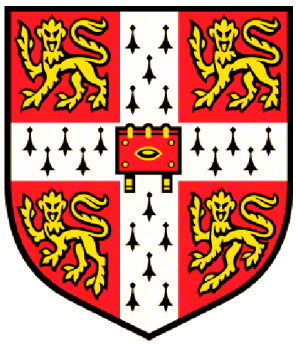
Gravitational Waves:

Understanding Black Holes

Christopher James Moore
Corpus Christi College
& Institute of Astronomy
University of Cambridge



This dissertation is submitted for the degree of
Doctor of Philosophy

Supervisor: Jonathan R. Gair

30th March 2016

Gravitational Waves: Understanding Black Holes
Christopher James Moore

Summary

This thesis concerns the use of observations of gravitational waves as tools for astronomy and fundamental physics. Gravitational waves are small ripples in spacetime produced by rapidly accelerating masses; their existence has been predicted for almost 100 years, but the first direct evidence of their existence came only very recently with the announcement in February 2016 of the detection by the LIGO and VIRGO collaborations.

Part I of this thesis presents an introduction to gravitational wave astronomy, including a detailed discussion of a wide range of gravitational wave sources, their signal morphologies, and the experimental detectors used to observe them.

Part II of this thesis concerns a particular data analysis problem which often arises when trying to infer the source properties from a gravitational wave observation. The use of an inaccurate signal model can cause significant systematic errors in the inferred source parameters. The work in this section concerns a proposed technique, called the Gaussian process marginalised likelihood, for overcoming this problem.

Part III of this thesis concerns the possibility of testing if the gravitational field around an astrophysical black hole conforms to the predictions of general relativity and the cosmic censorship hypothesis. It is expected that the gravitational field should be well described by the famous Kerr solution. Two approaches for testing this hypothesis are considered; one using X-ray observations and one using gravitational waves. The results from these two approaches are compared and contrasted.

Finally, the conclusions and a discussion of future prospects are presented in part IV of this thesis.

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University of similar institution except as declared in the Preface and specified in the text. All concepts, results and data obtained by others are properly referenced. The use of "we" in the main text reflects my stylistic preference and nothing more.

Some of the results presented here have been published previously. These are:

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

- **Moore C. J.**, Taylor S. R. and Gair J. R. (2015)
  Estimating the sensitivity of pulsar timing arrays
  Classical and Quantum Gravity **32**, 5, 055004

- **Moore C. J.** (2015)
  The sensitivity of pulsar timing arrays
  Journal of Physics Conference Series **610**, 1, 012018

which largely constitute the content of part I of this thesis;

- **Moore C. J.** and Gair J. R. (2014)
  Novel Method for Incorporating Model Uncertainties into Gravitational Wave

Parameter Estimates
Physical Review Letters **113**, 25, 251101

- **Moore C. J.**, Berry C. P. L., Chua A. J. K. and Gair J. R. (2015)
  Improving gravitational-wave parameter estimation using Gaussian process regression
  Physical Review D **93**, 6, 064001

- Gair J. R. and **Moore C. J.** (2015)
  Quantifying and mitigating bias in inference on gravitational wave source populations
  Physical Review D **91**, 12, 124062

- **Moore C. J.**, Berry C. P. L., Chua A. J. K. and Gair J. R. (2016)
  Fast methods for training Gaussian processes on large data sets
  Royal Society Open Science **3**, 5

which largely constitute the content of part II of this thesis;

- **Moore C. J.** and Gair J. R. (2015)
  Testing the no-hair property of black holes with x-ray observations of accretion disks
  Physical Review D **92**, 2, 024039

- **Moore C. J.** and Gair J. R. (in preparation)
  Testing the no-hair property of black holes with gravitational wave observations of extreme-mass-ratio inspirals

which largely constitute the content of part III. Explicit citation to any of these papers indicates work done in collaboration. This is done in Sec. 6.4.2, where the original derivation of Eq. 6.62 is due to Jonathan Gair; I subsequently verified this result and extended it to Eq. 6.74, and the numerical results in Sec. 6.4.3.

This dissertation is less than 60000 words in length, including footnotes, captions and the bibliography. The word count evaluated using `TeXcount`, it contains 56749 words.

Acknowledgements

Completing a Ph.D. is a long and difficult process, I am grateful to many people for their help and support during that process.

Firstly I would like to thank my supervisor, Jonathan Gair. He has helped me throughout my Ph.D., and parts of my undergraduate degree before that. It is scarcely an exaggeration to say that he has taught me everything that I know about gravitational waves.

I must also thank a number of students at the IoA, with whom I have shared office space, enlightening and amusing conversations, along with trips to the pub. In particular I would like to thank Christopher Berry, Stephen Taylor, Robert Cole, Michael Coughlin, Tom Callister, Alvin Chua, and Deyan Mihaylov.

I am also grateful for the help of the administrative staff, both in the Institute of Astronomy and Corpus Christi College. Particular thanks should go to Debbie Peterson, Margaret Harding, Jan Leaver, and Vasily Belokurov.

I've benefited greatly from the academic environment in the department, and Cambridge more widely. I would like to thank Donald Lynden-Bell for many interesting discussions on various aspects of general relativity and differential geometry, Ulrich Sperhake for useful discussions regarding numerical relativity, and Priscilla Canizares for her help understanding the gravitational self force.

During the later part of my Ph.D. I joined the LIGO Scientific Collaboration, and I have benefited from many discussions with the other members. I would like to thank several members of the group in Birmingham, including Christopher Berry, Will Farr, Ilya Mandel, and Alberto Vecchio. I would also like to thank the Compact Binary Coalescence group as a whole for useful comments on a number of draft papers.

Finally I would like to thank my Mum. Proofreading parts of some of my papers was only the last and least of the things that she has done for me over the years.

Contents

Part I

Introduction

Chapter 1

Gravitation

> Spacetime tells matter how to move;
> matter tells spacetime how to curve.
>
> John A. Wheeler

Of the four fundamental forces now known to modern physics, only two were known prior to the quantum revolution of the 20th century. These two *classical* forces are electromagnetism and gravitation. Despite the fact that we now appreciate it is predominantly the electromagnetic force that governs the behaviour of the familiar objects that surround us, it was the gravitational force which was first formulated in a recognisably modern way by Sir Isaac Newton in the 17th century.

Our current best theory of gravity is the theory of *general relativity*, developed by Einstein in the early part of the 20th century. In this chapter a brief historical introduction to general relativity will be presented, focussing on the prediction of the existence of gravitational waves by Einstein (1916, 1918).

1.1 Newton's gravity

In the early part of the 17th century Johannes Kepler published a number of observations regarding the motions of the planets around the sun. Firstly, regarding the motion of an individual planet (he originally considered Mars) he observed that the motion was elliptical with the sun as one focus, and (expressed in modern terminology) that the planet had a constant angular momentum (Kepler, 1609). He later published a series of further observations regarding the relationships between the orbits of different planets;

in particular he observed that the square of the period of a planet's orbit is proportional to the cube of the semi-major axis (Kepler, 1619).

Newton realised that an attractive radial force on a planet which scaled inversely with the square of distance to the sun would explain Kepler's earlier observations. Kepler's later observations gave a hint of the universal nature of gravity, but Newton went much further in proposing his universal law of gravitation (Newton, 1687):

> Any two bodies in the Universe attract each other with a force that is directly proportional to the product of their masses and inversely proportional to the square of the distance between them.

For two objects of masses m_1 and m_2 at positions $\vec{x}_1$ and $\vec{x}_2$ the force experienced by m_2 due to m_1 is given by

$$\vec{F} = -\frac{Gm_1m_2}{|\vec{x}_2 - \vec{x}_1|^3}(\vec{x}_2 - \vec{x}_1) \ , \tag{1.1}$$

where $G = 6.67408(31) \times 10^{-1}\,\mathrm{m^3 kg^{-11} s^{-2}}$ is experimentally determined, and is called Newton's gravitational constant[1].

This law can be recast in modern notation as a "scalar field theory". The gravitational force on object m_2 is proportional to the gradient of the scalar *gravitational potential*, $\phi(\vec{x})$, generated by all other masses; $F_2 = m_2 \nabla|_{\vec{x}=\vec{x}_2}\phi$. The gravitational potential satisfies Poisson's equation sourced by the mass density, ρ, with the boundary condition that ϕ decays as $1/|\vec{x}|$, or faster, as $|\vec{x}| \to \infty$;

$$\nabla^2\phi(\vec{x}) = 4\pi G\rho(\vec{x}) \ . \tag{1.2}$$

Equation 1.1 is proportional to the gradient of the Green's function for this field theory, and therefore satisfies the required boundary conditions. The field Eq. 1.2 should be interpreted as being valid throughout all of space separately at each moment in time. The time evolution of the system is governed by a separate set of laws, Newton's laws of motion, involving all of the forces in the system as well as the gravitational force defined above.

The analogous field equations for the other fundamental force known to classical physics, electromagnetism, were formulated nearly two centuries later by Maxwell (1865). These comprise a set of coupled partial differential equations for the electric, $\vec{E}$, and

[1] `http://physics.nist.gov/cuu/Constants/index.html`

magnetic, $\vec{B}$, vector fields sourced by charge, ρ_e, and current, $\vec{J}$, density distributions.

$$\begin{aligned} &\text{(a)} \quad \nabla \cdot \vec{E}(t,\vec{x}) = \tfrac{\rho_e(t,\vec{x})}{\epsilon_0} \qquad \text{(c)} \quad \nabla \cdot \vec{B}(t,\vec{x}) = 0 \\ &\text{(b)} \quad \nabla \times \vec{E}(t,\vec{x}) = -\tfrac{\partial \vec{B}(t,\vec{x})}{\partial t} \quad \text{(d)} \quad \nabla \times \vec{B}(t,\vec{x}) = \mu_0 \left(\vec{J}(t,\vec{x}) + \epsilon_0 \frac{\partial \vec{E}(t,\vec{x})}{\partial t} \right) \end{aligned} \tag{1.3}$$

The (defined) constants μ_0 and ϵ_0 are the permeability and permittivity of free space respectively. In a statics problem the curl of $\vec{E}$ vanishes and the electric field can be written as the gradient of a scalar *electric potential*, $\vec{E} = \nabla\phi_e$. This is analogous to the gravitational potential, indeed Eq. 1.3(a) written in terms of ϕ_e bears a striking resemblance to Eq. 1.2 for ϕ.

However, in contrast to gravity, the coupling of $\vec{E}$ to the time derivative of $\vec{B}$ (and visa versa) allows time-dependent solutions to Eq. 1.3, even in the absence of time-varying current and charge sources. In fact, two linearly independent plane wave solutions exist in vacuum. These *electromagnetic wave* solutions propagate with phase and group velocities equal to $c = (\mu_0\epsilon_0)^{-1/2}$, and the electric and magnetic fields are orthogonal and both transverse to the direction of propagation.

Physicists were relatively slow to fully appreciate the symmetry hidden in Maxwell's equations. For example, the question of against which frame of reference the speed of these waves should be measured led to various searches for the postulated *ether*; this was supposed to have defined a preferred inertial frame against which motion could be absolutely determined. The most famous of these experimental searches was performed by Michelson and Morley (1887), who used an interferometer to measure speed of light in different directions in an effort to detect the motion of the Earth through the ether. The experiment found no variation of the speed of light. Many physicists, including Fitz Gerald (1889), Lorentz (1892), Larmor (1897) and Poincaré (1900), proposed partial explanations for this null result based around ideas of length contraction and time dilation. These ideas were broadly motivated by considerations of the transformation properties of Maxwell's equations between different inertial frames.

But it was Einstein (1905) who first postulated that all observers should measure the same speed for light and that all the laws of physics should be the same in all inertial frames. These ideas are expressed most clearly in the geometric picture of *spacetime* proposed by Minkowski (1909); each event in 4-dimensional spacetime may be ascribed coordinates x^μ, which may conveniently be taken to be $x = (t, \vec{x})$. The constancy of the speed of light leads to the notion of an invariant interval between any two events;

expressed in infinitesimal form in terms of the coordinates x^μ this interval is given by

$$\mathrm{d}s^2 = \eta_{\mu\nu}\mathrm{d}x^\nu\mathrm{d}x^\nu \equiv -\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 \ , \tag{1.4}$$

where the Minkowski metric, $\eta_{\mu\nu}$ has been defined[2]. Einstein's postulate that the laws of physics should be the same expressed in all inertial frames is equivalent to demanding that they be invariant under *Lorentz transformations* of the coordinates, $x^\mu \to x'^\mu = \Lambda(x)$, which leave invariant the Minkowski line element in Eq. 1.4. Einstein called this the theory of *special relativity*.

The field equation for Newtonian gravity in Eq. 1.2 is manifestly not Lorentz invariant. Several attempts were made in the early 20th century to modify the field equation for Newtonian gravity to make it compatible with special relativity. For example Nordström (1912, 1913) proposed replacing the Laplacian on the left-hand side with the Lorentz invariant d'Alembertian and the density on the right-hand side with the Lorentz invariant rest mass density, ρ_0,

$$\partial_\mu\partial^\mu\phi(t,\vec{x}) \equiv \partial_t^2\phi(t,\vec{x}) - \nabla^2\phi(t,\vec{x}) \stackrel{?}{=} -4\pi G\rho_0 \ . \tag{1.5}$$

Such simple theories run into a variety of problems; for example, the field equation in Eq. 1.5 coupled with the simple Lorentz invariant 4-force law $F_\mu = \partial_\mu\phi$ implies that the potential ϕ is constant along the worldline of every particle, which is clearly overly restrictive. In addition Nordström's theory was incompatible with the strong equivalence principle. These difficulties caused Einstein to abandon attempts to construct a scalar theory of gravity; it is now known that is possible to rectify these problems with scalar theories of gravity, but Einstein's intuition on this point ultimately proved to be correct.

The quest to construct a theory which was invariant under general coordinate transformations (diffeomorphism invariance, as well as the more specific Lorentz invariance) and compatible with the equivalence principle led Einstein to abandon the idea of gravity as a force completely, and instead formulate gravity as an intrinsic property of spacetime. Einstein called this theory of gravity the *general theory of relativity*. Although Newton's original concept of a gravitational force is abandoned in the new theory, it is interesting that a field theory interpretation of gravity remains valid; albeit with a much more complicated multicomponent field!

[2]Minkowski originally used an imaginary time coordinate so that the metric was that of 4-dimensional Euclidien space, we adopt the equivalent modern convention of a mixed metric signature here.

1.2 Einstein's gravity

In general relativity the model for spacetime is a manifold, $\mathcal{M}$, with a metric, $\mathbf{g}$ (signature $+2$) defined on it. The manifold is 4-dimensional and may be covered by a collection of coordinate charts x^μ with associated coordinate bases $e_\mu = \partial/\partial x^\mu \equiv \partial_\mu$. Matter, including the electromagnetic field, is represented as various fields (e.g. tensor fields, including scalar and vector fields) defined on $\mathcal{M}$, which we denote collectively by $\psi(x)$.

A manifold with a metric singles out a preferred (torsion-free) covariant derivative, or connection[3], with respect to which the metric is covariantly constant: $\nabla \mathbf{g} = 0$. This is called the Levi-Civita connection, the components of the Levi-Civita connection coefficients in a coordinate basis are given by

$$(\nabla_X Y)^\mu = X^\nu e_\nu(Y^\mu) + \Gamma^\mu_{\nu\rho} Y^\nu X^\rho, \tag{1.6}$$

$$\text{where} \quad \Gamma^\mu_{\nu\rho} = \frac{1}{2} g^{\mu\sigma} \left(\partial_\rho g_{\sigma\nu} + \partial_\nu g_{\sigma\rho} - \partial_\sigma g_{\nu\rho}\right) \ . \tag{1.7}$$

The dynamics of the matter fields is governed by an action $S_{\text{matter}} = \int_{\mathcal{M}} \mathrm{d}^4\mathcal{V}\, \mathcal{L}_{\text{matter}}$, where $\mathrm{d}^4\mathcal{V}$ is the volume element on $\mathcal{M}$ and $\mathcal{L}_{\text{matter}}(\psi, \nabla\psi)$ is the matter Lagrangian. Demanding that this be invariant under general diffeomorphisms gives rise to a conserved current called the stress energy tensor, $\nabla^\mu T_{\mu\nu} = 0$. The stress energy tensor is defined as the following functional derivative,

$$T_{\mu\nu} = \frac{-2}{\sqrt{-\det(g)}} \frac{\delta\left(\mathcal{L}_{\text{matter}} \sqrt{-\det(g)}\right)}{\delta g^{\mu\nu}} \ . \tag{1.8}$$

Given a metric and the Levi-Civita connection, the *Riemann curvature* tensor may be defined on $\mathcal{M}$; for vector fields X, Y, and Z the Riemann tensor is defined as

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z \ , \tag{1.9}$$

where the commutator of two vector fields may be defined in terms of its components in a coordinate chart $[X,Y]^\mu = X^\nu \partial_\nu Y^\mu - Y^\nu \partial_\nu X^\mu$. In a particular coordinate chart the

[3] A connection, or covariant derivative, ∇, on a manifold $\mathcal{M}$ is a map from a pair of vector fields X, Y to a vector field $\nabla_X Y$ which is linear in X, distributive in Y, and satisfies Leibniz's rule for a scalar field f; $\nabla_X(fY) = f\nabla_X Y + Y\nabla_X f$, where $\nabla_X f \equiv X(f)$. The covariant derivative can be extended to a covector, u, by the Leibniz rule, $(\nabla_X u)(Y) \equiv \nabla_X(u(Y)) - u(\nabla_X Y)$. The covariant derivative can be further extended to a tensor of arbitrary rank by continuing to demand that the Leibniz property holds in a similar way.

components of the Riemann tensor are given by

$$R^{\mu}{}_{\nu\rho\sigma} = \partial_\rho \Gamma^{\mu}_{\nu\sigma} - \partial_\sigma \Gamma^{\mu}_{\nu\rho} + \Gamma^{\tau}_{\nu\sigma}\Gamma^{\mu}_{\tau\rho} - \Gamma^{\tau}_{\nu\rho}\Gamma^{\mu}_{\tau\sigma} \ . \tag{1.10}$$

The components of the Riemann tensor satisfy the *Bianchi identity*,

$$\nabla_\tau R^{\mu}{}_{\nu\rho\sigma} + \nabla_\rho R^{\mu}{}_{\nu\sigma\tau} + \nabla_\sigma R^{\mu}{}_{\nu\tau\rho} = 0 \ . \tag{1.11}$$

The Ricci curvature tensor and scalar may be defined in a coordinate basis as the contractions $R_{\mu\nu} = R^{\rho}{}_{\mu\rho\nu}$, and $R = g^{\mu\nu}R_{\mu\nu}$, of the Riemann tensor respectively. Finally the Einstein tensor may also be defined in a coordinate basis as

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} \ . \tag{1.12}$$

The Einstein tensor is covariantly conserved, $\nabla_\mu G^{\mu\nu} = 0$; this is a purely geometric relation which follows from contracting the Bianchi identity satisfied by the Riemann tensor.

The field equations of general relativity were first published by Einstein (1915*a*)[4]. In general relativity the gravitational field is represented as the curvature of spacetime. The Einstein field equations relate the local curvature of spacetime, characterised by the Einstein tensor, to the stress energy tensor of the matter fields in the spacetime.

$$G_{\mu\nu}(x) = \frac{8\pi G}{c^4}T_{\mu\nu}(\psi(x)) \tag{1.13}$$

The conservation of energy and momentum (including that of the gravitational field), $\nabla_\mu T^{\mu\nu} = 0$, can also be derived from these equations as a consequence of the contracted Bianchi identities, $\nabla^\mu G_{\mu\nu} = 0$. These equations govern the motion of the matter fields on spacetime. In particular, the fact that a test particle with worldline $x^\mu(\tau)$, with stress tensor $T^{\mu\nu}(z) = \partial_\tau z^\mu \partial_\tau z^\nu \delta^4(z, x(\tau))$, follows a geodesic of the metric $g_{\mu\nu}$ can be derived from the above equations. This state of affairs is perhaps best summarised by the famous quote at the beginning of this chapter, "Spacetime tells matter how to move; matter tells spacetime how to curve" by Wheeler (2000).

It is easy to verify that the flat Minkowski metric $\eta_{\mu\nu}$ satisfies the vacuum field equations. In c.1916 Einstein was faced with the formidable task of trying to solve the field equations Eq. 1.13 in more general situations. It is no surprise that he quickly

[4] As our focus here is not cosmological we will neglect any cosmological constant term, $\Lambda g_{\mu\nu}$ on the left-hand side of these equations.

considered the linearised theory about the background Minkowski solution. In this analysis the metric is written in terms of a perturbation h_μ, and terms $\mathcal{O}(h^2)$ are discarded;

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = (\eta_{\mu\nu} + h_{\mu\nu})\,\mathrm{d}x^\mu \mathrm{d}x^\nu + \mathcal{O}(h^2)\ . \tag{1.14}$$

To the relevant order in the metric perturbation indices may be raised and lowered using the background Minkowski metric (so $h_{\mu\nu}$ may be thought of a tensor field defined on the background spacetime), so the trace of the perturbation is given by $h = \eta^{\mu\nu}h_{\mu\nu}$. It will be convenient to define the trace-reversed metric perturbation

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}h\eta_{\mu\nu}\ , \tag{1.15}$$

in terms of which the Einstein equations in Eq. 1.13 become

$$-\frac{1}{2}\partial^\rho\partial_\rho\bar{h}_{\mu\nu} + \frac{1}{2}\partial^\rho\partial_\mu\bar{h}_{\nu\rho} + \frac{1}{2}\partial^\rho\partial_\nu\bar{h}_{\mu\rho} - \frac{1}{2}\partial^\rho\partial^\sigma\bar{h}_{\rho\sigma} = \frac{8\pi G}{c^4}T_{\mu\nu} \tag{1.16}$$

The ability to change coordinates, diffeomorphism invariance, means that there is a considerable gauge freedom in the linearised Einstein equations. In particular we may make the following infinitesimal coordinate transformation, $x^\mu \to x'^\mu = x^\mu + \xi^\mu$, where ξ^μ is an arbitrary (small) smooth vector field. Under this transformation the metric perturbation transforms as[5]

$$h_{\mu\nu} \to h'_{\mu\nu} = h_{\mu\nu} - \partial_\mu\xi_\nu - \partial_\nu\xi_\mu\ , \tag{1.17}$$

and the trace-reversed metric perturbation transforms as

$$\bar{h}_{\mu\nu} \to \bar{h}'_{\mu\nu} = \bar{h}_{\mu\nu} - \partial_\mu\xi_\nu - \partial_\nu\xi_\mu + \eta_{\mu\nu}\partial_\rho\xi^\rho\ . \tag{1.18}$$

This gauge freedom may be used to impose the harmonic (or Lorenz[6]) gauge condition $\partial^\mu\bar{h}_{\mu\nu} = 0$; from Eq. 1.18 it can be seen that $\partial^\mu\bar{h}'_{\mu\nu} = \partial^\mu\bar{h}_{\mu\nu} - \partial^\mu\partial_\mu\xi_\nu$ which can be set to zero and solved separately for each component ξ_ν using a Green's function. In this gauge the linearised Einstein equations become

$$-\frac{1}{2}\partial^\rho\partial_\rho\bar{h}_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu} \tag{1.19}$$

[5]The transformation law for the metric perturbation may be obtained by considering $h_{\mu\nu}$ to be a tensor field defined on the background spacetime; Eq. 1.17 is then just the usual law for the transformation of the components of a tensor field. Alternatively, Eq. 1.17 may be obtained from the invariance of the theory under a one parameter family of diffeomorphisms defined using the vector field ξ^μ. For a discussion of these two viewpoints see Carroll (2004).

[6]Named for the resemblance to the eponymous gauge choice in electromagnetism.

It should be noted that the Lorenz gauge condition does not completely specify the gauge; there still exists the freedom to make additional transformations of the form $x^\mu \to x'^\mu = x^\mu + \xi^\mu$, where $\partial^\mu \partial_\mu \xi^\nu = 0$, as these preserve the Lorenz gauge condition.

1.3 Gravitational waves

The time independence of Newton's gravity in Eq. 1.2 precludes the existence of time-dependent fields, unless of course there is a time-dependent mass distribution somewhere in space sourcing them. In contrast, Einstein's equations permit time-dependent solutions even in a complete, global vacuum; $T_{\mu\nu} = 0$ everwhere. In fact there is a much stronger analogy between the Einstein equations and Maxwell's equations than there is with Newton's theory of gravity. Like Maxwell's equations, the linearised Einstein equations admit two linearly independent, transverse plane wave solutions which propagate at the speed of light; these gravitational waves were first studied by Einstein (1916, 1918)[7]

Using the linearised Einstein equations in the Lorenz gauge, Eq. 1.19, the form of the metric perturbation can be further restricted by using the residual gauge freedom discussed above. Firstly, we choose ξ^0 such that $\bar{h}' = 0$; from Eq. 1.18 it can be seen that $\bar{h}' = \bar{h} + 2(\partial_0 \xi^0 - \partial_i \xi^i)$ which can be set to zero and integrated to find ξ^0. In this *traceless gauge* we have $\bar{h}_{\mu\nu} = h_{\mu\nu}$, from now on we assume this gauge choice has been made and drop the overbar from our notation. Secondly, we choose ξ^i such that $h_{0i} = 0$; from Eq. 1.18 it can be seen that $h'_{0i} = h_{0i} - \partial_0 \xi_i - \partial_i \xi_0$ which can be set to zero and integrated to find ξ_i while still preserving $\partial_\mu \partial^\mu \xi^\nu = 0$. The Lorenz gauge condition implies that $\partial^0 h_{00} = 0$, as we are interested in time-dependent solutions we choose to set $h_{00} = 0$. In summary we now have in the *transverse-traceless gauge*,

$$h_{0\mu} = 0\,, \quad h^i{}_i = 0\,, \quad \text{and} \quad \partial^j h_{ij} = 0\,. \tag{1.20}$$

From now on we will always assume that the metric perturbation is written in the transverse-traceless gauge.

