UNIVERSITY OF BIRMINGHAM

LISA Data Analysis:

The search for long-lived gravitational wave signals

by

Diganta Bandopadhyay

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

Institute for Gravitational Wave Astronomy

School of Physics and Astronomy

College of Engineering and Physical Sciences

University of Birmingham

May 2025

Abstract

The detection of gravitational waves from the merger of binary black holes in 2015 opened a new window into the universe, and the Laser Interferometer Space Antenna will once-again revolutionize our understanding of the cosmos by unveiling the low-frequency gravitational wave universe. Among the many unique challenges that LISA presents, this thesis focuses on the several distinct challenges within the field of LISA data analysis. One of the most demanding aspects is the search for long-lived gravitational wave signals, expected from stellar-mass binary inspirals and extreme-mass ratio inspirals. These signals are expected to be much quieter than the noise in the LISA data and persist in the datastream for a significant fraction of the mission lifetime.

This thesis follows the development of a novel semi-coherent search strategy for stellar-mass binary inspiral signals in LISA data. This search method is initially developed in a simplified form, searching for single signals within noiseless data. Subsequently, the method is extended to search for multiple signals in the presence of noise, and is successfully applied to a LISA data challenge which contains several simulated stellar-mass binary signals. A complementary framework is developed to estimate the significance of each candidate signal, this is used to determine an approximate signal-to-noise ratio threshold for the search.

Acknowledgements

This thesis and the work presented within it would not have been possible without the support of several people. Firstly, I would like to thank my supervisor Chris Moore, for his guidance and patience throughout the many years (and for correcting my constant spelling and grammar mistakes), I have learnt a great deal from you. I would also like to thank all the members of the gravitational wave group at Birmingham, especially Alberto for his sage advice on several academic and non-academic matters over the last few years. I am also very grateful to both of my examiners: Dr. Michalis Agathos and Dr. Clément Bonnerot for their valuable feedback on the thesis.

All of my fellow PhD students in the department were a great source of both support and procrastination. In particular, I would like to thank Alex for the endless Spar snack runs, Alice and Simona for forcing me to occasionally go outside for 'two steps in the sun' (even when it was 10 degrees celsius) and Shaun for the amusing jokes. Special thanks to Christian for both invigorating academic discussions and for being a great source of memes, you made the writing up process far more entertaining (and arguably longer by distracting me with more interesting work). There are also several friends outside the department who have supported me over the years, including: Mike, for basically handholding me through first year maths (*honestly I may not have survived that year without you*), Guy, for being a great friend and listening to me complain about my work for hours on end [1] (*I'm pretty sure at this point you could have a second PhD in gravitational waves*) and Giovanni, for being my partner in crime for pretty much every physics project (*One day*

[1] Also the TV was not wasting power.

upsampling will work). A special thanks must also go to all my friends who I met online, from the FFXIV community: Morn, Ursuva, Sheal and Lilith, thank you for the several entertaining late night raiding sessions, they helped keep me sane (*One day we will finish an ultimate*).

I would also like to thank my family, for their love and support over the years. Thank you to Ani, for being a great brother and also patiently tutoring me through GCSE maths, I would not have been able to pursue physics without your help. I am incredibly grateful to my parents for their constant support. Some of my earliest memories are of my mum tirelessly working away at her thesis, through natural disasters, family tragedies and everything else under the sun. While at the time I claimed that I would never do a PhD (as frankly it looked very painful), I am sure that without her example of perseverance I would not be the person I am today. I would also like to thank my dad for his encouragement and support, even if I am jealous that you graduated first!

Finally, a significant thanks has to go to my partner, Chloe, for putting up with me when I was up at 2am setting off code, for listening to me incoherently ramble about statistics and code for hours on end and for always being there to offer love and encouragement when I needed it most.

Contents

1 Introduction to gravitational waves **1**

1.1 Theoretical overview 2

1.2 Leading order waveform model 4

1.2.1 Transforming to gravitational wave frequency and SI units 10

1.3 Geodesic deviation 11

2 LISA sources **14**

2.1 Massive black hole binaries 14

2.2 Ultra-compact binaries 19

2.3 Stellar-mass binary inspirals 21

2.4 Extreme mass ratio inspiral 23

3 LISA data analysis **26**

3.1 Search and parameter estimation 26

3.1.1 Matched filtering 27

3.1.2 Template banks 31

3.1.3 Parameter estimation 36

3.1.4 Overlapping sources $\rightarrow$ Global fit 41

3.1.5 Noise properties and characterizing significance 45

3.1.6 Stationary phase approximation 47

3.2 LISA response function 50

3.3 Time-Delay Interferometry 59

3.4 Time-frequency analysis 61

4 A proof-of-concept semi-coherent search 65

4.1 Introduction 65

4.2 Semi-Coherent methods 70

4.3 Case study: GW170817 77

4.4 Stellar-mass binary black holes 82

4.5 Particle Swarm Optimization 89

4.6 Future work 94

5 GPU-accelerated search for stellar-mass binary inspiral signals in LISA 98

5.1 Introduction 99

5.2 Search Statistics 101

5.2.1 Coherent search statistic 101

5.2.2 Semi-coherent search statistic 102

5.2.3 Sampling distributions 103

5.3 Search Structure 106

5.3.1 Particle swarm optimization 107

5.3.2 Multi-Peak tracking 107

5.3.3 Transition to MCMC inference 109

5.4 Search implementation details 109

5.5 Properties of the search statistic 114

5.5.1 Sampling distributions of the semi-coherent search statistics 115

5.5.2 PP-type plots for the semi-coherent likelihood 116

5.6 Search results 119

5.7 Computational cost 123

5.8 Conclusion 125

6 Searching for stellar-origin binary black holes in LISA Data Challenge 1b: Yorsh126

6.1 Introduction 127

6.2 Data 128
6.3 Methods 130
6.4 Results 133
6.5 Computational Cost 137
6.6 Discussion 138

7 Conclusions **139**

A Chapter 4 appendix **142**

A.1 Tempering vs semi-coherent 142
A.2 PSO configuration 145
A.3 Parameter transforms 146
A.4 Quadrature integration methods 147

B Chapter 5 appendix **151**

B.1 Linearly chirping signal 152
B.2 PSO & Inference parameters 153
B.3 GW190521-like search 154
B.4 Detection probability 155

Bibliography **157**

1. Introduction to gravitational waves

The detection of gravitational waves (GWs) in 2015 by the LIGO detectors ushered in a new era of astronomy and provided yet another experimental confirmation for Einstein's theory of General Relativity [1, 2, 3]. While the existence of gravitational waves had been inferred in 1974 by observations of the decay in the orbital period of the Hulse-Taylor binary pulsar [4], the direct detection of GWs was a key milestone in the field. Since then the field of GW astronomy has grown rapidly, with more than 100 confident detections of compact binary sources, spanning binary black holes, binary neutron stars and black hole-neutron star binaries [5].

While the focus of this thesis will be on the detection and characterization of specific classes of GW signals, this chapter is dedicated to providing a brief theoretical overview of GWs. We will start by discussing how a small perturbation on a flat spacetime metric propagates in a way analogous to electromagnetic waves; travelling at the speed of light, with two independent and transverse modes of polarisation. We will then derive the leading-order expression for a GW signal from a compact binary under some simple assumptions. While the results of this calculation are not sufficiently accurate to be used in detailed analysis, it provides a useful starting point for understanding the basic physics of the problem, and will also be useful for computing order-of-magnitude estimates throughout the thesis.

1.1 Theoretical overview

Within the theory of General Relativity, spacetime is represented as a 4-dimensional manifold, with its geometry described by an associated metric $g_{\mu\nu}$. The dynamics of this theory are governed by the Einstein-Hilbert action which is defined as:

$$S_{\mathrm{EH}} = \frac{1}{16\pi} \int \mathrm{d}^4x \sqrt{-g} R, \tag{1.1}$$

where R is the Ricci scalar, computed from $g_{\mu\nu}$, which is a measure of the curvature of the manifold, and g is the determinant of the metric. The total action is the sum of the Einstein-Hilbert and terms representing all of the other matter fields. Varying the total action with respect to $g_{\mu\nu}$ leads to the Einstein field equations:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi T_{\mu\nu}, \tag{1.2}$$

where $R_{\mu\nu}$ is the Ricci tensor, which can be calculated by contracting over the Riemann tensor $R_{\mu\nu\alpha\beta}$. The term $T_{\mu\nu}$ is called the stress-energy tensor and arises from the variation of all the matter terms in the action with respect to the metric $g_{\mu\nu}$. Note that for the rest of this subsection, we use a metric with signature (eigenvalues) of $(-1, 1, 1, 1)$ and use natural units throughout, i.e. $G = c = 1$.

In flat spacetime, in the absence of any matter, the metric $g_{\mu\nu} = \eta_{\mu\nu}$, where $\eta_{\mu\nu}$ is the Minkowski metric, this describes a region of spacetime without any gravitational fields. We can add a small perturbation to this metric $h_{\mu\nu}$, such that $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$. Assuming $h_{\mu\nu} \ll 1$ (this requires all gravitational fields to be weak), $R, R_{\mu\nu}, R_{\mu\nu\alpha\beta}$ can be expanded to linear order in $h_{\mu\nu}$ which can be substituted into Eqn. 1.2 to obtain[1]:

$$-\Box \bar{h}_{\mu\nu} + \partial_\alpha \partial_\mu \bar{h}^\alpha_\nu + \partial_\nu \partial^\alpha \bar{h}_{\mu\alpha} - \eta_{\mu\nu} \partial_\alpha \partial_\beta \bar{h}^{\alpha\beta} = 16\pi T_{\mu\nu}. \tag{1.3}$$

[1]Switching from $h_{\mu\nu}$ to its trace-reversed form $\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu} h$, where $h = h^\mu_\mu$. Additionally, we will use η as the metric to raise and lower indices, such that $h = \eta^{\mu\nu} h_{\mu\nu}$.

Here there is some gauge freedom, which allows for transformation of the coordinates for the problem, without changing the physics. For the system at hand, it is convenient to impose the Lorentz gauge, which requires that $\partial_\alpha h^{\alpha\beta} = 0$, enforcing this condition leads to the following simplified form of the equation above:

$$\Box \bar{h}_{\mu\nu} = -16\pi T_{\mu\nu}, \tag{1.4}$$

where $\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h$ is the trace-reversed form of the metric perturbation and $\Box = \partial^\mu\partial_\mu$ is the d'Alembertian operator in flat space which can be written out explicitly in Cartesian coordinates as $\Box = (-\frac{\partial^2}{\partial t^2} + \boldsymbol{\nabla}^2)$. Hereafter, we use $h_{\mu\nu}$ to denote $\bar{h}_{\mu\nu}$ for brevity. In what follows, we will work in a gauge where $h^\mu_\mu = 0$ anyway, therefore the trace reversal condition is automatically satisfied.

Equation 1.4 can be studied in two situations: the production of GWs, where $T_{\mu\nu} \neq 0$, and the propagation of GWs through a vacuum, where $T_{\mu\nu} = 0$. In this thesis we are interested in the latter case, in which Eqn. 1.4 simplifies to the familiar wave equation:

$$\Box h_{\mu\nu} = \left(-\frac{\partial^2}{\partial t^2} + \boldsymbol{\nabla}^2 \right) \bar{h}_{\mu\nu} = 0. \tag{1.5}$$

In analogy to Maxwell's equations for electromagnetic waves, this admits plane wave solutions of the form:

$$h_{\mu\nu} = A_{\mu\nu} e^{i(k_\alpha x^\alpha)}, \tag{1.6}$$

where $A_{\mu\nu}$ are constant amplitudes. The condition from Eqn. 1.5 imposes $\partial_\gamma\partial^\gamma h_{\mu\nu} = (ik_\alpha)(ik^\alpha) = 0$. The null wavevector implies that GWs travel at the speed of light. Additionally, to satisfy the Lorenz gauge condition, $\partial_\nu h^{\mu\nu} = 0$ (it is always possible to impose this gauge condition), we must have $A_{\mu\nu}k^\nu = 0$. Physically, this implies that the wavevector is orthogonal to the perturbation, as $h_{\mu\nu}k^\nu \sim A_{\mu\nu}k^\nu = 0$, i.e. GWs are transverse.

One more gauge choice is required to fix the remaining degrees of freedom and reduce $h_{\mu\nu}$ to its simplest form. This can be done by imposing the transverse-traceless gauge

condition, which imposes the following two conditions in addition to the Lorentz gauge conditions mentioned previously:

$$h_{0\mu} = 0, \;\; h^i_i = 0, \tag{1.7}$$

where we have used the roman indices $\{i, j, k\}$ to denote the three spatial coordinates. The first term, $h_{0\mu}$ requires the time components of the perturbation to be zero, while the second term h^i_i requires the spatial part of the tensor to be traceless. Additionally, recall that since both $g_{\mu\nu}$ and $\eta_{\mu\nu}$ are symmetric, $h_{\mu\nu}$ must also be symmetric. With the additional step of choosing the wavevector k parallel to the z axis, we can simplify the form of the metric perturbation to only contain two independent components:

$$A_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & A_+ & A_\times & 0 \\ 0 & A_\times & -A_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{1.8}$$

where A_+ and $A_\times$ are the two independent polarisation amplitudes of the GW in the plane perpendicular to $\hat{z}$. Notice the similarity between GWs and electromagnetic waves, both are transverse waves, with two polarization states and travelling at a velocity c.

1.2 Leading order waveform model

In this section we will derive the leading-order expression for a GW under some simplifying assumptions. We will make use of the following two equations which together are often referred to as the quadrupole formula [2, 6]:

$$\left.\frac{dE}{dt}\right|_{\text{GW}} = \frac{1}{5} < \dddot{Q}_{ij} \dddot{Q}^{ij} > \quad \text{and} \quad h_{ij} = \frac{2}{D_L} \ddot{Q}_{ij}, \tag{1.9}$$

where <> denotes an orbit-averaged quantity and D_L is the luminosity distance to the source. Notice in the second equation, the amplitude of the perturbations drop as D_L^{-1}; in the case of GWs, the detectors measure a strain (See Sec. 1.3), which is related to the luminosity of the source as $|\frac{\Delta L}{L}| \sim h \sim D_L^{-1}$, while in the case of EM observations the flux measured scales as D_L^{-2}.

The reduced quadrupole moment Q_{ij} is defined (in the non-relativistic limit) as [6]:

$$Q_{ij} = M_{ij} - \frac{1}{3}\delta_{ij}M_k^k, \tag{1.10}$$

where M_{ij} is the second mass moment tensor (and M_k^k is the trace of this tensor) defined as the following (for a set of N point masses with positions $\boldsymbol{x}_k$ and masses m_k):

$$M_{ij} = \sum_{k=1}^{N} x_{ki}x_{kj}m_k. \tag{1.11}$$

A closely related quantity to this, the moment of inertia tensor from classical mechanics, is defined as:

$$I_{ij} = \sum_{k=1}^{N}(|\boldsymbol{x}_k|^2\delta_{ij} - x_{ki}x_{kj})m_k, \tag{1.12}$$

where $|\boldsymbol{x}_k|^2$ is the distance of the k point mass from the origin. The second mass moment and the moment of inertia are related by the following expression:

$$I_{ij} = M_k^k\delta_{ij} - M_{ij}. \tag{1.13}$$

In this subsection, we will consider a binary system of two point masses, m_1 and m_2, in a circular orbit such that both the orbital angular momentum vector $\vec{L}$ is parallel to $\hat{z}$. The gravitational two-body problem can be treated in the centre of mass frame as a one-body problem with a stationary central mass $M = m_1 + m_2$ and an orbiting reduced mass $\mu = (m_1m_2)/(m_1 + m_2)$, μ orbits M at a distance R which represents the distance between m_1 and m_2 in the original problem. In this one-body formulation, the equations

describing the orbit of μ are given by:

$$x(t) = R\cos(2\pi ft + \phi_0) \quad \text{and} \quad y(t) = R\sin(2\pi ft + \phi_0), \tag{1.14}$$

where ϕ_0 is the initial orbital phase of the orbit and f is the orbital frequency. To slightly simplify the expression for M_{ij}, we will assume $\phi_0 = \pi/2$. In this centre of mass frame one body formulation, assuming we are considering μ to be a point particle mass, M_{ij} is:

$$M_{ij} = \mu x_i x_j, \tag{1.15}$$

where x_i corresponds to the ith component of the position of μ. Therefore M can be written as:

$$M_{ij} = \frac{\mu R^2}{2}\begin{pmatrix} 1-\cos(4\pi ft) & -\sin(4\pi ft) & 0 \\ -\sin(4\pi ft) & 1+\cos(4\pi ft) & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{1.16}$$

This is then used to calculate Q_{ij} and its derivatives:

$$Q_{ij} = \mu R^2 \begin{pmatrix} \frac{1}{6} - \frac{\cos(4\pi ft)}{2} & -\frac{1}{3} - \frac{\sin(4\pi ft)}{2} & 0 \\ -\frac{1}{3} - \frac{\sin(4\pi ft)}{2} & \frac{1}{6} + \frac{\cos(4\pi ft)}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{1.17}$$

$$\dddot{Q}_{ij} = 32\pi^3 f^3 \mu R^2 \begin{pmatrix} -\sin(4\pi ft) & \cos(4\pi ft) & 0 \\ \cos(4\pi ft) & \sin(4\pi ft) & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{1.18}$$

Note that this assumes the orbit is closed with a *constant* orbital frequency f. This is not physically reasonable as the binary radiates energy in the form of GW which increases f. This will be addressed later in this section.

Assuming the orbital frequency $f(t)$ does not change significantly over one orbit, the orbit averaged trigonometric quantities simplify as $<\sin^2(4\pi ft)>=<\cos^2(4\pi ft)>=1/2$. This allows for computation of GW luminosity in Eqn. 1.9:

$$\left.\frac{dE}{dt}\right|_{\rm GW} = \frac{1}{5} < \dddot{Q}_{11}^2 + \dddot{Q}_{22}^2 + \dddot{Q}_{12}^2 + \dddot{Q}_{21}^2 >= \frac{2048}{5}\mu^2 R^4 \pi^6 f^6. \tag{1.19}$$

Before proceeding further it is worth considering that both R and f are functions of time. As energy is radiated away from the binary in the form of GW, the orbit shrinks, causing R to decrease and f to increase. In this derivation, we will assume that the orbit evolves adiabatically, i.e. that energy is radiated sufficiently slowly such that $R(t)$ and $f(t)$ vary slowly compared to the orbital timescale i.e. $\dot{f}/f^2 \ll 1$. We will use the orbit averaged luminosity in Eqn. 1.19 to compute the evolution of $R(t)$ and $f(t)$[2].

In Newtonian mechanics, the total energy in a Keplerian orbit is given by $E = -\frac{M\mu}{2R}$. The rate of change of energy in the orbit can be related to the change of the orbital frequency using

$$\left.\frac{dE}{dt}\right|_{\rm orbit} = -\frac{1}{3(2\pi)^{2/3}}\frac{M^{4/3}}{R^2} f^{-5/3}\frac{df}{dt}. \tag{1.20}$$

Assuming the orbit evolves adiabatically only via the emission of GWs, this can be equated with Eqn. 1.19. With an additional step of using Keplers 3rd law ($(2\pi ft)^2 R^3 = M$) to write R in terms of f, we can obtain the following form for the evolution of the orbital frequency

$$\frac{df}{dt} = \frac{96}{5}(2\pi)^{\frac{8}{3}}\pi f^{11/3}\mu M^{2/3}. \tag{1.21}$$

At this point is useful to introduce another variable, chirp mass $\mathcal{M}_c = (m_1 m_2)^{\frac{3}{5}}/(m_1 + m_2)^{\frac{1}{5}} = (M^2\mu^3)^{1/5}$. The frequency evolution $f(t)$ can be obtained by integrating this expression:

$$\int_\infty^f \frac{df'}{f'^{\frac{11}{3}}} = \int_{t_c}^t (2\pi)^{\frac{8}{3}}\frac{96}{5}\mathcal{M}_c^{\frac{5}{3}} dt'. \tag{1.22}$$

The lower limit of this integral at $t = t_c$ and $f \to \infty$ corresponds to the time of the merger, while the upper limits corresponds to some time t, which is $t_0 - t$ before the

[2]Although, due to the Newtonian approximation we are under, we can relate the two using Keplers 3rd law to simplify the derivation.

merger. Near the point of binary merger, the Keplerian orbits which are used throughout the derivation are no longer valid, so this model should not be trusted close to the merger. However for the large majority of this thesis, we are concerned with the inspiral of stellar-mass binaries, so this model is sufficient as an order-of-magnitude estimator.

$$f(t) = \frac{5^{\frac{3}{8}}}{16\pi}\mathcal{M}_c^{-\frac{5}{8}}(t_c - t)^{-\frac{3}{8}} \tag{1.23}$$

It will be useful to simply rearrange this to express the time to merger, in terms of the orbital frequency of the binary at that time:

$$t_c - t = \frac{5}{(16\pi)^{\frac{8}{3}}}\mathcal{M}_c^{-\frac{5}{3}} f^{\frac{8}{3}}. \tag{1.24}$$

This expression will be useful for computing estimates for the time-to-merger of various binaries in Sec. 2.

The waveform h_{ij} can be computed by substituting the expression for $\ddot{Q}_{ij}$ into Eqn. 1.9.

$$h_+(t) = \frac{16\mu R^2(t)\pi^2 f^2(t)}{D_L}\cos(\phi(t)) \quad \text{and} \quad h_\times(t) = \frac{16\mu R^2(t)\pi^2 f^2(t)}{D_L}\sin(\phi(t)). \tag{1.25}$$

Where the time-domain phase $\phi(t) = \int 4\pi f \mathrm{d}t$. Substituting Eqn. 1.23 into this, the phase is:

$$\phi(t) - \phi_c = \int_t^{t_c} 4\pi f(t')\mathrm{d}t' = -\frac{4}{5^{\frac{5}{8}}}\mathcal{M}_c^{-\frac{5}{8}}(t_c - t)^{\frac{5}{8}}, \tag{1.26}$$

where ϕ_c is the phase at the time of coalescence. Note we had to explicitly compute the accumulated phase using an integral over time instead of simply using $\phi = 4\pi f t$ as the orbital frequency f is time dependent, as discussed earlier.

A frequency-domain variant of this model is presented in Sec. 3.1.6, where the stationary phase approximation (SPA) is used to obtain the Fourier domain waveform, the SPA is a good approximation in the early inspiral when the orbital evolution of the binary is slow.

There is also a simple way to compute a rough estimate of the orbital frequency at which the merger of two compact objects will occur, using Kepler's 3rd law to compute the frequency at which the separation of the two binaries is equal to the sum of their respective radii. In the scenario where both compact objects are black holes, this can be taken to be the sum of their Schwarzschild radii. While this is a *(very)* crude estimate (especially since in the merger phase the two bodies are in a strong gravitational field, and thus the dynamics are not well described by Keplerian orbital mechanics) it is sufficient for an order-of-magnitude estimate for the merger frequency. In a similar way, the inspiral is usually considered to end when the orbital radius is equal to the inner-most stable circular orbit (ISCO) of the binary, in the Schwarzschild test-particle limit (small compact object orbiting a massive object) this corresponds to 3 times the Schwarzschild radius of the massive object, this can also be generalized to more comparable mass ratios [6, 7].

The derivation above assumes that we are considering the leading order quadrupolar radiation, which corresponds to the $(l, |m|) = (2, 2)$ in the spin weighted spherical harmonic basis. These multipole modes arise from solving Eqn. 1.4 in the case where $T_{\mu\nu} \neq 0$: this corresponds to the generation of GWs, in this setup the waves are 'extracted' far from the source, the solution of this wave equation involves the expansion of Greens functions in the (spin-weighted) spherical harmonic basis $Y_{lm}(\theta, \phi)$ [6]. While only considering $(l, |m|) = (2, 2)$ is sufficient for the majority of this thesis, it is worth noting that in general there are higher order multipoles, e.g. $l = 3, m = 1$; however these modes only carry a significant fraction of the total energy in the waveform for highly unequal mass ratio binaries or those with significant spin-orbit precession [8], neither of which are considered in this thesis. Generally, binaries with more asymmetry (which may be caused by eccentricity, spin-orbit precession or mass ratio) will have a more complicated structure of radiation on the sphere, which necessitates the inclusion of higher order multipoles in the waveform model. In cases where the higher modes are significant, they allow more precise parameter estimation of the binary parameters; notably, measuring higher modes can constrain the degeneracy between luminosity distance and inclination of the binary

orbit with respect to the line of sight of the observer [8].

1.2.1 Transforming to gravitational wave frequency and SI units

While in the previous section for convenience we have made use of natural units throughout, it will be useful to express some of these quantities that we have defined in terms of SI units, so will reinsert the appropriate factors of G and c. Additionally, we also make the transformation from f being the orbital frequency, to the GW frequency, as can be seen from Eqn. 1.17, for a monochromatic GW, the waveform $h(t) \sim \cos(4\pi f_{\rm orb} t)$, this implies that $f_{\rm GW} = 2f_{\rm orb}$. From here on out, we will instead denote f to be $f_{\rm GW}$.

In this form, Eqs. 1.23 and 1.24 can be written as:

$$f(t) = \frac{5^{\frac{3}{8}}}{8\pi} \frac{c^{\frac{15}{8}}}{G^{\frac{5}{8}}} \mathcal{M}_c^{-\frac{5}{8}} (t_c - t)^{-\frac{3}{8}} \quad \text{and} \quad t_c - t = \frac{5}{(8\pi)^{\frac{8}{3}}} \frac{c^5}{G^{\frac{5}{3}}} \mathcal{M}_c^{-\frac{5}{3}} f^{-\frac{8}{3}} \tag{1.27}$$

Similarly the phase evolution in Eqn. 1.26 can be written as:

$$\phi(t) - \phi_c = -\frac{2}{5^{\frac{5}{8}}} \frac{c^{\frac{15}{8}}}{G^{\frac{5}{8}}} \mathcal{M}_c^{-\frac{5}{8}} (t_c - t)^{\frac{5}{8}}. \tag{1.28}$$

Additionally, the energy radiated away as GWs is given by $dE/dt = G/(5c^9) < \dddot{Q}_{ij} \dddot{Q}^{ij} >$, and the waveform is $h_{ij} = 2G/(D_L c^6) \ddot{Q}_{ij}$. Using this, the $(+, \times)$ polarizations can be expressed as:

$$h_+(t) = \frac{4\mu G R^2(t) \pi^2 f^2(t)}{D_L c^4} \cos(\phi(t)) \quad \text{and} \quad h_\times(t) = \frac{4\mu G R^2(t) \pi^2 f^2(t)}{D_L c^4} \sin(\phi(t)). \tag{1.29}$$

The amplitude in the expressions above can slightly be simplified by once again using Keplers third law to relate $R(t)$ to $f(t)$, simplifying the right hand side of the expressions down to only one time dependent quantity $f(t)$ which has been derived earlier:

$$h_+(t) = \frac{4G^{\frac{5}{3}}\mathcal{M}_c^{\frac{5}{3}} f^{\frac{2}{3}}(t)\pi^{\frac{2}{3}}}{D_L c^4}\cos(\phi(t)) \quad \text{and} \quad h_\times(t) = \frac{4G^{\frac{5}{3}}\mathcal{M}_c^{\frac{5}{3}} f^{\frac{2}{3}}(t)\pi^{\frac{2}{3}}}{D_L c^4}\sin(\phi(t)). \tag{1.30}$$

1.3 Geodesic deviation

In this section we briefly discuss the detection principle used by GW detectors. We will consider the case of a GW passing through a pair of free-falling test masses, this derivation closely follows that presented in Ref. [9]. Initially both of these test masses are considered to be at rest, therefore $dx^\mu/dt = (1, 0, 0, 0)$ (in natural units). The geodesic equation for each test particle is then:

$$\frac{d^2x^\mu}{d\tau^2} = -\Gamma^\mu_{\nu\sigma}\frac{dx^\nu}{d\tau}\frac{dx^\sigma}{d\tau} = -\Gamma^\mu_{00}\frac{dx^0}{d\tau}\frac{dx^0}{d\tau} = -\Gamma^\mu_{00}, \tag{1.31}$$

$$\Gamma^\mu_{00} = \frac{1}{2}g^{\mu\kappa}(\partial_0 g_{\kappa 0} + \partial_0 g_{0\kappa} - \partial_\kappa g_{00}). \tag{1.32}$$

Since $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, and η is constant, we only need to examine the derivatives of $h_{\mu\nu}$. In the form of $h_{\mu\nu}$ given in Eqn. 1.8 (Lorenz and the transverse traceless gauge), the only non-zero components of $h_{\mu\nu}$ are the spatial components h_{ij}, the terms of the metric perturbation involved in the Christoffel symbols above are all 0. Therefore, $\frac{d^2x^\mu}{d\tau^2} = 0$, which physically means that these test particles (in these coordinates) remain at rest as a GW passes through them.

We now use the result that the proper distance between the two test masses is a gauge invariant quantity. We will refer to the proper distance between the two test masses as $L(t)$ (assigning it time dependence anticipating that it will vary as a function of time). The Euclidean distance between these two objects is defined as:

$$L^2(t) = \delta_{ij}d^i d^j = \Delta x^2 + \Delta y^2 + \Delta z^2. \tag{1.33}$$

This quantity can also be computed directly from the metric:

$$L^2(t) = g_{ij}\xi^i\xi^j = (\delta_{ij} + h_{ij}(t))\xi^i\xi^j, \tag{1.34}$$

where ξ is the 3-vector pointing from one test mass to the other. Since the free particles remain at rest, ξ^i is constant. Therefore we can write $d^i(t) = \xi^i + \epsilon^i(t)$, where $\epsilon^i(t)$ is the small perturbation to the distance between the two test masses we expect to see due to the passing GW. Using this, we can expand $\delta_{ij}d^id^j$ to retain terms only to first order in ϵ:

$$L^2(t) = \delta_{ij}\xi^i\xi^j + 2\delta_{ij}\xi^i\epsilon^j(t). \tag{1.35}$$

This can then be equated with the expression for $L^2(t)$ in terms of the metric and simplified to obtain an expression for the perturbation $\epsilon^i(t)$:

$$\begin{aligned} 2\delta_{ij}\xi^i\epsilon^j(t) &= h_{ij}\xi^i\xi^j, \\ 2\delta_{ij}\epsilon^j(t) &= h_{ij}\xi^j, \\ \epsilon^k(t) &= \frac{1}{2}h^k_j\xi^j. \end{aligned} \tag{1.36}$$

Since $d^i = \xi^i + \epsilon^i(t)$ and $L^2(t) = \delta_{ij}d^id^j = d_id^i$;

$$L^2(t) = \left[\xi_i + \frac{1}{2}h_{ij}(t)\xi^j\right]\left[\xi^i + \frac{1}{2}h^i_j(t)\xi^j\right] \approx \xi_i\xi^i + \frac{1}{2}h_{ij}(t)\xi^i\xi^j + \frac{1}{2}h^i_j(t)\xi_i\xi^j,$$

where we have only kept terms to first order in h. The first term is the Euclidean square distance between the two test masses L_0^2, and the second and third terms are the same and can be combined as follows:

$$L^2(t) = L_0^2 + h_{ij}(t)\xi^i\xi^j, \tag{1.37}$$
$$L(t) = L_0\left[1 + h_{ij}(t)\frac{\xi^i\xi^j}{L_0^2}\right]^{1/2} \approx L_0\left[1 + \frac{1}{2}h_{ij}(t)\frac{\xi^i\xi^j}{L_0^2}\right].$$

This is the main result that shows how the proper distance between two test masses changes as a GW passes through them (at leading order). This is the fundamental concept on which GW interferometers are based, measuring deviations in arm lengths. This principle will be used extensively in Sec. 3.2 to derive the instrument response function for LISA, Sec. 3.3 will cover how this detection principle is used when the arm-lengths of the interferometer are not equal.

2. LISA sources

LISA is expected to observe GWs emitted by a wide range of astrophysical sources. This section describes some of these sources, and the key astrophysical insights that we expect to glean from LISA observations. This thesis is primarily focused on the analysis of compact binary sources and these are the sources we discuss in the most detail here. We also discuss some other astrophysical LISA sources in detail, but we neglect possible stochastic sources of GWs, including cosmological backgrounds originating from the early universe [10].

2.1 Massive black hole binaries

Massive black hole (MBH) binaries are one of the most important sources of GWs that LISA will target. MBHs usually reside in the centre of galaxies; our own galaxy, the Milky Way hosts the MBH *Sagittarius A** at its centre, with a mass of about $4.3 \times 10^6\,\mathrm{M}_\odot$[16]. The mass of *Sagittarius A** was initially estimated using observations of stellar orbits close to the central MBH, these observations indicated the orbits can be explained by a single central massive object [17, 18]. One of the formation channels for supermassive black holes is thought to be hierarchical mergers of smaller MBHs, this implies that MBH binaries must regularly form and merge, in order to obtain the supermassive black holes observed at the centres of some galaxies with masses $\sim 10^9\,\mathrm{M}_\odot$ [19] (the MBH mergers are thought to be associated with the mergers of their host galaxies). LISA will be sensitive to GWs emitted during the late inspiral and merger of these MBH binaries with component masses in the range $[10^5, 10^7]\,\mathrm{M}_\odot$, the lower limit of this range is set by the rough requirement

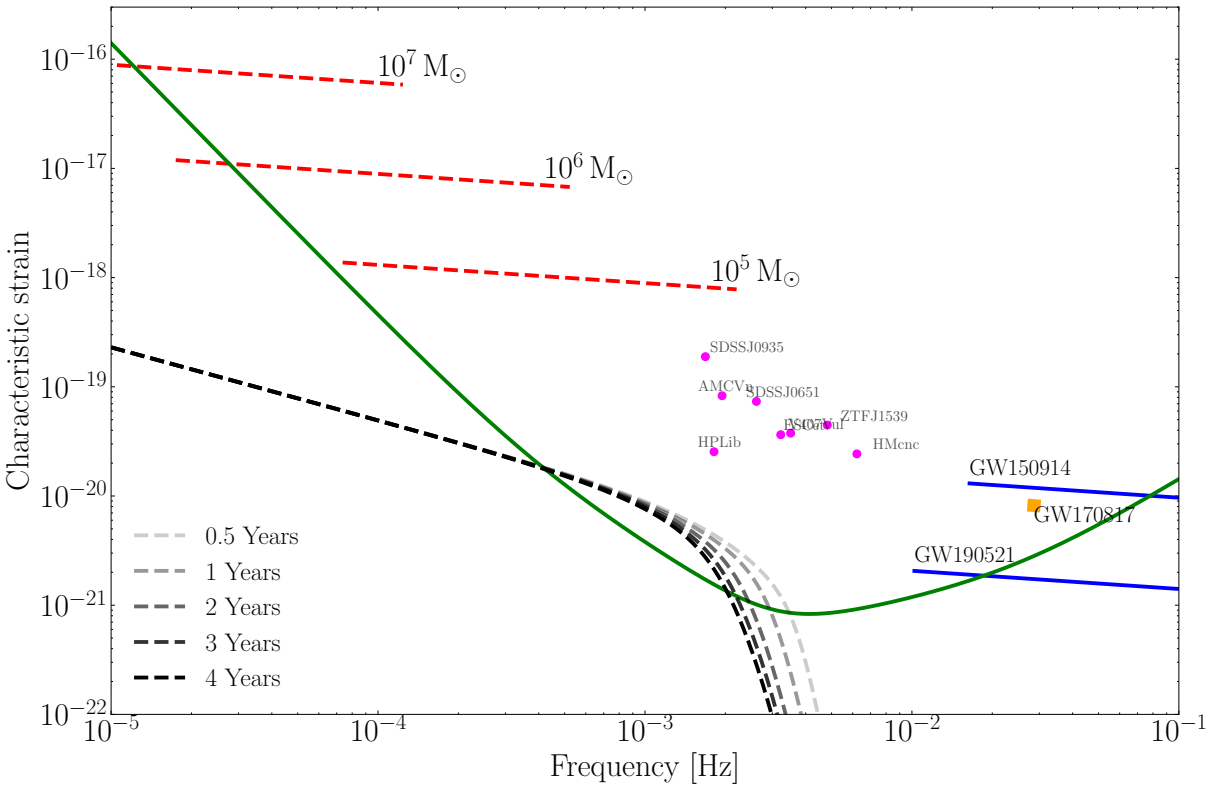

Figure 2.1: Reproduction of Fig. 1 from [11], showing the characteristic strain as a function of frequency for a variety of astrophysical sources, all of them using the leading order waveform derived in Chapter. 1. The characteristic strain corresponds to a frequency-dependent rescaling of h (and the power-spectral-density (PSD)), in the case of broad-band sources the area between the characteristic strain and the PSD corresponds to the signal-to-noise ratio [12]. Dashed red lines denote massive black hole inspiral and merger. The dashed black lines denote the confusion noise arising from the double white dwarf galactic foreground (neglecting the periodic modulation of the background, see discussion in Sec. 2.2), with the level of shading indicating the suppression of the confusion noise as a function of time [13]. The magenta scatter points are a handful of loud verification binaries (see discussion in Sec. 2.2) we expect to see [14, 15]. Solid blue lines are used to mark the inspiral of a few stellar-mass binary events, both placed 2 years before the merger at the start of the LISA mission. The solid orange box indicates inspiral track of GW17087 (due to its very low chirp mass, it does not evolve significantly in frequency over the 2 year mission lifetime assumed for this plot). EMRI signals have not been considered for this plot as the leading order waveform model is completely invalid for these sources.

that these binaries should be transient, sweeping $[10^{-4}, 10^{-3}]$ Hz in the LISA band over a timescale much less than the total mission time. Binaries with (chirp) masses below this range are usually referred to as intermediate mass binaries which may be present in the LISA datastream over a period of time comparable to the mission lifetime [20]. Meanwhile, the upper limit of the mass range is set by the requirement that the binary

must merge with a GW frequency that is detectable by LISA, i.e. $\gtrsim 10^{-4}\,\mathrm{Hz}$, binaries with masses above $10^7 - 10^8\,\mathrm{M}_\odot$ will merge outside the LISA frequency band, see Fig. 2.1.

While LISA will observe the merger of MBH binaries, it is worth noting here that the early inspiral of these binaries radiates GWs in the nHz frequency band, which is inaccessible to LISA, but can be probed by pulsar timing arrays (PTAs). While it is not (yet) possible to individually detect and characterize MBH binaries using this method, the inspiral of numerous MBH binaries across the universe are expected to form a stochastic GW background detectable by PTAs [21]. Evidence for this background has recently been detected by a range of PTAs [22, 23]. Constraints on the stochastic background from PTAs can be used to estimate the merger rate of MBH binaries observable by LISA [24].

MBH binaries are expected to be the loudest sources of GWs that will be observed by LISA, this is primarily due to the very large mass of the binary (leading to a very large amplitude for the GW, see Eqn. 1.29), additionally LISA is most sensitive at the frequencies which correspond to the mergers of these binaries, see Fig. 2.1. Since these sources are intrinsically very loud, mergers of MBH binaries close by at $z \lesssim 6$ are expected to be detectable with signal-to-noise ratio (SNR) $\mathcal{O}(1000)$ [25]; while at large distances of $z \sim 20$ these sources remain detectable with SNR $\mathcal{O}(10)$ [26, 27]. In comparison, the SNR for current ground-based observations of GWs emitted during stellar-mass binary mergers are $\mathcal{O}(10)$. Note that at $z \sim 20$ the luminosity distance is $D_L \approx 230\,\mathrm{Gpc}$ and the universe age is $\sim 200\,\mathrm{Myr}$[1]. This is particularly exciting as EM observations are opaque past the epoch of reionization which occurred in the redshift range $z \approx 6 - 15$ [29] (in rough agreement with this redshift range, the furthest quasar observed via EM observations is at $z \approx 7.6$ [30]). The ability to probe structure formation in the early universe at redshifts inaccessible to EM observations is a key advantage of LISA GW observations. Recent James Webb Space telescope observations have discovered massive galaxies in the early universe at redshift $z \sim 9 - 13$ [31, 32], LISA may be able to probe the mergers of these

[1] Assuming $H_0 = 69.6, \Omega_M = 0.286, \Omega_\Lambda = 0.714$ in a flat universe [28].

early galaxies.

The large SNR of these sources is also a key advantage for parameter estimation, resulting in extremely precise estimates of both the intrinsic and extrinsic parameters of the binary[2]. As an example, the evolution of the spin vector for both bodies in the MBH binary can be tightly constrained throughout the time the binary is radiating GW in the LISA band [33]. The large SNR also highlights the importance of accounting for additional physical effects such as gas effects and eccentricity [34], as for these extremely loud systems neglecting these effects can lead to biases in the inferred parameters of the system. These binary sources can also act as a "standard bright siren" for probing the expansion history of the universe. One of the parameters of a GW waveform which is well measured in the case of MBH binaries is the luminosity distance D_L [35] (in the scenario that higher-order multipoles are significant, D_L can be constrained at the $\sim 0.01\%$ level [33]), this is a function of redshift, and can be obtained from the distance-redshift relation which accounts for the expansion rate of the universe. Since the merger of a MBH binary can exhibit an EM counterpart (e.g. if it occurs in a gas-rich environment), if the spectroscopic redshift of the source can be measured this can be combined with the GW inferred luminosity distance to constrain the expansion history of the universe [36, 37]. While in principle this can be done for any source type with an EM counterpart from which a spectroscopic redshift can be measured, this works best with detections with a tight constraint on D_L, making loud MBH binaries ideal for this. While the bright-siren method requires a coincident GW and EM detection, there is also the possibility of testing cosmology with the dark-siren method, which only uses the GW signal [38]. This method uses a galaxy catalogue to find galaxies within the sky localization of the GW detection, and then marginalizes over all the galaxies in the catalogue to obtain a redshift measurement for the source, consequently this method is strongly dependent on a precise sky localization from the GW signal [39]. It is worth highlighting that the accuracy of

[2]Intrinsic parameters describe the source in a frame co-moving with the binary, while extrinsic parameters describe the location and orientation of the binary with respect to the observer. This distinction is less well motivated for LISA sources compared to LVK sources, see discussion in Sec. 4.2

the dark siren method strongly depends on having complete galaxy catalogs, which are difficult to obtain in practice for the early universe, specially at the extreme redshifts where we expect to observe MBH binaries [40].

In the late inspiral of a MBH binary, the binary may be within a circumbinary disk, which can in-turn cause mini disks to form around the individual MBHs, accreting matter onto them. The friction in the mini disks heats up the gas to a temperature where it is expected to radiate strongly in the X-ray band, this radiation may be observable by EM telescopes, this X-ray signal may also be modulated periodically by the binary orbit [41], these features can be used as pre-merger signatures of an MBH binary and may be used in combination with GW early-warning searches (such as that in Ref. [42]) to detect such a source before merger. While the existence of a detectable EM signal at merger is more uncertain, there are scenarios in which the merger (or post-merger phase) of a MBH binary can produce various signatures in the EM band, see e.g. [43, 44, 45].

The high SNR of these sources also makes them ideal for various tests of fundamental physics, one of the most notable through the process of black hole spectroscopy. The black hole remnant from a binary will emit a GW signal in the post-merger phase which is known as the *ringdown*, this signal is a superposition of damped sinusoids known as quasi-normal modes (QNMs), with the frequencies and damping times of these modes being completely determined by the mass and spin of the remnant black hole (due the no-hair theorem for Kerr black holes)[46]. Detecting multiple of these modes within the ringdown signal can constrain deviations from the no-hair theorem. While this has been done for ground-based observations of BBH mergers [14], the SNR of these detections is much lower than that expected for MBH binary signals in LISA, this increased SNR may allow for the detection of more QNMs leading to a more precise test of the no-hair theorem [47].

2.2 Ultra-compact binaries

Ultra-compact binaries are the most numerous astrophysical sources of GWs that LISA is expected to observe, these binaries are usually comprised of two degenerate (i.e. white dwarf or neutron star) objects, with masses $\mathcal{O}(1)\,\mathrm{M}_\odot$. While the complete astrophysical population of these binaries is expected to occupy a wide range of orbital periods, LISA will observe the subset of these binaries with short orbital periods, below 1 hour. These binaries will emit GWs in the frequency range $\sim [10^{-4}, 10^{-3}]\,\mathrm{Hz}$, this is the frequency range where LISA is most sensitive. The GWs emitted by these sources are expected to be monochromatic to a good approximation, primarily due to the very small mass of the binary, this leads to a small evolution in the orbital period of these binaries over LISA's expected mission lifetime (assuming this is a detached binary that is only evolving via GW radiation). The small fraction binaries within this source category which are eccentric and reside at high frequencies ($\sim 10^{-3}\,\mathrm{Hz}$) are expected to exhibit a small $\dot{f} \sim 1\,\mu\mathrm{Hz/yr}$ (see Eqn. 10 in Ref. [48]), this can be treated as a rough upper bound for the frequency evolution of UCBs. The most common source types within this category will be the double-white-dwarf (DWD) binaries, followed by neutron star-white dwarf (NSWD) binaries, black hole-white dwarf binaries (BHWD), binary neutron stars (BNS) and neutron star-black hole binaries (NSBH) (See Table 4 in [49] for relative abundances).

Sources from this category are extremely numerous in LISA observations, there are expected to be $\sim \mathcal{O}(10^8)$ DWD binaries radiating GWs in the frequency range $f \gtrsim 10^{-4}\,\mathrm{Hz}$ [49, 50, 51]. However, most of these sources will be quiet and will not be detectable individually within the LISA data, instead they will combine to form a stochastic signal which is commonly referred to as the confusion noise foreground [52, 53], mostly originating from sources within the Milky Way and surrounding areas. Although it will generally not be possible to identify and characterize individual quiet binaries within the confusion noise signal, useful astrophysical information can be extracted from the stochastic background alone, e.g. the amplitude and spectral shape of the confusion noise signal can be used to

constrain the shape and stellar mass content of the Milky Way [54, 55]. It is worth highlighting that the amplitude of the stochastic confusion noise is time dependent; varying periodically due to the motion of the LISA detector around the sun [56].

A small subset of these binaries ($\sim 10^4$) are expected to be individually detectable [57], while this corresponds to a small fraction of the total UCB population, this source category is expected to contain more resolvable sources than all others. The binaries that will be detectable are likely to be the most massive and at higher orbital frequencies (closer to merger), see Fig. 2 in [57]. As these signals are present in the datastream throughout the entire LISA mission, over time the sky localization for these sources will improve, making these sources ideal candidates for electromagnetic follow-up. The surface temperature of a WD star is $\sim 10^4 - 10^5$K, WDs on the cooler end of this temperature range radiate mainly in the optical wavelengths, while extremely hot WD are found using Ultraviolet/X-ray observations. Electromagnetic observations of UCBs in the local group of galaxies have discovered some of these sources. The orbital parameters of these sources measured by EM observations have been used to predict the GW signal emitted by these binaries [58]. Since we have knowledge of the parameters describing these sources prior to LISA flying and we know they will be present in the data, these sources are referred to as *verification binaries*, the latest EM observations have characterized ~ 50 such verification binaries which should be detectable in LISA data [59, 60].

While most UCBs are expected to be circular, a handful are expected to have measurably eccentric orbits, specifically some of the NSWD binaries. Within this source-type, there is a formation channel in which the WD forms first, before the NS, and the NS supernova kick imparts eccentricity into the binary orbit [61]. Binaries with sufficiently high eccentricities at high GW frequencies ($\gtrsim 10^{-3}$ Hz) can be detected by LISA, with a confident measurement of eccentricity it is possible to break the usual degeneracy between the chirp mass and mass ratio of the binary, inferring the component masses directly, this can be useful for distinguishing between different types of compact objects [48].

Since these binaries are on relatively tight small period orbits, it is usually assumed

that the orbital evolution is completely driven by the emission of GW, leading to a positive $\dot{f}$ for the GW. While this is true for a binary which is detached, if there is any mass transfer between the two objects $\dot{f}$ can instead be negative, and this can be detected in the GW signal [62, 63].

2.3 Stellar-mass binary inspirals

The stellar-mass binary black hole sources which emit GWs in the ground-based detector frequency range of $[1, 100]\,\mathrm{Hz}$ during their merger will be present in LISA data much earlier in their orbital evolution, in their inspiral phase. These sources generally contain a pair of black holes with masses in the range $[5, 100]\,\mathrm{M}_\odot$. The time taken for these signals to evolve through the LISA frequency band can span $\sim \mathcal{O}(1 - 100)$ years, depending on several factors such as the mass of the binary and the orbital frequency of the binary at the time LISA begins observations. At frequencies $f \sim 10^{-3}\,\mathrm{Hz}$, these sources have a small $\dot{f}$, and are close to monochromatic. While more massive sources which are closer to merger, e.g. at $f \gtrsim 10^{-2}\,\mathrm{Hz}$, emit GWs which evolve through the LISA frequency band over $\sim \mathcal{O}(1)$ years, and eventually exit the upper end of the frequency band, merging and emitting GWs in the ground-based detector band soon after this. These are also the sources which will likely accumulate enough SNR over the LISA mission lifetime to be individually detectable.

These are expected to be the only type of gravitational wave sources which will cross both the space and ground-based detector frequency bands. In principle for any stellar-mass source which is present in the LISA datastream, if we are prepared to wait long enough, the source will eventually emit GWs in the ground-based detector band. However, the label of a *multiband* source is reserved for those that accumulate enough SNR in both (space and ground-based) detectors to be individually detected in both detectors. Combining the independent measurements of the source parameters from the two detector types will allow for extremely precise constraints on the orbital parameters of the source

[64, 65].

The possibility of detecting orbital eccentricity is another prospect of LISA observations of these sources. While there have been significant efforts to detect GW sources with orbital eccentricity in the ground-based detectors [66, 67], as of yet a confident measurement of non-zero eccentricity has only been claimed for a NS-BH binary source [68]. While there are certain scenarios which can excite large eccentricities in binaries close to merger [69, 70], the large majority of binaries are expected to be quasicircular at this stage as it is likely that GW radiation has circularized the orbit [3].However, there may still be detectable residual eccentricity for some of these sources in the early inspiral when they are observed by LISA. Small eccentricities of $e \sim 10^{-3}$ can change the orbital phase evolution of a stellar-mass binary inspiral in the LISA band significantly, resulting in a measurable effect on the waveform [72]. Eccentricity measurements of stellar-mass binary inspirals can be used to determine the formation channel for the binary [73]. The main formation channels for stellar-mass binary inspirals are known as the isolated or dynamical channels, in the isolated channel case the main sequence progenitors are assumed to have evolved as a binary in isolation before collapsing into a BBH system at the end of their stellar lifetimes, these binaries are not expected to have a significant eccentricity[4]. Whereas a binary formed within a globular cluster via the dynamical interactions may exhibit significant eccentricity in the LISA frequency band [72].

The stellar-mass binary black hole mergers that have been observed thus far do not have any associated electromagnetic counterparts, EM emission is usually expected only when the binary contains matter such as a neutron star (as was the case for GW170817, which had an associated gamma-ray burst [74]). However, there is a plausible scenario in which an EM counterpart could be produced from a stellar-mass binary black hole merger, from the shock heating of a circumbinary disk due to the recoil of the remnant black hole

[3]This expectation comes from Ref. [71] where an eccentric binary is shown to lose angular momentum through gravitational wave emission at a rate faster than it loses energy, leading to a circularization of the orbit (i.e $\mathrm{d}e/\mathrm{d}t$ is always negative). A full derivation of this is also provided in Ref. [6].

