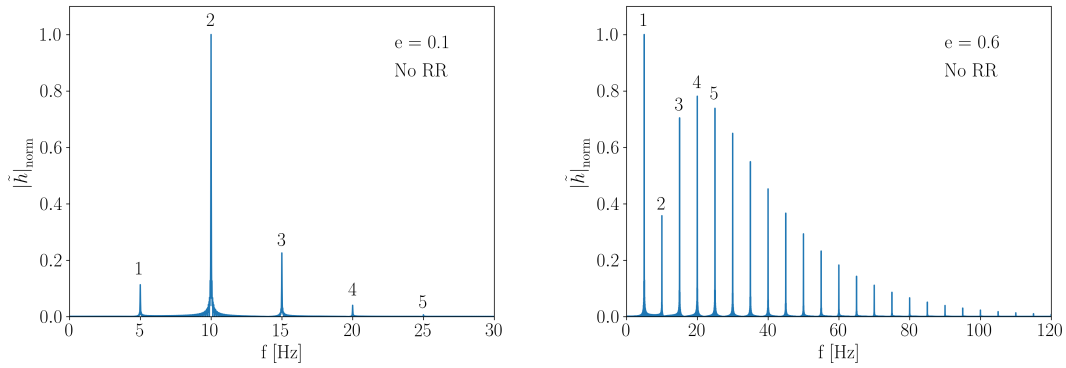

Figure 5.2: The normalized Fourier amplitude of the numerically evolved GW signal in the absence of radiation reaction, obtained by numerically solving Eq. (5.3b) and discretely Fourier transforming Eq. (5.5) for a $(10, 10)M_\odot$ binary black hole (BBH) system with a mean orbital frequency of 5 Hz and orbital eccentricity of 0.1 (left) and 0.6 (right). Observe that the Fourier amplitude naturally splits into harmonics of the mean orbital frequency, where we have labeled the first five. For systems with larger orbital eccentricities, there are many harmonics of comparable strength present.

reaction, we have access to the following Fourier series [217] ,

$$\cos\phi = -e + \frac{2}{e}(1-e^2)\sum_{j=1}^{\infty} J_j(je)\cos jl \ , \tag{5.8a}$$

$$\sin\phi = \sqrt{1-e^2}\sum_{j=1}^{\infty}\left[J_{j-1}(je) - J_{j+1}(je)\right]\sin jl \ , \tag{5.8b}$$

$$\left(\frac{a}{r}\right)^2\cos\phi = \sum_{j=1}^{\infty} j\left[J_{j-1}(je) - J_{j+1}(je)\right]\cos jl \ , \tag{5.8c}$$

$$\left(\frac{a}{r}\right)^2\sin\phi = \sum_{j=1}^{\infty} j\left[J_{j-1}(je) + J_{j+1}(je)\right]\sin jl \ , \tag{5.8d}$$

where $J_j(x)$ are Bessel functions of the first kind. Combining Eqs. (5.7) and (5.8) and rearranging to match the form given in Eq. (5.6) yields the harmonic amplitudes:

$$C_+^{(j)} = \frac{2}{e^2}\Big\{c_{2\beta}(1+c_\iota^2)e(1-e^2)j(J_{j+1}(je) - J_{j-1}(je))$$

$$- \left[c_{2\beta}(1+c_\iota^2)(e^2-2) + e^2 s_\iota^2\right] J_j(je) \Bigg\} \,, \tag{5.9a}$$

$$S_+^{(j)} = \frac{4}{e^2}(1+c_\iota^2)s_{2\beta}\sqrt{1-e^2}\Bigg\{ eJ_{j-1}(je) - \left[1+(1-e^2)j\right] J_j(je) \Bigg\} \,, \tag{5.9b}$$

$$C_\times^{(j)} = \frac{4}{e^2} s_{2\beta} c_\iota \Bigg\{ 2e(1-e^2) j J_{j-1}(je) - 2\left[1+(1-e^2)j - \frac{e^2}{2}\right] J_j(je) \Bigg\} \,, \tag{5.9c}$$

$$S_\times^{(j)} = \frac{8}{e^2} c_{2\beta} c_\iota \sqrt{1-e^2}\Bigg\{ eJ_{j-1} - \left[1+(1-e^2)j\right] J_j(je) \Bigg\} \,, \tag{5.9d}$$

with the notation $c_\theta \equiv \cos\theta$ and $s_\theta \equiv \sin\theta$. These expressions are exact, and thus, the waveform in Eq. (5.6) is valid to all eccentricities.

Figure 5.2 shows the normalized amplitude of the Fourier transform of the GW signal in the absence of radiation reaction (i.e. for a system whose mean orbital frequency and eccentricity remain constant). As the figure shows, the Fourier amplitude is composed of harmonics of the mean orbital frequency F. For the small eccentricity case shown on the left panel, the second harmonic is clearly dominant. However, as the eccentricity is increased, as shown on the right panel, the first harmonic of the mean motion dominates and many harmonics are of comparable strength.

Figure 5.2 demonstrates that one *cannot* specify a time domain quantity, such as the orbital eccentricity or the mean orbital frequency at a unique GW frequency. The presence of multiple harmonics demands that at any given time an eccentric binary emits GWs at several different GW frequencies. For example, although the eccentricity of the emitting binary is 0.6 at all times on the right panel of Fig. 5.2, this system emits GWs with significant power at 5, 10, 15, 20 Hz, etc. As such, there is no one-to-one mapping between eccentricity and GW frequency, and one cannot

unambiguously define an eccentricity as a signal "enters band." A much more sensible statement is to refer to the orbital eccentricity at a given value of the mean orbital frequency, which is uniquely defined.

5.2.2 Radiation Reaction and GW Fourier Response

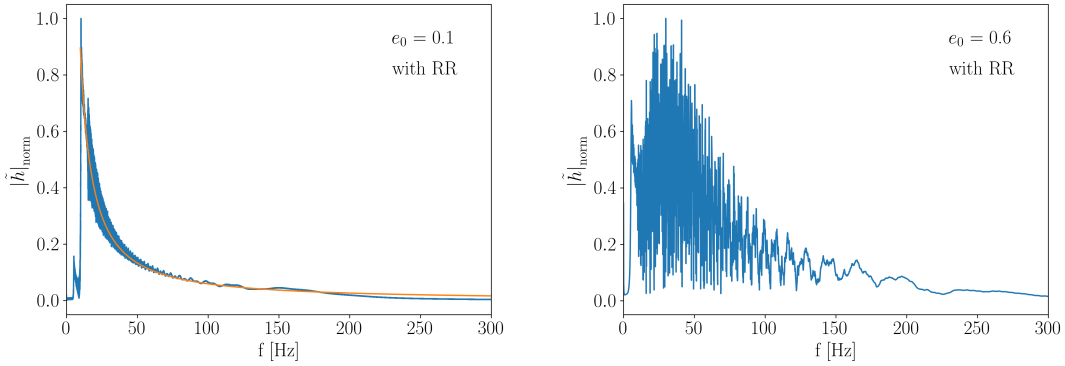

Figure 5.3: The normalized Fourier amplitude of the numerically evolved GW signal in the presence of radiation reaction obtained by numerically solving Eqs. (5.3b), (5.11), and (5.12) for a $(10, 10)M_\odot$ binary black hole (BBH) system with an initial mean orbital frequency of 5 Hz at an initial orbital eccentricity of 0.1 (left) and 0.6 (right). In the low eccentricity case on the left panel, the amplitude is similar to that of a quasi-circular GW (shown in orange), but there is some interference between harmonics above 15 Hz that leads to small oscillations about the quasi-circular spectrum. In the moderate eccentricity case shown on the right panel, the amplitude displays considerable oscillations from the interference of many harmonics of comparable strength.

Let us now consider the effect of radiation reaction on the emitted GWs in the frequency domain. In General Relativity, GWs carry away energy and momentum from the binary, and in response, e and F vary with time. At leading-order in the PN expansion, the equations for r, $\dot{r}$, ϕ, and $\dot{\phi}$, Eqs. (5.2)-(5.3), remain the same, as does Kepler's equation, but the mean anomaly l now obeys

$$l = \int^t 2\pi F(t')dt'. \tag{5.10}$$

The time evolution of the mean orbital frequency and eccentricity were first derived in [218], and are given by

$$\frac{dF}{dt} = \frac{\eta}{2\pi m^2}(2\pi m F)^{11/3}\left(\frac{96+292e^2+37e^4}{5(1-e^2)^{7/2}}\right) \tag{5.11}$$

and

$$\frac{de}{dt} = -\frac{\eta}{m}(2\pi m F)^{8/3} e\left(\frac{304+121e^2}{15(1-e^2)^{5/2}}\right). \tag{5.12}$$

These equations can be combined to form

$$\frac{dF}{de} = -3\frac{F}{e}\left[\frac{96+292e^2+37e^4}{(1-e^2)(304+121e^2)}\right], \tag{5.13}$$

which is separable and easily solved

$$F\sigma(e)^{3/2} = C_0 \tag{5.14}$$

with the definition

$$\sigma(e) = \frac{e^{12/19}}{1-e^2}\left(1+\frac{121}{304}e^2\right)^{870/2299}. \tag{5.15}$$

The constant C_0 is set by the initial conditions $F_0\sigma(e_0)^{3/2} = C_0$, where e_0 is the orbital eccentricity when the mean orbital frequency is F_0.

Applying the stationary phase approximation (SPA), reviewed in Appendix 5.A, to the time domain harmonic decomposition of the GW signal, Eq. (5.6), yields

$$\begin{aligned}\tilde{h}^{\rm SPA}_{+,\times} =& -\frac{m\eta}{2R}\sum_{j=1}^{\infty}\frac{(2\pi m F(t^*_j))^{2/3}}{\sqrt{j\dot{F}(t^*_j)}}\left[C^{(j)}_{+,\times}(t^*_j) + iS^{(j)}_{+,\times}(t^*_j)\right] \\ &\times e^{i\psi_j}\Theta\left(f-jF_0\right)\Theta\left(jF_{\rm LSO}-f\right),\end{aligned} \tag{5.16}$$

where the Fourier phase for each harmonic, ψ_j, is given by

$$\psi_j = 2\pi f t_j^* - jl(t_j^*) - \pi/4 \ . \tag{5.17}$$

Here t_j^* is the "stationary" time for the j^{th} harmonic which relates the mean orbital frequency (F) to the Fourier frequency (f) through the stationary phase condition

$$jF(t_j^*) = f \ . \tag{5.18}$$

Equation 5.18 supports the last conclusion of Sec. 5.2.1: an eccentric binary emits GWs at all integer multiples of its mean orbital frequency, and the mapping between time and GW frequency is harmonic dependent, and thus, not one-to-one.

The SPA waveform model of Eq. (5.16) contains a Heaviside function, $\Theta(x)$, because of the finiteness of GW emission in the time domain. A binary that is formed at t_0 will emit GWs until it merges, but the PN model is not valid once the orbital velocities become a non-negligible fraction of the speed of light. As is customary in the GW literature, we thus terminate the time-domain PN waveforms at the eccentric analogue of the innermost stable circular orbit of a point-particle in a Schwarzschild spacetime: the Last Stable Orbit (LSO). The mean orbital frequency at the LSO, F_{LSO}, is defined by [219–221]

$$F_{\text{LSO}} = \frac{1}{2\pi m} \left(\frac{1 + e_{\text{LSO}}}{6 + 2e_{\text{LSO}}} \right)^{3/2} . \tag{5.19}$$

When one computes the Fourier transform of this time-domain PN model in the SPA, the Heaviside function persists, as we review in Appendix 5.A.

Figure 5.3 shows the normalized Fourier amplitude of the GW signal when radiation reaction is included for systems with the same initial condition as shown

in Fig. 5.2. In the low eccentricity case ($e_0 = 0.1$) shown on the left panel, the amplitude is very close to the well known $f^{-7/6}$ trend of quasi-circular GW models. The rapid oscillations appearing past 15 Hz are due to interference between different harmonics. In the high eccentricity case ($e_0 = 0.6$) shown on the right panel, this trend is lost, and instead one finds rapid and large oscillations due to many harmonics of comparable strength interfering with one another. This result makes it clear that a faithful representation of the Fourier transform of eccentric signals must necessarily include several harmonic terms oscillating at different Fourier frequencies.

The harmonic structure of the signal can be more easily appreciated through a Q-transform, as shown for example in Fig. 5.4 for the same BBH system as that used in the right panel of Fig. 5.3. The Q-transform is a wavelet transform where the basis wavelets are Gaussian-windowed complex exponentials. Since these wavelets are localized both in time and frequency, the Q-transform produces a time-frequency representation of the GW signal. A large value of Q localizes the wavelets more in frequency, while a low Q localizes the wavelets more in time, and so in Fig. 5.4 we use a Q of 40. The harmonic structure shown in Eq. (5.16) manifests itself in Fig. 5.4 as several different tracks in time-frequency. At later times, higher harmonics become subdominant as their amplitude is proportional to the orbital eccentricity, which has significantly decayed.

The general structure of the Fourier-domain waveform in Eq. (5.16) is common across all current eccentric models. The main differences arise in how one treats (i) the mean anomaly l and the stationary time t_j^* as functions of frequency, which appear in the Fourier phase, as well as (ii) the harmonic amplitudes and the choice of truncation of the sum. At higher PN order, other differences arise, such as the precise way in which periastron precession is modeled and the inclusion of modifications to Kepler's equations at 2PN order. As a first step toward the construction of a new,

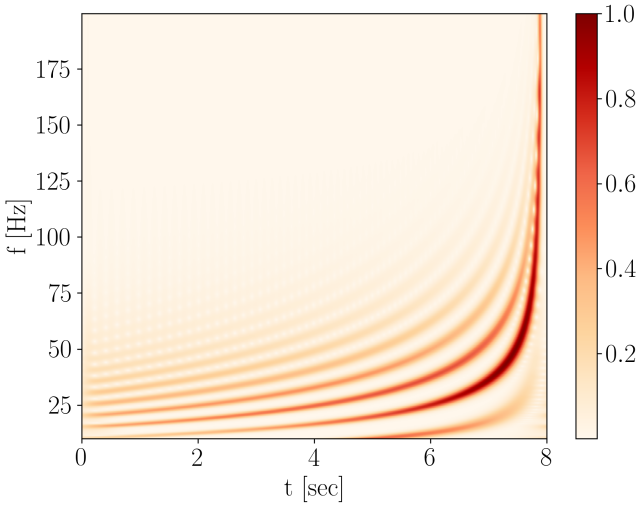

Figure 5.4: The normalized amplitude in a time-frequency representation (Q-transform) of the same system as that used for the right panel of Fig. 5.3. The presence of many harmonics in the signal leads to several tracks in time-frequency space. At later times the higher harmonics have considerably smaller amplitude, as a consequence of eccentricity decay.

analytic Fourier-domain waveforms for arbitrarily eccentric binaries, we will refrain from going to higher PN order here, but we will review current analytic models in Sec. 5.4.1.

5.3 Measures to Compare Eccentric Waveforms

Before proceeding with a description of current analytic models, and the development of a new one, it will be useful to first describe how to validate and compare different models. A useful data analysis measure to quantify the agreement between two waveforms h_1 and h_2 is through the match

$$M = \max_{t_c, l_c} \frac{(h_1|h_2)}{\sqrt{(h_1|h_1)(h_2|h_2)}} \,. \tag{5.20}$$

This statistic is the normalized inner product between two waveforms, a "signal" h_1 and a "model" h_2, maximized over a time shift t_c and phase shift l_c, which in the models we consider here arise in the phase functions t and l of Eq. (5.17) as constants of integration. The inner products are defined by

$$(h_1|h_2) = 4\mathrm{Re}\int_0^\infty \frac{\tilde{h}_1^* \tilde{h}_2}{S_n(f)} df \,, \tag{5.21}$$

where $S_n(f)$ is the noise power spectral density of the detector and Re is short-hand for the real part. In this paper, we use the design-aLIGO spectral noise density (zero-detuned, high-power) noise curve, which assumes stationary Gaussian noise [222].

When the value of the match is unity, the model perfectly represents the signal to within a time and phase offset, while the more different they are, the lower the match becomes. What value of the match is then high enough for the model to be a "faithful" representation of the signal? To set this threshold, we demand that the systematic error from mismodeling is smaller than the statistical error. Following the detailed discussion in Appendix G of [209], this requirement translates to

$$1 - M < \frac{D}{2\rho^2} \,, \tag{5.22}$$

where D is the (effective) dimension of the model, 10 in our case , and ρ is the signal to noise ratio (SNR) defined by $\rho^2 = (h|h)$.

For quasi-circular GW templates, techniques have been developed to rapidly maximize Eq. (5.20) over a time and phase offset. For eccentric templates, however, this maximization is complicated by the harmonic dependence of the phase offsets arising from l_c. For the remainder of this section, we will thus first review the maximization techniques valid in the quasi-circular limit, and then we will extend them to the case of eccentric templates, assuming the different harmonics are mutually

orthogonal. Lastly, we investigate the error incurred using this maximization scheme as we relax our assumption of orthogonality.

5.3.1 Review of Quasi-Circular Match Maximization

For quasi-circular GW templates, only the $j = 2$ harmonic is non-vanishing at leading PN order. Let us then write the template model as $h = \hat{h}e^{i\phi_c - 2\pi i f t_c}$, where $\hat{h}$ is a complex function of frequency and $\phi_c = 2l_c$. Suppose now that we have some data d that is perfectly represented by our model aside from an orbital phase shift and a time shift: $d = \hat{h}e^{i\phi_0 - 2\pi i f t_0}$, where ϕ_0 and t_0 are inherent to the data and we do not have access to their values. Our task is then to develop an algorithm to find the values of (ϕ_c, t_c) that will maximize the overlap between d and h, where in this case we know the solution is simply $(\phi_c, t_c) = (\phi_0, t_0)$.

The inner product we wish to maximize is

$$(h|d) = 4\mathrm{Re}\left[e^{-i\phi_c}\int_0^\infty \frac{|\hat{h}|^2}{S_n(f)}e^{-2\pi i f t_0 + i\phi_0}e^{2\pi i f t_c}df\right]. \tag{5.23}$$

Without specifying t_c or ϕ_c we are able to construct the function

$$\tilde{G}(f) = \frac{|\hat{h}|^2}{S_n(f)}e^{-2\pi i f t_0 + i\phi_0}\,, \tag{5.24}$$

whose inverse Fourier transform $\mathcal{F}^{-1}[\cdot]$ with respect to t_c,

$$G(t_c) = \mathcal{F}^{-1}[\tilde{G}(f)] = \int_0^\infty \frac{|\hat{h}|^2}{S_n(f)}e^{-2\pi i f t_0 + i\phi_0}e^{2\pi i f t_c}df\,, \tag{5.25}$$

appears in the above inner product

$$(h|d) = 4\mathrm{Re}\left[e^{-i\phi_c}G(t_c)\right]. \tag{5.26}$$

The quantity $G(t_c)$ is a complex number for any value of t_c and the factor of $e^{-i\phi_c}$ rotates the argument of the real operator in Eq. (5.26) in the complex plane, but does not change the *magnitude* of $G(t_c)$. Thus, we can maximize over ϕ_c by taking the magnitude of $G(t_c)$:

$$\max_{\phi_c}(h|d) = 4|G(t_c)|. \tag{5.27}$$

The overlap maximized over both ϕ_c and t_c, the match M, is found by searching for the maximum value of the array returned by $G(t_c)$ in the inverse Fourier transform over t_c, or simply

$$M = 4\max_{t_c}|G(t_c)|\,. \tag{5.28}$$

In this method the values of t_c and ϕ_c that maximize the match need not be computed explicitly to evaluate the match, and thus, we refer to this method as *implicit maximization.*

In order to explicitly find the (t_c, ϕ_c) pair that maximizes the match, one can begin by identifying the time corresponding to the value of $t_c = t_{\rm max}$ that maximizes $|G(t_c)|$. The quantity $G(t_{\rm max})$ is a complex number that is rotated off the real axis by ϕ_0 (easily verified upon inspection of Eq. (5.25) with $t_c = t_0 = t_{\rm max}$). Thus we find the value of the ϕ_c that maximizes the inner product ($\phi_{\rm max}$) via:

$$\phi_0 = \phi_{\rm max} = \arctan\left[\frac{{\rm Im}(G(t_{\rm max}))}{{\rm Re}(G(t_{\rm max}))}\right]. \tag{5.29}$$

The pair that maximize the match is then $(t_{\rm max}, \phi_{\rm max})$ and one calculates the match explicitly a posteriori. Since in this case $t_{\rm max}$ and $\phi_{\rm max}$ need to be found explicitly to evaluate the match, we refer to this method as *explicit maximization.*

These two analytic methods to maximize the match are very similar in the quasi-circular case, but as we shall see, this is not the case for eccentric templates.

Moreover, the implicit maximization method is computationally faster, as it does not require the evaluation of $t_{\max}$ or $\phi_{\max}$. We shall see next how all of this comes to play for eccentric templates.

5.3.2 Eccentric Match Maximization

The maximization of the match between an eccentric template and an eccentric signal is complicated by the harmonic dependence of the phase offsets arising from the mean anomaly at coalescence. Let us then consider the template h and the data d to be of the form

$$d = \sum_{k=1}^{\infty} \hat{h}_k e^{ikl_0 - 2\pi i f t_0}, \tag{5.30}$$

and

$$h = \sum_{j=1}^{\infty} \hat{h}_j e^{ijl_c - 2\pi i f t_c}. \tag{5.31}$$

Where $\hat{h}_j$ and $\hat{h}_k$ are complex functions of frequency associated with the j^{th} and k^{th} harmonics of mean orbital frequency. We are then tasked with maximizing

$$\begin{aligned}(h|d) &= 4\mathrm{Re}\left[\int_0^{\infty} \sum_{j,k=1}^{\infty} \frac{\hat{h}_j^* \hat{h}_k}{S_n(f)} e^{(ikl_0 - ijl_c)} e^{-2\pi i f t_0} e^{2\pi i f t_c} df\right] \\ &= \sum_{j,k=1}^{\infty} (h_j|d_k)\,. \end{aligned} \tag{5.32}$$

Without making any assumptions, the only way to maximize the above exactly is either through a grid search on t_c and l_c or through another numerical maximization scheme, such as a hill-climber algorithm. These methods are computationally expensive and slower than analytic techniques when the latter exist.

The analytic maximization techniques used in the previous subsection do not immediately extend to maximization of the match for eccentric waveforms, but we generalize them under the assumption that the harmonics are mutually orthogonal.

In the absence of radiation reaction, the different harmonics are exactly mutually orthogonal, $(h_j|d_k) = 0$ for $j \neq k$, because they are delta functions centered at integer multiples of mean orbital frequency. In the presence of radiation-reaction, this is not exactly the case any longer, but let us continue to assume it is so, and then check at the end the amount of error introduced by this approximation.

Working in the mutual orthogonal approximation, Eq. (5.32) becomes

$$\begin{aligned}(h|d) &= \sum_{j,k=1}^{\infty}(h_j|d_k) \approx \sum_{j=1}^{\infty}(h_j|d_j)\,, \\ &\approx 4\mathrm{Re}\left[\sum_{j=1}^{\infty} e^{ijl_c}\int_0^{\infty}\frac{|\hat{h_j}|^2}{S_n(f)}e^{-2\pi f t_0+ijl_0}e^{2\pi f t_c}df\right]\,. \end{aligned} \tag{5.33}$$

Without specifying t_c or l_c, we are able to construct

$$\tilde{G}_j(f) = \frac{|\hat{h_j}|^2}{S_n(f)}e^{-2\pi i f t_0+ijl_0}\,, \tag{5.34}$$

whose inverse Fourier transform $\mathcal{F}^{-1}[\cdot]$ with respect to t_c,

$$G_j(t_c) = \mathcal{F}^{-1}\left[\tilde{G}_j(f)\right] = \int_0^{\infty}\frac{|\hat{h_j}|^2}{S_n(f)}e^{-2\pi i f t_0+ijl_0}e^{2\pi i f t_c}df\,, \tag{5.35}$$

appears in the inner product of Eq. (5.33)

$$(h|d) = 4\sum_{j=1}^{\infty}\mathrm{Re}\left[e^{ijl_c}G_j(t_c)\right]\,. \tag{5.36}$$

This expression can now be maximized over l_c and t_c in a way analogous to the quasi-circular case. Each individual term in the sum of Eq. (5.36) can be individually maximized on l_c by taking the absolute value of the argument of the real operator because the factor e^{ijl_c} only rotates $G_j(t_c)$ in the complex plane, but

leaves its *magnitude* unchanged. Thus we can maximize over l_c by computing

$$\begin{aligned}\max_{l_c}(h|d) &= 4\sum_{j=1}^{\infty}\max_{l_c}\left\{\mathrm{Re}\left[e^{ijl_c}G_j(t_c)\right]\right\}\\ &= 4\sum_{j=1}^{\infty}|G_j(t_c)|. \end{aligned} \tag{5.37}$$

To maximize on t_c we take the maximum value of the array that results from taking the sum of the absolute values of the inverse Fourier transforms appearing above. Again, in this method of maximization one never explicitly calculates $l_{\max}$ to compute the match, and thus, we refer to it as the *eccentric implicit maximization* method. This method runs the risk of *overestimating* the match, as we will show later in Sec. 5.3.3 because of the mutual orthogonality assumption.

Let us now consider the explicit maximization method. In order to explicitly find $l_{\max}$ and $t_{\max}$, we begin by associating $t_{\max}$ with the maximum value of the sum of the absolute values of the inverse Fourier transforms appearing in Eq. (5.37). Inspection of Eq. (5.35) reveals that $G_j(t_{\max})$ is a complex number which is rotated off the real axis by jl_0 (again since $t_c = t_0 = t_{\max}$, the integrand appearing in Eq. (5.35) is purely real aside from the factor of e^{ijl_0}, which rotates it off the real axis). The maximum l_c value is then found using trigonometry:

$$jl_0 = jl_{\max} = \arctan\left\{\frac{\mathrm{Im}[G_j(t_{\max})]}{\mathrm{Re}[G_j(t_{\max})]}\right\}. \tag{5.38}$$

Unlike in the quasi-circular case, however, the above procedure only produces $jl_{\max}$, which is degenerate with $l_{\max} \to l'_{\max} + 2\pi/j$. One is thus forced to either break this degeneracy by computing Eq. (5.38) explicitly with two different harmonics or try each of the j degenerate guesses for $l_{\max}$ (Since the $j = 1$ harmonic is not degenerate, one could just use this harmonic.) In practice, we have found that the most reliable

method is to compute Eq. (5.38) for 3 different harmonics ($j = 2, 3, 4$), break the degeneracy, try each $l_{\rm max}$, and take the one that leads to the best match. As in the quasi-circular case, we here explicitly find the value of $(l_{\rm max}, t_{\rm max})$ that maximize the match, and thus, we refer to it as the *eccentric explicit maximization* method. As expected, this method is a bit slower than the implicit maximization, but is better behaved as we relax the assumption of orthogonality between harmonics.

5.3.3 Orthogonality Investigation

Let us begin by examining how much orthogonality is broken when including radiation-reaction by computing the quantity

$$\Delta_{i,j} = \max_{l_c} \frac{(h_i|h_j)}{\sqrt{(h_i|h_i)(h_j|h_j)}}. \tag{5.39}$$

The above inner product between different harmonics h_i and h_j is maximized over any relative phase shift. If the value of $\Delta_{i,j}$ is 0 for all i and j, then the different harmonics are orthogonal and the above approximate methods of maximization are both *exact*. In the numerics that ensue, we use our new eccentric model to evaluate Δ_{ij}, which we refer to as the *Newtonian eccentric Fourier domain* model (NeF) and will be described in detail in Sec. 5.4. The results below, however, should be representative of the orthogonality between harmonics of other PN models as well, as long as they're being evaluated in their domain of validity in eccentricity.

Table 5.1 shows $\Delta_{i,j}$ for the first 5 harmonics of the GWs emitted by a $(10, 10)M_\odot$ binary black hole (BBH) system with a few different initial conditions, where the harmonics are generated with the NeF model. We adopt slightly different frequency resolutions, with choices made to correspond to the resolution of a discrete Fourier transform of a waveform with a period determined by the initial conditions, sampled at 8192Hz in the time domain, and zero-padded such that its length is the nearest

$\Delta_{i,j}$	2	3	4	5
1	0.0035	0.00041	0.00023	0.0043
2	-	0.0013	0.000046	0.00011
3	-	-	0.0010	0.00022
4	-	-	-	0.00039

$\Delta_{i,j}$	2	3	4	5
1	0.0043	0.00074	0.00016	0.0068
2	-	0.0025	0.00024	0.000088
3	-	-	0.000046	0.000083
4	-	-	-	0.00015

$\Delta_{i,j}$	2	3	4	5
1	0.03	0.016	0.011	0.0080
2	-	0.090	0.037	0.022
3	-	-	0.033	0.013
4	-	-	-	0.0077

Table 5.1: The value of $\Delta_{i,j}$ for a $(10, 10)M_\odot$ system with $e_0 = 0.3$ and $F_0 = 3$ Hz (left table), $e_0 = 0.6$ and $F_0 = 3$ Hz (center table) and $e_0 = 0.9$ and $F_0 = 3$ Hz (right table). The first row and column of each table give the values of i and j, respectively. The diagonal terms are 1 and $\Delta_{i,j}$ is symmetric, thus redundant entries have been omitted. The frequency resolution of the inner product appearing in $\Delta_{i,j}$ is $\delta f = 0.0078$ Hz (left), $\delta f = 0.03125$ Hz (center) and $\delta f = 2$ Hz (right). Observe that orthogonality is weakly violated in the low and moderate-eccentricity cases, and less weakly violated in the high-eccentricity case.

power of 2. For the systems with lower eccentricity ($e_0 = 0.3, 0.6$), the values of $\Delta_{i,j}$ are fairly small, of $\mathcal{O}(10^{-4})$. However, for the higher eccentricity system ($e_0 = 0.9$), $\Delta_{i,j}$ is larger with values of $\mathcal{O}(10^{-2})$. This high value is partly due to the low frequency resolution, but it is consistent with the resolution of some of the matches that one would realistically find using waveforms of short duration.