We now look for wavelike solutions of these equations $h_{ij} = H_{ij} \exp(\mathrm{i} k_\mu x^\mu)$, for fixed wavevector k^μ, constant spatial tensor H_{ij}, and it is understood that only the real part is to be considered. The Einstein equations imply that k^μ is null (i.e. gravitational waves

[7] Einstein later temporarily concluded that gravitational waves did not exist, and that their appearance was purely a gauge effect. This work was initially rejected for publication and only later published elsewhere with this claim corrected (Einstein and Rosen, 1937).

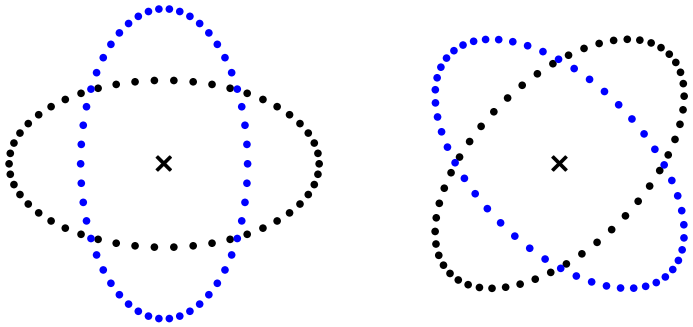


Figure 1.1: A circular ring of test particles in the xy plane is distorted into an ellipse by the passage of a linearly polarised, plane-fronted gravitational wave travelling in the z direction. There are two such independent polarisations, indicated here in the two diagrams above. The black and blue ellipses indicate the ring of particles at two extreme points in the wave cycle separated by a phase difference of π.

travel at the speed of light), we take $[k^\mu] = (\omega, 0, 0, k)$ with $\omega/k = c$. The most general form of H_{ij} that satisfies the transverse-traceless gauge conditions is a linear combination of the two basis tensors, or polarisation states:

$$\left[H^+_{ij}\right] = \begin{pmatrix} +1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} , \quad \left[H^\times_{ij}\right] = \begin{pmatrix} 0 & +1 & 0 \\ +1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} . \tag{1.21}$$

A general linearly-polarised gravitational wave is any real combination of these two polarisation states; a complex combination forms an elliptically polarised plane gravitational wave. The effect of two linearly polarised gravitational waves on a ring of test particles is sketched in Fig. 1.1.

The generation of gravitational waves provides a further analogy with electromagnetism; whilst electromagnetic waves are generated by accelerating charges, gravitational waves are generated by accelerating masses. To see this more clearly, the linearised versions of Einstein's equations, Eq. 1.16, in the presence of matter can be solved using Green's function to write the metric perturbation as an integral over the source;

$$h_{\mu\nu} = \frac{4G}{c^4} \int \mathrm{d}^3\vec{x}' \, \frac{T_{\mu\nu}(t - |\vec{x} - \vec{x}'|/c, \vec{x}')}{|\vec{x} - \vec{x}'|} . \tag{1.22}$$

In the case of a compact supported source of size $\sim R$, moving at speed $v \sim \omega R \ll c$ evaluating Eq. 1.22 in the far-zone, $|\vec{x}| \gg c/\omega$, gives the *quadrupole formula*,

$$\bar{h}_{ij}(t, \vec{x}) = \frac{2G}{c^4 |\vec{x}|} \ddot{I}_{ij}(t - r/c) \ , \tag{1.23}$$

where the trace-free quadrupole moment tensor has been defined as

$$I^{ij} = \int \mathrm{d}^3 x \ \frac{T^{00}}{c^2} (x^i x^j - \frac{1}{3} \delta^{ij} x^k x_k) \ . \tag{1.24}$$

The quadrupole formula was first obtained by Einstein (1916).

There has been an ongoing effort to detect gravitational waves almost ever since Einstein first predicted their existence. Part of the motivation for this search is the desire to confirm or refute one of the key predictions of general relativity. To some extent this desire was satisfied by the discovery of a binary pulsar by Hulse and Taylor (1974) which was losing energy in precisely the manner that would be expected if it was radiating gravitational waves according to the quadrupole formula in Eq. 1.23. This constituted the first indirect detection of gravitational waves and was awarded the Nobel Prize in Physics (1993). However, efforts to directly detect gravitational waves continued. The main motivation for this effort was the opportunity to use gravitational wave observations as a new tool for astronomy. As will be discussed in the remainder of this thesis, gravitational wave observations offer an exciting new way to study compact objects and black holes as well as testing Einstein's theory of general relativity in strong and dynamical gravitational fields.

Chapter 2

Gravitational Wave Detectors

After predicting the existence of gravitational waves, Einstein (1916) concluded that they would be too weak to be directly detected and that the energy radiated via gravitational waves would have a "practically vanishing value". It is not hard to understand how he came to this point of view; applying Eq. 1.23 to the sun Jupiter system gives a maximum strain at the inner edge of the wave zone of $\mathcal{O}(10^{-27})$[1]. Measuring such strains seemed impossible.

However, gravitational waves have now been detected, so why was Einstein wrong on this point? Partly, the answer lies in the impressive technological progress of the past century. But more importantly, a succession of astronomical observations have revealed systems that can radiate gravitational waves far more efficiently than anything Einstein would have been able to imagine in 1916. In the 30s the work of Baade and Zwicky (1934*a*,*b*) lead to the prediction that a highly compact *neutron star* could form in the supernova explosion at the end point of stellar evolution. More gradually it was realised that *black holes* could also form in a similar process. In the 70s the discovery of the Hulse-Taylor binary pulsar (Hulse, 1994) proved that neutron stars could form binaries which were losing energy via gradational waves at a rate sufficient to cause merger within a Hubble time. When it gets close to merger the Hulse-Taylor pulsar will radiate gravitational waves with an amplitude (measured at the inner edge of the wave zone) 22 orders of magnitude greater than the sun-Jupiter system!

[1]Einstein formulated his argument about the undetectability of gravitational radiation in terms of the energy radiated. Here the point is illustrated using the strain amplitude as this is the quantity that majority of modern gravitational wave detectors discussed below measure.

Efforts to directly detect gravitational waves began in the 60s with resonant bar detectors; this work was pioneered by Weber (1960). Some detection claims were even published which are now nearly universally believed to be false. However, these claims served to spur on the development of the more sensitive interferometric detectors during the 70s. This development has lead the groundbreaking sensitivity of the current generation of ground based detectors (including Advanced LIGO (The LIGO Scientific Collaboration and The Virgo Collaboration, 2016*b*)) and the first detection of gravitational waves (The LIGO Scientific Collaboration and the Virgo Collaboration, 2016*c*). This amazing detection also constitutes the first proof that binary black holes can form and merge within a Hubble time.

In this chapter we will summarise some signal analysis preliminaries. These results will be useful to quantify the sensitivity of different detectors. This chapter also includes some significant new results from the following papers: Moore (2015); Moore, Cole and Berry (2015); Moore, Taylor and Gair (2015). We will then briefly discuss a range of possible methods of detecting gravitational waves; from the early efforts of the resonant bar detectors, through modern interferometers, to pulsar timing arrays and future space-based detectors. The work in the current chapter, as well as Chapters 3 and 4, is partly based on the following published papers, Moore, Cole and Berry (2015), Moore, Taylor and Gair (2015), and Moore (2015).

2.1 Signal analysis preliminaries

As discussed in Sec. 1.3, a plane gravitational wave has two independent polarisation states denoted + and ×; a general signal can be described as a linear combination of the two, $h_{\mu\nu} = A_+ h^+_{\mu\nu} + A_\times h^\times_{\mu\nu}$. The sensitivity of a detector to these depends upon the relative orientations of the source and detector. The output of any gravitational wave detector(s) $s(t)$ contains a superposition of noise $n(t)$ and (possibly) a signal $h(t)$,

$$s(t) = n(t) + h(t) \, . \tag{2.1}$$

The scalar signal $h(t)$ is formed from a geometric projection of the metric perturbation onto the detector, and a convolution with a frequency dependent term which accounts for the different response (both amplitude and phase) of the instrument at different frequencies

$$h(t) = \left((P^{\mu\nu} h_{\mu\nu}) * R \right)(t). \tag{2.2}$$

Different types of detector will have very different noise properties, they will also have different geometric, P, and frequency, R, responses; some explicit examples of these will be derived and discussed below for interferometers, pulsar timing arrays (PTAs) and resonant bar detectors.

We shall frequently have recourse to work with the Fourier transform of the signal; we use the conventions that

$$\tilde{x}(f) \;=\; \mathcal{F}\,\{x(t)\}\,(f) \;=\int_{-\infty}^{\infty} \mathrm{d}t\; x(t)\exp(-2\pi \mathrm{i} f t)\;, \tag{2.3}$$

$$x(t) \;=\mathcal{F}^{-1}\,\{\tilde{x}(f)\}\,(t) \;=\int_{-\infty}^{\infty} \mathrm{d}f\; \tilde{x}(f)\exp(2\pi \mathrm{i} f t)\;. \tag{2.4}$$

For simplicity, it is assumed that the noise in the gravitational wave detector is stationary and Gaussian (with zero mean); under these assumptions the noise is fully characterised via the one-sided noise PSD, $S_n(f)$,

$$\langle \tilde{n}(f)\tilde{n}^*(f')\rangle = \frac{1}{2}\delta(f-f')S_n(f)\;, \tag{2.5}$$

where angle brackets $\langle \ldots \rangle$ denote an ensemble average over many noise realisations (Cutler and Flanagan, 1994*a*). In reality, we have only a single realisation to work with, but the ensemble average can be replaced by a time average for stationary stochastic noise. The procedure is to measure the noise over a sufficiently long duration T and then compute the Fourier transform $\tilde{n}(f)$ with a frequency resolution $\Delta f = T^{-1}$; this is repeated many times to give an average. The noise PSD $S_n(f)$ has units of inverse frequency.

Since the gravitational wave signal and detector output are both real, it follows that $\tilde{h}(-f) = \tilde{h}^*(f)$ and $\tilde{n}(-f) = \tilde{n}^*(f)$; therefore, $S_n(f) = S_n(-f)$. The fact that $S_n(f)$ is an even function means that Fourier integrals over all frequencies can instead be written as integrals over positive frequencies only, e.g., Eq. 2.10 and Eq. 2.15; it is for this reason that $S_n(f)$ is called the *one-sided* PSD.[2]

When integrated over all positive frequencies, the PSD gives the mean square noise amplitude. Starting by taking the time average of the square of the detector noise:

$$\begin{aligned}\overline{|n(t)|^2} \;&= \lim_{T\to\infty}\tfrac{1}{2T}\quad \int_{-T}^{T}\mathrm{d}t\; n(t)n^*(t) \\ &= \lim_{T\to\infty}\tfrac{1}{2T}\quad \int_{-T}^{T}\mathrm{d}t\;\int_{-\infty}^{\infty}\mathrm{d}f\;\int_{-\infty}^{\infty}\mathrm{d}f'\;\tilde{n}(f)\tilde{n}^*(f')\exp\left(2\pi\mathrm{i}ft\right)\exp\left(-2\pi\mathrm{i}f't\right)\end{aligned} \tag{2.6}$$

[2] An alternative convention is to use the two-sided PSD $S_n^{(2)}(f) = S_n(f)/2$.

$$= \lim_{T\to\infty} \tfrac{1}{2T} \int_{-T}^{T} \mathrm{d}t \int_{-\infty}^{\infty} \mathrm{d}f \int_{-\infty}^{\infty} \mathrm{d}f' \, \mathcal{F}\{n(\tau)\}(f) \left[\mathcal{F}\{n(\tau)\}(f')\right]^* \times \exp(2\pi \mathrm{i} f t) \exp(-2\pi \mathrm{i} f' t) \ , \quad (2.7)$$

where we have substituted in using the definitions of the Fourier transform and its inverse. A property of Fourier transforms is that a time-domain translation by amount t is equivalent to a frequency-domain phase change $2\pi f t$; if $\mathcal{F}\{n(\tau)\}(f) = \tilde{n}(f)$, then $\mathcal{F}\{n(\tau - t)\}(f) = \tilde{n}(f)\exp(2\pi \mathrm{i} f t)$. Therefore, the exponential factors in Eq. 2.7 may be absorbed as

$$\overline{|n(t)|^2} = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \mathrm{d}t \iint_{-\infty}^{\infty} \mathrm{d}f \mathrm{d}f' \, \mathcal{F}\{n(\tau - t)\}(f) \left[\mathcal{F}\{n(\tau - t)\}(f')\right]^* . \quad (2.8)$$

Since the noise is a randomly varying signal, we can use the ergodic principle to equate a time average, denoted by $\overline{(\ldots)}$, with an ensemble average, denoted by $\langle \ldots \rangle$. The noise is stationary, consequently, its expectation value is unchanged by the time-translation performed above. Therefore, using Eq. 2.5, the mean square noise amplitude is given by

$$\overline{|n(t)|^2} = \int_{-\infty}^{\infty} \mathrm{d}f \int_{-\infty}^{\infty} \mathrm{d}f' \, \langle \tilde{n}(f)\tilde{n}^*(f') \rangle \quad (2.9)$$

$$= \int_{-\infty}^{\infty} \mathrm{d}f \int_{-\infty}^{\infty} \mathrm{d}f' \, \frac{1}{2} S_n(f) \delta(f - f')$$

$$= \int_{0}^{\infty} \mathrm{d}f \, S_n(f) \ . \quad (2.10)$$

Given a detector output, the challenge is to extract the signal. There is a well known solution to this problem that involves constructing a Wiener optimal filter (Wiener, 1949*b*). Let $K(t)$ be a real filter function with Fourier transform $\tilde{K}(f)$. Convolving this with the detector output gives a contribution from the signal and a contribution from the noise,

$$(s * K)(\tau) = \int_{-\infty}^{\infty} \mathrm{d}t \ [h(t) + n(t)] \, K(t - \tau) \approx \mathcal{S} + \mathcal{N} \ . \quad (2.11)$$

The quantity $\mathcal{S}$, which includes contributions from both the signal and the noise, is defined as the expectation of the convolution in Eq. 2.11 when a signal is present, maximised by varying the offset to achieve the best overlap with the data. Since the expectation of pure noise is zero it follows that

$$\mathcal{S} = \int_{-\infty}^{\infty} \mathrm{d}t \ h(t)K(t) = \int_{-\infty}^{\infty} \mathrm{d}t \ h(t)K^*(t) = \int_{-\infty}^{\infty} \mathrm{d}f \ \tilde{h}(f)\tilde{K}^*(f) \ . \quad (2.12)$$

The squared contribution from noise $\mathcal{N}^2$ is defined as the mean square of the convolution in Eq. 2.11 when no signal is present,

$$\begin{aligned}\mathcal{N}^2 &= \int_{-\infty}^{\infty} \mathrm{d}t \int_{-\infty}^{\infty} \mathrm{d}t' \, K(t)K(t') \left\langle n(t)n(t') \right\rangle \\ &= \int_{-\infty}^{\infty} \mathrm{d}t \int_{-\infty}^{\infty} \mathrm{d}t' \, K(t)K^*(t') \int_{-\infty}^{\infty} \mathrm{d}f \int_{-\infty}^{\infty} \mathrm{d}f' \, \left\langle \tilde{n}(f)\tilde{n}^*(f') \right\rangle \exp\left[2\pi \mathrm{i}(ft - f't')\right] \\ &= \int_{-\infty}^{\infty} \mathrm{d}f \, \frac{1}{2} S_n(f) \tilde{K}(f) \tilde{K}^*(f) \, , \end{aligned} \tag{2.13}$$

using the definition of $S_n(f)$ from Eq. 2.5. Hence the signal-to-noise ratio (SNR) ϱ is given by

$$\varrho^2 = \frac{\mathcal{S}^2}{\mathcal{N}^2} = \frac{\left(\frac{1}{2} S_n(f) \tilde{K}(f) \middle| \tilde{h}(f)\right)^2}{\left(\frac{1}{2} S_n(f) \tilde{K}(f) \middle| \frac{1}{2} S_n(f) \tilde{K}(f)\right)}, \tag{2.14}$$

where we have introduced the inner product between signal $\tilde{A}$ and $\tilde{B}$ as (Finn, 1992)

$$\left(\tilde{A}(f) \middle| \tilde{B}(f)\right) = 4\Re \left\{ \int_0^{\infty} \mathrm{d}f \, \frac{\tilde{A}^*(f)\tilde{B}(f)}{S_n(f)} \right\} . \tag{2.15}$$

The optimum filter is that function $\tilde{K}(f)$ which maximises the SNR in Eq. 2.14. From the Cauchy–Schwarz inequality, it follows that the optimum filter is

$$\tilde{K}(f) = \frac{\tilde{h}(f)}{S_n(f)} \, . \tag{2.16}$$

This is the Wiener filter, which may be multiplied by an arbitrary constant since this does not change the SNR. Using this form for $\tilde{K}(f)$, the expectation of the squared SNR is

$$\varrho^2 = \int_0^{\infty} \mathrm{d}f \frac{4|\tilde{h}(f)|^2}{S_n(f)} = \left(\tilde{h}(f) \middle| \tilde{h}(f)\right) . \tag{2.17}$$

In order to construct the Wiener filter, it is necessary to know *a priori* the form of the signal, $\tilde{h}(f)$, for this reason the Wiener filter is sometimes called the *matched* filter.

Whilst the magnitude of the Fourier transform of the signal $|\tilde{h}(f)|$ provides a simple quantification of the gravitational wave amplitude as a function of frequency, it has one main deficiency. For an inspiralling source, the instantaneous amplitude can be orders of magnitude below the noise level in a detector; however, as the signal continues over many orbits, the SNR can be integrated up to a detectable level. It is useful to have a quantification of the amplitude that accounts for this effect; we shall now describe how this can be achieved.

2.1.1 Characteristic strain

The characteristic strain h_c is designed to include the effect of integrating an inspiralling signal. Its counterpart for describing noise is the noise amplitude h_n. These are defined as

$$[h_c(f)]^2 = 4f^2 \left|\tilde{h}(f)\right|^2 , \tag{2.18}$$

$$[h_n(f)]^2 = fS_n(f) , \tag{2.19}$$

such that the SNR in Eq. 2.17 may be written

$$\varrho^2 = \int_{-\infty}^{\infty} \mathrm{d}\,(\log f)\ \left[\frac{h_c(f)}{h_n(f)}\right]^2 . \tag{2.20}$$

The strain amplitudes $h_c(f)$ and $h_n(f)$ are dimensionless. Using this convention, when plotting on a log–log scale, the area between the source and detector curves is related to the SNR via Eq. 2.20. This convention allows the reader to integrate by eye to assess the detectability of a given source (see Fig. 4.1).

An additional advantage of this convention is that the values on the strain axis for the detector curve $h_n(f)$ have a simple physical interpretation: they correspond to the root-mean-square noise in a bandwidth f. One downside to plotting characteristic strain is that the values on the strain axis $h_c(f)$ do not directly relate to the amplitude of the waves from the source. Another disadvantage is that applying Eq. 2.18 to a monochromatic source gives a formally undefined answer. The correct identification of characteristic strain for a monochromatic source is the amplitude of the wave times the square root of the number of periods observed (see Sec. 2.2.1).

2.1.2 Power spectral density

A second commonly used quantity for sensitivity curves is the square root of the PSD or the amplitude spectral density (see Fig. 4.2). When discussing a detector, rearranging Eq. 2.19 gives

$$\sqrt{S_n(f)} = h_n(f) f^{-1/2} ; \tag{2.21}$$

by analogy, we can define an equivalent for source amplitudes

$$\sqrt{S_h(f)} = h_c(f) f^{-1/2} = 2f^{1/2} \left|\tilde{h}(f)\right| , \tag{2.22}$$

where we have used Eq. 2.18. Both $\sqrt{S_n(f)}$ and $\sqrt{S_h(f)}$ have units of $\mathrm{Hz}^{-1/2}$. The root PSD is the most frequently plotted quantity in the literature.

The PSD, as defined by Eq. 2.5, has the nice property, demonstrated in Eq. 2.10, that integrated over all positive frequencies it gives the mean square amplitude of the noise in the detector. However, in one important regard it is less appealing than characteristic strain: the height of the source above the detector curve is no longer directly related to the SNR.

2.1.3 Energy density

A third way of describing the amplitude of a gravitational wave is through the energy carried by the radiation. This has the advantage of having a clear physical significance. The energy density is most commonly used in sensitivity curves showing stochastic backgrounds of gravitational waves (see Sec. 2.2.2).

The energy in gravitational waves is described by the Isaacson stress–energy tensor (Misner et al., 1973, section 35.15)

$$T_{\mu\nu} = \frac{c^4}{32\pi G}\left\langle \partial_\mu \bar{h}_{\alpha\beta} \partial_\nu \bar{h}^{\alpha\beta} \right\rangle \ , \tag{2.23}$$

where the angle brackets denote averaging over several wavelengths or periods, and $\bar{h}_{\alpha\beta}$ is the transverse-traceless metric perturbation. The energy density ρc^2 is given by the T_{00} component of this tensor. Consequently (cf. Berry and Gair, 2013),

$$\rho c^2 = \frac{c^2}{16\pi G} f_0 \int_{-\infty}^{\infty} \mathrm{d}f \ (2\pi f)^2 \, \tilde{h}(f)\tilde{h}^*(f) \tag{2.24}$$

$$= \int_0^{\infty} \mathrm{d}f \ \frac{\pi c^2}{4G} f f_0 S_h(f) \ , \tag{2.25}$$

where Eq. 2.22 has been used, and f_0 is a constant determined by the size of the region averaged over in Eq. 2.23. There is no natural way to fix f_0, different choices are more convenient for different types of source (e.g. monochromatic waves vs stochastic backgrounds). Here we choose to let $f_0 = f$, then the integrand in Eq. 2.25 is taken as the definition of the spectral energy density, (Hellings and Downs, 1983)

$$S_{\mathrm{E}}(f) = \frac{\pi c^2}{4G} f^2 S_h(f) \ ; \tag{2.26}$$

this is a reasonable (but not unique) definition for the energy per unit volume of space, per unit frequency for a generic burst-like signal. A corresponding expression for the noise can be formulated by replacing $S_h(f)$ with $S_n(f)$.

Cosmological studies often work in terms of the dimensionless quantity $\Omega_{\rm GW}$, the energy density per logarithmic frequency interval, normalised to the critical density of the Universe $\rho_{\rm c}$,

$$\Omega_{\rm GW}(f) = \frac{f S_{\rm E}(f)}{\rho_{\rm c} c^2} \,. \tag{2.27}$$

The critical density is

$$\rho_{\rm c} = \frac{3H_0^2}{8\pi G} \,, \tag{2.28}$$

where H_0 is the Hubble constant, commonly parametrized as

$$H_0 = h_{100} \times 100 \text{ km s}^{-1} \text{ Mpc}^{-1}. \tag{2.29}$$

The reduced Hubble parameter h_{100} has nothing to do with strain. The most common quantity related to energy density to be plotted on sensitivity curves is $\Omega_{\rm GW} h_{100}^2$, see Fig. 4.3, as this removes sensitivity to the (historically uncertain) measured value of the Hubble constant.

This quantity has one aesthetic advantage over the others: it automatically accounts for there being less energy in low frequency waves of the same amplitude. However, unlike characteristic strain, the area between the source and detector curves is no longer simply related to the SNR.

2.1.4 Relating the descriptions

The dimensionless energy density in GWs $\Omega_{\rm GW}$, spectral energy density $S_{\rm E}$, one-sided PSD S_h, characteristic strain $h_{\rm c}$ and frequency-domain strain $\tilde{h}(f)$ are related via

$$H_0^2 \Omega_{\rm GW}(f) = \frac{8\pi G}{3c^2} f S_{\rm E}(f) = \frac{2\pi^2}{3} f^3 S_h(f) = \frac{2\pi^2}{3} f^2 \left[h_{\rm c}(f)\right]^2 = \frac{8\pi^2}{3} f^4 \left|\tilde{h}(f)\right|^2 \,, \tag{2.30}$$

using Eq. 2.18, Eq. 2.22, Eq. 2.26, Eq. 9.20 and Eq. 2.28. Corresponding expressions for the noise are obtained by substituting $S_n(f)$ for $S_h(f)$, $h_n(f)$ for $h_{\rm c}(f)$ and $\tilde{n}(f)$ for $\tilde{h}(f)$.

2.2 Types of source

Gravitational wave signals can be broadly split into three categories: those from well-modelled sources, for which we have a description of the expected waveform; stochastic backgrounds, for which we can describe the statistical behaviour; and unmodelled (or

poorly-modelled) transient sources. The classic example of a well-modelled source is the inspiral of two compact objects, this is discussed in Sec. 2.2.1. Stochastic backgrounds can either be formed from many overlapping sources, which could be modelled individually, or from some intrinsically random process, these are discussed in Sec. 2.2.2. An example of an unmodelled (or poorly-modelled) transient source is a supernova; searches for signals of this type are often called burst searches and are discussed in Sec. 2.2.3.