[4]While the binary may have gained some orbital eccentricity due to the supernovae kicks, this has probably been radiated away by the time the binary is in either the LISA or LIGO bands.

produced in a merger [75, 76]. One of the reasons no counterparts have yet been observed may be the lack of an early warning system, the mergers observed by LIGO/VIRGO usually last for milliseconds in the detector band, so it is difficult to detect the GW signal and rapidly coordinate EM observations. LISA observations of these sources will allow for a much longer lead time on the warning that these sources are about to merge, in some cases on the timescale of months. Multimessenger observations such as these also have some ability to discriminate between formation channels, as a circumbinary disk is not expected to survive in the dynamical formation channel scenario, e.g. within a cluster where interactions are frequent.

The accurate tracking of the waveform phase for these sources will also allow for tests of fundamental physics, such as tests for dipolar gravitational radiation which would accelerate the inspiral of the binary [77]. Multiband observations of these sources can also constrain deviations from general relativity which cause the propagation speed of GW to differ from c [78].

2.4 Extreme mass ratio inspiral

Extreme mass ratio inspiral (EMRI) systems comprise a supermassive black hole, usually at the centre of a galaxy with mass $\sim 10^5 - 10^6\,\mathrm{M}_\odot$ in a binary with a compact object of mass $\sim 1 - 100\,\mathrm{M}_\odot$, such as stellar-mass black holes. These systems will appear as long-lived signals in LISA, with a rich frequency harmonic structure; these sources undergo numerous orbits in the strong gravitational field of the MBH, as a result, the GW waveforms contain a very rich structure of harmonics of integer multiples of the three fundamental frequencies of the orbit. These three fundamental frequencies are usually denoted as $(\Omega_r, \Omega_\theta, \Omega_\phi)$, respectively describing the orbit of the motion in the radial, polar and azimuthal directions.

The orbital trajectory of the compact object in an EMRI is expected to be extremely complex, with the compact object experiencing a very large number of orbits, with high

eccentricity and inclination [79]. As a result the compact object's trajectory effectively maps out the spacetime around the massive black hole over many orbits [80, 81]. Due to the extreme asymmetry and high eccentricity of the binary, the emitted GW is expected to have significant power in $\mathcal{O}(100)$ harmonics, complicating both the waveform modelling and the search [82, 83].

Current EMRI event rate predictions vary significantly, in the range 10-1000 (per year), mainly driven by uncertainties in the astrophysical massive black hole populations [84]. Since MBHs are (in combination with stellar mass compact objects) the progenitors for EMRI systems, a good understanding of the MBH population is a pre-requisite for an accurate prediction of the event rate. These sources also require the gravitational binding of a small compact object to the massive black hole, thus the fraction of MBH that are hosted within galaxies with dense star forming cusps will affect the EMRI rate and this is another origin of uncertainty which is carried through to the event rate [85]. LISA observations will constrain the EMRI event rate, in-turn constraining these astrophysical population models.

EMRIs can act as a probe for AGN disk physics. The central MBH for EMRIs are expected to often be found within AGN disks where the secondary compact object for an EMRI system can be captured from within the disk. Compact objects originating from this formation channel may experience gas torques while inspiralling into the MBH, this will change the orbital trajectory of the secondary object and in-turn cause a phase shift in the GW waveform relative to a source in a vacuum environment. Due to the large number of orbits the compact object will experience in the LISA band, the cumulative phase shift from the gas torque can leave a significant and detectable imprint on the observed signal. While this may complicate modelling these systems (since one must also model the gas torque on the binary in addition to GR driven orbital evolution), if such a signal can be successfully modelled and found, the phase evolution of the waveform will encode the torque experienced by the compact object from which one can infer astrophysical properties of the AGN disk such as the disk viscosity. This would likely be one of the

first direct measurements from a GW signal of the environment around the source. There has been recent work towards this in Ref. [86] where the authors conduct hydrodynamical simulations of an EMRI system with a compact object within an AGN disk and observe the phase shift in the resulting waveform due to the gas torque, Ref. [87] evaluates the detectability of this type of phase shift in the signal, establishing a method for using it to discern between two different AGN disk prescriptions.

This is also expected to be the source type with the best precision in the inferred intrinsic parameters of the binary. The long evolution of many frequency modes imposes strict constraints on the intrinsic parameters. Specifically, in the case of an EMRI the three fundamental frequencies of motion and their evolution are tightly constrained. The fundamental frequencies and their evolution are determined by the intrinsic parameters of the binary, therefore strict constraints on the frequencies in-turn place precise constrains on the intrinsic parameters of the source [88, 89]. While this complicates the search for EMRI signals, assuming such a signal is found and characterized, the source parameters can be used to directly test the no-hair theorem prediction for Kerr black holes with unprecedented precision [84].

3. LISA data analysis

In this chapter some of the key methods that are used in the analysis of GWs are discussed, with a focus on techniques that are currently used in the analysis of LVK data and how these can be translated into the distinct analysis problems that LISA data will bring.

3.1 Search and parameter estimation

Current analysis of GW data from ground-based detectors can be typically divided into two phases: search and parameter estimation. The search phase answers two closely related questions: *Is there a statistically significant signal in the data?* and *What are the approximate parameters describing the astrophysical source?* Meanwhile, the parameter estimation phase refines the source parameter estimates, and establishes robust statistical uncertainties on the inferred parameters.

The search can broadly be interpreted as an attempt to match an observed time-series signal from the detector d (which we suspect may contain a GW signal) with a GW waveform model that we generate with a set of parameters $\boldsymbol{\theta} = \{\theta_1, \theta_2, ..., \theta_N\}$. Generally it is assumed the data contains additive stochastic noise in the form $d = h(\boldsymbol{\theta}_*) + n$, where h is the GW waveform model evaluated at some unknown true parameters $\boldsymbol{\theta}_*$ and where n is the instrumental noise.

3.1.1 Matched filtering

This section derives the matched filter, which is the optimal filter for detecting a signal in additive noise. This is an optimal filter in the sense that it maximizes the signal-to-noise ratio (SNR) for a known signal in stationary Gaussian noise. We will focus on the case of stationary Gaussian noise, which is a good approximation for current ground-based detector observations.

The goal is to find a filter $F(t)$ that maximizes the SNR when applied to a dataset $d(t)$ with a model $h(\boldsymbol{\theta}_*)(t)$. This filter is assumed to be a linear, time-invariant filter which can be applied to d as a convolution of the following form:

$$\rho_{\mathrm{MF}}(t) = \int_{-\infty}^{\infty} d(t)F(\tau - t)d\tau = \int_{-\infty}^{\infty} (h(t) + n(t))F(\tau - t)d\tau,$$

where $\rho_{\mathrm{MF}}(t)$ is the filtered value at time t. The integral can be split into two parts, one for the convolution of the filter with the signal h ($\rho^s_{\mathrm{MF}}(t)$) and one for the noise ($\rho^n_{\mathrm{MF}}(t)$). The SNR of the filter is then given by the ratio of power in the signal to the squared expectation of the power in the noise:

$$\left(\frac{S}{N}\right)_t = \frac{|\rho^s_{\mathrm{MF}}(t)|^2}{|\rho^n_{\mathrm{MF}}|^2}.$$

The convolution theorem allows us to express this integral as a product of Fourier transforms;

$$\rho^s_{\mathrm{MF}}(t) = \int_{-\infty}^{\infty} \tilde{F}(f)\tilde{h}(f)e^{2\pi i f t}df. \tag{3.1}$$

Although we can calculate the waveform model $h(t)$ or $\tilde{h}(f)$, it is not possible to compute an exact analytic closed form expression for the noise case since it is a random process. Instead, the power in the filtered noise must be considered in a statistical sense, instead of directly being able to evaluate $\rho^n_{\mathrm{MF}}(t)$, we must instead consider its root-mean-square expectation.

$$\rho^n_{\rm MF}(t) = \int_{-\infty}^{\infty} \tilde{\rho}^n_{\rm MF}(f) e^{2\pi i f t} {\rm d}f = \int_{-\infty}^{\infty} \tilde{F}(f)\tilde{n}(f) e^{2\pi i f t} {\rm d}f. \tag{3.2}$$

We proceed by computing the root-mean-square expectation of $\rho^n_{\rm MF}(t)$:

$$E[|\rho^n_{\rm MF}(t)|^2] = E\left[\left| \int_{-\infty}^{\infty} \tilde{F}(f)\tilde{n}(f) e^{2\pi i f t} {\rm d}f \right|^2\right],$$

this can be expanded out in terms of the double integrals as follows:

$$E[|\rho^n_{\rm MF}(t)|^2] = E\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tilde{F}(f)\tilde{F}^*(f')\tilde{n}(f)\tilde{n}^*(f') e^{2\pi i (f-f') t} {\rm d}f {\rm d}f'\right].$$

This expression can be simplified using the properties of stationary Gaussian noise. Stationary Gaussian noise can be characterized using a power spectral density (PSD) $S_n(f)$ [1]:,

$$E[n^*(f)n(f')] = \frac{1}{2}\delta(f-f')S_n(f). \tag{3.3}$$

The delta function is a consequence of the noise being stationary, leading to the noise being uncorrelated at different frequencies. Using this to evaluate $E[n(f)n(f')]$ in the above integral, we find:

$$E[|\rho^n_{\rm MF}(t)|^2] = \int_{0}^{\infty}\int_{0}^{\infty} \frac{1}{2}\tilde{F}(f)\tilde{F}^*(f)\delta(f-f')S_n(f) e^{2\pi i (f-f') t} {\rm d}f {\rm d}f'. \tag{3.4}$$

Integrating over f' the delta function collapses to $f = f'$, and we are left with:

$$E[|\rho^n_{\rm MF}(t)|^2] = \frac{1}{2}\int_{0}^{\infty} |\tilde{F}(f)|^2 S_n(f) {\rm d}f. \tag{3.5}$$

Therefore the SNR of the matched filter is given by:

[1] There are two conventions for the PSD, one-sided or two-sided, referring to support in only the positive frequencies or in both positive and negative frequencies (since $n(t)$ is real, $S_n(f) = S_n(-f)$). Throughout this thesis we use the one-sided PSD.

$$\left(\frac{S}{N}\right)_t = \frac{2|\int_{-\infty}^{\infty}\tilde{F}(f)\tilde{h}(f)e^{2i\pi ft}\mathrm{d}f|^2}{\int_0^{\infty}|\tilde{F}(f)|^2 S_n(f)\mathrm{d}f}. \tag{3.6}$$

The aim is to obtain a form $F(t)$ such that this SNR is maximized. Here we use the Cauchy-Schwarz inequality, which states that for any two functions $a(f)$ and $b(f)$:

$$\left|\int_{-\infty}^{\infty}\tilde{a}(f)\tilde{b}^*(f)|\mathrm{d}f\right|^2 \leq \int_{-\infty}^{\infty}|\tilde{a}(f)|^2\mathrm{d}f\int_{-\infty}^{\infty}|\tilde{b}(f)|^2\mathrm{d}f.$$

This is analogous to the dot product of any two vectors being less than the product of their magnitudes, i.e. $|\boldsymbol{a}.\boldsymbol{b}| \leq |\boldsymbol{a}||\boldsymbol{b}|$.

Applying this to maximize the numerator of Eqn. 3.6, let $a(f) = \tilde{F}(f)\sqrt{S_n(f)}e^{2\pi ift}$ and $b(f) = \tilde{h}(f)/\sqrt{S_n(f)}$, then the inequality collapses to[2]:

$$\left(\frac{S}{N}\right)_t \leq 4\int_0^{\infty}\frac{|\tilde{h}(f)|^2}{S_n(f)}\mathrm{d}f. \tag{3.7}$$

This inequality can be made into an equality by selecting $F(f) = \tilde{h}^*(f)e^{-2\pi ift}/S_n(f)$. We have arrived at the matched filter which maximizes the SNR of a signal.

Applying this to the data $d(t)$, Fourier transformed into $\tilde{d}(f)$ (this involves substituting $F(f) = h^*(f)e^{-2\pi ift}/S_n(f)$ into Eqn. 3.1), we obtain the following expression for the filtered data[3]:

$$\rho^2_{\mathrm{mf}} = 4\mathrm{Re}\left[\int_0^{\infty}\frac{\tilde{h}^*(f)\tilde{d}(f)}{S_n(f)}\mathrm{d}f\right] = \langle d|h\rangle\,, \tag{3.8}$$

where we have also labelled this expression as $\langle d|h\rangle$, this is the noise-weighted inner-product between the data and the waveform template, this bra-ket notation will be useful in the following sections.

A common use of the noise-weighted inner product is to construct a detection statistic. While there are several choices for this, the following statistic forms the basis for most of

[2]Since the PSD is symmetric wrt positive and negative frequencies and the integrand is Hermitian symmetric, restricting the integral to positive frequencies leads to an extra factor of 2 in the numerator.

[3]Notice numerator here is also Hermitian symmetric.

the state-of-the-art detection pipelines, e.g. [90]:

$$\rho = \frac{\langle d|h\rangle}{\sqrt{\langle h|h\rangle}}. \tag{3.9}$$

This quantity can be interpreted as a normalized version of $\rho_{\rm mf}$. Notice that in the case of $\rho_{\rm mf}$, constant amplitude prefactors in the template h (such as those arising from luminosity distance) would output larger values of $\rho_{\rm mf}$, however ρ is invariant to changes in the amplitude of the template. Conveniently, this also allows us to rescale h by any constant, which will be useful later in the construction of the template bank.

We are interested in examining the probability distribution of ρ in the case that the data contains only noise and contrasting this to the case where the data also includes a signal which is exactly the waveform template.

First considering the case where $d = n$, the numerator of Eqn. 3.9 is a random variable. Since h is assumed to be constant throughout this analysis, $\langle n|h\rangle$ must also be a Gaussian random variable, therefore it is completely characterized by a mean and variance. The noise is assumed to be zero mean Gaussian noise, $E[n(f)] = 0$, therefore,

$$E[\langle n|h\rangle] = 4\mathrm{Re}\left[\int_0^\infty \frac{\tilde{h}^*(f)E[\tilde{n}(f)]}{S_n(f)}\mathrm{d}f\right] = 0. \tag{3.10}$$

Since $E[\langle n|h\rangle] = 0$, $\mathrm{Var}[\langle n|h\rangle] = E[\langle n|h\rangle^2] - E[\langle n|h\rangle]^2 = E[\langle n|h\rangle^2]$. Combining these results, and using Eqn. 3.3 to simplify the resulting integrals gives $\mathrm{Var}[\langle n|h\rangle] = \langle h|h\rangle$. We are now armed to consider the distribution of ρ in the noise only and noise plus signal cases.

In the noise only case, ρ is a Gaussian random variable distributed according to $\mathcal{N}(0, 1)$. In the noise plus signal case, $\langle d|h\rangle$ can be decomposed into two terms with $\langle d|h\rangle = \langle n|h\rangle + \langle h|h\rangle$. As shown above the first term is a Gaussian random variable with zero mean and variance $\langle h|h\rangle$, while the second term is not a random variable since h is held fixed in this analysis, assuming $\langle h|h\rangle = A^2$ being a measure of the amplitude of the gravitational wave h, ρ is then distributed according to $\mathcal{N}(A, 1)$. At this point it is

common to define $\sqrt{\langle h|h\rangle} = A = \rho_{\text{opt}}$ as the *optimal* SNR of a given template.

Using this, given a datastream d and a template h, a value for ρ can be computed and the probability of obtaining a value of ρ greater than this from $\mathcal{N}(0,1)$ is defined as the false alarm probability. This is a finite integral of a standard Gaussian and therefore can evaluated using error functions. It is worth noting that for current GW detections, the power spectral density describing the instrumental noise is not always Gaussian, therefore additional data-quality checks are incorporated into the search/detection pipelines. Examples of this include the χ^2 test [91] which seeks to verify that the frequency evolution of the signal power matches that expected from the template.

3.1.2 Template banks

We have outlined how given a template h which exists in the data, we can filter the signal out of the data using the matched filter and characterize the significance of the detection in a simple way. However, the task of selecting the source parameters $\boldsymbol{\theta}$ which match the data with $h(\boldsymbol{\theta})$ still remains to be solved.

In principle, aided with infinite computing power, one could generate waveforms for all possible source parameters, and then use the matched filter to select the best match. In reality, we are limited to a finite set of templates, this is known as a template bank. In this section we will denote this as $\{h_i | i = 1, 2, \ldots, N_{\text{bank}}\}$ where N_{bank} is the total number of templates in the bank. For a given datastream d, the process of searching for a signal in the data comes down to computing the matched filter SNR ρ for every template in the bank, and computing the corresponding false alarm probabilities to assess significance.

The templates are placed in corresponding locations in parameter space $\{\boldsymbol{\theta}_i\}$, this space must be covered with sufficient density to ensure that no signal is missed, while also keeping N_{bank} small enough such that the search is computationally feasible.

A natural way to define the similarity (or distance) between two gravitational wave signals is using the *match* M, which is defined as the normalized noise-weighted inner product between the waveform evaluated at two different locations in parameter space θ_1

and θ_2:

$$M(h(\theta_1), h(\theta_2)) = \frac{\langle h(\theta_1)|h(\theta_2)\rangle}{\sqrt{\langle h(\theta_1)|h(\theta_2)\rangle \langle h(\theta_1)|h(\theta_2)\rangle}}.$$

The match varies in the range $0 \leq M \leq 1$ for orthogonal to parallel signals (where parallel signals are identical upto amplitude). The unnormalized version of the match appears when evaluating ρ for $d = n + h_i$ with template h_j when $i \neq j$, in this case,

$$\rho = \frac{\langle n + h_i|h_j\rangle}{\sqrt{\langle h_j|h_j\rangle}} = \frac{\langle n|h_j\rangle}{\sqrt{\langle h_j|h_j\rangle}} + \frac{\langle h_i|h_j\rangle}{\sqrt{\langle h_j|h_j\rangle}} = \frac{\langle n|h_j\rangle}{\sqrt{\langle h_j|h_j\rangle}} + \rho_{i,\mathrm{opt}} M(h_i, h_j),$$

where $\rho_{i,\mathrm{opt}}$ is the optimal SNR of the ith template which is present in the data, the match term here can be interpreted as the fraction of the optimal SNR of template i that is retained when filtering the data with template j (in the notation used here). Consider the parameters θ_* and a small distance away from this point $\theta_* + \Delta\theta$. Since $M(h(\theta), h(\theta, \Delta\theta = 0)) = 1$, we can perform a Taylor approximation around $\Delta\theta = 0$,

$$M(h(\theta), h(\theta + \Delta\theta)) \approx 1 + \frac{\partial M}{\partial \theta^i}\bigg|_{\Delta\theta^k=0} \Delta\theta^i + \frac{1}{2}\frac{\partial^2 M}{\partial\theta^i\theta^j}\bigg|_{\Delta\theta^k=0} \Delta\theta^i\Delta\theta^j.$$

Since by definition $\Delta\theta^k = 0$ is a peak for M, the first derivative term goes to 0, and we are left with:

$$M(\theta, \Delta\theta) \approx 1 + \frac{1}{2}\frac{\partial^2 M}{\partial\theta^i\partial\theta^j}\bigg|_{\Delta\theta^k=0} \Delta\theta^i\Delta\theta^j. \tag{3.11}$$

Define a parameter space metric as $g_{ij}(\theta)$ as:

$$g_{ij}(\theta) = -\frac{1}{2}\left(\frac{\partial^2 M}{\partial\theta^i\theta^j}\right)\bigg|_{\Delta\theta^k=0}. \tag{3.12}$$

The metric can be computed numerically at a given point θ in parameter space using finite difference approximations for the derivative, or by using the modern alternative of automatic differentiation [92]. It is also possible (for some waveforms) to transform to a set of coordinates θ' in which the metric $g_{ij}(\theta')$ is flat or very close to flat [93].

The metric can be related to the mismatch $1 - M$ as:

$$1 - M \approx g_{ij}(\theta)\Delta\theta^i\Delta\theta^j, \tag{3.13}$$

where $g_{ij}(\theta)\Delta\theta^i\Delta\theta^j$ can be interpreted as a proper distance squared in parameter space. However it is worth keeping in mind that $g_{ij}(\theta)$ is generally a local quantity, and will change as one moves around in parameter space. Assuming templates are placed in an N dimensional parameter space, within a uniform lattice with proper length dl, in the worst-case scenario of trying to detect a signal, it must fall $dl/2$ away from all the corners of the lattice, with $g_{ij}(\theta)\Delta\theta^i\Delta\theta^j = N(dl/2)^2$.

Since g_{ij} is also related to the match, $1 - \min(M) \approx N(dl/2)^2$. Given a minimum match that the user desires for the template bank, dl can be selected.

$$dl = 2\sqrt{\frac{1 - \min(M)}{N}}. \tag{3.14}$$

Note that the min operator here is finding the smallest match between the data and all the templates in the bank. Within the considered lattice, each template would then occupy a volume dl^N. The total proper volume of this space can be computed as $V = \int \sqrt{\det|g_{ij}(\theta)|}\mathrm{d}^N\boldsymbol{\theta}$, leading to the total number of templates needed to cover the space as V/dl^N:

$$\mathcal{N} = \frac{\int \sqrt{\det|g_{ij}(\boldsymbol{\theta})|}\mathrm{d}^N\boldsymbol{\theta}}{\left[2\sqrt{\frac{1-\min(M)}{N}}\right]^N}. \tag{3.15}$$

While this lattice-based approach was used in early template-bank designs [94, 95, 96], modern techniques tend to use a stochastic placement of templates in parameter space, which is better suited to high dimensional parameter spaces with curved degeneracies (correlations between parameters) than the lattice approach [97]. The template bank approach to the search for compact binary mergers in LVK data has been extremely successful, detecting binary black hole, binary neutron star and neutron star black hole

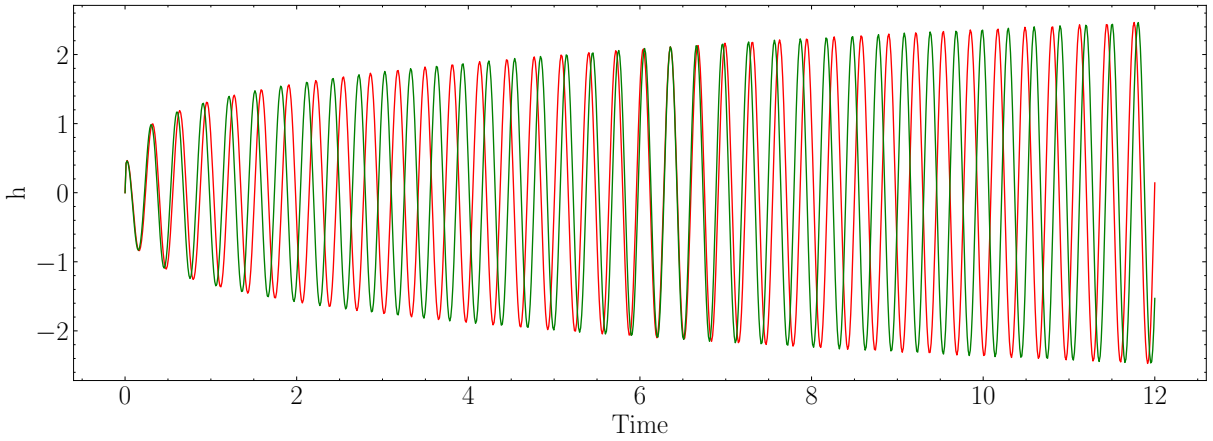

Figure 3.1: Toy model of a linearly chirping sinusoidal wave, the red curve can be considered $h(\theta_*)$ and the green curve as $h(\theta_* + \Delta\theta)$. Here the only (unknown) waveform parameter is $\{\theta\} = \{f_0\}$, i.e the initial frequency of the wave. Both waveforms have the same amplitude (as a function of time) and frequency derivative, and are both initially phase aligned.

mergers. Additionally, the template based search has also been shown to work well for massive black hole binary mergers in simulated LISA data [98], capable of detecting and characterizing binaries before merger [42].

However, this search method becomes prohibitively expensive when looking for signals that are long-lived. This is a well studied issue within the continuous wave community, where there is an ongoing search for long lived, monochromatic GWs emitted from deformed isolated spinning neutron stars [99]. This is closely related to the problem that will plague the search for both stellar-mass binary inspirals and EMRI signals in LISA data, this will be discussed in detail within Chapters. 4, 5 and 6, here we provide some brief intuitive reasoning for this.

Recall that the proper distance in template space is related to the metric by $ds^2 = g_{ij}(\theta)d\theta^i d\theta^j$, and the aim is to place templates that are a calibrated (fixed) proper distance dl apart from one another. Considering the metric about a specific point θ_*, $g_{ij}(\theta_*) = -1/2(\partial^2 M/\partial\theta^i\partial\theta^j)|_{\theta_*}$, therefore $ds \sim \sqrt{|(\partial^2 M/\partial\theta^i\partial\theta^j)|_{\theta_*}|}$.

We will use a toy problem here of a linearly chirping sinusoidal wave, this is qualitatively a good simplified representation for the GW signal emitted during the early inspiral of a binary black hole (but not for the merger). The metric at a given point in parameter

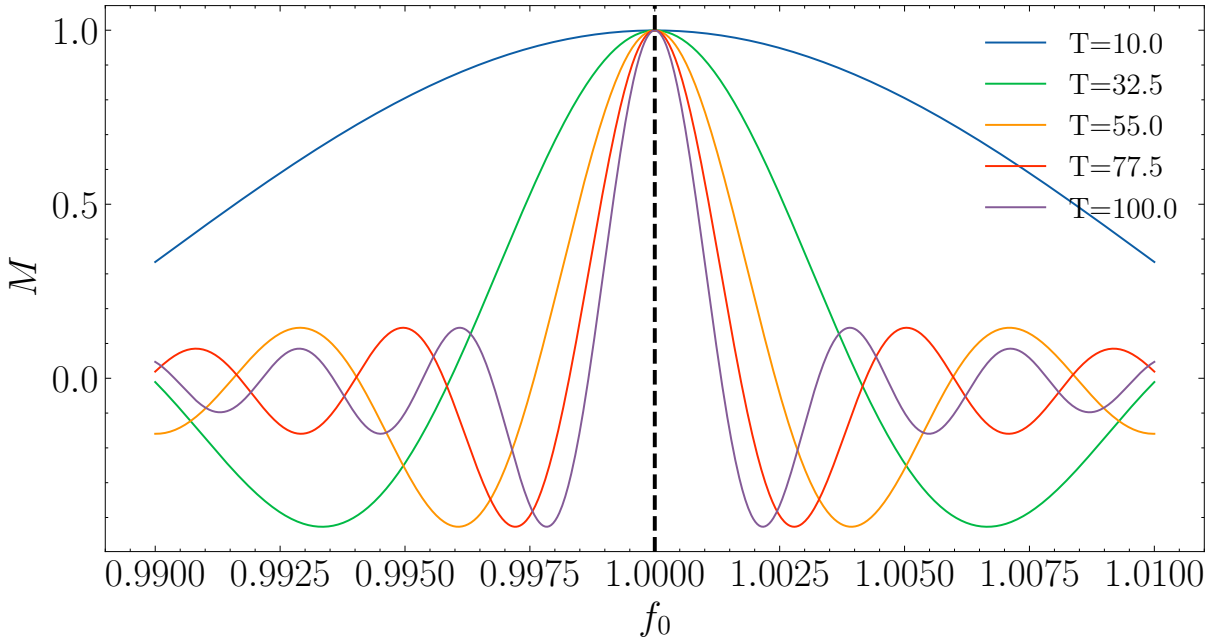

Figure 3.2: Match between a linear chirp with $f_0 = 1$ and a waveform with varying initial frequency, for increasing observation times.

space θ_* is a function of the second derivative of the match between waveform $h(\theta_*)$ and $h(\theta_* + \Delta\theta)$ for a small $\Delta\theta$, this is visualised in Fig. 3.1. For signals with many orbits such as these, even if two waveforms are close in initial frequency, a small deviation in the frequency will cause dephasing over many orbits, leading to a low match. Thus the match is very sensitive to parameters which control the phase of the wave, such as the frequency and frequency derivative. This is shown in Fig. 3.2 where the match against a reference waveform is plotted as a function of the initial frequency of the wave, for increasing observation times with increasing number of orbits. In this simplified toy problem, templates would only need to be placed in the parameter f_0 since it is assumed all other parameters of the wave are known, in this case the proper distance in ds is related to the second derivative of the curves shown in Fig. 3.2. Observing a larger number of orbits leads to a greater sensitivity in the match to f_0 and thus increases ds. Since the aim is to place templates a fixed proper distance dl between them, this increases the number of templates.

The result of this can be seen in the density with which templates are placed for current LVK searches, in the (m_1, m_2) space the low mass region is extremely densely

populated, with generally reduced density at higher masses, see for e.g. Fig. 3 in [100], Fig. 2 in [101] and Fig. 2 in [102]. This is a direct result of the much larger number of orbits that are observed for the inspiral of lower mass sources (such as binary neutron star and neutron star-black hole binaries), as they chirp slower through the $10 - 100\,\mathrm{Hz}$ band that the LVK detectors operate in.

The number of templates that would be needed to cover the astrophysical parameter space for stellar-mass inspiral signals in LISA data has been estimated using a Monte-Carlo approximation of the integral in Eqn. 3.15, requiring $\sim 10^{40}$ templates [103], similar early estimates were made for the cost of the EMRI search [88]. This computational cost is prohibitive, highlighting the need for alternative search strategies.

3.1.3 Parameter estimation

Once a candidate signal has been identified from a search, and is classed as a statistically significant detection, the next step is to infer the source parameters. The search already performs some level of inference and provides a rough location in parameter space where the signal is located, but this *parameter estimation* phase aims to characterize a robust Bayesian posterior probability over all the source parameters.

Bayes theorem is as follows:

$$P(\boldsymbol{\theta}|d) = \frac{P(d|\boldsymbol{\theta})P(\boldsymbol{\theta})}{\mathcal{Z}}, \tag{3.16}$$

where $P(\boldsymbol{\theta}|d)$ is the *posterior* probability over source parameters given the data and an assumed model for the GW, $P(d|\boldsymbol{\theta})$ is the *likelihood* of the data given the source parameters, $P(\boldsymbol{\theta})$ is the *prior* probability over the source parameters and $\mathcal{Z}$ is the *evidence*, corresponding to $\mathcal{Z} = \int P(d|\boldsymbol{\theta})P(\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}$. The likelihood is often denoted as L in the literature and the prior as $\pi(\boldsymbol{\theta})$.

The prior $P(\boldsymbol{\theta})$ is informed by both astrophysics and the limitations of the GW waveform model used for the analysis. An example of a model limitation is that the simple

waveform model derived in Chapter. 1 of this thesis is most accurate for binaries with mass ratios approaching unity. As an example of an astrophysical prior, the component masses for a DWD analysis would be restricted to a maximum of around $1.4\,\mathrm{M}_\odot$ corresponding to the Chandrasekhar limit. Other astrophysical priors include a cosmological distance prior that is uniform in the comoving volume at a given redshift (corresponding to a specific distance).

Under the assumptions of Gaussian stationary noise, characterized by a PSD $S_n(f)$, the likelihood function L in the frequency domain is given by:

$$P(d|\boldsymbol{\theta}) = L(d|\boldsymbol{\theta}) \propto e^{-\frac{1}{2}\langle d-h(\boldsymbol{\theta})|d-h(\boldsymbol{\theta})\rangle}. \tag{3.17}$$

Note that we will switch from the integral form to the discretized version of the noise-weighted inner product from here, i.e.:

$$\left\langle \tilde{d}(f)\middle|\tilde{h}(f)\right\rangle = 4\mathrm{Re}\left[\sum_{i=0}^{N-1}\frac{\tilde{d}^*(f)\tilde{h}(f)}{S_n(f)}\Delta f\right], \tag{3.18}$$

where $\Delta f = 1/T_{\mathrm{obs}}$ is the natural frequency resolution of the data [104]. In the case of multiple detectors, it can usually safely be assumed that noise is uncorrelated between the different detectors, therefore the total multi-detector likelihood is the product of the likelihoods for each detector.

From here onwards, we will deal with the likelihood in its natural logarithm form:

$$\log L(d|\boldsymbol{\theta}) = -\frac{1}{2}\langle d-h(\boldsymbol{\theta})|d-h(\boldsymbol{\theta})\rangle + \mathrm{const}, \tag{3.19}$$

where we have neglected to include the constant as it is only relevant when $S_n(f)$ is not known and is being inferred, i.e. when $S_n = S_n(f, \boldsymbol{\theta})$, in almost all situations within this work we consider cases where $S_n(f)$ is assumed to be known; in practice the PSD can be inferred from a periodogram of signal-free data surrounding the event[4]. In this form the multiple detectors can be simply combined by summing the log-likelihoods from each

[4]The LISA global fit is a situation in which is not true; the PSD is inferred alongside the signal.

detector.

As a way of intuitively connecting the search and parameter estimation phases, it can be shown that the matched filter SNR ρ is equivalent to the maximum likelihood estimator for L. Thus, the search phase provides a good starting point for the parameter estimation.

The aim of the parameter estimation phase is mainly to map the posterior distribution $P(\boldsymbol{\theta}|d)$. As the posterior spans a high-dimensional parameter space, it is not feasible to compute it exhaustively. Instead, stochastic sampling-based methods are used to explore it, these produce a discerete set of samples that are distributed according to the posterior. Examples of commonly used sampling methods include Markov Chain Monte Carlo (MCMC) based algorithms such as Metropolis-Hastings [105] and slice sampling [106], popular alternatives to MCMC algorithms include variants of nested sampling [107]. The resulting discrete set of posterior samples are conventionally visualized using a corner plot, which shows one and two-dimensional marginalized subsets of the posterior.

The process of marginalization is equivalent to integrating out nuisance parameters that are not of interest in the context of inference. A good example in the case of a BBH merger may be the phase of coalescence ϕ_c, this is parameter that shifts the waveform phase at the time of merger. Marginalisation can be performed in two ways, numerically or analytically. For a posterior $P(\boldsymbol{\theta}|d)$, the marginal distribution $P(\theta_i)$ can be found by:

$$P(\theta_i) = \int P(\boldsymbol{\theta}|d) \prod_{j \neq i} \mathrm{d}\theta_j. \tag{3.20}$$

Since stochastic sampling produces a set of discrete samples drawn from the full posterior, marginalisation across the dimensions θ_j is simply equivalent to taking only the samples in the θ_i dimension and disregarding θ_j for each sample. This is equivalent to the Monte-Carlo estimate of Eqn. 3.20 over the posterior samples. This is a way of numerically obtaining the marginalised posterior. As an alternative, the likelihood can also be directly marginalised analytically over such nuisance parameters. When this is possible, the main benefit is that the marginalised likelihood has a lower number of dimensions and thus the sampling for parameter estimation is less computationally expensive, while still taking

the uncertainities in the marginalised parameters into account. Analytically marginalised likelihoods exist for common nuisance parameter such as phase at coalescence [104]. While marginalisation is the Bayesian approach to handling nuisance parameters, maximisation is the frequentist analog to this and is often used in the search phase [108].

While stochastic sampling of the posterior is the conventional method for inference in GW astronomy, a popular modern alternative is based on Simulation-Based Inference [109]. These methods were originally conceptualized for cases where the likelihood is either analytically or computationally intractable, e.g. [110]. This method directly infers the posterior from the data, and is able to produce posterior samples rapidly without the costly exploration of parameter space involved in the previously mentioned "classical" stochastic sampling methods. In exchange for near-instant inference at runtime, a costly 'offline'[5] training step is required, where a normalizing flow is trained to learn the mapping between the data and the posterior. A normalizing flow is a set of successive functions (each with tunable parameters) that can analytically map from a simple probability distribution (such as a Gaussian) to a more complex distribution [111]. Simulation based inference methods are robust to more general noise properties than those assumed in the standard Whittle likelihood, such as non-stationary and non-Gaussian noise, this is done simply by including simulations of this more complicated noise spectrum within the training data [112]. Similar methods have recently been tested in a LISA context, see e.g. [113, 114].

Fisher matrix

This section presents a simple derivation of the Fisher Information matrix (FIM) which is a common tool in preliminary GW data analysis studies to roughly estimate source parameter uncertainties without resorting to full inference via stochastic sampling.

Assume we wish to investigate the likelihood for a source with parameters θ_* in the neighbourhood of θ_*, the first step is to use a linear signal approximation, where it is assumed that changes to the waveform in response to a small change $\Delta\theta$ in the source

[5]Offline meaning this can be done before the inference is run on real data, analogous to the template construction step for template banks.

parameters are small and linear:

$$h(\theta_* + \Delta\theta) \approx h(\theta_*) + \left.\frac{\partial h}{\partial \theta^i}\right|_{\theta=\theta_*} \Delta\theta^i.$$

The data is assumed to contain noise and the signal, $d = h(\theta_*) + n$. In this form using the linear signal approximation, the log-likelihood can be written as:

$$\begin{aligned}\log L(\theta) &= -\frac{1}{2}\left\langle n - \left.\frac{\partial h}{\partial \theta^i}\right|_{\theta=\theta_*}\Delta\theta^i \middle| n - \left.\frac{\partial h}{\partial \theta^j}\right|_{\theta=\theta_*}\Delta\theta^j \right\rangle \\ &= -\frac{1}{2}\left[\langle n|n\rangle - 2\left\langle n \middle| \left.\frac{\partial h}{\partial \theta^i}\right|_{\theta=\theta_*} \right\rangle \Delta\theta^i + \left\langle \left.\frac{\partial h}{\partial \theta^i}\right|_{\theta=\theta_*} \middle| \left.\frac{\partial h}{\partial \theta^j}\right|_{\theta=\theta_*} \right\rangle \Delta\theta^i\Delta\theta^j \right].\end{aligned}$$

The first term $\langle n|n\rangle$ is constant with respect to changes in θ and thus can be ignored when investigating the behaviour of L as a function of the source parameters. The second term can be rewritten as $\partial/\partial\theta^i[\langle n|h\rangle]$, since $\langle n|h\rangle \sim \mathcal{N}(0,1)\rho_{\rm opt}$, this term is a random variable drawn from the distribution $\mathcal{N}(0,1)\partial\rho_{\rm opt}/\partial\theta^i$. In the large SNR limit the signal will (instantaneously) obey $|n| \ll |h|$, therefore the likelihood is dominated by the final term:

$$\log L(\theta) \approx -\frac{1}{2}\left[\left\langle \left.\frac{\partial h}{\partial \theta^i}\right|_{\theta=\theta_*} \middle| \left.\frac{\partial h}{\partial \theta^j}\right|_{\theta=\theta_*} \right\rangle \Delta\theta^i\Delta\theta^j \right]. \tag{3.21}$$

In this form the likelihood is a Gaussian distribution with a covariance matrix given by $(\Gamma^{-1})_{ij}$, where Γ is:

$$\Gamma_{ij} = \left\langle \left.\frac{\partial h}{\partial \theta^i}\right|_{\theta=\theta_*} \middle| \left.\frac{\partial h}{\partial \theta^j}\right|_{\theta=\theta_*} \right\rangle, \tag{3.22}$$

where Γ_{ij} is denoted as the Fisher information matrix. Fisher matrices allow for a quick estimate of source parameter uncertainty and correlations at one location in parameter space, avoiding the computational burden of performing a full parameter estimation via stochastic sampling. This has been used in several papers to perform early estimates to determine the detectability of certain physical effects, eg. [115, 116, 117]. It is also useful in calibrating the proposal distribution for stochastic sampling, leading to faster convergence to the posterior when using stochastic sampling algorithms such as Metropolis-Hastings

MCMC, e.g. see Ref. [118] and Sec. 5.4.3 of Ref. [119].

However there are a number of limitations to the Fisher matrix, one of them being that it is only valid for the case of high SNR, where the likelihood can be well approximated to a Gaussian. The term $\langle n|\partial h/\partial\theta^i\rangle$ is no longer subdominant in the moderate-low SNR case, thus Eqn. 3.22 can no longer be regarded as a good approximant for the likelihood. While most MBHB mergers observed by LISA will likely be sufficiently high in SNR, sources such as stellar-mass inspirals will likely be quieter and thus the Fisher matrix may not be a good approximation for these sources. Another key limitation of the FIM is that it is only able to probe linear correlations in source parameters, due to the linear signal approximation, this is a problem for cases where there are strong non-linear correlations between source parameters.

A practical limitation of the FIM is that it requires waveform derivatives with respect to the source parameters, which are not readily available (at least in an analytic form) for all waveform models. These derivatives are often approximated using finite difference schemes, however it can be difficult to ensure that these derivatives are numerically stable and accurate (ensuring an appropriate step-size is chosen). The recent advent of automatic differentiation has made this problem less severe, as it allows for the computation of accurate derivatives without the need of finite difference schemes [120].

3.1.4 Overlapping sources → Global fit

LISA data will bring about a unique challenge that has not been encountered thus far in the ground-based detector era, the problem of many signals overlapping in both time and frequency. In current LVK detectors, GW signals are short-lived and well separated in time, the longest observed signal to date - from a binary neutron star merger - lasted approximately 100 seconds [121]. LVK data either contains only noise or noise and a single GW signal, as a result the search and parameter estimation must scan over the parameter space of one signal. LISA data will contain many overlapping signals (in both time and frequency) from the moment that the detector begins observations. As

discussed in Chapter 2 $\mathcal{O}(10^8)$ monochromatic ultra-compact binaries will be present in the LISA datastream throughout the mission, potentially alongside several stellar-mass binary inspiral and EMRI signals, with the occasional $\mathcal{O}$(month) long signals from the merger of massive black hole binaries.

Consequently, LISA data analysis requires simultaneous inference over the parameters of all present signals - an unknown and potentially large number - within the data. This is known as the global fit problem [122]. The dimensionality of the full global fit is mostly driven by the UCBs in the data, due to the large number ($O(10^4)$) that we expect to be individually resolvable, each with 8 parameters (assuming the small fraction of eccentric sources are ignored).

The problem of disentangling multiple sources from a single datastream is not unique to GW astronomy, this problem was initially named *the cocktail party problem*[6], describing a cocktail party where many people are speaking at the same time and the listener aims to disentangle these voices [123]. The characteristics of each person's voice being different in properties like pitch, volume and accent allowed the listener to effectively disentangle the voices. Analogously, in the case of sources observable by LISA there are several phenomenological differences between the signals emitted by different sources which allows for a simplification of the global fit problem.

Within the global fit, there are two main types of correlations that must be taken into account: inter-source correlations and intra-source correlations. Correlations here refer to two signals which have a non-negligible match M, this leads to correlations in their source parameter posteriors. Inter-source correlations are correlations between different astrophysical sources (e.g. between a UCB and a massive black hole merger), while intra-source correlations are correlations between different signals from the same source category (e.g. correlations between different UCB signals).

Intra-source correlations will likely not be a large problem for the quasi-monochromatic

[6]The LISA data challenges, respectively called *Radler, Yorsh, Sangria, Spritz* and the upcoming *Mojito* are names of famous cocktails, perhaps as a reference to the original cocktail party problem. Analogous to the increasing alcohol content of the cocktails, the data challenges get progressively more difficult.

long-lived slowly chirping EMRIs and stellar-mass binary inspirals due to the same argument as that presented in Sec. 3.1.2. Unless signal parameters are extremely close together, the match between the two signals within either of these source types is expected to be small, this was demonstrated for EMRI signals in Ref. [89]. In the case of UCBs, these sources are expected to appear as extremely narrowband sources in the frequency domain, with the width of the signal primarily being controlled by the doppler shift due to the periodic motion of LISA. Therefore two UCBs will likely only be strongly correlated when they have similar GW frequencies, state-of-the art search and global fit pipelines for these sources split the frequency range in which UCBs are expected into small segments, and search for multiple UCBs simultaneously within each segment. This significantly reduces the computational cost of the global fit as one must only simultaneously update the source parameters of UCBs within the same frequency segment [124]. It is important to note that this does not mean the UCBs in a given segment are not correlated with those in a different segment, but rather that they are likely to be weakly correlated. A block-Gibbs sampling approach is well suited to this type of inference problem where subsets of source parameters are expected to be more strongly correlated than others [125]. An early implementation of this sampling method for the global fit is presented in Ref. [126]. In the case of MBHBs, the signals are no longer compact in frequency (compared to the UCBs), but they enter the LISA frequency band and merge within $\sim$ months, therefore these signals are compact in time with the majority of the SNR accumulated near/at merger. Thus searches for these signals are performed in small segments of time, to minimise the chance of dealing with correlated signals, e.g. see Sec. 5.B of Ref. [63].

Inter-source correlations are generally less well studied in the literature, the most investigated source pair correlations are between UCBs and MBHBs. This is because the majority of galactic binaries are expected to radiate at frequencies $\lesssim 10^{-3}\,\mathrm{Hz}$, this is the rough frequency at which MBHB mergers will occur, therefore ignoring the presence of UCBs when analysing a section of data suspected to contain an MBHB may lead to significant biases in the inference of the MBHB source parameters, this will be more

important for the analysis of quiet MBHB signals. The current state-of-the-art global fit implementations consider data containing noise and MBHB and UCB sources (with the additional task of estimating a noise PSD, as part of the inference) [63, 127, 126, 128]. It was found that overestimating the power-spectral-density of the noise in the early parts of the global fit (to include detectable UCBs) did not hinder detection of the MBHB signals, however since only relatively loud MBHB signals were present in the simulated datastreams this result is unlikely to hold for quieter mergers.

As previously discussed in the special case of the long-lived quasi-monochromatic signals, the EMRIs and stellar-mass binary inspirals, they are unlikely to be confused with MBHBs or UCBs as the signal morphology in the frequency domain is completely different. For this reason it is possible that searches for these sources can be performed independently to the search for MBHB and UCBs. However, the possibility of confusion between one frequency mode of an EMRI and a stellar-mass inspiral needs to be investigated as this can cause potential problems for the search of both sources.

A lower-dimension version of the global fit problem is already faced in current LVK analysis, in the form of modelling short glitches which are coincident with short transient burst-like signals. Short glitches are not extremely problematic when the astrophysical GW signal in the detector is long (such as BNS), as the SNR accumulates over a long timescale, however on a shorter timescale glitches can be problematic as they can mimic the signal morphology of a short burst-like GW signal. In this case it is necessary to model both the glitch and signal simultaneously, in a mini global fit-like framework. This is what is done in `BayesWave` [129], where parameter estimation is performed using a trans-dimensional MCMC algorithm (Reversible Jump MCMC [130]), this sampling method allows for the number of parameters in the model to change, in this case it can fit a variable number of glitches to a datastream, with corresponding posteriors over parameters describing each glitch.

3.1.5 Noise properties and characterizing significance

As discussed in Sec. 3.1.1, usually the noise within the data is assumed to be stationary. This requires that $E[\langle n(t)|n(t+\tau)\rangle]$ is only a function of τ the time difference. Equivalently, the noise can be fully described by a PSD $S_n(f)$ which depends on the frequency alone. This is one of the key assumptions behind the Whittle likelihood. If this assumption is voided, and the noise is non-stationary, the PSD becomes $S_n(f, f')$, a function of two frequencies, and the noise weighted inner product in Eqn. 3.8 becomes a sum over the two frequencies, which is computationally expensive.

The LISA noise is made up primarily of two contributions, the acceleration (low-frequency) noise and the interferometer (high-frequency) noise. In the case of LVK detectors, the low frequency noise is dominated by seismic noise, while the LISA low frequency noise is dominated by test-mass acceleration; since the test-masses inside each spacecraft are supposed to be free-falling, any acceleration of the test-mass due to factors such as radiation pressure or thermal noise reduces the sensitivity of the detector to low frequency GWs [25]. The test-mass acceleration noise was demonstrated to be below the required-level by LISA Pathfinder [131]. In the high frequency regime, the LISA noise budget is dominated by interferometer noise, this includes contributions from the instrument readout and the laser shot noise (which dominates the interferometer noise at high frequencies), see Fig. 7.2 in Ref. [25] for the relative contributions of the various noise sources. It is worth emphasizing that the noise arising from fluctuations in the laser frequency is dominant over the whole LISA spectrum (See Fig. 1.2 in Ref. [132]), the other noise sources discussed above are subdominant in comparison; the suppression of this laser frequency noise using time-delay interferometry will be discussed further in Sec. 3.3. The LISA data is expected to be sampled at 4 Hz by the instrument [133], however the data is commonly downsampled to a rate of 0.2 Hz for science analysis, this corresponds to a maximum Nyquist frequency of 0.1 Hz. Additionally, frequencies above this are not expected to contribute significant SNR to the detection of signals as the noise PSD is

steeply growing in this frequency range. As a result, most analyses of high frequency sources use a maximum frequency of 0.1Hz.

While stationarity is a good assumption for noise in the LVK detectors over the typical $\lesssim$ 1s merger timescale, signals present in LISA data can persist over the whole mission lifetime; the noise will not be stationary over these long timescales. In addition to changes in the instrumental noise spectrum, the astrophysical confusion noise from galactic UCBs (which is often modelled as part of the noise PSD) will also change over the mission lifetime, as discussed in Sec. 2.2. Additionally, the LISA datastream is also expected to contain gaps; there will be scheduled gaps during the process of repointing the satellites, and there may be unscheduled gaps due to the need to perform emergency maintenance on the detector [134]. Such gaps violate the stationarity assumption and introduce discontinuities that must be handled in data analysis pipelines. One of the most promising solutions to these problems is the time-frequency approach to the analysis, this will be discussed in Sec. 3.4, in summary, this approach allows for easy handling of data gaps and allows one to approximate a non-stationary timeseries as stationary within small time scales [129, 56].

In the current LVK detectors, signals are short, well separated and the large majority of the data contains only noise[7]. Thus a periodogram of the whole (long) data is a good approximation for the PSD describing the instrumental noise. This will not be possible for LISA since the data will always have GW signals present within it. Consequently, the noise PSD must be inferred jointly with signal parameters as part of the global fit, it cannot be easily estimated from a periodogram of signal-free data.

Characterizing the significance of a detection is another part of the analysis that will require different methods in LISA than those currently used in ground based detectors. In the LVK detectors, the significance of a detection is quantified by a false-alarm rate (FAR), this is the rate at which a noise feature could mimic a GW signal with the detected SNR in all the detectors at once. The process for computing this quantity involves an

[7]With the notable exception of the theorised continuous wave signals.

'offline' step of shifting the data in each detector by more than the expected travel time of the GW signal between the geographical locations of the detectors, ensuring that in this shifted artificial data, no GW signals can be present and any coincident 'detections' must arise from noise transients. This artificial data stream is then searched. Through searching a series of these time-shifted datastreams a distribution of 'false triggers' can be built up, with a corresponding matched filter SNR for each. In the 'online' part of a real low-latency search, the SNR of any detected signal can be compared against the 'background' distribution of triggers to find the rate at which a noise trigger with the same or greater SNR can occur. Practically this is done by comparing the SNR of the detected search trigger to that corresponding to a maximum threshold FAR [90]. Since the LISA data will be signal dominated the time-slide method will not be applicable, e.g. time shifted interferometer channels will still contain the same almost monochromatic DWD signals. In summary, non-stationary noise, data gaps, and the lack of signal-free data necessitate new approaches to both noise modelling and significance estimation in the LISA era.

3.1.6 Stationary phase approximation

Motivated by the previous discussions of the convenience of frequency domain analyses when the noise is stationary and Gaussian, we seek analytic frequency domain representations of waveforms. In this chapter we outline one avenue for obtaining this with the leading order Newtonian waveform that is derived in Chapter 1.

The stationary phase approximation (SPA) is a common method to deal with oscillatory integrals, isolating and only considering the part of the integrand which varies slowly and contributes significantly, discarding the oscillatory part. In the context of GW modelling it is a useful tool to analytically compute an approximant for the Fourier transform of a time-domain waveform [135], assuming it satisfies some criteria, discussed in this section.