Let us now examine the error incurred by our analytic maximization over t_c and l_c due to the absence of orthogonality. Assuming the maximum inner product is obtained when $t_c = t_0$ and $l_c = l_0$, we find by inspection of Eq. (5.32) that the maximized inner product takes the form

$$M_{\text{exact}} = 4 \sum_{k,j=1}^{\infty} \text{Re} \left[\int_0^\infty \frac{\hat{h}_j^* \hat{h}_k}{S_n(f)} e^{il_0(k-j)} df \right]. \tag{5.40}$$

When we consider the inner product maximized via the eccentric implicit maximization, Eq. (5.37), it is straightforward to show that we are calculating the following in

the absence of orthogonality:

$$M_{\rm implicit} = 4\sum_{j=1}^{\infty} {\rm Abs}\left[\sum_{k=1}^{\infty} e^{ikl_0}\int_0^{\infty}\frac{\hat{h}_j^*\hat{h}_k}{S_n(f)}df\right]. \tag{5.41}$$

Comparing Eqs. (5.41) and (5.40) reveals that the implicit maximization can potentially *overestimate* the value of the match and that it can even become greater than 1. Therefore, in the high eccentricity limit this maximization technique becomes inaccurate, but is still a useful approximation.

Let us now consider the error incurred by using the eccentric explicit maximization. In this case, there is no way to overestimate the match and the overlap is still properly normalized. However, we can still incur an error because the $l_{\rm max}$ derived from the explicit maximization method in Eq. (5.38) assumes mutual orthogonality, and thus, it may not be an accurate estimate of the true $l_{\rm max}$. Since $G_j(t_{\rm max})$ appears in Eq. (5.38), it is useful to consider this quantity in the absence of orthogonality:

$$G_j^{\rm exact}(t_{\rm max}) = \int_0^{\infty}\frac{|\hat{h}_j|^2}{S_n(f)}e^{ijl_0}df + \sum_{k\neq j}^{\infty}\int_0^{\infty}\frac{\hat{h}_j^*\hat{h}_k}{S_n(f)}e^{ikl_0}df. \tag{5.42}$$

Adopting the notation

$$\alpha_{j,k} = \int_0^{\infty}\frac{\hat{h}_j^*\hat{h}_k}{S_n(f)}e^{ikl_0}df, \tag{5.43}$$

we rewrite Eq. (5.42) as

$$G_j^{\rm exact}(t_{\rm max}) = \alpha_{j,j}\left(1 + \sum_{k\neq j}^{\infty}\frac{\alpha_{j,k}}{\alpha_{j,j}}\right). \tag{5.44}$$

We can now identify a small parameter

$$\epsilon \equiv \sum_{k\neq j}^{\infty}\frac{\alpha_{j,k}}{\alpha_{j,j}} \ll 1\,, \tag{5.45}$$

and expand Eq. (5.38) to first order in ϵ to obtain

$$jl_{\rm max}^{\rm exact} = \arctan\left[\frac{{\rm Im}[G_j^{\rm exact}(t_{\rm max})]}{{\rm Re}[G_j^{\rm exact}(t_{\rm max})}\right]$$
$$\approx jl_0 + \sin(jl_0)\cos(jl_0)[{\rm Im}(\epsilon) - {\rm Re}(\epsilon)]. \tag{5.46}$$

We then see that the error in $l_{\rm max}$ using the explicit maximization method is proportional to the difference in the imaginary and real parts of ϵ, which is always much less than unity.

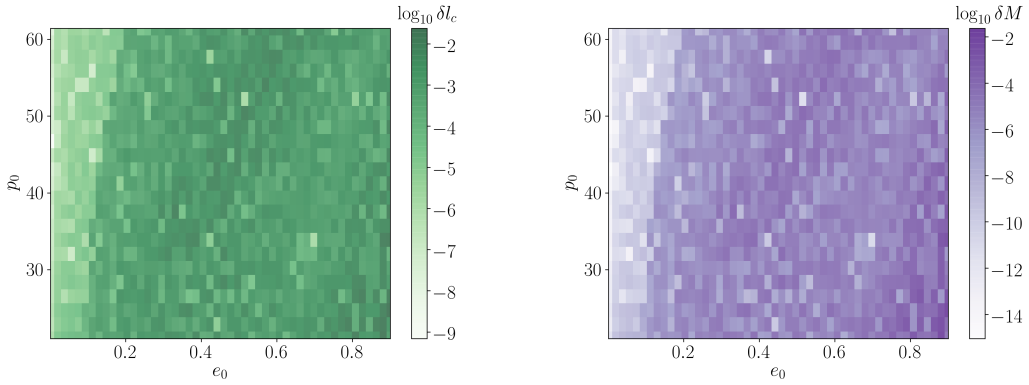

Figure 5.5: The log of the error between the $l_{\rm max}$ as predicted by Eq. (5.38) and the maximum l_c found by `Mathematica`'s built-in maximization on the left. On the right the error in the match as predicted by the explicit maximization technique and `Mathematica`'s built-in maximization is shown. The system is a BBH binary and the initial conditions are indicated on the x and y axis. The error in the match resulting from explicit maximization is greater as eccentricity increases, but the mean of the error in match is $\sim 10^{-6}$, which is much less than our estimated numerical accuracy.

The left panel of Fig. 5.5 shows the error between $l_{\rm max}^{\rm explicit}$, as predicted by the explicit maximization in Eq. (5.38), and $l_{\rm max}^{\rm exact}$, as predicted by `Mathematica`'s built in maximization routine, while the right panel shows the resulting error in the match. The matches here are between NeF and an "exact" time domain waveform (described in detail in Sec. 5.4.3) obtained by numerically solving the orbital dynamics, given

in Eqs. (5.3b), (5.11), and (5.12), and inserting these solutions into the plus and cross GW polarizations given in Eqs. (5.4) and (5.5). For a majority of the initial conditions the explicit maximization is correct to a few (3-5) decimal places. The resulting error in match is also in about the 4^{th}-8^{th} digit, with a mean error of 10^{-6}. We thus conclude that the explicit method of maximization is highly accurate.

5.4 Eccentric Models

With all of this maximization discussion under control, let us now discuss eccentric waveform models. We begin by reviewing current models in the context of the general Fourier response presented in Eq. (5.16) of Sec. 5.2.2. The main result of this work, the derivation of our *Newtonian eccentric Fourier domain model* (NeF), is presented in Sec. 5.4.2. Lastly, we present the *numerically evolved Newtonian time domain model* (NeT), which we will treat as "exact" for the purpose of comparison with our frequency domain model.

5.4.1 Previous Work

In the Post-Circular formalism (PC), the waveform model is given by Eq. (5.16), but expanding all quantities in a low eccentricity expansion. The functions appearing in the phase, ψ_j, are re-expressed via the chain rule as integrals over mean orbital frequency:

$$t - t_c = \int^F \frac{dF'}{\dot{F}(F', e(F'))} \ , \tag{5.47}$$

$$l - l_c = 2\pi \int^F \frac{F'}{\dot{F}(F', e(F'))} dF' \ . \tag{5.48}$$

The subscript c denotes the respective quantity at the time of coalescence. Since the time derivative of the mean orbital frequency appearing in the denominator

of the above integrands depends on the orbital eccentricity, one must obtain the orbital eccentricity as an explicit function of frequency, $e(F)$. Equation (5.14) is transcendental for $e(F)$, but it can be inverted in the low eccentricity limit as a bivariate expansion in e_0 and $\chi = F_0/F$. One then substitutes this inversion in the integrand appearing in Eqs. (5.47) and (5.48) to analytically perform the integration.

In the PC formalism, the amplitudes of each harmonic appearing in Eq. (5.9) are also expanded in the low eccentricity limit. Since the j^{th} harmonic amplitude scales, to leading order in a low eccentricity expansion, as e^{j-2} except for the $j = 1$ harmonic, the number of harmonics kept in the sum appearing in Eq. (5.16) is controlled by the self consistency of the low eccentricity expansion. The end result is an analytic waveform of the form of Eq. (5.16) in which all pieces are series expansions in e_0 and χ. The advantage of such models is their computational efficiency, but the main disadvantage is that they become inaccurate at moderate eccentricities.

The PC formalism has been implemented to several PN orders. The framework was introduced at Newtonian order in Ref. [210], keeping eccentric corrections up to $\mathcal{O}(e_0^8)$, and as a result, the sum in Eq. (5.16) was truncated at $j = 10$. This work was extended to 2PN order in the phasing by [211], with the PN amplitude corrections kept at Newtonian order and the PN phase corrections truncated at $\mathcal{O}(e_0^6)$. A further extension to 3PN in the phasing was done in [212], keeping only the second ($j = 2$) harmonic, the amplitude at Newtonian order in the quasi-circular limit and the PN phase corrections truncated at $\mathcal{O}(e_0^2)$. None of these models ought to be accurate for moderately eccentric systems, although a precise analysis in terms of the match and relative to exact, numerical PN waveforms have not yet been carried out.

Several other extensions of the PC framework also exist. The enhanced PC (ePC) model introduced by [223] leverages the results of [210] and the quasi-circular part of the Fourier phase (known to 3.5PN order) in order to construct a 3.5PN

eccentric model. However, this model is not constructed in a PN consistent manner. Recently [224] incorporated spin into a PC-like model, using a semi-analytic approach by computing the phase functions numerically at 3PN order. As such, the Fourier phases are accurate for all eccentricities, but the amplitudes appearing in Eq. (5.16) are computed through a low-eccentricity expansion. As in the more vanilla PC models, these extensions are also valid only for small eccentricities, with e.g. [224] claiming matches greater than 0.99 provided $e_0 \lesssim 0.3$.

One reason for the inaccuracy of all PC models is the low-eccentricity inversion of the orbital eccentricity as a function of mean orbital frequency, with comparison against numerical inversions sometimes used as a rough gauge of the regime of validity of the model. Section 5.5 studies this error quantitatively and in great detail, but let us here summarize the main results. Consider introducing inaccuracies in the exact solutions for $e(F)$ and ψ_j parametrically, and then studying the loss in match as the magnitude of the inaccuracies and the orbital eccentricity is increased. Such a study would reveal that even for mildly eccentric binaries ($e_0 \sim 0.3$), the relative error in $e(F)$ must be below $\sim 10^{-4}$ and the relative error in ψ_j must be below $\sim 10^{-3}$ for the match to remain above 99%. These results strongly suggest that an accurate Fourier domain model must represent $e(F)$ and ψ_j very accurately to avoid a large loss in match.

One potential source of confusion that occurs throughout the literature of the PC models is the interpretation of the parameter $\chi = F/F_0$. In order to evaluate the model at a given GW frequency, χ must be evaluated at the stationary point, $\chi(F(t_j^*))$, via the stationary phase condition $jF(t_j^*) = f$, yielding $\chi = f/(jF_0)$. What is often done is to choose $j = 2$ in this relation so that $\chi = f/f_0$ with $f_0 := 2F_0$ identified as "the frequency when the signal enters band." In reality, this is the GW frequency when the $j = 2$ harmonic enters band, and although this harmonic dominates the

entire signal when the eccentricity is truly small, it does not even when $e_0 \sim 0.1$, as shown in Fig. 5.2. This observation strongly suggests that an accurate Fourier domain model must represent the Fourier phase as a function of the harmonic index.

As far as we know, there is only a single alternative to the small eccentricity approximation of the PC framework if one wishes to obtain analytic Fourier waveforms, but it comes at the cost of computational expense. In Ref. [213], the stationary time and the orbital phase in Eq. (5.16) are solved at Newtonian order by changing the integration variable from time to eccentricity ($dt = de/\dot{e}$). The resulting integrals then yield hypergeometric functions which depend on the orbital eccentricity, and to invert these and express eccentricity as a function of orbital frequency, one resorts to numerical methods. Reference [213], however, does not discuss how to truncate the sum appearing in Eq. (5.16) or how a formal model ought to be constructed. Reference [215] extends this method to 1PN order, but a 1PN term in the integral for the time to coalescence is there approximated as unity, when in reality it varies from 1 to 0.94. This leads to a 2% error in $t(e)$, which Section 5.5.1 shows is too large of an error for faithful modeling. Regardless of these issues, these alternative models are computationally expensive because the amplitudes are left exact as infinite sums, and costly hypergeometric functions appear in the Fourier phase.

5.4.2 NeF Model

Let us now introduce the main result of this work: an analytic Fourier-domain waveform model that is valid to eccentricities as high as 0.9 and that we will refer to as the *Newtonian eccentric Fourier domain model* (NeF) throughout this work. This model is defined by the SPA of the harmonically-decomposed time-domain signal in Eq. (5.16), where notice that the amplitude coefficients are *not* expanded in small

eccentricity, unlike what is done in the PC model. The model is then fully defined if we can

(i) Provide an accurate analytic representation for the harmonic-dependent Fourier phase $\psi_j(e)$,

(ii) Separately attempt to solve for the orbital eccentricity e as a function of the mean orbital frequency F.

Once we have $\psi_j(e)$ and $e(F)$, we can then express the harmonic-dependent Fourier phase as a function of the GW frequency through the stationary phase condition in Eq. (5.18).

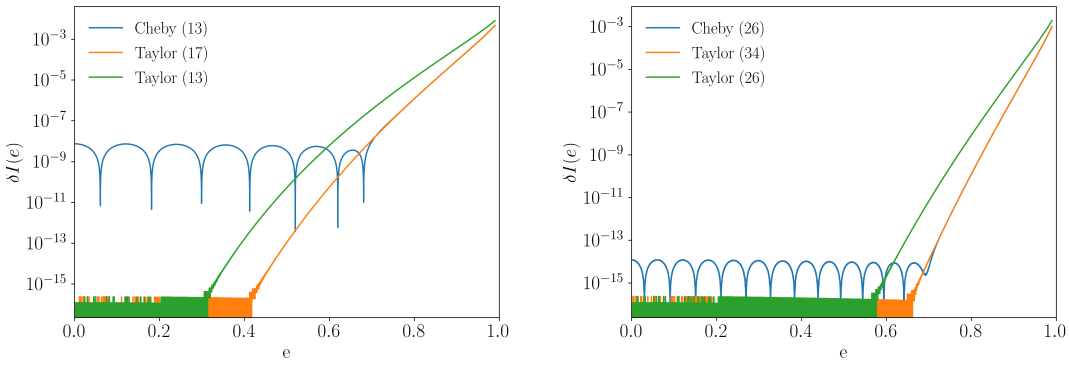

Figure 5.6: The error in $I(e)$ between the exact solution and the Chebyshev resummation, the Taylor expansion which is resummed, and a Taylor expansion which contains as many terms as the Chebyshev resummation. On the left, we keep 13 terms in the Chebyshev expansion, and 17 in the Taylor expansion which is resummed (as indicated in parenthesis in the legend). On the right we double the number of terms kept in both. We can achieve higher accuracy when keeping more terms, however, we find that a 13 term Chebyshev resummation is sufficient for our purposes.

We begin by evaluating the different pieces that make up $\psi_j(e)$ in the SPA, as for example given in Eq. (5.17). Following the work of [213], we use the chain rule

$dt = de/\dot{e}$ to write

$$t - t_c = \int_0^e \frac{de'}{\dot{e}(e')} \ , \tag{5.49}$$

$$l - l_c = 2\pi \int_0^e \frac{F(e')}{\dot{e}(e')} de' \ . \tag{5.50}$$

Using the solution for $F(e)$ given in Eq. (5.14), these integrals can be evaluated in closed form and they yield

$$\begin{aligned} t - t_c &= -\frac{m}{\eta} \frac{15}{(2\pi m F_0)^{8/3} \sigma(e_0)^4} \int_0^e \frac{\sigma(e')^4 (1 - e'^2)^{5/2}}{e'(304 + 121 e'^2)} de' \\ &= -\frac{m}{\eta} \frac{15}{304 (2\pi m F_0)^{8/3} \sigma(e_0)^4} I_t(e) \ , \end{aligned} \tag{5.51}$$

and

$$\begin{aligned} l - l_c &= -\frac{30\pi}{(2\pi m F_0)^{5/3} \sigma(e_0)^{5/2}} \int_0^e \frac{[\sigma(e')(1 - e'^2)]^{5/2}}{e'(304 + 121 e'^2)} de' \\ &= -\frac{30\pi}{(2\pi m F_0)^{5/3} \sigma(e_0)^{5/2}} I_l(e) \ , \end{aligned} \tag{5.52}$$

where we have defined

$$I_t(e) = \frac{19}{48} e^{48/19} F_1 \left(\frac{24}{19}; -\frac{1181}{2299}, \frac{3}{2}; \frac{43}{19}; -\frac{121}{304} e^2, e^2 \right) \ , \tag{5.53}$$

$$I_l(e) = \frac{19}{30} e^{30/19} {}_2F_1 \left(\frac{124}{2299}, \frac{15}{19}; \frac{34}{19}; -\frac{121}{304} e^2 \right) . \tag{5.54}$$

Here F_1 in I_t is the ApellF1 hypergeometric function and ${}_2F_1$ in I_l is the generalized hypergeometric function.

The Fourier phase of the j^{th} harmonic is then

$$\psi_j(e) = j l_c - 2\pi f t_c - j \frac{15}{304\eta} \left(\frac{1}{2\pi m F_0} \right)^{5/3}$$

$$\times \sigma(e_0)^{-5/2} \left\{ \frac{1}{\sigma[e(t_j^*)]^{3/2}} I_t(e) - I_l(e) \right\}. \tag{5.55}$$

We can simplify the Fourier phase using the identity for ApellF1 hypergeometric functions

$$\begin{aligned} F_1(\alpha;\beta,\beta';\gamma;x,y) &= (1-x)^{-\beta}(1-y)^{-\beta'} \\ &\times F_1\left(\gamma-\alpha;\beta,\beta';\gamma;\frac{x}{x-1},\frac{y}{y-1}\right). \end{aligned} \tag{5.56}$$

which then yields

$$\begin{aligned} \psi_j &= jl_c - 2\pi f t_c \\ &- j\frac{15}{304\eta}\left(\frac{1}{2\pi m F_0}\right)^{5/3} \sigma(e_0)^{-5/2} e^{30/19} I(e)\ , \end{aligned} \tag{5.57}$$

where

$$\begin{aligned} I(e) &= \frac{19}{48(1+\frac{121e^2}{304})^{124/2299}} \\ &\times F_1\left(1;-\frac{1181}{2299},\frac{3}{2};\frac{43}{19};\frac{121e^2}{304+121e^2},\frac{e^2}{e^2-1}\right) \\ &- \frac{19}{30}\ {}_2F_1\left(\frac{124}{2299},\frac{15}{19};\frac{34}{19};-\frac{121e^2}{304}\right)\ . \end{aligned} \tag{5.58}$$

The above equation for the phase is exact, and thus valid for all eccentricities, but it suffers from the fact that the hypergeometric functions are computationally costly to evaluate, especially F_1. One could of course create a look-up table for these functions to remove the computational cost, but as we shall see, there is a better, analytic approach. After exploring different representations of hypergeometric functions, we find that $I(e)$ is well-approximated by a Chebyshev resummation of

a Taylor expansion about small eccentricity. After the collecting of like terms in eccentricity in the Chebyshev resummation, we are able to approximate the frequency dependence of the phase as

$$I(e) \approx \sum_{n=0}^{n=12} C_n e^{2n}, \tag{5.59}$$

where the C_n coefficients and a brief review of Chebyshev resummations is provided in Appendix 5.B. The first few C_n, evaluated to double precision, are

$$C_0 = -0.2375000017589697 \ , \tag{5.60a}$$

$$C_1 = -0.3006152896315175 \ , \tag{5.60b}$$

$$C_2 = -0.009023468528641264 \ , \tag{5.60c}$$

$$C_3 = -0.03715292002990571 \ , \tag{5.60d}$$

$$C_4 = -0.01543183701515415 \ . \tag{5.60e}$$

Using a simple timing study implemented in `Mathematica`, we find that the resulting approximation is $\sim 10^4$ times faster than evaluating the exact result in Eq. (5.58).

Figure 5.6 shows the relative error between the exact form of $I(e)$ and the Chebyshev resummation keeping 13 terms (left panel) and 26 terms (right panel), and between the exact form and the Taylor series which is Chebyshev resummed or the Taylor series with the same number of terms as what is kept in the Chebyshev resummation. By relative error, we explicitly mean $\delta I(e) = 1 - I_{\rm app}(e)/I(e)$, where $I_{\rm app}(e)$ is the approximate solution for $I(e)$, which is the exact solution. Observe that the simple Taylor expansion does surprisingly well at representing the exact function. Observe also that the Chebyshev resummation is capable of representing the exact function in a wider range of eccentricity to a relative accuracy better than 10^{-8} and

10^{-13} in the 13-term and 26-term case respectively. We will see in Secs. 5.5 and 5.6.2 that the level of error in the left panel is tolerable to obtain matches above 99%. Nonetheless, the method can be easily extended to higher order if higher accuracy is desired for implementation in more sensitive, third-generation detectors.

The Fourier phase in Eq. (5.57) is a function of the orbital eccentricity, but the latter must be mapped to mean orbital frequency F, so that it can be further mapped to GW frequency f via the stationary phase condition $jF[e(t_j^*)] = f[e(t_j^*)]$. Reasonably, one is tempted to invert the transcendental equation for $F(e)$ in Eq. (5.14) in a low eccentricity approximation. However, Section 5.5.2 will show that the low eccentricity inversion of $F(e)$ fails at $e_0 \gtrsim 0.3$, regardless of the number of terms kept. In fact, it is this inversion for $F(e)$ which is probably responsible for the failure of all PC models.

Let us then discuss the analytic inversion of this function. The condition we must invert is

$$\sigma(e) = \zeta \,, \tag{5.61}$$

where recall that $\sigma(e)$ is defined in Eq. (5.15), we have defined $\zeta := (jC_0/f)^{2/3} = \sigma(e_0)(jF_0/f)^{2/3} = \sigma(e_0)(F_0/F)^{2/3}$, and we have used that the constant $C_0 = F_0\sigma(e_0)^{3/2}$ ensures that $F(e_0) = F_0$. Defining the inverse function κ such that $\kappa[\sigma(e)] = e$, the solution is then simply

$$e(f) = \kappa(y) = \kappa\left[\left(\frac{f}{jC_0}\right)^{3/2}\right] . \tag{5.62}$$

Since the inverse function κ is *system-independent*, this function can be obtained once and only once by any means at our disposal.

Let us first discuss analytic inversions. We were able to obtain two analytic representations of $e(\zeta)$ that meet the error tolerance that will be laid out in Sec. 5.5

(relative error of $\mathcal{O}(10^{-6})$). The first is obtained by introducing a $\bar{\sigma}_3(e) \approx \sigma(e)$, whose inverse $\bar{\kappa}_3(\zeta)$ can be found algebraically (the subscript "3" arises from the construction of other approximate $\bar{\sigma}_n$ which are reviewed in Appendix 5.C). The approximate inverse function $\bar{\kappa}_3(\zeta)$ is given by

$$\begin{aligned}\bar{\kappa}_3(\zeta) = \Bigg\{ & -\frac{1}{43923 + 277248\zeta^{19/6}} \Big[1824\zeta^{19/6} \\ & \times \left(\frac{283475(2)^{2/3}}{(38)^{1/3}\alpha} - 152\right) \\ & +4\left(18392 - \left(\frac{2}{19}\right)^{1/3} \frac{2225432}{\alpha} - 19^{2/3}\alpha\right)\Big]\Bigg\}^{1/2} , \end{aligned} \tag{5.63}$$

where we have defined

$$\begin{aligned}\alpha = & \left\{2154218176 + 12750\zeta^{19/12}\left(14641 + 92416\zeta^{19/6}\right)\right. \\ & \times \left[51\left(1216 + 11475\zeta^{19/6}\right)\right]^{1/2} - 7650\zeta^{19/6} \\ & \left. \times \left(-41031947 + 117830400\zeta^{19/6}\right)\right\}^{1/3} . \end{aligned} \tag{5.64}$$

Using this $\bar{\kappa}_3(\zeta)$, we can approximate $e(\zeta)$ to relative error of $\mathcal{O}(10^{-3})$ for sources with e_0 as high as 0.9. In order to further decrease the error we numerically fit the difference using a function of the form

$$e(\zeta) \approx \bar{\kappa}_3(\zeta)\left(1 + \frac{ae^{-b\zeta^{-c}}}{\zeta^d} + \sum_n^{30} H_n L_n(\zeta)\right) , \tag{5.65}$$

where $L_n(\zeta)$ are Legendre polynomials and the constants a, b, c, d, H_n are fitted for and presented in Appendix 5.C.

A second analytic inversion that is faster than the first can be obtained through

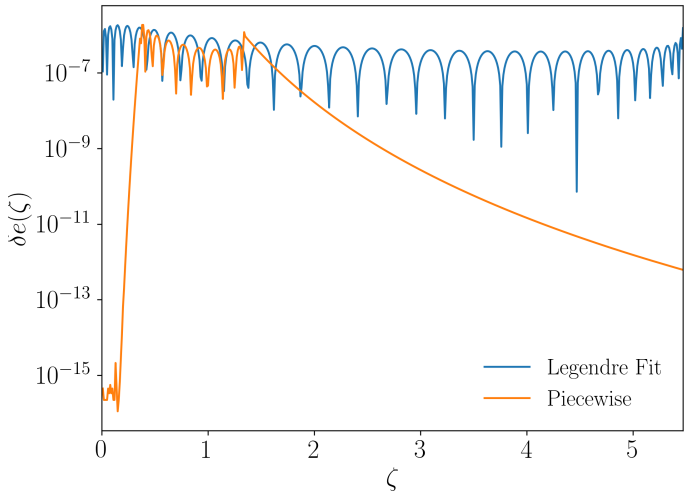

Figure 5.7: The relative error, $\delta e(\zeta)$, between a numerical solution and the Legendre fit presented in Eq. (5.65) (shown in blue) and the piecewise function in Eq. (5.66). The piecewise representation is ~ 30 times faster than the Legendre fit and covers the entire domain of $\zeta \in [0, \infty]$, with less error.

a piecewise ansatz

$$e(\zeta) \approx \begin{cases} e_{\text{Low}}(\zeta) & \zeta \leq \zeta_{\text{Low}} \\ e_{\text{Mid}}(\zeta) & \zeta_{\text{Low}} < \zeta < \zeta_{\text{High}} \\ e_{\text{High}}(\zeta) & \zeta \geq \zeta_{\text{High}} \end{cases} \tag{5.66}$$

The $e_{\text{Low}}(\zeta)$ and $e_{\text{High}}(\zeta)$ are found by inverting Eq. (5.61) in the $\zeta \ll 1$ and $\zeta \gg 1$ limits through a Taylor expansion, while $e_{\text{Mid}}(\zeta)$ is built through a numerical fit. The bounds ζ_{Low} and ζ_{High} correspond to the ζ at which the two inversions exceed $\mathcal{O}(10^{-6})$ relative error with an exact solution. The details of the construction of this function and the values of the coefficients are given in Appendix 5.C.

Although these analytic representations are sufficiently accurate for our purposes, they are not the fastest and most accurate solution to the problem. As we indicated above, Eq. (5.61) is system-independent, so any inversion that is accurate

and fast will do. In particular, due to the presence of the unit step functions in Eq. (5.16), each harmonic samples $e(f)$ at frequencies $f \in [jF_0, jF_{\rm LSO}]$, and since by the stationary phase condition $f = jF$, any harmonic samples $e(F)$ at orbital frequencies $F \in [F_0, F_{LSO}]$. Thus one only needs to solve for $e(F)$ once, and the result can be used to sample $e(f)$ for any harmonic. A fully-numerical possibility is to either solve de/dF numerically (constructed by differentiating Eqs. (5.11) and (5.12)) or to sample $F(e)$ sufficiently discretely, swapping the columns and then interpolating the result with a cubic spline. Either of these two methods is extremely fast and it can be carried out to double precision.