2.2.1 Inspirals

Inspiralling binaries may be the most important gravitational wave source. They spend a variable amount of time in each frequency band. If ϕ is the orbital phase, then the number of cycles generated at frequency f can be estimated as

$$N_{\rm cycles} = \frac{f}{2\pi}\frac{{\rm d}\phi}{{\rm d}f} = \frac{f^2}{\dot{f}} \ , \tag{2.31}$$

where an overdot represents the time derivative and $\dot{\phi} = 2\pi f$. The squared SNR scales with $N_{\rm cycles}$, so it would be expected that $h_{\rm c}(f) \approx \sqrt{N_{\rm cycles}}|\tilde{h}(f)|$.

The form for $h_{\rm c}$ can be derived from the Fourier transform in the stationary-phase approximation. Consider a source signal with approximately constant (root-mean-square) amplitude h_0 and central frequency f'. In this case,

$$h(t) = \sqrt{2}h_0 \cos\left[\phi(t)\right] \ , \tag{2.32}$$

$$\tilde{h}(f) = \frac{h_0}{\sqrt{2}}\int_{-\infty}^{\infty} {\rm d}t \ \exp\left\{2\pi{\rm i}\left[\frac{\phi(t)}{2\pi t} - f\right]t\right\} + \exp\left\{-2\pi{\rm i}\left[\frac{\phi(t)}{2\pi t} + f\right]t\right\} \ . \tag{2.33}$$

Without loss of generality, we can assume an initial phase of zero, such that $\phi(0) = 0$. The largest contribution to the integral comes from where the argument of the exponentials is approximately zero. For the first term, this occurs when $f = f'$, then the term in brackets is

$$\left[\frac{\phi(t)}{2\pi t} - f\right]_{f=f'} = f' + \dot{f}'t + \mathcal{O}\left(t^2\right) - f' = \dot{f}'t + \mathcal{O}\left(t^2\right). \tag{2.34}$$

The higher-order terms cause the expoential to oscillate rapidly, such that these terms integrate to zero and may be neglected. Performing a similar expansion about $f = -f'$ for the second term in Eq. 2.33, and then evaluating the Gaussian integrals gives

$$\begin{aligned}\tilde{h}(f) &\simeq \frac{h_0}{\sqrt{2}}\int_{-\infty}^{\infty} {\rm d}t \ \exp\left(2\pi{\rm i}\dot{f}'t^2\right) + \exp\left(-2\pi{\rm i}\dot{f}'t^2\right) \\ &\simeq \frac{h_0}{\sqrt{2\dot{f}'}}.\end{aligned} \tag{2.35}$$

From Eq. 2.18 and Eq. 2.35, the characteristic strain for inspiralling sources is given by Finn and Thorne (2000)

$$h_{\rm c}(f) = \sqrt{\frac{2f^2}{\dot{f}}} h_0 \ . \tag{2.36}$$

Equation 2.18 should be considered as the definition of characteristic strain and Eq. 2.36 a consequence of it for inspirals. Equation 2.36 is the relation between $h_{\rm c}(f)$ and the instantaneous root-mean-square amplitude h_0 for an inspiralling source; for other types of source a new relation satisfying Eq. 2.20 has to be found.

2.2.2 Stochastic backgrounds

Another important source of gravitational waves is that of stochastic backgrounds, which can be produced from a large population of unresolvable sources. These can be at cosmological distances, where it is necessary to distinguish the frequency in the source rest frame $f_{\rm r}$ from the measured frequency f; the two are related through the redshift z via $f_{\rm r} = (1+z)f$. The comoving number density of sources ν producing the background is also a function of redshift; if the sources producing the background are all in the local Universe, then simply set $\nu(z) = \nu_0 \delta(z)$ and replace $d_{\rm L}(z)$ with d in all that follows, where $d_{\rm L}(z)$ and d are respectively the luminosity and comoving distances to the source, $d_{\rm L}(z) = (1+z)d$.

We shall assume that the individual sources are binaries, in which case the number density of sources is also a function of the component masses. It is convenient to work in terms of the chirp mass, defined as $\mathcal{M} = \mu^{3/5} M^{2/5}$, where μ is the reduced mass and M is the total mass of the binary. The comoving number density of sources shall be represented by $\nu(z, \mathcal{M})$.

Equation 2.30 gives an expression for the energy density in gravitational waves per logarithmic frequency interval,

$$f S_{\rm E}(f) = \frac{\pi c^2}{4G} f^2 \left[h_{\rm c}(f)\right]^2 \ . \tag{2.37}$$

Note that for consistency, we will continue to use the definitions of energy density in gravitational waves given in Sec. 2.1.3, although alternative conventions may be more convenient when discussing stochastic signals alone. The total energy emitted in the logarithmic frequency interval $\mathrm{d}\left(\log f_{\rm r}\right)$ by a single binary in the population is

$[\mathrm{d}E_{\mathrm{GW}}/\mathrm{d}(\log f_{\mathrm{r}})]\,\mathrm{d}(\log f_{\mathrm{r}})$; the energy density may be written as

$$fS_{\mathrm{E}}(f) = \int_0^\infty \mathrm{d}z\, \frac{\mathrm{d}\nu}{\mathrm{d}z} \frac{1}{(1+z)} \frac{1}{[d_{\mathrm{L}}(z)]^2} \frac{\mathrm{d}E_{\mathrm{GW}}}{\mathrm{d}\,(\log f_{\mathrm{r}})}\ , \tag{2.38}$$

where the factor of $(1+z)^{-1}$ accounts for the redshifting of the energy.

For simplicity, consider the background to comprise of binaries in circular orbits, with frequencies $f_{\mathrm{GW}} = f_{\mathrm{r}}/2$, which are far from their last stable orbit. The energy radiated may then be calculated using the quadrupole approximation (Peters and Mathews, 1963). The energy in gravitational waves from a single binary per logarithmic frequency interval is

$$\frac{\mathrm{d}E_{\mathrm{GW}}}{\mathrm{d}\,(\log f_{\mathrm{r}})} = \frac{G^{2/3}\pi^{2/3}}{3}\mathcal{M}^{5/3} f_{\mathrm{r}}^{2/3} \tag{2.39}$$

between minimum and maximum frequencies set by the initial and final radii of the binary orbit respectively. Here, we assume that the maximum and minimum frequencies are outside of the range of our detector and hence can be neglected. Using Eq. 2.37, Eq. 2.38 and Eq. 2.39, an expression for characteristic strain can now be found (Sesana et al., 2008)

$$[h_{\mathrm{c}}(f)]^2 = \frac{4G^{5/3}}{3\pi^{1/3}c^2} f^{-4/3} \int_0^\infty \mathrm{d}z \int_0^\infty \mathrm{d}\mathcal{M}\, \frac{\mathrm{d}^2\nu}{\mathrm{d}z\,\mathrm{d}\mathcal{M}} \frac{1}{[d_{\mathrm{L}}(z)]^2} \left(\frac{\mathcal{M}^5}{1+z}\right)^{1/3}\ . \tag{2.40}$$

From (2.40) it can be seen that the characteristic strain due to a stochastic background of binaries is a power law in frequency with spectral index $\alpha = -2/3$. The amplitude of the background depends on the population statistics of the binaries under consideration via $\nu(z, \mathcal{M})$. The power law is often parametrised as

$$h_{\mathrm{c}}(f) = A\left(\frac{f}{f_0}\right)^{\alpha}\ , \tag{2.41}$$

and constraints are then placed on A. In practice, this power law also has upper and lower frequency cut-offs related to the population of source objects.

A stochastic background from other sources, such as cosmic strings or relic gravitational waves from the early Universe, can also be written in the same form as Eq. 2.41, but with different spectral indices: $\alpha = -7/6$ for cosmic strings or α in the range -1 to -0.8 for relic gravitational waves (Jenet et al., 2006).

An alternative method for graphically representing the sensitivity of a detector to stochastic backgrounds, called the *power-law-integrated sensitivity curve*, was suggested by Thrane and Romano (2013). This method accounts for the there being power across all

frequencies in the sensitivity band by integrating the noise-weighted signal over frequency. As our aim here is to present stochastic backgrounds alongside other types of sources for comparison, we will not use this approach.

2.2.3 Burst sources

Some sources of gravitational waves can produce signals with large amplitudes, greater than the detector noise. The typical duration of such a signal is short, of the order of a few wave periods, and so there is not time to accumulate SNR in each frequency band as for inspirals. As a consequence, waveform models are not required for detection; we simply rely on identifying the excess power produced by these burst sources. Typically, we may be looking for signals from core-collapse supernovae (Ott, 2009), the late stages of merging compact binaries (Abadie et al., 2012*b*), cosmic strings (Aasi et al., 2014*a*), or more generally, signals from any unexpected or poorly modelled sources.

Burst searches are often carried out using time–frequency techniques. The data stream from a detector is temporally split into segments, the length of which can be tuned to give greater sensitivity to particular sources. Each segment is then transformed into the frequency domain, whitened and normalised to the noise spectrum of the detector to produce a time–frequency plot. Potential gravitational wave signals are identified by searching for clusters of pixels that contain an excess of power (e.g., Abadie et al., 2012*a*).

The presence of excess power across a number of pixels eliminates modelled noise sources, but such a cluster may also be caused by atypical noise within a detector. We can improve our confidence of a gravitational wave signal by making use of information obtained from other detectors. Signals across a network of detectors should have compatible arrival times (given the sky direction) as well as consistent amplitudes, frequencies and shapes of the waveform. Different pipelines are currently in use that analyse the signal consistency in different ways: both coincidence searches (Chatterji et al., 2004) and fully coherent methods (Klimenko et al., 2008*b*) are used.

An important aspect of burst search algorithms is to accurately estimate the noise properties within each time segment. To this extent, null data streams can be constructed that are insensitive to real gravitational wave signals. In order to estimate the false alarm rate, the data from different detectors can be shifted in time to remove any genuine coincident signals (Cannon, 2008). These time-shifts are then analysed to simulate the potential occurence of coincident noise events. The algorithms are tuned using

time-shifted data to ensure there is no bias in the final search.

As discussed in Sec. 2.2.1, the expected relation between $h_{\rm c}(f)$ and a typical waveform $\tilde{h}(f)$ is

$$h_{\rm c}(f) = \sqrt{N_{\rm cycles}}\left|\tilde{h}(f)\right| \ , \tag{2.42}$$

where $N_{\rm cycles}$ is the number of cycles of radiation generated by the source, which is of order unity for bursts.

An alternative characterisation of the signal amplitude commonly used for burst sources is the root-sum-square of the waveform polarisations:

$$h_{\rm rss}^2 = \int {\rm d}t \ |h_+(t)|^2 + |h_\times(t)|^2 \ . \tag{2.43}$$

For a linearly polarised wave, with $\tilde{h}(f)$ constant across the bandwidth Δf, this is approximately related to the characteristic strain via

$$h_{\rm rss} \simeq \left|\tilde{h}(f)\right| \sqrt{\Delta f} \ , \tag{2.44}$$

where we have neglected the detector response functions (see Sec. 2.4), which are of order unity. In this work, we favour a constant $h_{\rm c}(f)$ rather than $h_{\rm rss}$ for consistency with the other types of source where the bandwidth is detector specific.

2.3 Resonant detectors

Generally speaking gravitational waves interact only very weakly with matter; nevertheless interactions do occur at some level, and energy may be absorbed in this process. A resonant detector seeks to detect gravitational waves by measuring the vibrations excited in a material object due to the passage of a gravitational wave.

The field of experimental gravitational wave physics began in the 60s with the construction of the first cylindrical bar resonant detectors by Weber (1960). The sensitivity of resonant bars to gravitational waves has been improved greatly since these early efforts (see, for example, Astone et al. (2010)) however no direct detection has been made via this approach. As well as the classic *bar*-shaped resonant detectors, other geometries for resonant detectors have also been considered; for example the spherical MiniGrail detector (de Waard et al., 2006).

In this section we will briefly describe the operating principles of an idealised resonant bar detector. In particular the geometric and frequency response function defined in

Eq.2.2 will be derived for this ideal case. The principle sources of noise limiting the sensitivities of these detectors will be discussed. For a more detailed discussion of resonant bar detectors see either the original work of Weber (1960), or for a modern treatment the textbook by Maggiore (2007). Finally, we conclude by briefly discussing various attempts to detect gravitational waves by measuring the resonantly excited oscillations in naturally occurring objects.

2.3.1 Operating principles of a resonant bar detector

A typical bar detector consists of a thin cylinder of length L, placed along the z-axis so that the ends are at coordinates $z = \pm L/2$. It will turn out that the most sensitive frequency of such a detector is set by the sound travel time along the length of the bar, $f \sim v_s/L$. Because the speed of sound is much less than the speed of light the response of the detector may be safely analysed in the long wavelength limit; $\lambda \gg L$.

From the equation for geodesic deviation the force on an element of the bar with mass $\mathrm{d}m$ is given by the second derivative of the metric perturbation expressed in the transverse-traceless gauge.

$$\mathrm{d}F_i = \frac{1}{2}\ddot{h}_{ij}x^j \mathrm{d}m\,. \tag{2.45}$$

The longitudinal displacement, $u(t, z)$, of the bar satisfies a wave equation (with natural speed v_s) driven by the force due to gravitational waves.

$$\left(\partial_t^2 - v_s^2\partial_z^2\right) u(t, z) = \frac{1}{2}z\ddot{h}_{zz}\,. \tag{2.46}$$

The free ends of the bar are unable to transmit any elastic energy to the surroundings and so the appropriate boundary conditions for this equation are the Neumann conditions $\partial_z u(t, z) = 0$ at $z = \pm L/2$.

The fundamental mode of oscillation of the bar is given by $u_0(t, z) = \xi(t)\sin(\pi z/L)$ where $\xi(t)$ satisfies

$$\left(\partial_t^2 + \omega_0^2\right)\xi(t) = \frac{2L}{\pi^2}\ddot{h}_{zz}\,, \tag{2.47}$$

and $\omega_0 = \pi v_s/L$ is the natural frequency of the bar. In reality of course, the oscillations of the bar will be slightly damped. In this case Eq. 2.47 becomes

$$\left(\partial_t^2 + \gamma\partial_t + \omega_0^2\right)\xi(t) = \frac{2L}{\pi^2}\ddot{h}_{zz}\,. \tag{2.48}$$

The damping is described by a *quality factor* $Q = \omega/\gamma$. Experimentally the bars are designed to have minimal damping, or as high a Q-factor as possible.

It can be seen from Eq. 2.48 that incident gravitational waves drive oscillations in the bar. The main difficulty in designing a practical resonant bar detector is detecting these very small oscillations; i.e. designing an amplification and read out system. Joseph Weber pioneered an approach that involved attaching a *resonant-transducer*[3] to the ends of the bar. The small oscillations in the bar were used to drive comparatively large oscillations of the transducers, which in turn drove oscillations in a piezoelectric crystal. The voltage produced by the piezoelectric crystal was then measured.

Equation 2.48 describing oscillations in the bar may be Fourier transformed to give

$$\left(-\omega^2 - \mathrm{i}\gamma\omega + \omega_0^2\right)\tilde{\xi}(\omega) = \frac{2L\omega^2}{\pi^2}\tilde{h}_{zz} \tag{2.49}$$

$$\tilde{\xi}(\omega) = \frac{2L\omega^2}{\pi^2\left((\omega_0^2 - \omega^2) - \mathrm{i}\gamma\omega\right)}\tilde{h}_{zz} \equiv \tilde{R}(\omega)\tilde{h}_{zz}\,. \tag{2.50}$$

This constitutes the frequency response (defined in Eq. 2.2) for the resonant bar, for a high Q bar this response is strongly peaked around the resonant frequency $\omega = \omega_0$. Of course, we don't actually want the response function of the bar, what we really need is the response function of the output of the transducer attached to the end of the bar. This will be given by the convolution of the response function of the bar with the response function of the amplification and read out system. This can make a significant difference, for the low mass resonant transducer described above, the effect on the response function $\tilde{R}(\omega)$ is to split the single narrow resonant peak in two, $\omega_\pm = \omega_0 \pm \Delta\omega$. A discussion of the details of the modification to the response function from the amplification and read out systems will depend upon the details of the experimental setup, and is beyond the scope of this introduction; for more detail see Maggiore (2007).

As well as the frequency response, we may also examine the geometric response function of an ideal resonant bar detector (also defined for the general detector in Eq. 2.2). From Eq. 2.48 it can be seen that it is the component of the strain along the bar, h_{zz}, that excites the oscillations. For a linearly polarised plane gravitational wave travelling along the direction $\vec{k} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ with polarisation angle ψ (i.e. $h_{ij} = e^+_{ij}\cos\psi + e^\times_{ij}\sin\psi$) the relevant component of the strain is given by

$$F(\theta, \phi, \psi) \propto l^i l^j h_{ij}(\theta, \phi, \psi)\,, \tag{2.51}$$

[3] A resonant transducer is a low mass oscillator tuned to the frequency of gravitational waves being searched for. Small oscillations driving the transducer near resonance result in larger output oscillations of the transducer.

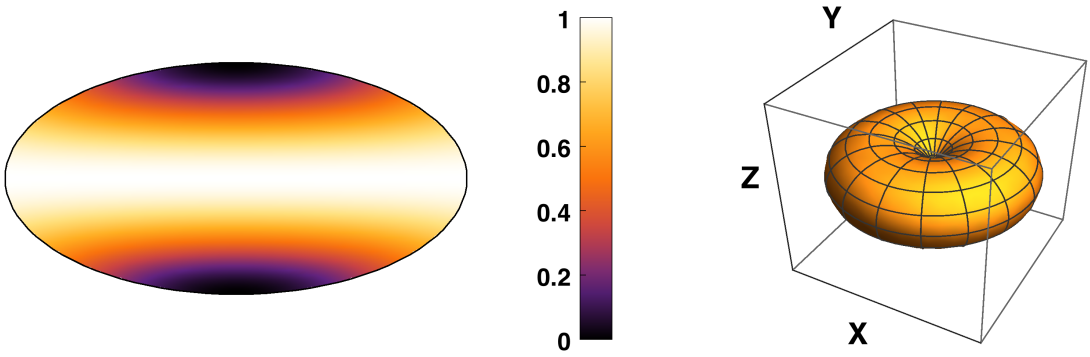


Figure 2.1: The angular response function of an resonant bar detector shown both as a surface plot and in an Aitoff–Hammer projection. The quantity that is plotted is is the polarisation average, $\left[(1/2\pi)\int \mathrm{d}\psi\ (F^{+})^2 + (F^{\times})^2\right]^{1/2}$. The response is a function of two sky angles, θ and ϕ, and varies between 0 and 1. The bar lies along the z-axis, along the zero in the response function.

where $\vec{l} = (0, 0, 1)$ is the unit vector along the bar. The polarisation averaged response function, normalised such that $0 \leq F \leq 1$, is plotted as a function on the sphere in Fig. 2.1. It can be seen that the bar detector is sensitive to gravitational waves that cross the axis of the bar.

Of course the sensitivity of a detector is limited by the noise present in the read out. For most resonant detectors the sensitivity is primarily limited by two noise sources: thermal noise, and read out noise. The read out noise is that noise associated with the amplification of the output voltage from the piezoelectric crystal to a detectable level, this amplification is done in practice with a superconducting quantum interference device (SQUID). The read out noise is approximately white across the bandwidth of interest. Thermal noise is due to the inevitable oscillations of the atoms (or molecules) in the bar when held at a finite temperature. These can be mistaken for the very small bulk motion of the bar caused by a gravitational wave. Thermal noise may be minimised by designing the bar to have as high a Q-factor as possible; a high Q-factor means that large bulk motions in the bar persist for a long time before being damped out into thermal vibrations of the bar, conversely this means that thermal vibrations couple very weakly to the bulk motions that we are trying to measure. Thermal noise may also be minimised by cooling the detectors to cryogenic temperatures.

The current best strain sensitivities achieved by resonant detectors are around

$S_h \sim 10^{-21}\,\mathrm{Hz}^{-1/2}$ at frequencies around $f \sim 920\,\mathrm{Hz}$. Comparable sensitivities have been achieved by several groups, notably the NAUTILUS[4], EXPLORER[5], and AURIGA[6] experiments.

2.3.2 Naturally occurring resonant detectors

In the previous section it was shown that gravitational waves may excite measurable oscillations in a specially constructed cylindrical bar. Gravitational waves may also excite detectable oscillations in naturally occurring objects. There have been a number of efforts aimed at detecting, or placing limits on, a stochastic background of gravitational waves across a range of frequencies by measuring the oscillations in a number of naturally occurring systems.

For example. gravitational waves may excite free surface oscillations in the Earth, which in principle may be detected via a network of broadband seismometers placed around the globe. Using this technique a limit was placed on the energy in the background in the frequency range $(0.05 - 1)\,\mathrm{Hz}$ by Coughlin and Harms (2014*b*). The limit obtained for the stochastic background was $\Omega_{\mathrm{GW}} < 1.2 \times 10^8$, where Ω_{GW} is the energy density per logarithmic frequency interval, normalised to the critical density of the Universe (see Sec. 2.1.4). A similar approach has also been applied to lunar seismic activity using data from the Apollo seismic array (Coughlin and Harms, 2014*a*); this yields an improved limit of $\Omega_{\mathrm{GW}} < 1.2 \times 10^5$ in the frequency range $(0.1 - 1)\,\mathrm{Hz}$.

As well as terrestrial and lunar seismology, *astro-seismology* may also be used to detect gravitational waves. A background of gravitational waves may resonantly interact with any star (including our own sun) and excite low order acoustic modes; these can in principle be detected by high precision measurements of the surface velocity of the star. Astro and helio-seismological observations may therefore be able to detect a background of gravitational waves in the frequency range $(0.2 - 5)\,\mathrm{mHz}$. The sensitivity of such an approach is still rather uncertain, however Lopes and Silk (2014) suggest that limits in the range $\Omega_{\mathrm{GW}} < (10^{-10} - 10^{-4})$ may be possible.

Both the astro and terrestrial seismology approaches aim to detect the oscillations in a single object. However two objects gravitationally bound in an orbit will also be affected by the passage of a gravitational wave with a wavelength that roughly matches

[4] `http://www.roma1.infn.it/rog/nautilus/`

[5] `http://www.roma1.infn.it/rog/explorer/`

[6] `http://www.auriga.lnl.infn.it`

the binaries' orbital radius, r. A stochastic background of gravitational waves at these frequencies would cause the orbital properties (radius and eccentricity) to change with time in a measurable way. The absence of a detection of this effect in precise timing observations of binary pulsars over a long baseline period of time was used by Hui et al. (2013) to constrain the energy density in the stochastic background at frequencies around $f \sim c/r \approx 0.1\,\mathrm{mHz}$. The limit obtained via this approach was $\Omega_{\mathrm{GW}} < 80$.

2.4 Interferometric detectors

Gravitational wave interferometers are conceptually easier to understand than the resonant detectors discussed above. However, because of the need to make the interferometer large (several kilometres for the current advanced generation detectors) to achieve astrophysically interesting sensitivities their development began after that of the resonant bars. To some extent the development of interferometers was spurred on during the 70s by the need to confirm or refute the earlier claims of detections by Joseph Weber.

2.4.1 Operating principles of an interferometric detector

All of the Earth bound man-made detectors discussed in this section utilise the principle of interferometry. Such detectors work by taking a beam of monochromatic light and splitting it into two beams travelling at an angle to each other. Each beam is passed in to an optical cavity where it undergoes a number of round trips before being recombined to form an interference pattern. The ends of the cavity are, in the ideal case, freely floating test masses which move relative to each other in response to a passing gravitational wave, this effect is measured by observing the changing interference pattern.

In this section space-based gravitational wave detectors, and pulsar timing arrays (PTAs) will also be discussed. The physical principles underpinning the operation of these are similar to the ground-based interferometers although with significant technical differences. For a detailed discussion of the method of *time-delay interferomety* vital to the space-based efforts to detect gravitational waves see Tinto and Dhurandhar (2005). A brief discussion of the principles of PTAs will be given in Sec. 2.4.4 below; for a review of the efforts to detect gravitational waves using PTAs see, for example, Manchester (2010).

The response of a detector to an incident plane-fronted gravitational wave depends upon the relative orientations of the detector and the incoming wave. Let us choose the

origin of our coordinate system to be the beam-splitter of the interferometer, and l_1^i and l_2^i to be unit 3-vectors pointing along the two arms. In the absence of noise the output of the detector is the difference in strain between the two arms (Thorne, 1987)

$$h(t) = \frac{1}{2} h_{ij} \left(l_1^i l_1^j - l_2^i l_2^j \right) \ , \tag{2.52}$$

where h_{ij} are the spatial components of the metric perturbation. Let $\hat{r}^i$ be the unit 3-vector pointing towards the source of the GWs, with spherical polar angles (θ, ϕ) relative to some axes fixed to the detector, and let p^i and q^i be unit vectors orthogonal to $\hat{r}^i$. We can now define the basis tensors

$$H_{ij}^{+} = p_i p_j - q_i q_j \ , \tag{2.53}$$

$$H_{ij}^{\times} = p_i q_j + q_i p_j \ . \tag{2.54}$$

There remains a freedom in the coordinates described, a rotation of p^i and q^i through an angle ψ about $\hat{r}^i$ known as the polarization angle. For a single frequency component, the strain induced by a gravitational wave may be written as

$$h_{ij} = A_+ H_{ij}^{+} \cos\left(2\pi f t\right) + A_\times H_{ij}^{\times} \cos\left(2\pi f t + \Delta\phi\right) \ , \tag{2.55}$$

where A_+ and $A_\times$ are the amplitudes of the two polarisation states. Combining Eq. 2.52 and Eq. 2.55 allows the detector output to be written as

$$h(t) = F^{+}(\theta, \phi, \psi) A_+ \cos\left(2\pi f t\right) + F^{\times}(\theta, \phi, \psi) A_\times \cos\left(2\pi f t + \Delta\phi\right) \ , \tag{2.56}$$

where the response functions inherit their angular dependence from the choice of coordinates

$$F^{+}(\theta, \phi, \psi) = \tfrac{1}{2} H_{ij}^{+} \left(l_1^i l_1^j - l_2^i l_2^j \right) \ , \tag{2.57}$$

$$F^{\times}(\theta, \phi, \psi) = \tfrac{1}{2} H_{ij}^{\times} \left(l_1^i l_1^j - l_2^i l_2^j \right) \ . \tag{2.58}$$

The response function of a two-arm interferometric detector is quadrupolar, an example is plotted in Fig. 2.2.