The time-domain waveform $h(t)$ can be decomposed as the following, $h(t) = A(t)e^{i\phi(t)}$,

where $A(t)$ and $\phi(t)$ are real-valued functions of time, the Fourier transform of this is:

$$\tilde{h}(f) = \int_{-\infty}^{\infty} h(t)e^{-2\pi i f t}\mathrm{d}t = \int_{-\infty}^{\infty} A(t)e^{i\phi(t)}e^{-2\pi i f t}\mathrm{d}t = \int_{-\infty}^{\infty} A(t)e^{i[\phi(t)-2\pi f t]}\mathrm{d}t.$$

Defining $\Psi(t) = \phi(t) - 2\pi f t$, the integrand is oscillatory where the difference between $\phi(t)$ and $2\pi f t$ is large, this leads to contributions to $\tilde{h}(f)$ cancelling. The only significant contribution is near the stationary point of $\Psi(t)$, where $\partial\Psi/\partial t = 0$ in this region the integrand phase varies slowly. This can be used to relate $h(f)$ at a certain frequency f to a time at which the integrand is stationary, known as the stationary t_s:

$$\left.\frac{\partial\Psi}{\partial t}\right|_{t=t_s} = \left.\frac{\partial}{\partial t}\big(\phi(t) - 2\pi f t\big)\right|_{t=t_s} = 0 \Rightarrow \left.\frac{\partial\phi}{\partial t}\right|_{t=t_s} = 2\pi f. \tag{3.23}$$

The integrand phase can be Taylor expanded as follows around t_s:

$$\begin{aligned}\Psi(t) &= \Psi(t_s) + \left.\frac{\partial\Psi}{\partial t}\right|_{t=t_s}(t-t_s) + \frac{1}{2}\left.\frac{\partial^2\Psi}{\partial t^2}\right|_{t=t_s}(t-t_s)^2 + \cdots, \\ &= \phi(t_s) - 2\pi f t_s + \frac{1}{2}\left.\frac{\partial^2\phi}{\partial t^2}\right|_{t=t_s}(t-t_s)^2,\end{aligned} \tag{3.24}$$

where the first derivative term vanishes at the stationary time. This expression is used to approximate the phase of the integrand around the stationary point. Using this, the Fourier transform is approximated as:

$$\tilde{h}(f) \approx \int_{-\infty}^{\infty} A(t)e^{i\left[\phi(t_s)-2\pi f t_s + \frac{1}{2}\left.\frac{\partial^2\phi}{\partial t^2}\right|_{t=t_s}(t-t_s)^2\right]}\mathrm{d}t. \tag{3.25}$$

The amplitude of the gravitational wave $A(t)$ is assumed to be varying on a timescale much larger than the phase, such that $A(t)$ can be safely approximated by $A(t_s)$. Under this assumption, the integral takes the form

$$\tilde{h}(f) \approx A(t_s)e^{i\left[\phi(t_s)-2\pi f t_s\right]}\int_{-\infty}^{\infty} e^{\frac{i}{2}\left.\frac{\partial^2\phi}{\partial t^2}\right|_{t=t_s}(t-t_s)^2}\mathrm{d}t,$$

Recognising that $f = 1/2\pi\,\partial\phi/\partial t, \rightarrow \partial^2\phi/\partial t^2 = 2\pi\partial f/\partial t$ using the time-frequency relation discussed earlier, the integral can be written as:

$$\tilde{h}(f) \approx A(t_s)e^{i\left[\phi(t_s)-2\pi f t_s\right]}\int_{-\infty}^{\infty} e^{\frac{i}{2}\frac{\partial f}{\partial t}\Big|_{t=t_s}w^2}\mathrm{d}w, \tag{3.26}$$

where $w = t - t_s$. This is a standard Gaussian integral and the following identity can be used:

$$\int_{-\infty}^{\infty} e^{\frac{i}{2}Cx^2}\mathrm{d}x = \sqrt{\frac{2\pi i}{C}} = \sqrt{\frac{2\pi}{C}}e^{\pm i\frac{\pi}{4}}, \tag{3.27}$$

for $C > 0$. Using the identity leads us to the SPA approximation for $\tilde{h}(f)$:

$$\tilde{h}(f) \approx A(t_s)e^{i\left[\phi(t_s)-2\pi f t_s \pm \frac{\pi}{4}\right]}\left[\frac{\partial f}{\partial t}\bigg|_{t=t_s}\right]^{-\frac{1}{2}}. \tag{3.28}$$

Practically, $\partial f/\partial t$ can be analytically computed from the time domain waveform using $\partial^2\phi/\partial t^2$.

The SPA can be used to approximate the Fourier transforms of $h_+, h_\times$ given in Sec. 1.2.1. Applying this to signals of the form $A(t)\cos(\phi(t))$ leads to the following expression for $h(f)$ (switching between the two polarisation states simply corresponds to a $\pi/2$ phase shift.):

$$\tilde{h}(f) = \sqrt{\frac{5}{24}}\frac{\mathcal{M}_c^{\frac{5}{6}}G^{\frac{5}{6}}}{D_L f^{\frac{7}{6}}\pi^{\frac{2}{3}}}\exp\left[2\pi f t_c + \phi_c - \frac{3}{128}\frac{c^5}{(\pi f G\mathcal{M}_c)^{\frac{5}{3}}}\right]. \tag{3.29}$$

Waveforms considered in this thesis are all evaluated directly in the Fourier domain and although the models used in Chapters. 4, 5 and 6 are more physically complete than that derived above, their leading order behaviour (in the inspiral phase) is well described by this result.

The SPA is a good approximation for the Fourier transform of a signal when the frequency of the signal is broadband, slowly evolving with time and the amplitude is changing on a much larger timescale than the frequency. Note that while the derivation

we have outlined is for a signal with one harmonic, the presence of multiple harmonics does not necessarily break the validity of the SPA, since it can be applied to each harmonic individually, each harmonic obtains its own time-frequency correspondence specified by the SPA.

This approach fails when the amplitude evolves on a similar timescale to the frequency, this is the case for the merger GW signals of black hole binaries, in these cases the SPA is no longer a good approximation to the Fourier domain waveform. Another requirement of the SPA is that the time-frequency map is monotonic, i.e. that there is a one-one correspondence between time and frequency, in certain cases e.g. EMRI signals the time-frequency correspondence is not monotonic even in the inspiral phase [82]. The SPA also fails for stellar-mass binary inspirals when there is significant spin-induced orbital precession, as the time-frequency map can be non-monotonic [136].

In some cases where the SPA fails, the closely related method of shifted uniform asymptotic (SUA) can be used to obtain a good approximation for the Fourier transform [137].

3.2 LISA response function

The leading-order GW model derived in Chapter 1 is presented in the h_+, $h_\times$ polarisation basis, this is a coordinate frame suited to the GW, however there remains an additional step in projecting the waveform onto the LISA detector, this projection is often labelled as a detector response. In the case of current ground based detectors, this projection is the angular rotation of the waveform required to project it onto the arms of the interferometer. This projection for current ground based detectors makes use of the long-wavelength approximation, which is valid when the wavelength of the GW is much larger than the arm length of the detector. Under this approximation, the GW is assumed to be a plane wave which arrives at all parts of the detector arm at the same time, the time delay arising from the propagation of the GW (along the arm) is negligible compared to the period of

the GWs. However, this will not be the case for the high frequency sources LISA expects to observe (EMRI and SmBBH inspirals), thus a response which models the time-delays that arise from the propagation of the GW within the laser arms is required. Additionally, detectors situated on the Earth do not move significantly over the course of short single compact binary coalescence observations. In the case of LISA, the motion of the detector must be taken into account, as a consequence of these two factors the response function is time dependent.

The two satellites (sending and receiving the laser) will be following paths parameterised by $\boldsymbol{x_s}$ and $\boldsymbol{x_r}$, with spatial components indexed by $i \in \{1, 2, 3\}$. The time of transmission and arrival of the laser at the respective spacecraft are denoted by (t_s, t_r).

The 'unperturbed' photon path can be parameterised by the 3-vector $\boldsymbol{x_0}(\lambda) = \boldsymbol{x_s} + \lambda(\boldsymbol{x_r} - \boldsymbol{x_s})$, where $\lambda \in (0, 1)$ is a parameter describing the position on the path. Additionally, the time derivative of the 3-vector $\dot{\boldsymbol{x}}_0 = \partial \boldsymbol{x_0}/\partial\lambda = \boldsymbol{x_r} - \boldsymbol{x_s}$. The 4-vector which describes this path is $x_0(\lambda) = (t_s + \lambda(t_r - t_s), \boldsymbol{x_0}(\lambda))$, with corresponding derivative $\dot{x}_0 = (t_r - t_s, \boldsymbol{x_r} - \boldsymbol{x_s})$. Since x_0 is the unperturbed path, it must be a null vector with respect to $\eta_{\mu\nu}$, thus:

$$\eta_{\mu\nu}\dot{x}_0^\mu \dot{x}_0^\nu = 0.$$

The perturbation to the full metric $g_{\mu\nu}$ as a result of a passing GW is parameterised in the same way as in Chapter 1, introducing a order counting parameter ϵ:

$$g_{\mu\nu} = \eta_{\mu\nu} + \epsilon h_{\mu\nu} + \cdots .$$

Recall that throughout this thesis we are using the TT gauge for the GW, in these coordinates when a GW passes through two particles in Minkowski flat space the positions of the detectors do not change as a function of time, see discussion in Sec. 1.3. As a result of this, in these coordinates the full unperturbed photon path (3-vector) is equivalent to the complete path: $\boldsymbol{x}(\lambda) = \boldsymbol{x_0}(\lambda)$. The complete photon path must be null with respect

to the full metric, therefore:

$$g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu = \eta_{\mu\nu}\dot{x}_0^\mu\dot{x}_0^\nu + \epsilon h_{\mu\nu}\dot{x}_0^\mu\dot{x}_0^\nu + \mathcal{O}(\epsilon^2) = 0. \tag{3.30}$$

The first term can be expanded as the following $\eta_{\mu\nu}\dot{x}_0^\mu\dot{x}_0^\nu = -\dot{t}^2 + |\dot{\boldsymbol{x}}_\mathbf{0}|^2$ [8], where $|\dot{\boldsymbol{x}}_\mathbf{0}|^2 = |\boldsymbol{x_r} - \boldsymbol{x_s}|^2$ is the (unperturbed) square distance between the two spacecraft. Using this,

$$\dot{t} = \left[|\dot{\boldsymbol{x}}_\mathbf{0}|^2 + \epsilon h_{\mu\nu}\dot{x}_0^\mu\dot{x}_0^\nu\right]^{\frac{1}{2}} = |\dot{\boldsymbol{x}}_\mathbf{0}|\left[1 + \frac{\epsilon h_{\mu\nu}\dot{x}_0^\mu\dot{x}_0^\nu}{|\dot{\boldsymbol{x}}_\mathbf{0}|^2}\right]^{\frac{1}{2}}. \tag{3.31}$$

Since the second term within the brackets is expected to be a small perturbation on the distance between the two spacecraft, this expression can be Taylor expanded to the leading order[9]:

$$\dot{t} \approx |\dot{\boldsymbol{x}}_\mathbf{0}|\left[1 + \frac{1}{2}\frac{\epsilon h_{\mu\nu}\dot{x}_0^\mu\dot{x}_0^\nu}{|\dot{\boldsymbol{x}}_\mathbf{0}|^2}\right]. \tag{3.32}$$

The path length travelled by the photon L can be calculated as:

$$L = \int_0^1 \dot{t}\mathrm{d}\lambda = \int_0^1 |\dot{\boldsymbol{x}}_\mathbf{0}|\left[1 + \frac{1}{2}\frac{\epsilon h_{\mu\nu}\dot{x}_0^\mu\dot{x}_0^\nu}{|\dot{\boldsymbol{x}}_\mathbf{0}|^2}\right]\mathrm{d}\lambda = L_0\int_0^1 \left[1 + \frac{1}{2}\frac{\epsilon h_{\mu\nu}\dot{x}_0^\mu\dot{x}_0^\nu}{|\dot{\boldsymbol{x}}_\mathbf{0}|^2}\right]\mathrm{d}\lambda. \tag{3.33}$$

Since $|\dot{\boldsymbol{x}}_\mathbf{0}|$ is a constant and not a function of λ, we denote it with L_0 and place it outside the integral. The second term in the integral is the leading order correction to the unperturbed path length arising from the passing GW. The integral can be seperated into two parts as follows:

$$L = L_0 + \frac{1}{2L_0}\int_0^1 h_{\mu\nu}(\lambda)\dot{x}_0^\mu\dot{x}_0^\nu\mathrm{d}\lambda. \tag{3.34}$$

This can be simplified further once again by imposing the gauge conditions described

[8] Using natural units.

[9] Notice the similarity between this equation and Eqn. 1.37, the beginning of both derivations follows the same path.

in Chapter. 1, where the GW metric perturbation $h_{\mu\nu}$ is only nonzero for the spatial elements, therefore $h_{\mu\nu} \to h_{ij}$. Additionally, recall the derivative terms $\dot{x}_0^\mu = x_r^\mu - x_s^\mu$, for convenience we will denote this as $\boldsymbol{r^i}$ to indicate the (un-normalized) vector between the two spacecraft (since the non-spatial components of $h_{\mu\nu}$ are zero, it collapses the 4-vector $\dot{x}_0^\mu$ to the 3-vector $\boldsymbol{r^i}$). This can be used to simplify the expression further, leading to:

$$\begin{aligned} L &= L_0 + \frac{L_0}{2}\int_0^1 h_{ij}(\lambda)\frac{\boldsymbol{r^i r^j}}{L_0^2}\mathrm{d}\lambda, \\ &= L_0 + \frac{L_0}{2}\int_0^1 h_{ij}(\lambda)\hat{\boldsymbol{r}}^{\boldsymbol{i}}\hat{\boldsymbol{r}}^{\boldsymbol{j}}\mathrm{d}\lambda, \end{aligned} \tag{3.35}$$

where we have defined $\hat{\boldsymbol{r}}^{\boldsymbol{i}} = \boldsymbol{r^i}/L_0$ as the unit vector between the two spacecrafts.

As shown in Chapter 1, the GW signal h_{ij} can be modelled as a transverse plane wave, travelling with a wavevector $\boldsymbol{k}$;

$$h_{ij}(t, \boldsymbol{x}) = H_{ij}e^{i(\omega t - \boldsymbol{k}\cdot\boldsymbol{x})}, \tag{3.36}$$

where H_{ij} and ω are respectively the amplitude and the angular frequency of the GW. Define the waveform phase as $\phi = \omega t(\lambda) - \boldsymbol{k}\cdot\boldsymbol{x}(\lambda)$, however since $\boldsymbol{k} = \omega\hat{\boldsymbol{k}}$ (in natural units), and ω is assumed to be constant, we define $\zeta = \phi/\omega = t(\lambda) - \hat{\boldsymbol{k}}\cdot\boldsymbol{x}(\lambda)$. Additionally, $t(\lambda) = \lambda(t_r - t_s) + t_s$, therefore;

$$\begin{aligned} \frac{\partial\zeta}{\partial\lambda} &= \frac{\partial t}{\partial\lambda} - \hat{\boldsymbol{k}}\cdot\frac{\partial\boldsymbol{x}}{\partial\lambda}, \qquad \frac{\partial t}{\partial\lambda} = t_r - t_s = L_0, \qquad \frac{\partial\boldsymbol{x}^\alpha}{\partial\lambda} = L_0\hat{r}^\alpha \\ &= L_0\big(1 - \hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}\big). \end{aligned} \tag{3.37}$$

A change of variables can be made within Eqn. 3.35, $\lambda \to \zeta$. A corresponding change of limits imposes the following bounds on the integral:

$$\lambda = 0 \rightarrow \zeta_{\lambda=0} = t_s - \hat{\boldsymbol{k}} \cdot \boldsymbol{x_s}, \tag{3.38}$$
$$\lambda = 1 \rightarrow \zeta_{\lambda=1} = t_r - \hat{\boldsymbol{k}} \cdot \boldsymbol{x_r}.$$

We denote the difference between the two limits as $\zeta_{\lambda=1} - \zeta_{\lambda=0} = L_0 - \hat{\boldsymbol{k}} \cdot \boldsymbol{r} = \Delta\zeta$, this will be useful later. Upon substituting the limits, the integral reduces to;

$$\begin{aligned} L &= L_0 + \frac{L_0}{2} \hat{\boldsymbol{r}}^{\boldsymbol{i}} \hat{\boldsymbol{r}}^{\boldsymbol{j}} \left[L_0 (1 - \hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{r}}) \right]^{-1} \int_{\zeta_{\lambda=0}}^{\zeta_{\lambda=1}} h_{ij}(\zeta) \mathrm{d}\zeta, \\ &= L_0 + \frac{1}{2} \frac{\hat{\boldsymbol{r}}^{\boldsymbol{i}} \hat{\boldsymbol{r}}^{\boldsymbol{j}}}{1 - \hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{r}}} \int_{\zeta_{\lambda=0}}^{\zeta_{\lambda=1}} h_{ij}(\zeta) \mathrm{d}\zeta. \end{aligned} \tag{3.39}$$

Generally the unit vector between the two spacecraft $\hat{\boldsymbol{r}}$ varies as a function of time as the spacecraft move throughout the mission on their own orbits. Incorporating this into the integral leads to the following leading order expression for the change in path length due to the passing GW:

$$\Delta L(t) = \frac{1}{2} \frac{\hat{\boldsymbol{r}}^{\boldsymbol{i}}(t) \hat{\boldsymbol{r}}^{\boldsymbol{j}}(t)}{1 - \hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{r}}(t)} \int_{\zeta_{\lambda=0}}^{\zeta_{\lambda=1}} h_{ij}(\zeta) \mathrm{d}\zeta. \tag{3.40}$$

Following on from the discussions in earlier sections of this chapter, we aim to calculate the response function of LISA directly in the frequency domain for computational convenience.

The Fourier and inverse Fourier transform of $h(\zeta)$ are defined as:

$$\tilde{h}_{ij}(f) = \int_{-\infty}^{\infty} h_{ij}(\zeta) e^{-2\pi i f \zeta} \mathrm{d}\zeta, \qquad h_{ij}(\zeta) = \int_{-\infty}^{\infty} \tilde{h}_{ij}(f) e^{2\pi i f \zeta} \mathrm{d}f.$$

The integral in the expression for $\Delta L(t)$ above can be expanded as:

$$\int_{\zeta_{\lambda=0}}^{\zeta_{\lambda=1}} h_{ij}(\zeta)\mathrm{d}\zeta = \int_{\zeta_{\lambda=0}}^{\zeta_{\lambda=1}} \int_{-\infty}^{\infty} \tilde{h}_{ij}(f) e^{2\pi i f\zeta}\mathrm{d}f\mathrm{d}\zeta, \tag{3.41}$$
$$= \int_{-\infty}^{\infty} \tilde{h}_{ij}(f)\left[\int_{\zeta_{\lambda=0}}^{\zeta_{\lambda=1}} e^{2\pi i f\zeta}\mathrm{d}\zeta\right]\mathrm{d}f.$$

The inner integral evaluates to:

$$\int_{\zeta_{\lambda=0}}^{\zeta_{\lambda=1}} e^{2\pi i f\zeta}\mathrm{d}\zeta = \left[\frac{e^{2\pi i f\zeta}}{2\pi i f}\right]_{\zeta_\lambda=1-\Delta\zeta}^{\zeta_{\lambda=1}}, \tag{3.42}$$
$$= \frac{e^{2\pi i f\zeta_{\lambda=1}}}{2\pi i f}[1 - e^{-2\pi i f\Delta\zeta}].$$

Therefore, the original equation (Eqn. 3.41) can be written as:

$$\int_{\zeta_{\lambda=0}}^{\zeta_{\lambda=1}} h_{ij}(\zeta)\mathrm{d}\zeta = \int_{-\infty}^{\infty} \frac{\tilde{h}_{ij}(f)}{2\pi i f} e^{2\pi i f\zeta}[1 - e^{-2\pi i f\Delta\zeta}]\mathrm{d}f. \tag{3.43}$$

Substituting this into the expression for $\Delta L(t)$ leads to:

$$\frac{\Delta L(t)}{L_0} = \frac{1}{2L_0}\frac{\hat{\boldsymbol{r}}^i(t)\hat{\boldsymbol{r}}^j(t)}{1 - \hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}(t)}\int_{-\infty}^{\infty} \frac{\tilde{h}_{ij}(f)}{2\pi i f} e^{2\pi i f\zeta}[1 - e^{-2\pi i f\Delta\zeta}]\mathrm{d}f. \tag{3.44}$$

This can be rewritten in the following form:

$$\frac{\Delta L(t)}{L_0} = \frac{\hat{\boldsymbol{r}}^i(t)\hat{\boldsymbol{r}}^j(t)}{2}\int_{-\infty}^{\infty} \tilde{h}_{ij}(f)\tau(f,t,\boldsymbol{k})e^{2\pi i f\zeta}\mathrm{d}f, \tag{3.45}$$

where $\tau(f,t,\boldsymbol{k})$ is a transfer function operating on the Fourier domain waveform. This result describes the strain on one laser arms of the LISA detector as a function of time. The integral in the above expression acts as an inverse Fourier transform to bring the frequency domain detector strain back into the time domain.

Substituting the limits explicitly for ζ in terms of $\boldsymbol{k}, \hat{\boldsymbol{r}}$ and L_0:

$$\begin{aligned}\tau(f,t,\boldsymbol{k}) &= \frac{1-e^{-2\pi i f L_0(1-\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}(t))}}{2\pi i f L_0(1-\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}(t))}, \\ &= \frac{1}{2i}\frac{[e^{i\pi f L_0[1-\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}(t)]}-e^{-i\pi f L_0[1-\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}(t)]}]}{\pi f L_0[1-\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}(t)]}e^{-i\pi f L_0[1-\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}(t)]}, \\ &= \mathrm{sinc}(\pi f L_0[1-\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}(t)])e^{-i\pi f L_0[1-\hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{r}}(t)]}.\end{aligned} \tag{3.46}$$

Note that the sinc in the above expression is defined as $\mathrm{sinc}(x) = \frac{\sin(x)}{x}$. This matches the expression for the instrument response presented in Refs. [138, 139] among others. Note that in some cases the prefactor to the integral which contains the terms modelling the orbit of the spacecraft are also included in the transfer function.

While our aim was to derive a response function that is purely a function of frequency, unfortunately we have ended up with parts of the response, namely $\hat{\boldsymbol{r}}^{\boldsymbol{i}}(t)$ which are time dependent. Practically, applying this response usually involves the computation of a time-frequency map, like that obtained via the SPA for binaries in their inspiral phase.

The response derived here is a leading order 'local' response, it is valid for scenarios in which the modulations of the GW signal arising within the response are slowly varying, and also assumes $\tilde{h}(f)$ is varying slowly. We will demonstrate that our derived Fourier domain response is a leading order approximation in what follows, this derivation is detailed in the Sec III of Ref. [138].

It is assumed that the waveform in the time domain is $h(t)$ and that there is a general time dependent delay $d(t)$ applied to the waveform, such that $h_d(t) = h(t + d(t))$. The time dependent modulation caused by the response is given by $F(t)$ such that the response of the detector to the delayed GW in the time domain is $s(t) = F(t)h_d(t)$ (this is the most general form of the response and encapsulates any arbitrary time-dependent modulation and delay). The aim is to find a convenient form of calculating $\tilde{s}(f) = \int s(t)e^{-2\pi i f t}\mathrm{d}t$, this can be expanded as

$$\begin{aligned}\tilde{s}(f) &= \int \mathrm{d}t e^{-2\pi i f t} F(t) h(t + d(t)), \\ &= \int \mathrm{d}t e^{-2\pi i f t} F(t) \int \mathrm{d}f' e^{2\pi f'(t+d(t))} \tilde{h}(f'), \\ &= \int \mathrm{d}f' \tilde{h}(f') \int \mathrm{d}t e^{-2\pi i f t} F(t) e^{2\pi f'(t+d(t))}.\end{aligned} \tag{3.47}$$

With a substitution of variables as $f'' = f - f'$, this expression can be rearranged to:

$$\begin{aligned}\tilde{s}(f) &= -\int df'' \tilde{h}(f - f'') \int \mathrm{d}t e^{-2\pi i f'' t} F(t) e^{2\pi i (f - f'') d(t)}, \\ &= -\int df'' \tilde{h}(f - f'') \tilde{G}(f - f'', f'').\end{aligned} \tag{3.48}$$

Which is a generalised convolution in the Fourier domain between the waveform and a response kernel G, which is defined as (the arrow corresponds to Fourier transforming the time-domain response kernel):

$$G(f, t) = e^{2\pi i f d(t)} F(t) \to \tilde{G}(f, f'') = \int \mathrm{d}t e^{-2\pi i f'' t} e^{2\pi i f d(t)} F(t). \tag{3.49}$$

This convolution is at the heart of why computing the full response in the Fourier domain is expensive, while in the time-domain the response is simply a multiplication therefore it is cheap to compute. Physically, $F(t)$ and $d(t)$ respectively represent the amplitude modulations and delays incorporated into the response, since these are both slowly varying quantities, typically varying on the timescale of $\sim$ weeks or slower, the convolution kernel $\tilde{G}(f, f'')$ is compact, limited to a maximum frequency corresponding to the timescale of these modulations. Here a local approximation in frequency is made, since f'' is small for significant support in $\tilde{G}$. Since f'' is small, $\tilde{h}(f - f'')$ can be safely Taylor expanded around f to obtain (only expanding phase terms, since amplitude evolves slower than phase in the inspiral):

$$\psi(f - f'') \approx \psi(f) - \left.\frac{\partial\psi}{\partial f}\right|_f (f - f'' - f) = \psi(f) - f''\left.\frac{\partial\psi}{\partial f}\right|_f, \tag{3.50}$$

$$\tilde{h}(f - f'') \approx A(f)\exp\left\{-i\left[\psi(f) - f''\left.\frac{\partial\psi}{\partial f}\right|_f\right]\right\}. \tag{3.51}$$

Here, $\psi(f)$ and $A(f)$ respectively represent the phase and amplitude of the gravitational wave, in the form $h(f) = A(f)\exp(i\psi(f))$. Inserting this local approximation for the waveform into the expression for $\tilde{s}(f)$ leads to (the kernel G is left at 0 order and not expanded further):

$$\tilde{s}(f) = \tilde{h}(f)\int \mathrm{d}f''\tilde{G}(f, f'')\exp\left[if''\left.\frac{\partial\psi}{\partial f}\right|_f\right], \tag{3.52}$$

Recognizing that this integral can be rewritten as an inverse Fourier transform for $\tilde{G}(f, f'')$, we arrive at the expression;

$$\tilde{s}(f) = \tilde{h}(f)G\left(f, \frac{1}{2\pi}\left.\frac{\partial\psi}{\partial f}\right|_f\right). \tag{3.53}$$

Note the similarity between the above expression and Eqn. 3.45, both expressions do not contain a convolution and instead are evaluated by multiplication of the Fourier domain waveform with a response transfer function evaluated at a given time. Notice the derivative of the phase in analogy to the SPA. An intuitive way of understanding the process in going from Eqn. 3.48 to 3.53 is the following: $\tilde{G}(f - f'', f'')$ is compact in f'' due to the expected long timescale of modulations $G(f - f'', f'') \to \tilde{G}(f, f'')$, as a result $\tilde{h}(f - f'') \to \tilde{h}(f)$, in the case of inspiral signals, $\tilde{h}(f)$ can be related to $h(t)$ via the SPA. As a result, the response collapses to a simple multiplication of G at the SPA-time.

Going beyond the leading order local response that is presented here involves the inclusion of higher order derivatives in the Taylor expansion of both the waveform and the response kernel G, this is detailed in Ref. [138].

3.3 Time-Delay Interferometry

In the current ground-based detectors, Michelson Interferometry is used to detect variations in the optical path length of lasers. The Michelson interferometer measures optical path differences by splitting a laser beam into two paths, reflecting and recombining the beams, causing the two returning beams to interfere. In the case that the two split beams travelled exactly the same optical path distance, they would interfere constructively, however if the two beams had travelled different distances they can interfere destructively (depending on the difference in the optical path lengths). The variation of the interference pattern as a function of time is then used to infer the variation in the path length, which can in-turn be linked to the GW passing through the detector.

Arm lengths in the LVK detectors are equal, one of the main benefits of this configuration is the cancellation of laser noise. In practice, lasers are not able to constantly produce electromagnetic radiation at a single wavelength (limited coherence length), there are occasional fluctuations in the frequency of the laser light. In the case arm lengths are equal this does not cause a problem; although the instantaneous frequency of the laser may vary, the accumulated phase in each of the arms is the same since the arm lengths are equal. However since LISA will not have equal arm lengths, this raises an issue as the accumulated phase arising from the laser frequency fluctuations will be different in each arm. These fluctuations may be mistaken as a passing GW signal which is causing changes in the path length. Time delay interferometry (TDI) is a solution to this problem [140].

In essence, TDI is a method of combining multiple interferometer channels to create a new 'virtual' channel which is free of the effects of laser frequency noise when the arm lengths are unequal. We will demonstrate this with a simple two arm example [141], illustrated in Fig. 3.3. For this example, assume that on recombination the two beams are interfered with the outgoing laser with phase ϕ. The respective phases at recombination for ϕ_1 and ϕ_2 in the two arms are given by:

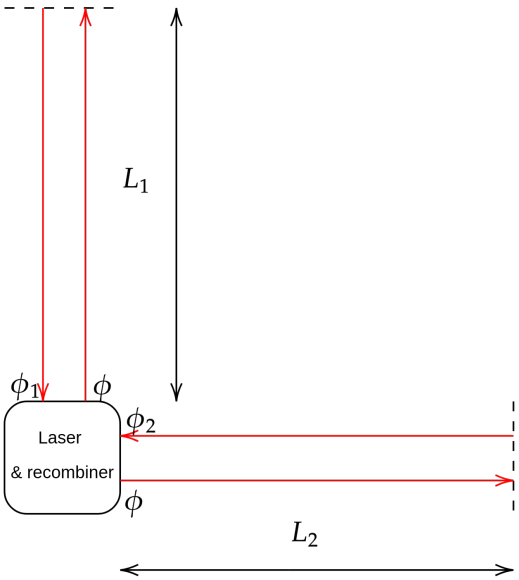

Figure 3.3: A Michelson interferometer with two arms, the laser light in both arms is assumed to be from the same source. Red lines indicate the laser optical paths, and the dashed black lines represent the mirrors. The arm lengths L_1 and L_2 are not generally assumed to be equal. The outgoing laser in both arms has phase ϕ and the incoming lasers beams respectively have phases ϕ_1 and ϕ_2.

$$\phi_1(t) = \phi(t - 2L_1) - \phi(t), \qquad \phi_2(t) = \phi(t - 2L_2) - \phi(t) \tag{3.54}$$

Notice that in the equal arm case, that $\phi_1 - \phi_2 = 0$ and we return to the default Michelson interferometer. However when $L_1 \neq L_2$, the recombined phases do not cancel. Thus any transient laser frequency fluctuations in the unequal arm case can cause a phase difference in the recombined laser light which can be mistaken as a passing GW signal. In order to mitigate this, a new 'virtual' interferometer (TDI) channel is created by combining the two channels and incorporating further delays in postprocessing:

$$X(t) = [\phi_2(t - 2L_1) - \phi_2(t)] + [\phi_1(t - 2L_2) - \phi_1(t)]. \tag{3.55}$$

Substituting the expressions for ϕ_1 and ϕ_2 into this leads to a cancellation of the phase in the unequal arm length case. The TDI channel we have constructed is analogous to the 'first generation' TDI channels, which assume that the arm-lengths are unequal but static, however it is well known that the arm lengths between each of the satellites will

not be static, but will rather slowly vary over the period of a year. In this case the TDI channel we have constructed would not be sufficient to suppress the laser frequency noise as it relies on the commutativity of delays; expanding Eqn. 3.55 leads to $X(t) \sim \phi(t - 2L_1 - 2L_2) - \phi(t - 2L_2 - 2L_1)$, when the arm lengths are unequal and varying as a function of time, the two terms do not cancel. TDI variables accounting for (slowly) varying arm lengths are known as 'second generation' variables [142][10].

The Michelson TDI variables for LISA are labelled X, Y, Z, however these are not the channels in which data analysis is conventionally conducted, in this form the channels are not statistically independent. Thus any analyses over these channels will have to consider the correlations between them, increasing computational cost; instead, a linear combination of the TDI channels are used to form the A, E, T variables. This change of basis orthogonalises the noise covariance matrix for X, Y, Z, allowing for the analysis of each channel independently [141]. This transformation is given by:

$$\begin{pmatrix} A \\ E \\ T \end{pmatrix} = \begin{pmatrix} \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}. \tag{3.56}$$

The transformed variables A, E, T are only orthogonal for equal arm length orbits and assumes the noise level in each arm is identical [144].

3.4 Time-frequency analysis

In this section we briefly discuss the advent of time-frequency methods in GW data analysis and motivate their use for the analysis of certain LISA sources. As previously discussed LISA observations of a single source are long-lived, over this time the sensitivity of the detector will likely vary significantly, therefore all the machinery detailed within this chapter which assumes a detector with stationary noise is rendered invalid. The

[10]TDI generation 1.5 is an intermediate TDI generation between 1 and 2 in that it accounts for the rotation of the satellite constellation but assumes the arm lengths are not time-varying [143].

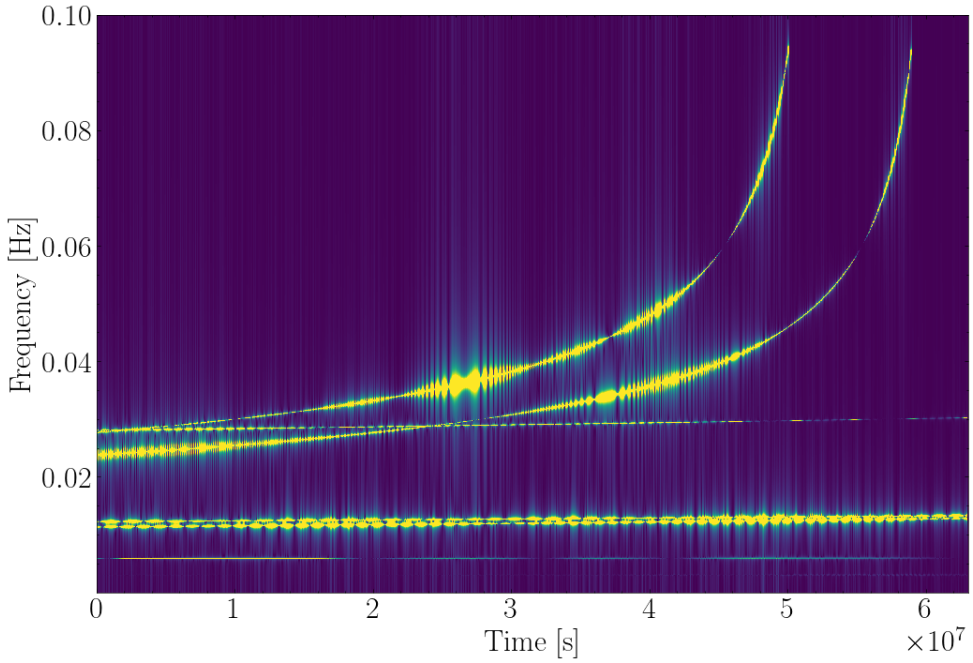

Figure 3.4: Time-frequency representation of a set of stellar-mass binary (noiseless) signals. These are the signals present in the Yorsh dataset used in Chapter. 6.

time-frequency approach to GW analysis is a method for dealing with this problem.

While the noise spectrum of the detector is not expected to be stationary throughout the whole mission, it is expected to remain stationary over a short timescale. Although there are expected to be a range of non-stationary contributions to the noise spectrum, it is expected that the shortest of these will vary (conservatively) on the timescale of $\sim$ 10 hours to days [145, 146]. Therefore, the whole LISA (non-stationary) datastream can in principle be divided up into smaller chunks of data with stationary noise, and each of these chunks can be analysed using the methods suited for datastreams with stationary noise spectra. This is equivalent to computing many short Fourier transforms (SFTs) to create a spectrogram of the data, as demonstrated in Fig. 3.4. Slowly varying noise spectrums can be dealt with within this framework by assigning a different power spectral density in each SFT (column of Fig. 3.4). Additionally, dealing with gaps within the time-frequency framework is also simple. Since the LISA datastream is expected to contain both scheduled and unscheduled gaps [134], data analysis frameworks need to be able to handle gapped data. In the time-frequency framework, SFTs (i.e columns in the

spectrogram) which contain gaps can be (crudely) ignored, this will work well for data gaps which are of a similar size in time as that of the SFT chunk (for much shorter gaps ignoring the whole SFT segment unnecessarily discards a lot of data).

An additional benefit of this approach is the improvement in computational cost when dealing with sources that are long lived and slowly chirping. The source types that are expected to benefit most from this approach are the stellar-mass binary inspiral and EMRI signals. These sources are broadband and are also likely to be in the LISA frequency band for close to the whole mission lifetime (i.e these sources are neither compact in time or frequency), using a fast-Fourier transform grid to represent such a signal would require $\mathcal{O}(10^7)$ points in the frequency domain, this leads to a very large computational cost for the analysis of these signals (see Appendix A.4). The cost is mainly driven by the evaluation of both the waveform and response on such a large grid. The time-frequency framework offers a solution to this problem; since these signals are compact in the time-frequency domain (see single source tracks in Fig. 3.4) the waveform and response can both be evaluated only on the pixels along the time-frequency track of the source, reducing the grid to a dimensionality of $\mathcal{O}(10^4)$. Note that these benefits are unique to the slowly evolving broadband sources, other sources such as the merger of massive black hole binaries are not expected to be compact in the time-frequency domain, thus waveform and response evaluations cannot be accelerated for this source type. Fig. 3.4 also shows that the time-frequency tracks of these sources do not perfectly occupy one frequency bin for each SFT, but are rather spread out over a small set of frequency bins. This can be caused by two main reasons; the first is that the SFT chunk size (in time) is too large and thus the GW frequency of the source has evolved significantly within the chunk (this is the case for the signals which reach the high frequencies), or these are the effects of spectral leakage.

The efficacy of the time-frequency framework in the presence of non-stationary noise has been demonstrated for stellar-mass binary parameter estimation in Ref. [147]. This framework has also been shown to be effective for the analysis of the cyclo-stationary galactic foreground in Ref. [56]. Further computational acceleration is possible within

this framework in the form of a semi-analytical approximation to the noise-weighted inner product which is particularly effective for long inspirals such as those discussed in this thesis [148].

4. A proof-of-concept semi-coherent search

Overview

This chapter is an edited version of [149]:

Diganta Bandopadhyay and Christopher J. Moore. LISA stellar-mass black hole searches with semi-coherent and particle-swarm methods, published in Physical Review D, Volume 108(8):084014, October 2023.

At the time of publication this was the first paper which trialed the use of a semi-coherent strategy in searching for stellar-mass inspiral signals within simulated LISA data. A semi-coherent variant of the Whittle likelihood is motivated as a potential objective function for the search. The properties of this likelihood are tested in an LVK context for the analysis of GW170817, and then applied to a (noiseless) search for simulated stellar mass binary signals in LISA data. The majority of this study uses stochastic sampling to sample the widened posterior, but a particle swarm optimization (PSO) algorithm is briefly introduced and demonstrated as a viable alternative to sampling.

I wrote the code used for the analysis in this study, produced all the figures, and wrote the large majority of the text, with advice from my supervisor and co-author Christopher J. Moore.

4.1 Introduction

Among the sources that LISA will observe, stellar-mass binary black hole (SmBBH) [150, 64] and extreme-mass-ratio inspiral (EMRI) [84, 151, 79] sources are of particular interest

in this study. SmBBHs will eventually merge in the LIGO/Virgo [152, 153] frequency band; events similar to GW150914 [1] and GW190521 [154] would have appeared as quiet, long-lived LISA sources, had the instrument been operating several years previously.

SmBBH systems are observed by LISA relatively early in their inspiral, when the orbital separation is much larger than the Schwarzschild radius of either BH. At this stage in the inspiral the orbital velocity is small, $v \ll c$, and these systems are weak sources of GWs with the GW frequency evolving slowly (i.e. approximately adiabatically) and the source completing many orbits in the LISA frequency band. EMRI signals meanwhile are observed late in their inspiral, when the orbital separation is comparable to the radius of the larger BH. These are highly relativistic sources with $v \lesssim c$. However, the extreme mass-ratio of these systems means that they are also weak sources of GWs with slowly evolving frequencies. Again, this leads to a large number of orbits being completed in the LISA band. Although SmBBHs and EMRIs are physically very different, both will appear as long-lived, broadband signals in LISA with $\gtrsim 10^5$ observable GW cycles. From a data analysis perspective, the main difference between the two source types is that SmBBH signals are dominated by a single-frequency harmonic, whereas EMRI signals may have significant contributions from many harmonics. Both sources promise exciting new possibilities for multimessenger [65, 64] and fundamental physics [155].

The search phase will be extremely challenging for SmBBH and EMRI signals due to the size of the parameter space that must be explored (see discussion in Sec. 3.1.2). Ideas for EMRI search strategies have been investigated in Refs. [156, 157, 158, 159, 160]. Once a search identifies a candidate detection, the parameter estimation phase is tackled using a well-established Bayesian framework that maps out the posterior distribution on the waveform parameters. Parameter estimation for both EMRI [161, 82] and SmBBH [65, 162, 147, 163, 164] signals has been previously demonstrated. The holy grail of source characterization for LISA is the global fit which aims to simultaneously estimate the source parameters of all the signals observed by LISA. The global fit is made tractable via the prior identification of regions in parameter space where signals might exist; this

prior identification is the role of the search phase. This primary search phase is an open problem for SmBBH and EMRI signals [103, 88] and is the subject of this chapter.

Long-lived signals, such as SmBBH and EMRIs, undergo a large number of orbits. This allows certain parameters that control the GW frequency (notably the binary chirp mass, $\mathcal{M}_c$) to be measured with unprecedented precision. Among the current GW detections, the closest analog to an SmBBH or EMRI signal is the binary neutron star (BNS) GW170817 [121]. This low-mass ($\mathcal{M}_c \sim 1\,M_\odot$) source completed ~ 3000 cycles in the detector frequency band, compared to just ~ 10 for the high-mass ($\mathcal{M}_c \sim 30\,M_\odot$) GW150914 binary BH [1]. The longer signal translates to a more precise measurement of the chirp mass; for GW170817 the fractional error was $\delta\mathcal{M}_c/\mathcal{M}_c \sim 10^{-3}$ [121] whereas for the much shorter GW150914 it was $\sim 10^{-1}$. In contrast, for the extremely long SmBBH and EMRI systems observed in LISA, the fractional error on the chirp mass is expected to be several orders of magnitude smaller, depending on the source parameters; for example, in the case of a GW190521-like source observable by LISA, the fractional uncertainty is predicted to be 10^{-5} [164, 65]. The precision of these measurements drives the requirements for the search; for systems where we can measure the source parameters with greater precision, the search must cover the parameter space with a correspondingly finer resolution. SmBBH and EMRI signals in LISA represent a completely new challenge, orders of magnitude more difficult than those encountered to date in GW astronomy. This calls for completely new analysis tools and methods.

Searches for compact binary coalescences in LIGO-Virgo data have been successfully conducted using template banks since the very first detection [165]. As discussed in Sec. 3.1.2, this approach to the search becomes computationally intractable as the signals being searched for become longer.

It is interesting to note that, if one was prepared to wait several years, one could rely on some future-generation ground-based detector to observe the final merger of the SmBBH systems. This could be used as a trigger to go back and perform parameter estimation on the archival LISA data without the need for a full search. Such *archival searches*

have been demonstrated for quiet SmBBH mock LISA signals [166]. It is worth noting that even in the archival targeted search scenario, localizing a signal in the constrained parameter space can still be challenging [167]. We do not want to rely solely on archival searches for several reasons: ground-based detectors will not operate with a 100% duty cycle and will therefore miss a fraction of events; the prospect of an advanced warning of a GW event alongside a sky localization can be invaluable for multimessenger astronomy [65, 64, 168]; and archival searches will not be possible at all for EMRIs.

Several approaches to the SmBBH and EMRI search problem have been proposed, although none are fully developed. One family of approaches involves splitting the data into multiple time or frequency segments and searching each individually. It is not necessarily expected to be possible to confidently detect a signal in a single segment, but by suitably combining the results of searches across segments a detection can be achieved. This type of method can be described as incoherent, or semi-coherent, because the model used is not required to accurately describe the signal phase evolution across the entire observation [88]. Semi-coherent methods relax the stringent requirements on the phase accuracy of the models; therefore, another attractive aspect of semi-coherent methods is the prospect of being able to use a simpler, computationally cheaper waveform (e.g. a lower-dimensional model, perhaps neglecting some of the physics) for the search. Semi-coherent methods are already used in searches for continuous GWs in LIGO/Virgo data [169]. Another approach, specific to EMRIs, is harmonic matching, where several discrete frequency harmonics of the signal are first identified individually before being later combined into a single detection (see Fig. 5.8 of Ref. [170] and Ref. [171]). In practice, this is challenging as individual harmonics are quieter than the full signal and are therefore harder to disentangle from instrumental noise and the numerous other overlapping sources. In principle, semi-coherent and harmonic matching techniques can be used in combination. We also note the existence of machine-learning based approaches to the search problem; Ref. [172] demonstrated the detection of high signal-to-noise ratio (SNR) EMRIs using convolutional neural networks, but without the ability to provide information on the source parameters.

This study focuses on the use of a semi-coherent approach, in combination with a particle swarm optimization (PSO) algorithm to make progress towards a realistic search algorithm for SmBBH signals in LISA. PSO is a stochastic optimization algorithm (see, e.g. Refs. [173, 174]), variants of which can be tailored to be well-suited to the identification of multiple, widely separated peaks in the likelihood surface [175, 176]. It is our hope that this property will also make it suitable for EMRI searches (this will be explored in future work). PSO methods have previously been used in a LISA context for galactic double white dwarf binaries [177, 178]. We show that PSO can successfully locate the source parameters for an SmBBH signal when coupled with a hierarchical approach, iteratively exploring semi-coherent likelihoods with a decreasing number of segments. The work done in this study is still in the early prototyping stage for this search strategy; due to this, several simplifying assumptions have been made in the analyses in Secs.4.4 and 4.5. These include the use of a simulated data-stream without the presence of noise (i.e. working with zero-noise injections) and without other confusing sources such as galactic binaries. We also assume the LISA instrumental noise will be stationary and use a constant power spectral density (PSD), neglecting the expected cyclo-stationary (noise with statistical properties which vary periodically in time)[1] and non-stationary components. Our data-stream is also devoid of gaps and glitches, which will exist in the real LISA data-stream.

In this chapter the `TaylorF2` waveform model is used for both injection and recovery of the SmBBH signals. This model of a SmBBH inspiral can be described using 11 parameters: chirp mass $\mathcal{M}_c$, mass ratio q, time to coalescence t_c (from start of LISA mission), initial orbital phase ϕ, luminosity distance of the source D_L, sky position of the source in ecliptic coordinates (λ, β), polarisation of the gravitational wave ψ, inclination of the orbital plane with respect to the line of sight of the observer i and the magnitude of the spin vectors of the two binaries χ_1 and χ_2 which are aligned with the orbital angular momentum[2] Although, note that we switch to a different parameterisation for the analysis

[1]This corresponds to the contribution from the double-white-dwarf galactic foreground, which is expected to be a stochastic source of gravitational waves, with an amplitude that is modulated with a 1 year period, set by the orbital period of the LISA instrument, see discussion in Sec. 2.2.

[2]In general each black hole can have a 3D spin vector that is misaligned with the orbital angular

in this chapter in order to make the posteriors more Gaussian-like and easier to sample, this parameterisation is detailed in Sec. 4.4.

The semi-coherent methods that are used in this study are introduced in Sec. 4.2. Sec. 4.3 explores the properties of the semi-coherent likelihood by using it to reanalyze the GW170817 BNS event. Sec. 4.4 explores the properties of the semi-coherent likelihood for SmBBH sources. In Sec. 4.5 we introduce PSO as a search method which is able to locate the source parameters for a SmBBH signal. Sec. 6.6 discusses the further work required to develop this into a full search and possible extensions of this method to explore the extremely multi-modal likelihood surfaces expected from EMRI signals. Throughout this chapter we work in natural units where $G = c = 1$.

4.2 Semi-Coherent methods

In this section we describe the semi-coherent data analysis methods used in this study and contrast them with the conventional, fully-coherent analysis more commonly used in GW astronomy.

Parameter estimation uses the following log-likelihood which constructed from the noise weighted inner product (Eqn. 3.8) and is explored as a function of the model parameters;

$$\begin{aligned} \log L(d|\theta) &= -\frac{1}{2}\langle d-h|d-h\rangle + c \\ &= -\frac{1}{2}\langle d|d\rangle - \frac{1}{2}\langle h|h\rangle + \langle d|h\rangle + c. \end{aligned} \tag{4.1}$$

The number of free parameters is the dimensionality of the parameter vector, $\dim(\theta)$. The log-normalization c does not depend on θ. On the second line, $\log L$ is split into three terms: $\langle d|d\rangle$ is constant (in that is doesn't depend on θ) and can be neglected; $\langle h|h\rangle = \rho^2$ is the optimal squared signal-to-noise ratio (SNR) and is approximately constant over small regions of parameter space. Therefore, $\langle h|d\rangle$ is the key quantity that controls the

momentum, however in this chapter we limit the analysis to aligned spin systems.

shape of the likelihood surface.

For most calculations in this chapter it will be assumed that the signal contains a single mode, by which we mean the model can be decomposed into amplitude and phase as $\tilde{h}(f) = A(f)e^{i\Phi(f)}$ where one of the waveform parameters is an orbital phase angle ϕ which enters as $\Phi(f) \to \Phi(f) + \phi$ (the angle ϕ is one component of the parameter vector θ). If this is the case, then it is possible to analytically maximize the noise weighted inner product $\langle h|d\rangle$ term with respect to ϕ. We define the overlap as this phase-maximized inner product;

$$\begin{aligned}\mathcal{O}(d,h) \equiv &\max_{\phi} \left\langle d\middle|he^{i\phi}\right\rangle \\ =&4\left|\sum_{\alpha}\int_{f_{\min}}^{f_{\max}} \frac{d_\alpha(f)h^\dagger_\alpha(f)}{S_\alpha(f)}\mathrm{d}f\right|. \end{aligned} \tag{4.2}$$

Note, the overlap is simply the magnitude of a complex inner product. An intuitive way to understand the form of the above equation is that the absolute operator is simply estimating and applying a constant phase shift to the model that maximizes the real part of the expression. If the model contains multiple modes, the maximization with respect to ϕ must be done numerically. The coherent overlap in Eqn. 4.2 is a function of $\dim(\theta) - 1$ free parameters, not including the phase angle ϕ.

While maximizing over the phase angle will affect the shape of the likelihood (and hence the posterior distribution) in the other parameters and is less desirable than marginalizing over it, the difference is expected to be small, especially for long signals such as SmBBHs and EMRIs. This is demonstrated explicitly in Sec. 4.3 for the BNS signal GW170817 and is also found to be the case in Sec. 4.4 for SmBBH signals. We have also verified that the phase maximized likelihood evaluated with a template which has a random value of initial orbital phase is numerically equivalent to the coherent likelihood evaluated with the *true* injected phase.