Figure 5.7 shows the relative error between the numerical solution for $e(\zeta)$, the fit developed in Eq. (5.65), and the piecewise representation of $e(\zeta)$ in Eq. (5.66). Both analytic representations of $e(\zeta)$ maintain a relative error below $\mathcal{O}(10^{-6})$ which is what we will show in Sec. 5.5 is needed for applications in which one is concerned with keeping a high value of the match. However, the evaluation of the Legendre (piecewise) representations is roughly 100 (3) times slower than the evaluation of the numerical solution. For this reason, we employ the numerical solution when constructing the NeF model henceforth due to its computational speed advantages.

5.4.3 NeT Model

The validation of the NeF model requires its comparison to another model that we consider more accurate or exact. We will take this to be an eccentric extension of the quasi-circular TaylorT4 approximant [225], which we will refer to as the *Newtonian eccentric Time domain model* (NeT). This is obtained by numerically solving the set of differential equations in Eqs. (5.3b), (5.11) , and (5.12) for $\phi(t)$, $F(t)$, and $e(t)$, respectively, and then inserting these expressions in the plus and cross polarizations given by Eqs. (5.4) and (5.5). The time-domain polarizations are then discretely

Fourier transformed to obtain a representation in the frequency domain.

The integration of the equations of motion is done as follows. We start all numerical integrations with initial conditions $\phi(t_0) = -\pi/2$, $F(t_0) = F_0$, and $e(t_0) = e_0$. This choice is such that binary is initially moving towards and close to pericenter passage. For highly eccentric systems ($e \approx 0.9$), much of the GW power is emitted during pericenter passage, and the eccentricity can decrease significantly between passages if the system is compact enough. Thus, our choice of initial conditions is such that the binary begins emitting GWs with significant power at an initial orbital eccentricity of e_0. For systems with small initial eccentricity or large initial separation (low initial mean orbital frequency) the choice of ϕ_0 becomes less important. We stop our numerical evolution at the last stable orbit (LSO), which is given by Eq. (5.19). Two independent implementations of this integration were developed, one in `Mathematica` with precision and accuracy goals of 13, and one in `C` using the Implicit Bulirsch-Stoer method with a supplied Jacobian for adaptive step-size control. These implementations lead to $h(t)$, $n(t)$, and $e(t)$ that agree to $\sim 10^{-8}$ relative error.

The discrete Fourier transform is done as follows. First, the time-domain functions are sampled at a rate of 8192 Hz, leading to N samples. Then, this discretely sampled time-series is zero padded on both edges with pads of equal length, such that the new time-series is of length 2^p, with p the smallest integer such that $2^p > N$. The new time series is then discrete Fourier transformed.

5.5 Error Analysis of SPA Ingredients

In this section, we investigate the loss in match, maximized via the prescription given in Sec. 5.3, due to errors in the inversion of $F(e)$ and in the calculation of $\psi_j(e)$ to inform analytic models on tolerable errors. We then show that the standard low

eccentricity inversion of $F(e)$ leads to a significant loss of match when $e_0 \sim 0.3$ for a BBH system.

$\log_{10}(\epsilon)$	$e_0 = 0.1$		$e_0 = 0.3$		$e_0 = 0.6$		$e_0 = 0.9$	
-16	0.9877	(0.9877)	0.9847	(0.9847)	0.9825	(0.9825)	0.9828	(0.9828)
-8	0.9877	(0.9877)	0.9847	(0.9847)	0.9825	(0.9825)	0.9828	(0.9828)
-6	0.9877	(0.9877)	0.9847	(0.9847)	0.9825	(0.9825)	0.9824	(0.9825)
-5	0.9877	(0.9877)	0.9847	(0.9847)	0.9820	(0.9822)	0.9743	(0.9761)
-4	0.9872	(0.9876)	0.9844	(0.9846)	0.9765	(0.9790)	0.9046	(0.9134)
-3.5	0.9854	(0.9871)	0.9790	(0.9826)	0.9312	(0.9546)	0.8459	(0.8852)
-3	0.9670	(0.9833)	0.9469	(0.9705)	0.8641	(0.9189)	0.7623	(0.8389)
-2.75	0.9395	(0.9773)	0.9108	(0.9605)	0.8185	(0.8952)	0.7189	(0.8086)
-2.5	0.8776	(0.9639)	0.8500	(0.9447)	0.7614	(0.8682)	0.6532	(0.7783)
-2.25	0.7886	(0.9438)	0.7633	(0.9223)	0.6853	(0.8444)	0.5887	(0.7502)
-2	0.6703	(0.9168)	0.6506	(0.8923)	0.5890	(0.8105)	0.5137	(0.7130)

Table 5.2: Match for different values of ϵ when using e_1 (not in parenthesis) and e_2 (in parenthesis) as approximate inversions. The first column lists the value of ϵ, the other columns list the value of the match for a given e_0. For each of the values shown here, the corresponding system is a BBH system with an initial dimensionless semi-latus rectum of $p_0 = 50$. For large eccentricities, the match begins to decrease as $\epsilon \sim 10^{-5}$.

In order to investigate the loss in match due to these inaccuracies, we require an exact solution for the analytic Fourier response. To create this exact model we interpolate the eccentricity dependent term of the phase, $I(e)$, given in Eq. (5.58), and we generate a numerical solution for $e(F)$ by numerically solving Eq. (5.14). With these solutions in hand, we have an exact (to machine precision), numerical solution for the Fourier response in the SPA, and it is this exact solution that we will use to investigate inaccuracies in ψ_j and $e(F)$. For consistency with our faithfulness study

$\log_{10}(\epsilon)$	$e_0 = 0.1$		$e_0 = 0.3$		$e_0 = 0.6$		$e_0 = 0.9$	
-16	0.9877	(0.9877)	0.9847	(0.9847)	0.9825	(0.9825)	0.9828	(0.9828)
-8	0.9877	(0.9877)	0.9847	(0.9847)	0.9825	(0.9825)	0.9828	(0.9828)
-6	0.9877	(0.9877)	0.9847	(0.9847)	0.9825	(0.9825)	0.9827	(0.9827)
-5	0.9877	(0.9877)	0.9847	(0.9847)	0.9823	(0.9824)	0.9809	(0.9812)
-4	0.9874	(0.9877)	0.9847	(0.9847)	0.9820	(0.9823)	0.9311	(0.9416)
-3.5	0.9874	(0.9874)	0.9830	(0.9841)	0.9670	(0.9743)	0.8904	(0.9057)
-3	0.9862	(0.9858)	0.9680	(0.9781)	0.9067	(0.9399)	0.8276	(0.8693)
-2.75	0.9655	(0.9825)	0.9431	(0.9695)	0.8647	(0.9202)	0.7714	(0.8459)
-2.5	0.9299	(0.9753)	0.9028	(0.9587)	0.8156	(0.8957)	0.7269	(0.8158)
-2.25	0.8617	(0.9603)	0.8369	(0.9419)	0.7561	(0.8724)	0.6540	(0.7819)
-2	0.7661	(0.9390)	0.7445	(0.9180)	0.6755	(0.8423)	0.5887	(0.7522)

Table 5.3: Match for different values of ϵ when approximating the Fourier phase with I_1 (not in parenthesis) and I_2 in parenthesis. The first column lists the value of ϵ, and the other columns list the value of the match for a given e_0. For each of the values shown here, the corresponding system is a BBH system with an initial dimensionless semi-latus rectum of $p_0 = 50$.

in Sec. 5.6.2, we truncate the the sum over harmonics in Eq. (5.16) at a sufficiently high harmonic index such that the match between it and the NeT model is ~ 0.99, as detailed in Sec.5.6.1 and 5.6.2.

5.5.1 Loss in match due to inaccurate $e(F)$ and ψ_j

To investigate the loss in match due to the inaccurate inversion of $F(e)$, we begin by numerically solving Eq. (5.14) and refer to this solution as $e(F)$. To simulate inaccuracies, we construct two different expressions for $e(F)$ using the exact numerical

solution:

$$
\begin{aligned}
e_1 &= e(F)\,(1-\epsilon)\,, \\
e_2 &= e(F)\left[1-\epsilon\left(\frac{e(F)}{e_0}\right)^2\right]\,,
\end{aligned}
\tag{5.67}
$$

with ϵ a constant. In e_1, the relative error with respect to the exact solution is a constant ϵ, while in e_2 it is $\epsilon\,[e(F)/e_0]^2$, decreasing as $e(F)^2$. As before, by relative error we here mean $1-[e_{\rm app}/e(F)]$, where $e_{\rm app}=e_1$ or e_2 is the approximate solution.

We calculate the value of the match for different values of ϵ, given a BBH system with initial dimensionless semi-latus rectum of $p_0=50$ and at various different initial eccentricities ($e_0=0.1, 0.3, 0.6,$ and 0.9). We increase epsilon to reflect larger values of relative error between the approximate and exact solutions. The values of these matches are shown in Table 5.2. In the e_1 case, the low eccentricity cases begin to show a significant decrease in match when $\epsilon\sim 10^{-3.5}$. This implies that a more accurate inversion of $F(e)$ is required than that quoted by many PC models if one is interested in applications that require such high values of the match, such as parameter estimation. In the moderately eccentric case, the match starts to decrease significantly when $\epsilon\sim 10^{-4}$, while in the large eccentricity case even when $\epsilon\sim 10^{-5}$. In the e_2 case, the trend is similar, but the decrease in match is less sharp for larger ϵ. The match will be more sensitive for systems that are longer lived (lower mass), but these results set roughly tolerable inaccuracy in the inversion of $F(e)$.

We take a similar approach to analyze the loss in match due to inaccuracies in the Fourier phase, $\psi_j(e)$. We interpolate the eccentricity dependent part of the phase appearing in Eq. (5.58) and refer to this as $I(e)$. To simulate inaccuracies, we

construct two approximate expressions for $I(e)$ using the exact solution

$$I_1 = I(e)(1-\epsilon)\,,$$
$$I_2 = I(e)\left[1-\epsilon\left(\frac{e}{e_0}\right)^2\right]\,, \tag{5.68}$$

with ϵ a constant. In the above, we use a numerical solution for $e(F)$ to isolate inaccuracies in the latter from those in the functional solution for $I(e)$. As in the $e(F)$ case, the relative error in I_1 is a constant ϵ, while that in I_2 decreases as e^2.

Table 5.3 shows the match when using the approximate I_1 and I_2 expressions in the Fourier phase. Observe that the Fourier phase can tolerate more inaccuracy than the inversion of $e(F)$. This is not surprising given that $I(e)$ is multiplied by an overall factor of $e^{30/19}$. In the low eccentricity case, the match begins to decay significantly when $\epsilon \sim 10^{-2.75}$, while in the large eccentricity case, the match deteriorates even when $\epsilon \sim 10^{-4}$. The greater sensitivity to error in the higher eccentricity case can be explained by the appearance of the overall factor of harmonic index j multiplying $I(e)$, and thus, also multiplying the error. As the higher harmonics become more important (i.e. for moderate to large eccentricities), the phase must be approximated more accurately.

5.5.2 Error in the low eccentricity inversion of $F(e)$

Let us conclude by reviewing the standard low eccentricity inversion of $F(e)$ as given in Eq. (5.14), and investigate its associated loss in match. Our task is then to solve Eq. (5.61) for $e(F)$ perturbatively in $e_0 \ll 1$. We then wish to invert Eq. (5.61) in the limit of small ζ, which corresponds to a low eccentricity and large frequency expansion, since $\zeta \ll 1$ requires $e_0 \ll 1$ and $(F_0/F) \ll 1$. For a system with initial eccentricity e_0, $\zeta \in [\sigma(e_0)(F_0/F_{\rm LSO})^{2/3}, \sigma(e_0)]$. We invert Eq. (5.61) by proposing a

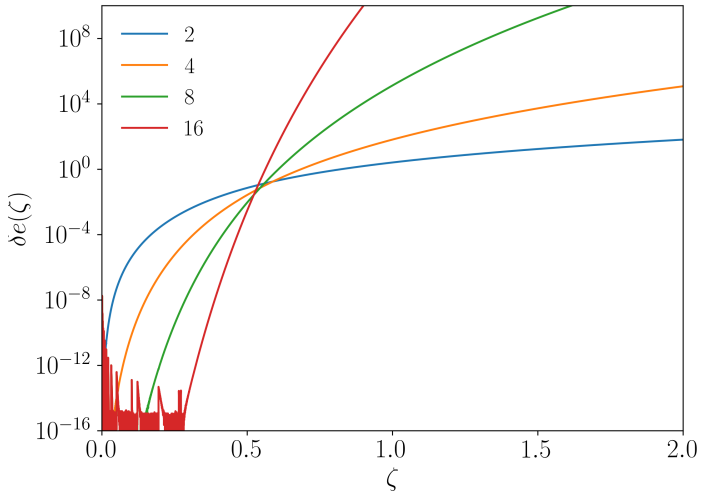

Figure 5.8: The relative error between a numerically solved $e(\zeta)$ and an $e(\zeta)$ obtained by inverting $\zeta(e)$ in the low eccentricity limit. Here $\delta e(\zeta) = 1 - e_{\rm app}(\zeta)/e(\zeta)$, where $e_{\rm app}(\zeta)$ is the approximate inversion of $F(e)$ and $e(\zeta)$ is the exact solution. The different colored lines represent different values of $n_{\rm max}$ (number of terms kept in the Taylor expansion). For $\zeta \sim 0.5$ the error in $e(\zeta)$ becomes considerable which corresponds to $e_0 \sim 0.3$ regardless of the number of terms kept. This suggests that even keeping even more terms in the approximate inversion of $\zeta(e)$ will not increase its domain of validity.

solution of the form

$$e(\zeta) \approx \zeta^{19/12} \sum_{n=0}^{n_{\max}} M_n \zeta^{19n/6}. \tag{5.69}$$

This proposed solution for $e(\zeta)$ is then substituted into $\sigma(e)$ appearing in Eq. (5.61) and expanded in small ζ. The coefficients, M_n, are determined by demanding that $\sigma[e(\zeta)] = \zeta$.

Figure 5.8 shows the relative error between the solution in Eq. (5.69) with different values of $n_{\max}$ and a numerical solution for $e(\zeta)$. Regardless of the number of terms kept in the approximate inversion, the error becomes considerable as $\zeta \sim 0.5$ which corresponds to an initial orbital eccentricity of $e_0 \sim 0.3$. This suggests that the PC models (where all quantities are expanded in low eccentricity) will become unfaithful near $e_0 \sim 0.3$ due to the approximate inversion of $F(e)$. Table 5.4 shows the match when using the low-eccentricity expansion of $e(\zeta)$ for different values of $n_{\max}$. Increasing $n_{\max}$ mitigates the loss in match until the inversion is pushed past $e_0 = 0.3$, at which point the value of the match drops to nearly 0 in one case as the inversion becomes unphysical. While keeping more terms may increase the accuracy at low ζ, the solution becomes useless for $e_0 \gtrsim 0.35$.

5.6 Faithfulness Analysis of NeF Model

In this section, we will determine the faithfulness of the NeF model by comparing it to the NeT model. All matches in this paper are computed with the response function $h(t) = F_+ h_+(t) + F_\times h_\times(t)$, where the plus and cross polarizations are given by the different waveform models (either NeF or NeT), while $F_{+,\times}$ are the antenna patterns of the detector

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

$i_{\max}$	$e_0 = 0.05$	$e_0 = 0.10$	$e_0 = 0.15$	$e_0 = 0.20$	$e_0 = 0.25$	$e_0 = 0.30$	$e_0 = 0.35$
2	0.9911	0.9870	0.9766	0.9531	0.9242	0.8921	0.8527
4	0.9911	0.9877	0.9908	0.9878	0.9799	0.9671	0.9465
8	0.9911	0.9877	0.9908	0.9903	0.9875	0.9806	0.9724
16	0.9911	0.9877	0.9908	0.9904	0.9884	0.9831	0.0057
exact	0.9911	0.9877	0.9908	0.9904	0.9884	0.9841	0.9886

Table 5.4: The value of the match for different values of $i_{\max}$ when inverting $F(e)$ in a low eccentricity approximation. The first column lists the value of $n_{\max}$ (the "exact" in the last row indicates a numerical solution for $e(F)$), and the other columns list the value of the match for a given e_0. For each of the values shown here, the corresponding system is a BBH system with an initial semilatus rectum of $p_0 = 50$. The value of 0.0057 for the 16 term inversion at an e_0 of 0.35 is due to $e(\zeta)$ returning eccentricities less than 0. Presumably, as more terms are kept in the inversion $e(\zeta)$ is very accurate for small ζ and then becomes unusable for $\zeta \sim 0.5$ as indicated in Fig. 5.8 which corresponds to $e_0 \sim 0.3$. The loss in match becomes fairly large for initial eccentricities greater than $e_0 \sim 0.35$ regardless of the number of terms kept in the inversion.

$$- \cos\theta \sin 2\Phi \sin 2\psi, \tag{5.70a}$$

$$F_\times(\theta, \Phi, \psi) = F_+(\theta, \Phi, \psi - \pi/4). \tag{5.70b}$$

Here θ and Φ are the sky angles associated with the orientation of the detector and ψ is the polarization angle that defines the rotation from the wave's polarization basis to that defined by the arms of the detector. All throughout, we choose $\theta = \Phi = \psi = \beta = \iota = 3\pi/7$ as an arbitrary location and source orientation, where recall that β and ι are the polar angles describing the polarization axes.

Before we can study the faithfulness of the NeF model of Sec. 5.4.2, we must first decide the appropriate number of harmonics that ought to be kept in Eq. (5.16). Section 5.6.1 addresses this in detail. With this at hand, Section 5.6.2 studies the faithfulness of the NeF model with the optimal number of harmonics kept against the

NeT model.

5.6.1 Maximum Number of Harmonics Needed

We begin by studying the optimal number of harmonics that need to be kept in order for the match to be above 99%. We do so by computing the match between the NeT model (which contains all harmonics) and a harmonically-decomposed time domain model (obtained by solving $\dot{F}$, $\dot{e}$, $\dot{l}$ and using Eq. (5.6)). We then demand that we keep as many harmonics in the latter such that its match against NeT exceeds 0.99. This value of the match ensures the statistical error exceeds the systematic error due to mis-modeling for sources of roughly SNR 20 or less, as discussed in Sec. 5.3.

Figure 5.9 shows the number of harmonics needed in the time domain harmonic decomposition for the match against NeT to exceed 0.99. Results are shown for three different fiducial systems, a $(1.4, 1.4)M_{\odot}$ binary neutron star (BNS) system, a $(1.4, 10)M_{\odot}$ neutron star - black hole (BH-NS) system, and a $(10, 10)M_{\odot}$ binary black hole (BBH) system. For systems with initial eccentricities less than ~ 0.6 (and for even larger eccentricities at larger initial separations), we only require 10 or less harmonics to achieve this threshold match. We thus expect that one need only keep just as many harmonics in the NeF model. This result is very promising as the biggest computational expense of the NeF model is the inclusion of higher harmonics.

Attempts have been made to analytically estimate the number of harmonics that one must keep in order to recover 99% *signal power*. Reference [199] argued that since the power in GWs is peaked near ω_p (orbital frequency at pericenter) and the fractional power in GWs is less than 10^{-3} past $5\omega_p$ [226], one can then approximate the number of harmonics needed to capture 99% of signal power via

$$j_{\max} = \frac{5\omega_p}{\omega_{orb}} = 5\frac{(1+e_0)^{1/2}}{(1-e_0)^{3/2}}. \tag{5.71}$$

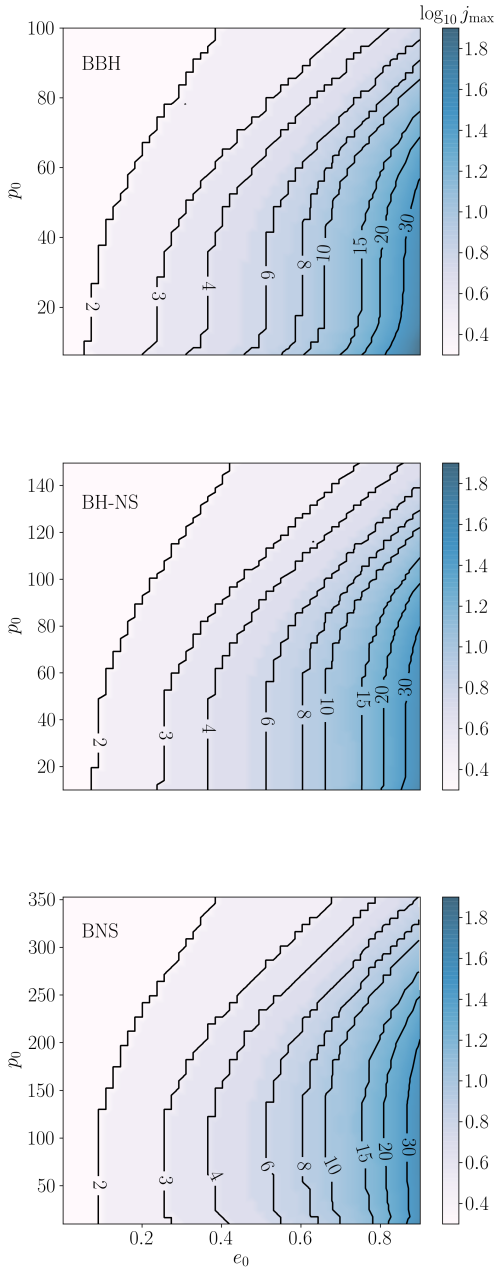

Figure 5.9: The number of harmonics required to achieve a match greater than 0.99 as a function of e_0 and p_0 (the initial orbital eccentricity and dimensionless semilatus rectum) for a BBH (top), BH-NS (middle), and BNS (bottom) system. For systems with initial eccentricities less than ~ 0.6 (and even larger eccentricities at larger initial separations) we only require 10 or less harmonics to achieve a match of 0.99.

Reference [227] takes a different approach and uses [218] to calculate the power in a given harmonic. With this at hand, they then compute the harmonic with the most power, $j_{\rm m}$, for a collection of different eccentricities, which they finally fit to obtain

$$j_{\rm m} = 2\frac{(1+e)^{1.1954}}{(1-e^2)^{3/2}}\,. \tag{5.72}$$

Equation (5.72) correctly limits to the $j_{\rm m} = 2$ quasi-circular harmonic as $e \to 0$, while Eq. (5.71) reduces to $j_{\rm max} = 5$ in this limit.

Figure 5.10 compares Eq. (5.71) to Eq. (5.72) and to a cross-section of our results for a fixed $p_0 = 20$ and $p_0 = 80$, using a BBH system. Observe that Eq. (5.71) overestimates the number of harmonics needed regardless of initial separation. This is because this equation is derived in the limit of an unbound orbit, and thus, it behaves poorly for moderate eccentricities. Also, this equation is designed to recover 99% of the signal *power*, whereas the match, and the inner product, are weighted by the spectral noise density of the detector; therefore, 99% signal power does not imply a 99% match (the threshold for our choice of $j_{\rm max}$).

Equation (5.72) does remarkably well at predicting the number of harmonics needed when the initial separation is small. This result is particularly surprising because Eq. (5.72) predicts the harmonic at which the power spectrum peaks, but referring back to Fig. 5.2, although the amplitude peaks at the $4^{\rm th}$ harmonic, the higher harmonics are still relatively strong. One would therefore reasonably expect that significantly more harmonics should be kept than just the peak harmonic. However, the eccentricity of the binary is rapidly decaying, and as a result, the signal power in higher harmonics also rapidly decays. Coincidentally, the harmonic at which the power spectrum peaks at the initial orbital eccentricity gives a fair estimate for the number of harmonics needed for faithful matches. However, as the initial separation

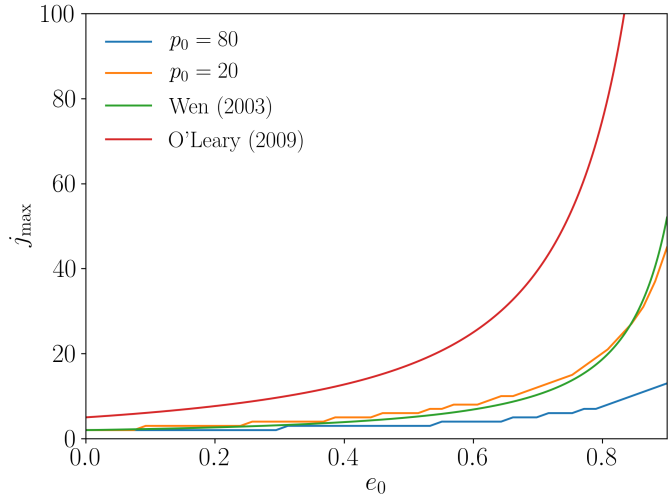

Figure 5.10: Number of harmonics needed as predicted by Eq. (5.71) in red and Eq. (5.72) in green. In orange and blue we plot the number of harmonics needed for a fixed p_0 using the BBH results of Fig. 5.9 for a close initial dimensionless semi-latus rectum (20) and a wider initial dimensionless semi-latus rectum (80). For close initial separation, Eq. (5.72) gives a good estimate for the number of harmonics needed. As the initial separation increases, one needs even less harmonics than predicted by Eq. (5.72). Equation (5.71) overestimates the number of harmonics needed in all cases.

increases, even Eq. (5.72) overestimates the number of harmonics needed. For such large initial separations, the power of the higher harmonics has sufficiently decreased by the time they enter the sensitivity band, and thus, they affect the match even less. It is also worth reiterating that this comparison was done requiring a 0.99 match (which is driven by the desire to reduce systematic biases for signals with roughly SNR 20 or less), but to achieve higher matches (considering sources with higher SNR) one would need to keep more harmonics.

5.6.2 Faithfulness of the SPA

Finally, in this section we estimate the faithfulness and the domain of validity of the NeF model of Sec. 5.4.2. To do so, we truncate the NeF model at the maximum

harmonic index found in Sec. 5.6.1. We then compute the match between the NeF and the NeT models. Since the harmonic index at which we choose to truncate our frequency domain model corresponds to the number of harmonics needed to achieve a match of at least 0.99 in the numerical time domain harmonic decomposition, the best match the NeF model can possibly achieve is also 0.99. Of course, this maximum could be improved by keeping more harmonics, but we expect the results we obtain will not qualitatively change.

The NeT and NeF models are discussed in detail in Sec. 5.4.3 and Sec. 5.4.2 respectively, but we go over a few key details here related to the match calculations. Recall that the NeT model is evolved from an initial eccentricity and semi-latus rectum (e_0, p_0) to the LSO, which corresponds to a mean orbital frequency F_0 to F_{LSO}. This time series is then zero padded, such that the length of the new time series is a power of 2, with the result then discrete Fourier transformed. The NeF model, on the other hand, is sampled at the frequencies associated with the discrete Fourier transform of the NeT model. In order to do, so we numerically invert $F(e)$ in order to map the Fourier frequency, f, to the mean orbital frequency, F, which is then further mapped to the orbital eccentricity, e, which appears explicitly in the harmonic amplitudes and phases.

Once both models are sampled we compute the match with all integrals over frequency evaluated as discrete sums from $f \in [0, f_{ny}]$, and then maximized as described in Sec. 5.3. We do not window the NeF model, but we do use a square window for the NeT model. Two independent versions of this code were implemented, one in `Mathematica`, and one in `C` and found to agree in the match to within $\sim 0.1\%$ absolute error. We take this to be an estimate of the numerical error of the match presented here.

Figure 5.11 shows the match for two of the three systems considered in the

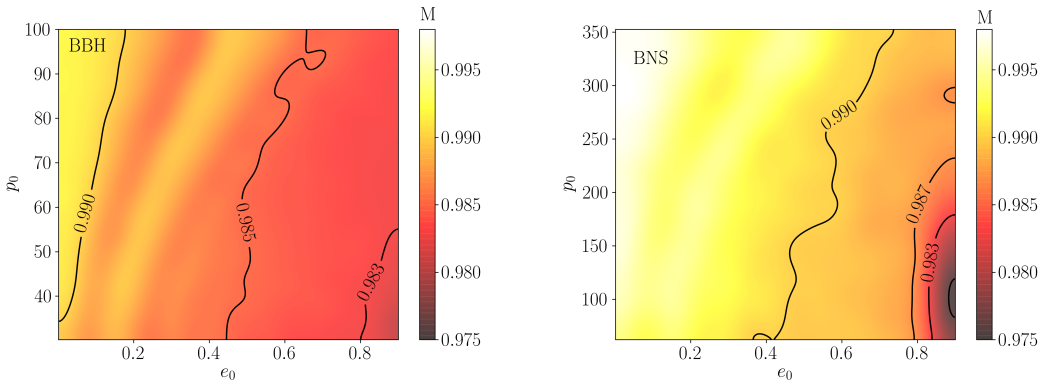

Figure 5.11: Match between the NeT model described in Sec. 5.4.3 and the NeF model as described in Sec. 5.4.2 for a BBH (top) and BNS (bottom) system (a BH-NS case is presented in Fig. 5.1). The match is given as a function of the initial condition (e_0, p_0) on the x and y axis. The maximum match is 0.99, and for a vast majority of the parameter space, NeF very faithfully represents NeT. For the shorter lived source (BBH), the match behaves more poorly than its longer lived counterparts due to finite time effects. The boundaries of the fringes correspond to a boundary where the model demands including more harmonics and thus the match displays a sharp increase as the next harmonic is included.

last section (BNS, BBH, BH-NS), for different initial eccentricities and dimensionless semilatera recta. Figure 5.1 shows the corresponding figure for the BH-NS case. All $\sim 2,500$ matches are Gaussian filtered in order to obtain smooth contours in the figure. In reality, some of the matches presented in Fig. 5.11 may be above or below the value of the corresponding contour, but the contours represent the general features.