Equations 2.57 and 2.58 constitute the geometric response defined in Eq. 2.2 for a "L-shaped" interferometer. In contrast to a resonant detector the frequency response is relatively mild function of frequency; in the ideal case of a detector comprising two test masses the response is completely achromatic, i.e. $\tilde{R}(f) = \text{const}$. In practice the

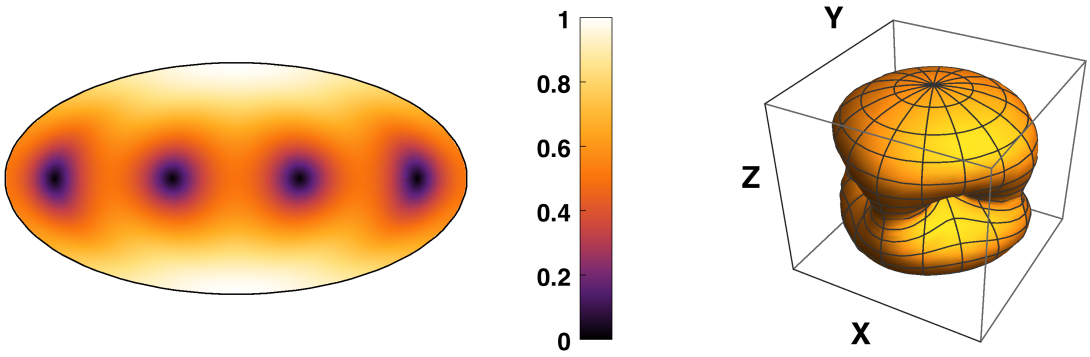


Figure 2.2: The angular response function of an interferometric detector shown both as a surface plot and in an Aitoff–Hammer projection. The quantity that is plotted is is the polarisation average, $\left[(1/2\pi)\int \mathrm{d}\psi\ (F^{+})^2 + (F^{\times})^2\right]^{1/2}$. The response is a function of two sky angles, θ and ϕ, and varies between 0 and 1. The two detector arms lie in the x–y plane either side of one of the zeros in the response.

response function for the interferometer may be measured by driving the mirrors at a number of known frequencies, and then interpolated using (for example) cubic splines. The uncertainties arising from this process may be marginalised over during the analysis. For a discussion of how this was done for the first detection of gravitational waves from GW150914 by the aLIGO detectors see The LIGO Scientific Collaboration and the Virgo Collaboration (2016*a*) and The LIGO Scientific Collaboration and the Virgo Collaboration (2016*e*).

Throughout the rest of this section the phrase "detector sensitivity" is taken to refer to the polarisation and sky averaged sensitivity F, where

$$F^2 = \int_0^{2\pi} \frac{\mathrm{d}\psi}{2\pi} \int_0^{2\pi} \frac{\mathrm{d}\phi}{2\pi} \int_0^{\pi} \frac{\sin\theta\,\mathrm{d}\theta}{2} \left[F^{+}\left(\theta,\phi,\psi\right)^2 + F^{\times}\left(\theta,\phi,\psi\right)^2\right] . \tag{2.59}$$

For a single 90°-interferometer, such as LIGO, the sky and polarisation averaged response is $F = \sqrt{1/5} \approx 0.447$.

A detector may consist of several interferometers. Let F_a be the averaged response of the a-th interferometer, the average response of a network of k detectors is obtained by adding in quadrature,

$$F^2_{\mathrm{Total}} = \frac{1}{k} \sum_{a=1}^{k} F_a^2 . \tag{2.60}$$

The averaging in Eq. 2.59 assumes a uniform distribution of polarisation angles ψ. This is the case for a stochastic background; however, for a non-inspiralling circular binary, the polarisation is a function of the two spherical polar angles (ι, ξ) specifying the orientation of the binary's orbital angular momentum. Here, ι is the inclination angle, the polar angle between the orbital angular momentum and the line joining the source to the detector $(-\hat{r})$ and ξ is the azimuthal angle around the same line. In this case, to characterise the detector sensitivity we average over all four angles $(\theta, \phi, \iota, \xi)$. If the binary is inspiralling, then the polarisation depends on still more parameters which need to be averaged over. These more complicated averages all have the property that they depend on both the detector and the source, hence they are unhelpful for our present purpose separating the source amplitude from the detector sensitivity. Additionally, the different averages do not work out to be so different from each other: Finn and Chernoff (1993) calculated the sensitivity for a detector with the LIGO geometry averaged over the four angles $(\theta, \phi, \iota, \xi)$ as $\sqrt{4/25} = 0.4$ times peak sensitivity, which should be compared with the value $\sqrt{1/5} \approx 0.447$ above. The effect of replacing the true sensitivity with the sky-averaged sensitivity for the LISA detector was considered in detail by Vallisneri and Galley (2012); they also found there is a small difference when considering an entire population of sources. For the remainder of this chapter the three-angle average defined in Eq. 2.59 is used.

2.4.2 Ground-based detectors: kHz

Ground-based detectors are the most numerous. A collection of interferometric detectors are listed in Tab. 2.1, these are sensitive to GWs in the frequency range $\mathcal{O}(10\text{–}10^3)$ Hz. They all simulate free-floating test masses by suspending a mass from a pendulum system with natural frequency far removed from than that of the GW. Their sensitivity curves include narrow lines that arise from noise sources in the instrument, including resonances in the suspension system and electrical noise at multiples of 60 Hz: these have been removed from the plots in Chapter 4 for clarity. The detectors fall broadly into three categories: first-generation detectors, which have already operated; second-generation detectors currently under construction; and third-generation detectors at the planning stage.

The most notable ground-based detectors are LIGO and Virgo, which work in collaboration, supported by GEO600. LIGO, Virgo and GEO600 have completed science

runs as first-generation detectors (e.g., Abadie et al., 2010*a*). As LIGO and Virgo are currently being upgraded their initial configurations are now referred to as Initial LIGO (iLIGO) and Initial Virgo (iVirgo) respectively. The upgraded, second-generation versions are referred to as Advanced LIGO (aLIGO) and Advanced Virgo (AdV) respectively. LIGO has two observatories: one at Hanford, Washington, which has two detectors; and another at Livingston, Louisiana. There is an agreement to move one of the upgraded Hanford detector systems to a location in India (Iyer et al., 2011; Unnikrishnan, 2013). The GEO600 detector is not subject to a major upgrade plan, but since summer 2009 it has been enhanced by a series of smaller improvements, notably improving high-frequency sensitivity (GEO-HF; Willke et al., 2006). The advanced detectors should start operation in the next couple of years, with LIGO-India following further in the future.

TAMA300 is a Japanese first-generation detector. Its successor, currently under construction, is the Kamioka Gravitational Wave Detector (KAGRA), formerly the Large-scale Cryogenic Gravitational wave Telescope (LCGT), which is located underground in the Kamioka mine. It employs more sophisticated noise-reduction techniques than LIGO or Virgo, such as cryogenic cooling.

The Einstein Telescope (ET) is an ambitious proposal to construct an underground third-generation detector. Its location would provide shielding from seismic noise, allowing it to observe frequencies of (10–10^4) Hz.

Table 2.1: Summary of ground-based laser interferometers.

Detector	Country	Arm length	Approximate date	Generation
GEO600[a]	Germany	600 m	2001–present	First
TAMA300[b]	Japan	300 m	1995–2003	First
iLIGO[c]	USA	4 km	2004–2010	First
Virgo[d]	Italy	3 km	2007–2010	First
aLIGO[e]	USA	4 km	2015–present	Second
AdV[f]	Italy	3 km	*est.* 2016	Second
KAGRA[g]	Japan	3 km	*est.* 2018	Second
ET[h]	—	10 km	*est.* 2025	Third

[a]Grote (2010), [b]Ando (2002), [c]Abbott et al. (2009), [d]Accadia et al. (2012), [e]Harry (2010), [e]The LIGO Scientific Collaboration and The Virgo Collaboration (2016*b*), [f]Acernese et al. (2009), [g]Somiya (2012), [h]Hild et al. (2011).

We use an interpolation to the data published on `https://wwwcascina.virgo.infn.it/advirgo/` (2013) for the AdV sensitivity curve, an interpolation to the data for version D of the KAGRA detector published on `http://gwcenter.icrr.u-tokyo.ac.jp/en/researcher/parameter` (2013) and analytic fits to the sensitivity curves from Sathyaprakash and Schutz (2009) for the remaining detectors.

2.4.3 Space-based detectors: mHz

Space-based detectors work on similar principles to ground based detectors, but with the test masses residing inside of independent, widely separated satellites. Space-based detectors are sensitive to lower frequency GWs than their ground-based counterparts; this is partly because space-based detectors can have much longer arms, and partly because they are unaffected by seismic noise which limits the low frequency performance of ground-based detectors.

The canonical design for a space-based detector is the Laser Interferometer Space Antenna (LISA), which is sensitive to millihertz GWs. LISA would consist of three satellites flying in a triangular constellation with arms of length 5×10^9 m in a 1 AU orbit around the Sun, trailing the Earth by 20°. The laser arms in a LISA-like detector are not a cavity, the light only travels once along each arm. eLISA is a rescoped version of LISA designed to probe the same frequency range, while proposals such as the Advanced Laser

Interferometer Antenna (ALIA), Big Bang Observer (BBO) and Deci-hertz Interferometer Gravitational wave Observatory (DECIGO) are designed to probe decihertz GWs.

2.4.3.1 LISA and eLISA

The instrumental noise curves for LISA are approximated by the analytic fit given by Sathyaprakash and Schutz (2009), which we use for the plots in Chapter 4. When observing individual sources with LISA there is an additional contribution to the noise from a background of unresolvable binaries. This is not included here as we consider the background as a source of GWs (see Sec. 3.2.2). eLISA is a rescoped version of the classic LISA mission, the main differences are shorter arms (10^9 m instead of 5×10^9 m), two laser arms instead of three, and a different orbit (drifting away from Earth instead of 20° Earth trailing). The effect of these changes is a slightly reduced peak sensitivity and a shift to higher frequencies. We use an analytic fit to the instrumental noise curve given by Amaro-Seoane et al. (2013).

2.4.3.2 DECIGO, ALIA and BBO

These missions are designed to probe the decihertz region of the gravitational wave spectrum; they are considerably more ambitious than the LISA or eLISA mission and their launches will be further into the future. We use a simple analytic fit to the sensitivity curve for ALIA (Bender et al., 2013), while for DECIGO and BBO, fits to the sensitivity curves given by Yagi and Seto (2011) are used.

2.4.4 Using the galaxy as a detector: nHz

There is a current global effort under way to detect very low frequency ($f \approx \mathrm{yr}^{-1}$) gravitational waves (GWs) via the precision timing of a network of galactic millisecond pulsars. These detection efforts exploit the exquisite rotational stability of millisecond pulsars to track any deviations of the pulse time-of-arrivals (TOAs) from deterministic timing-models. A wave propagating between the Earth and a pulsar will induce a perturbation to the space-time metric along the Earth-pulsar line-of-sight, leading to a change in the proper separation, and consequently a shift in the perceived pulsar rotational frequency (Burke, 1975; Detweiler, 1979; Estabrook and Wahlquist, 1975; Sazhin, 1978). Subtracting a deterministic timing-model (which describes the pulsar's spin, spin-down rate, etc.) from the TOAs gives a set of *timing-residuals*, which encode

all unmodeled phenomena, whether they are noise processes or GWs. Utilising a network (or "array") of these pulsars allows us to cross-correlate the data-streams and leverage the fact that GWs will be influencing all pulsars (Foster and Backer, 1990; Hellings and Downs, 1983), whilst intrinsic pulsar noise processes will not. There are three separate pulsar timing array (PTA) efforts underway: the European Pulsar Timing Array (EPTA)[7] (Kramer and Champion, 2013), the Parkes Pulsar Timing Array (PPTA)[8] (Hobbs, 2013) and the North American Nanohertz Observatory for Gravitational waves (NANOGrav)[9] (McLaughlin, 2013). There are also ongoing efforts to combine the techniques and data from all three PTAs within the umbrella consortium of the International Pulsar Timing Array (IPTA)[10] (Manchester and IPTA, 2013).

PTAs can be thought of as naturally occurring interferometers with galactic-scale arm lengths. Accordingly, they are sensitive to much lower frequencies than the detectors previously discussed. Each pulsar is a regular clock and the measured pulse arrival time can be compared against a prediction, leaving a residual which includes the effects of passing GWs. Using an array of these pulsars spread across the sky allows us to correlate residuals between different pulsars, to exploit the fact that GWs influence all pulsars whereas intrinsic pulsar noise does not. The correlation between different pulsars depends only on their angular separation on the sky, and has a distinctive shape, known as the Hellings and Downs curve (Hellings and Downs, 1983). A derivation of this relationship is presented in this section.

The redshift of the rate of arrival of pulses for a pulsar at a distance L from the Solar-System barycentre (SSB), in the direction of the unit spatial vector $\hat{p}$ induced by a gravitational wave travelling in direction of the unit vector $\hat{\Omega}$ is (Anholm et al., 2009)

$$z(t,\hat{\Omega}) = \frac{1}{2}\frac{\hat{p}^j\hat{p}^i}{1+\hat{\Omega}\cdot\hat{p}}\left[h_{ij}^{\text{Pulsar}}\left(t-\frac{L}{c},\hat{\Omega}\right) - h_{ij}^{\text{Earth}}(t,\hat{\Omega})\right] = \frac{1}{2}\frac{\hat{p}^j\hat{p}^i}{1+\hat{\Omega}\cdot\hat{p}}\Delta h_{ij}(t,\hat{\Omega})\ . \tag{2.61}$$

The redshift includes two terms: the pulsar term and the Earth term. The pulsar term is often neglected in PTA analysis as it can be considered as an extra noise term which averages to zero across the array. The experimentally measured quantity is not the

[7] http://www.epta.eu.org/

[8] http://www.atnf.csiro.au/research/pulsar/ppta/

[9] http://nanograv.org/

[10] http://www.ipta4gw.org/

redshift but the timing residual, the two are related via

$$R(t,\hat{\Omega}) = \int_0^t \mathrm{d}t'\; z(t',\hat{\Omega}) \; . \tag{2.62}$$

All of the pulsars, and the Earth, are subject to the same metric-perturbation field. This may be expressed in terms of its Fourier transform

$$h_{ij}(t,\vec{r}) = \sum_{A=+,\times} \int \mathrm{d}f \iint_{\mathbb{S}_2} \mathrm{d}\hat{\Omega}\; \tilde{h}_A(f,\hat{\Omega}) e^A_{ij}(\hat{\Omega}) \exp\left[2\pi i f\left(t - \frac{\hat{\Omega}\cdot\vec{x}}{c}\right)\right] \; , \tag{2.63}$$

where $e^A_{ij}(\hat{\Omega})$ is the A polarisation basis tensor for direction $\hat{\Omega}$, and $\vec{r}$ is the spatial position. Choosing the SSB as the origin of our coordinate system, so the pulsar is at position $L\hat{p}$, gives

$$\Delta h_{ij}(t,\hat{\Omega}) = \sum_{A=+,\times} \int \mathrm{d}f\; \tilde{h}_A(f,\hat{\Omega}) e^A_{ij}(\hat{\Omega}) \exp\left(2\pi \mathrm{i} f t\right) \left\{\exp\left[-2\pi \mathrm{i} f L\left(1+\hat{p}\cdot\hat{\Omega}\right)\right] - 1\right\} \; . \tag{2.64}$$

From Eq. 2.61 and Eq. 2.64, the Fourier transform of the redshift $\tilde{z}(f,\hat{\Omega})$ can be identified as

$$\tilde{z}(f,\hat{\Omega}) = \left\{\exp\left[-2\pi \mathrm{i} f L\left(1+\hat{p}\cdot\hat{\Omega}\right)\right] - 1\right\} \sum_{A=+,\times} \tilde{h}_A(f,\hat{\Omega}) F^A(\hat{\Omega}) \; , \tag{2.65}$$

where

$$F^A(\hat{\Omega}) = \frac{e^A_{ij}(\hat{\Omega})\hat{p}^j\hat{p}^i}{2\left(1+\hat{\Omega}\cdot\hat{p}\right)} \; . \tag{2.66}$$

The function $F^A(\hat{\Omega})$ may be regarded as the PTA equivalent of the detector response functions defined in general in Eq. 2.2 and plotted for a resonant bar and an interferometer in Figs. 2.1 and 2.2. The stochastic background of GWs is fully characterised by the one-sided PSD via the expectation value

$$\left\langle \tilde{h}^*_A(f,\hat{\Omega}) \tilde{h}_{A'}(f',\hat{\Omega}') \right\rangle = \frac{1}{2} S_h(f) \delta^{(2)}(\hat{\Omega},\hat{\Omega}') \delta_{AA'} \delta(f-f') \; , \tag{2.67}$$

where $\delta^{(2)}(\hat{\Omega},\hat{\Omega}')$ is the delta-function on the sphere. From Eq. 2.65 and Eq. 2.67, the expectation of the product of signals from two different pulsars in directions $\hat{p}_1$ and $\hat{p}_2$ may be evaluated as

$$\left\langle \tilde{z}_1(f) \tilde{z}^*_2(f') \right\rangle = \frac{1}{2} S_h(f) \delta(f-f') \Gamma(f) \; , \tag{2.68}$$

where

$$\begin{aligned} \Gamma(f) \;\; = \sum_{A=+,\times} \;\; & \iint_{\mathbb{S}^2} \mathrm{d}\hat{\Omega}\; \left\{\exp\left[2\pi \mathrm{i} f L_1\left(1+\hat{\Omega}\cdot\hat{p}_1\right)\right] - 1\right\} \\ & \times \left\{\exp\left[-2\pi \mathrm{i} f L_2\left(1+\hat{\Omega}\cdot\hat{p}_2\right)\right] - 1\right\} F^A_1(\hat{\Omega}) F^A_2(\hat{\Omega}) \; . \end{aligned} \tag{2.69}$$

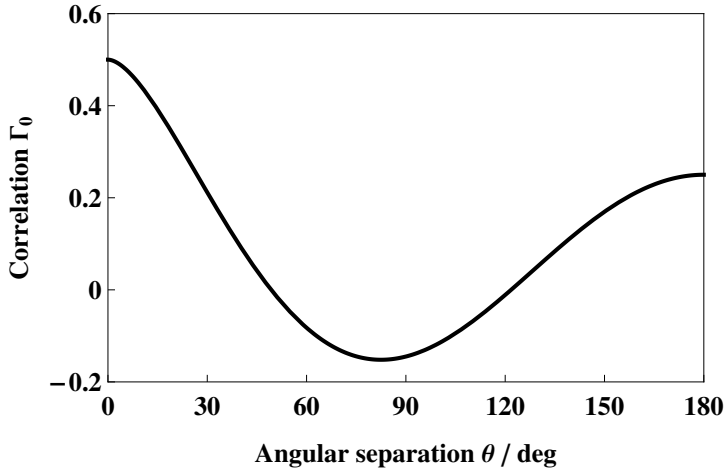


Figure 2.3: The Hellings and Downs (1983) curve, the correlation between the timing residuals caused by gravitational waves of two pulsars separated on the sky by an angle θ.

The overlap function $\Gamma(f)$ tends to a constant value in the limit that the distances to the pulsars are large compared to the wavelength of GWs; PTAs operate in this limit (Mingarelli and Sidery, 2014), so the overlap may be approximated as a constant,

$$\Gamma(f) \approx \Gamma_0 = \sum_{A=+,\times} \iint_{\mathbb{S}^2} \mathrm{d}\hat{\Omega}\; F_1^A(\hat{\Omega}) F_2^A(\hat{\Omega}) \; . \tag{2.70}$$

Neglecting the exponential terms in the overlap is the frequency-domain equivalent of neglecting the pulsar term in Eq. 2.61. The integral may be evaluated to give an expression depending only on the angle θ between the two pulsars; this is the famous Hellings and Downs curve, shown in Fig. 2.3,

$$\Gamma_0 = \frac{1}{2} + \frac{3\xi}{2}\left(\ln\xi - \frac{1}{6}\right) \; , \tag{2.71}$$

where $\xi = (1 - \cos\theta)/2$.

There is a discrepancy between the treatment of PTA sensitivity curves here and the higher frequency detectors discussed in Secs. 2.4.2 and 2.4.3. When observing a long-lived source, such as an inspiral, with a high frequency detector, the convention was to define a *characteristic* strain to satisfy Eq. 2.20. Here, the convention is to leave the strain untouched and instead adjust the PTA sensitivity curve with observation time, again to satisfy Eq. 2.20. This discrepancy is an unfortunate result of the conventions in use by the different gravitational wave communities; however, it is natural given

the sources under observation. When observing a transient source, such as a burst or inspiral, which changes within the lifetime of the detector, it is natural to consider the detector as performing constantly while the signal changes. However, when observing a monochromatic source or a stochastic background, which is unchanging over the detector lifetime, it is more natural to consider the source as being fixed and the sensitivity of the detector gradually improving. All that is required by the definition in Eq. 2.20 is that the ratio $h_{\rm c}(f)/h_n(f)$ is constant.

The sensitivity of a given PTA to a particular source of gravitational waves depends upon the number of pulsars in the array and the noise in their timing residuals. In the following sections the sensitivity to a plane monochromatic gravitational wave is derived for a simplified model of a PTA using both a frequentist and a Bayesian approach. For analogous derivations of the sensitivity of a PTA to a stochastic background of gravitational waves see Moore, Taylor and Gair (2015).

2.4.4.1 Frequentist detection of a monochromatic source

The frequentist method of detection involves defining a detection statistic $\mathcal{S}$. The SNR of this statistic is defined as the expectation in the presence of a signal divided by the root mean square (rms) value in the absence of a signal. A detection is claimed if the SNR in a particular realisation of the experiment exceeds a predetermined threshold value $\varrho_{\rm th}$.

The noise in the timing-residuals is assumed to be white, Gaussian, and uncorrelated between each pulsar. Let $s_x(t)$ be the smooth function from which the discretely sampled timing risiduals in pulsar x are drawn. The real data from a PTA contain noise and a signal

$$\mathbf{s}(t) = \mathbf{n}(t) + \mathbf{h}(t) \ , \quad \text{where } \mathbf{s}(t)^{\rm T} = \left(s_1(t), s_2(t), \ldots, s_{N_p}(t)\right) \ . \tag{2.72}$$

The noise satisfies $\left\langle \tilde{n}_x(f)\tilde{n}_y^*(f')\right\rangle = (1/2)\delta(f-f')\delta_{xy}S_{n,x}$, where $S_{n,x} = 2\sigma_x^2\delta t_x$, $1/\delta t_x$ and σ_x are the cadence and the rms error in the timing-residuals in pulsar x. Correlating the residuals with a symmetric filter matrix $\mathbf{K}(t)$ defines a statistic and associated SNR

$$\mathcal{S} = \int {\rm d}t\, \mathbf{K}^\dagger(t)\mathbf{s}(t) = \int {\rm d}f\, \tilde{\mathbf{K}}^\dagger(f)\tilde{\mathbf{s}}(f) \ , \tag{2.73}$$

$$\varrho^2 = \frac{\mu^2}{\sigma^2} = \frac{\langle \mathcal{S}\rangle^2_{s=h+n}}{\langle \mathcal{S}^2\rangle_{s=n} - \langle \mathcal{S}\rangle^2_{s=n}} \ . \tag{2.74}$$

Using the identical Gaussian properties of the noise in each pulsar it is straightforward to show that the expectation value of $\mathcal{S}$ in the presence of a signal and the variance of $\mathcal{S}$

in the absence of a signal are given respectively by

$$\mu = \int \mathrm{d}f\, \tilde{\mathbf{K}}^{\dagger}(f)\tilde{\mathbf{h}}(f)\,, \tag{2.75}$$

$$\sigma^2 = \int \mathrm{d}f\, \frac{S_n}{2}\tilde{\mathbf{K}}^{\dagger}(f)\tilde{\mathbf{K}}(f)\,. \tag{2.76}$$

The SNR, Eq. (2.74), is then given by the following inner product

$$\varrho^2 = \frac{\left(\frac{2\tilde{\mathbf{h}}(f)}{S_n}|\tilde{\mathbf{K}}(f)\right)^2}{\left(\tilde{\mathbf{K}}(f)|\tilde{\mathbf{K}}(f)\right)}\,, \quad \text{where } \left(\tilde{\mathbf{A}}|\tilde{\mathbf{B}}\right) = \int \mathrm{d}f\, \frac{S_n}{2}\tilde{\mathbf{A}}^{\dagger}(f)\tilde{\mathbf{B}}(f)\,. \tag{2.77}$$

The optimal filter is the matrix $\mathbf{K}(t)$ which maximises the SNR in Eq. (2.77). It follows from the Cauchy-Schwarz inequality that the optimal filter (up to an arbitrary factor) and its corresponding SNR are given by

$$\tilde{\mathbf{K}}(f) = \frac{2\tilde{\mathbf{h}}(f)}{S_n}\,, \quad \text{where } \varrho^2 = \sum_x 4\int_0^{\infty} \mathrm{d}f\, \frac{\left|\tilde{h}_x(f)\right|^2}{S_n}\,. \tag{2.78}$$

In order to evaluate the optimal, or "matched", filter $\tilde{\mathbf{K}}(f)$, the waveform $\tilde{\mathbf{h}}(f)$ must be known to sufficient accuracy. A monochromatic source is a simple example of a waveform which can be extracted using matched filtering. This analysis is different to that performed in Moore, Taylor and Gair (2015); in that paper a cross correlation analysis was adopted as it was desired to compare the analysis with that used for a stochastic background search, in that case one is forced to used a cross-correlation statistic because the stoachastic signal in each pulsar is buried in the pulsar red-noise. The two searches have comparable sensitivities.