We now proceed to split the inner product into N segments. The segmented inner product between two sets of time series a_α and b_α is calculated in terms of the N separate

frequency integrals,

$$[a|b]_n^N = \sum_\alpha 4\mathrm{Re}\int_{f_n}^{f_{n+1}} \frac{\tilde{a}_\alpha(f)\tilde{b}_\alpha^\dagger(f)}{S_\alpha(f)}\mathrm{d}f. \tag{4.3}$$

We emphasize that we are segmenting our data in the frequency domain; each segment involves data taken at all times. This is to be contrasted with what was envisaged in, for example, Fig. 1 of Ref. [88], where the data was segmented in time. For slowly inspiraling sources such as SmBBHs which are well approximated by a stationary phase approximation the two approaches are equivalent. Here the $f_0 = f_{\rm low}$, $f_N = f_{\rm high}$, and the intermediate frequency boundaries f_n are ordered as $f_n < f_{n+1}$. For now the frequency boundaries are only required to be ordered, and we will discuss a method for selecting these later in this section. Since there is no phase maximization incorporated into this inner product yet, the sum across all segments is equal to the standard noise-weighted inner product,

$$\sum_{n=0}^{N-1}[a|b]_n^N = \langle a|b\rangle\,. \tag{4.4}$$

So far we have not actually done anything except (arbitrarily) splitting the integral in Eqn. 3.8 into a number of sub integrals. However, we now generalize by allowing the waveform model to be different in each segment. The most extreme approach is to allow all of the waveform model parameters to differ in every segment; in the frequency range $f_n < f < f_{n+1}$, the waveform model is $h(\theta_n)$. In this case the total number of model parameters is now $N\mathrm{dim}(\theta)$. Note that the waveform model is discontinuous at the segment boundaries. Combining the phase-maximized and segmented inner products, we define a semi-coherent overlap with N segments. The phase-maximized inner products for each segment follows the same form as Eqn. 4.2 truncated at the appropriate frequency boundaries for that segment:

$$\hat{\mathcal{O}}_N(d,h) = \sum_{n=0}^{N-1}\max_{\phi_n}[d|he^{i\phi_n}]_n^N \tag{4.5}$$

Therefore, there are now N phase parameters in our model, one per segment. We maximize over all of these phase parameters independently, each of which individually are unphysical, a combination of which correspond to the orbital phase ϕ. The semi-coherent overlap in Eqn. 4.5 is a function of the $N(\dim(\theta) - 1)$ free parameters $\{\theta_0, \theta_1, \ldots, \theta_{N-1}\}$, not including the phase angles ϕ_n.

The purpose of introducing all the additional parameters θ_n is to make the model less sensitive to any of the parameters individually. For example, in a coherent analysis a small change in, say, the chirp mass parameter, $\mathcal{M}_c$, may be enough to alter the phase evolution of the signal and cause the coherent overlap to drop significantly, $\mathcal{O} \sim 0$, but the same small change in the chirp mass in just the first segment, $\mathcal{M}_{c,0}$, will have a much smaller effect on the semi-coherent overlap, $\hat{\mathcal{O}}_N \sim 1$. It is intended that this drop in required precision will make it easier to perform the initial search. It may also have the benefit of allowing for the use of less accurate waveform models.

These benefits come at a cost; the total number of free parameters is increased. This increase in the flexibility of the model necessarily leads to a decrease in sensitivity. This flexible model is more likely to be able to fit well a signal containing only noise. For this reason, a search algorithm using the semi-coherent likelihood has an increased false alarm rate (compared to a similar search using the normal, coherent likelihood) and must therefore raise the detection threshold accordingly [179, 88]. In the context of continuous GW searches with ground-based interferometers, it is known that in the limit of a large number of segments ($N \to \infty$) the sensitivity of an idealized semi-coherent search loses sensitivity and is lower than that of an idealized coherent search by a factor $\propto N^{\frac{1}{4}}$ [180]. One possible approach would be to lower the threshold sensitivity in the early stages of the search, accepting a larger number of false positives, which are then followed up and can be vetoed in later stages of the search. This work is intended to be a proof-of-concept for the method and thus we use a noiseless datastream, the effect of noise on this search is discussed in Chapter 5.

Note that in the case of $N = 1$ segments, the semi-coherent overlap recovers the

standard, coherent phase-maximized result;

$$\hat{\mathcal{O}}_{N=1}(a,b) = \mathcal{O}(a,b). \tag{4.6}$$

Swapping the semi-coherent overlap into the expression for the log-likelihood in Eqn. 4.1, we reach our definition of the semi-coherent log-likelihood;

$$\begin{aligned}\log\hat{L}_N(d|\theta_0,\theta_1,\ldots,\theta_{N-1}) = &-\frac{1}{2}\langle d|d\rangle - \frac{1}{2}\langle h|h\rangle \\ &+ \hat{\mathcal{O}}_N(d,h).\end{aligned} \tag{4.7}$$

Note $\langle h|h\rangle$ remains a standard inner product because (for a single mode waveform) $\langle h|h\rangle$ does not depend on ϕ. This will not be exactly true for a signal that contains many modes, such as an EMRI.

What has been described so far is an extreme approach to a semi-coherent analysis where all the model parameters are allowed to vary between segments. This may not be necessary; it is typically the phase evolution that is most important as the overlap is most sensitive to this. For a signal with many frequency modes such as an EMRI, one might imagine a search keeping the majority of the parameters (e.g. sky position, distance, BH masses and spin, and the orbital shape parameters p, e and ι; see, for example, Ref. [181]) constant between segments, while introducing extra phase angles in each segment that can be maximized over. As we are concerned here with single-mode waveforms, for the remainder of this chapter we restrict to the case where only a single orbital phase parameter ϕ_n is allowed to vary between segments, and these are maximized over analytically as in Eqn. 4.2. Therefore, our semi-coherent likelihood becomes

$$\log\hat{L}_N(d|\theta) = -\frac{1}{2}\langle d|d\rangle - \frac{1}{2}\langle h|h\rangle + \hat{\mathcal{O}}_N(d,h), \tag{4.8}$$

which is a function of just $\dim(\theta) - 1$ parameters. Note, that in the case of $N = 1$

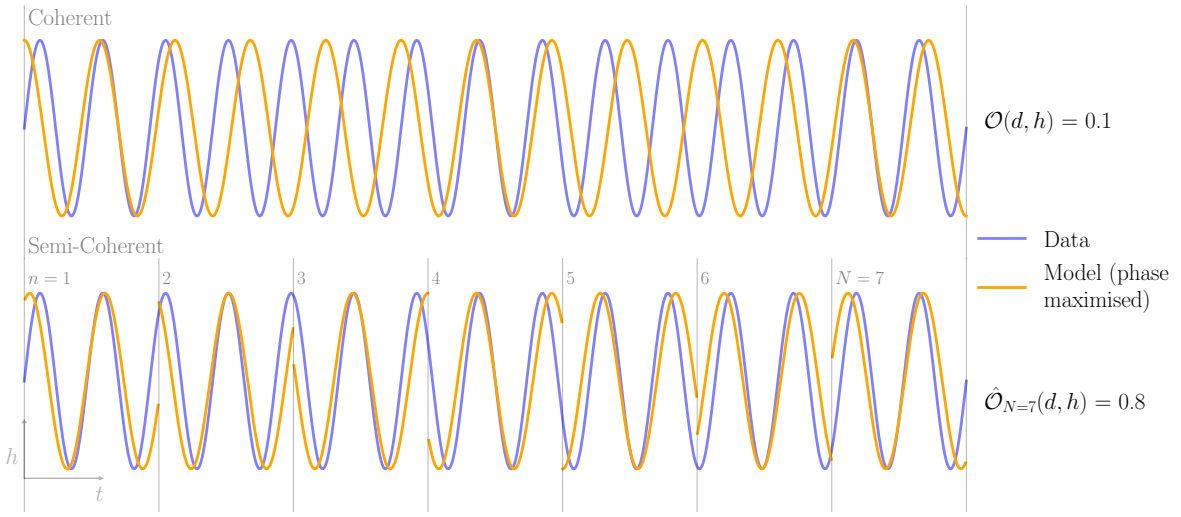

Figure 4.1: A sketch illustrating the semi-coherent method. For simplicity, time-domain sinusoidal signals are used; the data (purple) $d = \sin(2\pi f' t)$, where $f' T_{\rm obs} = 15$, and the model (orange) $h = \sin(2\pi f t)$, where $(f - f')T_{\rm obs} = 2.5$. *Top panel:* a coherent analysis; the model is compared against the data, trying to coherently maximize the overlap $\mathcal{O}(d,h)$ across the entire observation period, $T_{\rm obs}$. *Bottom panel:* an $N = 7$ segment semi-coherent analysis; the model phase is varied in each segment (shown in gray) independently to maximize the overlap, leading to discontinuities at the segment boundaries. In the coherent analysis, the model drifts out of phase with the data leading to a low overlap. The extra freedom in the semi-coherent analysis partially compensates for this leading to a larger semi-coherent overlap, $\hat{\mathcal{O}}_{N=7}(d,h) \gg \mathcal{O}(d,h)$. The amplitudes are normalized such that $\langle d|d\rangle = 1$.

segments, the semi-coherent likelihood is related to the standard, coherent likelihood via

$$\log \hat{L}_{N=1}(d|\theta) = \max_{\phi} \log L(d|\theta). \tag{4.9}$$

Eqn. 4.8 is the definition of $\hat{L}$. We name this quantity the *semi-coherent likelihood* emphasizing the connection with L in Eqn. 4.1. However, $\hat{L}$ is not a likelihood in the usual sense. Because it is a function of the data it can be regarded as a new statistic that is introduced here as part of a new proposed search strategy.

Fig. 4.1 illustrates the semi-coherent approach for a pair of simple sinusoidal waves with similar, but not identical frequencies. In the figure the idea is illustrated in the time domain, although our analysis in the following sections will segment the data in the frequency domain. The signal and data gradually drift out of phase with each other over many cycles resulting in a low coherent overlap. In the semi-coherent analysis the

model frequency is kept constant across the entire range of the observation but the phase angle is allowed to vary between segments; this partially compensates for the difference in frequency with the data and the semi-coherent overlap is much higher than the vanilla inner product. The semi-coherent overlap is less sensitive to variation of the parameters that affect the frequency and phase evolution of the signals.

There is a freedom in our definition of the semi-coherent likelihood corresponding to the choice of the segment boundaries f_n. Perhaps the simplest option is uniform segmentation with $f_n = f_{\rm min} + (n/N)(f_{\rm max} - f_{\rm min})$. However, SmBBH and EMRI systems spend a disproportionately large amount of time at lower frequencies. Therefore, uniform segmentation would result in the majority of the signal being contained in a small number of segments. Another option is logarithmic segmentation with $\log(f_n/f_{\rm min}) = (n/N)\log(f_{\rm max}/f_{\rm min})$. This results in more segments at lower frequencies, however the signal is still not necessarily split equally between the segments. Therefore, we choose to define our segment boundaries with respect to the signal that is being analyzed. We opt to define segments such that they contain an equal squared SNR. The squared SNR in segment n is

$$\rho_n^2 \equiv \sum_\alpha 4 \int_{f_n}^{f_{n+1}} \frac{|\tilde{h}_\alpha(f)|^2}{S_\alpha(f)} {\rm d}f = \frac{\rho^2}{N}; \tag{4.10}$$

this is an implicit equation for the segment boundaries f_n under the equal squared SNR segmentation scheme. The total (optimum) SNR is given by $\rho^2 = \sum_n \rho_n^2$. The downside of this approach is that the segment boundaries $f_n(\theta)$ now depend on the source parameters, and must be recomputed at each evaluation of the semi-coherent likelihood.

It is hoped that the semi-coherent likelihood will lead to wider posterior distributions, particularly on the parameters which strongly influence the phase of the GW signal (these are referred to as *phasing parameters*). For the signals observed by ground-based detectors the phasing parameters can be identified with the *intrinsic* source parameters. However, this identification breaks down for LISA where extrinsic parameters (such as the sky position) also effect the phase (e.g. via the direction-dependent Doppler shift due to the

motion of the detector). It is shown in Sections 4.3 and 4.4 that the semi-coherent analysis method does indeed broaden the posteriors on the most important phasing parameters while leaving the posteriors on the other parameters largely unchanged.

Another well-known method for broadening posterior distributions is tempering or simulated annealing [182, 183]; this involves raising the likelihood to a power β, where $\frac{1}{\beta}$ is commonly called the annealing temperature. This can be used as a method of accelerating sampling for highly multimodal probability distributions, because it makes it easier for many stochastic methods to traverse the likelihood surface. This has been used in several parameter estimation studies for sources in LISA data [128, 147, 164]. Tempering or annealing modifies the likelihood surface in a way that is somewhat similar to the semi-coherent approach, in that it reduces the severity of secondary maxima. However, there exists a clear distinction between the two methods. The semi-coherent method smooths the log-likelihood surface around the injection, *removing* secondary peaks in the log-likelihood surface, whereas tempering raises the "floor" value of the log-likelihood at large distances from the peak which has the effect of gradually congealing secondary peaks in the likelihood. Tempering preserves multi-modality in the log-likelihood surface while the semi-coherent method eradicates it. See appendix A.1, and Fig. A.1 therein for a comparison between tempering and semi-coherent methods.

4.3 Case study: GW170817

The BNS signal GW170817 is the longest GW signal observed to date. In some respects it is the closest thing we currently have to a SmBBH or EMRI signal (albeit, still with orders of magnitude fewer wave cycles). Therefore, reanalyzing GW170817 is a gentle way to prepare and build up to analyzing SmBBH and EMRI signals. In this section we explore the properties of the semi-coherent likelihood in Eqn. 4.8 by using it to reanalyze GW170817.

This case study is intended to build intuition for the semi-coherent likelihood, ensuring

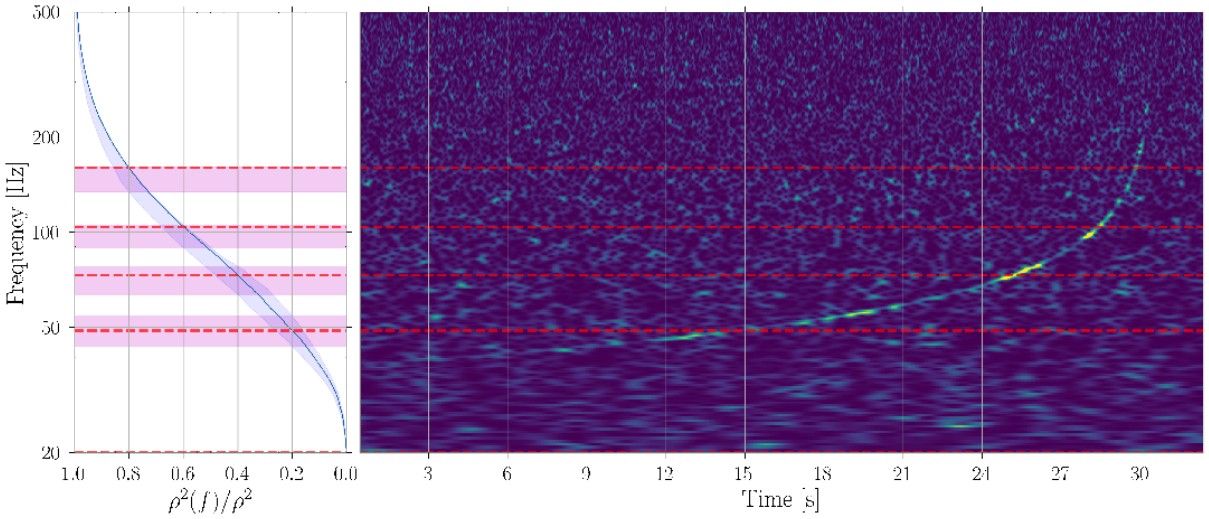

Figure 4.2: Illustration of the $N = 5$ segment semi-coherent analysis of GW170817. *Right panel:* The Q-transform time-frequency scan of the 32 s of analysis data from the LIGO Livingston detector (other two instruments not shown). The Q-transform is a time-frequency representation of the data, intuitively this can be thought of as similar to the short-Fourier-transform decomposition described in Sec. 3.4, with the notable change of using variable time-windows which allows for better resolution of high and low-frequency behaviour [184]. The characteristic chirp of the BNS is clearly visible. *Left panel:* The cumulative (normalized) squared network SNR. The rate at which the SNR accumulates with frequency depends on the source parameters; the blue line and shaded region show the result for the maximum likelihood and 90% credible region computed using the posterior samples from Sec. 4.3. The $N = 5$ segment semi-coherent likelihood splits the frequency range $[20, 1000]$ Hz at 4 intermediate frequencies such that an equal squared SNR accumulates in each segment. These four frequencies (median and 90% credible regions) are shown with horizontal red lines.

that the peaks, although broader, remain consistent with the vanilla likelihood (i.e. L in Eqn. 4.1). This also gives us an opportunity to explore how the semi-coherent likelihood behaves in the presence of real detector noise. Although a real search will aim to just locate the peaks in the semi-coherent likelihood, in this section the full likelihood distribution is explored using stochastic sampling, in order to gain a better understanding of the tails of the semi-coherent likelihood surface.

In the following analysis we use the waveform model `IMRPhenomPv2_NRTidalv2` [185, 186]. This is a fast, frequency domain, phenomenological waveform model built on the quasi-circular, spin-precessing binary BH model `IMRPhenomPv2` [187, 188]. Tidal effects are expected to be significant in the late inspiral of a BNS system and `IMRPhenomPv2_NRTidalv2`

accounts for this through the inclusion of tidal deformability parameters for both compact objects; these are parametrized by $\tilde{\Lambda}$ and $\delta\tilde{\Lambda}$ [189]. The `PhenomPv2` waveform model is in turn constructed from the spin-aligned `PhenomD` [190, 191] model which contains only the $(l, |m|) = (2, 2)$ mode. `PhenomPv2` "twists" this model in a way that mimics the effects of spin-orbit precession. It is the $(2, 2)$ mode of the `PhenomD` waveform that the `NRTidalv2` model modifies, incorporating an amplitude and phase correction which originate from the tidal interactions between the two neutron stars in the binary.

The data for the following analysis span 32 seconds in the GPS time range [1187008852.4, 1187008884.4]s. The PSD is estimated using a Welch periodogram (as implemented in GWpy [192]) using 1024 seconds of off-source data in the GPS time range [1187007316.4, 1187008340.4]s, which is offset from the trigger by 512 s to avoid any possible contamination from the long-lived signal. The time-series data, sampled at 4096 Hz, was obtained from the Gravitational Wave Open Science Center [193]. Data from the Livingston, Hanford and Virgo detectors are used. The Livingston data contains a prominent glitch just before the merger (see Fig. 2 of Ref. [121]). The glitch has been modeled and removed using `Bayeswave` [129]. Specifically the glitch-subtracted data is obtained from Ref. [194]. The data was analyzed using $f_{\rm min} = 20\,{\rm Hz}$ to $f_{\rm max} = 1000\,{\rm Hz}$.

The likelihoods used in this case study segment the data and model using the equal square SNR scheme discussed in Sec. 4.2. At each new evaluation of the likelihood, i.e. at each proposed set of parameters θ, the segment boundaries f_n must be recomputed from Eqn. 4.10. This is done using the cumulative squared SNR as a function of frequency;

$$\rho^2(f) = \sum_\alpha 4 \int_{f_{\rm min}}^{f} \frac{|\tilde{h}_\alpha(f)|^2}{S_\alpha(f)} {\rm d}f, \tag{4.11}$$

where the total square SNR $\rho^2 \equiv \rho^2(f_{\rm max})$. This curve is used to divide the total square SNR into segments that contain equal ρ_n^2. This process is illustrated in Fig. 4.2.

Likelihood maximization with respect to phase is performed analytically, as shown in Eqn. 4.2. This analytic maximization is possible because the `PhenomD` model, from which

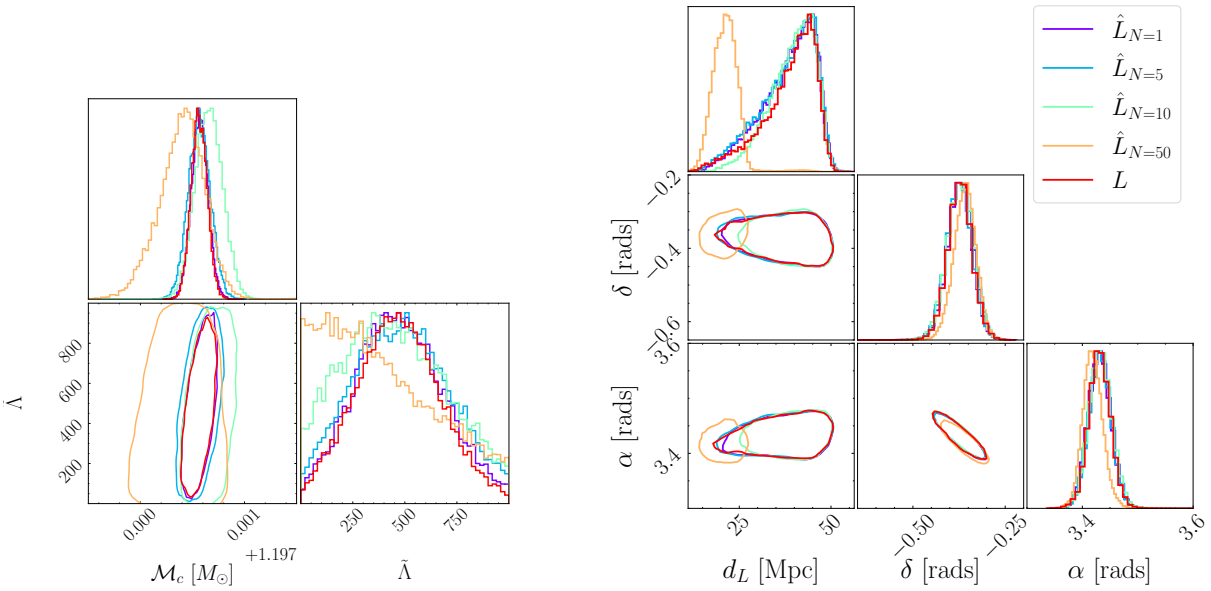

Figure 4.3: Posterior distributions obtained with various semi-coherent likelihoods $\hat{L}_N(d|\theta)$ with N in the range 1 to 50, and the vanilla likelihood $L(d|\theta)$. *Left panel:* posterior distributions for a selection of the phasing parameters ($\mathcal{M}_c$ and $\tilde{\Lambda}$). These are intrinsic parameters describing the frequency (or phase) evolution of the source. *Right panel:* posterior distributions on a selection of non-phasing parameters (luminosity distance d_L, declination δ, and right ascension, α). These are extrinsic parameters describing the location of the source. In all 2D plots the contours show the 90% confidence regions.

our waveform is constructed, is a single mode $(2, 2)$ waveform. It has been verified that the likelihood is unchanged if the phase maximization is instead performed numerically.

Stochastic sampling of the posterior distribution was performed using the `dynesty` [195] nested sampler [107] as implemented in the `Bilby` [196] library. However, we note that a custom log-likelihood function is used, which implements the semi-coherent $\log \hat{L}_N$ described in Sec. 4.2. The priors used for the following analyses were those in Ref. [197], with the exceptions of the following parameters: the priors on the sky position angles (right ascension α and declination δ) are uniform over the whole sky ($\cos\delta \in [-1, 1]$ and $\alpha \in [0, 2\pi]$), the dimensionless tidal deformability parameters are uniform over the range $\tilde{\Lambda} \in [0, 1000]$ and $\delta\tilde{\Lambda} \in [-5000, 5000]$.

Fig. 4.3 shows a set of posterior distributions obtained with both the semi-coherent and vanilla likelihoods. The corner plots show subsets of the source parameters, consisting of phasing (left) and non-phasing (right) parameters. Two important phasing parameters in

this case are the chirp mass ($\mathcal{M}_c$) and the dimensionless tidal deformability parameter ($\tilde{\Lambda}$). Examples of non-phasing parameters are those that define the 3D location of the source. Posterior distributions on the phasing parameters broaden as the number of segments used in the semi-coherent likelihood is increased, while the distributions for the non-phasing parameters are not significantly affected. The exception to this is the luminosity distance posterior for the $N = 50$ segment likelihood, which is biased, and is not consistent with lower segment posteriors (which are themselves consistent with literature [121, 197]). We have verified that the distance posterior varies continuously between the $N = 10$ and $N = 50$ cases shown. The bias in the distance will not be problematic for the purpose of a search. Fig. 4.3 displays several posteriors obtained from the same underlying data, which exhibits identical noise characteristics, however the likelihoods are different; the small shifts between the likelihood peaks can likely be attributed to the semi-coherent nature of each likelihood interacting with the noise.

The semi-coherent likelihood partitions the data into a number of segments. A natural theoretical maximum number of such segments is set by the number of orbits in the signal. Beyond this number, each frequency segment covers $\sim$ one orbit and orbital phase maximization ceases to be meaningful. A practical, useful maximum number of segments will be a correction factor ($\gamma < 1$) multiplied by the number of orbits in the signal. In the case of GW170817, the signal has ~ 3000 orbits in band over the whole ~ 100 seconds the signal is present in the detector [121]. The semi-coherent limit for the number of segments has been verified; we checked that the semi-coherent analysis with ~ 500 segments produces posteriors that are extremely broad, multi-modal, and not consistent with the literature (these results are not shown here), for GW170817 $\gamma \sim 1/6$. SmBBH/EMRI systems have $\gtrsim 10^5$ orbits in band, so the natural upper limit of segments will be much higher. A far more conservative upper limit on the number of segments is set by the sensitivity of the search in data which contains noise, see discussion in Chapter 5.

The key takeaway from this case study is the semi-coherent likelihood does broaden the posterior distributions of parameters that control the GW phase, while not affecting

the non-phasing parameters. Additionally, while this analysis method breaks down when the number of segments approaches the number of orbital cycles, we hope this will not be an issue for the SmBBH/EMRI signals as both source types will undergo a much larger number of orbits, $\sim 10^5$, within the LISA frequency band.

4.4 Stellar-mass binary black holes

We now apply semi-coherent likelihoods to the analysis of a LISA SmBBH signal. In contrast to the analyses in Sec. 4.3 which worked with real noisy data, all of the semi-coherent analyses in this section and Sec. 4.5 are performed on zero-noise mock injections; this means that the (fully-coherent) likelihood surface is peaked at the injected source parameters. As we are in the early prototyping phase of testing this search strategy, we simulate a noise-less data-stream to simplify the search. We do not expect the inclusion of noise to change the results significantly, posteriors obtained from noisy data will be shifted relative to the true parameters, however the shifted posterior will still be consistent with the no-noise posterior.

The analysis in this section uses parameter estimation methods that fully explore the likelihood distributions (including their low-probability tails) mirroring the analysis performed in the previous section for GW170817. This is done to build an understanding of the properties of the semi-coherent likelihood. The sampling iterates through a sequence of semi-coherent likelihoods with steadily decreasing number of segments, N, progressively localizing the signal to smaller regions in parameter space, thereby mimicking a search process. However, a real search would not use sampling methods that waste time exploring the tails of the distributions at early stages. Sec. 4.5 repeats this procedure using an optimizer (as opposed to a sampler) as part of a more realistic search algorithm.

We inject a fiducial SmBBH source to test our sampling and optimization methods. The injected source parameters are given in Tab. 4.1. The sampling is performed over the following parameters with flat priors: chirp mass $\mathcal{M}_c$, time to merger t_c, dimensionless

Table 4.1: Injection parameters and priors for both sampling and optimization conducted in sections 4.4 and 4.5. Parameters above the line are those that are sampled in, using flat priors over the ranges shows, those below the line are *derived* parameters defined in the text. All masses are given as detector frame quantities.

Parameter	Injection	Prior range: $\theta_{\rm min}$–$\theta_{\rm max}$
$\mathcal{M}_c\,[M_\odot]$	62.46453697	$[61.46, 63.46]$
$t_c\,[\text{months}]$	38.04	$[t_c-1, t_c+1]$
$\delta\mu$	0.27	$[0, 0.7]$
$\lambda\,[\text{rad}]$	2.0	$[0, 2\pi]$
$\sin\beta$	0.3	$[-1, 1]$
$\sqrt{A_{\rm left}}\,[\text{pc}^{-1/2}]$	3.73×10^{-5}	$[0, 10^{-4}]$
$\sqrt{A_{\rm right}}\,[\text{pc}^{-1/2}]$	4.44×10^{-5}	$[0, 10^{-4}]$
χ_1	-0.58	$[-1, 1]$
χ_2	-0.17	$[-1, 1]$
$\phi_{\rm left}\,[\text{rads}]$	6.04	$[0, 2\pi]$
$\phi_{\rm right}\,[\text{rads}]$	2.24	$[0, 2\pi]$
$d_L\,[\text{Mpc}]$	300	-
$m_1\,[M_\odot]$	95	-
$m_2\,[M_\odot]$	55	-
$\phi\,[\text{rads}]$	1	-
$\psi\,[\text{rads}]$	-2.52	-
$\iota\,[\text{rads}]$	1.66	-
ρ	11.44	-

mass difference $\delta\mu$, ecliptic longitude λ, sine of ecliptic latitude $\sin\beta$, square root of left and right-handed circularly polarized GW amplitudes $A^{1/2}_{\rm left,right}$, dimensionless aligned spin magnitudes χ_1 and χ_2, and phases for the left- and right-handed GW polarizations $\phi_{\rm left,right}$. These are related to the more familiar component mass parameters m_1, m_2, phase and polarization angles ψ, ϕ, inclination ι, and luminosity distance d_L, via the equations in appendix A.3.

The wide (i.e. uninformative) prior ranges $\Delta\theta \equiv \theta_{\rm max} - \theta_{\rm min}$ are chosen to be representative of a search, Much narrower priors are typically used in parameter estimation studies, e.g. Refs. [65, 163]. Most parameters are allowed to vary over their full physical ranges, the exceptions are the important phasing parameters, $\mathcal{M}_c$ and t_c. The priors on these parameters are wide enough to cover a sizable fraction ($\sim 1/50$ in both dimensions) of the LISA discovery space described at the beginning of Sec. 6.1. We envision eventually using multiple ($\sim 50^2 = 2500$) such searches to tile the full parameter space.

Computationally efficient post-Newtonian waveforms for the inspiral phase of SmBBH systems, incorporating the effects of eccentricity and spin-precession, are available; see for example Refs. [198, 137, 199]. These low-order post-Newtonian waveforms are expected to be sufficiently faithful for the analysis of SmBBH sources in LISA [200]. These waveforms are computationally fast which makes their use for searching and sampling over the large parameter space feasible. As a proof of concept, this analysis uses the simple `TaylorF2` waveform (as implemented in `Balrog`[3]) for spin-aligned, non-eccentric (i.e. quasicircular) binaries. Our methods are expected to generalize easily to waveforms which incorporate additional physics, such as spin-orbit precession and eccentric orbits. The `TaylorF2` waveform includes only the $(\ell, |m|) = (2, 2)$ spherical harmonic mode; this is expected to be sufficient as higher modes are strongly suppressed early in the inspiral. The waveform model produces the polarizations $\tilde{h}_{+,\times}(f)$ as a function of the source parameters, θ.

It is also necessary to model the response of the LISA instrument to the two GW polarizations $\tilde{h}_{+,\times}(f)$. The three satellites in the LISA constellation are connected by six laser links. The measured phase time series in these links are expected to contain large-amplitude laser noise. Therefore, six links are combined into three output channels in a process called time-delay interferometry [205] that is designed to suppress this laser noise below the level of other, secondary noise sources. The LISA response is modeled using a rigid adiabatic approximation [206] which is used produce the TDI outputs $\tilde{h}_{X,Y,Z}(f)$. The response derived from the rigid adiabatic approximation has previously been tested for SmBBH sources in LISA [163]. The TDI outputs are then transformed [207] to the noise-orthogonal TDI channels $\tilde{h}_{\alpha}(f)$, where $\alpha \in \{A, E, T\}$ which are used for the likelihood evaluation.

The BNS analysis in Sec. 4.3 used a uniformly-sampled fast Fourier transform (FFT) frequency grid. However, SmBBH signals are broadband ($\Delta f \sim 10^{-2}\,\mathrm{Hz}$), and LISA observations are long duration ($T \sim 10^{8}\,\mathrm{s}$), which would result in an FFT frequency

[3] `Balrog` is a package being developed for waveform generation and parameter estimation for LISA sources, including supermassive binary BH mergers [201], double white dwarfs [202, 203, 204, 59] and SmBBH inspirals [163, 65].

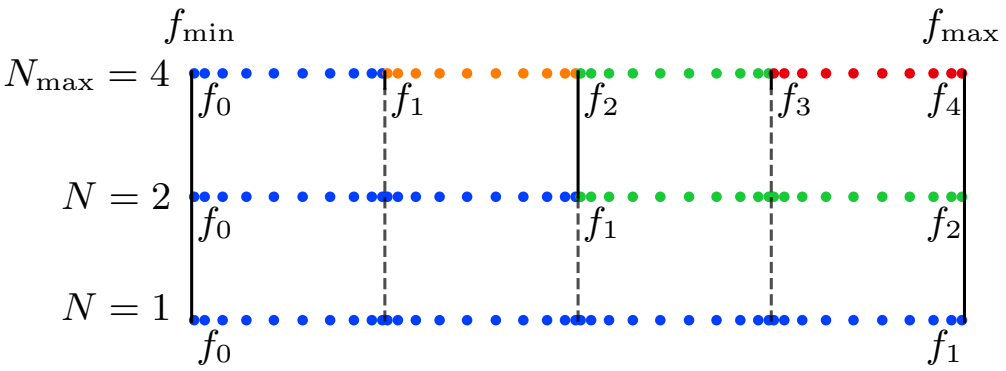

Figure 4.4: Illustration of the segmentation method used for SmBBH sources with the Clenshaw-Curtis quadrature integration rule. In this illustrative example, the highest number of segments is $N_{\rm max} = 2^2 = 4$, and the other possible numbers of segments ($N = 2^1 = 2$ and the fully coherent $N = 2^0 = 1$) are constructed by the union of pairs of quadrature grid. At each level, colors and solid vertical lines indicate the semi-coherent frequency segments, color dots represent the location of the quadrature integration nodes (in this illustrative example there are 10 nodes per quadrature grid), and dashed vertical lines represent the end of a quadrature integration grid (but not of a semi-coherent segment).

grid with $\sim 10^6$ nodes. A likelihood calculated using this frequency grid would be too slow to be used in a search. Instead, `Balrog` uses Clenshaw-Curtis quadrature [208] to accelerate evaluations of the inner product, using $\sim 10^2$ frequency nodes, vastly reducing the computational cost of likelihood evaluations [65, 163, 201]. We highlight a caveat that quadrature methods are not suitable for evaluating likelihoods on realistic data containing noise, this is further discussed in more detail in appendix A.4.

As described in Sec. 4.2, we aim to split the data into frequency segments containing equal square SNR. However, this is complicated by the fact we are no longer integrating using a uniformly-spaced FFT frequency grid. To circumvent this, we use multiple quadrature grids adapted to our semi-coherent frequency segments. Specifically, (i) we choose the maximum number of segments that will be used to be a power of two, $N_{\rm max} = 2^a$. (ii) We select a reference waveform with parameters chosen in the center of the prior ranges. The reference waveform is evaluated once (before the search) on the uniformly-sampled FFT grid and this is used to find the segment boundaries $f_0, f_1, \ldots, f_{N_{\rm max}}$ as described in Sec. 4.2. (iii) We then construct $N_{\rm max}$ irregularly-spaced quadrature frequency grids for

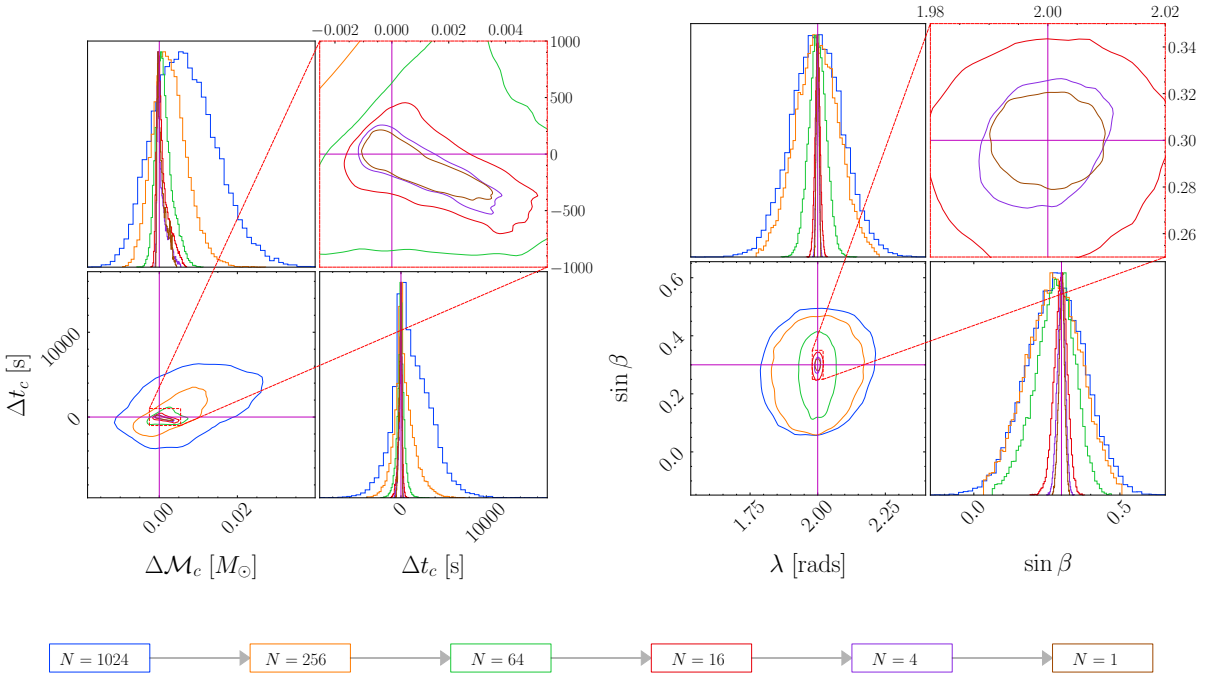

Figure 4.5: Iterative sampling results for the semi-coherent likelihood with varying number of segments, N. 1D marginalized posteriors are shown for chirp mass and time-to-merger parameters (left) and sky position parameters (right). (The axes on the left-hand plot are shifted such that the injection parameters are at the origin.) The semi-coherent likelihood with large N broadens the likelihood peak significantly in the parameters which effect the GW phasing; for long-lived LISA sources, such as SmBBHs, this includes the sky position parameters. 2D contours show the 90% confidence interval. Results are shown for $N \in \{1024, 256, 64, 16, 4, 1\}$ segments. In both corner plots the top-right panel shows a zoom in on the region around the injection parameters.

these segments. These frequency grids are then fixed and the data and model (for any parameters in the prior range) are evaluated on this sparse grid. For any power-of-two number of segments, $N = 2^b$ with $b \leq a$, $\hat{L}_N$, can be evaluated by pairing together these segments. This construction is illustrated in Fig. 4.4 with $N_{\max} = 4$. Since, for computational efficiency, the same frequency grid is used for all sources within the prior, most sources have only approximately equal squared SNR per segment. For our fiducial source, prior range and choice of $N_{\max}$, we have verified that ρ_n^2 varies between segments by a factor $\lesssim 3$.

The semi-coherent likelihood, as defined in Eqn. 4.8, is evaluated using the segmented

inner products computed using Clenshaw-Curtis quadrature (see Appendix A.4),

$$[a|b]_n^N = \sum_\alpha 4\mathrm{Re} \sum_i \frac{w_{i,n}\tilde{a}_\alpha(f_{i,n})\tilde{b}^\dagger_\alpha(f_{i,n})}{S_\alpha(f_{i,n})}. \tag{4.12}$$

Here, $f_{i,n}$ is the i^{th} node in the n^{th} frequency quadrature grid and $w_{i,n}$ are the associated quadrature weights. If the number of nodes and the span of each quadrature grid is the same, then $w_{i,n} = w_i$. We find empirically that a search starting with $N_{\max} = 1024$ segments, with 11 nodes per quadrature grid, performs well for this source. We note that this is reasonably consistent with the rough early estimate of $N \gtrsim 100$ for the minimum number of segments required for an EMRI search made in Ref. [88].

While we choose to use quadrature rules, as illustrated in Fig. 4.4, the semi-coherent method is more general and could be adapted to work with other techniques such as heterodyning/relative binning [209].

We simulate a 4 year LISA mission. Inner products are evaluated between the frequency limits $[f_{\text{low}}, f_{\text{high}}] = [0.0056, 0.1]\,\text{Hz}$. The source in Tab. 4.1 is 3.17 years from merger when LISA observations begin, and is initially radiating at a frequency above f_{low}. After exiting the LISA frequency band, the source merges in ~ 3.5 days. A simple analytical model based on the latest LISA science requirements (SciRD) was used for the noise PSD. The functions $S_\alpha(f)$ are the sum of the analytic approximations to the instrumental noise curve in Ref. [210] and galactic binary confusion noise in Eqn. 4 of Ref. [84] scaled to a mission duration of 4 years. These are used to construct PSDs for each of the noise-orthogonal TDI channels A, E and T.

Within the search region set by the prior ranges, we use an iterative search strategy. Initially, the sampler is tasked with exploring the $\hat{L}_{N=N_{\max}}$ likelihood surface; this is expected to exhibit the broadest features which makes finding the peak possible. Once the optimizer/sampler has converged, the number of segments in the likelihood is reduced to $N = N_{\max}/4$; we find that reducing the number of segments by a factor of four at each stage is reasonably efficient for our fiducial source. The prior ranges on the phasing

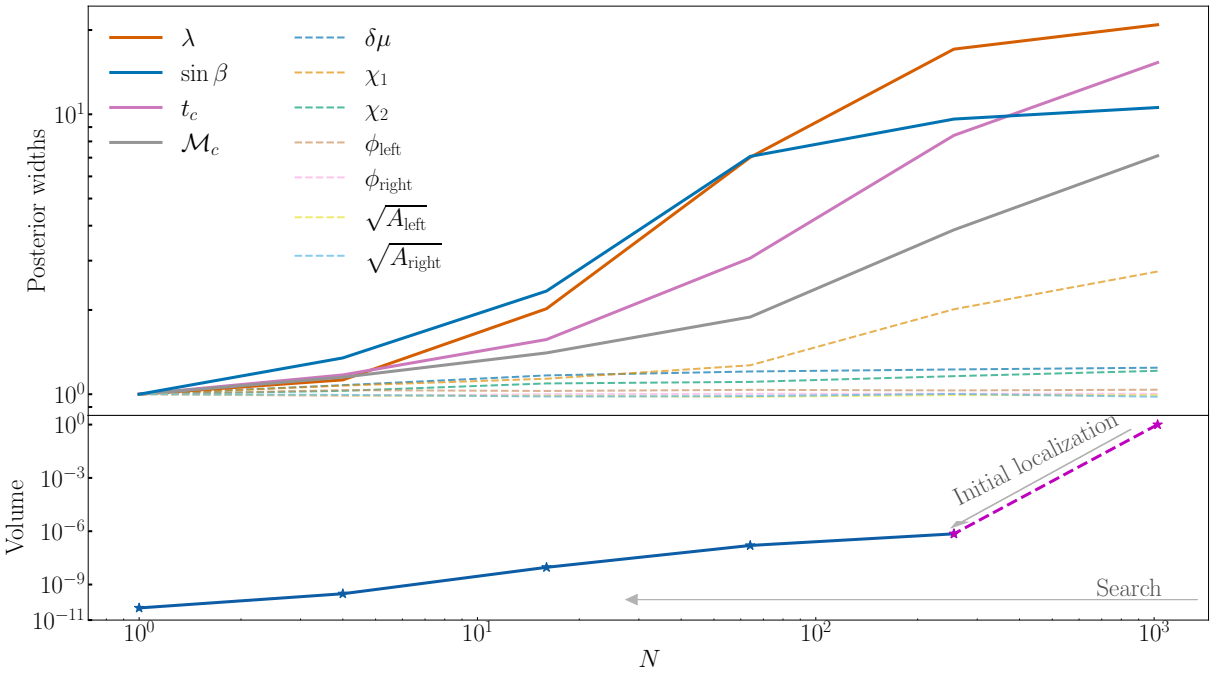

Figure 4.6: *Top*: The standard deviation, or width, of the 1-dimensional marginalized posterior distributions obtained with the semi-coherent likelihood using different numbers N of segments. Posterior widths (normalized by dividing by the width of $N = 1$ segment posterior) are shown for all parameters, with the most important phasing parameters highlighted using solid lines. *Bottom*: the volume of the prior hyper-rectangle used for the analysis at each number of segments. Note, only the priors on the four important phasing parameters are changed between each iteration. The volumes are normalized to the initial prior volume at $N = 1024$ segments. The initial prior (for $N = 1024$) is *not* a measure of the posterior size at this segment, instead it is a prior we have chosen that is large enough to represent the search problem while still remaining tractable for a proof of concept study. We label the transition: $N : 1024 \rightarrow 256$ as the 'initial localization' and distinguish it from the other segment transitions using a dashed magenta line.

parameters are also reduced, with the new bound on each parameter calculated as the 98% confidence interval of the 1D marginal posteriors from the previous stage. This simple approach shrinks the prior using progressively smaller hyper-rectangles. The sampler is now tasked with exploring the new $\hat{L}_N$ likelihood surface with smaller N. This process is repeated, reducing N and shrinking the prior ranges, until the sampler has explored the fully coherent $\hat{L}_{N=1}$ likelihood surface. Sampling was performed using the `CPNest` nested sampling package [211].

The sequence of posteriors for a selection of the phasing parameters are shown in Fig. 4.5. Constraints on these parameters improve throughout the iterative process.

Fig. 4.6 shows the width of the 1D marginal posterior distributions for all parameters as a function of the segment number N, alongside the prior volume at different stages of the search.

For the BNS GW170817 observed in LIGO/Virgo, the sky position and time-of-merger parameters are not strongly impacted by the semi-coherent analysis (see Fig. 4.3). These are generally referred to as extrinsic parameters and they don't impact the phasing of the GW signal. However, for SmBBHs in LISA this is not the case. The time-to-merger parameter controls the frequency of the source at the start of observations and the sky position affects the observed frequency via a periodic Doppler shifting caused by the detector motion. Therefore, posteriors on these parameters narrow during the search process (see Fig. 4.5).

4.5 Particle Swarm Optimization

In the previous section we tested the iterative semi-coherent search by sampling the likelihoods; this is unnecessarily inefficient for a search. In this section a stochastic optimization algorithm is used to locate and track the peak of the semi-coherent likelihood with varying N without wasting time exploring the tails of the likelihood distribution. Here, a Particle Swarm Optimizer (PSO) is used to do this, although the semi-coherent likelihood is general and can be used with any optimization algorithm.