Observe that the NeF model matches the NeT model extremely well in nearly all cases. As the initial eccentricity increases and the initial semi-latus rectum decreases, however, the match begins to deteriorate from 0.99 (which is nearly perfect due to the truncation of the harmonic sum) to 0.985 and 0.98. This is the result of finite time effects in the DFT of NeT and windowing. Finite time effects result in a loss of match (see [228] and [229]) and this is exacerbated by the fact that eccentric binaries

evolve faster than their quasi-circular counterparts (for the same initial mean orbital frequency). This reasoning is supported by the fact that the match is higher for longer lived sources (BH-NS and BNS) than shorter lived ones (BBH) for the same initial eccentricity and mean orbital frequency. If one were interested in a model with a higher value of the match (i.e. one faithful for higher SNR sources), one could include more harmonics in the SPA sum to obtain a higher match than that required in Sec. 5.6.1. One could also attempt to incorporate finite time effects into the SPA itself, as done in [229] or minimize finite time effects by applying the same window functions to both models.

In the BNS system, the match begins to drop steeply to as low as 0.975 near $e_0 \sim 0.85$ even for relatively large initial separations. The loss of match here is not due to finite time effects, but rather it is due to the inaccuracy of the approximation to the Fourier phase in Eq. (5.59). One could improve this match by increasing the number of terms kept in the Chebyshev resummation and in the Taylor expansion. Alternatively, one could choose a different number of terms in the representation of the Fourier phase for sources with different lifetimes, as longer lived sources are more sensitive to error in the representation of the phase. For shorter lived sources, the phase could likely be approximated even more cheaply than done here. We leave a more thorough investigation of such an implementation to future work.

The sudden jumps and drops in the value of the match that appear most prominently near the low eccentricity region of Fig. 5.11 are due to the harmonic nature of the model. For a fixed number of harmonics kept, the value of the match begins to decrease as the initial eccentricity increases. When the model demands the inclusion of another harmonic (informed by the results of Sec. 5.6.1), the value of the match sharply increases. As such, the sudden jumps correspond to the contours of $j_{\max}$ shown in Fig. 5.9.

5.7 Conclusions & Future Work

We have taken the first steps toward the construction of a Fourier domain model that is valid for parameter estimation to arbitrary eccentricity. In particular, we explicitly calculated an efficient analytic frequency domain model to Newtonian order, the NeF model, which we show is faithful to very high initial eccentricity ($e_0 \sim 0.9$) through match calculations against a numerical time-domain model to Newtonian order, the NeT model (see Fig. 5.1 and 5.11). As byproducts of this analysis, we also derived new techniques to rapidly maximize the match over phase and time at coalescence for eccentric signals, and we analyzed the most important source of errors in the NeF model. Given its analytic features, we have analytic control over the NeF model, which thus allows us to straightforwardly improve it by keeping higher order terms in the analytic expansions.

As a proof-of-concept, we have here presented the NeF model to leading PN order, so it is not yet useful for data analysis. Even once higher PN terms are included, the model as presented here is likely still not accurate enough for a LISA-like detector, which will require a higher match for high signal-to-noise ratio events. Both of these problems, however, can be straightforwardly solved by (i) extending this NeF model to higher PN order and (ii) retaining more terms in the Chebyshev representation of the Fourier phase and more harmonics in the SPA amplitude. All of this future work will make for very interesting extensions of the NeF model.

The least straightforward extension is the inclusion of higher PN order terms. Much of the work required for this, however, already exists. The fluxes and time derivatives of the orbital elements have been computed in [230–233] to 3PN order. The time domain harmonic decomposition becomes troublesome at higher PN order, as Kepler's equation are modified at 2PN order. A method to treat the time domain

harmonic decomposition at 3PN, however, is introduced in [234]. Therefore, the extension of the NeF model to higher PN order requires the careful assembly of pieces that are already present in the literature.

With a higher PN waveform in hand, one could in the future conduct many useful studies. For example, one could investigate the ability of aLIGO, LISA, or 3G detectors to measure the parameters and any biases of the model given eccentric signals. Since eccentric binaries are thought to be the result of specific astrophysical scenarios, it could be possible to study whether the detection of eccentric GWs can constrain different formation scenarios by recovering e_0 and F_0.

Another useful future study is to incorporate effects of modified gravity theories into the new model. If a modified gravity theory modifies the eccentricity evolution of a binary, such a theory could be constrained more stringently than with quasi-circular events only. This is because eccentricity excites additional harmonics in the signal, which contain useful information that could be leveraged to break degeneracies and constrain General Relativity. One such approach has already been taken by extending the parameterized post-Einsteinian (ppE) framework [235] to high-eccentricity bursts [236].

Perhaps the most difficult and rewarding goal is an inspiral-merger-ringdown (IMR) type hybrid that incorporates the effects of orbital eccentricity of moderate and large magnitude. Currently, the only IMR model that incorporates orbital eccentricity is that of [237]. This model is constructed in the time domain using PN theory, results of numerical relativity, and a machine learning algorithm to tune the model. It would be interesting to compare that model to this future hybrid model, once one is developed.

Acknowledgments

B.M. was supported by the Joan L. Dalton Memorial Fellowship in Astronomy from Montclair State University. T. R. appreciates the support of the NASA grant NNX16AB98G. N. Y. acknowledges support from NASA grants NNX16AB98G and 80NSSC17M0041. Computational efforts were performed on the Hyalite High Performance Computing system, which is supported by University Information Technology at Montana State University.

5.A Application of the Stationary Phase Approximation

In this Appendix we review the application of the stationary phase approximation (SPA) as applied to the time domain harmonic decomposition of the GW signal given in Eq. (5.6). For a more detailed and general presentation of the approximation see [238]. We begin by considering the time domain GW signal generated by the $j^{\rm th}$ harmonic:

$$h_j(t) = \left[A_j(t)e^{-ijl(t)} + \text{c.c.}\right]\Theta(F(t) - F_0)\Theta(F_{\rm LSO} - F(t))\,, \tag{5.73}$$

where c.c. stands for the complex conjugate and the amplitude is given by

$$A_j(t) = -\frac{m\eta}{2R}(2\pi mF)^{2/3}\left[C^{(j)}_{+,\times}(t) + iS^{(j)}_{+,\times}(t)\right]\,. \tag{5.74}$$

The unit step function, $\Theta(x)$, arises because the binary is assumed to be formed at a time t_0 with mean orbital frequency F_0 and the emission is terminated at the LSO when the binary has mean orbital frequency $F_{\rm LSO}$ (as is customary in the GW literature).

We wish to approximate the Fourier transform of the GW signal given by

$$\tilde{h}_j(f) = \int_{-\infty}^{\infty} \left\{ A_j(t) e^{i[2\pi f t - j l(t)]} + A_j^*(t) e^{i[2\pi f t + j l(t)]} \right\} \times \Theta\left(F(t) - F_0\right) \Theta\left(F_{\rm LSO} - F(t)\right) dt \,. \tag{5.75}$$

In the above integrand, the amplitude varies much more slowly than the mean anomaly ($\dot{A}_j/A_j \ll j\dot{l}$) appearing in the phase. Thus, for most values of t this integrand is rapidly oscillating. However, there exists a time where the phase of the integrand is roughly constant and thus strongly contributes to the integral. This time is called the stationary time (denoted t_j^*) and it occurs when the stationary phase condition is satisfied, i.e. when the first derivative of the argument of the complex exponential vanishes. Thus, the stationary phase condition is

$$2\pi f \mp j\dot{l}(t_j^*) = 0 \,, \tag{5.76}$$

where the minus sign corresponds to the stationary phase condition of the first term in the integrand, and the plus corresponds to the second. Since the time derivative of the mean anomaly is proportional to the mean orbital frequency, a positive quantity, the stationary phase condition is never satisfied for the second term in the integrand and so it does not contribute strongly to the integral. The stationary phase condition that is satisfied provides a mapping between the Fourier frequency f and the mean orbital frequency F for the $j^{\rm th}$ harmonic: $f = jF$.

To approximate the integral in Eq. (5.75), we Taylor expand the amplitude to leading order and the phase to first order (as the amplitude is more slowly varying than the phase) about the stationary time and evaluate the unit step functions at the

stationary time:

$$\tilde{h}_j(f) \approx A_j(t_j^*) e^{i[2\pi f t_j^* - jl(t_j^*)]} \Theta(f - jF_0)\Theta(jF_{\rm LSO} - f) \times \int_{-\infty}^{\infty} \exp\{-ij\frac{\ddot{l}(t_j^*)}{2}(t - t_j^*)^2\} dt\,. \tag{5.77}$$

We carry out the above integral and obtain

$$\tilde{h}_j(f) \approx A_j(t_j^*)\sqrt{\frac{2\pi}{|j\ddot{l}(t_j^*)|}} e^{i(2\pi f t_j^* - jl(t_j^*) - \pi/4)} \times \Theta\left(f - jF_0\right)\Theta\left(jF_{\rm LSO} - f\right)\,. \tag{5.78}$$

Rewriting the above in terms of more familiar quantities we have

$$\tilde{h}_j(f) \approx -\frac{m\eta}{2R}\frac{(2\pi m F(t_j^*))^{2/3}}{\sqrt{j\dot{F}(t_j^*)}}\left[C_{+,\times}^{(j)}(t_j^*) + iS_{+,\times}^{(j)}(t_j^*)\right] \times e^{i\psi_j}\Theta\left(f - jF_0\right)\Theta\left(jF_{\rm LSO} - f\right)\,, \tag{5.79}$$

with

$$\psi_j = 2\pi f t_j^* - jl(t_j^*) - \pi/4. \tag{5.80}$$

To recover the full frequency response due to many harmonics, we simply sum the above over the harmonic index j.

5.B Chebyshev Resummation

In this appendix we review the approximation of the Fourier phase and provide the coefficients of the resulting series in eccentricity to 16 digits of precision. Since the exact coefficients are unwieldy rationals, they are provided in a supplementary

notebook. We begin by approximating $I(e)$ in Eq. (5.58) through a Maclaurin series

$$I(e) \approx \sum_{n=0}^{n_{\max}} A_{2n} e^{2n}. \tag{5.81}$$

We use an even basis function for the expansion about small eccentricity because $I(e)$ is even. The coefficients A_n are simply the coefficients of the Taylor expansion, namely

$$A_n = \frac{1}{n!} \left. \frac{d^n I}{de^n} \right|_{e=0} . \tag{5.82}$$

The first seventeen non-zero A_n are given by

$$\begin{aligned}
A_0 &= -0.237\,500\,000\,000\,000\,0 \ , \\
A_2 &= -0.300\,615\,876\,322\,989\,4 \ , \\
A_4 &= -0.008\,991\,117\,053\,859\,665 \ , \\
A_6 &= -0.037\,851\,834\,031\,183\,92 \ , \\
A_8 &= -0.007\,615\,853\,560\,080\,763 \ , \\
A_{10} &= -0.009\,769\,540\,332\,607\,544 \ , \\
A_{12} &= -0.004\,556\,366\,050\,156\,979 \ , \\
A_{14} &= -0.003\,968\,026\,824\,286\,266 \ , \\
A_{16} &= -0.002\,630\,835\,486\,794\,598 \ , \\
A_{18} &= -0.002\,115\,104\,845\,741\,527 \ , \\
A_{20} &= -0.001\,610\,593\,129\,877\,186 \ , \\
A_{22} &= -0.001\,304\,358\,739\,038\,224 \ , \\
A_{24} &= -0.001\,055\,578\,416\,146\,812 \ , \\
A_{26} &= -0.000\,875\,581\,422\,324\,481\,2 \ ,
\end{aligned}$$

$$A_{28} = -0.000\,732\,963\,908\,943\,822\,6 \ ,$$
$$A_{30} = -0.000\,621\,975\,182\,560\,585\,6 \ ,$$
$$A_{32} = -0.000\,532\,723\,793\,460\,911\,2 \ . \tag{5.83}$$

We then wish to resum this Taylor expansion using Chebyshev polynomials (see [239] for a discussion of the convergence properties of Chebyshev series). The Chebyshev polynomials, $T_n(x)$, are defined for $x \in [-1, 1]$, and thus, to make them a function of the eccentricity and preserve their orthogonality we rescale them as $T_n(s)$ where $s = 2e^2 - 1$, which guarantees $s \in [-1, 1]$ for $e \in [0, 1]$. The first few Chebyshev polynomials are

$$T_0(s) = 1 \ , \tag{5.84a}$$
$$T_1(s) = 2e^2 - 1 \ , \tag{5.84b}$$
$$T_2(s) = 8e^4 - 8e^2 + 1 \ , \tag{5.84c}$$
$$T_3(s) = 32e^6 - 48e^4 + 18e^2 - 1 \ , \tag{5.84d}$$

which can be inverted to find different powers of the eccentricity as functions of Chebyshev polynomials

$$e^0 = T_0(s) \ , \tag{5.85a}$$
$$e^2 = \frac{1}{2}(T_1(s) + T_0(s)) \ , \tag{5.85b}$$
$$e^4 = \frac{1}{8}(T_2(s) + 4T_1(s) + 3T_0(s)) \ , \tag{5.85c}$$
$$e^6 = \frac{1}{32}(T_3(s) + 6T_2(s) + 15T_1(s) + 10T_0(s)) \ . \tag{5.85d}$$

We now seek a resummation of Eq. (5.81) into the form

$$I(e) \approx \sum_{k=0}^{k_{\max}} B_k T_k(s)\,, \tag{5.86}$$

for some set of coefficients B_k. We do so by inserting the monomials of e in terms of the Chebyshev polynomials, Eq. (5.85), into Eq. (5.81) and collecting terms order by order in the Chebyshev polynomials. The first terms for B_k are

$$\begin{aligned}
B_0 &= -0.411\,345\,910\,752\,926\,6\ , \\
B_1 &= -0.186\,724\,296\,957\,817\,7\ , \\
B_2 &= -0.016\,970\,870\,078\,114\,17\ , \\
B_3 &= -0.005\,185\,894\,691\,512\,303\ , \\
B_4 &= -0.001\,448\,030\,529\,666\,706\ , \\
B_5 &= -0.000\,467\,082\,998\,906\,841\,1\ , \\
B_6 &= -0.000\,146\,724\,477\,933\,540\,0\ , \\
B_7 &= -0.000\,043\,850\,639\,900\,567\,47\ , \\
B_8 &= -0.000\,011\,991\,016\,195\,612\,29\ , \\
B_9 &= -2.925\,622\,159\,845\,224 \times 10^{-6}\ , \\
B_{10} &= -6.210\,389\,911\,462\,248 \times 10^{-7}\ , \\
B_{11} &= -1.117\,172\,525\,811\,035 \times 10^{-7}\ , \\
B_{12} &= -1.649\,269\,175\,725\,618 \times 10^{-8}\ , \\
B_{13} &= -1.913\,379\,819\,810\,920 \times 10^{-9}\ , \\
B_{14} &= -1.632\,587\,267\,118\,429 \times 10^{-10}\ , \\
B_{15} &= -9.096\,722\,174\,897\,557 \times 10^{-12}\ ,
\end{aligned}$$

$$B_{16} = -2.480\,688\,474\,424\,701 \times 10^{-13} \ . \tag{5.87}$$

Clearly if the Maclaurin series is truncated at $n_{\max}$, and the Chebyshev series is truncated at the same order, then the two approximations are identical. However, we find that we can keep less terms in the Chebyshev series than the Maclaurin series to accurately approximate the phase. Specifically, we find that we obtain a sufficiently accurate representation of $I(e)$ if we use a Maclaurin series truncated at $n_{\max} = 16$ and a Chebyshev resummation truncated at $k_{\max} = 12$.

After truncating the Chebyshev series in Eq. (5.86) at $k_{\max} = 12$, it is more computationally efficient to collect like terms in eccentricity when implementing this function, lest one evaluates the same power of eccentricity many times. After collecting the like terms in eccentricity we are left with

$$I(e) \approx \sum_{n=0}^{n=12} C_n e^{2n}. \tag{5.88}$$

We now provide the coefficients $C_{n>5}$ in decimal form to 16 digits:

$$\begin{aligned}
C_5 &= 0.041\,993\,340\,084\,657\,63 \ , \\
C_6 &= -0.223\,130\,740\,649\,105\,5 \ , \\
C_7 &= 0.608\,665\,210\,424\,159\,5 \ , \\
C_8 &= -1.161\,566\,561\,696\,015 \ , \\
C_9 &= 1.469\,991\,268\,471\,017 \ , \\
C_{10} &= -1.216\,042\,897\,081\,895 \ , \\
C_{11} &= 0.595\,816\,296\,413\,753\,2 \ , \\
C_{12} &= -0.138\,350\,726\,016\,453\,3 \ .
\end{aligned} \tag{5.89}$$

The coefficients $C_{n\in(1,4)}$ were already provided in Eq. (5.60). All of the A_n, B_k, and C_n are provided in their exact rational form in a supplemental notebook.

5.C Inversion of F(e)

Let us review our attempts at inverting

$$\sigma(e) = \zeta, \tag{5.90}$$

where

$$\sigma(e) = \frac{e^{12/19}}{1-e^2}\left(1+\frac{121}{304}e^2\right)^{870/2299}. \tag{5.91}$$

Recall that $\zeta = \sigma(e_0)(F_0/F)$ and sources with initial conditions (e_0, F_0) will sample values of $\zeta \in [\sigma(e_0)(F_0/F_{\rm LSO}), \sigma(e_0)]$. In Sec. 5.5.2 we invert Eq. (5.90) in the low eccentricity and late frequency regime (i.e. $\zeta \ll 1$) and we find that this inversion leads to a significant loss in match for sources which sample $\zeta \sim 0.5$ which corresponds to sources with $e_0 \sim 0.3$. Since the focus of this work is to provide a model which is useful for parameter estimation (i.e. with matches ~ 0.99) and valid for arbitrary eccentricity, we must investigate other representations of $e(\zeta)$.

Two models for $e(\zeta)$ meet the accuracy goals laid out in Sec. 5.5: (i) a model which is composed of a controlling factor obtained by introducing an approximate $\sigma(e)$ into Eq. (5.90) and algebraically solving and then fitting the remaining error and (ii) a piecewise representation of $e(\zeta)$, which is composed of two Taylor expansions about $\zeta \ll 1$ and $\zeta \gg 1$, and an efficient numerical fit of the function in the range in ζ for which the error in the Taylor expansions is too large for faithful modeling. The piecewise representation is the best of the two methods in speed, accuracy, and domain of validity. All of the coefficients listed in this Appendix are provided in a

supplementary notebook.

As a first attempt to solve Eq. (5.90), we introduce

$$\bar{\sigma}(e) = \zeta, \tag{5.92}$$

where $\bar{\sigma}(e)$ is a simpler rational polynomial similar to $\sigma(e)$ which admits an exact

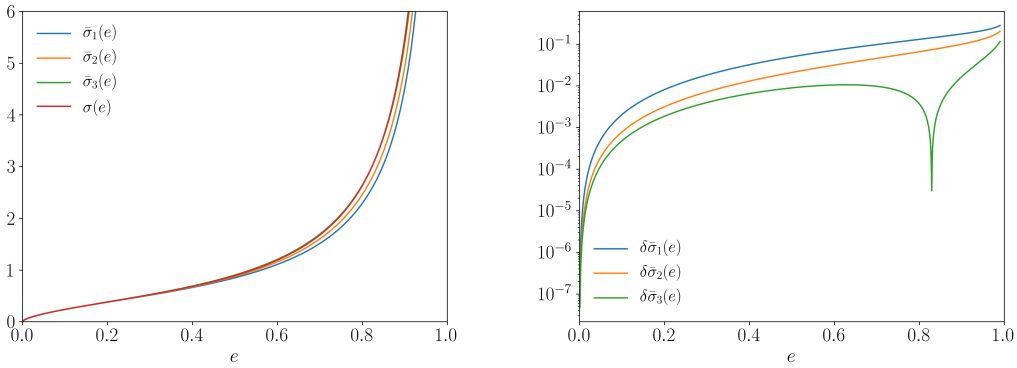

Figure 5.12: The value of $\sigma(e)$, $\bar{\sigma}_1(e)$, $\bar{\sigma}_2(e)$, $\bar{\sigma}_3(e)$ (left), and their relative errors (right). Observe that $\bar{\sigma}_3(e)$ most closely represents $\sigma(e)$.

solution. Let us define the inverse function $\bar{\kappa}(\zeta)$ such that $\bar{\kappa}[\bar{\sigma}(e)] = e$. We then seek to leverage this inverse function, which we can solve for exactly, to find an approximate solution for $e(\zeta)$ by multiplying by a Taylor series:

$$e(\zeta) \approx \bar{\kappa}(\zeta) \left(\sum_{k=0}^{k_{\max}} D_k (\zeta - \zeta_0)^k \right) . \tag{5.93}$$

The factor which is raised to the 870/2299 power in Eq. (5.91) varies from 1 to 1.135, so we first neglect this term. With this term set to 1, we observe that the resulting equation can be solved if the factor of $(1-e^2)$ appearing in the denominator is raised to the power 18/19. We then raise the entire resulting equation to the power 19/6 and we are left with a polynomial equation which can be readily algebraically solved.

This corresponds to a $\bar{\sigma}(e)$

$$\bar{\sigma}_1(e) = \frac{e^{12/19}}{(1-e^2)^{18/19}}\,. \tag{5.94}$$

We further note that we can incorporate the factor of $(1+\frac{121}{304}e^2)^{870/2299}$ if we approximate it as $(1+\frac{121}{304}e^2)^{3n/19}$. We generate two more $\bar{\sigma}(e)$ with different choices of n

$$\bar{\sigma}_2(e) = \frac{e^{12/19}}{(1-e^2)^{18/19}}\left(1+\frac{121}{304}e^2\right)^{6/19}, \tag{5.95}$$

$$\bar{\sigma}_3(e) = \frac{e^{12/19}}{(1-e^2)^{18/19}}\left(1+\frac{121}{304}e^2\right)^{12/19}. \tag{5.96}$$

$$\tag{5.97}$$

In Fig. 5.12 we plot the values of $\sigma(e)$, the approximate $\bar{\sigma}(e)$, and the relative error associated with the $\bar{\sigma}(e)$. Since $\bar{\sigma}_3(e)$ most closely represents $\sigma(e)$, we provide its inverse function in Eq. (5.63) of Sec. 5.4.2. In Fig. 5.13 we plot a numerical solution $\kappa(\zeta)$, $\bar{\kappa}_3(\zeta)$, and the associated relative error for $\zeta \in [0, 5.47]$ which corresponds to systems with e_0 as high as 0.9. As shown in Sec. 5.5, we require relative error of $\mathcal{O}(10^{-6})$ in order to faithfully model sources with arbitrary eccentricity. The relative error between $\bar{\kappa}_3(\zeta)$ and the numerical solution is $\mathcal{O}(10^{-3})$, so we conclude that we must further improve the inversion if we wish to have something accurate enough for our purposes.

In order to decrease the relative error further, we multiply $\bar{\kappa}_3(\zeta)$ by a Taylor series as written in Eq. (5.93). The coefficients D_k are determined by substituting this $e(\zeta)$ into $\sigma[e(\zeta)]$ and demanding $\sigma[e(\zeta)] = \zeta$ upon Taylor expansion. After trying many Taylor expansions about different ζ_0, we find that the resulting solution behaves better than $\bar{\kappa}_3(\zeta)$ in a small region near ζ_0, but leads to considerably more error than

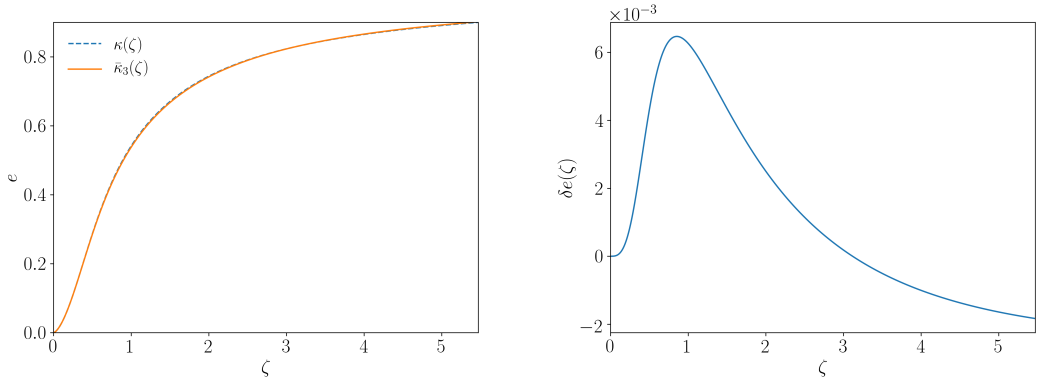

Figure 5.13: The value of a numerical solution for $\kappa(\zeta)$, $\bar{\kappa}_3(\zeta)$ as given in Eq. (5.63) (left), and the relative error between the exact solution and $\bar{\kappa}_3(\zeta)$ (right). We see that the two inversions are nearly indistinguishable by eye, but the relative error is $\mathcal{O}(10^{-3})$.

$\bar{\kappa}_3(\zeta)$ for most of the domain. We conclude that this approach does not lead to an accurate enough inversion of $F(e)$ to be useful for applications that require a high value of the match.

CHAPTER SIX

DETECTING GRAVITATIONAL WAVE BURSTS WITH LISA IN THE PRESENCE OF INSTRUMENTAL GLITCHES

Contributions of Authors and Co-Authors

Manuscript in Chapter 6

Author: Travis Robson

Contributions: Developed model and code for study. Wrote initial draft of manuscript.

Co-Author: Dr. Neil Cornish

Contributions: Conceived of the study design. Edited manuscript.

Manuscript Information Page

Travis Robson, Neil Cornish

Physical Review D

Status of Manuscript:

Submitted to a peer-reviewed journal

Abstract

The Laser Interferometer Space Antenna (LISA) will open a rich discovery space in the milli-Hertz gravitational wave band. In addition to the anticipated signals from many millions of binary systems, this band may contain new and previously un-imagined sources for which we currently have no models. To detect unmodeled and unexpected signals we need to be able to separate them from instrumental noise artifacts, or glitches. Glitches are a regular feature in the data from ground based laser interferometers, and they were also seen in data from the LISA Pathfinder mission. In contrast to the situation on ground, we will not have the luxury of having multiple independent detectors to help separate unmodeled signals from glitches, and new techniques have to be developed. Here we show that unmodeled gravitational wave bursts can be detected with LISA by leveraging the different way in which instrument glitches and gravitational wave bursts imprint themselves in the time-delay interferometery data channels. We show that for signals with periods longer than the light travel time between the spacecraft, the "breathing mode" or Sagnac data combination is key to detection. Conversely, for short period signals it is the time of arrival at each spacecraft that aids separation. We investigate the conditions under which we can distinguish the origin of signals and glitches consisting of a single sine-Gaussian wavelet and determine how well we can characterize the signal. We find that gravitational wave bursts can be unambiguously detected and characterized with just a single data channel (four functioning laser links), though the signal separation and parameter estimation improve significantly when all six laser links are operational.

6.1 Introduction

Gravitational wave astronomy has tremendous potential for discovery, as has been spectacularly demonstrated by the ground-based LIGO/Virgo observatories [1, 6]. The signals that have been detected to date have all been from binary systems, and are accurately modeled by theoretical templates. Going forward, it is hoped that entirely new classes of signals will be discovered, many of which we will not have templates for, either due to the difficulty in calculating the waveform (such as for supernovae), or from our ignorance about the existence of the source. Detecting signals of unknown morphology is challenging since the instruments themselves produce non-Gaussian transients, or glitches, that can be mistaken for signals of astrophysical origin.

The Laser Interferometer Space Antenna (LISA) [19], like its ground based cousins, will very likely be afflicted by glitches. Glitches were seen in data from the LISA Pathfinder mission [20, 21], and it is hard to imagine that they will be absent from the more complex LISA measurement system. Characterizing these glitches and accurately estimating their waveforms will be an important component of the LISA global data analysis program. Unlike the situation on ground, where the availability of multiple independent interferometers simplifies the task of separating glitches from signals, with LISA we will have a single instrument. Nor will we have any "off-source" data, free of loud gravitational wave signals, with which to perform a measurement of the instrument noise. With LISA the signal and noise measurement must be done simultaneously [60] as part of a global analysis.