From Eq. (2.61) it may be seen that the measured signal is proportional to the gravitational wave amplitude and a geometric factor depending on the sky positions of the pulsars, the sky position of the source and the source orientation. Since there are many pulsars in our PTA they effectively average this distribution over the sky position angles. For simplicity we set the source inclination and polarisation angles to zero, so we are calculating the sky-averaged sensitivity of the PTA to optimally orientated sources. It is straighforward to generalise this treatment to arbitrary source orientations. The sky-averaged value of the geometric factor in Eq. (2.61) is

$$\chi \;= \int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi} \frac{\sin\theta \mathrm{d}\theta\mathrm{d}\phi}{4\pi} \sqrt{\left(\frac{1}{2}\frac{\hat{p}_i\hat{p}_j\left(A^{+}H^{+}_{ij}+A^{\times}H^{\times}_{ij}\right)}{1+\hat{\Omega}\cdot\hat{p}}\right)^2} = \frac{1}{\sqrt{3}}\,, \tag{2.79}$$

where $\hat{p} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, $e^x = (1,0,0)$, $e^y = (0,1,0)$, $\hat{\Omega} = (0,0,1)$, $H_{ij}^{+} = \epsilon_{ij}^{+}\cos 2\psi + \epsilon_{ij}^{\times}\sin 2\psi$, $H_{ij}^{\times} = -\epsilon_{ij}^{+}\sin 2\psi + \epsilon_{ij}^{\times}\cos 2\psi$, $\epsilon_{ij}^{+} = e_i^x e_j^x - e_i^y e_j^y$, $\epsilon_{ij}^{\times} = e_i^x e_j^y + e_i^y e_j^x$, $A^{+} = (1+\cos^2\iota)/2$, $A^{\times} = \cos\iota$ and $\iota = \psi = 0$. Therefore for a monochromatic source of GWs with frequency f_0 the signal is given approximately by $\tilde{h}_x(f) \approx \tilde{h}_y(f) \approx (\chi h_c/f)\delta(f - f_0)$. Using Eq. (2.78), together with the finite-time delta function, $\delta_T(f) = \sin(\pi f T)/(\pi f)$, gives

$$\varrho^2 = 4N_p\chi^2 h_c^2 \int_{1/T}^{1/\delta t} \mathrm{d}f \, \frac{\delta_T^2(f - f_0)}{f^2 S_n} \, . \tag{2.80}$$

The PTA is sensitive to frequencies in the range $\sim 1/T$ up to the Nyquist frequency. Imposing a threshold for detection, $\varrho = \varrho_{\rm th}$ and rearranging gives h_c as a function of f_0. This is the desired sensitivity curve, and is shown as the red curve in the left panel of Fig. 2.5. The fact that the sensitivity curve tends to a constant value at low frequencies is obviously incorrect. The reason for this is that the loss of sensitivity which arises from fitting a deterministic timing-model to the raw TOAs has not been accounted for. Or, in Bayesian language, Eq. (2.80) assumes delta function priors on all of the pulsar timing model parameters.

Some insight into the shape of the sensitivity curve and the loss of sensitivity due to fitting for the pulsar timing-model may be gained by considering the inner product in the time domain. In Eq. (2.80), if the power of 4 were replaced by a power of 2 the fact that the noise is white would allow us to use Parseval's thoerem to change from a frequency integral to a time integral. By analogy, from Eq. (2.80) the SNR may be written approximately as

$$\varrho^2 \approx 2N_p\chi^2 h_c^2 \int_0^T \mathrm{d}t \, \frac{\sin^2(2\pi f t + \phi)}{\sigma^2 f^2 \delta t} \, . \tag{2.81}$$

For our PTA, at frequencies of $\approx 1\,\mathrm{yr}^{-1}$, this approximation holds to better than 10%. In the high frequency limit ($ft \gg 1$) the integral $\int \sin^4(2\pi f t + \phi)\,\mathrm{d}t \approx T/2$ and the sensitivity tends to

$$h_c^{\rm HIGH}(f) \approx \frac{\rho\sigma f}{\chi\sqrt{N_p}}\sqrt{\frac{\delta t}{T}} \, . \tag{2.82}$$

In the low frequency limit ($ft \ll 1$) the sine may be expanded as a power series

$$\sin(2\pi f t + \phi) \approx \sin(\phi) + 2\pi t f\cos(\phi) - 2\pi^2 f^2 t^2 \sin(\phi) - \frac{4}{3}\pi^3 f^3 t^3 \cos(\phi) + \mathcal{O}\left(f^4 t^4\right) \, . \tag{2.83}$$

The first term in this expansion is a constant offset from zero in the timing residuals; this type of signal is degenerate with the distance to the pulsar and hence will be "fit-out" of the data. Similarly the second term is a linear drift in the residuals and is degenerate with the pulsar spin (and also the line of sight component of the peculiar velocity). Finally the third term is degenerate with the pulsar spin-down rate. Higher order derivatives of the spin period do not need to be fit for independently as they are determined uniquely by the period and its first time derivative. Therefore at low frequencies the leading contribution, at $\mathcal{O}\left(f^3 t^3\right)$, gives

$$h_c^{\mathrm{LOW}}(f) \approx \sqrt{\frac{63\rho_{\mathrm{th}}^2}{32\pi^6}} \frac{1}{\chi\sqrt{N_p}} \sigma f^{-2} \sqrt{\frac{\delta t}{T}} T^{-3} \sec\phi \; . \tag{2.84}$$

The total sensitivity may be approximated by $h_c^{\mathrm{LOW}} + h_c^{\mathrm{HIGH}}$, i.e. as a two part power-law in f. This is shown as the blue curve in the left panel of Fig. 2.5, where the value of ϕ has been chosen to make $h_c^{\mathrm{LOW}} = h_c^{\mathrm{HIGH}}$ at a frequency of $2/T$. A threshold of SNR $\varrho_{\mathrm{th}} = 3$ has been used. Despite the apparent simplicity of this two part power law model it shows excellent agreement with the following Bayesian treatment.

2.4.4.2 Bayesian detection of a monochromatic source

One advantage of the Bayesian approach is that it provides a well motivated method for accounting for the loss of sensitivity due to fitting for the pulsar timing-model, as opposed to the rather *ad hoc* subtraction of a few terms from a Taylor series performed above.

In the Bayesian approach two competing hypotheses are considered: the noise and signal hypotheses. The noise hypothesis, $\mathcal{H}_n$, is that the data contain only contributions from noise and the timing-model while the signal hypothesis, $\mathcal{H}_h$, is that the data contain noise, timing-model and a signal. For each hypothesis, $i \in \{h, n\}$, the evidence may be calculated as

$$\mathcal{O}_i(\mathbf{s}) = \int \mathrm{d}\vec{\lambda}\; \mathcal{L}_i(\mathbf{s}, \vec{\lambda}_i) P_i(\vec{\lambda}_i) \; , \tag{2.85}$$

where $\vec{\lambda}_i$ is the vector of free parameters, $\mathcal{L}_i$ is the likelihood function, $\mathbf{s}$ is the measured data and P_i is the prior function, for hypothesis $\mathcal{H}_i$. From here on the dependence on the data is suppressed in our notation for compactness. The Bayes factor is then defined as the ratio of these evidences, and a detection is claimed if this exceeds some pre-determined threshold, $\mathcal{B} \equiv \mathcal{O}_h/\mathcal{O}_n > \mathcal{B}_{\mathrm{th}}$. A value of $\mathcal{B}_{\mathrm{th}} = 1000$ was used, this was

choosen to give roughly the same false alarm rate at the value of $\varrho_{\rm th} = 3$ used in Sec. 2.4.4.1.

In this section the physical signal, $h_x(t)$, is assumed to be that of a monochromatic source, and a quadratic timing model, $m_x(\vec{\Theta}_x, t)$, for each pulsar is assumed, with pulsar parameters $\vec{\Theta}_x$. In reality the timing model is more complex than a simple quadratic as it has to account for several effects, such as the pulsar's position, dispersion in the interstellar medium, peculiar motion, and orbital motion if the pulsar happens to be in a binary system (see Edwards et al., 2006; Hobbs et al., 2006, and references therein). However, a simple quadratic model serves here to illustrate the loss of sensitivity due to fitting for the distance to the pulsar, the pulsar spin and the pulsar spindown. The quadratic model also has the nice property that the pulsar timing-model parameters can be marginalised over analytically,

$$\begin{aligned}
\mathbf{s} &= (\overbrace{\underbrace{s_1(\delta t), s_1(2\delta t), \ldots, s_1(T)}_{T/\delta t}, \ldots, s_{N_p}(\delta t), s_{N_p}(2\delta t), \ldots, s_{N_p}(T)}^{N_p T/\delta t}) \,, \\
h_x(\vec{\Psi}, t) &= \tfrac{\chi h_c}{f} \sin\left(2\pi f t + \phi\right) \,, \quad \text{with source parameters } \vec{\Psi}^{\rm T} = \{h_c, f, \phi\} \,, \\
m_x(\vec{\Theta}_x, t) &= \vec{\Theta}_x^{\rm T} \cdot \vec{N} \,, \qquad\qquad \text{with } \vec{N}^{\rm T} = \{1, t, t^2\} \text{ and } \vec{\Theta}_x^{\rm T} = \{\alpha_x, \beta_x, \gamma_x\} \,.
\end{aligned} \tag{2.86}$$

where α_x is a constant phase offset, β_x is proportional to the pulsar's rotational frequency or peculiar velocity, and γ_x is proportional to its spindown rate or acceleration. Since the noise is Gaussian the log-likelihood for the noise hypothesis is given by

$$\log \mathcal{L}_n(\vec{\Theta}) = \log A - \frac{1}{2} \left(\mathbf{s} - \mathbf{m}(\vec{\Theta})\right)^{\rm T} \mathbf{\Sigma}_n^{-1} \left(\mathbf{s} - \mathbf{m}(\vec{\Theta})\right) \,, \tag{2.87}$$

where the covariance matrix is simply the scaled identity, $\mathbf{\Sigma}_n = \sigma^2 \mathbf{I}_{N_p T/\delta t}$, and A is a constant, absorbing determinant factors. Similarly the log-likelihood for the signal hypothesis is given by

$$\log \mathcal{L}_h(\vec{\Theta}, \vec{\Psi}) = \log A - \frac{1}{2} \left(\mathbf{s} - \mathbf{m}(\vec{\Theta}) - \mathbf{h}(\vec{\Psi})\right)^{\rm T} \mathbf{\Sigma}_n^{-1} \left(\mathbf{s} - \mathbf{m}(\vec{\Theta}) - \mathbf{h}(\vec{\Psi})\right) \,. \tag{2.88}$$

In both the noise and signal hypotheses identical, uniform priors on the timing-model parameters were assumed. In the signal hypothesis case there are also the priors on the source parameters to consider. Since we have adopted a very stringent detection threshold (Bayes factor of 1000) it is reasonable to expect the posterior to be strongly peaked at the true values independent of any (reasonable) prior used. Of course the data must be

used to find the position of this peak and this process will exhaust a certain amount of information in the data reducing the final evidence value. However, this reduction may be neglected in the limit of large final evidence. Numerically this approximation is equivalent to taking a delta-function prior on the source parameters positioned at the correct values, $P(\vec{\Psi}, \vec{\Theta}_x) \propto \delta^{(3)}(\vec{\Psi} - \vec{\Psi}')$, however it should be stressed that this is an analytic trick used to implement the approximation described and in practice the data will still be used to find the maximum in the posterior. (It may be the case that a very localised prior is used on the source parameters if a clear electromagnetic counterpart has been identified.)

From Eq. 2.86 it can be seen that the timing model is linear in the pulsar parameters, so the timing model may be expressed by a $(N_p T/\delta t) \times (3N_p)$ matrix, $\mathbf{M}$, known as the design matrix. This non-square matrix admits the usual unique singular value decomposition into the $(N_p T/\delta t) \times (N_p T/\delta t)$ matrix $\mathbf{U}$, the $(N_p T/\delta t) \times (3N_p)$ matrix $\mathbf{S}$ and the $(3N_p) \times (3N_p)$ matrix $\mathbf{V}$. The matrix $\mathbf{U}$ may be further uniquely decomposed into $\mathbf{F}$ and $\mathbf{G}$ where $\mathbf{G}$ is an $(N_p T/\delta t) \times (N_p T/\delta t - 3N_p)$ matrix.

$$\mathbf{m} = \mathbf{M}\vec{\Theta} \ , \quad \mathbf{M} = \mathbf{U}\mathbf{S}\mathbf{V}^{\dagger} \ , \quad \mathbf{U} = (\mathbf{F}, \mathbf{G}) \ . \tag{2.89}$$

The evidence integral for the noise hypothesis is a multivariate Gaussian in the pulsar timing-model parameters, $\vec{\Theta}_x$. This may be evaluated analytically (van Haasteren and Levin, 2013), and can be viewed as a projection of the data into the left null space of the design matrix.

$$\mathcal{O}_n = \int \mathrm{d}\vec{\Theta} \ \mathcal{L}_n(\vec{\Theta}) = \frac{\exp\left(-\frac{1}{2}\mathbf{s}^{\mathrm{T}}\mathbf{G}\left(\mathbf{G}^{\mathrm{T}}\boldsymbol{\Sigma}_n\mathbf{G}\right)^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{s}\right)}{\sqrt{(2\pi)^{n-m}\det\left(\mathbf{G}^T\boldsymbol{\Sigma}_n\mathbf{G}\right)}} \tag{2.90}$$

The evidence for the signal hypothesis may be similarly calculated by evaluating the following integral,

$$\mathcal{O}_h = \int \mathrm{d}\vec{\Psi} \ \delta(\vec{\Psi} - \vec{\Psi}') \int \mathrm{d}\vec{\Theta} \ \mathcal{L}_h(\vec{\Theta}_x, \vec{\Psi}). \tag{2.91}$$

However, as can be seen by comparing Eqs. (2.87) and (2.88), this is identical to the result in Eq. (2.90) with the transformation $\mathbf{s} \to \mathbf{s} - \mathbf{h}(\vec{\Psi}') \equiv \mathbf{s} - \mathbf{h}'$. The Bayes factor, defined as the ratio of the two evidences times the prior odds ratio, is then given by the following (where the prior odds has been set to unity),

$$\mathcal{B} \equiv \frac{\mathcal{O}_h}{\mathcal{O}_n} = \exp\left(-\frac{1}{2}\mathbf{h}'^{\mathrm{T}}\mathbf{G}\left(\mathbf{G}^{\mathrm{T}}\boldsymbol{\Sigma}_n\mathbf{G}\right)^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{h}' + \frac{1}{2}\mathbf{s}^{\mathrm{T}}\mathbf{G}\left(\mathbf{G}^{\mathrm{T}}\boldsymbol{\Sigma}_n\mathbf{G}\right)^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{s}\right) . \tag{2.92}$$

The measured data is given by $\mathbf{s} = \mathbf{h}' + \mathbf{m} + \mathbf{n}$. We also have by definition of the projection matrix $\mathbf{G}^{\mathrm{T}}\mathbf{m} = 0$. Averaging the Bayes factor over many realisations of Gaussian noise with gives the expectation value of the Bayes factor as $\overline{\mathcal{B}}$.

$$P(\mathbf{n})\mathrm{d}\mathbf{n} = \frac{\exp\left(-\frac{1}{2}\mathbf{n}^{\mathrm{T}}\boldsymbol{\Sigma}_n^{-1}\mathbf{n}\right)}{\sqrt{(2\pi)^{N_p T/\delta t}\det\left(\boldsymbol{\Sigma}_n\right)}}\,\mathrm{d}\mathbf{n}\,, \tag{2.93}$$

$$\overline{\mathcal{B}} = \int \mathrm{d}\mathbf{n}\; P(\mathbf{n})\mathcal{B} = \exp\left(\left(\mathbf{G}^{\mathrm{T}}\mathbf{h}'\right)^{\mathrm{T}}\left(\mathbf{G}^{\mathrm{T}}\boldsymbol{\Sigma}_n\mathbf{G}\right)^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{h}'\right)\,. \tag{2.94}$$

Hence the expected value of the Bayes factor is given by inner product of the signal, projected orthogonal to the quadratic timing model, with itself. This inner product may be written as integral in the time domain where the physical signal, $h(t)$, is replaced with the projected signal, $(Gh)(t)$. The projection is accomplished by explicitly choosing a basis of three orthogonal functions which span the space of the quadratic timing model.

$$\overline{\mathcal{B}} = \exp\left(N_p \int_0^T \mathrm{d}t\, \tfrac{(Gh)(t)^2}{2\sigma^2\delta t}\right) \quad \text{where,} \tag{2.95}$$

$$(Gh)(t) = h(t) - f_1(t)\frac{\int_0^T \mathrm{d}\tau\, f_1(\tau)h(\tau)}{\int_0^T \mathrm{d}\tau\, f_1(\tau)^2} - f_2(t)\frac{\int_0^T \mathrm{d}\tau\, f_2(\tau)h(\tau)}{\int_0^T \mathrm{d}\tau\, f_2(\tau)^2} - f_3(\tau)\frac{\int_0^T \mathrm{d}\tau\, f_3(t)h(\tau)}{\int_0^T \mathrm{d}\tau\, f_3(\tau)}$$

$$f_1(\tau) = \tfrac{\tau^2}{T} - \tau + \tfrac{T}{6}, \quad f_2(\tau) = \tau - \tfrac{T}{2} \quad \text{and} \quad f_1(\tau) = T.$$

Setting $\overline{\mathcal{B}} = \mathcal{B}_{\mathrm{th}}$ in Eq. 2.95 and rearranging gives an expression for $h_c(f, \phi)$, which defines the sensitivity curve (this expression is somewhat lengthy, see Appendix A). The final three terms in the expression for $(Gh)(t)$ arose from marginalising over the timing-model, neglecting these terms gives another expression, $H_c(f, \phi)$, which corresponds to the sensitivity without the loss due to a timing-model fit. Both of these are plotted in the right panel of Fig. 2.5 for different values of ϕ. This again illustrates the loss in sensitivity due to the requirement that we fit for free parameters in the timing-model. The black curves in Fig. 2.5 are the phase averaged sensitivities.

2.4.4.3 Numerical results for the sensitivity to a monochromatic source

The shape of the strain sensitivity curve is now reconstructed with numerical simulations of *Earth-term only* monochromatic signal injections. For our canonical PTA, we adopt the 36 pulsar network (see Fig. 2.4) of the first IPTA data challenge, where pulsars are timed fortnightly to 100 ns precision over a total baseline of 5 years. Injections are performed using the PALSimulation code, which is part of the PAL package[11] being

[11] `https://github.com/jellis18/PAL`

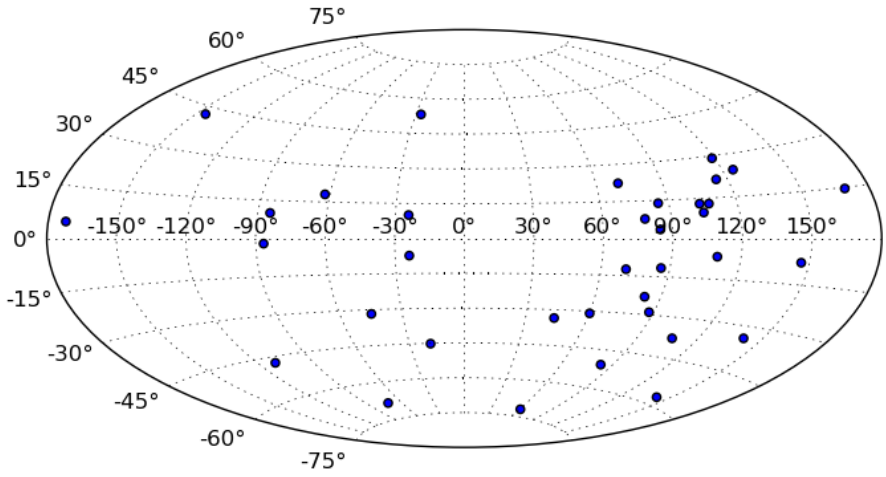


Figure 2.4: The sky positions, shown in gallactic coordinates using a Mollweide projection, of the 36 pulsars used in the first IPTA mock data challenge; `http://www.ipta4gw.org/?page_id=89`.

developed as a unifying suite of tools for pulsar timing analysis. These injections provide a set of simulated timing files, which along with associated pulsar parameter files, can be processed with the TEMPO2 pulsar-timing package (Edwards et al., 2006; Hobbs et al., 2006, 2009). The output from the TEMPO2 timing-model fit is a set of timing-residuals, and the design-matrix which describes the contribution of the deterministic timing-model parameters to each TOA observation. The likelihood model was analytically marginalised over uniform-prior timing-model parameters by projecting all quantities into the left null-space of the design matrix, equivalent to a linear operation on the timing-residuals and noise matrices (van Haasteren and Levin, 2013).

For a grid of PTA-band frequencies and GW-source distances, a systematic injection and recovery of varying SNR Earth-term only signals was performed. Searches were over the 7-dimensional parameter space of $\{\zeta, f, \theta, \phi, \iota, \psi, \phi_0\}$, where $\zeta = \mathcal{M}^{5/3}/D_L$ is a dimensionless strain-amplitude defined in terms of the binary chirp mass, $\mathcal{M}$, and luminosity distance, D_L; f is the gravitational wave frequency; (θ, ϕ) denote the sky-location of the source in spherical-polar coordinates; ι is the orbital inclination angle; ψ is the wave polarisation angle; and ϕ_0 is an initial orbital phase parameter. The angles ϕ_0, ψ and ι for the source were set to be zero; the sky position angles were set as $\phi = 1$ and $\theta = \pi/2 - 0.5$. For the sky positions of the 36 pulsars in the PTA the root mean square value of the geometric factor in the integrand of Eq. 2.61 is $\langle\chi\rangle = 0.51$. This is

in reasonable agreement with the expected value calculated in Eq. 2.79 of $1/\sqrt{3} \approx 0.58$. The chirp mass was set as $\mathcal{M} = 10^7 M_\odot$; and the luminosity distance was varried between 10^{-5} Mpc and 10 Mpc. This choice of source parameters ensures that the "chirping" timescale of the binary due to orbital shrinkage by GW-emission is much longer than the baseline of 5 yrs, whilst the range of distances scales the SNR of the injection from the regime of being completely undetectable to easily detectable.

Parameter estimation and evidence recovery are performed using the Bayesian inference package MultiNest (Feroz and Hobson, 2008; Feroz et al., 2009, 2013). The collection of recovered Bayesian evidence values were interpolated at each injected frequency to determine the characteristic strain-amplitude at which we exceed the predetermined detection threshold. The Bayes factor surface is shown in the left panel of Fig. 2.6, along with our numerically deduced Bayesian sensitivity curve in the right panel of Fig. 2.6. Comfortable qualitative agreement can be seen with the results of the simple frequentist/Bayesian analytic techniques shown in Fig. 2.5.

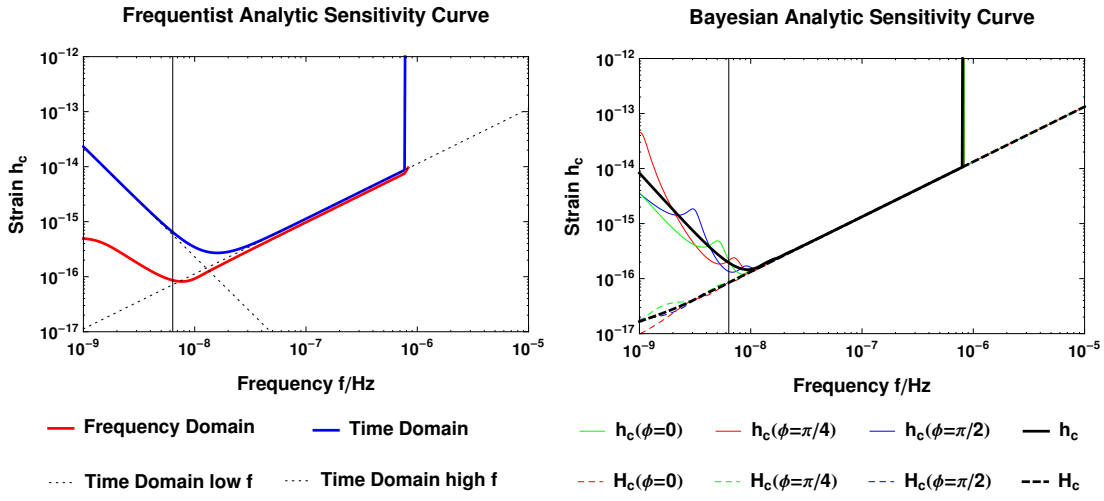


Figure 2.5: Sensitivity curves for the PTA discussed in the text to a monochromatic source. The left panel shows the prediction of the frequentist formula in Sec. 2.4.4.1, the right panel shows the prediction of the Bayesian formula in Sec. 2.4.4.2.