PSO [173, 174] is a stochastic optimization algorithm which uses a swarm of a large number N_p of particles to optimize an objective function over a high-dimensional parameter space. In this study we use it to optimize the semi-coherent log-likelihood, $\hat{L}_N$. Each of the N_p particles in the PSO swarm has a position vector in parameter space that is updated at each iteration, $\theta^{\mu}_{p,i}$: the index $\mu \in \{0, 1, \ldots, \dim(\theta) - 1\}$ labels the components of the source parameter vector; the index $p \in \{0, 1, \ldots, N_p - 1\}$ labels the particles in the swarm; and the index $i \in \{0, 1, \ldots\}$ labels the iteration of the algorithm. The rule for

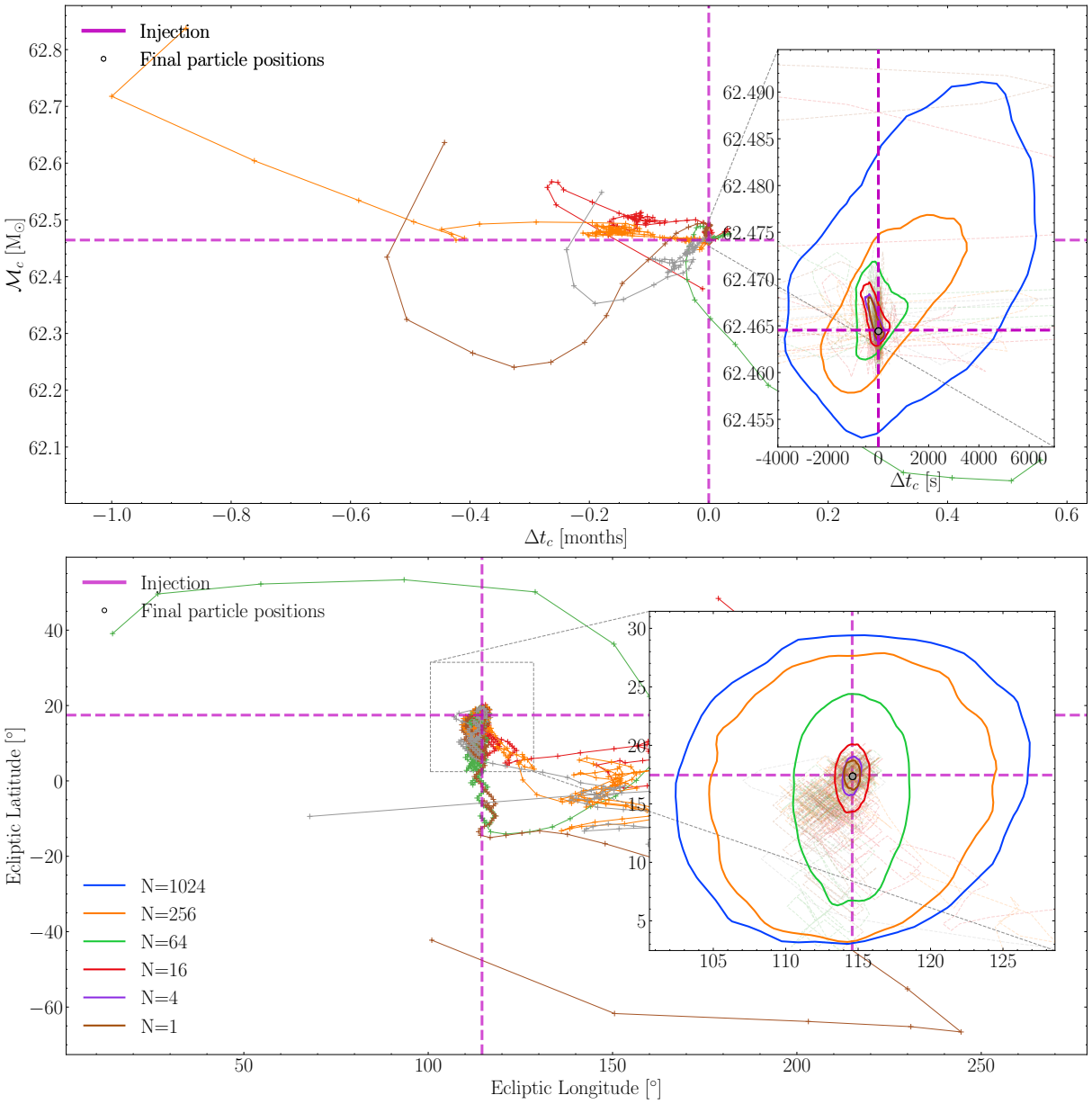

Figure 4.7: Tracks in parameter space for 5 randomly selected particles, showing evolution throughout the PSO optimization on the (*Top*) chirp mass, time to merger plane and (*Bottom*) ecliptic longitude and latitude plane. Inset plots zoom in around the injection parameters, overlaying 90% confidence intervals for posterior distributions obtained from analysis in Sec. 4.4.

updating the positions at each iteration is

$$\theta^{\mu}_{p,i+1} = \theta^{\mu}_{p,i} + v^{\mu}_{p,i}, \tag{4.13}$$

where $v^{\mu}_{p,i}$ is called the velocity and is calculated for each particle from the current state

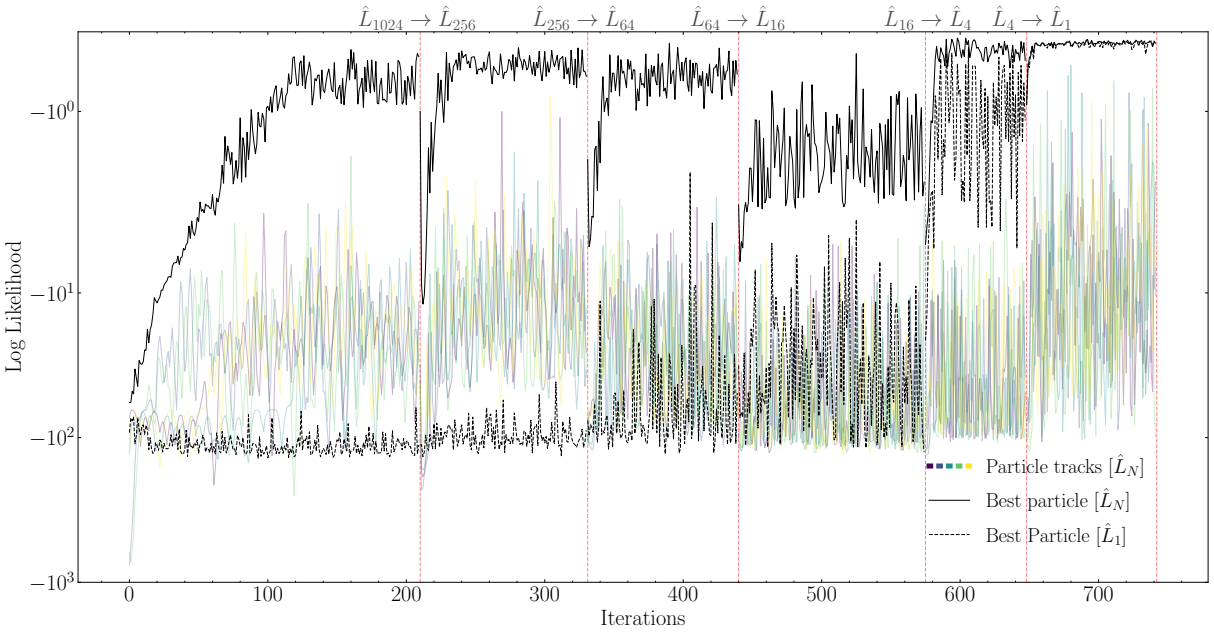

Figure 4.8: Log-likelihood evolution for the subset of particles plotted in Fig. 4.7 in solid lines. Dotted vertical lines indicate the iteration at which the optimizer switched to the next level in the hierarchical search. The colored lines show the track of the semi-coherent log-likelihood at the current number of segments, $\hat{L}_N$, for the 5 randomly selected particles (same particles as those shown in Fig. 4.7), throughout the entire evolution of the swarm. The black solid line shows the maximum semi-coherent log-likelihood of any particle in the swarm. (Note that the identity of the best particle changes repeatedly during the evolution, so the black line is not a particle track.) For comparison, the dotted black line shows the value of the coherent phase maximized log-likelihood $\hat{L}_{N=1}$ evaluated at the same parameter values as the solid black line. At early times the log-likelihood evaluated on the $N = 1024$ segment for the best particle increases quickly, however the coherent likelihood stays relatively flat; this shows the impact of using the semi-coherent likelihood with peaks that are much broader and easier to find in parameter space.

and past history of the whole swarm according to

$$v^{\mu}_{p,i+1} = \max(\epsilon^{\mu}, |u^{\mu}_{p,i}|)\frac{u^{\mu}_{p,i}}{|u^{\mu}_{p,i}|}, \tag{4.14}$$

where

$$\begin{aligned} u^{\mu}_{p,i} =& \Omega v^{\mu}_{p,i} + \\ & \Phi_P(r_P)^{\mu}_{p,i}(\Psi^{\mu}_{p,i} - \theta^{\mu}_{p,i}) + \Phi_G(r_G)^{\mu}_{p,i}(\Xi^{\mu}_{i} - \theta^{\mu}_{p,i}) \end{aligned} \tag{4.15}$$

and where $\Psi_{p,i}$ is the best (i.e. highest-likelihood) point visited by the p^{th} particle so far

in the evolution of the swarm,

$$\Psi^{\mu}_{p,i} = \theta^{\mu}_{p,I} \ , \quad \text{where} \quad I = \underset{i'<i}{\operatorname{argmax}} \hat{L}_N(d|\theta_{p,i'}), \tag{4.16}$$

and where Ξ_i is the best point visited by any particle

$$\Xi^{\mu}_{i} = \Psi^{\mu}_{P,i} \ , \quad \text{where} \quad P = \underset{p}{\operatorname{argmax}} \hat{L}_N(d|\Psi_{p,i}). \tag{4.17}$$

Eqn. 4.15 is called the PSO velocity rule and has been widely used in the literature. It includes 3 terms: the Ω term is called the *inertia*; the Φ_P term is called the *cognitive* term and acts to attract a particle back to the highest likelihood location that it has visited so far; and the Φ_G terms is called the *social* term and acts to attract the particle towards to the highest likelihood location that any particle in the swarm has visited so far. The quantities $(r_P)^{\mu}_{p,i}$ and $(r_G)^{\mu}_{p,i}$ are random numbers from $U[0,1]$ drawn independently for each particle, component and iteration.

Additional control over the behavior of the swarm can be achieved by varying the ϵ^{μ} parameters in Eqn. 4.14. These control the minimum velocity any particle can have along a particular dimension of parameter space. Imposing a minimum velocity mitigates against premature convergence to local optima, which is a common problem for PSO methods [212] when optimizing multi-modal objective functions. Strictly speaking, imposing a minimum velocity in this way also prevents the swarm from ever converging, because the particles can never stop moving. For this reason the ϵ^{μ} parameters must be decreased near the end of the search. Similar *velocity clamping* methods exist within the PSO literature, but these are usually aimed at restricting the maximum velocity of particles to prevent excessive exploration [213, 214].

Collectively, Ω, Φ_P, Φ_G, and ϵ^{μ} are referred to as the swarm hyperparameters. The behavior of the swarm can be altered by varying the hyperparameters. This allows us fine control over the rate of convergence and degree of exploration of the optimization algorithm, ideal for a semi-coherent search. Early in the search phase (high N) we want

the swarm to focus on exploring the parameter space to ensure the peak is found. We want to avoid at all costs the swarm getting stuck and wasting time optimizing secondary peaks. To achieve this the inertia, Ω, is set high and the social weight, Φ_G, is set low. Additionally, the minimum velocities, especially those in the most important phasing parameters are set to high values. Late in the search (low N), when a broad likelihood feature has already been identified, the aim is to refine the parameter values by optimizing and converging on the peak. To achieve this the inertia is decreased, and the social weight increased, and the minimum velocities are decreased. These different behaviors are often referred to as *exploration* and *exploitation* in the PSO literature. Tab. A.1 in appendix A.2 shows how all the PSO parameters vary throughout the search.

It is illustrative to compare PSO to a more well-known algorithm within the GW community, MCMC. PSO is similar in the sense that it is a stochastic algorithm where walkers move in a guided random way around parameter space. However, it differs in that it is an optimization, as opposed to a sampling algorithm and therefore tends to climb the likelihood surface without exploring the low-probability tails. PSO is also *not* a Markov-Chain because the velocity rule incorporates 'memory' of past positions through the Φ_P and Φ_G terms.

We use a swarm with $N_p = 15000$ particles. The initial particle positions are drawn from the prior, with velocities in parameter θ^μ drawn from the uniform distribution $U[-\Delta\theta/5, \Delta\theta/5]$ where $\Delta\theta$ is the prior range. As in Sec. 4.4, we use $N_{\max} = 1024$, and this likelihood surface is optimized over until the swarm is considered converged (see Appendix A.2). The optimizer is then configured for the next likelihood segment; dropping from $N_{\max} = 1024$ to $N = 256$. The particle positions are carried over from the final positions of the evolution at the previous level, mirroring the shrinking priors in the sampling. When dropping to a smaller number of segments the particle velocities are re-initialized by drawing from a zero-mean Gaussian distribution with a co-variance calculated from the final positions of all the particles in the previous segment. The swarm hyper-parameters are also changed to gradually transition the swarm from exploratory to

convergent behavior. This iterative process repeats until after $i_{\max}$ iterations the swarm has converged on the $N = 1$ phase-maximized coherent likelihood. The final value of $\theta_{\rm best} \equiv \Xi^{\mu}_{i_{\max}}$ is our estimate of the best-fitting parameters and constitutes the main result of the search.

The colored lines in Fig. 4.7 shows particle tracks in selected parameters for 5 randomly chosen particles from the swarm optimization throughout the evolution. Fig. 4.8 shows the corresponding log-likelihood evolution tracks for these 5 particles. Notice in Fig. 4.8, immediately after a step in the segment level, the log-likelihood curve (black solid line) drops significantly. Since the particle positions don't change as we move between segment levels, this drop in log-likelihood is due to the increased sensitivity to waveform phase of the semi-coherent likelihood with a smaller N. At each new level we observe that the function values get slightly worse in the first few iterations, before starting to improve. We attribute this to the particles initially exhibiting exploratory behavior due to the re-initialized random velocities. The simple PSO algorithm implemented in this study localizes the parameters of the injected signal with good accuracy, specially in the phasing parameters which are of interest in the context of establishing narrow priors around the posterior bulk for parameter inference. For the fiducial signal considered here, the search has an execution time $\sim$ 15 hours, although we obtain good estimates for phasing parameters such as chirp mass, time to merger and sky position parameters from the state of the optimizer at the end of the 1024 segment level, within $\sim$ hours. Parameter estimation on the vanilla likelihood has a tractable computational cost when paired with priors derived from this PSO search result, successfully sampling the vanilla posterior for the fiducial source in $\sim$ 20 hours using `CPnest`.

4.6 Future work

In this work we demonstrated how a search using a likelihood with a variable level of coherence in conjunction with a particle-swarm-based optimizer can be used to find a

SmBBH signal in mock LISA data. This demonstration has been performed in idealized data (i.e zero-noise, no gaps or glitches; see the discussion in Sec. 6.1). Further studies with realistic noise are needed to establish the sensitivity of this search at a fixed false alarm rate.

In this study we have demonstrated the search for a fiducial signal, searching within one 'tile' in parameter space. Scaling this up and tiling the whole parameter space of interest is left for future work. We have also searched for only one source, a more realistic search would aim to find multiple sources, although this is not expected to be a major problem as it is unlikely there will be more than one source per search 'tile' (see Tab. 1 of Ref. [215]). Precise tuning of our method (i.e. selecting the decreasing sequence of segment numbers between $N = N_{\max}$ and $N = 1$, and the PSO hyperparameters) across parameter space is also left as future work. Finally, this method has not been tested with more complete waveforms that incorporate additional physics eg. eccentricity, spin-orbit precession etc, however this is not expected to be a major obstacle.

There are a number of similarities between SmBBH and EMRI signals observed by LISA. Designing and implementing a successful SmBBH search is likely to be extremely good preparation for the EMRI search problem. However, these two sources are astrophysically distinct from one another; LISA observes the early inspiral of SmBBH systems, these systems are in the regime $v \ll c$, meanwhile the late stage inspiral of EMRI system places them in the $v \lesssim c$ regime. Due to this and the extreme asymmetry in the mass ratio of the EMRI system, EMRI signals consist of dozens of frequency modes (also known as harmonics or *voices*), each contributing a non-negligible fraction of the signal SNR. This is in stark contrast to SmBBH signals which are well described by just the leading order quadrupole frequency mode. The frequency evolution of each EMRI mode individually looks somewhat similar to a SmBBH inspiral. In addition to the many modes, EMRIs are generically expected to occur on eccentric and inclined orbits further complicating waveform generation [84]; consequently EMRI waveforms are expensive to evaluate. Increasingly accurate and efficient time-domain EMRI waveforms are starting to become

available [161] and these have started to be used for frequency-domain data analysis using a stationary phase approximation [82] for each frequency mode. However, it is still not clear which EMRI waveforms will be available (and at what computational cost) for use in searches during the LISA mission.

The semi-coherent method developed here would need to be augmented to deal with the numerous modes in EMRI signals. The segmentation of the signal done here in the frequency domain applied to an EMRI would analyze different modes at different times. Therefore, it might be necessary to analyze each mode individually with this type of frequency-domain semi-coherent method. An investigation of this multi-mode, semi-coherent analysis is left to future work.

Unlike the SmBBH likelihood, the EMRI likelihood surface exhibits an extreme degree of multi-modality, this was recently studied in detail within Ref. [89]. This likelihood has many spurious secondary peaks of comparable height to the primary likelihood peak around the true source parameters. It is worth emphasizing these peaks *do not* originate from noise artifacts within the data but rather due to alignment of frequencies (and frequency derivatives) between waveforms evaluated at different points in parameter space. The height of the secondary modes were found to only be weakly correlated with the euclidean distance between the secondary peak parameters and the injection parameters [89]. This makes the EMRI search even more challenging than the SmBBH case.

PSO is a highly flexible algorithm which can be tuned for this extremely degenerate likelihood surface. The velocity rule can be easily adapted to split a swarm into multiple sub-swarms. We propose the use of this multi-population particle swarm optimizer to explore this degenerate likelihood surface. Similar variants of particle swarm optimization for multi-modal objective functions have been studied in Refs. [175, 176]. Each swarm will be assigned to a peak and optimize across them in parallel, prioritizing the 'best' peaks. Similar methods are used in Ref. [216] to sort through multiple optima, using genetic algorithms to search for massive binary BH mergers in mock LISA data.

The initial configuration of such a multi-population swarm would be similar to that

of the vanilla version presented in this study, consisting of one exploratory swarm with a greater weight in Φ_P, causing the particles to cluster around the large number of optima. At the end of the first segment level, a clustering algorithm such as that suggested in Ref. [83] will be applied, clustering the single population swarm into multiple populations, each exploring one optimum. Over the course of the next segment level, each population will be treated as an individual swarm, optimizing over its own maxima. Clustering will be conducted at the end of each segment level, terminating swarms that are exploring sub-dominant peaks according to some 'veto' criteria, such as that suggested in Ref. [83] and redistributing the particles to the other swarms. One such 'veto' function, is proposed in Ref. [217], calibrated to suppress secondary maxima that arise due to phase matching of the dominant frequency mode.

In conclusion the semi-coherent search is a promising avenue for broadband, chirping sources which undergo many orbits in the LISA frequency band and warrants further investigation into its efficacy as a search pipeline for SmBBH and EMRI sources.

5. GPU-accelerated search for stellar-mass binary inspiral signals in LISA

Overview

This chapter is an edited version of Ref. [218]:

Diganta Bandopadhyay and Christopher J. Moore. GPU-accelerated semi-coherent hierarchical search for stellar-mass binary inspiral signals in LISA, published in Physical Review D, Volume 110:103026, November 2024.

This work follows on from that presented in Chapter 4 and crucially extends the search to data which contains Gaussian, stationary noise. Graphics processing units (GPUs) are used to accelerate the search in the presence of noise, this is an important technical advancement as some of the acceleration methods used in the previous chapter are rendered invalid when the datastream contains noise, see discussion in Appendix A.4. End-to-end searches are performed for a few simulated sources with wide search-like priors. The search presented here involves an initial optimization stage accompanied by a subsequent seeded MCMC stage that performs parameter estimation. A framework is established which can be used to estimate the false alarm probability of a given detection, at the time of publication this had not been considered for most LISA searches.

I wrote all of the code that is used in this chapter (released publicly in Ref. [219]) and performed all the analyses presented within, with input and guidance from my supervisor. I wrote and edited the text for all sections of this chapter, with guidance from my

supervisor.

5.1 Introduction

In this work we continue developing a semi-coherent approach to the search for SmBBH signals, building on the work in Ref. [149]. The approach adopted here is inspired by semi-coherent methods developed for searching for continuous GWs (e.g. from spinning asymmetric neutron stars) in LVK data [220, 221, 222]. The underlying reason for the large computational costs of a coherent analysis is the same for the SmBBH searches as it is for the LVK continuous wave searches; the long signal duration leads to extreme sensitivity to parameters that control the GW phase. Following our previous chapter, we adopt a hierarchical search scheme, gradually increasing the level of coherence (achieved by decreasing the number of segments) to narrow in on the true parameters of the source; as the number of segments decreases so does the false-alarm probability. As the number of segments is increased (and the segments get correspondingly shorter in time/frequency) there is a greater chance for a model to be fit to noise in the data, which decreases the sensitivity of the search and increases the threshold SNR.

One of the key differences between the work presented in Chapter 4 and that shown here is that we include simulated noise within our datastreams. This is important as in principle without this the search could be performed with an arbitrarily high number of segments, however in the presence of noise the number of segments must be limited to avoid a high false alarm probability. This work performs the search in the presence of noise and makes an effort to approximate the false alarm probability in a restricted region of parameter space.

Although our search pipeline does produce a crude posterior on the source parameters as an output, the focus here is very much on the search phase and identifying the presence of a source and estimating its statistical significance. We envision the crude posterior produced being used to seed a subsequent high-quality parameter inference with narrow

priors.

Besides our previous work, another approach to a hierarchical semi-coherent search has been developed by Ref. [223]. This study used a slightly different search statistic which analytically maximises over a number of extrinsic parameters, reducing the dimensionality and thus volume of parameter space that must be searched. They also used a smaller number of semi-coherent segments resulting in a lower detection threshold for the signal-to-noise ratio (SNR) but at an increased computational cost compared to the method presented here. Also, for computational reasons, Ref. [223] restricted the analysis to a single TDI channel, whereas this work uses all three.

In our previous work we demonstrated the feasibility of a semi-coherent search on simulated SmBBH signals with zero noise [149]. In this chapter, the semi-coherent search is further developed in the presence of simulated noise. Particle swarm optimization (PSO) is used as the stochastic search algorithm over source parameter space, however it is here improved to allow for tracking the multiple maxima that are expected when noise is present in the data stream (this feature will also be vital when such methods are applied to the EMRI search; see, for example, Ref. [89]). We also show that in the final stage of the search PSO particles can smoothly transition into Markov-chain Monte-Carlo (MCMC) walkers for parameter estimation. Searching over noisy data is computationally expensive, and this is solved here by hardware acceleration using graphics processing units (GPU).

This chapter is structured as follows. Sec. 5.2 defines the semi-coherent search statistics and discusses their properties in the presence of noise. Sec. 5.3 summarizes the structure of the search and details of PSO, including the changes made to adapt it to multi-peak optimization. Sec. 5.4 details the analysis pipeline, including the waveforms used, the LISA response, hardware acceleration etc. In Sec. 5.5 the sampling distributions of the semi-coherent search statistic discussed in Sec. 5.2 are verified, additionally a probability-probability plot is computed to check the statistical properties of the semi-coherent likelihood. The results of several successful mock searches are presented in

Sec. 5.6. Importantly, Sec. 5.6 also includes a discussion of how the false alarm probability (FAP) of a stochastic search can be estimated. Sec. 5.7 discusses the computational costs and hardware used to carry out the analyses in this chapter. Sec. 5.8 contains concluding remarks.

5.2 Search Statistics

This section briefly summarizes and presents the key expressions for the semi-coherent data analysis quantities used in this study. The coherent versions of these quantities have already been introduced in Chapter 3 and some of the semi-coherent quantities are discussed in our previous work in Chapter 4. This section builds on what was introduced in Ref. [149] and more details and further motivation for these definitions can be found there.

5.2.1 Coherent search statistic

Throughout this chapter, the analysis is conducted in the frequency domain. All time series (both models and data) are assumed to be represented in the frequency domain, unless explicitly stated otherwise. The dependence on frequency is suppressed in order to avoid clutter in our notation. Within this study we also perform analyses in the noise-orthogonal time delay interferometry channels: A, E and T [224, 141].

Here it is assumed that the data contains just a single source within the current region of parameter space being searched. This is likely to be a good assumption for SmBBH where the expected number of LISA sources is low [25]. Furthermore, for broadband, long-duration signals like SmBBHs that contain a large number of orbits, the overlap between different signals is usually extremely small and can be safely neglected.

The aim of a GW search is to detect a source, provide a rough estimate of the parameters, and estimate its statistical significance, e.g. by calculating its false alarm probability (FAP). This information can then be used by a subsequent parameter estimation phase of

the pipeline which refines the inference of θ_* by sampling from the full Bayesian posterior distribution $P(\theta|d) \propto L(d|\theta)\pi(\theta)$. In the search phase, the aim is not to perform sampling but instead to maximize a given detection statistic. Most detection statistics are based on the matched filter discussed in Sec. 3.1.1. Searches performed using data from the current ground-based detectors use variants of the matched filter, which efficiently maximize over parameters such as the time of merger and phase, reducing dimensionality of the parameter space (and thus the volume) for the search [225].

The matched filter (and derived quantities from this such as the Whittle likelihood) is referred to as a *coherent* quantity. This is because it are defined in terms of a single integral of the smooth waveform model $h(\theta)$ over all values of frequency. In the next section semi-coherent versions of these quantities are shown.

5.2.2 Semi-coherent search statistic

In Chapter. 4, we demonstrated an early proof-of-concept for a SmBBH search using a likelihood function which used a semi-coherent matched filter, in this following section we recall the key expressions from our previous study and analytically study how a semi-coherent search statistic should change in the presence of noise.

Our search will maximise a semi-coherent matched-filter statistic. In this chapter we use the semi-coherent search statistic proposed in Ref. [179], which is the natural semi-coherent generalization of the coherent matched filter:

$$\Upsilon_N(d,\theta) = \sum_{n=0}^{N-1} \frac{x_n^2}{[h(\theta)|h(\theta)]_n^N}, \tag{5.1}$$

$$\text{where} \quad x_n = \max_{\phi_n} [d|h(\theta)e^{i\phi_n}]_n^N.$$

Where $[a|b]_N^n$ is the semi-coherent inner product as defined in Eqns. 4.3 and 4.4. The functional form of this definition can be understood by recognizing that each term in the summation corresponds to a phase-maximized $\tilde{\rho}^2$ in that segment.

In our notation semi-coherent quantities are denoted either with a hat (e.g. $\hat{L}$) and/or

with a subscript capital N (e.g. Υ_N) which also indicates the number of segments used. The number of segments measures the adjustable level of coherence used in the analysis. If $N = 1$ then we recover a fully coherent quantity. For example, $\Upsilon_{N=1}$ is equal to the square of the phase-maximized value of the coherent detection statistic $\tilde{\rho}^2$.

There is some freedom in how the segment frequency boundaries f_n (for $n = 1, 2, \ldots, N-1$) are chosen. As discussed in Sec. 4.2, possible choices include spacing these uniformly (or log-uniformly) in frequency. However, here the segment boundaries are chosen such that each of the N segments contains an equal amount of squared SNR. As a consequence of this choice, the segment boundaries are themselves functions of the source parameters θ; changing the parameters of a source effects not only the total SNR, but also how SNR accumulates over the frequency band. Therefore, f_n is reevaluated for every new value of θ every time a semi-coherent quantity is to be evaluated. This differs from our previous work where the segment boundaries were fixed [149]. For a linearly chirping quasi-monochromatic source (i.e. $\dot{f}_{\mathrm{GW}} \approx \mathrm{const} \ll f_{\mathrm{GW}}/T$), and in the case of a white instrumental noise PSD (i.e. $S_\alpha(f) \approx \mathrm{const}$), choosing the segment boundaries based on an equal squared SNR per segment can be shown to be equivalent to splitting the signal in N equal length segments in the time domain (see Appendix B.1).

5.2.3 Sampling distributions

The search statistics (both coherent $\tilde{\rho}(d, \theta)$ and semi-coherent $\Upsilon_N(d, \theta)$) are functions of the noisy data d and are therefore random variables. The random element arises from the inclusion of noise in d, due to this the value of $\Upsilon_N(d, \theta)$ or $\tilde{\rho}(d, \theta)$ will vary at a fixed θ if the noise realization in the data (d) changes. If they are to be used in a search, it is important that the sampling distributions of these statistics (i.e. their probability distributions) are properly understood as these underpin the false alarm and detection probabilities for the search. The sampling distributions need to be known both in the case when the data contains only noise, i.e. when $d = n$, and also when the data contains the signal being searched for, $d = n + h(\theta)$.

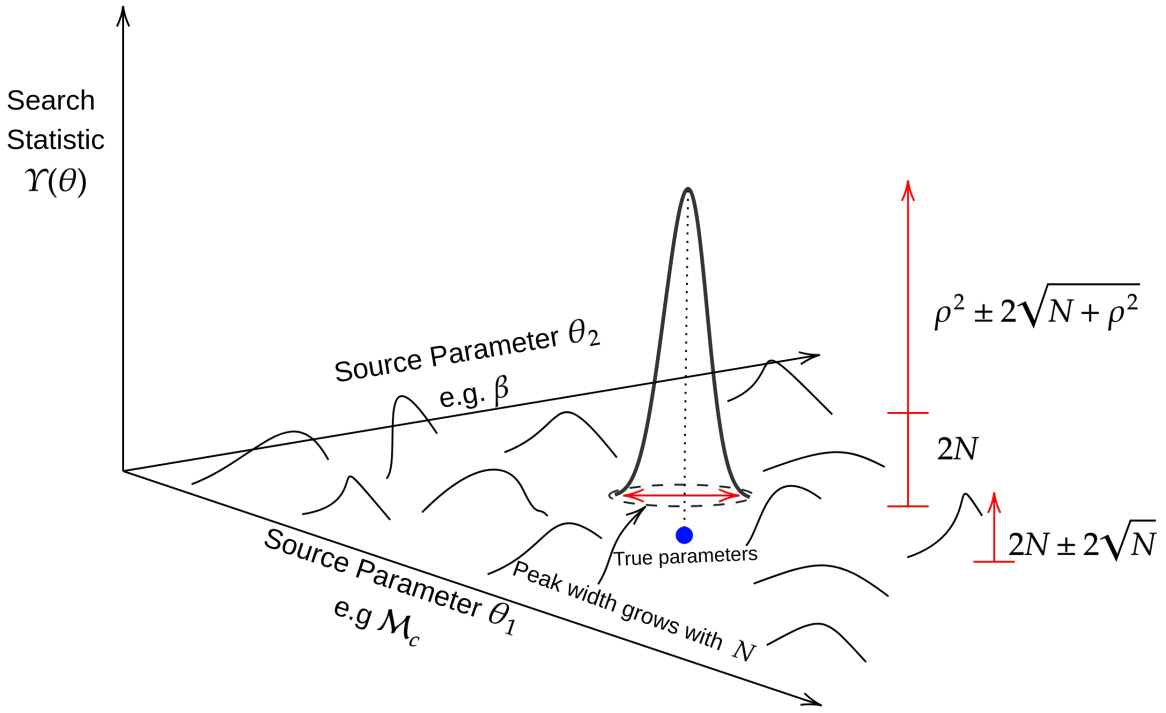

Figure 5.1: Sketch illustrating the variation of the semi-coherent search statistic $\Upsilon_N(d,\theta)$ with N segments over the parameter space θ. The statistic peaks near the true source parameters, θ_*, where the expected value $(2N+\rho^2)$ and standard deviation $(2\sqrt{N+\rho^2})$ of $\Upsilon_N(d,\theta_*)$ depends on the optimal squared SNR of the source and the number of semi-coherent segments, N. These values for the expectation and variance, given in Eqn. 5.3, are illustrated by the long red arrow. Away from the true parameter values, the search statistic fluctuates due to chance partial overlaps between the random noise and template. The values of the expectation $(2N)$ and standard deviation $(2\sqrt{N})$ of the statistic away from the true source parameters, given in Eqn. 5.2, are illustrated by the short red arrow.

It is well known that the standard matched filter $\tilde{\rho}(d,h)$ is normally distributed with unit variance and a mean given by the optimum SNR ρ (if there is no signal present, then the mean is zero). If the coherent detection statistic is maximized over phase (i.e., $\phi_{\max}(\tilde{\rho})$) then it can be shown that this statistic follows either a Rayleigh (in the only-noise case) or a Rice distribution (signal case) [226]. Since Υ_1 is equivalent to $\phi_{\max}(\tilde{\rho})^2$, it follows a standard chi-squared distribution with 2 degrees of freedom χ^2_2 (in the only-noise case) or a non-central chi-squared distribution with 2 degrees of freedom and a non-centrality parameter $\lambda = \rho^2$, $\chi^2_2(\lambda)$ (in the signal case), both distributions have unit scale parameter.

It is not possible to derive the corresponding analytic distributions for the semi-coherent search statistic Υ_N for finite N. However, in the limit of many segments, the

following approximations to these distributions were derived in Ref. [179] using the central limit theorem (CLT). In the presence of noise only, it was shown that approximately

$$\Upsilon_N \sim \mathcal{N}(\mu_0, \sigma_0^2), \tag{5.2}$$
$$\text{where } \mu_0 = N\mu_k \text{ and } \sigma_0^2 = 2N\sigma_k^2.$$

And in the presence of noise and signal, it was shown that approximately

$$\Upsilon_N \sim \mathcal{N}(\mu_1, \sigma_1^2), \tag{5.3}$$
$$\text{where } \mu_1 = A^2 + N\mu_k \text{ and } \sigma_1^2 = 2N\sigma_k^2 + 4A^2,$$

where $(\mu_k, \sigma_k) = (2.00, \sqrt{2})$ and $A = \rho$. These are referring to the expected distributions of Υ_N when it is evaluated with many independent realizations of the noise in the data-stream. Because these are based on the CLT, they are expected to be a good approximation in the limit of large N. In Ref. [179], these approximate distributions were tested by drawing and combining random samples from the analytic Rayleigh and Rice distributions on which they based and Eqs. 5.2 and 5.3 were found to provide good approximations for $N \gtrsim 70$.

In this chapter, these approximate distributions are tested in a more realistic way by simulating mock LISA time series containing Gaussian noise with and without signals. The numerical results are described in Sec. 5.5.1, the approximate distributions derived from the CLT are shown to be a good approximation even for N as a low as 10.

These probability distributions allow us to calculate the FAP of the search statistic at a fixed point in parameter space. However this is not indicative of the FAP for the whole search as that covers a large region in θ. The FAP is calculated by simulating searches over a large number of noise and noise + waveform data streams [225, 227]. A similar method is used to make a limited estimate of the FAP for this search in section 5.6.

At a higher number of segments, lower density of templates covering θ is traded for a higher FAP. Practically the higher FAP translates to the 'noise' maximas in Υ_N being of a

greater height and more comparable to the 'true' peak; this is described by equations 5.2 and 5.3 which is illustrated by Fig. 5.1. Since the height of the 'true' peak is proportional to the SNR of the injected signal, the noise peaks can be problematic for signals at low SNR. This is relevant for the SmBBH inspirals as these sources are expected to be quiet [166, 103]. This limits the maximum number of segments we can use for the search; using a very high number of segments increases the height of the noise peaks, thereby increasing the chance that a search may mistake such peaks in Υ_N as a signal.

5.3 Search Structure

The aim of the search is to maximise the matched-filter statistic with respect to θ. In LVK searches for compact binary coalescence signals a template bank approach is used. This involves generating a set of waveforms $\{h_0, h_1, \ldots, h_{N-1}\}$ at predetermined locations in parameter space $\{\theta_0, \theta_1, \ldots, \theta_{N-1}\}$, and computing the (coherent) inner product between the data and every template in the bank (maximized over phase and time offsets). The search identifies the template with the largest value of the statistic. As discussed earlier, for long-lived signals such as LISA SmBBH inspirals this procedure is not possible due to the number of templates that would be required.

Our proposed solution to this problem involves three changes with respect to the LVK searches. The first is the switch from the coherent to the semi-coherent inner product described in Sec. 5.2.2. The second is the hierarchical nature of the search where the coherence is gradually increased. Finally, the third is the use of a stochastic optimization method to maximize the detection statistic, instead of the deterministic template bank; this is described in this section.

This section describes the broad conceptual features of the search. Technical details specific to this particular implementation are presented in Sec. 5.4.

5.3.1 Particle swarm optimization

Following what is done in our previous work, we use PSO as the stochastic optimization algorithm to maximize the semi-coherent search statistic Υ_N [174]. For a detailed discussion of PSO, see Sec. 4.5.

The search proceeds in a hierarchical manner, with the level of coherence demanded gradually increasing as the swarm converges. This is achieved by decreasing in stages the number of segments N used in the objective function Υ_N. At each hierarchical stage, Υ_N is optimized until the swarm has reached a convergence criteria (see Appendix B.2) at which point the objective function is changed to $\Upsilon_{N'}$ (where $N' < N$). This sequence of decreasing N values is referred to as a ladder. Optimization at each level is repeated until the final step in the ladder ($N = 1$) has been reached and optimized. The coherent search statistic ($\Upsilon_{N=1}$) has the lowest FAP out of all levels in the ladder.

5.3.2 Multi-Peak tracking

As discussed in Sec. 5.2.3, using a high number of segments increases the FAP of a search using the statistic Υ_N. However, a large N is needed to efficiently find and localise signal peaks early in the search, especially when using broad priors.

The hierarchical search method presented here aims to solve this problem by starting with a large number of segments (initially accepting the associated high FAP) and tracking multiple source candidates through a sequence of decreasing values of N. As false positive candidates identified early in this process are tracked, the peak in the Υ_N surface becomes less significant (i.e. consistent with the background noise distribution) as N is decreased; this allows the false positive candidate to be vetoed later in the process. Although the hierarchical search may have a high FAP during the early stages, spurious sources are eliminated by the time $\Upsilon_{N=1}$ has been reached. Since $\Upsilon_{N=1}$ has similar FAP to $\tilde{\rho}$ for long signals, any signal that has a low FAP in $\Upsilon_{N=1}$ can be treated as being a statistically significant source (at the same level as if it were a trigger from searching using $\tilde{\rho}$).

To do this we need an optimizer that can track a large number of optima throughout the search, accounting for the possibility of changing (both increasing and decreasing) the number of optima being optimized. We do this using an implementation of a multi-swarm PSO algorithm. Similar ideas have been explored for multi-objective optimization in [176]. The optimizer is initialized with an (arbitrary) number of swarms each containing the same number of particles. Each swarm is evolved using the usual PSO rule described in Sec. 4.5. Once each swarm has converged on the $\Upsilon_{N_{\rm high}}$ surface, k-means clustering is run on the positions of all particles on all swarms to (re)group them. This allows for the merging of swarms which are optimizing the same peak or for the creation of new swarms when additional peaks are detected. The total number of particles is conserved throughout the search. The search statistic is then changed to $\Upsilon_{N'}$ ($N' < N$) as we move down a rung of the ladder. The search and re-clustering processes repeat until all swarms have converged on the $N = 1$ surface. The positions and $\Upsilon_{N=1}$ function values at the end of the search are the main results.

The use of multiple-swarms and k-means clustering is a significant development compared to our earlier work. See Appendix B.2 for the values of all the multi-swarm PSO algorithm hyper-parameters used for the search.

Note that PSO is not invariant with respect to a change of parameters $\theta \to \theta'(\theta)$. In particular, PSO is not invariant with respect to rotations in the parameter space. This can lead to undesirable features in the particle locations when optimizing simple, approximately Gaussian-looking peaks; see, for example, Fig. 2 of Ref. [228] and Ref. [229] for a discussion. We have found that this can be problematic for our search especially when Υ_N has non-linear correlations in parameter space. At the moment, this has been overcome simply by increasing the total number of particles used. However, we note there exist rotationally invariant versions of the PSO algorithm (e.g. [230]), these may help alleviate these problems. However, these algorithms are more complicated and have additional tuning parameters which need to be optimized. We do not investigate such algorithms here.

5.3.3 Transition to MCMC inference

The final aim of a search process is to provide rough estimates for the source parameters of the detected signals, these can then be used to 'seed' parameter estimation. At this early stage of development, we do this by directly initializing MCMC particles using some particles from the final state of the search. After the final search stage, a small number of particles from each swarm (that has passed a detection threshold) are extracted and used as initial positions for MCMC walkers. The posterior distribution sampled at this final stage uses the phase-maximized coherent likelihood $\log \hat{L}_1$. We have verified that for long-lived signals this produces posteriors extremely similar to the standard coherent likelihood (See Sec. 4.3).

Each PSO particle that transitions into an MCMC walker draws a luminosity distance D_L from the prior, the search statistic analytically maximizes over this parameter so it must be drawn for each particle (or walker) at the point of transition.

We use `Zeus` to sample the posterior associated with $\log \hat{L}_1$. `Zeus` is an MCMC sampler which uses ensemble slice sampling to efficiently traverse highly correlated parameter spaces [231, 232]. We found this to be particularly well-suited to some of the non-linear degeneracies and correlations present in the likelihood of SmBBH signals.

5.4 Search implementation details

In this section the technical details of the search and analysis pipeline as implemented here are presented. This includes details about the waveform, instrument response and hardware acceleration used in this study.

The `TaylorF2Ecc` waveform was used for injection, search and parameter estimation [233]. This is a frequency domain, post-Newtonian (PN) waveform approximant for the $(l, |m|) = (2, 2)$ mode, which includes contributions from the orbital eccentricity to the GW phase. This PN waveform is expected to be sufficiently accurate for the search and recovery of the early inspiral signal [200]. Binaries merging in the LIGO/VIRGO frequency

band are expected to have radiated away much of the orbital eccentricity, however the early inspiral of these systems – observable in LISA – may retain eccentricity [234], which contributes significantly to the waveform phase. This `TaylorF2Ecc` waveform neglects the BH spins. The component spins of SmBBH systems will not be well measured by LISA, although the aligned spin contribution to the GW phase may be moderately significant [163] and should be included in future work. This waveform has recently been widely used in the context of LISA; see, for example, Refs. [34, 235], similar waveforms are also used in Refs. [167, 223].

Early parameter estimation studies for SmBBH signals in LISA have mostly been limited to the case of zero noise. This is partly because the inclusion of noise makes the integrand for the $\langle d|h\rangle$ term in the likelihood a non-smooth function of frequency, which renders common methods of numerically evaluating this integral, such as quadrature integration, invalid; for a detailed discussion see Appendix A.4. We are forced to evaluate the waveform and sum in Eqn. 3.8 over the dense fast-Fourier transform (FFT) frequency grid. The number of points in this grid is $\sim [f_{\rm max} - f_{\rm min}]T_{\rm obs}$. Since SmBBH sources are broadband ($f_{\rm max} - f_{\rm min} \sim 10^{-1}\,{\rm Hz}$) and long lived ($T_{\rm obs} \sim$ years), this is typically a sum over millions of points. For other source types, such as MBBH (short duration, with $T_{\rm obs} \sim$ weeks) and DWD (narrow band, with $f_{\rm max} - f_{\rm min} \sim 10^{-6}\,{\rm Hz}$), the number of frequency points is much smaller.

A common method of accelerating the waveform evaluations is through the use of interpolation, evaluating the waveform on a sparse frequency grid and then interpolating onto the denser frequency grid where inner product or likelihood is computed. Similar approaches have previously been used in, for example, Refs. [162, 236, 237]. A decomposition of the waveform into amplitude and phase, in the form $h(f) = A(f)e^{i\phi(f)}$ is convenient for interpolation. The amplitude and (unwrapped) phase are smooth functions of frequency, making them easy to interpolate. Evaluating this interpolant onto the FFT frequency grid with millions of points is still an expensive operation which is here accelerated by using a GPU. We use the `BBHx` GPU-based implementation of cubic-spline interpolation

[236, 237].

The amplitude-phase decomposition is also a convenient representation to apply the LISA instrument response to the waveform. The LISA response can be represented using transfer functions which operate on each frequency mode and simulate the instruments response in the noise-orthogonal TDI channels $\alpha \in \{\mathrm{A}, \mathrm{E}, \mathrm{T}\}$ [224, 141].

$$h_\alpha(\theta, f) = \mathcal{T}_\alpha(\theta, f)\tilde{h}(\theta, f) \tag{5.4}$$

We use the `BBHx` implementation of the frequency domain LISA response, which performs interpolation for the transfer functions in a similar way [138]. The sparse frequency grid over which both the response and waveform are evaluated, is a downsampled version of the FFT frequency grid, by a factor of 1000.

The formalism used to derive the response breaks down in the limit of extremely slowly chirping signals. However, for SmBBH sources the response breaks down at $t_c \gtrsim 10 - 100$ years (see Fig. 3 from [138]), all sources considered in this chapter merge faster than this. It is expected that the large majority of SmBBH systems ($\sim 90\%$) merge with $t_c < 10$years, see Fig. 3 of Ref. [238].

The computation of all quantities related to signal inner products that involve summation over the FFT frequency grid (including $x_n(d, \theta)$, Υ_N and $\log\hat{L}_N$) were also GPU accelerated, significantly reducing the computational cost of both search and inference.

Throughout this work it has been assumed that the PSD that characterizes the noise spectrum in the data is known. The analytical approximation to the instrumental noise from Ref. [210] was used. The confusion noise arising from the unresolved galactic DWD population was not included as it is expected to be negligible (i.e. smaller than the instrumental noise) at frequencies $f_{\mathrm{GW}} \gtrsim 10^{-2}\,\mathrm{Hz}$ [13] where all the sources considered in this chapter reside. The noise is assumed to be Gaussian and is generated directly in the frequency domain from the known PSD. The simulated data stream has no gaps or glitches. Extending the search to data streams with these properties is left for future

work.

The actual LISA noise spectrum is expected to drift in time and to contain significant cyclostationary components associated with orbital modulations of the confusion noise [56]. However, the confusion noise dominates at lower frequencies $f \lesssim 3 \times 10^{-3}\,\mathrm{Hz}$; at the higher frequencies where the sources considered in this study reside, while the noise is not expected to be perfectly stationary, it will vary much less with time.

In this chapter the search is performed on data sets containing signals from two different binaries (a GW150914-like and GW190521-like source [1, 154]) injected at multiple different SNRs (controlled by the choice of the D_L parameter). Both sources are placed closer than the distance at which they were observed by the LVK in order to have a detectable SNR in LISA. We also arbitrarily pick the time to merger (controlled by the choice of the f_{low} parameter) such that these systems coalesce within ~ 4 years (i.e. within the nominal LISA mission lifetime). The fiducial "loud" source in this chapter is similar to GW150914, placed at a very close distance of 50Mpc[1]. In Sec. 5.6 we vary the distance to this source to evaluate the performance of this search at different SNRs. We simulate searches for signals at a range of SNRs, spanning the range $\rho \in [15, 38]$, focusing more on the subset that span $\rho \in [15, 22]$. While this SNR range is on the upper end of what LISA may observe, it is not extreme, see Fig. 6 of Ref. [238].

The injected source parameters and search priors on those parameters for the fiducial GW150914-like source are shown in Table 5.1, the two parameters not directly searched over are the luminosity distance D_L and the initial orbital phase ϕ_0. A similar table for GW190521 is shown in Appendix B.3. The luminosity distance adds a $1/D_L$ pre-factor to the waveform amplitude and is cancelled out of both $\tilde{\rho}$ and Υ_N. While in the definition of Υ_N, the phase is analytically maximized so it does not need to be directly searched over. After the search has concluded and $\Upsilon_{N=1}$ has been optimized, in the process of

[1]Ref. [239] demonstrated that using the LVK observed population to constrain LISA stellar-mass binary detection rates leads to a redshift limit of $z \lesssim 0.1$ corresponding to $D_L \lesssim 400\,\mathrm{Mpc}$, See Fig. 3 in Ref. [239]. While 50 Mpc is on the lower end of this range and a source at this distance would definitely be rare, it is not expected to be an extreme statistical anomaly among the detectable population, see Fig. 10 in [239], $z \approx 0.01$ has significant probability.

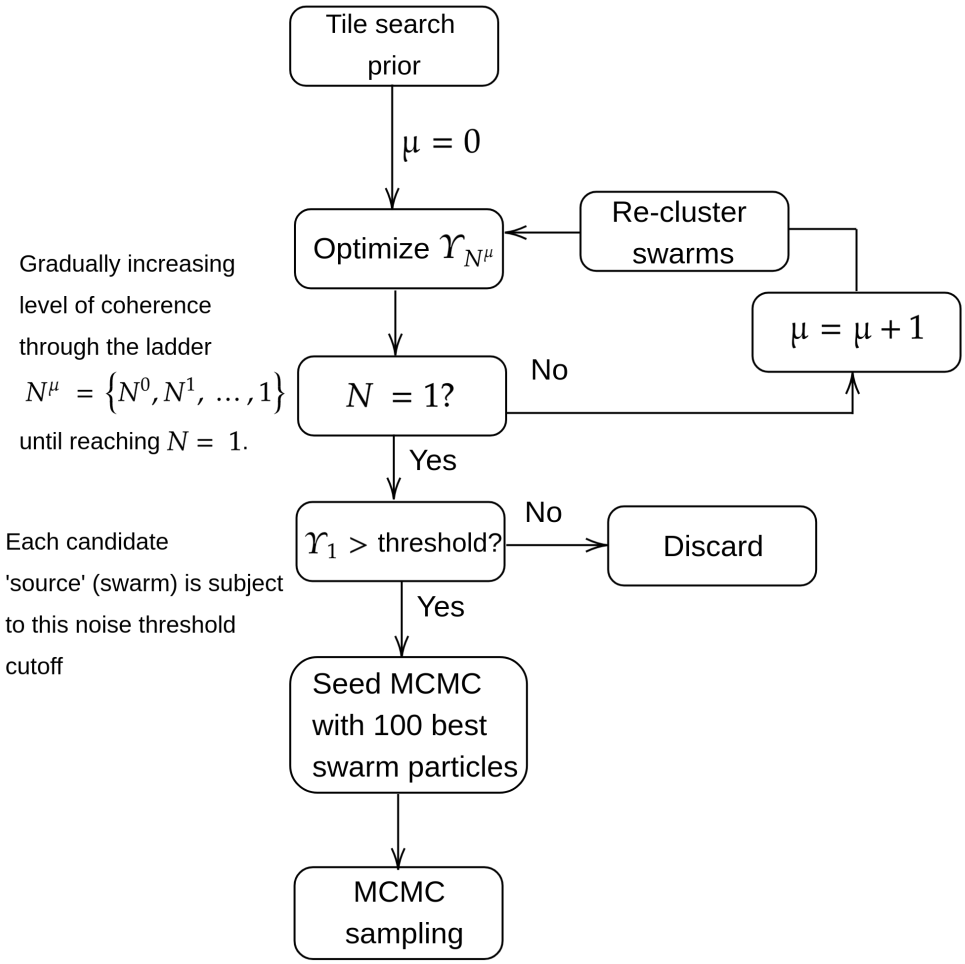

Figure 5.2: Schematic flowchart for the search pipeline within one $\mathcal{M}_c - t_c$ tile. The noise threshold used in this study for each swarm is $\Upsilon_1 \approx 100$, see top panel of Fig. 5.5.

automatically transitioning to inference, D_L is drawn from the prior in Table 5.1 for each PSO particle which switches behavior to an MCMC walker. Since in the inference stage $\log \hat{L}_1$ is being sampled, ϕ_0 is still maximized over.

We find empirically that this method is not able to search over the entire astrophysically reasonable parameter space for SmBBH signals at once. This is simply because, if the finite number of particles used here are spread too thinly across parameter space then the probability that a large peak is missed increases. Swarms of particles must be initialized with sufficient density over parameter space to find a peak in the search statistic and climb it, thus the search space needs to be broken into multiple 'tiles' each of which are searched independently. This can be done in parallel. We divide up the search

Table 5.1: Injection parameters and priors for the search of the GW150914-like fiducial source. Parameters in the top section of the table are those that are searched using uniform priors over the ranges shown, those in the middle section are injected but do not appear in the definition of the search statistic, and those in the bottom section are *derived* parameters and approximate prior ranges are given. At the point of transition between search and parameter estimation, each particle draws D_L uniformly from the prior range $[10, 200]$ Mpc. All masses are given in the detector frame. (Due to differences in the conventions between the waveform and the BBHx LISA response, the definition of the polarisation angle ψ acquires a negative sign.) Parameters that define the search tile are in highlighted rows. This injected source has an SNR of 38.

Parameter	Injection	Prior range: $[\theta_{\min}, \theta_{\max}]$
$\mathcal{M}_c\,[M_\odot]$	28.095555	$[27, 30]$
$f_{\rm low}$ [Hz]	0.018	$[0.0178, 0.0182]$
η	0.2471	$[0.15, 0.2495]$
λ [rad]	2.01	$[0, 2\pi]$
β [rad]	$\pi/4$	$[-\pi/2, \pi/2]$
i [rad]	2.498	$[0, \pi]$
ψ [rad]	-1.85	$[-\pi, 0]$
e_0	0.01	$[0.005, 0.1]$
ϕ_0 [rad]	0	-
D_L [Mpc]	50	-
$m_1\,[M_\odot]$	36	$\sim [32.44, 76.43]$
$m_2\,[M_\odot]$	29	$\sim [15.49, 32.96]$
t_c [months]	43.28	$\sim [38, 46]$

space into rectangular tiles in the $\mathcal{M}_c$ and t_c parameters. We have found empirically a tile dimension of $\sim 3\,\mathrm{M}_\odot$ and 6 months to be effective. Tiling the astrophysically reasonable search space of $\mathcal{M}_c \in [10, 100]\,\mathrm{M}_\odot$ and $t_c \in [0.5, 10]$ years using tiles of these dimensions would require around ~ 600 tiles. This is a rough estimate, and assumes each tile has similar dimensions, in reality for the high mass end of parameter space the tiles can be made bigger in both dimensions, since these undergo less orbital cycles. Fig. 5.2 shows a high level overview of the search process within one tile.

5.5 Properties of the search statistic

In this section the statistical properties of Υ_N and $\log \hat{L}_N$ are tested. In Sec. 5.5.1 we validate the theoretical sampling distributions of Υ_N under both only noise and noise plus

signal for various values of N. Then, in Sec. 5.5.2, the statistical properties of $\log \hat{L}_N$ are tested using probability-probability plots computed over many noise realizations.

5.5.1 Sampling distributions of the semi-coherent search statistics

We seek to validate the distributions of the random variables Υ_N given in Sec. 5.2.3, by computing Υ_N over many noise realizations and comparing the empirical distributions against those derived from the CLT (see Eqs. 5.2 and 5.3). We compare the histograms of these empirically obtained random variables against the theoretical probability density function. The GW source used for all computations in this subsection is the GW150914-like fiducial source in Table 5.1. The luminosity distance to the source is scaled to simulate signals with $\rho \sim \{10, 20, 30\}$.