Similar concerns led to the developments of burst and glitch characterization analyses for LIGO. One such analysis was the wavelet-based Bayesian algorithm BayesWave [43]. This algorithm has played the key role for model-independent

waveform reconstructions for most of the detected mergers seen by LIGO [1]. Its broad capabilities were best demonstrated by the binary neutron star merger. BayesWave's ability to characterize a loud instrumental glitch, obscuring a large fraction of the all-important late inspiral, allowed for an accurate reconstruction of a the astrophysical signal with the glitch removed [240], so that other analyses could properly characterize the binary neutron star's physical parameters [6]. For LISA, we wish to develop an algorithm to serve a similar purpose of analyzing glitches and bursts. Instrumental glitches in LISA studied here will fall into two categories: optical path and acceleration. Optical path glitches reflect non-Gaussian deviations in the optical path length of any of LISA's 6 laser links. Acceleration glitches result from disturbance to the acceleration of LISA's spacecraft. Laser phase noise glitches will be neglected in this study since they will be supressed in the time delay interferometry (TDI) data channels [241]. Glitches can be represented through a superposition of sine-Gaussian wavelets in each component of the instrument. Gravitational wave bursts can similarly be represented by a superposition of wavelets. The signal is referenced to the solar system barycenter (SSB) and then projected onto LISA by computing the instrument response. As a first step, we consider signals and glitches that are described by a single wavelet and defer the generalization to multi-wavelet fits to future work.

To investigate our ability to characterize glitches and bursts consisting of the single wavelet, we will use Bayesian probability theory to calculate our degree of belief in the parameters which describe the injected signal, as quantified by the posterior distribution. The duration of these signals ranges from tens of seconds to roughly a day. Their duration and frequency content, in relation to the light travel time between the spacecraft, will have important implications on our ability to characterize these signals and will also play a key role in our ability to distinguish whether the

data contains a glitch or burst. Glitches will enter the time delay interferometry (TDI) data channels with time delays of the light travel time between spacecraft. This time is about 8.3 seconds for the nominal $L = 2.5$ Gm separation, which sets the LISA response transfer frequency $f_* = c/(2\pi L)$ to be 19.1 mHz. Wavelets with frequencies below the transfer frequency will be harder to characterize and distinguish. An additional piece of the puzzle is which data channels the wavelets power crops up and in what proportion. Acceleration glitches enter the data stream by afflicting 2 different phase measurements while optical path glitches afflict only 1. Bursts, on the other hand, enter all phase measurements through time delays which depend on the various projected arm lengths depending on where the incident gravitational wave originates on the sky. While the power distribution is the most useful discriminant, the phasing becomes most important in the case of a malfunctioning LISA arm *i.e.* when when we would be left with only one data channel.

This work is organized as follows: Section 6.2 discusses the the waveforms for optical path and acceleration glitches and for gravitational wave bursts. Section 6.3 reviews Bayesian inference and then describes the Markov Chain Monte Carlo algorithm we employ to carry out the parameter estimation and model selection analyses in this paper. Section 6.4 shows how well we can characterize glitch and burst parameters and recover the injected waveform. In section 6.5 we explore under what conditions we are able to distinguish a glitch and burst, and identify what features of the signal are most responsible for making the distinction. We end with a discussion of future work to be carried out in section 6.6. Note that we work in units where $G = c = 1$.

6.2 Glitch and Burst Models

The LISA constellation consists of 3 spacecraft in the shape of a quasi-equilateral triangle trailing behind Earth. The spacecraft have a total of 6 laser links, 2 for each arm. The phase of each laser is measured by phasemeters onboard of the LISA spacecraft. A photon sent from spacecraft i, to spacecraft j is emitted at time $t - L_{ij}$, where L_{ij} is arm length connecting spacecraft i and j. The phase of this photon is measured at time t by the phasemeter on spacecraft j. This phase measurement can be approximated as [78]

$$\begin{aligned}\Phi_{ij}(t) = C_i(t - \mathrm{L}_{ij}) - C_j(t) + \psi_{ij}(t) + n^{\mathrm{o}}_{ij}(t) \\ - \hat{\mathbf{r}}_{ij}(t) \cdot \left(\mathbf{n}^{\mathrm{a}}_{ji}(t - L_{ij}) - \mathbf{n}^{\mathrm{a}}_{ij}(t)\right) \ .\end{aligned} \tag{6.1}$$

The noise in the laser phase itself is described by the terms C_i. The term ψ_{ij} describes the phase shift induced by the presence of gravitational waves. The term n^{o}_{ij} represents the contribution from the optical bench on spacecraft j that receives light from spacecraft i. The last term represents the contribution to the phase measurement incurred by the acceleration noise of the spacecraft. Note that in this simplified model, where we are neglecting higher order features of LISA's motion such as the flexing of the arms, the only component of the acceleration that is relevant is the differential acceleration along the line $\hat{\mathbf{r}}_{ij}$ connecting the center of mass of the two spacecraft.

These laser phase measurements are expected to be dominated by laser phase noise. Current estimates indicate that the laser phase noise will be roughly 10 orders of magnitude greater that the changes in phase induced by the gravitational waves of

interest [241]. The phase noise can be canceled using time delay interferometry (TDI). The TDI data combinations synthesize light paths of equal length by adding together phase measurements with time delays given by multiples of the instantaneous light travel times. This superposition cancels the laser phase noise. When higher order corrections to the spacecraft motion are taken into account the superposition of time delayed phase measurements become more complicated. Here we use the simpler first generation TDI data combinations. Three Michelson-like TDI channels can be formed from the signals extracted at each vertex of the observatory. These are denoted as X, Y, and Z. The X TDI channel is constructed as follows:

$$\begin{aligned} X(t) = \Phi_{12}(t - 3\mathrm{L}) - \Phi_{13}(t - 3\mathrm{L}) + \Phi_{21}(t - 2\mathrm{L}) \\ - \Phi_{31}(t - 2\mathrm{L}) + \Phi_{13}(t - \mathrm{L}) - \Phi_{12}(t - \mathrm{L}) \\ + \Phi_{31}(t) - \Phi_{21}(t) \ , \end{aligned} \tag{6.2}$$

where we have assumed that the LISA arm lengths are of constant length, i.e. $L_{ij}(t) = L$. The Y and Z channels are constructed through a cyclic permutation of the spacecraft labels—e.g. $1 \to 2$, $2 \to 3$, and $3 \to 1$ to construct the Y TDI channel. It is often convenient to work with following linear combinations of the X, Y, Z channels:

$$A = \frac{1}{3}(2X - Y - Z) \tag{6.3a}$$

$$E = \frac{1}{\sqrt{3}}(Z - Y) \tag{6.3b}$$

$$T = \frac{1}{3}(X + Y + Z) \ . \tag{6.3c}$$

Below the transfer frequency, the A and E channels synthesize two right angle interferometers with a relative orientation of 45°, and provide instantaneous measures

of the plus and cross polarization states of a gravitational wave. At these frequencies the T channel is mostly sensitive to the scalar breathing mode polarization state, which is absent in Einstein gravity, and thus provides a null channel that is useful for measuring a particular combination of the noise contributions. When the noise levels are equal on each spacecraft, the cross-spectral density of the noise in the A, E, T channels vanish [242, 243].

An arbitrary signal seen in the TDI data channels may be reconstructed by a superposition of sine-Gaussian wavelets. In this study we use Gabor-Morlet wavelets. In the time-domain they are given by

$$\Psi = Ae^{-(t-t_0)^2/\tau^2} \cos\left[2\pi f_0(t-t_0) + \phi_0\right] \ , \tag{6.4}$$

where A is the wavelet amplitude, t_0 and f_0 are the central time and frequency, the wavelet time scale is τ—related to the wavelet quality factor Q through $\tau = Q/2\pi f_0$—and ϕ_0 is the initial phase. Occasionally we will use the variable $\bar{\phi} = \phi_0 - 2\pi f_0 t_0$. The Fourier transform of the Gabor-Morlet wavelet is

$$\begin{aligned}\tilde{\Psi} = &\frac{\sqrt{\pi}\tau A}{2} e^{-i(2\pi f t_0 + \phi_0)} \\ &\times \left[e^{-(\pi\tau(f+f_0))^2} + e^{2i\phi_0} e^{-(\pi\tau(f-f_0))^2}\right] \ .\end{aligned} \tag{6.5}$$

In the Fourier domain we see in the large quality factor Q, or equivalently large τ regime, the second term in eqn. (6.5) is dominant. Ignoring the sub-dominant term, we can estimate the signal-to-noise ratio (SNR) in the case of white noise as

$$\rho^2 \approx \frac{4}{S_n(f_0)} \int_0^\infty \left(\frac{\sqrt{\pi}\tau A}{2}\right)^2 e^{-2(\pi\tau(f-f_0))^2} df$$

$$\approx \sqrt{\frac{\pi}{2}} \frac{A^2 \tau}{S_n(f_0)} \,, \tag{6.6}$$

where $S_n(f_0)$ is an appropriate noise power spectral density which has been assumed constant such that we may approximate the integral. This result will become useful later when we wish to estimate a reasonable bandwidth in the frequency domain to calculate these signals over.

6.2.1 Instrumental Glitches

To model instrumental glitches we inject a Gabor-Morlet wavelet into the appropriate term in eqn. (6.1). For example, a glitch in the optical path length pointing from spacecraft 1 to 2 is modeled as $\Phi_{12}(t) = n^{\mathrm{o}}_{12}(t) = \Psi(t)$. We will label such a glitch as Φ^{op}_{12}. An acceleration glitch associated with the proof mass on spacecraft 2 that is referenced against spacecraft 1 will appear in to phase measurements: $\Phi_{12}(t) = -\Psi(t)$ and $\Phi_{21}(t) = \Psi(t - L)$. This acceleration glitch will be denoted as Φ^{ac}_{12}. Laser phase glitches are neglected in this work since the TDI channels are constructed such that laser phase noise is canceled.

The X, Y, and Z TDI channels can be constructed for both optical path and accelerations glitches analytically in the frequency domain. For the optical path glitch Φ^{op}_{12} the response is

$$\tilde{X} = 2i\tilde{\Psi} e^{-2if/f_*} \sin \frac{f}{f_*} \tag{6.7a}$$

$$\tilde{Y} = -2i\tilde{\Psi} e^{-if/f_*} \sin \frac{f}{f_*} \tag{6.7b}$$

$$\tilde{Z} = 0 \,. \tag{6.7c}$$

Note that there is no response in the Z channel. The factor of $\sin f/f_*$ is due to differencing the disturbance by the time delay. The only other optical path glitch

that has no response in the Z channel is Φ_{21}^{op}, which produces the response

$$\tilde{X} = 2i\tilde{\Psi}e^{-if/f_*}\sin\frac{f}{f_*}\,, \tag{6.8a}$$

$$\tilde{Y} = -2i\tilde{\Psi}e^{-2if/f_*}\sin\frac{f}{f_*} \tag{6.8b}$$

$$\tilde{Z} = 0\,. \tag{6.8c}$$

We can already glean insight into how optical patch glitches can be identified. When all 6 laser links are functioning, none of the optical path glitches can be made to look like the other. For example, suppose we try to match the X channel response of Φ_{12}^{op} to that of Φ_{21}^{op}. This would require a time shift of $t + L$ i.e. a factor of e^{if/f_*} in the frequency domain. This time shift will of course shift the Y response in the opposite desired direction in time. We cannot find a transformation of wavelet parameters such that any optical path glitch looks like another when all 6 laser links functioning. If we are unfortunate enough to have only 2 functioning arms, we will be at a loss when attempting to distinguish these two glitches. That is if we have only the X channel, we will not be able to distinguish Φ_{12}^{op} from a time shifted Φ_{21}^{op}.

We must also contend with acceleration glitches. The acceleration glitch Φ_{12}^{ac} has the TDI response

$$\tilde{Y} = 4\tilde{\Psi}e^{-2if/f_*}\sin^2\frac{f}{f_*}\,. \tag{6.9}$$

Where both the X and Z channel are null. All acceleration glitches have a response in only 1 of the X, Y, and Z data channels. Acceleration glitches also have an additional suppression from the extra factor the transfer function $\sin f/f_*$. This is due to the acceleration glitch appearing in two phase measurements separated by the light travel time between spacecraft. With acceleration glitches however, we are unable to unambiguously determine their origin even when all 6 laser links are

functioning. There are perfect degeneracies between pairs of acceleration glitches. For example, the response to the acceleration glitch $\Phi_{32}^{\rm ac}$ is

$$\tilde{Y} = -4\tilde{\Psi}e^{-2if/f_*}\sin^2\frac{f}{f_*}\ . \tag{6.10}$$

has precisely the same form as the $\Phi_{12}^{\rm ac}$ except a shift in its initial phase (by π). In the scenario that we lose one arm of the constellation we will be no worse off with respect to distinguishing acceleration glitches. The response to glitches in other components are shown in table 6.1.

When generating these waveforms we wish economically sample an appropriate bandwidth. The signal-to-noise ratio for an optical path glitch, given one data channel, can be estimated as

$$\rho^2_{\rm est} = \frac{\sqrt{\pi/2}A^2\tau}{S_{X,M}(f_0)}\ , \tag{6.11}$$

in the large τ limit obtained from eqn. (6.6). $S_{X,M}$ is the Michelson-equivalent power spectral density defined as $S_{X,M} = S_X/4\sin^2(f/f_*)$. Similarly, the SNR for an acceleration glitch can be estimated as

$$\rho^2_{\rm est} = \frac{4\sqrt{\pi/2}A^2\tau\sin^2(f_0/f_*)}{S_{X,M}(f_0)}\ . \tag{6.12}$$

A bandwidth of $\Delta f = 4(\rho_{\rm est}/5)^2/\tau$ was used to capture in excess of 99.9% of the SNR in each glitch.

In addition to distributing power in different TDI channels, glitches in different components produce different phasing in the response. The phasing information depends critically on the frequency of the glitch, f_0 and the duration of the glitch τ. Higher frequency glitches get heavily modulated by the transfer functions making it

Table 6.1: This table contains the analytic first generation TDI variables for optical path and acceleration glitches. Note that optical path glitches occupy 2 of the XYZ TDI channels while the acceleration glitches only occupy 1. Acceleration glitches pick up an additional factor of the transfer function $\sin f/f_*$. A change in wavelet parameters, specifically the initial phase, leads to a perfect degeneracy between pairs of acceleration glitches when all three TDI channels are functioning. This is not the case for optical path glitches.

	$\tilde{X}$	$\tilde{Y}$	$\tilde{Z}$
$\Phi_{12}^{\rm op}$	$2i\tilde{\Psi}e^{-2if/f_*}\sin(f/f_*)$	$-2i\tilde{\Psi}e^{-if/f_*}\sin(f/f_*)$	0
$\Phi_{21}^{\rm op}$	$2i\tilde{\Psi}e^{-if/f_*}\sin(f/f_*)$	$-2i\tilde{\Psi}e^{-2if/f_*}\sin(f/f_*)$	0
$\Phi_{13}^{\rm op}$	0	$2i\tilde{\Psi}e^{-if/f_*}\sin(f/f_*)$	$-2i\tilde{\Psi}e^{-2if/f_*}\sin(f/f_*)$
$\Phi_{31}^{\rm op}$	0	$-2i\tilde{\Psi}e^{-if/f_*}\sin(f/f_*)$	$2i\tilde{\Psi}e^{-2if/f_*}\sin(f/f_*)$
$\Phi_{23}^{\rm op}$	$2i\tilde{\Psi}e^{-if/f_*}\sin(f/f_*)$	0	$-2i\tilde{\Psi}e^{-2if/f_*}\sin(f/f_*)$
$\Phi_{32}^{\rm op}$	$2i\tilde{\Psi}e^{-if/f_*}(f/f_*)$	0	$-2i\tilde{\Psi}e^{-2if/f_*}\sin(f/f_*)$
$\Phi_{12}^{\rm ac}$	0	$4\tilde{\Psi}e^{-2if/f_*}\sin^2(f/f_*)$	0
$\Phi_{21}^{\rm ac}$	$-4\tilde{\Psi}e^{-2if/f_*}\sin^2(f/f_*)$	0	0
$\Phi_{13}^{\rm ac}$	0	0	$-4\tilde{\Psi}e^{-2if/f_*}\sin^2(f/f_*)$
$\Phi_{31}^{\rm ac}$	$4\tilde{\Psi}e^{-2if/f_*}\sin^2(f/f_*)$	0	0
$\Phi_{23}^{\rm ac}$	0	0	$4\tilde{\Psi}e^{-2if/f_*}\sin^2(f/f_*)$
$\Phi_{32}^{\rm ac}$	0	$-4\tilde{\Psi}e^{-2if/f_*}\sin^2(f/f_*)$	0

easier to determine their origin. In Figure 6.1 the AET TDI channels for optical path glitches $\Phi_{12}^{\rm op}$ are displayed in red and acceleration glitches $\Phi_{21}^{\rm ac}$ in shown blue. The amplitudes of the optical and acceleration glitches were chosen for ease of comparison, while maintaining the correct relative amplitudes in the different TDI channels. The top row (case 1) displays glitches with the parameters $\tau = 0.2$ hours, and $f_0 = 2/\tau$ i.e.

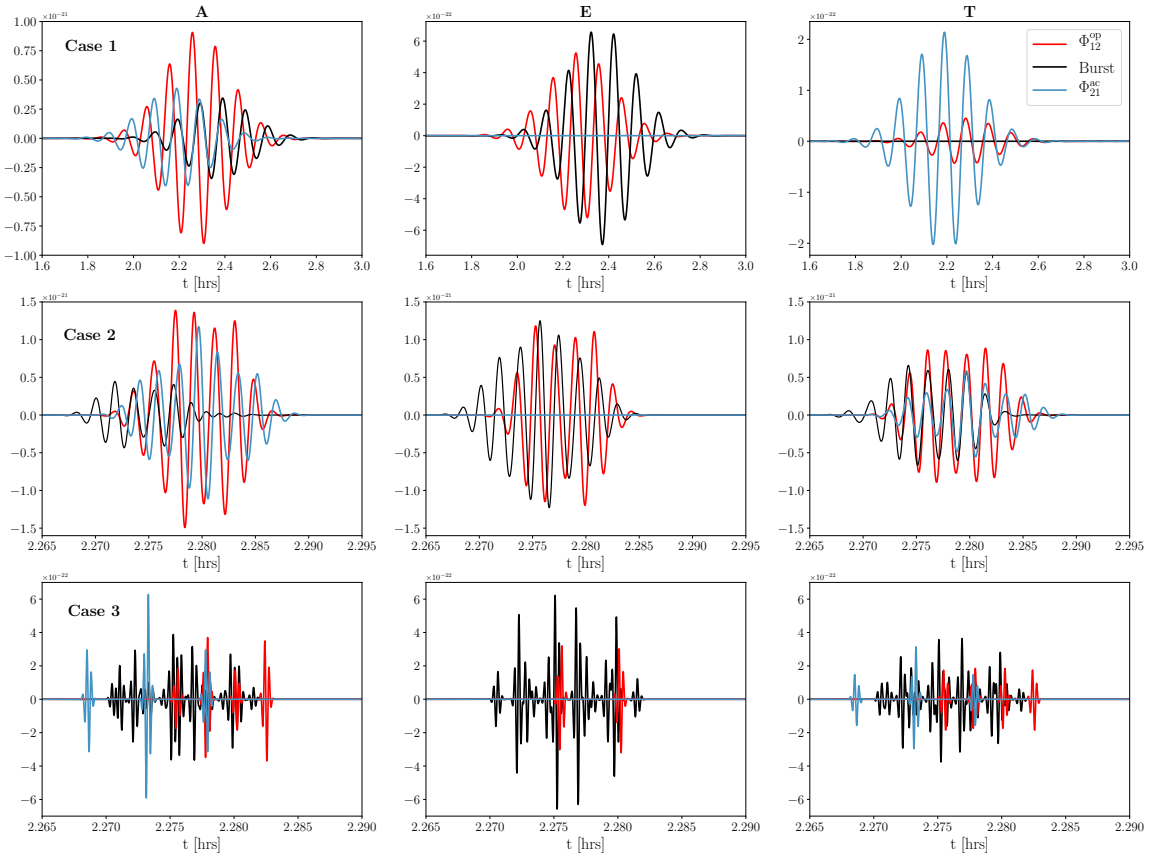

Figure 6.1: This figure displays the *AET* TDI channel responses for various glitches (Φ_{12}^{op} in red, and Φ_{21}^{ac} in blue) and gravitational wave bursts (in black). The top row shows wavelets with durations that are much longer than the light travel time between LISA spacecraft. The middle row shows wavelets with duration that is comparable to the light travel time. The bottom row shows wavelet with duration less than the light travel time, which leads to a clean separation of the glitch wavelets in the TDI channels. Note that the glitch wavelets only appear in a subset of the TDI channels.

2.7 mHz, placing this glitch well below the transfer frequency. These parameters give the glitch a quality factor of 12.6. Since the wavelet has a low frequency, its amplitude does not change substantially in a light travel time. This means that the construction of the TDI channels acts like a derivative of the input. In the middle row (case 2) the parameters of the wavelet are $\tau = L = 8.33$ seconds and $f_0 = 1.3/\tau = 156$ mHz ($Q \sim 8$). This wavelet has a temporal extent comparable to that of the light travel

time between spacecraft. This results in a waveform that is the superposition of two wavelets with a small time shift between them. Lastly, in the bottom row (cases 3) we see a wavelet of $\tau = 1$ second and $f_0 = 800$ mHz i.e. $Q = 5.1$. Here the frequency of the signal is substantially larger than the transfer frequency and the duration of the signal in time is much less than the light travel time, leading to a clean separation of the wavelets in the TDI channels. Note that in the low frequency regime the optical path glitch has a suppressed output in the T channel. We also see that the E channel response to the acceleration glitch is totally suppressed. This is because there is no Z or Y response for this specific acceleration glitch and the E channel has no X channel dependence.

6.2.2 Gravitational Wave Bursts

The optical path length change due to a gravitational wave signal in the laser link connecting the i^{th} and j^{th} spacecraft is given by

$$\delta\ell_{ij}(t) = \mathbf{D}_{ij} : \int_{\xi_i}^{\xi_j} \mathbf{h}(t)dt \ , \tag{6.13}$$

where the colon denotes full contraction between the tensors, i.e. $\mathbf{A} : \mathbf{B} = A^{jk}B_{jk}$. The time t is Solar System Barycenter (SSB) time, and $\xi_i = t_i - \hat{\mathbf{k}} \cdot \mathbf{x}_i$ is the wave variable defining surfaces of constant phase for the gravitational wave. The position of the i^{th} spacecraft is $\mathbf{x}_i$ and t_i is the time of emission of the laser photon from spacecraft i and t_j is the time of reception of the laser photon at spacecraft j. The detector tensor $\mathbf{D}$ is given by

$$\mathbf{D} = \frac{1}{2}\frac{\hat{\mathbf{r}}_{ij} \otimes \hat{\mathbf{r}}_{ij}}{1 - \hat{\mathbf{k}} \cdot \hat{\mathbf{r}}_{ij}} \ , \tag{6.14}$$

where $\hat{\mathbf{k}} = -(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ is the gravitational wave propagation direction—θ and ϕ designate the source's position in spherical polar coordinates in the SSB frame. The quantity $\hat{\mathbf{r}}_{ij}$ is the unit-separation vector between the LISA spacecraft pointing from spacecraft i to spacecraft j. In this study the LISA orbits are kept to leading order in eccentricity thereby fixing LISA arm length to be constant [78] $\mathrm{L} = |\mathbf{r}_{ij}|$ for all i, j combinations.

The gravitational wave tensor $\mathbf{h}$ is given by

$$\mathbf{h} = h_+(t)\mathbf{e}'_+(\psi,\theta,\phi) + h_\times \mathbf{e}'_\times(\psi,\theta,\phi) \ , \tag{6.15}$$

where $\mathbf{e}'_{+,\times}$ are the polarization tensors $\mathbf{e}_{+,\times}$ rotated by the polarization angle $\psi \in [0,\pi]$. In this work we assume that the gravitational waves are elliptically polarized such that, in the frequency domain, $\tilde{h}_\times = i\epsilon\tilde{h}_+$ parameterized by the ellipticity $\epsilon \in [0,1]$. We model the integrated gravitational wave polarizations as Gabor-Morlet wavelets such that $\int^t h_+(t')dt' = L\Psi(t)$. We may approximate the detector as static for the duration of a wavelet, since corrections would be on the order of $\tau/1\mathrm{yr}$. This means we may safely evaluate all terms associated with the position of the detector at the central time of the wavelet t_0 and assume the value is constant. The response to the wavelet in the frequency domain is then

$$\tilde{y}_{ij} = \frac{f}{f_*}\left(F_{ij}^{+\prime} + i\epsilon F_{ij}^{\times\prime}\right)\mathcal{T}_{ij}(f;\hat{\mathbf{k}})\tilde{\Psi}(f)e^{-2\pi i f\hat{\mathbf{k}}\cdot\mathbf{x}_i(t_0)} \tag{6.16}$$

where $y_{ij} = \delta\ell_{ij}/2L$ and $F_{ij}^{+,\times} = [\hat{\mathbf{r}}_{ij}\otimes\hat{\mathbf{r}}_{ij}] : \mathbf{e}_{+,\times}$. $\mathcal{T}_{ij}$, the transfer function, is given by

$$\mathcal{T}_{ij} = \frac{1}{4}e^{i\left(\frac{f}{2f_*}(1-\hat{\mathbf{k}}\cdot\hat{\mathbf{r}}_{ij})\right)}\,\mathrm{sinc}\left(\frac{f}{2f_*}(1-\hat{\mathbf{k}}\cdot\hat{\mathbf{r}}_{ij})\right) \ . \tag{6.17}$$

The wavelet has its central time shifted by the light travel time between the SSB origin and spacecraft i through the phase factor $e^{2\pi i f \hat{\mathbf{k}}\cdot\mathbf{x}_i(t_0)}$ present in eqn. (6.16). The ellipticity ϵ and polarization angle ψ simply modulate the amplitude of the response. We see that the sky angles modulate the amplitude too, but also enter into the phasing. As opposed to instrumental glitches, gravitational wave bursts will induce responses in all TDI channels. It is important to note though that for frequencies below the transfer frequency f_*, the gravitational wave response in the T channel is heavily suppressed [242]. This can be seen in the T channel response for case 1 in Figure 6.1 . The signal in each panel of Figure 6.1 represents a gravitational wave burst. The sky angles were chosen such that $\cos\theta = 0.23$ and $\phi = 2.31$. The polarization angle was 0.45 and the ellipticity 0.5. The wavelet parameters are precisely the same as those for the glitches in the panel the burst shares (give or take an amplitude factor or time shift for the sake of easy comparison). Note that for case 1 the signal response is distinctly different than for the glitch. Recall, that the glitches in this case were cleanly separated. This is because glitches enter the data stream with time delays equal to the light travel time between spacecraft, which is longer than their extent in time. Gravitational waves enter the data stream with time delays equal to the projected arm lengths. This can lead to foreshortened arms allowing for some overlap between the wavelets as see in case 3.

6.3 Bayesian Inference

With the glitch and burst models established we now turn to the methods used to infer the properties of the gravitational wave signals and instrument glitches and develop probability distributions for the parameters of the models. These probabilitiesnu are quantified by the posterior distribution $p(\vec{\lambda}_{\mathcal{M}}|\mathbf{s},\mathcal{M})$ which reflects our belief about a given set of parameters $\vec{\lambda}_{\mathcal{M}}$ which specify model $\mathcal{M}$ given data $\mathbf{s}$.