Here the sensitivity curve of a canonical PTA roughly equivalent to mock dataset Open1 in the recent IPTA data challenge has been calculated for a monochromatic wave.

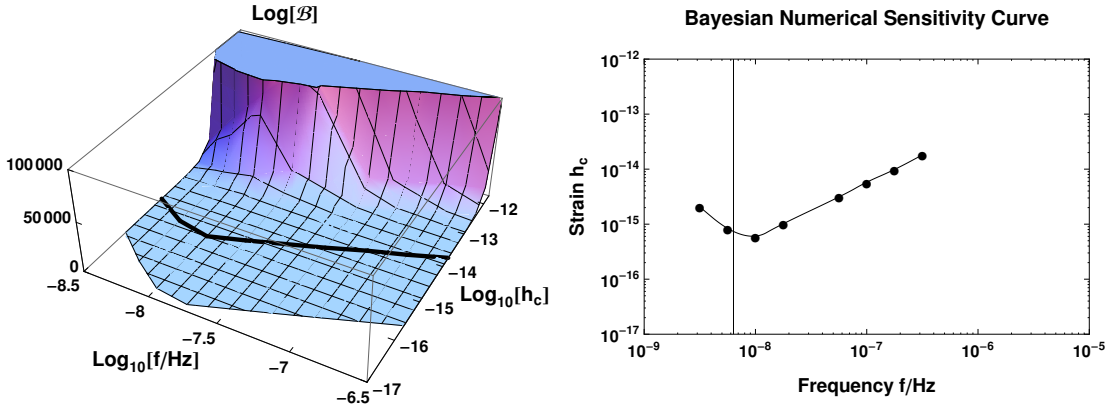


Figure 2.6: The left panel shows a plot of $\log(\mathcal{B})$ against amplitude and frequency. The black line is the contour $\mathcal{B} = \mathcal{B}_{\rm th}$. The black curve is identical to that plotted in the right panel which shows the numerically calculated sensitivity curve in Sec. 2.4.4.3.

These calculations have been performed in both Bayesian and frequentist frameworks and using both analytic and numerical techniques. The results show excellent qualitative agreement and good quantitative agreement up to a factor of a few. Additionally we have presented several simple analytic formulae for the sensitivity curves in both the Bayesian and frequentist pictures. Along the way to deriving these sensitivity curves simple analytic formulae for the frequentist signal-to-noise-ratio and the Bayesian evidence have also been derived.

The shape of the sensitivity curve of a PTA is different when analysing a stochastic background, in particular there is a sharp low frequency cutoff around $f \sim 1/T$ (for a detailed discussion see Moore, Taylor and Gair (2015).This difference is not specific to PTAs, it is also present in ground and space-based detectors. However the differences are particularly pronounced in the case of PTAs because they are most sensitive to frequencies $\sim 1/T$, where T is the total baseline observation time, and hence there are only a few complete wave cycles in the data. This is in contrast to, say, LIGO, which has a peak sensitivity of $\sim 100\,\mathrm{Hz}$, so a year's data contains $\sim 10^9$ cycles.

2.4.5 Atom interferometers

A possible alternative approach to using laser interferometry to detect gravitational waves, is to instead use light pulse atom interferometry. The approach works by taking a cloud

of atoms in free fall and applying a laser pulse at time $t = 0$ to place the atoms into two momentum different states. A second laser pulse in the opposite direction is then applied at time $t = T$. At time $t = 2T$, if the atoms have been freely falling in flat space, the atoms should be back in the original configuration. The passage of a gravitational wave in the time interval $0 \geq t \leq 2T$ will introduce a phase difference between the two atomic momentum states, which can in principle be measured when the atoms are recombined.

Ground-based atom interferometers have been proposed to probe the frequency range $(1 - 10)$ Hz and space-based atom interferometers to probe the range $(10^{-3} - 1)$ Hz. For details on these two techniques see, for example, Dimopoulos et al. (2009).

2.5 Using the universe as a detector

Primordial gravitational waves produce tensor fluctuations which could results in a measurable contribution to anisotropies in the cosmic microwave background (CMB) on large angular scales. Therefore observations of the CMB, in particular the B-mode component of the polarisation, provide a way to detect the lowest frequency gravitational waves (Seljak and Zaldarriaga, 1997).

In 2014 it was claimed that the BICEP2 telescope at the south pole had detected the B-modes in the CMB at a level parametrized by the tensor-to-scalar ratio of $r = 0.2^{+0.07}_{-0.05}$ (BICEP2 Collaboration, 2014). If confirmed this would have constituted strong evidence for the theory of inflation and would arguably have constituted the first detection of gravitational waves. However later work combining the BICEP2 data with other data sets (notably data from the Planck satellite) found that the observations were consistent with the presence of cosmic dust (BICEP2/Keck and Planck Collaborations, 2015; Planck Collaboration et al., 2014).

As the focus of the remainder of this thesis is on astrophysical rather than cosmological sources, the plots in Chapter 4 do not include these very low frequency gravitational wave sources and detectors.

Chapter 3

Gravitational wave sources

> The history of astronomy is a history of receding horizons.
>
> Edwin Hubble

All the sources described here are represented by shaded boxes in Chapter 4. Sources with short durations (i.e., burst sources) and sources that evolve in time over much longer timescales than our observations are drawn with flat-topped boxes for $h_c(f)$. Inspiraling binaries, or stochastic backgrounds of binaries, are drawn with a sloping top proportional to $f^{-2/3}$ for $h_c(f)$, which is result derived in Sec. 2.2.2. The width of the box gives the range of frequencies sources of a given type can have while remaining at a detectable amplitude. The question of the height of the box is more problematic; it would be desirable to normalise each box so that there was a fixed event rate, say one event per year with an amplitude lying within the box. However, for many of the sources considered, the event rate is subject to a large degree of uncertainty (and estimates change rapidly as our understanding of the astrophysics improves). Consequently, we take the more definite, but somewhat arbitrary, approach that for each type of source a fiducial event is nominated (with parameters or amplitude detailed in the relevant section below) and the amplitude of this event is used to position the top of the box. The parameters for each fiducial event are chosen such that the resulting amplitude is roughly consistent with the optimistic end of current predictions: the boxes indicate where we *could* plausibly find gravitational wave sources; the actual event rates could turn out to be substantially lower than this upper bound.

3.1 Sources for ground-based detectors

3.1.1 Compact binary coalescences

3.1.1.1 Neutron-star binaries

The late inspiral of a pair of neutron stars is a prime target for ground-based detectors. The expected event rate for this type of source is uncertain, but upper limits of $\gamma_{\mathrm{NS-NS}} = 1.3 \times 10^{-4}\ \mathrm{Mpc^{-3}\,yr^{-1}}$ are known (Abadie et al., 2012*c*), and the predicted rate is several orders of magnitude below this. Plotted in Chapter 4 are boxes labelled "*compact binary inspirals*" with amplitudes such that a ratio $h_{\mathrm{c}}/h_n = 16$ is produced for Advanced LIGO at peak sensitivity and a width between (3–300) Hz, corresponding to expected observable frequencies.

3.1.1.2 Stellar mass binary black holes

The inspiral and merger of a pair of stellar mass black holes in the mass range $(5-100)\,M_{\odot}$ is another prime target for ground-based detectors. Event rate estimates prior to the first operation run of advanced LIGO generally predicted a lower rate for detections of binary black holes than binary neutron stars (Abadie et al., 2010*b*); although there were large uncertainties due in part to the lack of any electromagnetic observations of these sources or their immediate progenitors. As it happened the first detection of gravitational waves was from a binary black hole with component masses $m_1 = 36^{+5}_{-4}\,M_{\odot}$ and $m_2 = 29^{+4}_{-4}\,M_{\odot}$ at a distance of $410^{+160}_{-180}\,\mathrm{Mpc}$ (The LIGO Scientific Collaboration and the Virgo Collaboration, 2016*c*). Plotted in Chapter 4 is a curve corresponding to a binary black hole with GW150914-like parameters.

3.1.2 Supernovae

3.1.2.1 Core collapse supernovae

Simulations of core-collapse supernovae show that GWs between $(10^2$–$10^3)$ Hz can be produced (Kotake et al., 2006). The signal undergoes $\mathcal{O}(1)$ oscillation and is hence burst-like. Dimmelmeier et al. (2002) calculate the average maximum amplitude of GWs for a supernova at distance r as

$$h_{\mathrm{max}} = 8.9 \times 10^{-21} \left(\frac{10\ \mathrm{kpc}}{r} \right) . \tag{3.1}$$

The event rate for supernovae is approximately $\gamma_{\rm SN} = 5 \times 10^{-4}\ {\rm Mpc}^{-3}\,{\rm yr}^{-1}$. The boxes labelled "*core collapse supernova*" plotted in Chapter 4 correspond to a distance $r = 300$ kpc (which includes the entirety of our galaxy, but no other galaxies of comparable size) with the frequency range quoted above. The LIGO and Virgo detectors have already placed bounds on the event rate for these sources (Abadie et al., 2012*a*).

3.1.2.2 Type Ia supernovae

Type IA supernova also are a possible source of gravitational waves, although in a lower frequency range $(0.1 - 10)$ Hz. The lower frequency is set by the light crossing time of the white dwarf, in contrast to the light crossing time of a neutron star in the case of core collapse supernova. The boxes labelled "*Type IA supernova*" plotted in Chapter 4 correspond to an event at a distance of 30 kpc with a gravitational wave signal calculated using the models of Seitenzahl et al. (2015). The event rate for sources this close (i.e. in our galaxy) is relatively low, of the order $0.01\,{\rm yr}^{-1}$.

The frequency of these sources means they fall in between the sensitivity bands of the second generation ground-based detectors and the proposed first generation space-based detectors. Although we include these sources here in the ground-based section, they may be detected by either (or both) third generation ground-based detectors or a second generation space-based detector.

3.1.3 Continuous waves from rotating neutron stars

Rotating neutron stars are a source of continuous GWs if they possess some degree of axial asymmetry (Aasi et al., 2013*a*; Abbott et al., 2007; Prix, 2009). The signals are near monochromatic with a frequency twice the rotation frequency of the neutron star, and are a potential source for ground-based detectors. The amplitude of the GWs depends upon the deformation of the neutron star, and its spin frequency. The magnitude of the distortion depends upon the neutron star equation of state and the stiffness of the crust, which are currently uncertain (Chamel and Haensel, 2008; Lattimer, 2012). Deformations can also be supported by internal magnetic fields (Haskell et al., 2008). Several known pulsars could be sources for the advanced detectors and upper limits from the initial detectors help to constrain the deformations.

The boxes labelled "*pulsars*" plotted in Chapter 4 correspond to the upper limits placed on a gravitational wave signal from the Crab pulsar (Aasi et al., 2014*b*), extrapolated

across a frequency range between (20–10^3) Hz. The extrapolation was performed using the scaling $h_0 \propto \sqrt{\dot{f}/f}$ (Aasi et al., 2014*b*) (this is the spin-down limit, which assumes that the only loss of energy from the system is due to gravitational wave emission) and Eq. 2.36. This gives $h_c \propto f^{1/2}$; higher frequency sources are observed at a louder SNR because they undergo more cycles.

3.2 Sources for space-based detectors

For a review of the gravitational wave sources for space-based missions see, for example, Amaro-Seoane et al. (2013), Gair et al. (2013) or Jennrich et al. (2011).

3.2.1 Massive black-hole binaries

Space-based detectors will be sensitive to equal-mass mergers in the range (10^4–10^7) $M_\odot$. Predictions of the event rate for these mergers range from $\mathcal{O}$(10–100) yr^{-1} for eLISA with SNRs of up to 10^3 (eLISA Consortium, 2013). The large range in the rate reflects our uncertainty in the growth mechanisms of the supermassive black hole population (Volonteri, 2010). A $10^6 M_\odot$ binary can be observed out to a redshift $z \sim 8$ with an SNR of 100 (eLISA Consortium, 2013). This fiducial source gives the amplitude of the boxes labelled "$\approx 10^6$ *solar mass binaries*" in Chapter 4. The range of frequencies is (3×10^{-4}–3×10^{-1}) Hz, extrapolated using the slope $h_c(f) \propto f^{-2/3}$ derived in Sec. 2.2.2.

3.2.2 Galactic white-dwarf binaries

For space-based detectors, these are the most numerous sources; they are also the only guaranteed source since several detectable systems (known as verification binaries) have already been identified by electromagnetic observations (Stroeer and Vecchio, 2006).

Galactic binaries divide into two classes: the unresolvable and the resolvable galactic binaries. The unresolvable binaries overlap to form a stochastic background as discussed in Sec. 2.2.2. The distinction between resolvable and unresolvable is detector specific; here we choose LISA. This boundary should not be too different for eLISA, but would move substantially for decihertz detectors. Plotted in Chapter 4 with the label "*unresolvable galactic binaries*" is the estimate of this background due to Nelemans et al. (2001) where

an observation time of one year has been assumed,

$$h_{\rm c}(f) = 5 \times 10^{-21} \left(\frac{f}{10^{-3}\ {\rm Hz}} \right)^{-2/3} \quad . \tag{3.2}$$

Estimates for the event rate of resolvable binaries centre around $\mathcal{O}(10^3)$ events for eLISA. The boxes plotted in Chapter 4 with the label "*resolvable galactic binaries*" have a ratio $h_{\rm c}/h_n = 50$ for eLISA at its peak sensitivity. The frequency range of the box is $\left(3 \times 10^{-4} – 10^{-2}\right)$ Hz, estimated from Monte Carlo population simulation results presented in Amaro-Seoane et al. (2013).

3.2.3 Extreme mass-ratio inspirals

EMRIs occur when a compact stellar-mass object inspirals into a supermassive black hole. There is extreme uncertainty in the event rate for EMRIs due to the poorly constrained astrophysics in galactic centres (e.g., Merritt et al., 2011); the best guess estimate is around 25 events per year with eLISA with SNR ≥ 20 (eLISA Consortium, 2013). The boxes labelled "*extreme mass ratio inspirals*" plotted in Chapter 4 have a characteristic strain of $h_{\rm c} = 3 \times 10^{-20}$ at 10^{-2} Hz, which corresponds to a $10 M_\odot$ black hole inspiralling into a $10^6 M_\odot$ black hole at a luminosity distance of 1 Gpc. The frequency width of the box is somewhat uncertain; EMRI events can occur around a black hole of any mass, and hence EMRIs can, in principal, occur at any frequency. The boxes in Chapter 4 are drawn with a width comparable to that of the LISA sensitivity curve.

3.3 Sources for PTAs

3.3.1 Supermassive black hole binaries

The main target for PTAs is a stochastic background of GWs produced by a population of supermassive black-hole binaries at cosmological distances (Sesana et al., 2008). Supermassive black holes are known to lie at the centres of most galaxies and black-hole mergers are associated with the mergers of their host galaxies (Ferrarese and Ford, 2005; Volonteri et al., 2003). The current best 95% confidence limit for the amplitude of the stochastic background is $h_{\rm c} = 2.7 \times 10^{-15}$ at a frequency of $f_0 = 1\ {\rm yr}^{-1}$ (Shannon et al., 2013). There is strong theoretical evidence that the actual background lies close to the current limit (Sesana, 2013). The distribution of source amplitudes with frequency is dependent on the chirp mass distribution of binaries, which is astrophysically uncertain.

Reflecting our lack of information about the true shape of this box, we draw it with a flat top in characteristic strain.

3.3.2 Stochastic background of supermassive binaries

Supermassive black-hole binaries at higher frequencies are inspiralling faster and, hence, there are fewer of them per frequency bin. At a certain frequency, these sources cease to be a background and become individually resolvable (Sesana et al., 2008; Sesana et al., 2009). It is currently unclear whether PTAs will be able to detect an individual binary or a stochastic background first. Plotted in Chapter 4 with the label "*stochastic background*" is a third of the current limit with a cut off frequency of $f = 1\ \mathrm{yr}^{-1}$. This is suggested by Monte Carlo population studies (Sesana et al., 2008), which give a range of plausible amplitudes the mean of which is plotted here. For the resolvable sources, labelled "$\approx$ *10^9 solar mass binaries*", the amplitude of the current limit is plotted between $\left(3 \times 10^{-9}\text{–}3 \times 10^{-7}\right)$ Hz.

3.4 Cosmological sources

In addition to the sources above, early Universe processes, such as inflation (Grishchuk, 2005) or a first-order phase transition (Binétruy et al., 2012), could have created GWs. More speculatively, it has been hypothesised that cosmic strings could also be a potential source (Aasi et al., 2014*a*; Binétruy et al., 2012; Damour and Vilenkin, 2005). These relic GWs allow us to explore energy scales far beyond those accessible by other means, providing insight into new and exotic physics. The excitement surrounding the tentative discovery by BICEP2 of the imprint of primordial GWs (generated during inflation) in the cosmic microwave background (BICEP2 Collaboration, 2014), and the subsequent flurry of activity, has shown the scientific potential of such cosmological GWs. These gravitational wave signals are so alluring because they probe unknown physics; this also makes them difficult to predict. Cosmological stochastic backgrounds have been predicted across a range of frequencies with considerable variation in amplitude. As a consequence of this uncertainty, although we could learn much from measuring these signals, we have not included them amongst the sources in Chapter 4.

The sensitivity curves plotted in Chapter 4 would change substantially for stochastic backgrounds. This is because searches for stochastic backgrounds utilise the cross

correlation of the outputs of multiple detectors for detection (Allen and Romano, 1999; Maggiore, 2000), instead of the individual outputs. Therefore, the sensitivity of a network of ground based detectors is much greater than that of any individual detector. The shape of the sensitivity curve for a stochastic background would also differ from the curves plotted here, for a discussion of this point see Thrane and Romano (2013).

Chapter 4

The spectrum of gravitational wave astronomy

A picture is worth a thousand words

Unknown

The plots in this section show all of the detectors and sources described in the main text. Clearer, interactive versions of these plots, allowing for removal of any of the curves, may be created and downloaded on-line, `http://rhcole.com/apps/GWplotter`.

All of the detector noise curves have any spikes (e.g. due to resonances in the suspension systems of ground-based detectors) removed for clarity.

For the pulsar timing arrays we choose to plot the sensitivity curves with the sharp cutoff at the low frequency end. For a discussion of the different possibilites see Sec. 2.4.4 and Moore, Taylor and Gair (2015).

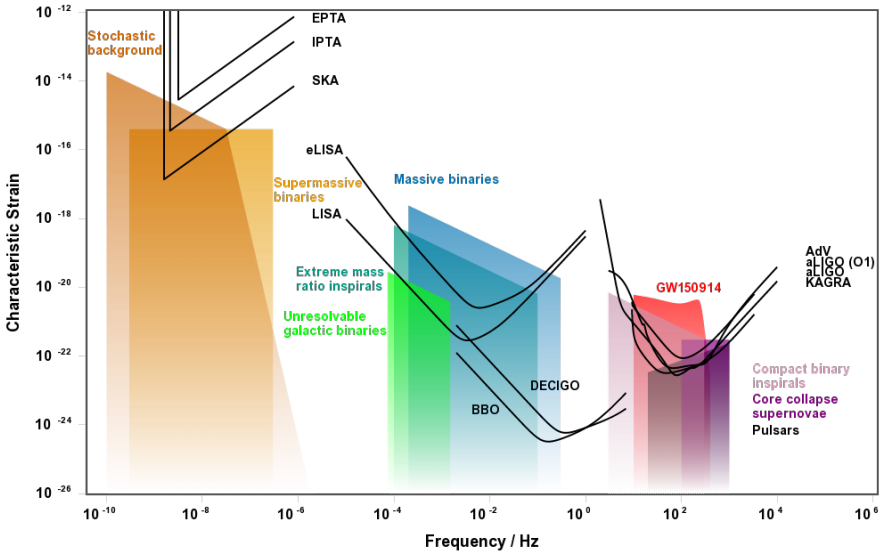


Figure 4.1: A plot of characteristic strain against frequency for a variety of detectors and sources.

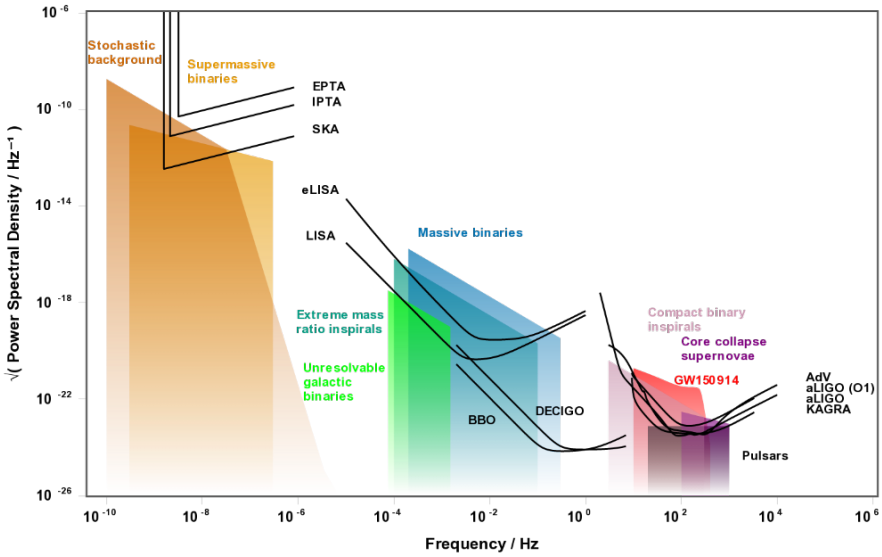


Figure 4.2: A plot of the square root of power spectral density against frequency for a variety of detectors and sources.

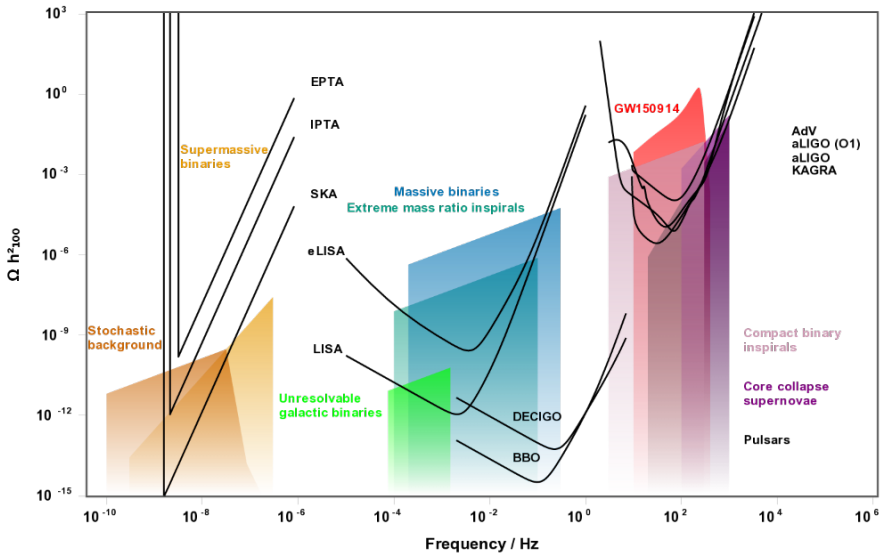


Figure 4.3: A plot of the dimensionless energy density in GWs against frequency for a variety of detectors and sources.

Part II

Data Analysis Methods

Chapter 5

Parameter estimation and model uncertainties

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Rev. Thomas Bayes

Once a signal has been detected in a gravitational wave experiment the next challenge is to infer the properties of the source that generated the waves. This type of inverse problem is common in the physical sciences and is typically very well tackled through the use of Bayesian methods.

These methods are most powerful when they make use of detailed and accurate models of the source. This is typically the situation we are faced with when analysing a compact binary coalescence source of gravitational waves, which consists of just two isolated compact objects and whose evolution is governed by known laws of motion; in this case Einstein's equations.

However, although models of the gravitational wave signal may be available, these models will always contain some errors. There are many possible sources of model error; errors may arise due to some approximation used in the construction of the model (for example, most current models of binary black holes are based on a finite number of terms in the post Newtonian approximation to general relativity), or they may be due to some physical effect which has been neglected by the model (most binary black hole models do not allow for eccentric orbits), or they may be numerical errors arising from solving the equations of the model to finite precision.

Imperfect models cause two main problems for gravitational wave data analysis; the *detection problem* and the *parameter estimation problem* which are described in detail below. The parameter estimation problem is the main focus of the work in this thesis. Recently a technique, called the *marginalised likelihood*, aimed at overcoming the problems associated with the use of imperfect models has been developed. In this chapter we present the details of the marginalised likelihood. The work in this part of the thesis, including chapters 5, 6, 7 and 8, is partly based on the following published papers Moore and Gair (2014), Moore, Berry, Chua and Gair (2016), Gair and Moore (2015), and the following paper in preparation Moore, Chua, Berry and Gair (2016).

5.1 Introduction

The era of advanced ground-based interferometric gravitational-wave (GW) detectors is here. The Advanced LIGO detectors (The LIGO Scientific Collaboration and The Virgo Collaboration, 2016*b*) in the USA started observing in September 2015, and the first direct detection was announced in February 2016 (The LIGO Scientific Collaboration and the Virgo Collaboration, 2016*c*). As discussed in Sec. 2.4.2, several other ground-based detectors are expected to join aLIGO over the coming years. The principal target sources of GWs for these detectors are the coalescences of pairs of compact objects, either neutron stars or black holes; the first detection was from a binary black hole. There is still a great uncertainty in the estimated event rate for these sources, typically over an order of magnitude in each direction; for a discussion of the expected event rates for systems containing a neutron star see Abadie et al. (2010*b*), for updated estimates of the binary black hole merger rate given the recent observation of GW150914 see The LIGO Scientific Collaboration and the Virgo Collaboration (2016*g*). Despite the remaining uncertainties, it is clear that over the coming years we can expect numerous gravitational wave detections of compact binary coalescences.