The well known Gaussian sampling distributions for $\tilde{\rho}$ are checked in the top panel of Fig. 5.3. In the noise only case $\tilde{\rho} \sim \mathcal{N}(0, 1)$, whereas in the noise+signal case, $\tilde{\rho} \sim \mathcal{N}(\rho, 1)$ where ρ is the optimal SNR of the signal. In Fig. 5.3 we plot $\tilde{\rho}^2$ since it is directly comparable with Υ_1. In the noise only case $\tilde{\rho}^2 \sim \chi^2_1$, a standard chi squared distribution with 1 degree of freedom. While in the noise+signal case $\tilde{\rho}^2 \sim \chi^2_1(\lambda, 1)$, a non-central chi squared distribution with non-centrality parameter $\lambda = \rho^2$ and unit scale parameter. The bottom panel of Fig 5.3 verifies the distributions given in Sec. 5.2.3 for Υ_N at $N \in \{1, 10, 50, 100\}$. The analytic approximations from the CLT are used as the theoretical distribution for $N \geq 10$, while for the Υ_1 case these are the theoretical χ^2 distributions derived in the same section. The increased FAP of Υ_N as a function of N can be seen in Fig. 5.3 as the overlap between the two distributions $\Upsilon_N(d = h + n, \theta)$ and $\Upsilon_N(d = n, \theta)$ increases with N. This is most clearly seen when comparing the black and red curves. The distributions of Υ_1 and $\tilde{\rho}^2$ are similar, indicating that either can be used as a good detection statistic. In our search, the significance of a trigger is determined at the final stage, when $N = 1$, using this statistic.

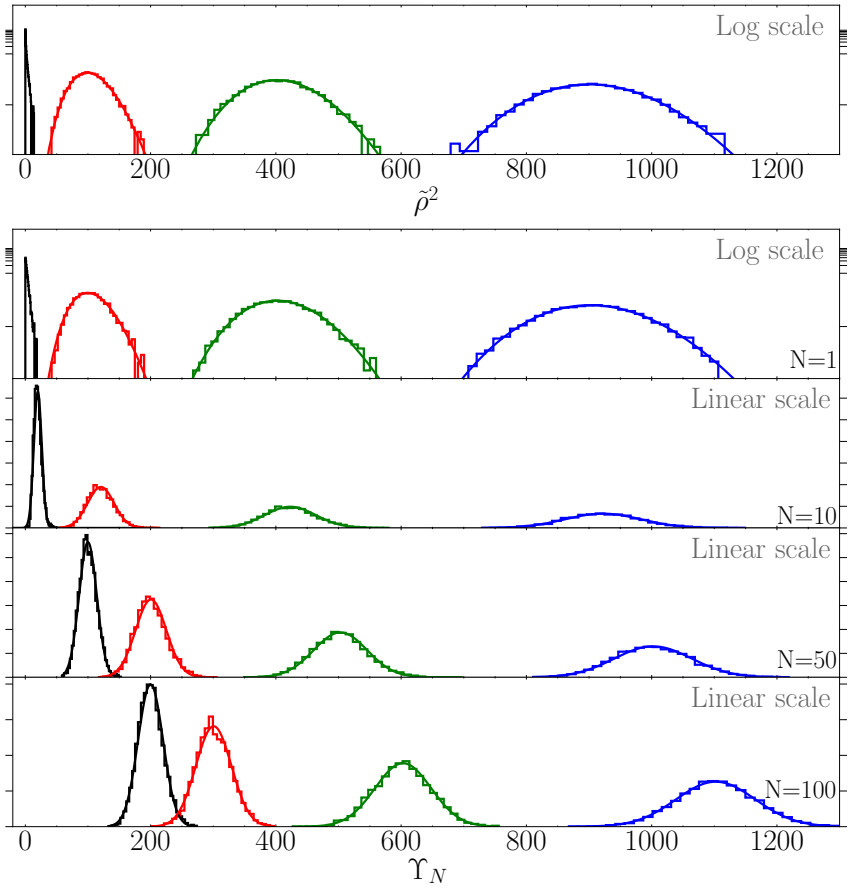

Figure 5.3: Distributions of the coherent $\tilde{\rho}^2$ (*Top panel*) and semi-coherent Υ_N (*Bottom panel*) matched filter statistics, sampled in the noise only ($d = n$; black) and signal ($d = h$) cases, and for varying injected SNRs: 10 (red), 20 (green) and 30 (blue). Each vertical panel represents a different number of segments N. The top panel plots $\tilde{\rho}^2$, instead of simply $\tilde{\rho}$, as this quantity is more comparable with Υ_1. The plots for the coherent $\tilde{\rho}^2$ and Υ_1 (top two plots) use a log-scaled y-axis while the lower three plots use a linear scale. The signal used for injections in this plot is the GW150914-like loud fiducial source from table 5.1. The analytic predictions for the distributions are plotted using smooth lines (except for the $\Upsilon_{N=1}$ and $\tilde{\rho}^2$ cases as they diverge at 0). In all cases 5000 independent noise realizations were used to construct the histograms.

5.5.2 PP-type plots for the semi-coherent likelihood

Probability-probability (PP) plots are a common statistical test in GW astronomy used to validate many ingredients in parameter estimation pipelines, including the likelihood function (see, for example, Ref. [104]).

Noiseless parameter estimation runs with appropriate priors converge on a posterior where the peak of the posterior distribution is at the injected parameters. The spread of

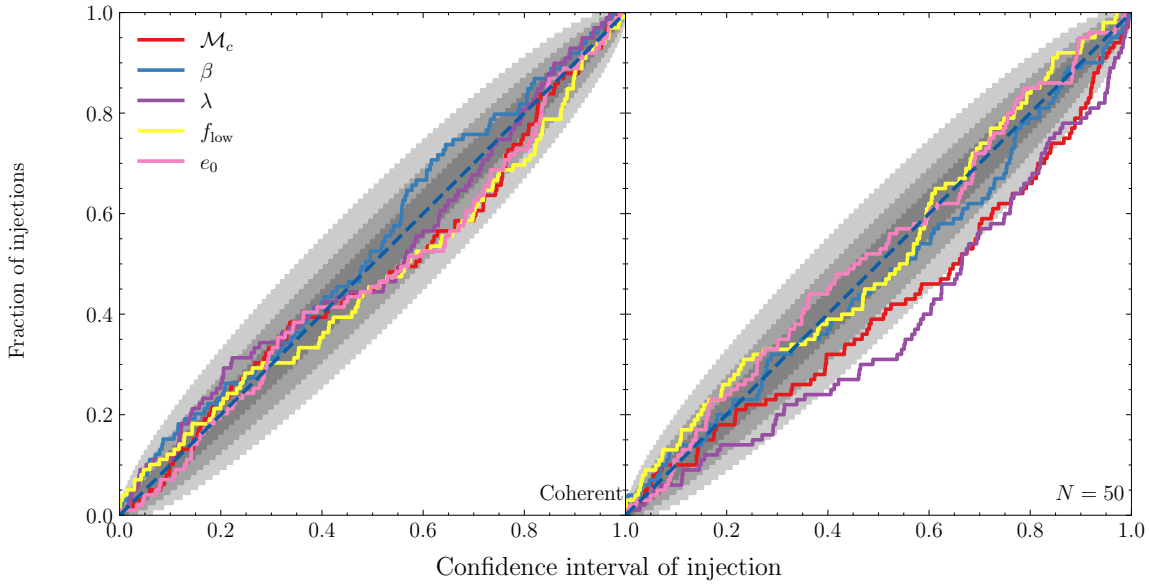

Figure 5.4: Probability-probability plot for the coherent (left panel) and $N = 50$ semi-coherent likelihoods (right panel), computed over 100 injection-inference runs of the GW150914-like source into simulated Gaussian noise. Shaded regions denote the 1, 2 and 3σ confidence intervals. The 5 colored curves show the 1-dimensional PP plots for the key phasing parameters in the legend. PP plots are shown for chirp mass ($\mathcal{M}_c$), ecliptic latitude (β), ecliptic longitude (λ), initial orbital frequency ($f_{\rm low}$) and initial eccentricity e_0.

the posterior distribution is expected to be consistent with shifts in the posterior peak due to noise in the data, i.e. if the no-noise posterior has 20% of the probability mass below a given θ_*, over many noise realizations, we expect that 20% of the time the peak will satisfy $\theta < \theta_*$. Measuring the shift between the peak of the noiseless posterior and that which includes noise is equivalent to measuring the confidence level (on the noise included posterior) on which the true injected parameters lie. A PP plot tests this over many noise realizations, by plotting the fraction of independent noise realisations that have the injected parameters at a given confidence level; in the limit of infinite noise realizations, this should tend to a perfectly diagonal line.

Here, this is used to check the properties of the semi-coherent likelihood $\log \hat{L}_N$. Peaks of the semi-coherent likelihood are (by design) wider than corresponding peaks in the usual (coherent) likelihood, this makes them easier to find in a search. It is also expected that the semi-coherent likelihood is more susceptible to noise fluctuations and that the width

of the peaks should be consistent with this increased scatter. If this is true, then the PP plot should remain approximately diagonal.

Each data-point in the PP plot corresponds to one full parameter estimation run conducted with either a coherent or semi-coherent likelihood. The fiducial GW150914-like signal is used for all parameter estimation runs, each data-stream contains an independent realization of the noise. Parameter estimation is performed on each data-stream with narrow priors using the `dynesty` nested sampler [195]. The 1D confidence interval the injection resides on is defined as

$$X = \int_{\{x|P(x)>P(x_*)\}} P(x)\mathrm{d}x, \tag{5.5}$$

where $P(x)$ is the one-dimensional marginalized posterior distribution on a parameter x and x_* corresponds to the injected value of that parameter. The discrete posterior samples from inference were fit with a Kernel Density Estimate (KDE), which provides a continuous approximation for $P(x)$. Then both $P(x_*)$ and $P(x_i)$ (posterior density for every posterior sample), can be estimated with the KDE. The integral required to compute the confidence level can then be approximated as a Monte-Carlo sum over the posterior samples:

$$X \approx \frac{1}{\hat{N}} \sum_{i=1}^{N} \mathbf{1}_{\{x|P(x_i)>P(x_*)\}}(x_i), \tag{5.6}$$

where $\mathbf{1}$ denotes the indicator function and $\hat{N}$ is the total number of posterior samples. This process is repeated over 100 independent parameter estimation runs to construct the PP plot. Finally, the fraction of injections that lie at various confidence levels are plotted on Fig. 5.4. The figure shows these results for two cases: a coherent and semi-coherent $N = 50$ posterior.

The parameter estimation runs used to populate Fig. 5.4 used the `dynesty` nested sampler instead of the MCMC sampler `Zeus`, this is because the automated parameter estimation can be seeded directly from the search, which provides good starting locations for the MCMC walkers, close to the maxima. Dynesty is only used to explore the likeli-

hoods in this section, it is not part of the search pipeline that is the main focus of this chapter.

We find that the diagonal property does approximately hold, see Fig. 5.4. The PP plots based on the semi-coherent likelihoods (e.g. $N = 50$) do exhibit a small "sag". We conjecture that this is due to the maximization (as opposed to marginalization) over the phase angles in each segment. However, this is a small effect (quantified by a maximum sag of $\sim 10\%$, see Fig. 5.4). But this effect is small enough so that our search is still able to track peaks accurately through the different hierarchical stages of the search.

It is worth highlighting the PP plots in Fig. 5.4 use only one source for injections (the GW150914-like source in Table 5.1) but with multiple noise realizations. Usually, PP plots should be generated with many sources drawn randomly from the prior. For computational reasons, here we use a single source at a fixed θ, this allows us to test the algorithm at a reasonable computational cost.

5.6 Search results

This section presents the results of a search for a number of GW150914-like sources injected at five SNRs (from 16 to 38, see legend of Fig. 5.5) into mock LISA data. We describe how to quantify the significance of detections by estimating the FAP. Additionally, Appendix B.3 presents similar results for the search of a GW190521-like source.

As described above, the search uses particle swarms to track multiple peaks in the semi-coherent search statistic Υ_N. The search is hierarchical in the sense that the number of segments N used is steadily decreased throughout the search. At the final coherent stage ($N = 1$), when the PSO algorithm converges to a number of peaks, this process stops. In a typical case, the number of peaks found by this process is ~ 10. However, many of these peaks do not correspond to real signals. The parameter values of these peaks are the candidate sources produced by the search. For each of our searches, the highest of these peaks (i.e. the one with the largest value of $\Upsilon_{N=1}$) is indicated by a solid

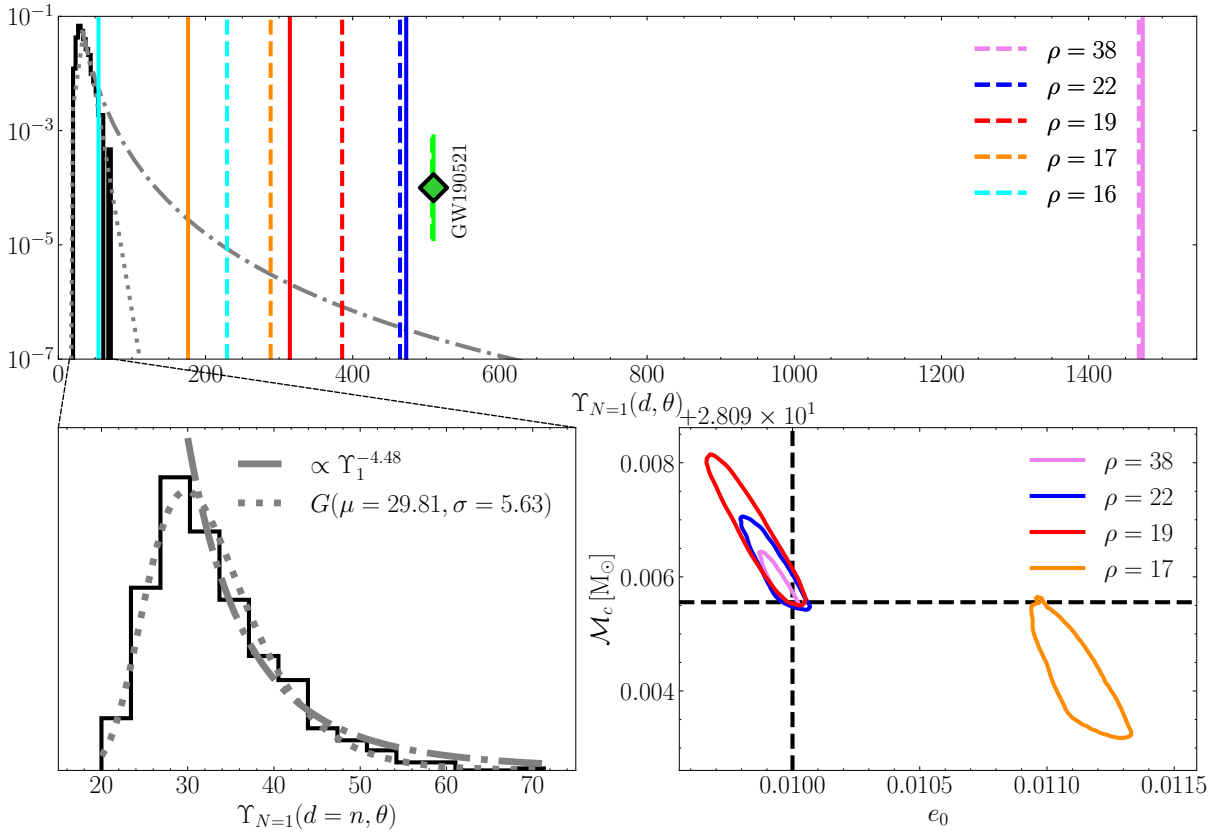

Figure 5.5: *Top:* The search result is the maximum value of the Υ_1 statistic, shown for each GW150914-like source (solid vertical colored lines). Also shown for comparison is the value of the search statistic evaluated at the injected parameters (dashed vertical colored lines). Also plotted is the background distribution of triggers in noise-only data (black histogram). The search result for a GW190521-like event is also shown as a short line, although this comes from a search in a very different region of parameter space and so this result is only intended to be qualitative, for the reasons explained below. Gray curves show extrapolations of the background noise distribution and can be used to estimate the FAP. Two such extrapolations are shown, these are intended to bracket the optimistic and pessimistic cases, and these are also shown in the bottom left panel. *Bottom left:* Distribution of background noise triggers, same as the black distribution plotted in the top panel. This histogram is constructed from 60 noise only data streams. The dash-dotted gray line indicates a conservative power-law fit to the upper tail of this distribution with PDF $\propto \Upsilon_{N=1}^{-4.48}$, while the dotted line indicates a fitted Gumbel distribution which decays faster than the power law and provides a more optimistic extrapolation of the distribution of background triggers at high Υ_1. This noise distribution was calculated from the search tile containing the GW150914-like source and therefore does not apply to the GW190521-like event. *Bottom right:* 90th percentile confidence contours on the chirp mass and eccentricity from the final MCMC stage of the search (not including the un-recovered $\rho = 16$ case). The characteristic lopsidedness of the posterior distribution is due to the degeneracy between $\mathcal{M}_c$ and η and the parameter space boundary at $\eta = 0.25$ (similar to that shown in Fig. 8 of [163]). For SNRs $\rho \geq 19$ the final MCMC stage of the search finds the true parameters within the 90th percentile confidence contour. However at the lower SNR $\rho = 17$ a significant bias is observed, this is because this is a crude MCMC.

vertical line in the top panel of Fig. 5.5.

If the injected signal is sufficiently loud, then the search succeeds in locating the source in parameter space. This can be seen from the results in Fig. 5.5. The top panel shows the highest value of the search statistic found by the particle swarms compared to the value of that statistic evaluated at the injected parameters; the optimized value is generally slightly higher due to noise fluctuations, similar to how the peak in the coherent log-likelihood is shifted from the true parameters when noise is present.

For SNRs $\rho \geq 19$ the search finds the correct source parameters. For SNRs $17 \leq \rho < 19$ the search often fails to find the exact peak in $\Upsilon_1(d, \theta)$, but it gets extremely close. In these cases the optimized value of the search statistic is less than the value at the injected parameters, and the parameters are somewhat offset from their true values. For lower SNRs $\rho \leq 16$ the search does not find the source.

It is necessary to assign a FAP to each of candidates found by the search. The FAP is determined by comparing the maximum optimized value of the $\Upsilon_{N=1}$ statistic found by the search to the noise background distribution. Unlike in the LVK case, where there are multiple independent instruments and large fractions of signal-free data, it is not possible to perform time slides of LISA data to determine this background. Instead, the background is computed by an injection campaign in which the search is performed on a large number of signal-free injections. These injections should ideally contain realistic LISA data but with no SmBBH sources present; here, for simplicity, 60 such searches were performed on time series containing only stationary, Gaussian noise. Exactly the same hierarchical, semi-coherent PSO search algorithm was applied to each of these signal-free injections and the peak values of the $\Upsilon_{N=1}$ statistics were recorded. This defines the noise background distribution of our search statistic and is plotted as a black histogram in the top and bottom left panels of Fig. 5.5.

The FAP for a given candidate with maximum optimized value of $\Upsilon^*_{N=1}$ is defined as the complementary cumulative density function of the noise background evaluated at $\Upsilon^*_{N=1}$. Similar techniques are used to determine the significance of candidates in LVK

Table 5.2: Estimates for the FAP of the searches for the different GW150914-like sources. In each case two estimates are provided, one using the power-law extrapolation of the noise background (conservative, i.e. higher FAP) and one using the Gumbel extrapolation (optimistic, i.e. lower FAP).

SNR ρ	FAP (lower)	FAP (upper)
16	7.4×10^{-2}	1.2×10^{-2}
17	1.3×10^{-3}	5.1×10^{-12}
19	1.6×10^{-4}	$< 10^{-12}$
22	4.1×10^{-5}	$< 10^{-12}$
38	7.8×10^{-7}	$< 10^{-12}$

searches [240, 241].

Given the noise background of 60 trials shown in Fig. 5.5 it is possible to bound the FAP of the candidates in the $\rho \geq 17$ searches as FAP $\lesssim 1/60$. In order to obtain better estimates of the FAP the background distribution needs to be extrapolated. This is done in two ways: first, a power law model was fit to the noise background above $\Upsilon_{N=1} \geq 30$, and secondly a Gumbel distribution was fitted to the entire background distribution. The power law represents a conservative extrapolation, the real distribution of triggers is expected to decay faster than this, see for example Fig. 4 of Ref. [1]. The Gumbel [242] distribution is more optimistic and decays faster (exponentially). The motivation for using the Gumbel distribution is that in the ideal case of Gaussian noise, $\Upsilon_{N=1}$ at a particular point in parameter space on the signal-free injections is known to be distributed as a χ^2 statistic. The maximum of many draws from a χ^2 distribution approaches the Gumbel distribution [242]. These two extrapolations are used to compute the FAP presented in Table 5.2.

The FAP numbers provided should just be considered as rough estimates. In addition to the extrapolation from a fairly small sample from the background distribution, computing an accurate FAP requires simulation of realistic signal-free injections. In this study, these were highly simplified in that they do not contain gaps, any other source types, glitches and the noise is stationary. However, the results of this study suggest that signals with $\rho \geq 20$ seem to be reliably detectable with this search pipeline while those at

$\rho \leq 16$ seem to be completely undetectable. This is roughly consistent with the detection probability of $\Upsilon_{N=100}$, which is expected to decay from 1 to 0 in the range $\rho \in [15, 20]$, see Fig. 7 in Ref. [179]. For intermediate SNRs, where the source is only detected some of the time, this detection probability is explored by performing multiple simulated searches with the results presented in Appendix B.4.

Once a significant candidate has been found, the final stage of the search involves an MCMC (see Sec. 5.3.3). This changes the behavior of the swarm of particles from optimizing and finding the peak of $\Upsilon_{N=1}$ to exploring, or sampling, the posterior distribution based on $\log \hat{L}_{N=1}$. It should be stressed that this pipeline is not intended to perform high-quality sampling of the posterior. Because the MCMC walkers are all seeded near the peak found by the PSO the subsequent sampling will be slow to properly explore the tails of the posterior and will likely miss entirely any secondary peaks. However, this relatively crude MCMC can still be useful because it gives an indication of the posterior width in each parameter and can be used to seed more detailed follow-up inference.

These parameter estimation results from the search of the GW150914-like $\rho = 38$ fiducial source are shown in Fig. 5.6. Additionally, posteriors on the chirp mass and eccentricity parameters are shown for all recovered sources in the bottom right panel of Fig. 5.5. Posteriors are only shown for the phasing parameters; i.e. parameters that significantly contribute to the GW phase. This final MCMC stage of the search performs well for loud sources, where the PSO reliably and accurately locates the peak. For quieter sources, such as $\rho = 17$ case, the parameter estimation results are biased even though the source is confidently detected.

5.7 Computational cost

The search described in this chapter is designed to be broken down into distinct tiles in $\mathcal{M}_c - t_c$ space. Each tile can be searched over simultaneously and we have shown example searches within a couple of these tiles. Each search in this chapter has used 1

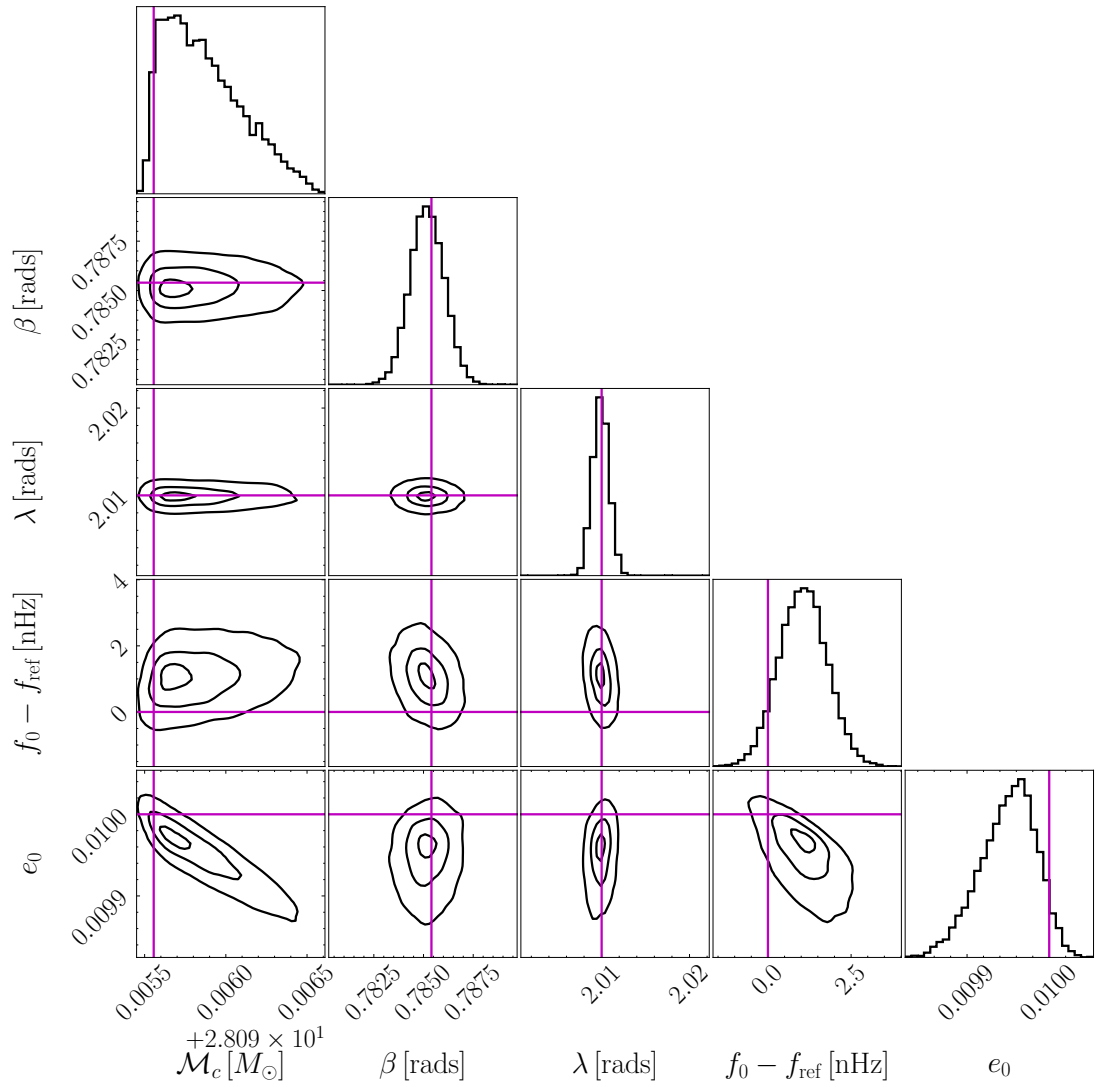

Figure 5.6: Parameter estimation results from the search for the GW150914-like fiducial source in Table 5.1. Plotted parameters here are: chirp mass ($\mathcal{M}_c$), ecliptic latitude (β), ecliptic longitude (λ), initial orbital frequency (f_0) and initial eccentricity (e_0). The $[10, 50, 90]$ percentile contours are plotted in black and the injected source parameters are marked by the magenta lines. Posteriors in the initial orbital frequency are plotted with respect to the reference frequency $f_{\rm ref} = 9 \times 10^{-3}$ Hz, which is also the injected value.

NVIDIA A100 GPU and 10 CPUs, we have made use of the `PyTorch` implementation of `multiprocessing` to parallelize the process of computing waveforms and search statistics on one GPU [243]. All the searches presented required $\sim 1 - 2$ days of compute time. This does not include the parameter estimation described in Sec. 5.3.3. For the small number of sources where this parameter estimation was performed this took an additional ~ 10 hours per candidate. Given that ~ 600 tiles will be needed to cover the parameter

space of interest and the expected small number of sources in LISA, while expensive, this search can be scaled up the required level. Additionally future hardware and algorithmic improvements will reduce the cost further.

5.8 Conclusion

In this chapter we have demonstrated a GPU accelerated semi-coherent hierarchical search for SmBBH inspirals in simulated LISA data containing stationary Gaussian noise. We have established the threshold SNR for this search to be around $\rho_{\text{threshold}} \approx 17$. This is roughly consistent with previous estimates of $\rho_{\text{threshold}} = 15$ for this threshold [103, 179]. Improvements to the optimization and search algorithm may push the threshold SNR down slightly further. We have also demonstrated a method for computing the background distribution of noise triggers from a stochastic search, and used it to place upper and lower limits on the FAP for triggers arising from a few simulated searches. Realistic simulated LISA data is required to make robust estimate of the FAP and detection probability for this type of search.

The work done in this chapter is limited to the sources with low eccentricity and no spin, however we have some promising early results that indicate the semi-coherent search should work well on aligned-spin sources. In addition to including spin, the main focuses of future work will be to extend this method to data-streams with non-stationary components, gaps and glitches, and also to search the LISA Data Challenge datasets which contain SmBBH signals within them. We will also explore lowering N_{high} and analytically maximizing over more of the extrinsic search parameters, as this may allow for a lower threshold SNR as demonstrated in Ref. [223]. See chapter 6 for an application of this method to the LISA data challenge *Yorsh* dataset.

6. Searching for stellar-origin binary black holes in LISA Data Challenge 1b: Yorsh

Overview

This chapter is an edited version of Ref. [244].

Diganta Bandopadhyay and Christopher J. Moore. Searching for stellar-origin binary black holes in LISA data challenge 1b: Yorsh, published in Physical Review D, Volume 111:044039, February 2025.

This work applies the search pipeline developed in Chapter 5 to the *Yorsh* LISA data challenge. Crucially the search pipeline was run completely independently of the injection pipeline, with a different waveform and instrument response model than that used to generate the injected data. Five out of eight of the sources within the data are found (with $\rho \gtrsim 12$), with seeded (coarse) parameter estimation results presented for each of source. For computational reasons this was not a completely blind search, however an implementation of this is currently being developed using the time-frequency approach described in Chapter 3 which is capable of performing a blind search on the *Yorsh* dataset within $\mathcal{O}$(weeks).

I performed all the analysis in this paper, applying the code developed in Chapter 5 to the *Yorsh* LDC dataset, and produced all the results in this paper. The text in this paper is jointly written between me and Christopher J. Moore.

6.1 Introduction

Following on from the development of the semi-coherent stellar mass search in Chapters. 4 and 5, this chapter focuses on the application of the search to the first 'realistic' scenario, in which the search pipeline is completely independent of the injection pipeline, and multiple sources are present in the datastream at once, alongside instrumental noise.

SmBBH searches have hitherto been run on simple datasets, specifically generated for the purpose of validating the search. Here, results are presented for the first attempts at a SmBBH search on the multipurpose LISA Data Challenges (LDCs) the community is using to prepare for LISA. Specifically, we use the 1b *Yorsh* LDC [245], a dataset designed for developing methods for analyzing complex, long-duration signals. This increases the difficulty because both the waveform and the LISA response model used in the search differ from those used to generate the data. Furthermore, because the search is performed by different people and codes from those that generated the data, it forces us to confront the practical problems of matching all the necessary conventions. *Yorsh* is not a *blind* data challenge, meaning the injected sources and their parameters are known.

The search output is a point estimate for the parameters of a candidate signal that maximize the semi-coherent statistic $\Upsilon_{N=1}$ [218]. The significance of the candidate is assessed against a background distribution of noise triggers. Confident detections are followed up with rapid, low-cost parameter estimation using an ensemble Markov chain Monte Carlo (MCMC) [231, 218]. This first estimate of the Bayesian posterior is intended to be used to initialize more detailed parameter estimation as part of a global fit [128, 63]. Sec. 6.4 contains the results of two versions of our search using different models for the instrument response and noise power spectral densities (PSDs). The best performing search confidently identifies five sources, including one with an injected signal-to-noise ratio (SNR) as low as 12.94. Rapid parameter estimation confirms these correspond to the five loudest injections with only small parameter biases expected due to differences in the waveform and response modeling.

In the rest of the chapter, Sec. 6.2 describes the SmBBH part of the *Yorsh* LDC. Sec. 6.3 briefly describes the search, focusing on the differences from Chapter. 5. Sec. 6.5 contains details of the computational costs of the search. Sec. 6.6 contains a discussion of the results.

6.2 Data

The data analyzed in this study was produced for the 1b *Yorsh* LDC [245]. The simulated data includes time series for the three second-generation time-delay interferometry (TDI-2) channels (X, Y, Z), 2 years in duration and downsampled to a 5 s cadence. Our search uses all of the approximately noise-orthogonal (A, E, T) TDI channels [207] obtained by a linear combination of (X, Y, Z). Note that although the GW response in the T channel is suppressed at low frequencies, and is consequently often neglected (see, e.g. Refs. [128, 63]), it is more important at the high frequencies $\gtrsim 10$ mHz where most SmBBH signals reside[1].

The simulated data includes idealized instrumental noise. The noise model is stationary and Gaussian and was used to generate data in the frequency domain from a known PSD before transforming to the time domain. The PSD was the same as that used in the *Sangria* LDC data challenge [246], differing slightly from that in the LISA science requirements (SciRD) [210]. Hereafter, this will be referred to as the *Yorsh* PSD. Galactic confusion noise [13, 247] was not included; while this will be a significant non-stationary noise source at lower frequencies $\lesssim 3$ mHz [13], it is unlikely to significantly affect the search for the loudest SmBBH sources at high frequencies $\gtrsim 10$ mHz (see Fig. 6.1)[2].

For the simulated data, the SmBBH injections were modeled using the quasi-circular, aligned-spin, frequency domainwaveform `IMRPhenomD` [248] (as implemented within `LISAbeta` [249]) transformed into the time domain via the stationary phase approximation.

[1] For example, source #8 has $\rho_A^2 = 276$, $\rho_E^2 = 292$ and $\rho_T^2 = 26$; the T channel contains $\sim 4\%$ of the injected squared SNR.

[2] In a recently updated version of *Yorsh*, galactic confusion noise is included, a search including this is currently underway.

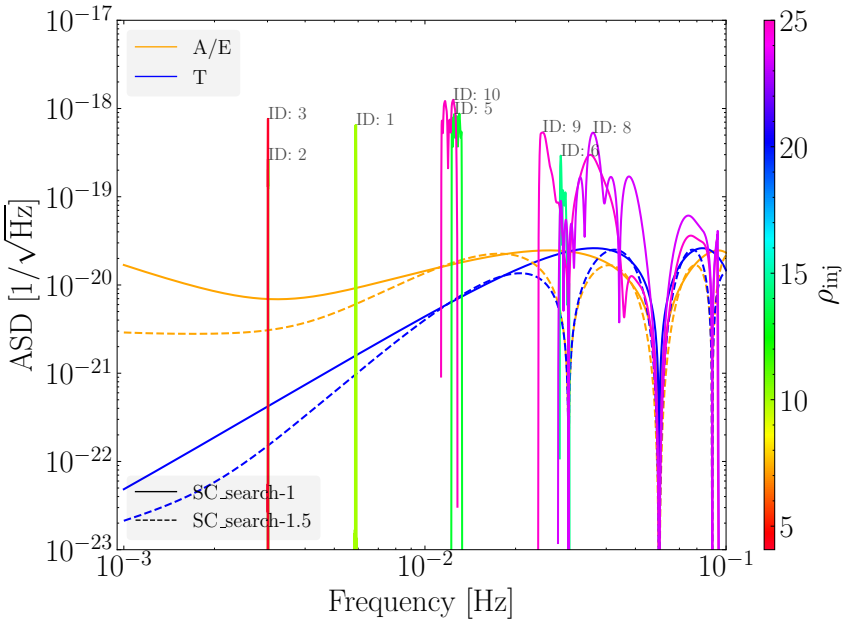

Figure 6.1: The high-frequency LISA bandwidth containing SmBBH signals. Orange (A, E) and blue (T) curves show the ASDs in the TDI channels. Different line styles show the ASDs for the two searches. Note, the `SC_search-1.5` ASD has three zeros in the range plotted while the `SC_search-1` ASD has one. Also plotted are representations of the signals for the sources in Table 6.1 colored by injected SNR; the waveform quantity plotted is $2|\tilde{A}(f)|\sqrt{f}$. Sources #2 and #3 overlap on this plot and cannot be distinguished.

The instrument response is modeled in the time domain, computed using `LISACode` [250]. There are eight SmBBH sources in *Yorsh*, numbered 1 to 10 (4 and 7 missing). Key parameters of the injections are summarized in Table 6.1 and the signals are plotted in Fig. 6.1. The source parameters cover a range of chirp mass, $\mathcal{M}_c$, time to merger t_c, and mass ratio, q. These parameters were chosen to facilitate the development of data analysis algorithms and were not intended to be an astrophysically realistic population. For example, sources #2 and #3 have $\mathcal{M}_c$ and t_c values much closer together than would normally be expected to allow for investigations into source confusion.

All injections have component spins aligned with the orbital angular momentum and are therefore non-precessing. LISA is not expected to be able to constrain the component spins of SmBBHs with high accuracy [162, 163, 147, 65] but selected results are presented in Sec. 6.4 for the effective aligned spin combination $\chi_{\rm eff}$.

Table 6.1 gives both the time to merger, t_c, and the GW frequency at the start of observations, $f_{\rm low}$. These are not independent; the search works with $f_{\rm low}$ directly and t_c

is a derived parameter, computed using `TaylorF2Ecc`.

Table 6.1 also reports the injected SNRs of the sources. These were calculated by taking a discrete Fourier transform of the injected signal after applying a Tukey window with parameter $\alpha = 0.1$ [251]. The same window was used to prepare the frequency domain data for the search.

6.3 Methods

The semi-coherent search algorithm (`SC_search`) used here closely follows that described in Chapter 5.

The search uses a custom implementation of the `TaylorF2Ecc` waveform model [233]. Note that this differs from the waveform used in the injections. This frequency-domain, $\ell = |m| = 2$ mode, post-Newtonian (PN) waveform approximant includes contributions to the GW phase from small orbital eccentricities, vital for the SmBBH population that LISA is expected to see [252]. Note that even though the injections in *Yorsh* were performed with quasi-circular waveforms, the search does not assume this and includes an additional parameter, e_0, for the binary eccentricity at the start of LISA observations. This waveform includes eccentric phase contributions up to 3 PN order expanded to leading order in e_0. Compared to what was used in Chapter 5, the waveform has been further extended to include leading-order contributions to the GW phase due to the aligned component spins. The corrections to the phase arising from the spin contribution can be found in Ref. [253] and references therein. All aligned spin terms up to 2.5 PN order were included. The new spin terms have been checked in the quasicircular $e_0 = 0$ limit by comparing with the `TaylorF2` implementation used in Ref. [163].

The search uses a fast model of the LISA response to convert the GW waveform, represented in the frequency domain using amplitude and phase, to time-delay interferometry (TDI) variables (A, E, T). Currently, there does not exist a fast GPU implementation of the second generation TDI-2 response suitable for use in this sort of analysis. Therefore,

two other models of the instrument response were used. Both searches use the frequency-domain response described in Refs. [35, 138] and implemented in `BBHx` [236]. `SC_search-1` used the first-generation TDI-1 variables (A, E, T) that assume a static LISA constellation. `SC_search-1.5` used the improved TDI-1.5 variables (A, E, T), which accounts for a rigid (i.e., constant armlength) rotation of the constellation. Neither used the TDI-2 variables used in the injections that account for relative motion between the spacecraft [254, 255].

The search assumes a known noise PSD, $S_\alpha(f)$ for $\alpha \in \{A, E, T\}$. `SC_search-1` used the SciRD model for the instrumental noise in the TDI-1 channels, from Ref. [210]. `SC_search-1.5` used the LDC *Yorsh* PSD for the TDI-2 channels (see Sec. 6.2). Double white dwarf confusion was not included as *Yorsh* does not include these sources. These PSDs are plotted as amplitude spectral densities (ASDs, $\sqrt{S_\alpha(f)}$) in Fig. 6.1.

As described in Chapter 5, the search does not cover the entire SmBBH parameter space at once. Instead, restricted search priors were used for $\mathcal{M}_c$ and $f_{\rm low}$; this part of parameter space is referred to as a search tile. The two-dimensional $\mathcal{M}_c$, $f_{\rm low}$ space can then be covered by $\approx 100 - 1000$ search tiles, which run independently and in parallel. In this study, to limit the computational cost, individual and separate search tile was used for each source in Table 6.1. The process of choosing the tile boundaries has not yet been automated and was therefore done by hand. The width of each tile in $\mathcal{M}_c$ was set to approximately $\sim 5\mathrm{M}_\odot$. The width of each tile in $f_{\rm low}$ was chosen such that the prior width on the derived parameter t_c was approximately equal to the actual time to merger of that source. Figure 6.2 shows the induced search priors on the derived parameter t_c for all eight sources. The exact values for the tile boundaries are given in Ref. [251].

Within each search tile, both the search and the subsequent rapid parameter estimation (if performed) use flat priors in all of the parameters described in Table 5.1. This includes the parameters $\mathcal{M}_c$ and $f_{\rm low}$ used to define the tile boundaries. Additionally, flat priors were used on the individual component spin magnitudes in the range $[-0.99, 0.99]$. Together with the flat priors on the chirp mass and symmetric mass ratio, this gives the

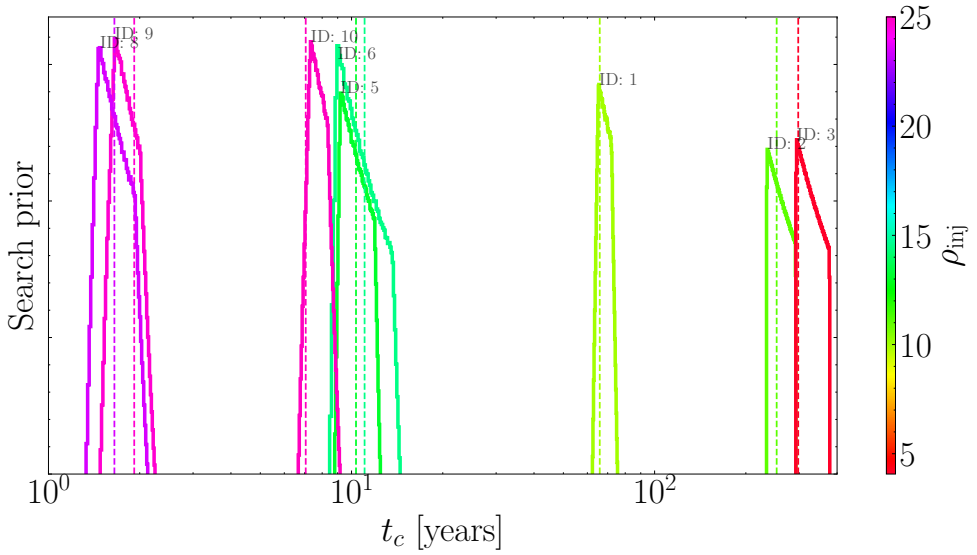

Figure 6.2: One-dimensional marginalized prior probability on the time to merger parameter t_c used in the search. Flat priors are used for the $\mathcal{M}_c$, $f_{\rm low}$ parameters within the restricted ranges described in the main text. When transformed to t_c, these priors give the distributions shown here. The injected values of t_c are shown for each source using the vertical dashed lines. Colors for all eight sources match those in Fig. 6.1.

non-flat prior on the $\chi_{\rm eff}$ parameter shown in the bottom right panel of Fig. 6.3.

Each search attempts to find the loudest candidate in its search tile. The statistical significance of these candidates is assessed by comparing the value of the $\Upsilon_{N=1}$ statistic against a background noise distribution. The background should be generated by running a large number of identical searches on simulated LISA datasets which do not contain SmBBHs. For more details, see the demonstration of this process in Ref. [218]. This process is computationally expensive and must be repeated for each tile to account for variations in the background distribution across parameter space. Therefore, the background noise distribution from Fig. 5 in Ref. [218] was used for all tiles in this study. Here, sources with $\Upsilon_{N=1} \geq 100$ are deemed to be confidently detected. This is a conservative threshold for a statistically significant source, see Figs. 5.3 and 5.5.

Confident detections were followed up with rapid parameter estimation. The rapid parameter estimation settings were identical to those in Chapter 5. The ensemble MCMC used 10^2 walkers, with 10^4 iterations.

Table 6.1: **Injection parameters**. Source parameters used to produce the injected signals in the *Yorsh* LDC. The final column reports the optimal SNR of each injection ($\rho^{\rm inj}$) computed using the injected waveforms (from the noiseless datastreams) and the LDC-2 PSD. All of the injected binaries were on quasi circular orbits, i.e. with zero injected eccentricity. Events marked with a star* are those that merge within the LISA mission duration and are often described as multiband sources.

ID	$\mathcal{M}_c^{\rm inj}$ $[M_\odot]$	$t_c^{\rm inj}$ [years]	$f_{\rm low}^{\rm inj}$ [mHz]	$d_L^{\rm inj}$ [Mpc]	$q^{\rm inj}$	$\chi_{\rm eff}^{\rm inj}$	$\rho^{\rm inj}$
#1	29.34741587	65.9176	5.85830665	159.9	0.91	0.50	10.91
#2	38.04622881	252.7789	3.00851783	94.5	0.83	−0.06	4.07
#3	34.51216704	297.7712	3.00698596	47.0	0.58	0.10	9.88
#5	27.41970433	10.3457	12.24273032	168.3	0.83	−0.55	12.94
#6	7.007404972	11.0420	28.02352272	17.3	0.88	−0.17	14.30
#8*	22.40969304	1.6501	27.65438527	34.0	0.59	0.002	24.37
#9*	26.08583360	1.9185	23.76783772	85.5	0.95	0.10	23.08
#10	39.14942200	7.0604	11.31112717	168.9	0.88	0.03	24.65

6.4 Results

The main results are summarized in Table 6.2, which reports the results of the searches, and Table 6.3, which reports the results of the rapid parameter estimation.

`SC_search-1` only finds two of the eight sources. This is partially expected given that it uses an inferior (lower-generation) TDI than that used in the injections. It is interesting that this search found source #5 with injected SNR 12.94 while missing the louder sources #6, #8 and #9. This is likely because source #5 was at lower frequency and where the differences between the TDI generations is less pronounced (see, e.g., the ASD curves in Fig. 6.1).

`SC_search-1.5` finds all five sources in the data challenge with injected signal-to-noise ratios $\gtrsim 12$. This is impressive and suggests a threshold SNR necessary for the detection of these SmBBH sources significantly lower than some previous estimates [218, 103, 256]. The expected number of SmBBH signals depends strongly on the threshold SNR, a lower threshold implying a higher source count. The threshold SNR will not be accurately known until a search pipeline has been fully developed and run on realistic LISA data (e.g., including realistic noise, gaps, and other sources) to estimate the background distribution of noise triggers. This is complicated by the fact that the threshold SNR

Table 6.2: **Search results**. Two searches were performed for each source using the different models for the LISA response and different PSDs described in the main text. Results from both searches are reported here. Sources that were confidently detected are indicated with a tick and their chirp mass and time to merger parameter values are reported as differences from the injected values; e.g., $\delta\mathcal{M}_c = \mathcal{M}_c^{\rm search} - \mathcal{M}_c^{\rm inj}$. The matched filter SNR, $\rho_{\rm mf}$, of the template that maximized the semi-coherent detection statistic is also reported. For sources that were not confidently detected, those indicated by a cross, the search did not find the injected source; i.e. the maximum-search-statistic-parameters did not correspond to the injection. Therefore, these non-detections represent sources which were missed by the search, not sources that were found and assigned a low significance.

	`SC_search-1`				`SC_search-1.5`			
ID	Found	$\delta\mathcal{M}_c\,[M_\odot]$	$\delta t_c\,[s]$	$\rho_{\rm mf}$	Found	$\delta\mathcal{M}_c, [M_\odot]$	$\delta t_c[s]$	$\rho_{\rm mf}$
#1	✗	-	-	-	✗	-	-	-
#2	✗	-	-	-	✗	-	-	-
#3	✗	-	-	-	✗	-	-	-
#5	✓	0.0015	42482	12.2	✓	0.0004	40689.	11.60
#6	✗	-	-	-	✓	0.0005	28031.	14.91
#8	✗	-	-	-	✓	0.0001	611.	21.33
#9	✗	-	-	-	✓	0.0020	1395.	23.53
#10	✓	0.0002	24297	26.54	✓	0.0006	22137.	25.77

Table 6.3: **Rapid parameter estimation results**. Here we report the results of the rapid parameter estimation performed by the simple MCMC algorithm which can be performed automatically at the end of a search when a source is found. For searches that did not find a candidate in Table 6.2, the corresponding cell in the table is left blank. For the chirp mass and time to merger, the values reported in the table are the widths (90% credible interval) of the posteriors which are all centered near the search values reported in Table 6.2. For the luminosity distance and the eccentricity, the median and 90% credible intervals are reported in the form of an asymmetric error bar. The (90% credible) sky area for each source, Ω_{90}, is also reported.

	`SC_search-1`					`SC_search-1.5`				
ID	$\Delta\mathcal{M}_c\,[M_\odot]$	Δt_c [hr]	d_L [Mpc]	Ω_{90} [deg^2]	$e_0\,[\times 10^{-3}]$	$\Delta\mathcal{M}_c\,[M_\odot]$	Δt_c [hr]	d_L [Mpc]	Ω_{90} [deg^2]	$e_0\,[\times 10^{-3}]$
#1	-	-	-	-	-	-	-	-	-	-
#2	-	-	-	-	-	-	-	-	-	-
#3	-	-	-	-	-	-	-	-	-	-
#5	0.01	21.80	121^{+30}_{-42}	25.42	$5.1^{+4.3}_{-5.1}$	0.01	21.63	237^{+61}_{-86}	25.87	$5.4^{+4.0}_{-5.4}$
#6	-	-	-	-	-	0.002	19.22	17^{+15}_{-4}	0.65	$6.0^{+3.1}_{-3.6}$
#8*	-	-	-	-	-	0.001	0.19	39^{+4}_{-3}	11.51	$3.3^{+1.1}_{-0.6}$
#9*	-	-	-	-	-	0.001	0.07	77^{+12}_{-8}	1.16	$3.2^{+0.4}_{-0.5}$
#10	0.01	9.7	84^{+15}_{-26}	1.05	$4.8^{+2.8}_{-4.8}$	0.01	10.32	156^{+34}_{-48}	1.98	$4.7^{+2.9}_{-4.7}$

will vary significantly across the large SmBBH parameter space, as a consequence of the background distribution (such as that computed in Fig. 5.5) shifting in different tiles in parameter space.

The `SC_search` results are summarized in Table 6.2. In cases where the maximum of the semi-coherent statistic $\Upsilon_{N=1}$ did not exceed the threshold, no confident candidate source was found; these are marked with a cross in Table 6.2 and no parameters are reported. When a source is not found this is typically because the source is too quiet or, in the case of `SC_search-1`, because a loud source (such as #6, #8 and #9) accumulates significant SNR at frequencies around $\sim 25\,\mathrm{mHz}$ where the simplified TDI-1 response used by `SC_search-1` gives a poor approximation to TDI-2 used for the injections and it is not possible to match the template signal. When a source is found, the parameters of the template that maximizes the detection statistic are reported in the table.

It is interesting to note that both searches made detections using TDI generations that differ from the injections. This shows that for some sources, particularly at lower frequencies, it is possible to perform LISA data analysis without the most accurate model for the instrument response. This will likely be true for other source types and may allow for the use of simpler and faster response models in carefully selected parts of a global fit.

Rapid parameter estimation was performed for all of the confident detections produced by both searches. The parameter estimation used the same model for the instrument response as the initial search. The rapid parameter estimation provides further confirmation that the high-confidence detections produced by the search do correspond to the injected sources in the data challenge. The posteriors are extremely narrow in several key parameters (notably, $\mathcal{M}_c$ and f_{low}) and are close to the injected values. An example posterior is shown in Fig. 6.3.

Several parameter biases are observed. This is expected due to the use of different waveform and instrument response models for the injection and recovery. The largest biases are in some of the distance parameters in `SC_search-1`; see, e.g. the source #10 in Fig. 6.3. The factor by which the distance is underestimated can be understood from the

definitions of the TDI-1 variables which are related to the TDI-1.5 variables by a factor $(1-\exp(-8\pi i f L))$ where L is the LISA armlength (see, for example, the implementation in `BBHx` [257]).