The posterior distribution is obtained via Baye's theorem:

$$p(\vec{\lambda}_{\mathcal{M}}|\mathbf{s},\mathcal{M}) = \frac{p(\mathbf{s}|\vec{\lambda}_{\mathcal{M}},\mathcal{M})p(\vec{\lambda}_{\mathcal{M}}|\mathcal{M})}{p(\mathbf{s}|\mathcal{M})} \ , \tag{6.18}$$

where $p(\vec{\lambda}_{\mathcal{M}}|\mathcal{M})$ is the prior distribution for the parameters $\vec{\lambda}_{\mathcal{M}}$, $p(\mathbf{s}|\vec{\lambda}_{\mathcal{M}},\mathcal{M})$ the likelihood of the data given the parameters, and $p(\mathbf{s}|\mathcal{M})$ is the evidence for the model $\mathcal{M}$. Along with the assumptions we have already made in the construction of the TDI channels, we further assume that, aside from the glitches modeled here, the noise is stationary and Gaussian. The likelihood function for the data then takes the form

$$p(\mathbf{s}|\vec{\lambda},\mathcal{M}) \propto \exp\left[-\frac{1}{2}\sum_I \Big(\mathbf{s}_I - \mathbf{h}_I(\vec{\lambda})|\mathbf{s}_I - \mathbf{h}_I(\vec{\lambda})\Big)\right] \ , \tag{6.19}$$

where the $\mathcal{M}$ subscript on the parameters has been dropped for simplicity. The sum is over TDI data streams $I = \{A, E, T\}$ (or just $I = \{X\}$ for some of our investigations). The noise-weighted inner product is defined as

$$(\mathbf{a}_I|\mathbf{b}_I) = 4\mathcal{R}\int_0^\infty \frac{\tilde{a}_I(f)\tilde{b}_I^*(f)}{S_{n,I}(f)}df \ . \tag{6.20}$$

The noise strain spectral density in these data channels are given by

$$S_{AE} = \frac{16}{3}\sin^2\frac{f}{f_*}\left[\left(2+\cos\frac{f}{f_*}\right)P_{\rm OMS} + 2\left(3+2\cos\frac{f}{f_*}+\cos\frac{2f}{f_*}\right)\frac{P_{\rm acc}}{(2\pi f)^4}\right]\frac{1}{(2{\rm L})^2} \tag{6.21a}$$

$$S_T = \frac{16}{3}\sin^2\frac{f}{f_*}\left[\left(1-\cos\frac{f}{f_*}\right)P_{\rm OMS} + \left(3-4\cos\frac{f}{f_*}+\cos\frac{2f}{f_*}\right)\frac{P_{\rm acc}}{(2\pi f)^4}\right]\frac{1}{(2{\rm L})^2} \tag{6.21b}$$

$$S_X = 4\sin^2\frac{f}{f_*}\left[4P_{\rm OMS} + 8\left(1+\cos^2\frac{f}{f_*}\right)\frac{P_{\rm acc}}{(2\pi f)^4}\right]\frac{1}{(2{\rm L})^2} \, . \tag{6.21c}$$

The noise in the A and E channels, S_{AE}, is the same and the noise in the T channel is S_T. The single-link optical metrology noise power $P_{\rm OMS}$ and single test-mass acceleration noise power $P_{\rm acc}$ are quoted in [59]. Another contribution to the measured noise comes from millions of unresolved galactic binaries [24] emitting gravitational waves to which LISA is sensitive. Estimates of the unresolved galactic binary confusion noise for various observation periods can also be found in the same reference.

6.3.1 Maximization over nuisance parameters

In a fully Bayesian analysis we would compute the joint posterior distributions all parameters in the model. To simplify the analysis and achieve more rapid convergence, we chose to eliminate certain nuisance parameters by analytically maximizing the likelihood with respect to these parameters using the F-statistic approach (we could have analytically *marginalized* over the nuisance parameters instead [244], but it is much faster and simpler to maximize). The F-statistic [40] provides a way to maximize the likelihood over the extrinsic parameters—A, ϕ_0 for a glitch, and A, ϕ_0, ψ, ϵ for a burst. Through the use of several filters, constructed from the burst or glitch wavelet with specific choices of extrinsic parameters, one may construct the maximized likelihood. To understand how to construct the F-statistic it is useful to consider the burst model in the large τ and low frequency limit

$$\begin{aligned} y_{ij} = &- \frac{f_0}{4f_*}\Big[F_{ij}^{+\prime}\sin\left(2\pi f_0 t_i + \bar{\phi}\right) \\ &+\epsilon F_{ij}^{\times\prime}\cos\left(2\pi f_0 t_i + \bar{\phi}\right)\Big] \ , \end{aligned} \tag{6.22}$$

where $t_i = t - \hat{\mathbf{k}} \cdot \mathbf{x}_i$. This signal may be deconstructed into four terms which consist of a constant amplitude dependent on extrinsic parameters multiplying a time-dependent factor, additionally dependent on the intrinsic parameters (f_0, τ, t_0, θ, ϕ)

$$y_{ij} = \sum_k a_k A^k(t) \ . \tag{6.23}$$

The four filters $A^k(t)$

$$A^1 = -\frac{f_0}{4f_*} F_{ij}^{+} \sin\left(2\pi f_0 t_i\right) \tag{6.24a}$$

$$A^2 = -\frac{f_0}{4f_*} F_{ij}^{\times} \sin\left(2\pi f_0 t_i\right) \tag{6.24b}$$

$$A^3 = -\frac{f_0}{4f_*} F_{ij}^{+} \cos\left(2\pi f_0 t_i\right) \tag{6.24c}$$

$$A^4 = -\frac{f_0}{4f_*} F_{ij}^{\times} \cos\left(2\pi f_0 t_i\right) \tag{6.24d}$$

may be constructed by inserting the following the extrinsic parameters, as described by table 6.2, into the burst waveform generator. The glitch F-statistic filters can be

Table 6.2: Plugging these parameters into the gravitational wave burst waveform generator will construct the filters eqns. (6.24). The resulting filters can then be used to maximize the likelihood over the extrinsic parameters.

Filter	A	$\bar{\phi}$	ψ	ϵ
A^1	1	0	0	0
A^1	1	0	$-\pi/4$	0
A^3	1	$\pi/2$	0	0
A^4	1	$\pi/2$	$-\pi/4$	0

constructed by the parameter choices: 1) $A = 1, \bar{\phi} = 0$ and 2) $A = 1, \bar{\phi} = -\pi/4$. The extrinsic parameter coefficients are

$$a_1 = A\left(\cos 2\psi \cos\bar{\phi} - \epsilon \sin 2\psi \sin\bar{\phi}\right) \tag{6.25a}$$

$$a_2 = A\left(-\sin 2\psi \cos\bar{\phi} - \epsilon \cos 2\psi \sin\bar{\phi}\right) \tag{6.25b}$$

$$a_3 = A\left(\cos 2\psi \sin\bar{\phi} + \epsilon \sin 2\psi \cos\bar{\phi}\right) \tag{6.25c}$$

$$a_4 = A\left(-\sin 2\psi \sin\bar{\phi} + \epsilon \cos 2\psi \cos\bar{\phi}\right) \ . \tag{6.25d}$$

The noise-weighted inner product of these filters with the data $N^k = (\mathbf{s}|\mathbf{A}^k)$ can be used to construct the maximized relative likelihood

$$\mathcal{F} = \log\mathcal{L} = \frac{1}{2}\left(M^{-1}\right)_{mn} N^m N^n \ . \tag{6.26}$$

The value $\mathcal{L}$ is the relative likelihood, i.e. the ratio between the likelihood assuming $\mathbf{h}$ contains a burst and the likelihood assuming there is no such signal, i.e. $\mathbf{h} = 0$, such that

$$\log\mathcal{L} = (\mathbf{s}|\mathbf{h}) - \frac{1}{2}(\mathbf{h}|\mathbf{h}) \ . \tag{6.27}$$

The results hold for summing over multiple data channels such as when we use the *AET* TDI channels. The matrix $M^{mn} = (\mathbf{A}^m|\mathbf{A}^n)$ is simply the inner product matrix of the filters. Although, in this study we do not make use of the extrinsic parameters which maximize the likelihood it may prove useful in a future to study to be able to calculate them. Inverting the equations for the filter returns the extremized extrinsic

parameters

$$A = \sqrt{\frac{1}{2}\left(s + \sqrt{pq}\right)} \tag{6.28a}$$

$$\epsilon = \frac{s - \sqrt{pq}}{2\left(a_1 a_4 - a_2 a_3\right)} \tag{6.28b}$$

$$\tan(2\psi) = \frac{a_1^3 + 2a_2 a_3 a_4 + a_1(a_2^2 + a_3^2 - a_4^2 + \sqrt{pq})}{a_1^2 a_2 + 2a_1 a_3 a_4 + a_2(a_2^2 - a_3^2 + a_4^2 + \sqrt{pq})} \tag{6.28c}$$

$$\tan\bar{\phi} = \frac{a_1^2 + a_2^2 - a_3^2 - a_4^2 + \sqrt{pq}}{-2(a_1 a_3 + a_2 a_4)} \tag{6.28d}$$

where $s = a_1^2 + a_2^2 + a_3^2 + a_4^2$, $p = (a_2 + a_3)^2 + (a_1 - a_4)^2$, and $q = (a_2 - a_3)^2 + (a_1 + a_4)^2$. For glitches the amplitude and phase can be extracted via

$$A = \sqrt{a_1^2 - a_2^2} \tag{6.29a}$$

$$\tan\bar{\phi} = \frac{-a_2}{a_1}\ . \tag{6.29b}$$

6.3.2 Markov Chain Monte Carlo

In this study we wish to characterize what we can learn about a wavelet present in the data. To accomplish this we marginalize the posterior distribution via the Markov Chain Monte Carlo (MCMC) algorithm. Suppose we inject a signal into our data $\mathbf{s}$. Upon choosing a model specified by the initial set of parameters $\vec{x}$ we generate a proposed set of parameters from a probability density $q(\vec{y}|\vec{x})$. The chance that we accept this new set of parameters $\vec{y}$ is given by the Hasting's ratio

$$H = \min\left\{1, \frac{p(\mathbf{s}|\vec{y},\mathcal{M})p(\vec{y}|\mathcal{M})q(\vec{x}|\vec{y})}{p(\mathbf{s}|\vec{x},\mathcal{M})p(\vec{x}|\mathcal{M})q(\vec{y}|\vec{x})}\right\}\ . \tag{6.30}$$

The sequence of parameters we accept, called a chain, constitute samples from the posterior distribution $p(\vec{\lambda}_{\mathcal{M}}|\mathbf{s},\mathcal{M})$. The MCMC we created used the F-statistic likelihood, extremizing the likelihood over the extrinsic parameters of the signal. This effectively reduces the search space of the MCMC, greatly improving its convergence, especially for the burst model which otherwise converges very slowly when the sky location is poorly-constrained.

For the MCMC developed in this study uniform priors were set for the parameter set $\{\log A,\ f_0,\ t_0,\ \log\tau,\ \bar{\phi},\ \cos\theta,\ \phi,\ \psi,\ \epsilon\}$. To aid in the convergence of the MCMC we used a mixture of proposal distributions. We utilized local Gaussian approximations to the posteriors through the Fisher matrix (which approximates the inverse covariance matrix)

$$\Gamma_{ij} = \sum_{\mathrm{I}} \left(\mathbf{h}_{\mathrm{I},i}|\mathbf{h}_{\mathrm{I},j}\right) \ , \tag{6.31}$$

where $\mathbf{h}_{I,i}$ represent the derivative of waveform (in the I^{th} data channel) with respect to the i^{th} parameter λ_i. These derivatives were calculated numerically using finite differencing of the waveforms discussed in section 6.2. We occasionally used proposals from the prior distribution. Since we are not currently developing a detection algorithm, only an MCMC which characterizes the signal, we used a targeting distribution to help the MCMC find appropriate central frequencies f_0 and decay factor τ. For f_0 and τ individually, this proposal consisted of a Gaussian distribution centered on the true parameter used to generate the injection. The width of Gaussian was chosen based on the Fisher matrix estimation for the error in that parameter. To improve the acceptance rate of this proposal distribution the Gaussian was mixed with a 20% by weight uniform distribution covering the prior range. Differential evolution [245] proposals were also used. Lastly, a time shift proposal was used to

help highly oscillatory wavelets where shifts, forwards or backwards, in the central time of the wavelet by the wavelet's period were proposed (with an appropriate shift in initial phase).

To further improve convergence, and to ensure a thorough exploration of parameter space—such as investigating the existence of secondary modes on the sky for bursts—parallel tempering [48] was utilized. During parallel tempering multiple chains are simulated simultaneously at different temperatures, i.e. their likelihoods are flattened $p(\mathbf{s}|\vec{\lambda}, \mathcal{M})^{\beta_j}$, where $\beta_j = 1/T_j$ is the inverse temperature fo the j^{th} chain. The cold chain, i.e. $T_0 = 1$, represents samples from the posterior distribution. The chains at various temperatures propose and accept new parameters just as before, but with the flattened likelihood. Occasionally, swaps of parameters between chains neighboring in temperature are proposed based on the probability

$$H_{\mathrm{PT}} = \min\left\{1, \frac{p(\mathbf{s}|\vec{\lambda}_j, \mathcal{M})^{\beta_{j+1}} p(\mathbf{s}|\vec{\lambda}_{j+1}, \mathcal{M})^{\beta_j}}{p(\mathbf{s}|\vec{\lambda}_j, \mathcal{M})^{\beta_j} p(\mathbf{s}|\vec{\lambda}_{j+1}, \mathcal{M})^{\beta_{j+1}}}\right\}. \tag{6.32}$$

Parallel tempering vastly improves convergence once a proper selection of temperatures is made. The maximum temperature is chosen such that the hottest chain freely explores the parameters' prior volume, while not so hot as to be redundant in the prior space exploration as cooler chains. In section 6.5.1 we see how parallel tempering additionally aids us in determining whether a glitch (and which one) or a burst best explains the data.

6.4 Parameter Estimation

The MCMC may now be used to address questions such as how well do we characterize the parameters of the signal, and how well do we recover the waveform itself. The central frequency f_0 and time damping factor τ are typically well

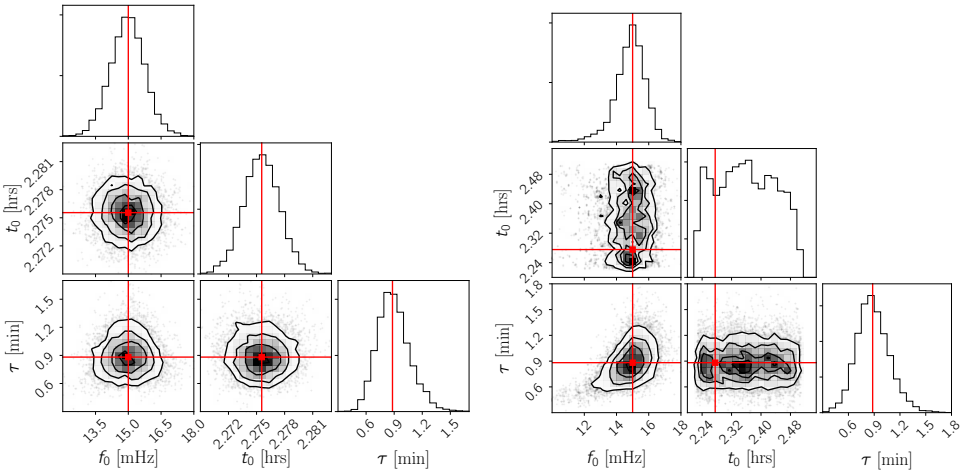

Figure 6.2: Marginalized posteriors for the parameters f_0, t_0, and τ are displayed for an optical path glitch Φ_{12}^{op} on the left panel and a gravitational wave burst on the right panel. The fully marginalized posteriors for f_0 and τ are similar for the two injections (while the joint posterior exhibits some correlation for the burst), but the central time posteriors differ significantly. The injected parameters are marked by the red lines.

determined for bursts and glitches. An example marginalized posterior for these parameters is seen in Figure 6.2. The left panel shows marginalized posterior distributions for an optical path glitch Φ_{12}^{op} and the right panel shows marginalized posteriors for the same parameters for a burst. The injected signals both had a signal-to-noise of 8. The signal-to-noise (SNR) is given by

$$\rho = \sum_I (h_I|h_I) \ . \tag{6.33}$$

They shared the parameter values $f_0 = 15$ mHz, $t_0 = 0.5T$ (where T is the observation period), and $\tau = 53$ seconds (giving the wavelets a quality factor of 5.0). The burst injection had the following additional parameters: $\cos\theta = 0.23$, $\phi = 2.31$, $\psi = 0.45$, and $\epsilon = 0.5$. We see that the fully marginalized posterior distributions for the central frequency f_0 and τ are rather similar for these two injections. However, the posteriors

for the central time t_0 are quite distinct in a significant way. One can show through a simple Fisher matrix calculation for a wavelet [43] that the standard deviation in t_0 for a wavelet scales as $1/\rho\tau$. The injected glitch has a measured standard deviation of 53 seconds while the Fisher matrix standard deviation estimates an error of 70 seconds, demonstrating agreement. The standard deviation in t_0 for the burst is 4.6 minutes, much larger than that of the glitch which must be attributed to the more complex response of a burst compared to a glitch.

The reason for the increase in error associated with the central time of the burst can be seen in Figure 6.3. There exists a substantial statistical dependence

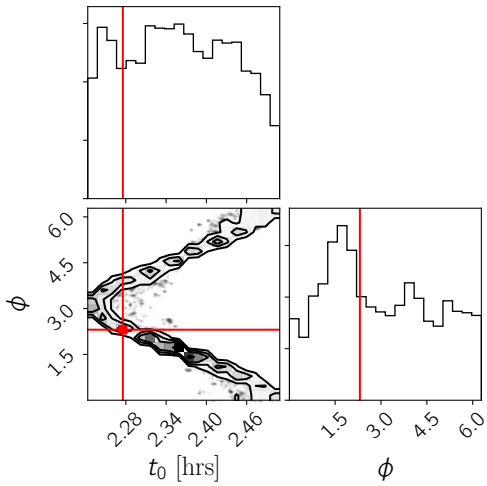

Figure 6.3: The figure displays the joint posterior for the azimuthal angle ϕ and the wavelet's central time t_0 as well as their fully marginalized posteriors. The injected parameters are again marked by the red lines. There exists a strong statistical dependence between these two parameters.

between the azimuthal sky angle ϕ and t_0. Without the central time constrained appropriately it turns out that we cannot determine the sky location, which is the case for this example burst. We can understand this by looking at the low frequency response to a GW burst signal. In this regime, the Michelson-equivalent A and E TDI channels are proportional to $\frac{f}{f_*}\tilde{\Psi}e^{-2\pi i f\hat{\mathbf{k}}\cdot\mathbf{x}_1}$ modulo overall constants that differ

between the channels. The T channel is null in this limit. We see that the sky angles enter the phasing through time shift factor multiplying the Fourier transform of the Gabor-Morlet wavelet. Since this factor is a time shift, the sky angles are almost perfectly degenerate with the central time of the wavelet. The likelihood is approximately constant under mappings of the azimuth sky location and central time that keep fixed the combination $t_0 - R\sin\theta\cos(2\pi f_m t_0 - \phi)$, where $f_m = 1/\mathrm{yr}$ is orbital modulation frequency and $R = 1$ AU. This relationship holds to leading order in the orbital eccentricity. Higher order corrections to the phasing incorporate additional information about the sky location in the form of the projected arm lengths $L_{ij} = L\left(1 - \hat{\mathbf{k}}\cdot\hat{\mathbf{r}}_{ij}\right)$.

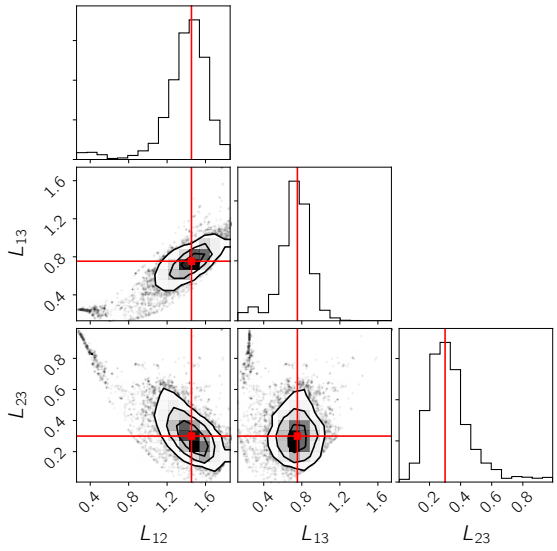

Figure 6.4: The posterior distributions for the quantities $1 - \hat{\mathbf{k}}\cdot\hat{\mathbf{r}}_{ij}$ are displayed here, i.e. LISA's projected arm lengths for a high frequency gravitational wave burst, normalized by the arm length $L = 2.5$ Gm. The red lines indicate the injected values for the GW burst signal.

In Figure 6.4 we see the posterior distributions for the projected arm lengths from a burst injection of the same sky location, polarization, ellipticity, and SNR as in the previous example, but now the central frequency is 50 mHz and $\tau = 16$ seconds.

This shorter envelope allows the central time to be measured and therefore the sky location to be better determined. The duration of the wavelet is more important in determining the sky location than the central frequency.

In Figure 6.5 we see the posterior distribution for the sky location for two different high frequency bursts. Both bursts in Figure 6.5 have the same sky location, which is denoted by the blue dot in the sky map on the left . The central frequency of each source is 50 mHz. The duration of the source shown on the left is $\tau = 16$ seconds (i.e. the same burst used in Figure 6.4) while the sky map on the right is for a source with $\tau = 2.8$ minutes. We see that the origin of the burst on the sky has been localized for the short duration burst. This is due to the tight constraint on the central time of the wavelet. When the central time of the wavelet is measured to better than the light travel time between spacecraft we begin to have the power to localize the wavelet on the sky. The source shown on the right had a longer duration and poorer constraint on the central time, and therefore a poorer constraint on the sky location. Interesting structure emerges on the sky posterior for the right source. The most important factors in determining the sky location are the measured values for the projected arm lengths. The projected arm lengths can be related to the sky locations via the relations

$$\begin{aligned}\hat{\mathbf{k}}\cdot\hat{\mathbf{r}}_{12}(t) = &-\frac{1}{2}\cos\theta\left[\cos\alpha+\cos(\alpha+\frac{\pi}{3})\right]\\ &+\frac{1}{8\sqrt{3}}\sin\theta\left[2\cos(2\alpha-\phi)-9\cos\phi+3\sqrt{3}\sin\phi+2\sin(\phi-2\alpha+\frac{\pi}{6})\right]\end{aligned} \tag{6.34a}$$

$$\hat{\mathbf{k}}\cdot\hat{\mathbf{r}}_{13}(t) = -\frac{1}{2}\cos\theta\left[\cos\alpha+\sin(\alpha+\frac{\pi}{6})\right]$$

$$+\frac{\sqrt{3}}{24}\sin\theta\left[2\cos(2\alpha-\phi)-9\cos\phi-3\sqrt{3}\sin\phi+2\sin(2\alpha-\phi+\frac{\pi}{6})\right] \tag{6.34b}$$

$$\hat{\mathbf{k}}\cdot\hat{\mathbf{r}}_{23}(t)=-\frac{\sqrt{3}}{2}\cos\theta\sin\alpha+\sin\theta\left[\sin(2\alpha-\phi+\frac{\pi}{6})-3\sqrt{3}\sin\phi+\sin(2\alpha-\phi-\frac{\pi}{6})\right]\ , \tag{6.34c}$$

where $\alpha = 2\pi f_m t$. When these values, $\hat{\mathbf{k}}\cdot\hat{\mathbf{r}}_{ij}(t)$ which are present in the phase, are well measured then the sky location can be determined. In the right panel of Figure 6.5 we see the curves on the sky defined by equations (6.34). The blue curve defines sky locations which give the same time delay (also same projected arm length) along the arm connecting the spacecrafts 1 and 2 specified by the true sky location of the injected burst. The red line displays the sky locations which maintain the same time delay between spacecraft 1 and 3 as the true sky location. The black line applies to the arm spanned by spacecraft 2 and 3. We see that these curves intersect at two specific sky locations, one of them coinciding with the true sky location by construction. The other intersection constituted a second mode that the MCMC explored. There exists one other mode which corresponds with a different central time that also provided a good fit to the data. This secondary mode is shifted by a half period in time from the true value. In addition to the time delays, the sky localization is also impacted by the antenna patterns which change the amplitude of the signal in each channel, but this is a weaker effect.

Lastly, we will also be concerned with the accuracy of our waveform reconstruction. This concern is especially warranted in future work where we will work with signals consisting of a superposition of wavelets. Figure 6.6 shows an example optical path glitch Φ_{12}^{op} with a signal-to-noise of 8, a central frequency of 15 mHz, τ of 2 minutes ($Q = 17.0$). The observation period was set to 4.55 hours. The dotted

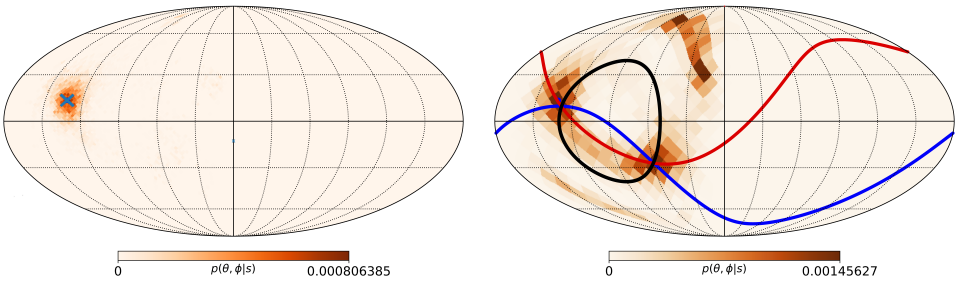

Figure 6.5: These figure displays the joint posterior for the sky location for two bursts of the same central frequency. The burst shown on the left has a short duration (τ = 17 seconds), while the burst shown on the right has a longer duration (τ = 2.8 minutes). The blue x on the left figure represents the true sky location, which is the same for both sources. Lines of constant projected arm length are shown in the sky map on the right. The blue line for spacecraft 1 and 2, the red line for 1 and 3 and the black line for 2 and 3. There are two sky locations which satisfy these constraints. There is an additional maximum away from the intersection of these lines that corresponds to a secondary mode with an overall half-period time shift.

black line denotes the signal corresponding to the injected parameters. The red lines denote waveforms for parameters sampled from the MCMC. Where the amplitude

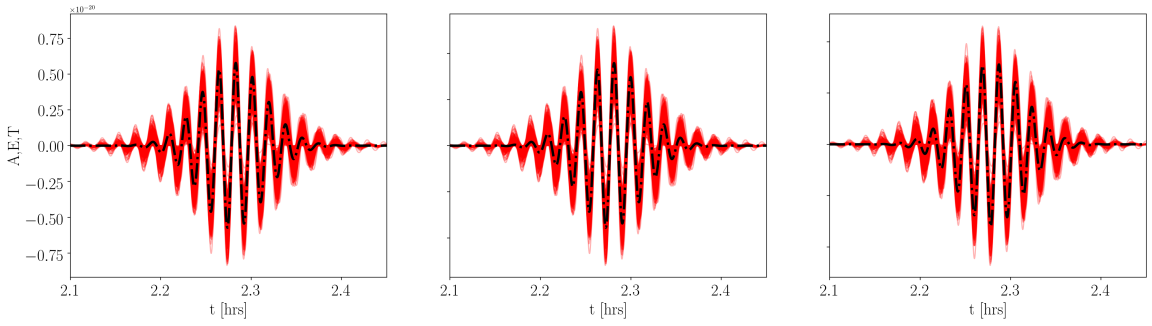

Figure 6.6: The left, center, and right panels display the *AET* TDI responses respectively for an optical path glitch injection denoted by the dot-dash black line. The red lines are MCMC samples of the waveform. The central frequency for this wavelet was 15 mHz, and $\tau = 2$ minutes.

of the wavelet is largest we see that the MCMC sampled wavelets hug the injected waveform more tightly; the errors in the wavelets are greater near the edges.

6.5 Model Selection

Now we will investigate whether we will be confused when identifying whether the data contains a glitch or signal. Recall that glitches entered the data stream with time delays of the arm light travel times. On the other hand, gravitational wave bursts instead enter the data stream with delays equal to the *projected* arm lengths $L(1 - \hat{\mathbf{k}} \cdot \hat{\mathbf{x}}_{ij})$. This impacts the phasing of the response, which can be used to infer the origin of glitch. Additionally, there are important differences between GW bursts and instrument glitches in where they place power in the TDI channels. Gravitational wave signals whose frequency is below the transfer frequency have a greatly diminished response in the T channel. The fact that gravitational waves are seen in at least two TDI channels while glitches enter in only 1 or 2 can be leveraged to distinguish the origin of a signal. In this section we first demonstrate through a simple argument that we do not expect to be confused between GW signals and glitches when the all six laser links are operational. Later in this section we will more rigorously demonstrate this conclusion through calculation of the Bayesian evidence. We will also study whether it is the phasing of the response or the power distribution in the TDI data channels that provides the greatest leverage for separating GW bursts from instrument glitches.

If one has access to the A, E, T data channels it is easy to make an argument that we will almost never confuse a glitch for a signal. Consider the following: noise in the data streams will affect our ability to match the true signal. A measurement of this match is the fitting factor (FF), i.e. a normalized (such that 1 indicates a perfect reconstruction of the signal) noise-weighted inner product, between the data and model waveform, maximized over all model parameters. The noise in the data leads to statistical deviations in the fitting factor, even if the true parameters and

model are used. The expected deviation from a fitting factor of 1 is described by [246]

$$1 - FF = \frac{D-1}{2\rho^2} \ , \tag{6.35}$$

where D is the dimension of the model.