The detection of gravitational wave signals is most efficient when we have accurate waveform models that can be matched to any signals in the (noisy) detector data.[1] For parameter estimation (PE), it is even more important that the template faithfully matches

[1]It is possible to detect GWs without templates by looking for coherent excess power in the detectors, e.g., Cornish and Littenberg (2015); Kanner et al. (2015); Klimenko et al. (2008*a*); The LIGO Scientific Collaboration and the Virgo Collaboration (2016*d*). This is effective for short-duration signals corresponding to high-mass binaries.

the true signal, as otherwise we could infer biased parameter values. Matching a template to data requires that the model waveform remains accurate over the entire duration of the signal; typically of the order of hundreds of seconds for neutron-star binaries and tens of seconds for black-hole binaries in the advanced-detector era (with frequency sensitivity down to 10 Hz). Although higher mass sources have shorter waveforms (in the detector band), these present more of a challenge for modelling as they have detectable merger and ringdown components. In contrast, binary neutron stars only have the (easier to model) inspiral part of the waveform in band. In this thesis, we are concerned with problems that arise from inaccurate models; therefore, we focus our attention on black-hole binaries where the issue of waveform uncertainty is most acute. However, the techniques we develop could equally be applied to neutron-star binaries or any other uncertain signal. Inaccurate waveform models are known to cause significant systematic errors when recovering source parameters from observations with both ground-based (Canitrot, 2001) and space-based detectors (Cutler and Vallisneri, 2007).

There are two problems that arise when using inaccurate signal models: the *detection* problem, and the *PE* problem. The detection problem is that the inaccurate model does not perfectly match to the physical waveform, leading to a loss of signal-to-noise and, hence, a lower chance of detection (for the same false alarm probability). The PE problem, which is the focus of the work in this thesis, is that the model waveform which has the best overlap with the physical signal in the data generally has parameter values offset from the true source parameters, leading to a *systematic* error in any parameter estimates.

Recently, Moore and Gair (2014) proposed a novel method of improving the detection and PE prospects of complicated physical phenomena in noisy data. The method applies generally to any situation where accurate models of the signal are available, but computational constraints mean that routine detection and PE tasks must be carried out with cheaper, less accurate, models.

In the case of binary black-holes, there are many different physical effects that should be included in waveform models, such as the merger and ringdown phases following the inspiral, the presence of generic spins and precession, eccentricity and higher-order modes. All these phenomena can, in principle, be simulated, thanks to recent rapid progress in numerical relativity (NR); see Baker et al. (2006); Campanelli et al. (2006); Pretorius (2005). However, NR simulations are extremely expensive, only a few hundred have been

performed to date (Aasi et al., 2014*c*; Mroué et al., 2013), and these typically consist of only the final few tens of orbits (although, see Szilágyi et al. (2015)). Detection and PE are therefore currently performed using less expensive waveform approximants. The existence of NR waveforms has permitted the calibration of semi analytic inspiral–merger–ringdown approximants such as the effective-one-body–NR (EOBNR) (Buonanno and Damour, 1999, 2000; Pan et al., 2011; Taracchini et al., 2014) or IMRPhenom (Santamaría et al., 2010) families, with recent efforts concentrating on including the effects of precession in these Hannam et al. (2014); Schmidt et al. (2012, 2015).[2] For some recent PE work with inspiral merger and ringdown waveform models see Ghosh et al. (2015); Graff et al. (2015); Veitch et al. (2015). Historically, PE has also used models based on the post-Newtonian formalism (Blanchet, 2014), such as the TaylorF2 and TaylorT2 waveforms Buonanno et al. (2009). Despite these lacking some of the relevant features, they are sufficiently quick to calculate that they can be simply used in PE algorithms.

To include uncertainty in waveform templates whilst minimising computational expense, we use Gaussian process regression (GPR) to estimate the effects of waveform errors. The method involves constructing a training set of the waveform differences between an expensive, accurate waveform and a cheaper, less accurate waveform. For the accurate waveforms it might be necessary to use some combination of NR and the NR calibrated approximants discussed above, depending on the numbers and length of the available NR simulations. The waveform difference is evaluated at a relatively small number of points in parameter space and stored for later use. GPR is then used to interpolate the difference across parameter space to give a best estimate and a corresponding uncertainty at a general point in parameter space. This interpolation provides a prior probability distribution on the waveform difference which is then used in marginalising the likelihood over waveform uncertainty. The result is an expression for the likelihood in terms of the cheaper waveform model, but with corrections coming from the training set. This marginalised likelihood is negligibly more complicated or computationally expensive to evaluate than the standard expression, but provides a better estimate of the true likelihood surface (and hence the posterior), factoring in our imperfect knowledge of the waveform. Therefore, we have not only built a relatively inexpensive waveform approximant that can include additional physics, but we have also

[2]See Littenberg et al. (2013) for a study of systematic error (or lack thereof) from using EOBNR waveforms with NR injections.

accounted for (marginalised over) the uncertainty in our new approximant.

If the standard likelihood with an approximate waveform is used for PE then, in general, biased parameter estimates are obtained.[3] It has recently been shown by Gair and Moore (2015) that, under certain conditions, this bias is completely removed by the marginalised likelihood, and, more generally, that the bias is always reduced by the marginalised likelihood.

The technique of GPR assumes that the data (waveform differences) in the training set have been drawn from a Gaussian process (GP) on the parameter space with a mean and a covariance function either specified *a priori* or estimated from the training set itself. The interpolation is then achieved by calculating the conditional probability for the GP at some new parameter point given the known training set values, the mean and the covariance. GPR provides a convenient non-parametric way to interpolate the waveform differences, and has the additional advantage that, by construction, it provides a Gaussian probability distribution for the unknown waveform difference which can be analytically marginalised over. This is important because it means no extra nuisance parameters are added to the PE task which would slow down an already expensive process.

5.2 Gaussian Process Regression

5.3 The method

In this section we detail how we incorporate waveform uncertainties into gravitational wave data analysis. In Sec. 5.3.1 we introduce the standard likelihood function and show how model uncertainties can be treated like nuisance parameters that can be integrated out (marginalised over). Performing this integration requires that a prior probability distribution is specified for the model uncertainties, this is constructed using GPR. This is introduced in Sec. 5.3.2, where we briefly summarise some key results pertaining to GPR; further details can be found in standard textbooks (see, for example, Adler (1981); MacKay (2003); Rasmussen and Williams (2006)). The result of the integration is the

[3]This is commonly assessed in the gravitational wave literature using probability–probability (P–P) plots; for a catalogue of events, these plot the cumulative fraction of events where the true parameter is found within the credible interval corresponding to a given probability. If the posteriors are well calibrated, then a proportion P should fall in the P credible interval, and the plot is a diagonal line. Introducing bias means that the line sags below the optimal diagonal.

marginalised likelihood presented in Eq. (5.28) which accurately encodes our state of knowledge of the signal parameters, given our imperfect waveform models and the noisy data.

5.3.1 The marginalised likelihood

We consider the scenario where we can construct two different waveform models, one accurate but computationally expensive, the other less accurate but quick to calculate. We use the parameters vector $\vec{\lambda}$ to fully characterise the signal; Latin indices from the beginning of the alphabet (a, b, ...) will be used to label the different components of this vector, and repeated indices should be summed over. The accurate waveforms will be referred to as the exact waveform $h(\vec{\lambda})$, although the method does not require that the accurate waveforms are perfect (see Sec. 5.4.3). The cheaper approximate waveform $H(\vec{\lambda})$ is related to $h(\vec{\lambda})$ by the waveform difference

$$H(\vec{\lambda}) = h(\vec{\lambda}) + \delta h(\vec{\lambda})\,. \tag{5.1}$$

The waveform templates may be calculated in either the time domain $h(t;\vec{\lambda})$ or the frequency domain $\tilde{h}(f;\vec{\lambda})$; the dependence of the waveform on time or frequency is suppressed in our notation for brevity.

In the context of modelling binary black-hole coalescences there are several highly accurate waveform approximants available, for example, NR waveforms (Aasi et al., 2014*c*) or spin EOBNR (SEOBNR) models (Pan et al., 2011; Taracchini et al., 2012, 2014). There are also multiple possibilities for the approximate waveform family, for example, the Taylor family of approximants of Buonanno et al. (2009). For the proof-of-principal numerical calculations in this section, we need to be able to perform mock PE runs with both waveform families so that we can assess our marginalisation technique does indeed offer a significant improvement. Therefore, we will pick both approximants to be quick to compute, rather than selecting on accuracy: our choice of waveform family is discussed in more detail in Sec. 7.1.

In a PE study, we wish to construct the posterior probability distribution, $p(\vec{\lambda}|s)$, for the signal parameters given the observed data (and any prior information we have about the source). From Bayes' theorem, the posterior is given by

$$p(\vec{\lambda}|s) = \frac{L'(s|\vec{\lambda})\pi(\vec{\lambda})}{\mathcal{Z}'(s)}\,, \tag{5.2}$$

where (keeping the notation of Moore and Gair (2014)) $L'(s|\vec{\lambda})$ is the likelihood, $\pi(\vec{\lambda})$ is the prior distribution on the parameters and $\mathcal{Z}'(s)$ is the normalising evidence

$$\mathcal{Z}'(s) = \int L'(s|\vec{\lambda})\pi(\vec{\lambda})\mathrm{d}\vec{\lambda}\,. \tag{5.3}$$

In a Bayesian analysis the evidence $\mathcal{Z}'(s)$ can be used as the detection statistic (by comparing it with the evidence for the null hypothesis to form the Bayes' factor) (Veitch and Vecchio, 2010), and the positions and widths of peaks in the posterior $p(\vec{\lambda}|s)$ are used to give the parameter estimates and associated uncertainties (Veitch, J. et al., 2015). For simplicity (although it is not necessary to do so), we assume throughout that $\pi(\vec{\lambda})$ is flat within the relevant region of parameter space. The single remaining challenge is to calculate the likelihood $L'(s|\vec{\lambda})$.

For a detector with stationary, Gaussian noise with power spectral density $S_n(f)$, the likelihood is given by (Cutler and Flanagan, 1994*b*)

$$L'(s|\vec{\lambda}) \propto \exp\left(-\frac{1}{2}\left\langle s - h(\vec{\lambda})\middle| s - h(\vec{\lambda})\right\rangle\right)\,. \tag{5.4}$$

Here the noise-weighted inner product has been defined as in Finn (1992)

$$\begin{aligned}\langle x|y\rangle &= 4\Re\left\{\int_0^\infty \mathrm{d}f\,\frac{\tilde{x}(f)\tilde{y}(f)^*}{S_n(f)}\right\}\\ &= 4\Re\left\{\sum_{\kappa=1}^{M}\delta f\,\frac{\tilde{x}(f_\kappa)\tilde{y}(f_\kappa)^*}{S_n(f_\kappa)}\right\}\,,\end{aligned} \tag{5.5}$$

where κ labels the M frequency bins with resolution δf. We define the norm of a waveform as

$$\|x\| = \sqrt{\langle x|x\rangle}\,, \tag{5.6}$$

for a signal this is equivalent to its SNR.

In practice it can be unfeasible to sample from the likelihood distribution in Eq. (5.4) because it is prohibitively expensive to calculate the exact waveforms $h(\vec{\lambda})$; instead, we must reply on the approximate waveforms to calculate an approximate likelihood,

$$L(s|\vec{\lambda}) \propto \exp\left(-\frac{1}{2}\left\|s - H(\vec{\lambda})\right\|^2\right)\,. \tag{5.7}$$

For a good approximant

$$L(s|\vec{\lambda}) \approx L'(s|\vec{\lambda})\,; \tag{5.8}$$

the natural way to improve this agreement is to construct (inevitably more expensive) approximants that have smaller waveform differences $\delta h(\vec{\lambda})$. Instead, the proposal of this

work is to replace $L(s|\vec{\lambda})$ with a new likelihood which accounts for the uncertainty in the waveforms. The alternative likelihood is

$$\mathcal{L}(s|\vec{\lambda}) \propto \int \mathrm{d}\left[\delta h(\vec{\lambda})\right]\ P[\delta h(\vec{\lambda})] \exp\left(-\frac{1}{2}\left\|s - H(\vec{\lambda}) + \delta h(\vec{\lambda})\right\|^2\right) . \tag{5.9}$$

This new likelihood has marginalised over the uncertainty in the waveform difference using the (as yet unspecified) prior on the waveform difference $P[\delta h(\vec{\lambda})]$.

The prior on the waveform difference should include the information available from the limited number of available accurate waveforms and could also encode our prior expectations about the signal, for example, that the approximate waveforms are most accurate at early times (or equivalently at low frequencies) when the orbiting bodies are well separated (Blanchet, 2014), but gradually become inaccurate as the bodies inspiral. At most points in parameter space, an accurate waveform is not available, and so it is necessary to interpolate the waveform difference across parameter space while simultaneously accounting for the error this introduces. It would seem that the problem rapidly becomes complicated, and even if a suitable prior could be constructed the computational time needed to evaluate $\mathcal{L}(s|\vec{\lambda})$ would make it impractical in most contexts.

Fortunately, the technique of GPR provides a natural way to interpolate the waveform differences across parameter space, incorporating all necessary prior information. GPR also has the additional property that it naturally returns an expression for $P[\delta h(\vec{\lambda})]$ which is a Gaussian in $\delta h(\vec{\lambda})$. Since the exponential factor in Eq. (5.9) is also Gaussian in $\delta h(\vec{\lambda})$, the functional integral can be evaluated analytically. This gives an analytic expression for $\mathcal{L}(s|\vec{\lambda})$ which can be evaluated in approximately the same computational time as $L(s|\vec{\lambda})$.

Henceforth, for brevity, the s dependence will be suppressed in all likelihoods, i.e. $L'(\vec{\lambda}) \equiv L'(s|\vec{\lambda})$, $L(\vec{\lambda}) \equiv L(s|\vec{\lambda})$, and $\mathcal{L}(\vec{\lambda}) \equiv \mathcal{L}(s|\vec{\lambda})$.

5.3.2 Gaussian process regression

Assume that we have access to accurate waveforms at a few values of the parameters $\{h(\vec{\lambda}_i) \mid i = 1, 2, \ldots, N\}$ and can cheaply compute approximate waveforms at the same parameter values. Our training set is the set of waveform differences

$$\mathcal{D} = \left\{ \left(\vec{\lambda}_i,\, \delta h(\vec{\lambda}_i)\right) \,\middle|\, i = 1,\, 2, \ldots,\, N \right\} , \tag{5.10}$$

where necessary the Latin indices i, j, ... will be used to label the different components of the training set (repeated indices are not summed over unless specified). It is now necessary to interpolate the training set to obtain the prior on the waveform difference first defined in Eq. (5.9),

$$P[\delta h] \equiv P(\delta h(\vec{\lambda})|\mathcal{D}, \mathcal{I}) \,, \tag{5.11}$$

where $\mathcal{I}$ is any other prior information we possess about the waveforms. The simplest and most natural choice for such a prior is to assume that the waveform difference is a realisation of a GP (a Gaussian is the maximum-entropy distribution given that we know a characteristic range of variation (Shannon, 1948)),

$$\delta h(\vec{\lambda}) \sim \mathcal{GP}(m(\vec{\lambda}), k(\vec{\lambda}, \vec{\lambda}')) \tag{5.12}$$

A GP can loosely be thought of as the generalisation of a Gaussian distribution to an infinite number of degrees of freedom. It is completely specified by the mean $m(\vec{\lambda})$ and covariance $k(\vec{\lambda}, \vec{\lambda}')$ functions in the same way as a Gaussian distribution is fully specified by a mean and variance. More formally, a GP is an infinite collection of variables, any finite subset of which are distributed as a multivariate Gaussian. For a set of parameter points $\{\vec{\lambda}_i\}$, including, but not limited to, the training set $\mathcal{D}$,

$$\left[\delta h(\vec{\lambda}_i)\right] \sim \mathcal{N}(\boldsymbol{m}, \boldsymbol{K}) \,, \tag{5.13}$$

where the mean vector and covariance matrix of this Gaussian distribution are fixed by the corresponding functions of the GP,

$$[\boldsymbol{m}]_i = m(\vec{\lambda}_i) \,, \quad [\boldsymbol{K}]_{ij} = k(\vec{\lambda}_i, \vec{\lambda}_j) \,, \tag{5.14}$$

with probability density function (here correcting the normalising prefactor written in Moore and Gair (2014) which mistakenly included a square root)

$$P\left(\left\{\delta h(\vec{\lambda}_i)\right\}\right) = \frac{1}{(2\pi)^N \left|\boldsymbol{K}\right|} \exp\left(-\frac{1}{2}\sum_{i,\,j}\left[\boldsymbol{K}^{-1}\right]_{ij}\left(\delta h(\vec{\lambda}_i)\middle|\delta h(\vec{\lambda}_j)\right)\right) . \tag{5.15}$$

The round brackets denote a new inner product with respect to some noise weighting $S_n'(f)$, which we leave unspecified for the moment;

$$\begin{aligned}(x|y) &= 4\Re\left\{\int_0^\infty \mathrm{d}f\, \frac{\tilde{x}(f)\tilde{y}(f)^*}{S_n'(f)}\right\} \\ &= 4\Re\left\{\sum_{\kappa=1}^{M} \delta f\, \frac{\tilde{x}(f_\kappa)\tilde{y}(f_\kappa)^*}{S_n'(f_\kappa)}\right\} .\end{aligned} \tag{5.16}$$

In writing down Eq. (5.15) and stipulating that the covariance function $k(\vec{\lambda}, \vec{\lambda}')$ has no dependence on frequency, we are effectively assuming that a) the parameter space structure of the model errors is frequency independent; and b) the typical size of errors has a frequency dependence proportional to $\sqrt{S_n'(f)}$. Under the assumption that waveform model errors are uncorrelated in frequency, the normalising factor in Eq. 5.15 should be raised to the power M; however, this assumption leads to model errors that average to zero over frequency and have only a small effect on the likelihood. The optimal means of incorporating frequency dependence would be to introduce an additional covariance function in frequency as well as the covariance in parameter space. This frequency covariance introduces a correlation length scale in frequency which can be learnt from the training set in exactly the same manner as we describe below for correlations in $\vec{\lambda}$. This correlation length scale reduces the number of independent frequencies from M to some new effective number $M_{\rm eff}$.

Performing this double GPR interpolation in f and $\vec{\lambda}$ is beyond the scope of the current work, although see Chapter 8. Instead, here we are in effect setting $M_{\rm eff} = 1$, giving the expression in Eq. 5.15; this is analogous to assuming that all the frequency bins of the noise-weighted waveform at a particular point in parameter space are perfectly correlated. Setting $M_{\rm eff} = 1$ gives the largest uncertainty of any fixed number of independent frequencies and is therefore a conservative choice. Despite these simplifications, our marginalised likelihood has many desirable properties (which we discuss and prove in Sec. 6), and performs well in the numerical example presented in Sec. 7. We will return to the more general problem of performing the extended GPR including frequency in the future.

Specifying how we compute the mean and variance for the GP determines how the waveforms are interpolated and fixes our prior for waveform uncertainty across parameter space. Our GP has a zero mean as we have chosen to interpolate the waveform difference rather than the waveform directly. By first subtracting off an approximate model we leave a quantity which is uncertain, but has no known bias. If we had some additional prior knowledge that the approximate waveform was systematically wrong across parameter space, then this should be added into the approximate model so that the zero-mean assumption becomes valid. Identical results for the marginalised likelihood could also be obtained by directly interpolating the accurate waveforms using a GP with a mean equal to the approximate waveforms; however, we choose to interpolate waveform differences

because zero-mean GPs are simpler to handle numerically.

Specifying the covariance function is central to GPR as it encodes our prior expectations about the properties of the function being interpolated. Possibly the simplest and most widely used choice for the covariance function is the squared exponential (SE) (Rasmussen and Williams, 2006)

$$k(\vec{\lambda}_i, \vec{\lambda}_j) = \sigma_f^2 \exp\left[-\frac{1}{2} g_{ab}(\vec{\lambda}_i - \vec{\lambda}_j)^a (\vec{\lambda}_i - \vec{\lambda}_j)^b\right] , \tag{5.17}$$

which defines a stationary, smooth GP. In Eq. (5.17), a scale σ_f and a (constant) metric g_{ab} for defining a modulus in parameter space have been defined. These are called *hyperparameters* and we denote them as $\vec{\theta} = \{\sigma_f, g_{ab}\}$, with Greek indices μ, ν, ... to label the components of this vector. We must have the number of hyperparameters being less than size of the training set. If the available accurate waveforms contain some uncertainty then this can also be included by adding a diagonal matrix $\boldsymbol{C}$ to Eq. (5.17), where the element C_{ii} (no summation) is the uncertainty in the accurate simulation at $\vec{\lambda}_i$; this is discussed further in Sec. 5.4.3.

The probability in Eq. (5.15) is referred to as the *hyperlikelihood*, or alternatively the *evidence* (as in Moore and Gair (2014)) for the training set; it is the probability that that particular realisation of waveform differences was obtained from a GP with a zero mean and specified covariance function. The hyperlikelihood depends only on the hyperparameters and the quantities in the training set, so we denote it as $Z(\vec{\theta}|\mathcal{D})$. The log hyperlikelihood is

$$\begin{aligned} \ln Z(\vec{\theta}|\mathcal{D}) \;=\; & -\frac{N}{2}\ln(2\pi) \\ & -\frac{1}{2}\sum_{i,j} \mathrm{inv}\left[k(\vec{\lambda}_i, \vec{\lambda}_j)\right]\left(\delta h(\vec{\lambda}_i)\middle|\delta h(\vec{\lambda}_j)\right) \\ & -\frac{1}{2}\ln\left|\det\left[k(\vec{\lambda}_i, \vec{\lambda}_j)\right]\right| . \end{aligned} \tag{5.18}$$

For all subsequent calculations the values of the hyperparameters are fixed to their optimum values $\vec{\theta}_{\rm op}$, defined as those which maximise the hyperlikelihood:

$$\left.\frac{\partial Z(\vec{\theta}|\mathcal{D})}{\partial \theta^\mu}\right|_{\vec{\theta}=\vec{\theta}_{\rm op}} = 0 . \tag{5.19}$$

Maximising the hyperlikelihood with respect to $\vec{\theta}$ is one of many approaches which could be taken. For example, a better motivated approach would be to consider the

hyperparameters as nuisance parameters in addition to the source parameters $\vec{\lambda}$, and marginalise over them while sampling an expanded likelihood,

$$\Lambda_{\rm expanded}(\vec{\lambda},\vec{\theta}|\mathcal{D}) \quad \propto \mathcal{L}(\vec{\lambda}|\vec{\theta},\mathcal{D})Z(\vec{\theta}|\mathcal{D}). \tag{5.20}$$

The disadvantage of this approach is that the hyperlikelihood is *much* more expensive to compute than the standard approximate likelihood and the inclusion of extra nuisance parameters also slows down any PE. In contrast, our proposed method of maximising the likelihood is a convenient heuristic which is widely used in other contexts (MacKay, 1999; Quiñonero-Candela et al., 2007; Snelson and Ghahramani, 2006) and allows all the additional computation to be done offline. It would be useful, in future work, to check explicitly that the different ways of dealing with the hyperparameters give consistent results in the context of gravitational wave source modelling.