Other small biases are observed, such as the e_0 posteriors which occasionally peak at non-zero values. Also, some $\chi_{\rm eff}$ posteriors (notably sources #8 and #9) peak away from the injected values. The shifts in the peaks of the marginal posteriors may be caused by a mismodelling of the response (as discussed in the previous chapters, the clear distinction between extrinsic and intrinsic parameters are not as clear for LISA as for LVK), both the instrument response, eccentricity and $\chi_{\rm eff}$ contribute significantly to the waveform phase, therefore a mismodelling of the response can lead to a shift in the phase which can be compensated for by a shift in $\chi_{\rm eff}$ or e_0. Additionally, since we are analysing one datastream, there is no way to test the statistical consistency of the posterior using a probability-probability plot, such as the process demonstrated in Sec. 5.5.2.

As expected, LISA SmBBH observations are not able to significantly constrain the component spins. An example of the posterior on $\chi_{\rm eff}$ for source #10 is shown in Fig. 6.3; although low values of $\chi_{\rm eff}$ are preferred, no parameter range is confidently excluded. LISA measurements also allow the time of merger to be predicted within a few hours. This is consistent with previous parameter estimation studies (see, e.g. Refs. [162, 163, 147, 65]) .

Table 6.3 also reports results for sky localizations obtained from the rapid parameter estimation analyses. As long-lived and high-frequency LISA sources, SmBBHs are typically quite well localized; all the recovered sky maps contain a single, roughly Gaussian posterior mode. The areas were calculated using a 2-dimensional kernel density estimate (KDE) of the posterior probability for the ecliptic longitude and sine latitude, $(\lambda, \sin\beta)$, obtained by performing a Monte Carlo integral with samples drawn from the KDE.

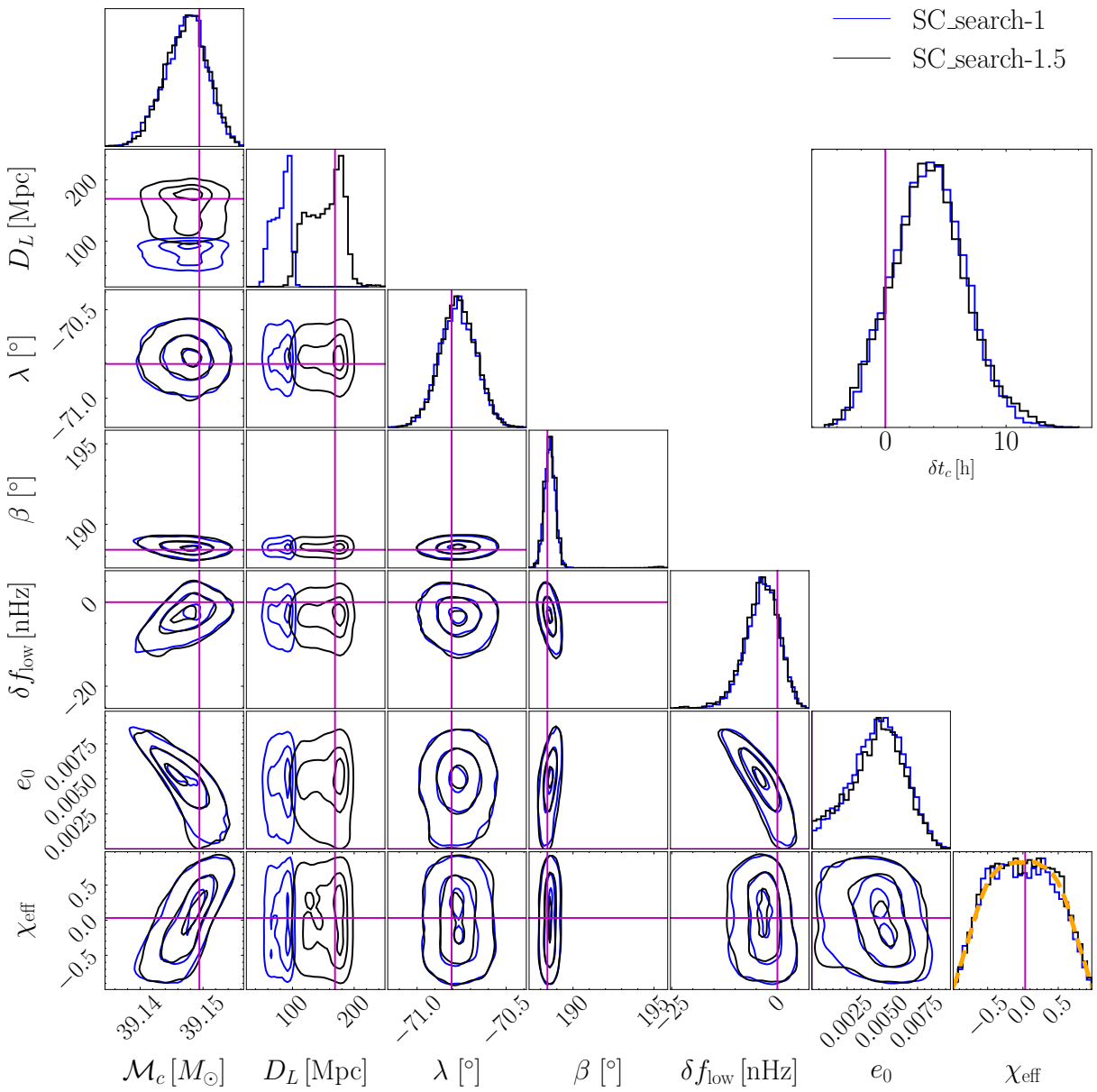

Figure 6.3: Rapid parameter estimation posteriors on selected parameters for source #10. Magenta lines indicate the injected parameters. Two-dimensional contours show the 10, 50, and 90% credible regions. The `SC_search-1` posterior significantly underestimates the distance to the source due to the use of a lower generation TDI, as explained in the text. The `SC_search-1.5` posterior is consistent with all the injected parameters. Also shown in the top right is the posterior on the derived parameter $\delta t_c = t_c - t_c^{\rm inj}$. The orange curve shows the 1-dimensional marginalized prior distribution on the $\chi_{\rm eff}$ parameter.

6.5 Computational Cost

The computational costs of the `SC_search` are similar to those described in Ref. [218]. Each tile used 1 NVIDIA A100 GPU and 10 CPUs and took $\mathcal{O}(1-3)$ days, with the highest

frequency searches taking the longest. The searches for different tiles were run in parallel. These costs may be further reduced with methodological improvements; therefore, this estimate should be treated as an upper bound for the cost of the SmBBH search.

6.6 Discussion

This chapter presented results from the first SmBBH searches in the LDCs. The most effective search successfully identifies all sources with injected SNRs $\gtrsim 12$. This is an encouraging result, as well as being somewhat unexpected given previous estimates for the threshold SNR, such as that in Chapter 5. This further highlights the need to robustly establish the background distribution(s) across the whole parameter space. The search also automates rapid parameter estimation and was able to provide crude estimates of the posterior distributions suitable for initializing more detailed parameter estimation, e.g. as part of a global fit.

The main limitation of the `SC_search` currently is the lack of a suitable fast implementation of the LISA response. This should be developed as a priority and will be incorporated into this search as soon as it is available.

The other limitations concern the idealized nature of the LDC data. The *Yorsh* data was generated without gaps or glitches, without other source classes (such as high-frequency double white dwarfs or extreme-mass-ratio inspirals), without uncertainty in the instrumental noise, and without non-stationary or non-Gaussian noise components. Subsequent data challenges have already started to relax some of these assumptions but so far do not include SmBBHs. This search will continue to be applied to future LDCs containing SmBBHs.

7. Conclusions

In this thesis, we have discussed various topics at the heart of LISA data analysis; beginning from the basic principles of gravitational wave detection and eventually following the development of a novel search strategy for the long-lived gravitational wave signals expected from the stellar-mass binary inspiral signals.

Chapter 1 introduced gravitational waves and derived a simple leading-order waveform for a binary inspiral, with a brief tangent on the detection principles used by current GW detectors. The results of this chapter are useful in that they provide the reader a qualitative understanding of the type of GW signal this thesis is concerned with. The content within Chapter 2 is a review of the compact binary sources that are expected to be detected by LISA observations. The phenomenology of the gravitational wave signals emitted by each compact binary source is discussed, along with any anticipated astrophysics and fundamental-physics insights expected from LISA observations. Chapter 3 introduces key aspects of GW data analysis, contrasting the applicability of current ground-based techniques to those required for the LISA mission. This chapter also provides an intuitive explanation of the intractable computational expense that is associated with applying template-bank techniques to the search for long-lived sources in LISA data. Additionally, the response of the LISA detector arms to a passing gravitational wave signal is derived, while this is not a new result, it is presented here in a form that is easy to follow. An independent implementation of this response is currently underway, and will be used to accelerate the search that is described in the subsequent chapters.

Chapter 4 marks the beginning of the original work presented in this thesis. In this

chapter we conceptualized a semi-coherent search framework for LISA data analysis, specifically for the stellar-mass binary search problem. This analysis framework is initially validated on the analysis of the binary neutron star event GW170817 as this is the closest analog to the LISA stellar-mass binary sources. The semi-coherent framework is then used to successfully search for a simulated binary black hole signal in noiseless data. The work within this chapter served as an early proof of concept for the semi-coherent search framework, successfully solving the prior localization problem for long-lived signals. Chapter 5 extended this work by augmenting the search pipeline through the use of both a multi-objective particle-swarm optimization algorithm and GPU acceleration. The GPU acceleration was a key aspect of this work as it made the search computationally tractable for datastreams which contain instrumental noise. This chapter also tackled the lack of a suitable method for computing detection significance for candidate signals found in LISA data, handling this problem by running searches on an ensemble of simulated noise-only data streams, and using the results of these searches to compute false-alarm probabilities for candidate detections. The search pipeline code was publicly released in [219].

Finally, Chapter 6 presents the results of applying the search pipeline to the *Yorsh* LISA data challenge. This study was a significant milestone as it tackled a variety of challenges: mismodeling of the power spectral density, GW waveform and LISA response. Despite this, the search pipeline was successfully able to detect and characterize (the loudest) 5 out of the 8 stellar-mass binary inspiral signals present within the data. Although this study only searched small sections of the parameter space, a tractable cost of the search is presented, extrapolating this to the full parameter space suggests a cost of $\sim$ months to complete the search. However, there has been ongoing work since this study, focused on accelerating the search pipeline through the use of the time-frequency approach described in Sec. 3.4. This has been successful in reducing the computational cost of the whole search (across the entire astrophysical parameter space) to $\sim$ 1 week, additionally this approach has been (empirically) shown to be effective in the presence of data gaps. The increased computational efficiency of the search pipeline also allows for a

more robust characterization of the false alarm probability of candidate signals.

In the future, I plan to apply the search to the upcoming LISA data challenge *Mojito*, which is expected to be released in late 2025. This data challenge will test interactions between the search and inference pipelines for several source types, within the framework of the global fit. Adaptation of this search strategy will be necessary to apply it to the more complex EMRI search problem, primarily due to the large number of harmonics present in the signal and the resulting degeneracy that this introduces in the parameter space.

A. Chapter 4 appendix

These appendices are slightly edited versions of appendices from Ref. [149]:

Diganta Bandopadhyay and Christopher J. Moore. LISA stellar-mass black hole searches with semi-coherent and particle-swarm methods, published in Physical Review D, Volume 108(8):084014, October 2023.

A.1 Tempering vs semi-coherent

A tempered version of a probability distribution p with temperature T is the new probability distribution

$$\tilde{p} = p^{1/T}. \tag{A.1}$$

In its application to GW data analysis, a tempered posterior aids the exploration of stochastic sampling algorithms and mitigates against sampling chains getting stuck in secondary maxima for multi-modal likelihood surfaces. Considering the simplest sampling algorithm, the Metropolis Hastings MCMC sampler, a proposal for a walker at position θ_i to move to θ_j is accepted with probability $\alpha = p(\theta_j)/p(\theta_i)$. In the situation where the walker is initially in a secondary maxima, it is unlikely to step out and into another well-separated global maxima. Parallel tempering broadens peaks by raising the log-likelihood floor, increasing the probability the walker will be able to jump between maxima and therefore explore the multi modalities in the surface. From Eqn. A.1 it can be seen that tempering is a re-scaling of the log-likelihood.

We contrast this with the semi-coherent method which broadens log-likelihood around the injection and combines secondary peaks smoothly, *removing* variations in the log-likelihood surface. Whereas the tempering approach *suppresses* variations in the log-likelihood by raising the troughs, however it does not remove the secondary peaks, just suppresses differences between the peaks and troughs. These suppressed variations result in a surface that is easier for a probabilistic sampling algorithms to explore, however the secondary peaks still exist in the surface, making it difficult for optimization algorithms (such as those conventionally used in the search/prior localization phase) to explore this surface. Meanwhile the semi-coherent approach results in a smooth surface around the injection, *removing* variations in the log-likelihood, this approach is much better suited for exploration by an optimizer.

A comparison between the tempered and semi-coherent likelihoods is shown in Fig. A.1.

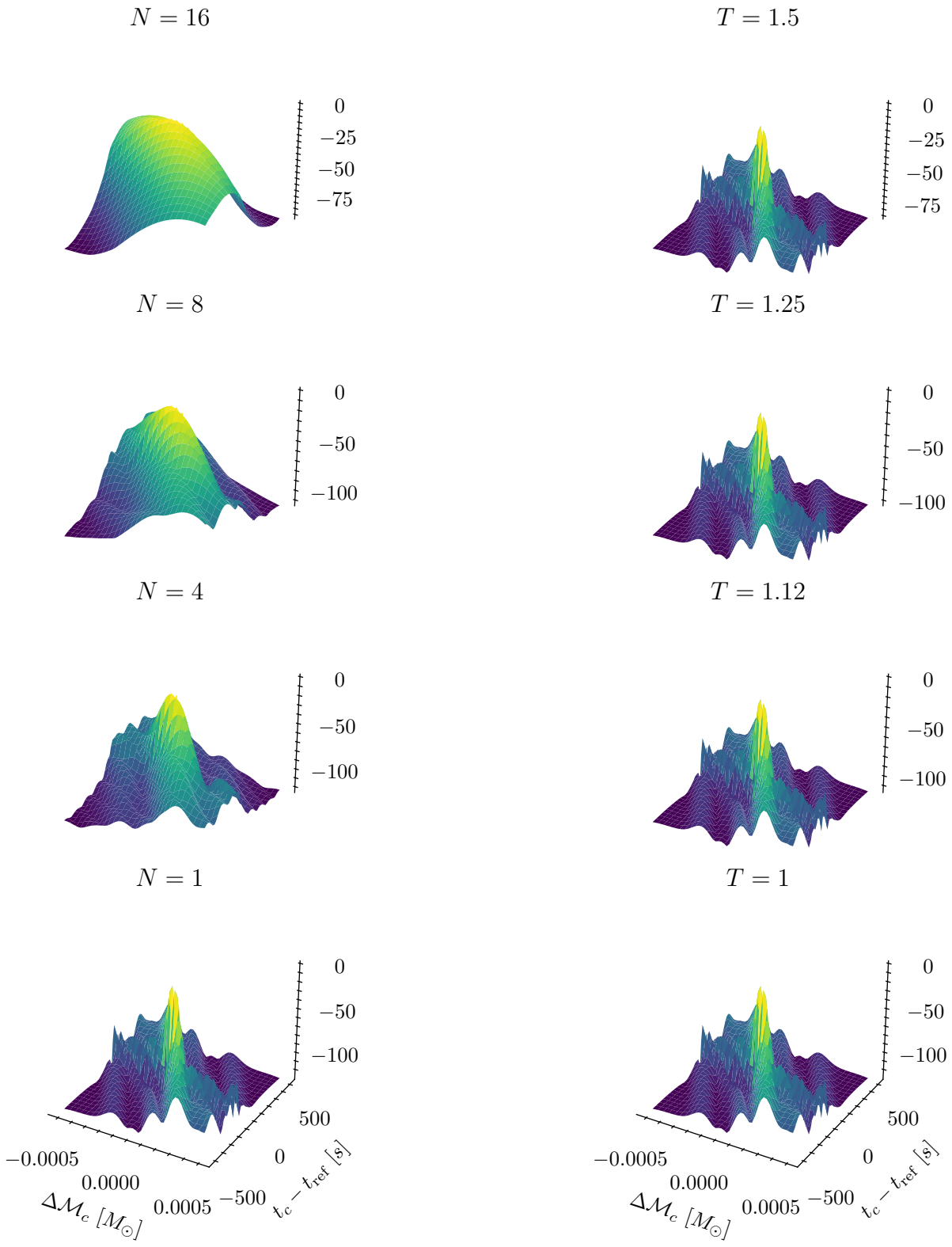
$N = 16$
$T = 1.5$
0
−25
−50
−75
$N = 8$
$T = 1.25$
0
−50
−100
$N = 4$
$T = 1.12$
0
−50
−100
$N = 1$
$T = 1$
0
−50
−100
500
0
−500
$t_c - t_{\mathrm{ref}}$ [s]
−0.0005
0.0000
0.0005
$\Delta\mathcal{M}_c$ [$M_\odot$]

Figure A.1

Figure A.1 *(previous page)*: Figure illustrating the difference between a tempered log-likelihood surface, $T^{-1}\log L(d|\theta)$, (*right column*) with variable temperature T and a semi-coherent log-likelihood, $\log\hat{L}_N$ (*left column*). The likelihood surfaces plotted here are for the SmBBH LISA signal described in Sec. 4.4 and Table. 4.1 (plotted as a function of the chirp mass and time to merger parameters with all other parameter fixed to their true values) although the trends shown are generic. Increasing N in the semi-coherent likelihood raises the floor of the likelihood surface, decreasing the peak-to-trough range of log-likelihood values. It also has the effect of congealing secondary-maxima together; the complicated structure of peaks troughs and ridges seen in the coherent ($N = 1$) are completely absent in the top $N = 16$ plot. In contrast, the tempered log-likelihood surface is simply a re-scaled version of the vanilla log-likelihood, this has the effect of only raising the log-likelihood floor. The values of T plotted here were chosen so that the range of log-likelihood values in the region plotted is similar between the left and right columns of plots.

A.2 PSO configuration

The PSO search used in this study used $N_p = 15000$ particles, with log-likelihood evaluations parallelized across 20 CPU cores. The optimization of each hierarchical segment (i.e. value of N) was allowed a maximum of 250 iterations, however the PSO swarm could stop and move to the next log-likelihood (i.e. $N \to N/4$) at iterations below this if they met the convergence criteria. In this study we set the convergence criteria as the absence of any significant improvement (with a tolerance of 0.01) in the best swarm function value in the last 50 iterations. Tab. A.1 shows the hyperparameters used for each stage of the hierarchical search.

We stress the PSO configuration and ladder of segments used in this study are tuned empirically and while it is sufficient at the level of a proof-of-concept study, further work is needed to provide concrete suggestions. It is likely such configurations will be source dependent, e.g. SmBBH systems will possibly have a very different optimal PSO configuration and segment ladder to EMRIs.

Table A.1: The PSO hyperparameters used throughout the analysis. In the first stage the PSO optimizes the semi-coherent likelihood with $N_{\max} = 1024$ segments using the hyperparameter settings in the first row of the table. In subsequent stages the number of segments is progressively reduced, e.g. to $N = 256$ in the second row, and the hyperparameters changed accordingly. The inertial weight Ω is reduced throughout the analysis and the social weight Φ_G is increased to gradually transition the behavior of the swarm from exploration to exploitation. The minimum velocities for each parameter ϵ^μ are also reduced during the run which also helps with the transition from exploration to exploitation. (The ϵ^μ parameters are dimensionfull and have the same units as the corresponding parameters in Tab. 4.1.)

Segment	Ω	Φ_P	Φ_G	ϵ^λ	$\epsilon^{\sin\beta}$	ϵ^{t_c}	$\epsilon^{\mathcal{M}_c}$	$\epsilon^{\delta\mu}$	ϵ^{χ_1}	ϵ^{χ_2}	$\epsilon^{\phi_{\rm left}}$	$\epsilon^{\phi_{\rm right}}$	$\epsilon^{\sqrt{A_{\rm left}}}$	$\epsilon^{\sqrt{A_{\rm right}}}$
1024	0.6	0.2	0.2	10^{-2}	10^{-2}	50	10^{-2}	10^{-1}	10^{-1}	10^{-1}	10^{-1}	10^{-1}	5×10^{-6}	5×10^{-6}
256	0.5	0.2	0.3	10^{-2}	10^{-2}	50	10^{-2}	10^{-1}	10^{-1}	10^{-1}	10^{-1}	10^{-1}	10^{-6}	10^{-6}
64	0.4	0.2	0.4	10^{-2}	10^{-2}	50	10^{-2}	5×10^{-2}	5×10^{-2}	5×10^{-2}	10^{-1}	10^{-1}	10^{-6}	10^{-6}
16	0.3	0.2	0.5	10^{-2}	10^{-2}	50	10^{-2}	10^{-2}	5×10^{-2}	5×10^{-2}	10^{-1}	10^{-1}	10^{-6}	10^{-6}
4	0.2	0.2	0.6	10^{-3}	10^{-3}	25	10^{-4}	10^{-2}	10^{-2}	10^{-2}	10^{-1}	10^{-1}	10^{-6}	10^{-6}
1	0.2	0.2	0.6	10^{-4}	10^{-4}	5	10^{-5}	10^{-3}	10^{-3}	10^{-3}	10^{-1}	10^{-1}	10^{-6}	10^{-6}

A.3 Parameter transforms

The following are the definitions of sampling parameters used in sections 4.4 and 4.5 for the SmBBH LISA analysis.

$$\mathcal{M}_c = \frac{(m_1 m_2)^{\frac{3}{5}}}{(m_1 + m_2)^{\frac{1}{5}}} \tag{A.2}$$

$$\delta\mu = \frac{m_1 - m_2}{m_1 + m_2} \tag{A.3}$$

$$\sqrt{A_{\rm left/right}} = \sqrt{\frac{1}{2d_L}}\Big(1 \pm \cos(\iota)\Big) \tag{A.4}$$

$$\phi_{\rm left/right} = \phi \mp 2\psi \tag{A.5}$$

A.4 Quadrature integration methods

The integral of a function $f(x)$ in the range $[x_{\min}, x_{\max}]$ can be approximated using the Riemann sum

$$\int_{x_{\min}}^{x_{\max}} f(x)\mathrm{d}x \approx \sum_{i=0}^{N_R-1} f(x_i)\Delta x, \tag{A.6}$$

where $\Delta x = (x_{\max} - x_{\min})/N_R$ and $x_i = x_{\min} + i\Delta x$. The set of points $\{x_i, i = 0, 1, \ldots, N_R\}$ constitute a discrete, uniform grid over which the integral is computed. Alternatively, this integral can be approximated using a quadrature rule;

$$\int_{x_{\min}}^{x_{\max}} f(x)\mathrm{d}x \approx \sum_{i=0}^{N_Q-1} w(x_i)f(x_i), \tag{A.7}$$

where the irregularly spaced x_i nodes are located at the roots of (suitably rescaled) Chebyshev polynomials and where i indexes the N_{Q} quadrature nodes. Quadrature numerical integration methods can achieve can typically achieve a given accuracy of approximation using a far smaller number of nodes ($N \ll N_{\mathrm{Riemann}}$) for smooth integrand functions f. The weights $w(x_i)$ only depend on the limits and can be pre-computed. These methods are commonly based on interpolating $f(x)$ over the domain using interpolation functions (Chebyshev polynomials in this case) which have known analytic integrals which are used to generate $w(x_i)$.

In this study, Clenshaw-Curtis quadrature is used to evaluate the semi-coherent likelihood in Eqn. 4.8 (specifically it is used to evaluate the noise weighted inner products). However, such methods are limited to integrands which are smooth over the integration domain due to the interpolation function usually being smooth. An example of smooth function where quadrature integration performs well is the $\langle h|h\rangle$ term in the log-likelihood (this is also equal to the squared SNR). Consider a signal $h(\theta) = A(f)e^{i\phi(f)}$, the squared

SNR is given by

$$\left\langle A(f)e^{i\phi(f)}\middle|A(f)e^{i\phi(f)}\right\rangle \sim \int |A(f)|^2 \mathrm{d}f, \tag{A.8}$$

where $A(f)$ is (usually) a smooth, slowly varying function over frequency (for clarity, we have omitted the factor of $S(f)$ and other terms that do not affect the argument here from Eqn. A.8). Thus this integral is well approximated by quadrature rules. Instead consider the case we have some data $d = h(\theta_*) = \tilde{A}(f)e^{i\tilde{\phi}(f)}$, where the parameters θ_* are the injected parameters of the source and we are performing a zero-noise injection. Consider the $\langle h|d\rangle$ term in the log-likelihood when $\theta \neq \theta_*$. The inner product is now

$$\langle h|d\rangle \sim \int A(f)\tilde{A}^\dagger(f)e^{-i(\phi(f)-\tilde{\phi}(f))}\mathrm{d}f. \tag{A.9}$$

The term $e^{-i(\phi(f)-\tilde{\phi}(f))}$ introduces oscillations into the integrand across the frequency domain. Assuming $\theta-\theta_*$ is small, $\phi(f)-\tilde{\phi}(f)$ will likely be small, in this regime quadrature rules are still valid. This is the case for parameter estimation of broadband signals shown in Refs. [65, 163, 201] where narrow priors are used. This method of evaluating the likelihood is only valid when $\theta - \theta_*$ is small. This sets a maximum size on each search tile. If the waveform model is evaluated at a location far from the injection in parameter space, $\phi(f)$ and $\tilde{\phi}(f)$ can be very different and the integrand oscillates rapidly over the frequency domain, quadrature rules are no longer valid. Note however that oscillatory integrands usually cancel to give small integrals. Introducing any source of rapid oscillations into the integrand will result in the quadrature approach failing. For example if the data d contains noise (which is discontinuous between frequency bins), the integrand for $\langle h|d\rangle$ is oscillatory and thus cannot be evaluated using quadrature integration. This issue with the likelihood computation was highlighted in Ref. [35]. We have verified the quadrature grid used in this study produces likelihoods that are sufficiently faithful to those evaluated on the FFT grid, for parameters within our prior bounds, in the zero-noise scenario.

Such problems with highly oscillatory integrals are not unique to quadrature rules, the

prevalent alternative to quadrature rules used in frequency domain analyses for mock LISA data is heterodyning/relative binning of the likelihood. This uses a template waveform (on the FFT frequency grid), and expresses a waveform model evaluated at another location in parameter space as a slowly varying difference between the model and the template waveform. This method is also limited in the distance one can travel in parameter space from the template before the model waveform becomes inaccurate [209]. Other methods evaluating this sort of oscillatory integral have been proposed, see Refs. [83, 35]. While it is not yet clear which method will be used for real, noisy LISA data, GPU hardware accelerated likelihoods are a promising avenue [236, 82] which circumvents the previously discussed problems by generating waveforms directly on the FFT frequency grid.

It is worth making a clear distinction between the type of quadrature method discussed here and those currently used in LVK analysis in the form of reduced-order quadrature [258]. Reduced-order models are a way of representing the waveform model in a basis of functions, such as:

$$\tilde{h}(f,\theta) = \sum_{i=1}^{N} a_i(\theta)\tilde{h}_i(f,\theta_i), \tag{A.10}$$

where $\tilde{h}_i(f)$ are the basis functions, usually these are waveforms evaluated at $\theta_i \neq \theta$. This corresponds to expressing a waveform evaluated at some position in parameter space θ using a linear combination of waveforms evaluated at other locations in parameter space (the selection of the basis waveforms is explained in further detail in Ref. [258]). Once a basis has been built, the process of evaluating a waveform at any given θ reduces to evaluating the coefficients $a_i(\theta)$, these can be evaluated efficiently using empirical-interpolation, see Sec. 4 of [258]. The empirical-interpolation method computes $a_i(\theta)$ by computing the *true* waveform (i.e not using the reduced basis) at a small number of frequencies/times, this sparse representation is then projected onto the basis element $\tilde{h}_i$ to compute $a_i(\theta)$. The crucial step here is that the waveform is only evaluated on a sparse

time/frequency grid. Using this representation of the waveform, the problematic term $\langle h|d\rangle$ is the following:

$$\langle h|d\rangle = \int \tilde{h}(f,\theta)\frac{\tilde{d}^*(f)}{S_n(f)}\mathrm{d}f \approx \int \left[\sum_i a_i(\theta)\tilde{h}_i(f,\theta_i)\right]\frac{\tilde{d}^*(f)}{S_n(f)}\mathrm{d}f, \tag{A.11}$$

$$\approx \sum_i a_i(\theta)\left[\int \tilde{h}_i(f,\theta_i)\frac{\tilde{d}^*(f)}{S_n(f)}\mathrm{d}f\right], \tag{A.12}$$

where the term inside the square brackets can be pre-computed prior to the analysis as a coefficient for every basis element. As a result, $\langle h|d\rangle$ collapses to:

$$\langle h|d\rangle = \sum_i a_i(\theta)\gamma_i, \tag{A.13}$$

where γ_i is the noise weighted inner product between the data and the basis elements. The similarity between the expression above and Eqn. A.7 is reason why this method is labelled as a quadrature rule. However, the key difference between what is shown here and the quadrature rule used in this chapter is that the inner product used to compute γ_i is being evaluated on the full dense frequency grid, thus this method is still valid when noise is present in the data.

B. Chapter 5 appendix

These appendices are slightly edited versions of appendices from Ref. [218]:

Diganta Bandopadhyay and Christopher J. Moore. GPU-accelerated semi-coherent hierarchical search for stellar-mass binary inspiral signals in LISA, published in Physical Review D, Volume 110:103026, November 2024.

Table B.1: Injection parameters and priors for the search of the GW190521-like fiducial source. Parameters in the top section of the table are those that are searched using uniform priors over the ranges shown, those in the middle section are injected but do not appear in the definition of the search statistic, and those in the bottom section are *derived* parameters and approximate prior ranges are given. At the point of transition between search and parameter estimation, each particle draws D_L uniformly from the prior range $[10, 300]$ Mpc. All masses are given in the detector frame. (Due to differences in the conventions between the waveform and the BBHx LISA response, the definition of the polarisation angle ψ acquires a negative sign.) Parameters that define the search tile are in highlighted rows. This source has an SNR of 21.45.

Parameter	Injection	Prior range: $[\theta_{\min}, \theta_{\max}]$
$\mathcal{M}_c\,[M_\odot]$	95.0209	$[93, 96]$
$f_{\rm low}$ [Hz]	0.0175	$[0.014, 0.019]$
η	0.234	$[0.15, 0.2495]$
λ [rad]	3.24	$[0, 2\pi]$
β [rad]	0.4	$[-\pi/2, \pi/2]$
i [rad]	2.0	$[0, \pi]$
ψ [rad]	-1.5	$[-\pi, 0]$
e_0	0.03	$[0.005, 0.1]$
ϕ_0 [rad]	0	-
D_L [Mpc]	200	-
$m_1\,[M_\odot]$	142	$[111.74, 244.58]$
$m_2\,[M_\odot]$	85	$[53.35, 105.47]$
t_c [months]	6.12	$\sim [5, 11]$

B.1 Linearly chirping signal

Throughout this paper, when signals and templates are split into semi-coherent segments, this splitting is done in the frequency domain such that each segment contains equal squared SNR. As discussed in the main text, for a simple source this can be related to splitting into equal duration segments in the time domain, this appendix demonstrates this. Consider a linearly chirping source with a constant amplitude,

$$h(t) = \cos\left(2\pi\left[ft + \frac{\dot{f}}{2}t^2\right]\right). \tag{B.1}$$

If the source is quasi-monochromatic (i.e. only increases its frequency in a narrowband across the observation period), $\dot{f}_{\rm GW} \ll f_{\rm GW}/T_{\rm obs}$, and the PSD can be approximated as constant across this narrow bandwidth (i.e. white noise), then the SNR can be computed in the time domain (Parseval's theorem). The cumulative square SNR up to a given time t is given by

$$\begin{aligned}
\rho^2(t) &\propto \int_0^t \cos^2\left(2\pi\left[ft' + \frac{\dot{f}}{2}t'^2\right]\right)\mathrm{d}t' \\
&\propto \frac{1}{2}\int_0^t \left[\cos\left(4\pi\left[ft' + \frac{\dot{f}}{2}t'^2\right]\right) + 1\right]\mathrm{d}t' \\
&\propto \frac{t}{2} + \frac{1}{2}\int_0^t \cos\left(4\pi\left[ft' + \frac{\dot{f}}{2}t'^2\right]\right)\mathrm{d}t'.
\end{aligned} \tag{B.2}$$

The second term evaluates to

$$\begin{aligned}
&\frac{1}{2\sqrt{\dot{f}}}\Bigg(\cos\left(\frac{2f^2\pi}{\dot{f}}\right)\left[-\mathcal{C}\left(\frac{2f}{\sqrt{\dot{f}}}\right) + \mathcal{C}\left(\frac{2(f+\dot{f}t)}{\dot{f}}\right)\right] + \\
&\qquad \sin\left(\frac{2f^2\pi}{\dot{f}}\right)\left[-\mathcal{S}\left(\frac{2f}{\sqrt{\dot{f}}}\right) + \mathcal{S}\left(\frac{2(f+\dot{f}t)}{\dot{f}}\right)\right]\Bigg),
\end{aligned} \tag{B.3}$$

where $\mathcal{C}$ and $\mathcal{S}$ are the Fresnel integrals. Fig. B.1 sketches this as a function of time. This is a highly oscillatory function that decays to 0 as $t \to \infty$. Therefore, for large observation

times this term is negligible and the squared SNR in Eqn. B.2 increases linearly with time.

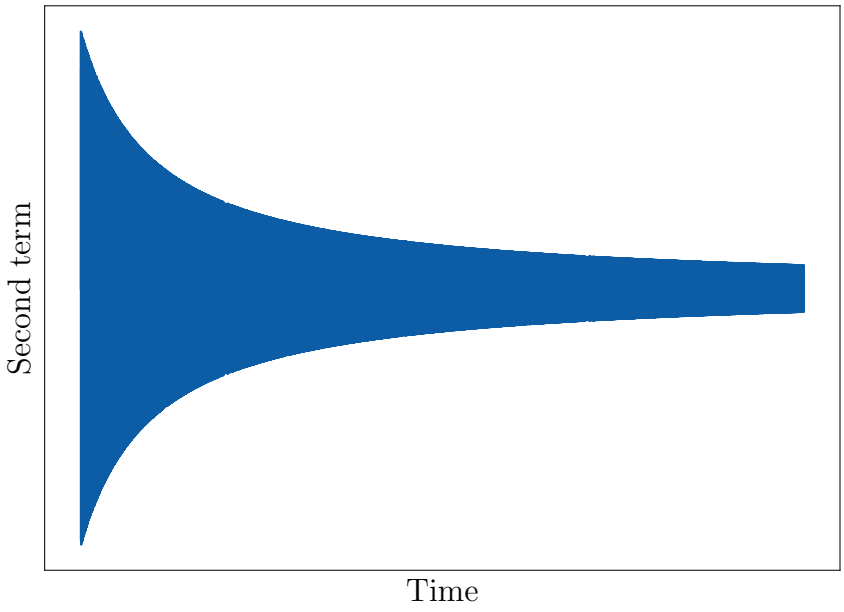

Figure B.1: Sketch showing the behaviour of the second term in the integral.

B.2 PSO & Inference parameters

In this appendix we summarise the parameters used for the PSO algorithm used in this paper. The search is initialized with 6 swarms, each with 10,000 particles and the total number of particles is kept constant throughout the search. All of the searches conducted in this chapter use the segment ladder $N = \{100, 50, 10, 1\}$. We have used $N_{\text{high}} = 100$ as this was estimated to be the minimum number of segments required for a search in Ref. [88] (this was computed for EMRIs, however the arguments still hold for SmBBH searches). A swarm is considered converged if the best value of the objective function in the history of the swarm does not improve by more than 2 in 50 iterations.

The minimum velocities in each dimension for each segment in the ladder are presented in Table B.2, alongside the 3 PSO hyper-parameters.

At the end of the optimization phase of the search, if a swarm passes the noise threshold, 100 PSO particles transition to MCMC walkers, evolving for 10,000 iterations.

Table B.2: PSO hyper-parameters and minimum velocities used throughout the search.

Number of segments (N)	Ω	Φ_P	Φ_G	$\epsilon^{\mathcal{M}_c}$	ϵ^{η}	ϵ^{β}	ϵ^{λ}	ϵ^{i}	ϵ^{ψ}	$\epsilon^{f_{\rm low}}$	ϵ^{e_0}
100	0.5	0.2	0.3	0.1	0.05	0.1	0.1	0.2	0.1	5×10^{-8}	0.1
50	0.5	0.2	0.4	0.01	0.05	0.01	0.01	0.2	0.1	1×10^{-8}	0.1
10	0.3	0.2	0.5	0.001	0.05	0.001	0.001	0.01	0.1	1×10^{-8}	0.1
1	0.72	1.193	1.193	0.0	0.0	0.0	0.0	0.0	0.0	0	0.0

B.3 GW190521-like search

This section contains results for a simulated search for a GW190521-like signal. The injected parameters and priors are shown in Table B.1. The results of the search are presented in Figs. 5.5 and B.2. For this search it is not possible to estimate the FAP because we did not perform a signal-free injection and search campaign in this search tile.

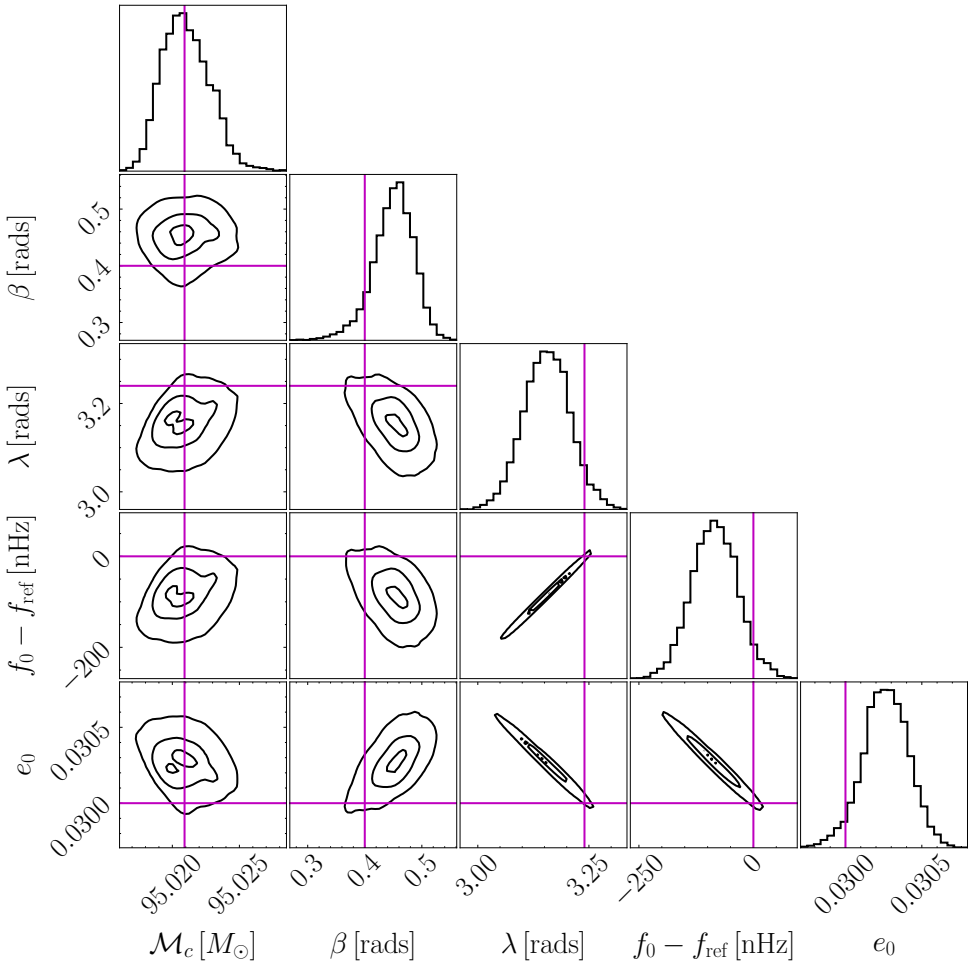

Figure B.2: Parameter estimation result from the search of the GW190521-like source with source parameters shown in Table B.1. Posteriors are plotted for the same parameters as in Fig. 5.6. For this source, $f_{\rm ref} = 8.75 \times 10^{-3}$ Hz.

B.4 Detection probability

For intermediate SNRs, the search only correctly identifies the signal some of the time. This detection probability is a function of SNR. This can be assessed by searching multiple injections at the same SNR in independent noise realizations. In this section we make a limited attempt of this by simulating 4 searches for signals at a given SNR, each with its own realization of the instrumental noise. This process is repeated across a range of SNRs. The results of this can be seen in Tables. B.3 and B.4. Based on these results the search seems to reliably find sources with $\rho \gtrsim 20$. For $\rho \lesssim 16$, the search fails. In the intermediate range of SNRs, there is a finite detection probability less than one. This should be caveated by the fact that this is all for one search tile in parameter space, as the search tile is moved to a different area in parameter space, the background noise distribution is likely to change, and therefore the threshold SNR for the search would also change.

Table B.3: Highest $\Upsilon_{N=1}$ from multiple searches conducted with the fiducial signal in Table 5.1, scaled to different SNR by varying D_L. The values in this table should be compared against the distribution of background triggers shown in the top panel of Fig. 5.5, confident candidates with FAP $\lesssim 1/100$ are shaded green, these represent successful searches. The cases where a confident candidate was not found are shaded in red. The corresponding match between the search result (waveform) against the injected source is shown in Table B.4.

SNR ρ	Run 1	Run 2	Run 3	Run 4
38	1450.1	1390.3	1422.3	1434.1
22	512.4	526.6	605.3	490.5
20	404.6	408.3	393.0	376.3
19	64.1	224.2	440.5	67.6
18	396.7	34.4	175.3	297.5
17	304.8	246.8	225.1	48.5
16	56.4	49.0	134.7	66.2

Table B.4: Matches against injected waveform for the highest significance triggers presented in Table B.3. Matches greater than 0.97 are shaded in green, those in the range $[0.5, 0.97]$ are shaded in yellow and those below 0.5 are highlighted in red.

SNR ρ	Run 1	Run 2	Run 3	Run 4
38	0.9991	0.9985	0.9988	0.999
22	0.9976	0.9874	0.9982	0.9987
20	0.9948	0.9921	0.9976	0.9995
19	0.003	0.7789	0.9993	0.2017
18	0.9962	0.0144	0.7078	0.9984
17	0.9938	0.9949	0.768	0.0
16	0.2424	0.1262	0.553	0.1634

Bibliography

[1] B. P. Abbott et al. Observation of Gravitational Waves from a Binary Black Hole Merger. PRL, 116(6):061102, February 2016.

[2] Albert Einstein. Näherungsweise Integration der Feldgleichungen der Gravitation. *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften*, pages 688–696, January 1916.

[3] B. P. Abbott et al. Tests of general relativity with GW150914. *Phys. Rev. Lett.*, 116(22):221101, 2016. [Erratum: Phys.Rev.Lett. 121, 129902 (2018)].

[4] J. M. Weisberg, J. H. Taylor, and L. A. Fowler. Gravitational waves from an orbiting pulsar. *Scientific American*, 245:74–82, October 1981.

[5] R. Abbott et al. GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo during the Second Part of the Third Observing Run. *Physical Review X*, 13(4):041039, October 2023.

[6] Michele Maggiore. *Gravitational Waves: Volume 1: Theory and Experiments*. 2007.

[7] Michele Maggiore. *Gravitational Waves: Volume 2: Astrophysics and Cosmology*. 2018.

[8] Cameron Mills and Stephen Fairhurst. Measuring gravitational-wave higher-order multipoles. PRD, 103(2):024042, January 2021.

[9] M. P. Hobson, G. P. Efstathiou, and A. N. Lasenby. *General Relativity: An Introduction for Physicists*. Cambridge University Press, 2006.

[10] Pierre Auclair et al. Cosmology with the Laser Interferometer Space Antenna. *Living Reviews in Relativity*, 26(1):5, December 2023.

[11] Pau Amaro-Seoane et al. Laser Interferometer Space Antenna. *arXiv e-prints*, page arXiv:1702.00786, February 2017.

[12] C. J. Moore, R. H. Cole, and C. P. L. Berry. Gravitational-wave sensitivity curves. *Classical and Quantum Gravity*, 32(1):015014, January 2015.

[13] Nikolaos Karnesis, Stanislav Babak, Mauro Pieroni, Neil Cornish, and Tyson Littenberg. Characterization of the stochastic signal originating from compact binary populations as measured by LISA. PRD, 104(4):043019, August 2021.

[14] Eliot Finch and Christopher J. Moore. Searching for a ringdown overtone in GW150914. PRD, 106(4):043005, August 2022.

[15] T Kupfer, V Korol, S Shah, G Nelemans, T R Marsh, G Ramsay, P J Groot, D T H Steeghs, and E M Rossi. Lisa verification binaries with updated distances from gaia data release 2. *Monthly Notices of the Royal Astronomical Society*, 480(1):302–309, 06 2018.

[16] Gravity Collaboration et al. Polarimetry and astrometry of NIR flares as event horizon scale, dynamical probes for the mass of Sgr A*. A&A, 677:L10, September 2023.

[17] S. Gillessen, F. Eisenhauer, S. Trippe, T. Alexander, R. Genzel, F. Martins, and T. Ott. Monitoring Stellar Orbits Around the Massive Black Hole in the Galactic Center. ApJ, 692(2):1075–1109, February 2009.

[18] A. M. Ghez, S. Salim, N. N. Weinberg, J. R. Lu, T. Do, J. K. Dunn, K. Matthews, M. R. Morris, S. Yelda, E. E. Becklin, T. Kremenek, M. Milosavljevic, and J. Naiman. Measuring Distance and Properties of the Milky Way's Central Supermassive Black Hole with Stellar Orbits. ApJ, 689(2):1044–1062, December 2008.

[19] Brandon C. Kelly and Andrea Merloni. Mass Functions of Supermassive Black Holes across Cosmic Time. *Advances in Astronomy*, 2012:970858, January 2012.

[20] Jillian Bellovary, Alyson Brooks, Monica Colpi, Michael Eracleous, Kelly Holley-Bockelmann, Ann Hornschemeier, Lucio Mayer, Priya Natarajan, Jacob Slutsky, and Michael Tremmel. Where are the Intermediate Mass Black Holes? *Bulletin of the American Astronomical Society*, 51(3):175, May 2019.

[21] R W Hellings and G S Downs. Upper limits on the isotropic gravitational radiation background from pulsar timing analysis. *Astrophys. J., Lett. Ed.; (United States)*, 265:2, 02 1983.

[22] Gabriella Agazie et al. The NANOGrav 15 yr Data Set: Evidence for a Gravitational-wave Background. *Astrophys. J. Lett.*, 951(1):L8, 2023.

[23] J. Antoniadis et al. The second data release from the European Pulsar Timing Array - III. Search for gravitational wave signals. *Astron. Astrophys.*, 678:A50, 2023.

[24] Nathan Steinle, Hannah Middleton, Christopher J. Moore, Siyuan Chen, Antoine Klein, Geraint Pratten, Riccardo Buscicchio, Eliot Finch, and Alberto Vecchio. Implications of pulsar timing array observations for LISA detections of massive black hole binaries. MNRAS, 525(2):2851–2863, October 2023.

[25] Monica Colpi et al. LISA Definition Study Report. *arXiv e-prints*, page arXiv:2402.07571, February 2024.

[26] Monica Colpi, Kelly Holley-Bockelmann, Tamara Bogdanovic, Priya Natarajan, Jillian Bellovary, Alberto Sesana, Michael Tremmel, Jeremy Schnittman, Julia Comerford, Enrico Barausse, Emanuele Berti, Marta Volonteri, Fazeel Khan, Sean McWilliams, Sarah Burke-Spolaor, Jeff Hazboun, John Conklin, Guido Mueller, and Shane Larson. Astro2020 science white paper: The gravitational wave view of massive black holes. *arXiv e-prints*, page arXiv:1903.06867, March 2019.

[27] Michael L. Katz and Shane L. Larson. Evaluating black hole detectability with LISA. MNRAS, 483(3):3108–3118, March 2019.

[28] C. L. Bennett, D. Larson, J. L. Weiland, and G. Hinshaw. The 1% Concordance Hubble Constant. ApJ, 794(2):135, October 2014.

[29] R. Barkana and A. Loeb. *In the Beginning: The First Sources of Light and the Reionization of the Universe*. Physics reports. Elsevier, 2001.

[30] Feige Wang, Jinyi Yang, Xiaohui Fan, Joseph F. Hennawi, Aaron J. Barth, Eduardo Banados, Fuyan Bian, Konstantina Boutsia, Thomas Connor, Frederick B. Davies, Roberto Decarli, Anna-Christina Eilers, Emanuele Paolo Farina, Richard Green, Linhua Jiang, Jiang-Tao Li, Chiara Mazzucchelli, Riccardo Nanni, Jan-Torge Schindler, Bram Venemans, Fabian Walter, Xue-Bing Wu, and Minghao Yue. A luminous quasar at redshift 7.642. *The Astrophysical Journal Letters*, 907(1):L1, January 2021.

[31] B. E. Robertson et al. Identification and properties of intense star-forming galaxies at redshifts z $>$ 10. *Nature Astronomy*, 7:611–621, May 2023.

[32] Ivo Labbé et al. A population of red candidate massive galaxies 600 Myr after the Big Bang. Nature, 616(7956):266–269, April 2023.

[33] Geraint Pratten, Patricia Schmidt, Hannah Middleton, and Alberto Vecchio. Precision tracking of massive black hole spin evolution with LISA. PRD, 108(12):124045, December 2023.

[34] Mudit Garg, Andrea Derdzinski, Shubhanshu Tiwari, Jonathan Gair, and Lucio Mayer. Measuring eccentricity and gas-induced perturbation from gravitational waves of LISA massive black hole binaries. MNRAS, 532(4):4060–4074, August 2024.

[35] Sylvain Marsat, John G. Baker, and Tito Dal Canton. Exploring the Bayesian parameter estimation of binary black holes with LISA. PRD, 103(8):083011, April 2021.

[36] Nicola Tamanini. Late time cosmology with LISA: Probing the cosmic expansion with massive black hole binary mergers as standard sirens. In *Journal of Physics Conference Series*, volume 840 of *Journal of Physics Conference Series*, page 012029. IOP, May 2017.

[37] Alberto Mangiagli, Chiara Caprini, Sylvain Marsat, Lorenzo Speri, Robert R. Caldwell, and Nicola Tamanini. Massive black hole binaries in LISA: Constraining cosmological parameters at high redshifts. PRD, 111(8):083043, April 2025.

[38] Walter Del Pozzo. Inference of the cosmological parameters from gravitational waves: application to second generation interferometers. *Phys. Rev. D*, 86:043011, 2012.

[39] Renjie Wang, Wen-Hong Ruan, Qing Yang, Zong-Kuan Guo, Rong-Gen Cai, and Bin Hu. Hubble parameter estimation via dark sirens with the lisa-taiji network. *National Science Review*, 9(2):nwab054, 04 2021.

[40] Rachel Gray, Ignacio Magaña Hernandez, Hong Qi, Ankan Sur, Patrick R. Brady, Hsin-Yu Chen, Will M. Farr, Maya Fishbach, Jonathan R. Gair, Archisman Ghosh, Daniel E. Holz, Simone Mastrogiovanni, Christopher Messenger, Danièle A. Steer, and John Veitch. Cosmological inference using gravitational wave standard sirens: A mock data analysis. PRD, 101(12):122001, June 2020.