Let us consider the scenario of using an acceleration glitch model, which crops up in the X channel only, when the data actually contains a gravitational wave burst. In the low frequency limit, were we would expect to be most confused, the burst does not have significant power in the T channel and also $A \sim h_+$, $E \sim h_\times$. Recall, in the frequency domain $\tilde{h}_\times = i\epsilon\tilde{h}_+$. The overlap (normalized noise-weighted inner product) between the acceleration glitch and burst is

$$\begin{aligned} &\frac{\sum_I (\mathbf{s}_I|\mathbf{h}_I)}{\sqrt{\left(\sum_I (\mathbf{s}_I|\mathbf{s}_I)\right)\left(\sum_J (\mathbf{h}_J|\mathbf{h}_J)\right)}} = \\ &\qquad \frac{\frac{2}{3}(\mathbf{A}_B|\mathbf{X}_{\rm ac})}{\rho\sqrt{\left(\frac{2}{3}\right)^2(\mathbf{X}_{\rm ac}|\mathbf{X}_{\rm ac}) + \left(\frac{1}{3}\right)^2(\mathbf{X}_{\rm ac}|\mathbf{X}_{\rm ac})}} \ , \end{aligned} \tag{6.36}$$

where ρ is the SNR of the burst injection. This overlap is maximized if somehow the acceleration glitch conspires to be proportional to the burst's A channel response $\mathbf{X}_{\rm ac} \propto \mathbf{A}_{\rm B}$. Let us also assume that the A channel response to the burst accounts for a fraction x of the squared SNR, i.e. $(\mathbf{A}_{\rm B}|\mathbf{A}_{\rm B}) = x\rho^2$. We then find that the overlap simplifies to $2\sqrt{x/5}$. If in the worst case scenario, all of the burst's SNR is in the A channel the largest fitting factor that can be obtained is 0.89—similar considerations for all other glitches demonstrate that this glitch is indeed the worst case in the regime under consideration.

Are we to be concerned by a fitting factor this large? To answer this question we can consider eqn. (6.35) to understand the statistical error in the fitting factor.

Inserting the value 0.89 into this equation results in an SNR of 4.3. We can loosely understand this SNR as the largest SNR possible for the burst that could result in a confusion of the origin of the data (i.e. whether it was a glitch or signal). So we see that, under some very general assumptions, it is only when a burst is marginally detectable that we might confuse it for an instrument glitch.

6.5.1 Thermodynamic Integration

To more rigorously find out which model best explains the data we must calculate the ratio of evidences $p(\mathbf{s}|\mathcal{M})$ for two given models. This quantity is known as the Bayes factor

$$B_{ij} = \frac{p(\mathbf{s}|\mathcal{M}_i)}{p(\mathbf{s}|\mathcal{M}_j)} \ . \tag{6.37}$$

In this subsection we calculate the Bayes factor for competing glitch and burst models for different injections such that we can understand our ability to distinguish a signal's origin and to determine when we cannot. We calculate the Bayes factor via thermodynamic integration [247]. Since we have utilized parallel tempering in our MCMC we can calculate the average log likelihood $E_\beta \left[\log p(\mathbf{s}|\vec{\lambda}, \mathcal{M})\right]$ for each temperature of the MCMC by simply calculating the sample mean of the log likelihood values for each sample in the chain. With these in hand one may calculate the evidence for a model via the integral

$$\ln p(\mathbf{s}|\mathcal{M}) = \int_0^1 d\beta E_\beta \left[\log p(\mathbf{s}|\vec{\lambda}, \mathcal{M})\right] \ . \tag{6.38}$$

We perform the integral using the methods described in Refs. [43,88]. The covariance matrix between the log likelihood values for each temperature are estimated and used to define a log likelihood for the thermodynamic integration integrand. [88]. The

integrand is fit by a cubic spline whose control points and locations are marginalized over via a reversible jump MCMC [248]. The MCMC gives us estimates for the evidence integral (upon integrating the cubic spline) and its associated error.

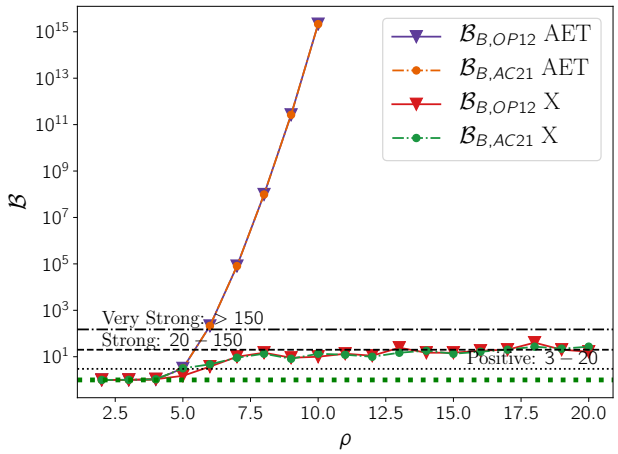

Figure 6.7: Bayes factors as function of signal to noise ratio. A burst was injected into the data stream with a central frequency of 15 mHz and $\tau = 2$ minutes. The purple and orange lines represent the Bayes factors when the *AET* TDI channels were used. Note that for the *AET* lines the Bayes factor lie on top of each other. The red and green lines represent the case when only the X TDI channel was used. Triangle markers denote the GW burst Bayes factors vs Φ_{12}^{op}, while circle markers represent GW burst Bayes factors vs Φ_{21}^{ac}. With the *AET* channels we gain confidence swiftly of the true model as SNR grows. The growth in Bayes factor is much slower when only the X channel is available.

Figure 6.7 shows the Bayes factor between the glitch and GW burst models for data containing a simulated burst. We use the notation $\mathcal{B}_{A,B} = p(\mathbf{s}|A)/p(\mathbf{s}|B)$ to represent the Bayes factor in the figure legends. Additionally, the label B is used to denote the burst model. The burst injection has the same values for θ, ϕ, ψ, and ϵ as the burst discussed in section 6.4. The other important parameters are the central frequency, set to 15 mHz, and τ set to 2 minutes (giving a quality factor of 17.0). The orange and purple lines denote the Bayes factor when the *AET* TDI channels are used and the red and green lines show the Bayes factor when only the

X channel was used. The lines marked by upside-down triangles represent the model comparison between the burst model and an optical path glitch between spacecraft 1 and 2. The lines marked with circles represent the Bayes factor between the burst model and an acceleration glitch between the two same spacecraft. The dashed green line represents a Bayes factor of 1, *i.e.* no preference between the two competing models. A Bayes factor between $3 - 20$ shows positive evidence [249] towards the true model. The evidence of the correct model is strong if the Bayes factor lies in the range $20 - 150$, and considered very strong if the Bayes factor is greater than 150. These regions are denoted by the various dashed horizontal black lines in figures 6.7 and 6.8. With the AET channel combination we see that the Bayes factor grows rapidly with signal-to-noise, and for SNRs greater than 5 we are confident that the signal is astrophysical. This supports our argument that GW burst and glitches are easily separated when we have the full collection of TDI channels. With just the X channel the prospects are not as good, and it is not until the signal reaches SNR 10 that it can be confidently distinguished from a glitch. In this low frequency regime we find that a burst injection recovered with an optical path glitch model gets biased central frequency and damping time scales. In Figure 6.8 we see the Bayes factors for a high frequency burst injection, where $f_0 = 50$ mHz and $\tau = 16.9$ seconds ($Q = 5.0$). We see that it is much easier to differentiate a gravitational wave burst from an acceleration glitch. Our ability to distinguish this burst from the optical path glitch Φ_{12}^{op} however is not enhanced as much as that of the acceleration glitch, but is still improved.

Lastly, we wish to how well we can differentiate models for glitch injections. For low frequency injections the story is similar in that with the full AET data stream we will be able to differentiate a glitch, both acceleration and optical path, through the distribution of power in the different data channels. When we only have the X

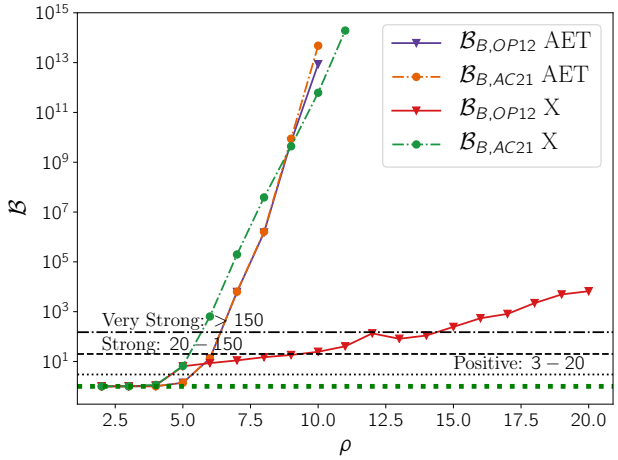

Figure 6.8: Bayes factors as function of signal to noise ratio for a high frequency burst injection. The lines are labeled according to the same scheme as in Figure 6.7. The GW burst is much easier to separate from an instrument glitch in this case.

channel discrimination once again becomes challenging until the SNR becomes large. In Figure 6.9 Bayes factors are displayed for high frequency glitch injections. The central frequency of the glitches was 50 mHz and $\tau = 11$ seconds ($Q = 3.5$). The blue lines represent Bayes factors for an optical path glitch injection $\Phi_{12}^{\rm op}$ and the red lines an acceleration glitch injection $\Phi_{21}^{\rm ac}$. We see that in this high frequency regime there will be little issue in discriminating the origin of the signal. This figure also suggests, as seen before, that there might be a more of a challenge discriminating this optical path glitch from a burst.

6.6 Discussion

To realize the full discovery potential of the LISA observatory we need to be in a position to detect unexpected and unknown signals. We have developed a forward model for a wavelet basis to represent instrumental glitches and gravitational wave bursts, as first step towards this goal. Ideally, to separate unmodeled signals from noise, we would have multiple independent LISA observatories. We have shown that

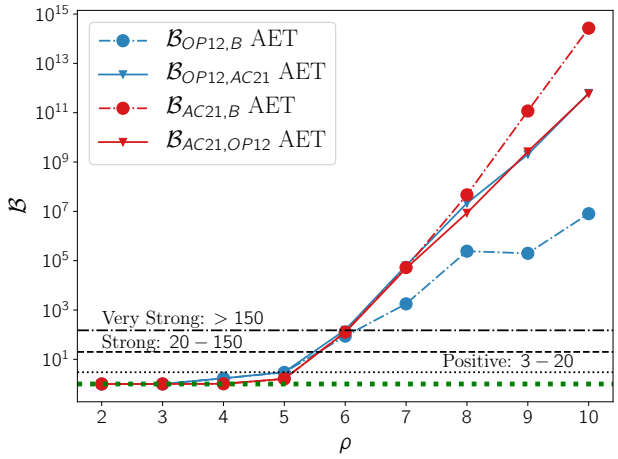

Figure 6.9: This figure displays Bayes factors for a high frequency glitch injections. The blue lines denote Bayes factors for an optical path glitch $\Phi_{12}^{\rm op}$ injection and the red lines denote a $\Phi_{21}^{\rm ac}$ injection.

the separation is possible with a single LISA detector, and even with a single TDI data channel, though the performance is much better when all three TDI channels are available. The properties of the signals and glitches can be recovered with good accuracy, though degeneracies in some parameters can degrade sky localization.

There are several extensions to that will need to be made to our analysis to handle generic glitches and signals. In our analysis we assumed the Gaussian noise level in each data channel were not only equal, but also that they were known. In reality power spectral density of the noise in each component will have to be estimated from the data, as was done in Refs. [242, 243]. We will also need to generalize the analysis to model non-stationary noise, a complication we know LISA will experience owed at least in part to the significant contribution to the noise by unresolved galactic binaries [24]. Our analysis in this paper took a quasi-Bayesian approach via the maximization of the likelihood over extrinsic parameters through the F-statistic. In the future a full marginalization will have to be done, though the the F-statistic could be used to produce very effective proposal distributions for the MCMC based on maps

of the F-statistic likelihood.

We will generalize our gravitational wave model such that the wave is not elliptically polarized. One last crucial extension is the use of multiple wavelets in the analysis [43]. Not only will we need to characterize multiple wavelets, but we will also need to marginalize over the number of wavelets in the data stream. Due to the sheer number of combinations of wavelets we expect in data stream we would expect the need to implement an effective Reversible Jump MCMC [248] to address the issue of determining an appropriate number of wavelets and for determining the evidence that a GW signal or an instrument glitch is present in the data. There may be additional information gathered by LISA in the form of instrument monitors. These could provide crucial information in characterizing glitches and assessing whether a glitch has indeed occurred.

Acknowledgments

TR and NJC appreciate the support of the NASA grant NNX16AB98G.

CONCLUSION

LISA is going to offer a wealth of science, but the fruit of the LISA mission will not be easy to harvest for a variety of reasons. The major obstacle to doing science which concerns us in this work is the development of a data analysis pipeline which models all facets of the LISA data stream *simultaneously*. The LISA data stream will contain a variety of astrophysical signals, Gaussian instrumental noise and glitches. Before we begin the development of specific pieces of the global analysis we must asses and study the expected population of relevant astrophysical sources. This assessment allows us to generate approximate goals for the global analysis pipeline. We must develop models for these signals and algorithms for characterizing the properties of the astrophysical sources. Since the parameter space for models of the data is daunting we must develop techniques by which to do the analysis fast. These speedups could come in the forms of smart algorithms and fast waveform generators. Additionally, we must be able to reconstruct waveforms for signals for which we have no model. This capability will aid us in the modeling of LISA's noise as there is guaranteed to be non-Gaussian features in the instrument noise for which we do not have models. Therefore, we need to incorporate the ability to characterize such deviations in noise in the global analysis. To reduce any bias in our characterization the astrophysical signals we will also have to model the noise levels of the detector and of the background noise from the tens of millions of unresolved galactic binaries. This simultaneous analysis must also ascertain how many sources are in the data and how complex of a model is needed to represent the source.

In this work we take several steps towards the development of such a global analysis. When considering an astrophysical source's signal as seen by LISA it is useful to generate sensitivity curves which quickly estimate the potential detection

of this source by LISA. In chapter 1 we lay out the foundations for calculating sensitivity curves for a variety of sources and estimating their signal-to-noise ratios. We generalized these curves to allow for sky dependence to be captured and we released the code for public consumption on GitHub. In making these estimates we must have an accurate estimate of the noise associated with the LISA data stream. In chapter 2 we provide estimates to the noise which results from the tens of millions of unresolved galactic binaries. These binaries are emitting gravitational waves to which LISA is sensitive, with too low signal-to-noise to be resolved, and will constitute a source of noise more powerful than LISA's instrumental noise. When calculating estimates for this contribution to the noise we simulate the process of resolving the loud binaries thus providing us with estimates for how many galactic binaries we can resolve. This process also predicts what we can learn about these binaries. In chapter 3 we theoretically address some of the complications that arise during a global analysis of the data. When one subtracts a resolved source from the data stream an error is introduced into the data because the estimated parameters are wrong. We find that this subtraction error results in signal characterization errors that are quite small in comparison to the natural variance induced by the instrumental noise. However, if we were to analyze the data source-by-source signals such as binary black holes which sweep across thousands of galactic binaries would have severely biased parameters. We show that when these signals are estimated simultaneously, instead of source by source, we merely get modestly inflated parameter errors.

Galactic binaries can posses relatively simple waveforms since standard processes that tighten or expand the binary, such as gravitational radiation or mass transfer, evolve the binary slowly, resulting in a nearly monochromatic signal modulated by LISA's motion. In chapter 4 we investigate the scenario when this picture gets complicated by the existence of a companion star to the galactic binary. This

hierarchical stellar system imparts center of mass acceleration to the inner binary, allowing us to extract information about the outer orbit. Since a substantial fraction of galactic binaries to which LISA will be sensitive will reside in hierarchical triples we expect LISA to provide a great opportunity to study the characteristics and evolution of hierarchical systems. The modeling of binary sources can also be complicated by orbital eccentricity. Chapter 5 develops a fast approximate model for the leading order gravitational wave signal for eccentric, non-spinning binaries, in the frequency domain. Since analysis is performed in the frequency domain creating a waveform already in the frequency domain typically results in a vast speed up in an analysis. Lastly, in chapter 6, we begin investigations in figuring out how to handle the case of an un-modeled waveform. We do this by investigating our ability to characterize a gravitational wave burst or instrumental glitch consisting of a single wavelet. Multiple wavelets can be used as a basis to represent a signal without a model. We additionally investigated our ability to distinguish wavelets of instrumental and astrophysical origin.

There remains a vast amount of work to be done towards the goal of developing a global analysis pipeline. We are currently developing an algorithm to solve a restricted version of the global analysis problem in which the signal consists of tens of millions isolated galactic binaries. This algorithm determines how many sources can be resolved and also estimates the combination of instrumental noise and unresolved galactic binary confusion noise. This code will be a significant first step towards the full global analysis code. In the future we will need to incorporate the ability to model a variety of types of sources simultaneously such as binary black holes or extreme mas ratio inspirals. Additionally, the code must be able to investigate more complex signals such as testing whether a galactic binary is part of a hierarchical stellar system, or perhaps whether it is eccentric. We must incorporate the ability to model

gravitational wave bursts and instrumental glitches which need to be represented by a superposition of wavelets. Currently, all of these analyses have taken place in the frequency domain. However, the confusion noise due to the unresolved galactic binaries will vary in time as the LISA antenna pattern moves over the galaxy, i.e. the confusion noise is non-stationary. A time-frequency representation of the data will therefore more appropriately represent the data. We must adapt our analyses and waveforms to this representation. Performing this all encompassing analysis will be challenging. Even with the development of waveforms directly in the domain of analysis and other various speed ups, the computational cost of solving the full global analysis problem will be substantial.

REFERENCES CITED

[1] Abbott B P e a (LIGO Scientific Collaboration and Virgo Collaboration) 2016 *Phys. Rev. Lett.* **116**(6) 061102 URL https://link.aps.org/doi/10.1103/PhysRevLett.116.061102

[2] Abbott B P e a (LIGO Scientific Collaboration and Virgo Collaboration) 2016 *Phys. Rev. Lett.* **116**(24) 241103 URL https://link.aps.org/doi/10.1103/PhysRevLett.116.241103

[3] Abbott B P e a (LIGO Scientific and Virgo Collaboration) 2017 *Phys. Rev. Lett.* **118**(22) 221101 URL https://link.aps.org/doi/10.1103/PhysRevLett.118.221101

[4] et al B P A 2017 *The Astrophysical Journal Letters* **851** L35 URL http://stacks.iop.org/2041-8205/851/i=2/a=L35

[5] Abbott B P e a (LIGO Scientific Collaboration and Virgo Collaboration) 2017 *Phys. Rev. Lett.* **119**(14) 141101 URL https://link.aps.org/doi/10.1103/PhysRevLett.119.141101

[6] Abbott B P e a (LIGO Scientific Collaboration and Virgo Collaboration) 2017 *Phys. Rev. Lett.* **119**(16) 161101 URL https://link.aps.org/doi/10.1103/PhysRevLett.119.161101

[7] Collaboration T L S, Collaboration T V, Collaboration T M, Collaboration T D E C G E, the DES Collaboration, Collaboration T D, Collaboration T L C O, Collaboration T V and Collaboration T M 2017 *Nature* **551** 85–88 URL http://dx.doi.org/10.1038/nature24471

[8] et al (LIGO Scientific Collaboration B P A and Collaboration) V 2017 *The Astrophysical Journal Letters* **850** L39 URL http://stacks.iop.org/2041-8205/850/i=2/a=L39

[9] et al (LIGO Scientific Collaboration B P A and Collaboration) V 2017 *The Astrophysical Journal Letters* **850** L40 URL http://stacks.iop.org/2041-8205/850/i=2/a=L40

[10] The LIGO Scientific Collaboration, the Virgo Collaboration, Abbott B P, Abbott R, Abbott T D, Acernese F, Ackley K, Adams C, Adams T, Addesso P and et al 2018 *ArXiv e-prints* (*Preprint* 1805.11581)

[11] The LIGO Scientific Collaboration, The Virgo Collaboration and Weinberg N N 2018 *ArXiv e-prints* (*Preprint* 1808.08676)

[12] Drout M R e a 2017 *Science* **358** 1570–1574 ISSN 0036-8075 (*Preprint* http://science.sciencemag.org/content/358/6370/1570.full.pdf) URL http://science.sciencemag.org/content/358/6370/1570

[13] et al A G 2017 *The Astrophysical Journal Letters* **848** L14 URL http://stacks.iop.org/2041-8205/848/i=2/a=L14

[14] Nissanke S, Vallisneri M, Nelemans G and Prince T A 2012 *The Astrophysical Journal* **758** 131 URL http://stacks.iop.org/0004-637X/758/i=2/a=131

[15] Del Pozzo W, Sesana A and Klein A 2018 *Monthly Notices of the Royal Astronomical Society* **475** 3485–3492 (*Preprint* /oup/backfile/content_public/journal/mnras/475/3/10.1093_mnras_sty057/1/sty057.pdf) URL http://dx.doi.org/10.1093/mnras/sty057

[16] Gair J R, Mandel I, Sesana A and Vecchio A 2009 *Classical and Quantum Gravity* **26** 204009 URL http://stacks.iop.org/0264-9381/26/i=20/a=204009

[17] Gair J R, Barack L, Creighton T, Cutler C, Larson S L, Phinney E S and Vallisneri M 2004 *Classical and Quantum Gravity* **21** S1595 URL http://stacks.iop.org/0264-9381/21/i=20/a=003

[18] Kupfer T, Korol V, Shah S, Nelemans G, Marsh T R, Ramsay G, Groot P J, Steeghs D T H and Rossi E M 2018 *Monthly Notices of the Royal Astronomical Society* **480** 302–309 (*Preprint* /oup/backfile/content_public/journal/mnras/480/1/10.1093_mnras_sty1545/1/sty1545.pdf) URL http://dx.doi.org/10.1093/mnras/sty1545

[19] Amaro-Seoane P e a 2017 *ArXiv e-prints* (*Preprint* 1702.00786)

[20] Armano M e a 2018 *Phys. Rev. Lett.* **120**(6) 061101 URL https://link.aps.org/doi/10.1103/PhysRevLett.120.061101

[21] Armano M e a 2016 *Phys. Rev. Lett.* **116**(23) 231101 URL https://link.aps.org/doi/10.1103/PhysRevLett.116.231101

[22] Ruiter A J, Belczynski K, Benacquista M, Larson S L and Williams G 2010 *The Astrophysical Journal* **717** 1006 URL http://stacks.iop.org/0004-637X/717/i=2/a=1006

[23] Nelemans G 2009 *Classical and Quantum Gravity* **26** 094030 URL http://stacks.iop.org/0264-9381/26/i=9/a=094030

[24] Cornish N and Robson T 2017 *Journal of Physics: Conference Series* **840** 012024 URL http://stacks.iop.org/1742-6596/840/i=1/a=012024

[25] Adams M R, Cornish N J and Littenberg T B 2012 *Phys. Rev. D* **86**(12) 124032 URL https://link.aps.org/doi/10.1103/PhysRevD.86.124032

[26] Nelemans G 2003 *AIP Conference Proceedings* **686** 263–272 (*Preprint* https://aip.scitation.org/doi/pdf/10.1063/1.1629441) URL https://aip.scitation.org/doi/abs/10.1063/1.1629441

[27] Robson T, Cornish N J, Tamanini N and Toonen S 2018 *ArXiv e-prints* (*Preprint* 1806.00500)

[28] Bender P L and Hils D 1997 *Classical and Quantum Gravity* **14** 1439 URL http://stacks.iop.org/0264-9381/14/i=6/a=008

[29] Cornish N J 2003 *ArXiv General Relativity and Quantum Cosmology e-prints* (*Preprint* gr-qc/0304020)

[30] Owen B J and Sathyaprakash B S 1999 *Phys. Rev. D* **60**(2) 022002 URL https://link.aps.org/doi/10.1103/PhysRevD.60.022002

[31] Usman S A, Nitz A H, Harry I W, Biwer C M, Brown D A, Cabero M, Capano C D, Canton T D, Dent T, Fairhurst S, Kehl M S, Keppel D, Krishnan B, Lenon A, Lundgren A, Nielsen A B, Pekowsky L P, Pfeiffer H P, Saulson P R, West M and Willis J L 2016 *Classical and Quantum Gravity* **33** 215004 URL http://stacks.iop.org/0264-9381/33/i=21/a=215004

[32] Cornish N J and Porter E K 2005 *Classical and Quantum Gravity* **22** S927 URL http://stacks.iop.org/0264-9381/22/i=18/a=S06

[33] Babak S, Baker J G, Benacquista M J, Cornish N J, Larson S L, Mandel I, McWilliams S T, Petiteau A, Porter E K, Robinson E L, Vallisneri M, Vecchio A, the Mock LISA Data Challenge Task Force, Adams M, Arnaud K A, Błaut A, Bridges M, Cohen M, Cutler C, Feroz F, Gair J R, Graff P, Hobson M, Key J S, Królak A, Lasenby A, Prix R, Shang Y, Trias M, Veitch J, Whelan J T and the Challenge 3 participants 2010 *Classical and Quantum Gravity* **27** 084009 URL http://stacks.iop.org/0264-9381/27/i=8/a=084009

[34] Arnaud K A, Babak S, Baker J G, Benacquista M J, Cornish N J, Cutler C, Larson S L, Sathyaprakash B S, Vallisneri M, Vecchio A and Vinet J 2006 *AIP Conference Proceedings* **873** 619–624 (*Preprint* https://aip.scitation.org/doi/pdf/10.1063/1.2405108) URL https://aip.scitation.org/doi/abs/10.1063/1.2405108

[35] Arnaud K A, Babak S, Baker J G, Benacquista M J, Cornish N J, Cutler C, Finn L S, Larson S L, Littenberg T, Porter E K, Vallisneri M, Vecchio A, Vinet J Y and Force T M L D C T 2007 *Classical and Quantum Gravity* **24** S551 URL http://stacks.iop.org/0264-9381/24/i=19/a=S18

[36] Cornish N J and Larson S L 2003 *Phys. Rev. D* **67**(10) 103001 URL https://link.aps.org/doi/10.1103/PhysRevD.67.103001

[37] Högbom J A 1974 *AAS* **15** 417

[38] Rubbo L J, Cornish N J and Hellings R W 2006 *AIP Conference Proceedings* **873** 489–493 (*Preprint* https://aip.scitation.org/doi/pdf/10.1063/1.2405089) URL https://aip.scitation.org/doi/abs/10.1063/1.2405089

[39] Crowder J and Cornish N J 2007 *Phys. Rev. D* **75**(4) 043008 URL https://link.aps.org/doi/10.1103/PhysRevD.75.043008

[40] Cornish N J and Crowder J 2005 *Phys. Rev. D* **72**(4) 043005 URL https://link.aps.org/doi/10.1103/PhysRevD.72.043005

[41] Littenberg T B 2011 *Phys. Rev. D* **84**(6) 063009 URL https://link.aps.org/doi/10.1103/PhysRevD.84.063009

[42] Babak S 2017 *Journal of Physics: Conference Series* **840** 012026 URL http://stacks.iop.org/1742-6596/840/i=1/a=012026

[43] Cornish N J and Littenberg T B 2015 *Classical and Quantum Gravity* **32** 135012 URL http://stacks.iop.org/0264-9381/32/i=13/a=135012

[44] Dupuis R J and Woan G 2005 *Phys. Rev. D* **72**(10) 102002 URL https://link.aps.org/doi/10.1103/PhysRevD.72.102002

[45] McHugh M P, Zalamansky G, Vernotte F and Lantz E 1996 *Phys. Rev. D* **54**(10) 5993–6000 URL https://link.aps.org/doi/10.1103/PhysRevD.54.5993

[46] Veitch J and Vecchio A 2008 *Classical and Quantum Gravity* **25** 184010 URL http://stacks.iop.org/0264-9381/25/i=18/a=184010

[47] Hastings W K 1970 *Biometrika* **57** 97–109 (*Preprint* /oup/backfile/content_public/journal/biomet/57/1/10.1093_biomet_57.1.97/1/57-1-97.pdf) URL http://dx.doi.org/10.1093/biomet/57.1.97

[48] Swendsen R H and Wang J S 1986 *Phys. Rev. Lett.* **57**(21) 2607–2609 URL https://link.aps.org/doi/10.1103/PhysRevLett.57.2607

[49] GREEN P J 1995 *Biometrika* **82** 711–732 (*Preprint* /oup/backfile/content_public/journal/biomet/82/4/10.1093/biomet/82.4.711/2/82-4-711.pdf) URL http://dx.doi.org/10.1093/biomet/82.4.711

[50] Cornish N J and Littenberg T B 2007 *Phys. Rev. D* **76**(8) 083006 URL https://link.aps.org/doi/10.1103/PhysRevD.76.083006

[51] Timpano S E, Rubbo L J and Cornish N J 2006 *Phys. Rev. D* **73**(12) 122001 URL https://link.aps.org/doi/10.1103/PhysRevD.73.122001

[52] Randall L and Xianyu Z Z 2018 *ArXiv e-prints* (*Preprint* 1805.05335)

[53] Katz M and Larson S 2018 Probing Massive Black Hole Populations and Their Environments with LISA *American Astronomical Society Meeting Abstracts #231* (*American Astronomical Society Meeting Abstracts* vol 231) p 255.13

[54] Willems B, Kalogera V, Vecchio A, Ivanova N, Rasio F A, Fregeau J M and Belczynski K 2007 *The Astrophysical Journal Letters* **665** L59 URL http://stacks.iop.org/1538-4357/665/i=1/a=L59

[55] Raghavan D, McAlister H A, Henry T J, Latham D W, Marcy G W, Mason B D, Gies D R, White R J and ten Brummelaar T A 2010 *The Astrophysical Journal Supplement Series* **190** 1 URL http://stacks.iop.org/0067-0049/190/i=1/a=1

[56] Tokovinin A 2014 *The Astronomical Journal* **147** 87 URL http://stacks.iop.org/1538-3881/147/i=4/a=87

[57] Benacquista M 2001 *AIP Conference Proceedings* **586** 793–795 (*Preprint* https://aip.scitation.org/doi/pdf/10.1063/1.1419658) URL https://aip.scitation.org/doi/abs/10.1063/1.1419658

[58] Kowalska I, Bulik T, Belczynski K, Dominik M and Gondek-Rosinska D 2011 *AAS* **527** A70 (*Preprint* 1010.0511)

[59] Cornish N and Robson T 2018 *ArXiv e-prints* (*Preprint* 1803.01944)

[60] Robson T and Cornish N 2017 *Classical and Quantum Gravity* **34** 244002 URL http://stacks.iop.org/0264-9381/34/i=24/a=244002

[61] Moore B, Robson T, Loutrel N and Yunes N 2018 *ArXiv e-prints* (*Preprint* 1807.07163)

[62] Robson T and Cornish N J 2018 *ArXiv e-prints* (*Preprint* 1811.04490)

[63] Audley H *et al.* 2017 (*Preprint* 1702.00786)

[64] Moore C J, Cole R H and Berry C P L 2015 *Class. Quant. Grav.* **32** 015014 (*Preprint* 1408.0740)

[65] Robson T, Cornish N and Chang L 2018 LISA sensitivity calculator github.com/eXtremeGravityInstitute/LISA_Sensitivity.git

[66] Robson T, Cornish N and Chang L 2018 Mybinder: LISA sensitivity mybinder.org/v2/gh/eXtremeGravityInstitute/LISA_Sensitivity.git/master

[67] Larson S L, Hiscock W A and Hellings R W 2000 *Phys. Rev.* **D62** 062001 (*Preprint* gr-qc/9909080)

[68] Cornish N J 2002 *Phys. Rev.* **D65** 022004 (*Preprint* gr-qc/0106058)

[69] Prince T A, Tinto M, Larson S L and Armstrong J W 2002 *Phys. Rev.* **D66** 122002 (*Preprint* gr-qc/0209039)

[70] Cornish N and Robson T 2017 Galactic binary science with the new LISA design *11th International LISA Symposium Zurich, Switzerland, September 5-9, 2016* (*Preprint* 1703.09858) URL http://inspirehep.net/record/1519956/files/arXiv:1703.09858.pdf

[71] Finn L S and Chernoff D F 1993 *Phys. Rev.* **D47** 2198–2219 (*Preprint* gr-qc/9301003)

[72] Ajith P *et al.* 2007 *Class. Quant. Grav.* **24** S689–S700 (*Preprint* 0704.3764)

[73] Hannam M 2014 *Gen. Rel. Grav.* **46** 1767 (*Preprint* 1312.3641)

[74] Schmidt P, Ohme F and Hannam M 2015 *Phys. Rev.* **D91** 024043 (*Preprint* 1408.1810)

[75] Chua A J K, Moore C J and Gair J R 2017 *Phys. Rev.* **D96** 044005 (*Preprint* 1705.04259)

[76] Chua A J K, Moore C J and Gair J R 2017–2018 EMRI Kludge Suite github.com/alvincjk/EMRI_Kludge_Suite

[77] Vallisneri M and Galley C R 2012 *Class. Quant. Grav.* **29** 124015 (*Preprint* 1201.3684)

[78] Rubbo L J, Cornish N J and Poujade O 2004 *Phys. Rev. D* **69**(8) 082003 URL https://link.aps.org/doi/10.1103/PhysRevD.69.082003

[79] Cutler C 1998 *Phys. Rev. D* **57**(12) 7089–7102 URL https://link.aps.org/doi/10.1103/PhysRevD.57.7089

[80] Danzmann K *et al* 2016 Laser interferometer space antenna Tech. rep. LISA Consortium URL https://www.elisascience.org/files/publications/LISA_L3_20170120.pdf

[81] Bender P, Brillet A, Ciufolini I, Cruise A, Cutler C, Danzmann K, Fidecaro F, Folkner W, Hough J, McNamara P, Peterseim M, Robertson D, Rodrigues M, Rüdiger A, Sandford M, Schäfer G, Schilling R, Schutz B, Speake C, Stebbins R, Sumner T, Touboul P, Vinet J, Vitale S, Ward H and Winkler W 1998 Laser interferometer space antenna for the detection and observation of gravitational

waves Tech. rep. NASA mPQ 233, Second Edition URL https://lisa.gsfc.nasa.gov/Documentation/ppa2.08.pdf

[82] Petiteau A, Hewitson M, Heinzel G, Fitzsimons E and Halloin H 2016 Lisa noise budget Tech. rep. LISA Consortium lISA-CST-TN-0001

[83] Nissanke S, Vallisneri M, Nelemans G and Prince T A 2012 *Astrophys. J.* **758** 131 (*Preprint* 1201.4613)

[84] Toonen S, Nelemans G and Portegies Zwart S 2012 A&A **546** A70 (*Preprint* 1208.6446)

[85] Nelemans G and Tout C A 2005 MNRAS **356** 753–764 (*Preprint* astro-ph/0410301)

[86] Roelofs G H A, Nelemans G and Groot P J 2007 MNRAS **382** 685–692 (*Preprint* 0709.2951)

[87] Carter P J, Marsh T R, Steeghs D, Groot P J, Nelemans G, Levitan D, Rau A, Copperwheat C M, Kupfer T and Roelofs G H A 2013 MNRAS **429** 2143–2160 (*Preprint* 1211.6439)

[88] Littenberg T B and Cornish N J 2015 *Phys. Rev. D* **91**(8) 084034 URL http://link.aps.org/doi/10.1103/PhysRevD.91.084034

[89] Arnaud K A *et al.* 2006 *AIP Conf. Proc.* **873** 619–624 [,619(2006)] (*Preprint* gr-qc/0609105)

[90] Arnaud K A *et al.* 2007 *Class. Quant. Grav.* **24** S529–S540 (*Preprint* gr-qc/0701139)

[91] Arnaud K A *et al.* 2007 *Class. Quant. Grav.* **24** S551–S564 (*Preprint* gr-qc/0701170)

[92] Babak S *et al.* (Mock LISA Data Challenge Task Force) 2008 *Class. Quant. Grav.* **25** 114037 (*Preprint* 0711.2667)

[93] Babak S *et al.* 2008 *Class. Quant. Grav.* **25** 184026 (*Preprint* 0806.2110)

[94] Babak S *et al.* (Mock LISA Data Challenge Task Force) 2010 *Class. Quant. Grav.* **27** 084009 (*Preprint* 0912.0548)

[95] Cutler C and Harms J 2006 *Phys. Rev. D* **73**(4) 042001 URL https://link.aps.org/doi/10.1103/PhysRevD.73.042001

[96] Crowder J and Cornish N J 2004 *Phys. Rev.* **D70** 082004 (*Preprint* gr-qc/0404129)

[97] Hils D, Bender P L and Webbink R F 1990 ApJ **360** 75–94

[98] Hils D and Bender P L 2000 *The Astrophysical Journal* **537** 334 URL http://stacks.iop.org/0004-637X/537/i=1/a=334

[99] Armstrong J W, Estabrook F B and Tinto M 1999 *The Astrophysical Journal* **527** 814 URL http://stacks.iop.org/0004-637X/527/i=2/a=814

[100] Vallisneri M 2008 *Phys. Rev. D* **77**(4) 042001 URL https://link.aps.org/doi/10.1103/PhysRevD.77.042001

[101] Cutler C and Flanagan É E 1994 *Phys. Rev. D* **49** 2658–2697 URL https://link.aps.org/doi/10.1103/PhysRevD.49.2658

[102] Takahashi R and Seto N 2002 *The Astrophysical Journal* **575** 1030 URL http://stacks.iop.org/0004-637X/575/i=2/a=1030

[103] Adams M R, Cornish N J and Littenberg T B 2012 *Phys. Rev.* **D86** 124032 (*Preprint* 1209.6286)

[104] Littenberg T B and Cornish N J 2015 *Phys. Rev.* **D91** 084034 (*Preprint* 1410.3852)

[105] Umstätter R, Christensen N, Hendry M, Meyer R, Simha V, Veitch J, Vigeland S and Woan G 2005 *Classical and Quantum Gravity* **22** S901 URL http://stacks.iop.org/0264-9381/22/i=18/a=S04

[106] Raghavan D, McAlister H A, Henry T J, Latham D W, Marcy G W, Mason B D, Gies D R, White R J and ten Brummelaar T A 2010 ApJS **190** 1–42 (*Preprint* 1007.0414)

[107] Tokovinin A 2014 AJ **147** 87 (*Preprint* 1401.6827)

[108] Fuhrmann K, Chini R, Kaderhandt L and Chen Z 2017 ApJ **836** 139

[109] Tokovinin A, Thomas S, Sterzik M and Udry S 2006 A&A **450** 681–693 (*Preprint* astro-ph/0601518)

[110] Pribulla T and Rucinski S M 2006 AJ **131** 2986–3007 (*Preprint* astro-ph/0601610)

[111] Lidov M L 1962 Planet. Space Sci. **9** 719–759

[112] Kozai Y 1962 AJ **67** 591

[113] Naoz S 2016 ARA&A **54** 441–489 (*Preprint* 1601.07175)

[114] Thompson T A 2011 ApJ **741** 82 (*Preprint* 1011.4322)

[115] Fang X, Thompson T A and Hirata C M 2017 (*Preprint* 1709.08682)

[116] Toonen S, Perets H B and Hamers A S 2018 *Astron. Astrophys.* **610** A22 (*Preprint* 1709.00422)

[117] Fabrycky D and Tremaine S 2007 ApJ **669** 1298–1315 (*Preprint* 0705.4285)

[118] Naoz S and Fabrycky D C 2014 ApJ **793** 137 (*Preprint* 1405.5223)

[119] Silsbee K and Tremaine S 2017 ApJ **836** 39 (*Preprint* 1608.07642)

[120] Antonini F, Toonen S and Hamers A S 2017 *ArXiv e-prints* (*Preprint* 1703.06614)

[121] Yunes N, Coleman Miller M and Thornburg J 2011 *Phys. Rev.* **D83** 044030 (*Preprint* 1010.1721)

[122] Meiron Y, Kocsis B and Loeb A 2017 *Astrophys. J.* **834** 200 (*Preprint* 1604.02148)

[123] Bonvin C, Caprini C, Sturani R and Tamanini N 2017 *Phys. Rev.* **D95** 044029 (*Preprint* 1609.08093)

[124] Inayoshi K, Tamanini N, Caprini C and Haiman Z 2017 *Phys. Rev.* **D96** 063014 (*Preprint* 1702.06529)

[125] Randall L and Xianyu Z Z 2018 (*Preprint* 1805.05335)

[126] Edwards R T, Hobbs G B and Manchester R N 2006 *Mon. Not. Roy. Astron. Soc.* **372** 1549–1574 (*Preprint* astro-ph/0607664)

[127] Tokovinin A 2014 AJ **147** 87 (*Preprint* 1401.6827)

[128] Sana H, Le Bouquin J B, Lacour S, Berger J P, Duvert G, Gauchet L, Norris B, Olofsson J, Pickel D, Zins G, Absil O, de Koter A, Kratter K, Schnurr O and Zinnecker H 2014 ApJS **215** 15 (*Preprint* 1409.6304)

[129] Evans N R, Carpenter K G, Robinson R, Kienzle F and Dekas A E 2005 AJ **130** 789–793 (*Preprint* astro-ph/0504169)

[130] Tokovinin A 2014 AJ **147** 86 (*Preprint* 1401.6825)

[131] Thompson T A 2011 ApJ **741** 82 (*Preprint* 1011.4322)

[132] Perets H B and Kratter K M 2012 ApJ **760** 99 (*Preprint* 1203.2914)

[133] Hamers A S, Pols O R, Claeys J S W and Nelemans G 2013 MNRAS **430** 2262–2280 (*Preprint* 1301.1469)

[134] Toonen S, Hamers A and Portegies Zwart S 2016 *Computational Astrophysics and Cosmology* **3** 6 (*Preprint* 1612.06172)

[135] Toonen S, Perets H B and Hamers A S 2018 A&A **610** A22 (*Preprint* 1709.00422)

[136] Antonini F and Perets H B 2012 ApJ **757** 27 (*Preprint* 1203.2938)

[137] Prodan S, Antonini F and Perets H B 2015 ApJ **799** 118 (*Preprint* 1405.6029)

[138] Stephan A P, Naoz S, Ghez A M, Witzel G, Sitarski B N, Do T and Kocsis B 2016 MNRAS **460** 3494–3504 (*Preprint* 1603.02709)

[139] VanLandingham J H, Miller M C, Hamilton D P and Richardson D C 2016 ApJ **828** 77 (*Preprint* 1604.04948)

[140] Hoang B M, Naoz S, Kocsis B, Rasio F A and Dosopoulou F 2018 ApJ **856** 140 (*Preprint* 1706.09896)

[141] Eggleton P P and Verbunt F 1986 MNRAS **220** 13P–18P

[142] Ivanova N, Chaichenets S, Fregeau J, Heinke C O, Lombardi Jr J C and Woods T E 2010 ApJ **717** 948 (*Preprint* 1001.1767)

[143] Naoz S, Fragos T, Geller A, Stephan A P and Rasio F A 2016 ApJ **822** L24 (*Preprint* 1510.02093)

[144] Perets H B and Fabrycky D C 2009 ApJ **697** 1048–1056 (*Preprint* 0901.4328)

[145] Kiseleva L G, Eggleton P P and Mikkola S 1998 MNRAS **300** 292–302

[146] Mazeh T and Shaham J 1979 A&A **77** 145–151

[147] Correia A C M, Laskar J, Farago F and Boué G 2011 *Celestial Mechanics and Dynamical Astronomy* **111** 105–130 (*Preprint* 1107.0736)

[148] Petrovich C 2015 ApJ **799** 27 (*Preprint* 1405.0280)

[149] Wu Y and Murray N 2003 ApJ **589** 605–614 (*Preprint* astro-ph/0303010)

[150] Naoz S, Farr W M, Lithwick Y, Rasio F A and Teyssandier J 2011 Nature **473** 187–189 (*Preprint* 1011.2501)

[151] Naoz S, Farr W M and Rasio F A 2012 ApJ **754** L36 (*Preprint* 1206.3529)

[152] Dawson R I and Chiang E 2014 *Science* **346** 212–216 (*Preprint* 1410.2604)

[153] Katz B and Dong S 2012 *ArXiv e-prints* (*Preprint* 1211.4584)

[154] Toonen S, Hollands M, Gänsicke B T and Boekholt T 2017 A&A **602** A16 (*Preprint* 1703.06893)

[155] Gültekin K, Miller M C and Hamilton D P 2006 ApJ **640** 156–166 (*Preprint* astro-ph/0509885)

[156] Seto N 2013 *Physical Review Letters* **111** 061106 (*Preprint* 1304.5151)

[157] Randall L and Xianyu Z Z 2018 *ArXiv e-prints* (*Preprint* 1802.05718)

[158] Klein Y Y and Katz B 2017 MNRAS **465** L44–L48 (*Preprint* 1609.03580)

[159] Hollands M A, Tremblay P E, Gaensicke B T, Gentile-Fusillo N P and Toonen S 2018 *ArXiv e-prints* (*Preprint* 1805.12590)

[160] Ransom S M, Stairs I H, Archibald A M, Hessels J W T, Kaplan D L, van Kerkwijk M H, Boyles J, Deller A T, Chatterjee S, Schechtman-Rook A, Berndsen A, Lynch R S, Lorimer D R, Karako-Argaman C, Kaspi V M, Kondratiev V I, McLaughlin M A, van Leeuwen J, Rosen R, Roberts M S E and Stovall K 2014 Nature **505** 520–524 (*Preprint* 1401.0535)

[161] Tauris T M and van den Heuvel E P J 2014 ApJ **781** L13 (*Preprint* 1401.0941)

[162] Sabach E and Soker N 2015 MNRAS **450** 1716–1723 (*Preprint* 1501.06787)

[163] Ford E B, Kozinsky B and Rasio F A 2000 ApJ **535** 385–401

[164] Harrington R S 1968 AJ **73** 190–194

[165] Will C M 2017 *Phys. Rev. D* **96**(2) 023017 URL https://link.aps.org/doi/10.1103/PhysRevD.96.023017

[166] Kozai Y 1962 AJ **67** 591

[167] Lidov M L 1962 Planet. Space Sci. **9** 719–759

[168] Naoz S and Fabrycky D C 2014 *The Astrophysical Journal* **793** 137 URL http://stacks.iop.org/0004-637X/793/i=2/a=137

[169] Seto N 2002 *Monthly Notices of the Royal Astronomical Society* **333**

[170] Marsh T R, Nelemans G and Steeghs D 2004 *Monthly Notices of the Royal Astronomical Society* **350**

[171] Cornish N J and Littenberg T B 2007 *Phys. Rev. D* **76**(8) 083006 URL https://link.aps.org/doi/10.1103/PhysRevD.76.083006

[172] Armstrong J W, Estabrook F B and Tinto M 1999 *The Astrophysical Journal* **527** 814 URL http://stacks.iop.org/0004-637X/527/i=2/a=814

[173] Robson T and Cornish N 2017 *Classical and Quantum Gravity* **34** 244002 URL http://stacks.iop.org/0264-9381/34/i=24/a=244002

[174] Cornish N and Robson T 2018 (*Preprint* 1803.01944)

[175] Moore B, Robson T, Loutrel N and Yunes N 2018 (*Preprint* 1807.07163)

[176] Cornish N J and Crowder J 2005 *Phys. Rev. D* **72**(4) 043005 URL https://link.aps.org/doi/10.1103/PhysRevD.72.043005

[177] Key J S and Cornish N J 2011 *Phys. Rev. D* **83**(8) 083001 URL https://link.aps.org/doi/10.1103/PhysRevD.83.083001

[178] Cutler C 1998 *Phys. Rev. D* **57**(12) 7089–7102 URL https://link.aps.org/doi/10.1103/PhysRevD.57.7089

[179] Chatziioannou K, Klein A, Cornish N and Yunes N 2017 *Phys. Rev. Lett.* **118**(5) 051101 URL https://link.aps.org/doi/10.1103/PhysRevLett.118.051101

[180] LIGO Scientific Collaboration, Aasi J, Abbott B P, Abbott R, Abbott T, Abernathy M R, Ackley K, Adams C, Adams T, Addesso P and et al 2015 *Classical and Quantum Gravity* **32** 074001 (*Preprint* 1411.4547)

[181] Acernese F *et al.* (VIRGO) 2015 *Class.Quant.Grav.* **32** 024001 (*Preprint* 1408.3978)

[182] LIGO-India http://www.gw-indigo.org

[183] Aso Y, Michimura Y, Somiya K, Ando M, Miyakawa O *et al.* 2013 (*Preprint* 1306.6747)

[184] Abbott, B P et al (LIGO Scientific Collaboration and Virgo Collaboration) 2016 *Phys. Rev. Lett.* **116**(6) 061102 URL http://link.aps.org/doi/10.1103/PhysRevLett.116.061102

[185] Abbott B P, Abbott R, Abbott T D, Abernathy M R, Acernese F, Ackley K, Adams C, Adams T, Addesso P, Adhikari R X and et al 2016 *Physical Review Letters* **116** 241103 (*Preprint* 1606.04855)

[186] Abbott B P, Abbott R, Abbott T D, Acernese F, Ackley K, Adams C, Adams T, Addesso P, Adhikari R X, Adya V B and et al 2017 *Physical Review Letters* **118** 221101 (*Preprint* 1706.01812)

[187] Abbott B P, Abbott R, Abbott T D, Acernese F, Ackley K, Adams C, Adams T, Addesso P, Adhikari R X, Adya V B and et al 2017 *Physical Review Letters* **119** 141101 (*Preprint* 1709.09660)

[188] Abbott B P, Abbott R, Abbott T D, Acernese F, Ackley K, Adams C, Adams T, Addesso P, Adhikari R X, Adya V B and et al 2017 *Physical Review Letters* **119** 161101 (*Preprint* 1710.05832)

[189] Abbott B P, Abbott R, Abbott T D, Acernese F, Ackley K, Adams C, Adams T, Addesso P, Adhikari R X, Adya V B and et al 2017 ApJ **851** L35 (*Preprint* 1711.05578)

[190] Nishizawa A, Sesana A, Berti E and Klein A 2017 MNRAS **465** 4375–4380 (*Preprint* 1606.09295)

[191] Nishizawa A, Berti E, Klein A and Sesana A 2016 *Phys. Rev. D* **94**(6) 064020 URL https://link.aps.org/doi/10.1103/PhysRevD.94.064020

[192] Seto N 2013 *Physical Review Letters* **111** 061106 (*Preprint* 1304.5151)

[193] Antognini J M, Shappee B J, Thompson T A and Amaro-Seoane P 2013 *ArXiv e-prints* (*Preprint* 1308.5682)

[194] Naoz S, Kocsis B, Loeb A and Yunes N 2013 *The Astrophysical Journal* **773** 187 (*Preprint* 1206.4316)

[195] Antonini F, Murray N and Mikkola S 2014 *The Astrophysical Journal* **781** 45 (*Preprint* 1308.3674)

[196] Rodriguez C L, Amaro-Seoane P, Chatterjee S and Rasio F A 2018 *Physical Review Letters* **120** 151101 (*Preprint* 1712.04937)

[197] Samsing J 2018 Phys. Rev. D **97** 103014 (*Preprint* 1711.07452)

[198] Blanchet L 2006 *Living Rev. Rel.* **9** 4

[199] O'Leary R M, Kocsis B and Loeb A 2009 MNRAS **395** 2127–2146 (*Preprint* 0807.2638)

[200] Samsing J, MacLeod M and Ramirez-Ruiz E 2013 (*Preprint* 1308.2964)

[201] Antonini F, Chatterjee S, Rodriguez C L, Morscher M, Pattabiraman B, Kalogera V and Rasio F A 2015 *ArXiv e-prints* (*Preprint* 1509.05080)

[202] Martel K and Poisson E 1999 *Phys. Rev. D.* **60** 124008 (*Preprint* gr-qc/9907006)

[203] Ajith P 2008 *Classical and Quantum Gravity* **25** 114033 (*Preprint* 0712.0343)

[204] Huerta E A and Brown D A 2013 *Phys. Rev.* **D87** 127501 (*Preprint* 1301.1895)

[205] Brown D A and Zimmerman P J 2010 *Phys. Rev. D.* **81** 024007 (*Preprint* 0909.0066)

[206] Cokelaer T and Pathak D 2009 *Classical and Quantum Gravity* **26** 045013 (*Preprint* 0903.4791)

[207] Favata M 2013 *ArXiv e-prints* (*Preprint* 1310.8288)

[208] Einstein Telescope www.et-gw.eu

[209] Chatziioannou K, Klein A, Yunes N and Cornish N 2017 Phys. Rev. D **95** 104004 (*Preprint* 1703.03967)

[210] Yunes N, Arun K G, Berti E and Will C M 2009 *Phys. Rev. D* **80**(8) 084001 URL http://link.aps.org/doi/10.1103/PhysRevD.80.084001

[211] Tanay S, Haney M and Gopakumar A 2016 *Phys. Rev.* **D93** 064031 (*Preprint* 1602.03081)

[212] Moore B, Favata M, Arun K G and Mishra C K 2016 *Phys. Rev.* **D93** 124061 (*Preprint* 1605.00304)

[213] Pierro V and Pinto I M 1996 *Nuovo Cimento B Serie* **111** 631–644

[214] Mikóczi B, Kocsis B, Forgács P and Vasúth M 2012 Phys. Rev. D **86** 104027 (*Preprint* 1206.5786)

[215] Mikóczi B, Forgács P and Vasúth M 2015 Phys. Rev. D **92** 044038 (*Preprint* 1502.00276)

[216] Moreno-Garrido C, Mediavilla E and Buitrago J 1995 MNRAS **274** 115–126

[217] Kovalevsky J 1967 *Introduction to Celestial Mechanics* (New York: Springer)

[218] Peters P C and Mathews J 1963 *Physical Review* **131** 435–440

[219] Cutler C, Kennefick D and Poisson E 1994 *Phys. Rev.* **D50** 3816–3835

[220] Levin J and Perez-Giz G 2008 Phys. Rev. D **77** 103005 (*Preprint* 0802.0459)

[221] Komorowski P G, Valluri S R and Houde M 2009 *Classical and Quantum Gravity* **26** 085001 (*Preprint* 0903.3684)

[222] Abbott B P, Abbott R, Abbott T D, Abernathy M R, Ackley K, Adams C, Addesso P, Adhikari R X, Adya V B, Affeldt C and et al 2017 *Classical and Quantum Gravity* **34** 044001 (*Preprint* 1607.08697)

[223] Huerta E, Kumar P, McWilliams S T, O'Shaughnessy R and Yunes N 2014 *Phys.Rev.* **D90** 084016 (*Preprint* 1408.3406)

[224] Klein A, Boetzel Y, Gopakumar A, Jetzer P and de Vittori L 2018 *ArXiv e-prints* (*Preprint* 1801.08542)

[225] Buonanno A, Iyer B R, Ochsner E, Pan Y and Sathyaprakash B S 2009 Phys. Rev. D **80** 084043 (*Preprint* 0907.0700)

[226] Turner M 1977 ApJ **216** 610–619

[227] Wen L 2003 *The Astrophysical Journal* **598** 419–430 (*Preprint* astro-ph/0211492)

[228] Droz S, Knapp D J, Poisson E and Owen B J 1999 *Phys.Rev.* **D59** 124016 (*Preprint* gr-qc/9901076)

[229] Damour T, Iyer B R and Sathyaprakash B S 2000 Phys. Rev. D **62** 084036 (*Preprint* gr-qc/0001023)

[230] Arun K, Blanchet L, Iyer B R and Qusailah M S 2008 *Phys.Rev.* **D77** 064034 (*Preprint* 0711.0250)

[231] Arun K, Blanchet L, Iyer B R and Qusailah M S 2008 *Phys.Rev.* **D77** 064035 (*Preprint* 0711.0302)

[232] Arun K G, Blanchet L, Iyer B R and Sinha S 2009 *Phys.Rev.* **D80** 124018 (*Preprint* 0908.3854)

[233] Loutrel N and Yunes N 2017 *Classical and Quantum Gravity* **34** 044003 (*Preprint* 1607.05409)

[234] Boetzel Y, Susobhanan A, Gopakumar A, Klein A and Jetzer P 2017 Phys. Rev. D **96** 044011 (*Preprint* 1707.02088)

[235] Yunes N and Pretorius F 2009 *Phys.Rev.* **D80** 122003 (*Preprint* 0909.3328)

[236] Loutrel N, Yunes N and Pretorius F 2014 Phys. Rev. D **90** 104010 (*Preprint* 1404.0092)

[237] Huerta E A, Moore C J, Kumar P, George D, Chua A J K, Haas R, Wessel E, Johnson D, Glennon D, Rebei A, Holgado A M, Gair J R and Pfeiffer H P 2018 Phys. Rev. D **97** 024031 (*Preprint* 1711.06276)

[238] Bender C M and Orszag S A 1999 *Advanced mathematical methods for scientists and engineers 1, Asymptotic methods and perturbation theory* (New York: Springer)

[239] Porter E K 2007 Phys. Rev. D **76** 104002 (*Preprint* 0706.0114)

[240] Pankow C *et al.* 2018 *Phys. Rev.* **D98** 084016 (*Preprint* 1808.03619)

[241] Tinto M, Estabrook F B and Armstrong J W 2002 *Phys. Rev. D* **65**(8) 082003 URL https://link.aps.org/doi/10.1103/PhysRevD.65.082003

[242] Adams M R and Cornish N J 2010 *Phys. Rev. D* **82**(2) 022002 URL https://link.aps.org/doi/10.1103/PhysRevD.82.022002

[243] Adams M R and Cornish N J 2014 *Phys. Rev. D* **89**(2) 022001 URL https://link.aps.org/doi/10.1103/PhysRevD.89.022001

[244] Thrane E and Talbot C 2018 (*Preprint* 1809.02293)

[245] Braak C J F T 2006 *Statistics and Computing* **16** 239–249 ISSN 1573-1375 URL https://doi.org/10.1007/s11222-006-8769-1

[246] Chatziioannou K, Klein A, Cornish N and Yunes N 2017 *Phys. Rev. Lett.* **118**(5) 051101 URL https://link.aps.org/doi/10.1103/PhysRevLett.118.051101

[247] Littenberg T B and Cornish N J 2009 *Phys. Rev. D* **80**(6) 063007 URL https://link.aps.org/doi/10.1103/PhysRevD.80.063007

[248] GREEN P J 1995 *Biometrika* **82** 711–732 URL http://dx.doi.org/10.1093/biomet/82.4.711

[249] Romano J D and Cornish N J 2017 *Living Reviews in Relativity* **20** 2 ISSN 1433-8351 URL https://doi.org/10.1007/s41114-017-0004-1