[41] Eduardo M. Gutiérrez, Luciano Combi, Scott C. Noble, Manuela Campanelli, Julian H. Krolik, Federico López Armengol, and Federico García. Electromagnetic Signatures from Supermassive Binary Black Holes Approaching Merger. ApJ, 928(2):137, April 2022.

[42] Gareth Cabourn Davies, Ian Harry, Michael J. Williams, Diganta Bandopadhyay, Leor Barack, Jean-Baptiste Bayle, Charlie Hoy, Antoine Klein, Hannah Middleton, Christopher J. Moore, Laura Nuttall, Geraint Pratten, Alberto Vecchio, and Graham Woan. Premerger observation and characterization of massive black hole binaries. *Phys. Rev. D*, 111:043045, Feb 2025.

[43] Daniel J. D'Orazio and Maria Charisi. Observational Signatures of Supermassive Black Hole Binaries. *arXiv e-prints*, page arXiv:2310.16896, October 2023.

[44] Bernard J. Kelly, Zachariah B. Etienne, Jacob Golomb, Jeremy D. Schnittman, John G. Baker, Scott C. Noble, and Geoffrey Ryan. Electromagnetic emission from a binary black hole merger remnant in plasma: Field alignment and plasma temperature. PRD, 103(6):063039, March 2021.

[45] Luke Major Krauth, Jordy Davelaar, Zoltán Haiman, John Ryan Westernacher-Schneider, Jonathan Zrake, and Andrew MacFadyen. Thermal X-ray signatures in late-stage unequal-mass massive black hole binary mergers. 3 2025.

[46] Olaf Dreyer, Bernard Kelly, Badri Krishnan, Lee Samuel Finn, David Garrison, and Ramon Lopez-Aleman. Black-hole spectroscopy: testing general relativity through gravitational-wave observations. *Classical and Quantum Gravity*, 21(4):787–803, February 2004.

[47] Swetha Bhagwat, Costantino Pacilio, Enrico Barausse, and Paolo Pani. Landscape of massive black-hole spectroscopy with LISA and the Einstein Telescope. PRD, 105(12):124063, June 2022.

[48] Christopher J. Moore, Eliot Finch, Antoine Klein, Valeriya Korol, Nhat Pham, and Daniel Robins. Discovering neutron stars with LISA via measurements of orbital eccentricity in galactic binaries. MNRAS, 531(2):2817–2829, June 2024.

[49] Pau Amaro-Seoane et al. Astrophysics with the Laser Interferometer Space Antenna. *Living Reviews in Relativity*, 26(1):2, December 2023.

[50] Astrid Lamberts, Sarah Blunt, Tyson B. Littenberg, Shea Garrison-Kimmel, Thomas Kupfer, and Robyn E. Sanderson. Predicting the LISA white dwarf binary population in the Milky Way with cosmological simulations. MNRAS, 490(4):5888–5903, December 2019.

[51] G. Nelemans, L. R. Yungelson, and S. F. Portegies Zwart. The gravitational wave signal from the Galactic disk population of binaries containing two compact objects. A&A, 375:890–898, September 2001.

[52] Peter L. Bender and Dieter Hils. Confusion noise level due to galactic and extragalactic binaries. *Classical and Quantum Gravity*, 14(6):1439–1444, June 1997.

[53] Neil Cornish and Travis Robson. Galactic binary science with the new LISA design. In *Journal of Physics Conference Series*, volume 840 of *Journal of Physics Conference Series*, page 012024. IOP, May 2017.

[54] M. Benacquista and K. Holley-Bockelmann. Consequences of Disk Scale Height on LISA Confusion Noise from Close White Dwarf Binaries. ApJ, 645(1):589–596, July 2006.

[55] Maria Georgousi, Nikolaos Karnesis, Valeriya Korol, Mauro Pieroni, and Nikolaos Stergioulas. Gravitational waves from double white dwarfs as probes of the milky way. MNRAS, 519(2):2552–2566, February 2023.

[56] Matthew C. Digman and Neil J. Cornish. LISA Gravitational Wave Sources in a Time-varying Galactic Stochastic Background. ApJ, 940(1):10, November 2022.

[57] Valeriya Korol, Elena M. Rossi, Paul J. Groot, Gijs Nelemans, Silvia Toonen, and Anthony G. A. Brown. Prospects for detection of detached double white dwarf binaries with Gaia, LSST and LISA. MNRAS, 470(2):1894–1910, September 2017.

[58] A. Stroeer and A. Vecchio. The LISA verification binaries. *Classical and Quantum Gravity*, 23(19):S809–S817, October 2006.

[59] Eliot Finch, Giorgia Bartolucci, Daniel Chucherko, Ben G. Patterson, Valeriya Korol, Antoine Klein, Diganta Bandopadhyay, Hannah Middleton, Christopher J. Moore, and Alberto Vecchio. Identifying LISA verification binaries among the Galactic population of double white dwarfs. MNRAS, 522(4):5358–5373, July 2023.

[60] Thomas Kupfer, Valeriya Korol, Tyson B. Littenberg, Sweta Shah, Etienne Savalle, Paul J. Groot, Thomas R. Marsh, Maude Le Jeune, Gijs Nelemans, Anna F. Pala, Antoine Petiteau, Gavin Ramsay, Danny Steeghs, and Stanislav Babak. LISA Galactic Binaries with Astrometry from Gaia DR3. ApJ, 963(2):100, March 2024.

[61] Valeriya Korol, Andrei P. Igoshev, Silvia Toonen, Nikolaos Karnesis, Christopher J. Moore, Eliot Finch, and Antoine Klein. Neutron star - white dwarf binaries: probing formation pathways and natal kicks with LISA. MNRAS, 530(1):844–860, May 2024.

[62] Joheen Chakraborty et al. Expanding the Ultracompacts: Gravitational-wave-driven Mass Transfer in the Shortest-period Binaries with Accretion Disks. ApJ, 977(2):262, December 2024.

[63] Michael L. Katz, Nikolaos Karnesis, Natalia Korsakova, Jonathan R. Gair, and Nikolaos Stergioulas. An efficient GPU-accelerated multi-source global fit pipeline for LISA data analysis. *arXiv e-prints*, page arXiv:2405.04690, May 2024.

[64] Alberto Sesana. Prospects for Multiband Gravitational-Wave Astronomy after GW150914. PRL, 116(23):231102, June 2016.

[65] Antoine Klein, Geraint Pratten, Riccardo Buscicchio, Patricia Schmidt, Christopher J. Moore, Eliot Finch, Alice Bonino, Lucy M. Thomas, Natalie Williams, Davide Gerosa, Sean McGee, Matt Nicholl, and Alberto Vecchio. The last three years: multiband gravitational-wave observations of stellar-mass binary black holes. *arXiv e-prints*, page arXiv:2204.03423, April 2022.

[66] Nihar Gupte, Antoni Ramos-Buades, Alessandra Buonanno, Jonathan Gair, M. Coleman Miller, Maximilian Dax, Stephen R. Green, Michael Pürrer, Jonas

Wildberger, Jakob Macke, Isobel M. Romero-Shaw, and Bernhard Schölkopf. Evidence for eccentricity in the population of binary black holes observed by LIGO-Virgo-KAGRA. *arXiv e-prints*, page arXiv:2404.14286, April 2024.

[67] A. G. Abac et al. Search for Eccentric Black Hole Coalescences during the Third Observing Run of LIGO and Virgo. ApJ, 973(2):132, October 2024.

[68] Gonzalo Morras, Geraint Pratten, and Patricia Schmidt. Orbital eccentricity in a neutron star - black hole binary. *arXiv e-prints*, page arXiv:2503.15393, March 2025.

[69] Pankaj Saini. Resolving the eccentricity of stellar mass binary black holes with next generation ground-based gravitational wave detectors. MNRAS, 528(1):833–842, February 2024.

[70] Isobel M. Romero-Shaw, Paul D. Lasky, and Eric Thrane. Signs of Eccentricity in Two Gravitational-wave Signals May Indicate a Subpopulation of Dynamically Assembled Binary Black Holes. *Astrophys. J. Lett.*, 921(2):L31, 2021.

[71] P. C. Peters and J. Mathews. Gravitational radiation from point masses in a Keplerian orbit. *Phys. Rev.*, 131:435–439, 1963.

[72] Atsushi Nishizawa, Emanuele Berti, Antoine Klein, and Alberto Sesana. eLISA eccentricity measurements as tracers of binary black hole formation. PRD, 94(6):064020, September 2016.

[73] Lisa Randall and Zhong-Zhi Xianyu. Eccentricity Without Measuring Eccentricity: Discriminating Among Stellar Mass Black Hole Binary Formation Channels. *arXiv e-prints*, page arXiv:1907.02283, July 2019.

[74] B. P. Abbott et al. Gravitational Waves and Gamma-Rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A. ApJ, 848(2):L13, October 2017.

[75] S. E. de Mink and A. King. Electromagnetic Signals Following Stellar-mass Black Hole Mergers. ApJ, 839(1):L7, April 2017.

[76] Shu-Xu Yi and K. S. Cheng. Where Are the Electromagnetic-wave Counterparts of Stellar-mass Binary Black Hole Mergers? ApJ, 884(1):L12, October 2019.

[77] Alexandre Toubiana, Sylvain Marsat, Stanislav Babak, Enrico Barausse, and John Baker. Tests of general relativity with stellar-mass black hole binaries observed by lisa. *Phys. Rev. D*, 101:104038, May 2020.

[78] Tessa Baker, Enrico Barausse, Anson Chen, Claudia de Rham, Mauro Pieroni, and Gianmassimo Tasinato. Testing gravitational wave propagation with multiband detections. J. Cosmology Astropart. Phys., 2023(3):044, March 2023.

[79] Jonathan R. Gair, Stanislav Babak, Alberto Sesana, Pau Amaro-Seoane, Enrico Barausse, Christopher P. L. Berry, Emanuele Berti, and Carlos Sopuerta. Prospects for observing extreme-mass-ratio inspirals with LISA. In *Journal of Physics Conference Series*, volume 840 of *Journal of Physics Conference Series*, page 012021, May 2017.

[80] Fintan D. Ryan. Gravitational waves from the inspiral of a compact object into a massive, axisymmetric body with arbitrary multipole moments. *Phys. Rev. D*, 52:5707–5718, Nov 1995.

[81] Fintan D. Ryan. Accuracy of estimating the multipole moments of a massive body from the gravitational waves of a binary inspiral. *Phys. Rev. D*, 56:1845–1855, Aug 1997.

[82] Lorenzo Speri, Michael L. Katz, Alvin J. K. Chua, Scott A. Hughes, Niels Warburton, Jonathan E. Thompson, Christian E. A. Chapman-Bird, and Jonathan R. Gair. Fast and Fourier: Extreme Mass Ratio Inspiral Waveforms in the Frequency Domain. *arXiv e-prints*, page arXiv:2307.12585, July 2023.

[83] Alvin J. K. Chua, Michael L. Katz, Niels Warburton, and Scott A. Hughes. Rapid generation of fully relativistic extreme-mass-ratio-inspiral waveform templates for lisa data analysis. *Phys. Rev. Lett.*, 126:051102, Feb 2021.

[84] Stanislav Babak, Jonathan Gair, Alberto Sesana, Enrico Barausse, Carlos F. Sopuerta, Christopher P. L. Berry, Emanuele Berti, Pau Amaro-Seoane, Antoine Petiteau, and Antoine Klein. Science with the space-based interferometer LISA. V. Extreme mass-ratio inspirals. PRD, 95(10):103012, May 2017.

[85] S. Sigurdsson and M. J. Rees. Capture of stellar mass compact objects by massive black holes in galactic cusps. MNRAS, 284(2):318–326, January 1997.

[86] A. Derdzinski, D. D'Orazio, P. Duffell, Z. Haiman, and A. MacFadyen. Evolution of gas disc-embedded intermediate mass ratio inspirals in the LISA band. MNRAS, 501(3):3540–3557, March 2021.

[87] Lorenzo Speri, Andrea Antonelli, Laura Sberna, Stanislav Babak, Enrico Barausse, Jonathan R. Gair, and Michael L. Katz. Measuring accretion-disk effects with gravitational waves from extreme mass ratio inspirals. *arXiv e-prints*, page arXiv:2207.10086, July 2022.

[88] Jonathan R. Gair, Leor Barack, Teviet Creighton, Curt Cutler, Shane L. Larson, E. Sterl Phinney, and Michele Vallisneri. Event rate estimates for LISA extreme mass ratio capture sources. *Classical and Quantum Gravity*, 21(20):S1595–S1606, October 2004.

[89] Alvin J. K. Chua and Curt J. Cutler. Nonlocal parameter degeneracy in the intrinsic space of gravitational-wave signals from extreme-mass-ratio inspirals. PRD, 106(12):124046, December 2022.

[90] Alexander H. Nitz, Tito Dal Canton, Derek Davis, and Steven Reyes. Rapid detection of gravitational waves from compact binary mergers with PyCBC Live. PRD, 98(2):024050, July 2018.

[91] Bruce Allen. χ^2 time-frequency discriminator for gravitational wave detection. PRD, 71(6):062001, March 2005.

[92] Adam Coogan, Thomas D. P. Edwards, Horng Sheng Chia, Richard N. George, Katherine Freese, Cody Messick, Christian N. Setzer, Christoph Weniger, and Aaron Zimmerman. Efficient gravitational wave template bank generation with differentiable waveforms. *Phys. Rev. D*, 106:122001, Dec 2022.

[93] Benjamin J. Owen and B. S. Sathyaprakash. Matched filtering of gravitational waves from inspiraling compact binaries: Computational cost and template placement. *Phys. Rev. D*, 60:022002, Jun 1999.

[94] Benjamin J. Owen. Search templates for gravitational waves from inspiraling binaries: Choice of template spacing. PRD, 53(12):6749–6761, June 1996.

[95] B. S. Sathyaprakash and S. V. Dhurandhar. Choice of filters for the detection of gravitational waves from coalescing binaries. *Phys. Rev. D*, 44:3819–3834, Dec 1991.

[96] S. V. Dhurandhar and B. S. Sathyaprakash. Choice of filters for the detection of gravitational waves from coalescing binaries. ii. detection in colored noise. *Phys. Rev. D*, 49:1707–1722, Feb 1994.

[97] I. W. Harry, B. Allen, and B. S. Sathyaprakash. Stochastic template placement algorithm for gravitational wave data analysis. PRD, 80(10):104014, November 2009.

[98] Connor R. Weaving, Laura K. Nuttall, Ian W. Harry, Shichao Wu, and Alexander Nitz. Adapting the PyCBC pipeline to find and infer the properties of gravitational waves from massive black hole binaries in LISA. *Classical and Quantum Gravity*, 41(2):025006, January 2024.

[99] R. Abbott et al. All-sky search for continuous gravitational waves from iso-

lated neutron stars using Advanced LIGO and Advanced Virgo O3 data. PRD, 106(10):102008, November 2022.

[100] Javier Roulet, Liang Dai, Tejaswi Venumadhav, Barak Zackay, and Matias Zaldarriaga. Template bank for compact binary coalescence searches in gravitational wave data: A general geometric placement algorithm. PRD, 99(12):123022, June 2019.

[101] Shio Sakon, Leo Tsukada, Heather Fong, James Kennington, Wanting Niu, Chad Hanna, Shomik Adhicary, Pratyusava Baral, Amanda Baylor, Kipp Cannon, Sarah Caudill, Bryce Cousins, Jolien D. E. Creighton, Becca Ewing, Richard N. George, Patrick Godwin, Reiko Harada, Yun-Jing Huang, Rachael Huxford, Prathamesh Joshi, Soichiro Kuwahara, Alvin K. Y. Li, Ryan Magee, Duncan Meacher, Cody Messick, Soichiro Morisaki, Debnandini Mukherjee, Alex Pace, Cort Posnansky, Anarya Ray, Surabhi Sachdev, Divya Singh, Ron Tapia, Takuya Tsutsui, Koh Ueno, Aaron Viets, Leslie Wade, Madeline Wade, and Jonathan Wang. Template bank for compact binary mergers in the fourth observing run of Advanced LIGO, Advanced Virgo, and KAGRA. PRD, 109(4):044066, February 2024.

[102] S. Babak, R. Biswas, P. R. Brady, D. A. Brown, K. Cannon, C. D. Capano, J. H. Clayton, T. Cokelaer, J. D. E. Creighton, T. Dent, A. Dietz, S. Fairhurst, N. Fotopoulos, G. González, C. Hanna, I. W. Harry, G. Jones, D. Keppel, D. J. A. McKechan, L. Pekowsky, S. Privitera, C. Robinson, A. C. Rodriguez, B. S. Sathyaprakash, A. S. Sengupta, M. Vallisneri, R. Vaulin, and A. J. Weinstein. Searching for gravitational waves from binary coalescence. PRD, 87(2):024033, January 2013.

[103] Christopher J. Moore, Davide Gerosa, and Antoine Klein. Are stellar-mass black-hole binaries too quiet for LISA? MNRAS, 488(1):L94–L98, September 2019.

[104] J. Veitch, V. Raymond, B. Farr, W. Farr, P. Graff, S. Vitale, B. Aylott, K. Blackburn, N. Christensen, M. Coughlin, W. Del Pozzo, F. Feroz, J. Gair, C. J. Haster,

V. Kalogera, T. Littenberg, I. Mandel, R. O'Shaughnessy, M. Pitkin, C. Rodriguez, C. Röver, T. Sidery, R. Smith, M. Van Der Sluys, A. Vecchio, W. Vousden, and L. Wade. Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library. PRD, 91(4):042003, February 2015.

[105] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of State Calculations by Fast Computing Machines. Journal of Chemical Physics, 21(6):1087–1092, June 1953.

[106] Radford M. Neal. Slice Sampling. *arXiv e-prints*, page physics/0009028, September 2000.

[107] John Skilling. Nested sampling for general Bayesian computation. *Bayesian Analysis*, 1(4):833–859, 2006.

[108] Piotr Jaranowski, Andrzej Królak, and Bernard F. Schutz. Data analysis of gravitational-wave signals from spinning neutron stars: The signal and its detection. PRD, 58(6):063001, September 1998.

[109] Maximilian Dax, Stephen R. Green, Jonathan Gair, Jakob H. Macke, Alessandra Buonanno, and Bernhard Schölkopf. Real-Time Gravitational Wave Science with Neural Posterior Estimation. PRL, 127(24):241103, December 2021.

[110] Guy Moss, Vjeran Višnjević, Olaf Eisen, Falk M. Oraschewski, Cornelius Schröder, Jakob H. Macke, and Reinhard Drews. Simulation-Based Inference of Surface Accumulation and Basal Melt Rates of an Antarctic Ice Shelf from Isochronal Layers. *arXiv e-prints*, page arXiv:2312.02997, December 2023.

[111] Danilo Jimenez Rezende and Shakir Mohamed. Variational Inference with Normalizing Flows. *arXiv e-prints*, page arXiv:1505.05770, May 2015.

[112] Jonas Wildberger, Maximilian Dax, Stephen R. Green, Jonathan Gair, Michael Pürrer, Jakob H. Macke, Alessandra Buonanno, and Bernhard Schölkopf. Adapting to noise distribution shifts in flow-based gravitational-wave inference. PRD, 107(8):084046, April 2023.

[113] Iván Martín Vílchez and Carlos F. Sopuerta. Efficient Massive Black Hole Binary parameter estimation for LISA using Sequential Neural Likelihood. *arXiv e-prints*, page arXiv:2406.00565, June 2024.

[114] James Alvey, Uddipta Bhardwaj, Valerie Domcke, Mauro Pieroni, and Christoph Weniger. Simulation-based inference for stochastic gravitational wave background data analysis. PRD, 109(8):083008, April 2024.

[115] Poulami Dutta Roy, Parthapratim Mahapatra, Anuradha Samajdar, and K. G. Arun. Identifying intermediate mass binary black hole mergers in AGN disks using LISA. *arXiv e-prints*, page arXiv:2503.11721, March 2025.

[116] Gabriel Andres Piovano, Andrea Maselli, and Paolo Pani. Constraining the tidal deformability of supermassive objects with extreme mass ratio inspirals and semi-analytical frequency-domain waveforms. *Phys. Rev. D*, 107:024021, Jan 2023.

[117] Gabriel Andres Piovano, Richard Brito, Andrea Maselli, and Paolo Pani. Assessing the detectability of the secondary spin in extreme mass-ratio inspirals with fully relativistic numerical waveforms. PRD, 104(12):124019, December 2021.

[118] Luigi Ferraioli, Edward K. Porter, and Eric Plagnol. Bayesian Inference for LISA Pathfinder using Markov Chain Monte Carlo Methods. *arXiv e-prints*, page arXiv:1211.7183, November 2012.

[119] John D. Veitch. *Applications of Markov Chain Monte Carlo methods to continuous gravitational wave data analysis*. PhD thesis, University of Glasgow, 2007.

[120] Francesco Iacovelli, Michele Mancarella, Stefano Foffa, and Michele Maggiore. GWFAST: A Fisher Information Matrix Python Code for Third-generation Gravitational-wave Detectors. ApJS, 263(1):2, November 2022.

[121] B. P. Abbott et al. GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. PRL, 119(16):161101, October 2017.

[122] Neil J. Cornish and Jeff Crowder. LISA data analysis using Markov chain Monte Carlo methods. PRD, 72(4):043005, August 2005.

[123] E. Colin Cherry. Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America*, 25(5):975–979, 09 1953.

[124] Stefan H. Strub, Luigi Ferraioli, Cédric Schmelzbach, Simon C. Stähler, and Domenico Giardini. Accelerating global parameter estimation of gravitational waves from Galactic binaries using a genetic algorithm and GPUs. PRD, 108(10):103018, November 2023.

[125] Jun S. Liu. The collapsed gibbs sampler in bayesian computations with applications to a gene regulation problem. *Journal of the American Statistical Association*, 89(427):958–966, 1994.

[126] Senwen Deng, Stanislav Babak, Maude Le Jeune, Sylvain Marsat, Éric Plagnol, and Andrea Sartirana. Modular global-fit pipeline for LISA data analysis. *arXiv e-prints*, page arXiv:2501.10277, January 2025.

[127] Stefan H. Strub, Luigi Ferraioli, Cédric Schmelzbach, Simon C. Stähler, and Domenico Giardini. Global Analysis of LISA Data with Galactic Binaries and Massive Black Hole Binaries. *arXiv e-prints*, page arXiv:2403.15318, March 2024.

[128] Tyson B. Littenberg and Neil J. Cornish. Prototype global analysis of LISA data with multiple source types. PRD, 107(6):063004, March 2023.

[129] Neil J. Cornish and Tyson B. Littenberg. Bayeswave: Bayesian inference for gravitational wave bursts and instrument glitches. *Classical and Quantum Gravity*, 32(13):135012, July 2015.

[130] Peter J. Green. Reversible jump markov chain monte carlo computation and bayesian model determination. *Biometrika*, 82(4):711–732, 12 1995.

[131] M. Armano et al. Beyond the required lisa free-fall performance: New lisa pathfinder results down to 20 μHz. *Phys. Rev. Lett.*, 120:061101, Feb 2018.

[132] Christian Edward Anthony Chapman-Bird. *Extracting astrophysics from long-lived gravitational wave signals with the Laser Interferometer Space Antenna*. PhD thesis, Glasgow U., 2024.

[133] LISA Science Study Team. Lisa science requirements document. May 2018.

[134] Ollie Burke, Sylvain Marsat, Jonathan R. Gair, and Michael L. Katz. Mind the gap: addressing data gaps and assessing noise mismodeling in LISA. *arXiv e-prints*, page arXiv:2502.17426, February 2025.

[135] Curt Cutler and Éanna E. Flanagan. Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral waveform? *Phys. Rev. D*, 49:2658–2697, Mar 1994.

[136] Antoine Klein, Yannick Boetzel, Achamveedu Gopakumar, Philippe Jetzer, and Lorenzo de Vittori. Fourier domain gravitational waveforms for precessing eccentric binaries. PRD, 98(10):104043, November 2018.

[137] Antoine Klein, Neil Cornish, and Nicolás Yunes. Fast frequency-domain waveforms for spin-precessing binary inspirals. PRD, 90(12):124029, December 2014.

[138] Sylvain Marsat and John G. Baker. Fourier-domain modulations and delays of gravitational-wave signals. *arXiv e-prints*, page arXiv:1806.10734, June 2018.

[139] Neil J. Cornish and Louis J. Rubbo. Lisa response function. *Phys. Rev. D*, 67:022001, Jan 2003.

[140] M. Tinto, G. Giampieri, R. W. Hellings, P. L. Bender, and J. E. Faller. Algorithms for unequal-arm Michelson interferometers. In *7th Marcel Grossmann Meeting on General Relativity (MG 7)*, pages 1668–1670, 7 1994.

[141] Massimo Tinto and Sanjeev V. Dhurandhar. Time-delay interferometry. *Living Reviews in Relativity*, 24(1):1, December 2021.

[142] Daniel A. Shaddock, Massimo Tinto, Frank B. Estabrook, and J. W. Armstrong. Data combinations accounting for LISA spacecraft motion. PRD, 68(6):061303, September 2003.

[143] Neil J. Cornish and Ronald W. Hellings. The effects of orbital motion on LISA time delay interferometry. *Classical and Quantum Gravity*, 20(22):4851–4860, November 2003.

[144] Olaf Hartwig, Marc Lilley, Martina Muratore, and Mauro Pieroni. Stochastic gravitational wave background reconstruction for a nonequilateral and unequal-noise lisa constellation. *Phys. Rev. D*, 107:123531, Jun 2023.

[145] M. Armano et al. Transient acceleration events in LISA Pathfinder data: Properties and possible physical origin. PRD, 106(6):062001, September 2022.

[146] Matthew C. Edwards, Patricio Maturana-Russel, Renate Meyer, Jonathan Gair, Natalia Korsakova, and Nelson Christensen. Identifying and addressing nonstationary LISA noise. PRD, 102(8):084062, October 2020.

[147] Matthew C. Digman and Neil J. Cornish. Parameter estimation for stellar-origin black hole mergers in LISA. PRD, 108(2):023022, July 2023.

[148] Rodrigo Tenorio and Davide Gerosa. SFTs: a scalable data-analysis framework

for long-duration gravitational-wave signals. *arXiv e-prints*, page arXiv:2502.11823, February 2025.

[149] Diganta Bandopadhyay and Christopher J. Moore. LISA stellar-mass black hole searches with semicoherent and particle-swarm methods. PRD, 108(8):084014, October 2023.

[150] Pau Amaro-Seoane and Lucía Santamaría. Detection of IMBHs with Ground-based Gravitational Wave Observatories: A Biography of a Binary of Black Holes, from Birth to Death. ApJ, 722(2):1197–1206, October 2010.

[151] Pau Amaro-Seoane, Jonathan R. Gair, Marc Freitag, M. Coleman Miller, Ilya Mandel, Curt J. Cutler, and Stanislav Babak. TOPICAL REVIEW: Intermediate and extreme mass-ratio inspirals—astrophysics, science applications and detection using LISA. *Classical and Quantum Gravity*, 24(17):R113, September 2007.

[152] LIGO Scientific Collaboration. Advanced LIGO. *Classical and Quantum Gravity*, 32(7):074001, April 2015.

[153] F. Acernese et al. Advanced Virgo: a second-generation interferometric gravitational wave detector. *Classical and Quantum Gravity*, 32(2):024001, January 2015.

[154] R. Abbott et al. GW190521: A Binary Black Hole Merger with a Total Mass of 150 $M_{\odot}$. PRL, 125(10):101102, September 2020.

[155] Christopher P. L. Berry, Scott A. Hughes, Carlos F. Sopuerta, Alvin J. K. Chua, Anna Heffernan, Kelly Holley-Bockelmann, Deyan P. Mihaylov, M. Coleman Miller, and Alberto Sesana. The unique potential of extreme mass-ratio inspirals for gravitational-wave astronomy. 3 2019.

[156] Stanislav Babak, Jonathan R. Gair, and Edward K. Porter. An algorithm for the detection of extreme mass ratio inspirals in LISA data. *Classical and Quantum Gravity*, 26(13):135004, July 2009.

[157] Neil J. Cornish. Detection strategies for extreme mass ratio inspirals. *Classical and Quantum Gravity*, 28(9):094016, May 2011.

[158] Jonathan Gair and Linqing Wen. Detecting extreme mass ratio inspirals with LISA using time frequency methods: II. Search characterization. *Classical and Quantum Gravity*, 22(18):S1359–S1371, September 2005.

[159] Linqing Wen and Jonathan R. Gair. Detecting extreme mass ratio inspirals with LISA using time frequency methods. *Classical and Quantum Gravity*, 22(10):S445–S451, May 2005.

[160] Jonathan Gair and Gareth Jones. Detecting extreme mass ratio inspiral events in LISA data using the hierarchical algorithm for clusters and ridges (HACR). *Classical and Quantum Gravity*, 24(5):1145–1168, March 2007.

[161] Michael L. Katz, Alvin J. K. Chua, Lorenzo Speri, Niels Warburton, and Scott A. Hughes. Fast extreme-mass-ratio-inspiral waveforms: New tools for millihertz gravitational-wave data analysis. PRD, 104(6):064047, September 2021.

[162] Alexandre Toubiana, Sylvain Marsat, Stanislav Babak, John Baker, and Tito Dal Canton. Parameter estimation of stellar-mass black hole binaries with LISA. PRD, 102(12):124037, December 2020.

[163] Riccardo Buscicchio, Antoine Klein, Elinore Roebber, Christopher J. Moore, Davide Gerosa, Eliot Finch, and Alberto Vecchio. Bayesian parameter estimation of stellar-mass black-hole binaries with LISA. PRD, 104(4):044065, August 2021.

[164] Laura Sberna, Stanislav Babak, Sylvain Marsat, Andrea Caputo, Giulia Cusin, Alexandre Toubiana, Enrico Barausse, Chiara Caprini, Tito Dal Canton, Alberto Sesana, and Nicola Tamanini. Observing GW190521-like binary black holes and their environment with LISA. PRD, 106(6):064056, September 2022.

[165] B. P. Abbott et al. GW150914: First results from the search for binary black hole coalescence with Advanced LIGO. PRD, 93(12):122003, June 2016.

[166] Becca Ewing, Surabhi Sachdev, Ssohrab Borhanian, and B. S. Sathyaprakash. Archival searches for stellar-mass binary black holes in LISA data. PRD, 103(2):023025, January 2021.

[167] Han Wang, Ian Harry, Alexander Nitz, and Yi-Ming Hu. Space-based gravitational wave observatories will be able to use eccentricity to unveil stellar-mass binary black hole formation. PRD, 109(6):063029, March 2024.

[168] Alberto Sesana. Multi-band gravitational wave astronomy: science with joint space- and ground-based observations of black hole binaries. In *Journal of Physics Conference Series*, volume 840 of *Journal of Physics Conference Series*, page 012018, May 2017.

[169] Keith Riles. Searches for Continuous-Wave Gravitational Radiation. *arXiv e-prints*, page arXiv:2206.06447, June 2022.

[170] A.I. Burke. *Extreme Precision and Extreme Complexity: Source Modelling and Data Analysis Development for the Laser Interferometer Space Antenna.* University of Edinburgh, 2021.

[171] Jonathan R. Gair, Ilya Mandel, and Linqing Wen. Improved time frequency analysis of extreme-mass-ratio inspiral signals in mock LISA data. *Classical and Quantum Gravity*, 25(18):184031, September 2008.

[172] Xue-Ting Zhang, Chris Messenger, Natalia Korsakova, Man Leong Chan, Yi-Ming Hu, and Jian-dong Zhang. Detecting gravitational waves from extreme mass ratio inspirals using convolutional neural networks. PRD, 105(12):123027, June 2022.

[173] J. Kennedy and R. Eberhart. Particle swarm optimization. In *Proceedings of*

ICNN'95 - International Conference on Neural Networks, volume 4, pages 1942–1948 vol.4, 1995.

[174] Y. Shi and R. Eberhart. A modified particle swarm optimizer. In *1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence (Cat. No.98TH8360)*, pages 69–73, 1998.

[175] D. Parrott and Xiaodong Li. Locating and tracking multiple dynamic optima by a particle swarm model using speciation. *IEEE Transactions on Evolutionary Computation*, 10(4):440–458, 2006.

[176] J. Kennedy. Stereotyping: improving particle swarm performance with cluster analysis. In *Proceedings of the 2000 Congress on Evolutionary Computation. CEC00 (Cat. No.00TH8512)*, volume 2, pages 1507–1512 vol.2, 2000.

[177] Yann Bouffanais and Edward K. Porter. Detecting compact galactic binaries using a hybrid swarm-based algorithm. PRD, 93(6):064020, March 2016.

[178] Xue-Hao Zhang, Soumya D. Mohanty, Xiao-Bo Zou, and Yu-Xiao Liu. Resolving Galactic binaries in LISA data using particle swarm optimization and cross-validation. PRD, 104(2):024023, July 2021.

[179] Alvin J. K. Chua, Christopher J. Moore, and Jonathan R. Gair. Augmented kludge waveforms for detecting extreme-mass-ratio inspirals. PRD, 96(4):044005, August 2017.

[180] Graham Woan. Semi-coherent searches for continuous gravitational waves, and the N1/4 law. `https://dcc.ligo.org/LIGO-T2100266/public`, 2021.

[181] Steve Drasco and Scott A. Hughes. Gravitational wave snapshots of generic extreme mass ratio inspirals. PRD, 73(2):024027, January 2006.

[182] Robert H. Swendsen and Jian-Sheng Wang. Replica Monte Carlo simulation of spin glasses. PRL, 57(21):2607–2609, November 1986.

[183] David J. Earl and Michael W. Deem. Parallel tempering: Theory, applications, and new perspectives. *Physical Chemistry Chemical Physics (Incorporating Faraday Transactions)*, 7(23):3910, January 2005.

[184] S. Chatterji, L. Blackburn, G. Martin, and E. Katsavounidis. Multiresolution techniques for the detection of gravitational-wave bursts. *Classical and Quantum Gravity*, 21(20):S1809–S1818, October 2004.

[185] Tim Dietrich, Anuradha Samajdar, Sebastian Khan, Nathan K. Johnson-McDaniel, Reetika Dudi, and Wolfgang Tichy. Improving the NRTidal model for binary neutron star systems. PRD, 100(4):044003, August 2019.

[186] Tim Dietrich, Sebastian Khan, Reetika Dudi, Shasvath J. Kapadia, Prayush Kumar, Alessandro Nagar, Frank Ohme, Francesco Pannarale, Anuradha Samajdar, Sebastiano Bernuzzi, Gregorio Carullo, Walter Del Pozzo, Maria Haney, Charalampos Markakis, Michael Pürrer, Gunnar Riemenschneider, Yoshinta Eka Setyawati, Ka Wa Tsang, and Chris Van Den Broeck. Matter imprints in waveform models for neutron star binaries: Tidal and self-spin effects. PRD, 99(2):024029, January 2019.

[187] Patricia Schmidt, Mark Hannam, and Sascha Husa. Towards models of gravitational waveforms from generic binaries: A simple approximate mapping between precessing and nonprecessing inspiral signals. PRD, 86(10):104063, November 2012.

[188] Mark Hannam, Patricia Schmidt, Alejandro Bohé, Leïla Haegel, Sascha Husa, Frank Ohme, Geraint Pratten, and Michael Pürrer. Simple Model of Complete Precessing Black-Hole-Binary Gravitational Waveforms. PRL, 113(15):151101, October 2014.

[189] Katerina Chatziioannou. Neutron-star tidal deformability and equation-of-state constraints. *General Relativity and Gravitation*, 52(11):109, November 2020.

[190] Sascha Husa, Sebastian Khan, Mark Hannam, Michael Pürrer, Frank Ohme, Xisco Jiménez Forteza, and Alejandro Bohé. Frequency-domain gravitational waves

from nonprecessing black-hole binaries. I. New numerical waveforms and anatomy of the signal. PRD, 93(4):044006, February 2016.

[191] Sebastian Khan, Sascha Husa, Mark Hannam, Frank Ohme, Michael Pürrer, Xisco Jiménez Forteza, and Alejandro Bohé. Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era. PRD, 93(4):044007, February 2016.

[192] D. M. Macleod, J. S. Areeda, S. B. Coughlin, T. J. Massinger, and A. L. Urban. GWpy: A Python package for gravitational-wave astrophysics. *SoftwareX*, 13:100657, 2021.

[193] R. Abbott et al. Open data from the first and second observing runs of Advanced LIGO and Advanced Virgo. *SoftwareX*, 13:100658, January 2021.

[194] Kent. Blackburn et al. LOSC CLN Data Products for GW170817. `https://dcc.ligo.org/P1700349/public`. [Accessed 17nd-April-2023].

[195] Joshua S. Speagle. DYNESTY: a dynamic nested sampling package for estimating Bayesian posteriors and evidences. MNRAS, 493(3):3132–3158, April 2020.

[196] Gregory Ashton, Moritz Hübner, Paul D. Lasky, Colm Talbot, Kendall Ackley, Sylvia Biscoveanu, Qi Chu, Atul Divakarla, Paul J. Easter, Boris Goncharov, Francisco Hernandez Vivanco, Jan Harms, Marcus E. Lower, Grant D. Meadors, Denyz Melchor, Ethan Payne, Matthew D. Pitkin, Jade Powell, Nikhil Sarin, Rory J. E. Smith, and Eric Thrane. BILBY: A User-friendly Bayesian Inference Library for Gravitational-wave Astronomy. ApJS, 241(2):27, April 2019.

[197] Romero-Shaw et al. Bayesian inference for compact binary coalescences with BILBY: validation and application to the first LIGO-Virgo gravitational-wave transient catalogue. MNRAS, 499(3):3295–3319, December 2020.

[198] Antoine Klein. EFPE: Efficient fully precessing eccentric gravitational waveforms for binaries with long inspirals. *arXiv e-prints*, June 2021.

[199] Thibault Damour, Achamveedu Gopakumar, and Bala R. Iyer. Phasing of gravitational waves from inspiralling eccentric binaries. PRD, 70(6):064028, September 2004.

[200] Alberto Mangiagli, Antoine Klein, Alberto Sesana, Enrico Barausse, and Monica Colpi. Post-Newtonian phase accuracy requirements for stellar black hole binaries with LISA. PRD, 99(6):064056, March 2019.

[201] Geraint Pratten, Antoine Klein, Christopher J. Moore, Hannah Middleton, Nathan Steinle, Patricia Schmidt, and Alberto Vecchio. LISA science performance in observations of short-lived signals from massive black hole binary coalescences. PRD, 107(12):123026, June 2023.

[202] Riccardo Buscicchio, Elinore Roebber, Janna M. Goldstein, and Christopher J. Moore. Label switching problem in Bayesian analysis for gravitational wave astronomy. PRD, 100(8):084041, October 2019.

[203] Elinore Roebber, Riccardo Buscicchio, Alberto Vecchio, Christopher J. Moore, Antoine Klein, Valeriya Korol, Silvia Toonen, Davide Gerosa, Janna Goldstein, Sebastian M. Gaebel, and Tyrone E. Woods. Milky Way Satellites Shining Bright in Gravitational Waves. ApJ, 894(2):L15, May 2020.

[204] V. Korol, S. Toonen, A. Klein, V. Belokurov, F. Vincenzo, R. Buscicchio, D. Gerosa, C. J. Moore, E. Roebber, E. M. Rossi, and A. Vecchio. Populations of double white dwarfs in Milky Way satellites and their detectability with LISA. A&A, 638:A153, June 2020.

[205] Massimo Tinto and Sanjeev V. Dhurandhar. Time-delay interferometry. *Living Reviews in Relativity*, 24(1):1, December 2021.

[206] Louis J. Rubbo, Neil J. Cornish, and Olivier Poujade. Forward modeling of space-borne gravitational wave detectors. PRD, 69(8):082003, April 2004.

[207] Thomas A. Prince, Massimo Tinto, Shane L. Larson, and J. W. Armstrong. LISA optimal sensitivity. PRD, 66(12):122002, December 2002.

[208] C. W. Clenshaw and A. R. Curtis. A method for numerical integration on an automatic computer. *Numerische Mathematik*, 2:197–205, 1960.

[209] Neil J. Cornish. Heterodyned likelihood for rapid gravitational wave parameter inference. PRD, 104(10):104054, November 2021.

[210] Stanislav Babak, Martin Hewitson, and Antoine Petiteau. LISA Sensitivity and SNR Calculations. *arXiv e-prints*, page arXiv:2108.01167, August 2021.

[211] Walter Del Pozzo and John Veitch. CPNest: Parallel nested sampling. Astrophysics Source Code Library, May 2022.

[212] Rie B. Larsen, Jerome Jouffroy, and Benny Lassen. On the premature convergence of particle swarm optimization. In *2016 European Control Conference (ECC)*, pages 1922–1927, 2016.

[213] R.C. Eberhart and Y. Shi. Comparing inertia weights and constriction factors in particle swarm optimization. In *Proceedings of the 2000 Congress on Evolutionary Computation. CEC00 (Cat. No.00TH8512)*, volume 1, pages 84–88 vol.1, 2000.

[214] Musaed Alhussein and Syed Irtaza Haider. Improved particle swarm optimization based on velocity clamping and particle penalization. In *2015 3rd International Conference on Artificial Intelligence, Modelling and Simulation (AIMS)*, pages 61–64, 2015.

[215] Koutarou Kyutoku and Naoki Seto. Concise estimate of the expected number of detections for stellar-mass binary black holes by eLISA. MNRAS, 462(2):2177–2183, October 2016.

[216] Antoine Petiteau, Yu Shang, Stanislav Babak, and Farhan Feroz. Search for spinning black hole binaries in mock LISA data using a genetic algorithm. PRD, 81(10):104016, May 2010.

[217] Alvin J. K. Chua. One-stop function for gravitational-wave detection, identification, and inference. PRD, 106(10):104051, November 2022.

[218] Diganta Bandopadhyay and Christopher J. Moore. Gpu-accelerated semicoherent hierarchical search for stellar-mass binary inspiral signals in lisa. *Phys. Rev. D*, 110:103026, Nov 2024.

[219] Diganta Bandopadhyay and Christopher J. Moore. `SC_Search` code. `https://github.com/dig07/SC_Search`, 2024.

[220] Keith Riles. Searches for continuous-wave gravitational radiation. *Living Reviews in Relativity*, 26(1):3, December 2023.

[221] Luca D'Onofrio and for the LIGO-Virgo-Kagra Collaboration. The targeted search for gravitational waves from known pulsars. *Journal of Physics: Conference Series*, 2429(1):012038, feb 2023.

[222] G. Ashton and R. Prix. Hierarchical multistage MCMC follow-up of continuous gravitational wave candidates. PRD, 97(10):103020, May 2018.

[223] Yao Fu, Yan Wang, and Soumya D. Mohanty. Hierarchical search method for gravitational waves from stellar-mass binary black holes in noisy space-based detector data. *arXiv e-prints*, page arXiv:2407.10797, July 2024.

[224] J. W. Armstrong, F. B. Estabrook, and Massimo Tinto. Time-delay interferometry for space-based gravitational wave searches. *The Astrophysical Journal*, 527(2):814, dec 1999.

[225] Samantha A. Usman, Alexander H. Nitz, Ian W. Harry, Christopher M. Biwer, Duncan A. Brown, Miriam Cabero, Collin D. Capano, Tito Dal Canton, Thomas

Dent, Stephen Fairhurst, Marcel S. Kehl, Drew Keppel, Badri Krishnan, Amber Lenon, Andrew Lundgren, Alex B. Nielsen, Larne P. Pekowsky, Harald P. Pfeiffer, Peter R. Saulson, Matthew West, and Joshua L. Willis. The PyCBC search for gravitational waves from compact binary coalescence. *Classical and Quantum Gravity*, 33(21):215004, November 2016.

[226] Alessandra Buonanno, Yanbei Chen, and Michele Vallisneri. Detection template families for gravitational waves from the final stages of binary black-hole inspirals: Nonspinning case. PRD, 67(2):024016, January 2003.

[227] Alexander H. Nitz, Thomas Dent, Tito Dal Canton, Stephen Fairhurst, and Duncan A. Brown. Detecting Binary Compact-object Mergers with Gravitational Waves: Understanding and Improving the Sensitivity of the PyCBC Search. ApJ, 849(2):118, November 2017.

[228] Jayanti Prasad and Tarun Souradeep. Cosmological parameter estimation using particle swarm optimization. PRD, 85(12):123008, June 2012.

[229] William M. Spears, Derek T. Green, and Diana F. Spears. Biases in Particle Swarm Optimization. *International Journal of Swarm Intelligence Research (IJSIR)*, 1(2):34–57, April 2010.

[230] Daniel N. Wilke, Schalk Kok, and Albert A. Groenwold. Comparison of linear and classical velocity update rules in particle swarm optimization: notes on scale and frame invariance. *International Journal for Numerical Methods in Engineering*, 70(8):985–1008, 2007.

[231] Minas Karamanis, Florian Beutler, and John A Peacock. zeus: A python implementation of ensemble slice sampling for efficient bayesian parameter inference. *arXiv preprint arXiv:2105.03468*, 2021.

[232] Minas Karamanis and Florian Beutler. Ensemble slice sampling: Parallel, black-

box and gradient-free inference for correlated & multimodal distributions. *arXiv preprint arXiv: 2002.06212*, 2020.

[233] Blake Moore, Marc Favata, K. G. Arun, and Chandra Kant Mishra. Gravitational-wave phasing for low-eccentricity inspiralling compact binaries to 3PN order. PRD, 93(12):124061, June 2016.

[234] Giulia Fumagalli, Isobel Romero-Shaw, Davide Gerosa, Viola De Renzis, Konstantinos Kritos, and Aleksandra Olejak. Residual eccentricity as a systematic uncertainty on the formation channels of binary black holes. *arXiv e-prints*, page arXiv:2405.14945, May 2024.

[235] Mudit Garg, Shubhanshu Tiwari, Andrea Derdzinski, John G. Baker, Sylvain Marsat, and Lucio Mayer. The minimum measurable eccentricity from gravitational waves of LISA massive black hole binaries. MNRAS, 528(3):4176–4187, March 2024.

[236] Michael L. Katz, Sylvain Marsat, Alvin J. K. Chua, Stanislav Babak, and Shane L. Larson. GPU-accelerated massive black hole binary parameter estimation with LISA. PRD, 102(2):023033, July 2020.

[237] Michael L. Katz. Fully automated end-to-end pipeline for massive black hole binary signal extraction from LISA data. PRD, 105(4):044055, February 2022.

[238] Davide Gerosa, Sizheng Ma, Kaze W. K. Wong, Emanuele Berti, Richard O'Shaughnessy, Yanbei Chen, and Krzysztof Belczynski. Multiband gravitational-wave event rates and stellar physics. PRD, 99(10):103004, May 2019.

[239] R. Buscicchio, J. Torrado, C. Caprini, G. Nardini, N. Karnesis, M. Pieroni, and A. Sesana. Stellar-mass black-hole binaries in LISA: characteristics and complementarity with current-generation interferometers. J. Cosmology Astropart. Phys., 2025(1):084, January 2025.

[240] Gareth S. Cabourn Davies and Ian W. Harry. Establishing significance of gravitational-wave signals from a single observatory in the PyCBC offline search. *Classical and Quantum Gravity*, 39(21):215012, November 2022.

[241] Cody Messick, Kent Blackburn, Patrick Brady, Patrick Brockill, Kipp Cannon, Romain Cariou, Sarah Caudill, Sydney J. Chamberlin, Jolien D. E. Creighton, Ryan Everett, Chad Hanna, Drew Keppel, Ryan N. Lang, Tjonnie G. F. Li, Duncan Meacher, Alex Nielsen, Chris Pankow, Stephen Privitera, Hong Qi, Surabhi Sachdev, Laleh Sadeghian, Leo Singer, E. Gareth Thomas, Leslie Wade, Madeline Wade, Alan Weinstein, and Karsten Wiesner. Analysis framework for the prompt discovery of compact binary mergers in gravitational-wave data. PRD, 95(4):042001, February 2017.

[242] Rodrigo Tenorio, Luana M. Modafferi, David Keitel, and Alicia M. Sintes. Empirically estimating the distribution of the loudest candidate from a gravitational-wave search. PRD, 105(4):044029, February 2022.

[243] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Köpf, Edward Yang, Zach DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. PyTorch: An Imperative Style, High-Performance Deep Learning Library. *arXiv e-prints*, page arXiv:1912.01703, December 2019.

[244] Diganta Bandopadhyay and Christopher J. Moore. Searching for stellar-origin binary black holes in lisa data challenge 1b: Yorsh. *Phys. Rev. D*, 111:044039, Feb 2025.

[245] LISA Data Processing Group. LISA Data Challenges. `https://lisa-ldc.lal.in2p3.fr`, accessed Nov 2024.

[246] LISA Data Processing Group. LDC Data Challenge 2a: Sangria. `https://lisa-ldc.lal.in2p3.fr/static/data/pdf/LDC-manual-Sangria.pdf`, 2020.

[247] Guillaume Boileau, Astrid Lamberts, Nelson Christensen, Neil J. Cornish, and Renate Meyer. Spectral separation of the stochastic gravitational-wave background for LISA in the context of a modulated Galactic foreground. MNRAS, 508(1):803–826, November 2021.

[248] Sebastian Khan et al. Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era. PRD, 93(4):044007, February 2016.

[249] Sylvain Marsat and John G. Baker. LISAbeta. `https://pypi.org/project/lisabeta/1.0.2/`, 2024.

[250] Antoine Petiteau, Gérard Auger, Hubert Halloin, Olivier Jeannin, Eric Plagnol, Sophie Pireaux, Tania Regimbau, and Jean-Yves Vinet. LISACode: A scientific simulator of LISA. PRD, 77(2):023002, January 2008.

[251] Diganta Bandopadhyay and Christopher J. Moore. LDC-Yorsch-stellar-mass-search. `https://github.com/dig07/LDC-Yorsh-stellar-mass-search`, 2024.

[252] Giulia Fumagalli, Isobel Romero-Shaw, Davide Gerosa, Viola De Renzis, Konstantinos Kritos, and Aleksandra Olejak. Residual eccentricity as a systematic uncertainty on the formation channels of binary black holes. PRD, 110(6):063012, September 2024.

[253] Marc Favata, Chunglee Kim, K. G. Arun, JeongCho Kim, and Hyung Won Lee. Constraining the orbital eccentricity of inspiralling compact binary systems with Advanced LIGO. PRD, 105(2):023003, January 2022.

[254] Olaf Hartwig and Martina Muratore. Characterization of time delay interferometry combinations for the LISA instrument noise. PRD, 105(6):062006, March 2022.

[255] Jean-Baptiste Bayle, Olaf Hartwig, and Martin Staab. Adapting time-delay interferometry for LISA data in frequency. PRD, 104(2):023006, July 2021.

[256] Xue-Ting Zhang, Natalia Korsakova, Man Leong Chan, Chris Messenger, and Yi-Ming Hu. Searching for gravitational waves from stellar-mass binary black holes early inspiral. *arXiv e-prints*, page arXiv:2406.07336, June 2024.

[257] Michael Katz. BBHx. `https://github.com/mikekatz04/BBHx/blob/master/src/Response.cu`, 2024.

[258] Priscilla Canizares, Scott E. Field, Jonathan R. Gair, and Manuel Tiglio. Gravitational wave parameter estimation with compressed likelihood evaluations. PRD, 87(12):124005, June 2013.