Having fixed the properties of the covariance function by examining the training set, we can now move on to using the GP as a predictive tool. The defining property of the GP is that *any* finite collection of variables drawn from it is distributed as a multivariate Gaussian in the manner of Eq. (5.15). Therefore, the set of variables formed by the training set plus the waveform difference at one extra parameter point $\delta h(\vec{\lambda})$ is distributed as

$$\begin{bmatrix}\delta h(\vec{\lambda}_i)\\ \delta h(\vec{\lambda})\end{bmatrix} \sim \mathcal{N}\left(\mathbf{0},\mathbf{\Sigma}\right), \quad \mathbf{\Sigma} = \begin{pmatrix}\boldsymbol{K} & \boldsymbol{K}_*\\ \boldsymbol{K}_*^{\rm T} & K_{**}\end{pmatrix}, \tag{5.21}$$

where $\boldsymbol{K}$ is defined in Eq. (5.14) and the vector $\boldsymbol{K}_*$ and scalar K_{**} are defined as

$$[\boldsymbol{K}_*]_i = k(\vec{\lambda}_i,\vec{\lambda}), \quad K_{**} = k(\vec{\lambda},\vec{\lambda}). \tag{5.22}$$

On the right-hand side of Eq. (5.21) all the quantities are known because the hyperparameters have been fixed to their optimum values, and on the left hand side all the quantities are known (from the training set) except for $\delta h(\vec{\lambda})$. Therefore, the conditional probability of the unknown waveform difference given the known differences in $\mathcal{D}$ can be found. This conditional probability is given by (e.g., MacKay (2003); Rasmussen and Williams (2006))

$$P[\delta h(\vec{\lambda})] = \frac{1}{2\pi\sigma^2(\vec{\lambda})\prod_{\kappa=1}^{M} S_n'(f_\kappa)} \exp\left(-\frac{\left(\delta h(\vec{\lambda})-\mu(\lambda)\middle|\delta h(\vec{\lambda})-\mu(\lambda)\right)}{2\sigma^2(\vec{\lambda})}\right), \tag{5.23}$$

where the GPR mean and its associated error have been defined as

$$\mu(\vec{\lambda}) \quad = \quad \sum_{i,j}[\boldsymbol{K}_*]_i\left[\boldsymbol{K}^{-1}\right]_{ij}\delta h(\vec{\lambda}_j), \tag{5.24}$$

$$\sigma^2(\vec{\lambda}) \;=\; K_{**} - \sum_{i,j} [\boldsymbol{K}_*]_i \left[\boldsymbol{K}^{-1}\right]_{ij} [\boldsymbol{K}_*]_j \; . \tag{5.25}$$

Furnished with the expression for $P[\delta h(\vec{\lambda})]$, the marginalised likelihood in Eq. (5.9) can now be evaluated. The integrand in Eq. (5.9) is the product of two Gaussians and can be calculated analytically,

$$\begin{aligned}\mathcal{L}(\vec{\lambda}) \propto & \frac{1}{1+\sigma^2(\vec{\lambda})\prod_{\kappa=1}^{M}\left(S_n'(f_\kappa)/S_n(f_\kappa)\right)} \\ & \times \exp\left(-\frac{1}{2}\left[s-H(\vec{\lambda})+\mu(\vec{\lambda})\middle|s-H(\vec{\lambda})+\mu(\vec{\lambda})\right]\right) \; .\end{aligned} \tag{5.26}$$

The square brackets denote a third inner product with respect to the new noise weighting $S_n''(f)$, where $S_n''(f,\vec{\lambda}) \equiv S_n(f) + \sigma^2(\vec{\lambda})S_n'(f)$,

$$\begin{aligned}[x|y] \;&=\; 4\Re\left\{\int_0^\infty \mathrm{d}f\, \frac{\tilde{x}(f)\tilde{y}(f)^*}{S_n''(f)}\right\} \\ &=\; 4\Re\left\{\sum_{\kappa=1}^{M} \delta f\, \frac{\tilde{x}(f_\kappa)\tilde{y}(f_\kappa)^*}{S_n''(f_\kappa)}\right\} \; .\end{aligned} \tag{5.27}$$

For the remainder of this section, for simplicity, we take $S_n'(f) = S_n(f)$ so the three signal inner products we have defined become $\langle\cdot|\cdot\rangle = (\cdot|\cdot) = [\cdot|\cdot]/(1+\sigma^2(\vec{\lambda}))$. With this simplifying assumption, the marginalised likelihood becomes (Moore and Gair, 2014)

$$\begin{aligned}\mathcal{L}(\vec{\lambda}) \propto & \frac{1}{1+\sigma^2(\vec{\lambda})} \\ & \times \exp\left(-\frac{1}{2}\frac{\left\|s-H(\vec{\lambda})+\mu(\vec{\lambda})\right\|^2}{1+\sigma^2(\vec{\lambda})}\right) \; .\end{aligned} \tag{5.28}$$

As mentioned earlier Eq. (5.15) issues that the waveform model errors are uncorrelated in frequency. The assumption that $S_n'(f) = S_n(f)$ additionally assumes that the typical size of the waveform error at a frequency f is given by $\sqrt{S_n(f)}$. This choice can be motivated to a certain extent by examining the hyperlikelihood in Eq. (5.18) which is used to train the GP. This hyperlikelihood contains the overlap matrix $(\delta h(\vec{\lambda}_i)|\delta h(\vec{\lambda}_j))$. Choosing $S_n'(f) = S_n(f)$ acts to downweight the correlations at frequencies we are insensitive to (ignoring errors we cannot measure) and hence the resulting hyperparameters give an interpolant which is tuned to better represent the waveform correlations at the frequencies to which we are most sensitive: we weight waveform errors based upon their impact on the likelihood. The assumption of frequency-independent models errors gives a value for

the GPR uncertainty σ^2 in Eq. (5.28) that is also frequency-independent. This can be shown to be a conservative choice in the sense that it gives broader and less informative posteriors.

In a follow-on study we will provide a proof of the conservative nature of this assumption and examine a number of different choices for the weighting function $S_n'(f)$, but we use the simplifying assumption $S_n'(f) = S_n(f)$ throughout the current section. Despite these simplifying assumptions, we find that the resulting likelihood in Eq. (5.28) performs well. In Appendix B we examine the sensitivity of the method to small changes in the noise curve $S_n(f)$ which will occur in real experiments.

In Eq. (5.28) the best fit waveform has shifted by an amount $\mu(\vec{\lambda})$; this is the best estimate of the waveform difference returned by the GPR. The quantity $H(\vec{\lambda}) + \mu(\vec{\lambda})$ can be regarded as a new waveform approximant built from the accurate and approximate waveforms with the aid of GPR. However, a bonus of this way of including the training set directly into the likelihood is that the extra uncertainty associated with using the GPR as an interpolant is automatically included via the broadening of the posterior caused by $\sigma^2(\vec{\lambda}) \geq 0$.

In this section we have explained how uncertainty in waveform models can be included in PE through use of a marginalised likelihood. We defined such a likelihood in Eq. (5.9), but the marginalisation requires a prior probability on the waveform uncertainty across parameter space. We construct this from a training set using GPR; the resulting prior is given in Eq. (5.23). Since this is of Gaussian form, we can marginalise analytically to produce the new likelihood Eq. (5.28). The properties of this marginalised likelihood are explored extensively throughout the remainder of this thesis.

In Secs. 5.4 and 6 we discuss theoretical properties of the GPR and marginalised likelihood respectively. A reader who is primarily interested in the PE results obtained with the likelihood in Eq. (5.28) may skip to Sec. 7.

5.4 The covariance function

In the previous section we described how waveform uncertainties could be marginalised out using a prior constructed by using GPR on a training set. The only aspect of this that is not prescribed by the training data is the choice of the covariance function. This plays an important role in determining the properties of a GP. In this section, we discuss the properties of different choices of the covariance function in GPR. The properties

of the covariance functions discussed in this section are known in the GPR literature, but are included here as they are not as widely appreciated in the gravitational wave community. The material presented in this section on the covariance function will be used in the interpretation of our results in chapter 6.

The only necessary requirements we have of a covariance function are that it is a positive definite; i.e. for *any* choice of points $\{\vec{\lambda}_i\}$ the covariance matrix $K_{ij} = k(\vec{\lambda}_i, \vec{\lambda}_j)$ is positive definite.

Throughout this work, GPs are assumed to have zero mean, and therefore be fully specified by the covariance function $k(\vec{\lambda}_1, \vec{\lambda}_2)$. However, the proofs regarding continuity and differentiability of GPs discussed in this section, and proved in Appendix C, are done without recourse to the zero-mean assumption. The covariance encodes all information available about the properties of the function being interpolated by the GPR. It is central to the GPR and hence also to the marginalised likelihood.

The covariance function (and the corresponding GP) is said to be *stationary* if the covariance is a function only of $\vec{\tau} = \vec{\lambda}_1 - \vec{\lambda}_2$, furthermore it is said to be *isotropic* if it is a function only of $\tau \equiv |\vec{\tau}| = |\vec{\lambda}_1 - \vec{\lambda}_2|$.[4] Isotropy of a GP implies stationarity. All of the GPs used for numerical calculations in this section are isotropic (and hence stationary) $k(\vec{\lambda}_1, \vec{\lambda}_2) \equiv k(\vec{\tau}) \equiv k(\tau)$, although the generalisation to non-stationary GPs is briefly discussed in Sec. 5.4.2.

An example of how the properties of the covariance function relate to the properties of the GP, and hence the properties of the resulting interpolant, is given by considering the *mean-square* (MS) continuity and differentiability of GPs. It can be shown that the first n_d MS derivatives of a GP are MS continuous (the GP is said to be n_d-times MS differentiable) if and only if the first $2n_d$ derivatives of the covariance function are continuous at the diagonal point $\vec{\lambda}_1 = \vec{\lambda}_2 = \vec{\lambda}_*$. For a stationary GP this condition reduces to checking the $2n_d$ derivatives of $k(\vec{\tau})$ at $\vec{\tau} = \vec{0}$, and for an isotropic GP checking the $2n_d$ derivatives of $k(\tau)$ at $\tau = 0$. A proof of this result, following Adler (1981), is given in Appendix C. It is the smoothness properties of the covariance function at the origin that determine the differentiability of the GP. This result is used in Sec. 5.4.2 when discussing different functional forms of covariance for use in GPR.

In this section, the effect of the choice of covariance function on the GPR are explored.

[4] We have yet to define a metric on parameter space with which to take the norm of this vector (see Sec. 5.4.1), but all that is required here is that a suitably smooth metric exists.

We consider three aspects that enter the definition of the covariance function:

(A) specifying the distance metric in parameter space g_{ab};

(B) specifying the functional form of the covariance with distance $k(\tau)$,

(C) and whether or not to include errors σ_n on the training set points.

Stages A and B cannot be completely separated; there exists an arbitrary scaling, α of the distance $\tau \to \alpha\tau$ which can be absorbed into the definition of the covariance, $k(\tau) \to k(\tau/\alpha)$. However, provided the steps are tackled in order, there is no ambiguity.

5.4.1 The metric g_{ab}

The first stage involves defining a distance τ between two points in parameter space. One simple way of doing this, and the way used in the SE covariance function in Eq. (5.17), is to define $\tau^2 = g_{ab}(\vec{\lambda}_1 - \vec{\lambda}_2)^a(\vec{\lambda}_1 - \vec{\lambda}_2)^b$, where g_{ab} are constant hyperparameters. This distance is obviously invariant under a simultaneous translation of $\vec{\lambda}_1 \to \vec{\lambda}_1 + \vec{\Delta}$ and $\vec{\lambda}_2 \to \vec{\lambda}_2 + \vec{\Delta}$; therefore, this defines a stationary GP. For a D-dimensional parameter space, this involves specifying $D(D+1)/2$ hyperparameters g_{ab}.

More complicated distance metrics (with a larger number of hyperparameters) are possible if the condition of stationarity is relaxed, i.e. $g_{ab} \to g_{ab}(\vec{\lambda})$. It was demonstrated by Paciorek and Schervish (2004) how, given a family of stationary covariance functions, a non-stationary generalisation can be constructed. A stationary covariance function can be considered as a kernel function centred at $\vec{\lambda}_1$; $k(\vec{\lambda}_1, \vec{\lambda}_2) \equiv k_{\vec{\lambda}_1}(\vec{\lambda}_2)$. Allowing a different kernel function to be defined at each point $\vec{\lambda}_1$, a new, non-stationary covariance function is[5]

$$k(\vec{\lambda}_1, \vec{\lambda}_2) = \int \mathrm{d}\vec{u}\; k_{\vec{u}}(\vec{\lambda_1}) k_{\vec{u}}(\vec{\lambda_2}) \; . \tag{5.29}$$

Applying this procedure to a D-dimensional SE function generates a non-stationary analogue (Paciorek and Schervish, 2004)

$$k(\vec{\lambda}_i, \vec{\lambda}_j) = \sigma_f \left|\mathcal{G}^i\right|^{1/4} \left|\mathcal{G}^j\right|^{1/4} \left|\frac{\mathcal{G}^i + \mathcal{G}^j}{2}\right|^{-1/2} \exp\left(-\frac{1}{2}\mathcal{Q}_{ij}\right) \; , \tag{5.30}$$

[5] To see that k is a valid covariance function consider an arbitrary series of points $\{\vec{\lambda}_i\}$, and the sum over training set points $I = \sum_{i,j} a_i a_j k(\vec{\lambda}_i, \vec{\lambda}_j)$; for k to be a valid covariance it is both necessary and sufficient that $I \geq 0$. Using the definition of k gives $I = \int \mathrm{d}\vec{u} \sum_{i,j} a_i a_j k_{\vec{u}}(\vec{\lambda_i}) k_{\vec{u}}(\vec{\lambda_j}) = \int \mathrm{d}\vec{u} \left(\sum_i a_i k_{\vec{u}}(\vec{\lambda_i})\right)^2 \geq 0$.

where

$$\mathcal{Q}_{ij} \quad = \quad (\vec{\lambda}_i - \vec{\lambda}_j)^a (\vec{\lambda}_i - \vec{\lambda}_j)^b \left(\frac{\mathcal{G}^i_{ab} + \mathcal{G}^j_{ab}}{2} \right)^{-1}, \tag{5.31}$$

and $\mathcal{G}^i_{ab} = \mathrm{inv}[g_{ab}(\vec{\lambda}_i)]$ is the inverse of the parameter-space metric at position $\vec{\lambda}_i$. Provided that the metric $g_{ab}(\vec{\lambda})$ is smoothly parameterised this non-stationary SE function retains the smoothness properties discussed earlier.

For the interpolation of waveform differences, it is easy to imagine the potential benefits of using non-stationary GPs. For example, in the case of the spin parameter, it could be imagined that the waveform difference considered as a function of the effective spin of the compact objects $\delta h(\chi)$ would vary on long length scales in χ for small values of the spin, but on much shorter scales for larger values of the spin.

The generalisation in Eq. (5.30) involves the inclusion of a large set of additional hyperparameters to characterise how the metric changes over parameter space; for example one possible parameterisation would be the Taylor series

$$g_{ab}(\vec{\lambda}) = g_{ab}(\vec{\lambda}_0) + \left(\vec{\lambda}^c - \vec{\lambda}^c_0\right) \left. \frac{\partial g_{ab}(\vec{\lambda})}{\partial \lambda^c} \right|_{\vec{\lambda}=\vec{\lambda}_0} + \ldots \tag{5.32}$$

with the hyperparameters $g_{ab}(\vec{\lambda}_0)$, $\partial g_{ab}(\vec{\lambda})/\partial\lambda^c$, and so on. As we see below, the inclusion of even a single extra hyperparameter can incur a significant Occam penalty (MacKay, 2003) which pushes the training set to favour a simpler choice of covariance function. For this reason we only consider stationary GPs. However, the generalisation to a non-stationary GP (perhaps in only a limited number of parameters, e.g., spin) should be investigated further in the future. In making this generalisation, one would have to be guided significantly by the prior expectations of which parameters to include and how to parameterise the varying metric.

An alternative to considering non-stationary metrics is instead to try and find new coordinates $\tilde{\lambda} \equiv \tilde{\lambda}(\vec{\lambda})$ such that the metric in these coordinates becomes (approximately) stationary. There could be hope for this approach, as a similar problem has been tackled by Owen (1996) in the context of template placement for gravitational wave searches. Here the problem is to find coordinates such that waveform templates placed on a regular grid in these coordinates have a constant overlap with each other. The waveform match can be viewed as defining a metric in parameter space, and hence the desired coordinates make this metric stationary. For a post-Newtonian inspiral signal, a set of chirp-time

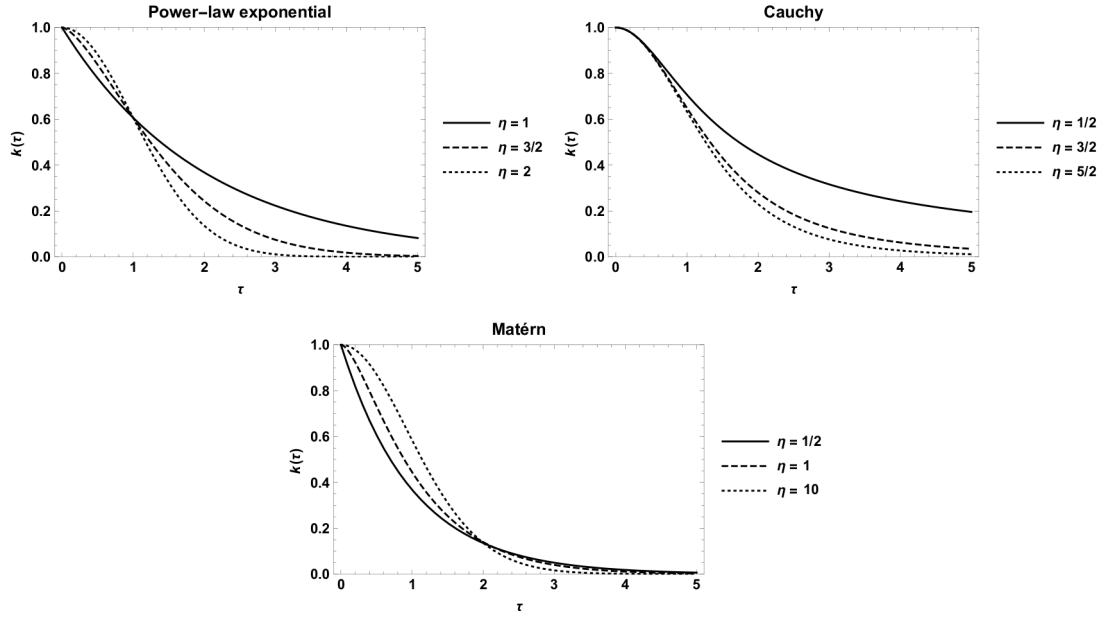


Figure 5.1: Plots of the different generalisations of the SE covariance function discussed in Sec. 5.4.2. The top-most panel shows the PLE function, the bottom-left panel shows the Cauchy function, and the bottom-right panel shows the Matérn function; in all cases the value of σ_f was fixed to unity. In each panel the effect of varying the additional hyperparameter is shown by the three curves. For the PLE covariance the case $\eta = 2$ recovers the SE covariance, while for the Cauchy and Matérn covariances the case $\eta \to \infty$ recovers the SE covariance.

coordinates were proposed by Sathyaprakash and Dhurandhar (1991) which make the metric nearly stationary. Metrics have also been calculated for inspiral–merger–ringdown models, for example IMRPhenomB (Kalaghatgi et al., 2015). While it could be possible to adapt the parameter-space metrics already calculated for different approximants for use in template placement algorithms to help in constructing our GPR training sets, we do not consider this approach further here.

Throughout the remainder of this section the metric components g_{ab} are treated as constant hyperparameters fixed to their optimum values, as discussed in Sec. 5.3.

5.4.2 The functional form of $k(\tau)$

The second stage of specifying the covariance function involves choosing the function of distance $k(\tau)$. In general whether a particular function $k(\tau)$ is positive definite (and hence is a valid covariance function) depends on the dimensionality D of the underlying space (i.e. $\vec{\lambda} \in \mathbb{R}^D$); however, all the functions considered in this section are positive definite for all D. Several choices for $k(\tau)$ are particularly common in the literature. These include the SE covariance function (which has already been introduced), given by

$$k_{\mathrm{SE}}(\tau) = \sigma_f^2 \exp\left(-\frac{1}{2}\tau^2\right) . \tag{5.33}$$

The *power-law exponential* (PLE) covariance function is given by

$$k_{\mathrm{PLE}}(\tau) = \sigma_f^2 \exp\left(-\frac{1}{2}\tau^\eta\right) , \tag{5.34}$$

where $0 < \eta \leq 2$. The PLE reduces to the SE in the case $\eta = 2$. The *Cauchy* function is given by

$$k_{\mathrm{Cauchy}}(\tau) = \frac{\sigma_f^2}{(1 + \tau^2/2\eta)^\eta} , \tag{5.35}$$

where $\eta > 0$. This recovers the SE function in the limit $\eta \to \infty$. And finally, the *Matérn* covariance function is given by Stein (1999)

$$k_{\mathrm{Mat}}(\tau) = \frac{\sigma_f^2 2^{1-\eta}}{\Gamma(\eta)} \left(\sqrt{2\eta}\,\tau\right)^\eta K_\eta \left(\sqrt{2\eta}\,\tau\right) , \tag{5.36}$$

where $\eta > 1/2$, and K_η is the modified Bessel function of the second kind (Watson, 1995). In the limit $\eta \to \infty$, the Matérn covariance function also tends to the SE.

Fig. 5.1 shows the functional forms of the covariance functions. They have similar shapes: they all return a finite covariance at zero distance which decreases monotonically with distance and tends to zero as the distance becomes large. In the case of interpolating waveform differences this indicates that the errors in the approximate waveform at two nearby points in parameter space are closely related, whereas the errors at two well separated points are nearly independent. The PLE, Cauchy and Matérn function can all be viewed as attempts to generalise the SE with the inclusion of one extra hyperparameter η, to allow for more flexible GP modelling. All three alternative functions are able to recover the SE in some limiting case, but the Matérn is the most flexible of the three. This can be seen from the discussion of the MS differentiability of GPs given at the beginning of this section.

The SE covariance function is infinitely differentiable at $\tau = 0$, and so the corresponding GP is infinitely MS differentiable. The PLE function is infinitely differentiable at $\tau = 0$ for the SE case when $\eta = 2$, but for all other cases it is not at all MS differentiable. In contrast, the Cauchy function is infinitely differentiable at $\tau = 0$ for all choices of the hyperparameter η. The Matérn function, by contrast, has a variable level of differentiability at $\tau = 0$, controlled via the hyperparameter η (Stein, 1999). The GP corresponding to the Matérn covariance function in Eq. (5.36) is n_d-times MS differentiable if and only if $\eta > n_d$. This ability to adjust the differentiability allows the same covariance function to successfully model a wide variety of data. In the process of maximising the hyperlikelihood for the training set over hyperparameter η, the GP *learns* the (non)smoothness properties favoured by the data, and the GPR returns a correspondingly (non)smooth function.

5.4.3 The inclusion of noise σ_n

Even the most accurate waveform models $h(\vec{\lambda})$ still contain some error with respect to the unknown true physical signal $h'(\vec{\lambda})$. This could be because the waveform model does not include all of the physics or because it is calculated using a method with finite accuracy. We can account for the error in our training set points by adding a noise variance term $\sigma_f^2 \sigma_{n,\,i}^2$ in the covariance function,

$$k(\vec{\lambda}_i, \vec{\lambda}_j) \to k(\vec{\lambda}_i, \vec{\lambda}_j) + \sigma_f^2 \sigma_{n,\,i}^2 \delta_{ij}\,, \tag{5.37}$$

which alters the covariance matrix in Eq. (5.14) correspondingly, but not the expressions in Eq. (5.22). Here $\sigma_{n,\,i}$ is the *fractional* error $\|h - h'\|/\|\delta h\|$ in each training set point, where the norm is taken with respect to the inner product in Eq. (5.5) and $\delta h = H - h$. This ensures that σ_f^2 is still an overall scale for the covariance function.

We do not maximise the hyperlikelihood over $\sigma_{n,\,i}^2$; this is because $\sigma_{n,\,i}$ is related to $\|h - h'\|$, which cannot be learnt from a training set containing the differences δh. The noise error is instead fixed at some overall error estimate for the accurate model, which is a conservative approach. We consider the simple case $\sigma_{n,\,i} = \sigma_n$ in this section; however, it is not necessary for all training set points to have the same error, as a training set might comprise different families of waveform models (e.g., a mix of different variants of (S)EOBNR or IMRPhenom waveforms, or NR waveforms with different numerical resolutions).

If the overall noise error is $\sigma_f \sigma_n$, then the GPR uncertainty at a training set point $\sigma(\vec{\lambda}_i)$ satisfies,

$$\sigma(\vec{\lambda}_i) \leq \sigma_f \sigma_n, \quad \forall i \in \{1, 2, \ldots, N\}. \tag{5.38}$$

This is because the different points in the training set are assumed to come from a correlated GP, and so nearby measurements also act to decrease the error.

There is also a more practical motivation for the inclusion of noise. Inversion of the covariance matrix in Eq. (5.14) can pose issues of numerical stability for large training sets. In general, as the number of points in the training set increases, the determinant of the covariance matrix decreases rapidly towards zero, such that the matrix is nearly singular (and hence the matrix is difficult to invert). The solution to this is to add a small fixed noise $\sigma_n^2 = J \ll 1$, or *jitter*, to each training set point as per Eq. (5.37). The eigenvalues of the new covariance matrix are then (approximately) the eigenvalues of the original matrix plus J. This prevents the determinant, the product of the eigenvalues, from becoming vanishingly small and dramatically improves the stability of the inversion. In effect, we are no longer requiring our interpolant to pass through every training set point; instead, we only ask it to pass close to each point (with the proximity determined by the value of J).

5.4.4 Compact support and sparseness

All of the covariance functions considered up until this point have been strictly positive;

$$k(\tau) > 0 \quad \forall\, \tau \in [0, \infty)\,. \tag{5.39}$$

When evaluating the covariance matrix for the training set K_{ij} this leads to a matrix where all entries are positive; i.e. a dense matrix. When performing the GPR it is necessary to maximise the hyperlikelihood for the training set with respect to the hyperparameters. This process involves inverting the dense matrix K_{ij} at each iteration of the optimisation algorithm. Although this procedure is carried out offline, it still can become prohibitive for large training sets. A related problem, as pointed out in Sec. 5.4.3 is that for large training sets the determinant of the covariance matrix is typically small which also contributes to making the covariance matrix hard to invert.

One potential way around these issues is to consider a covariance function with

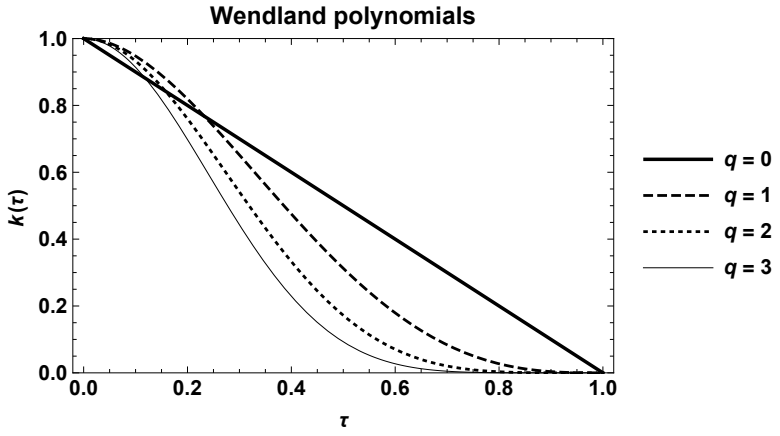


Figure 5.2: Plots of the first few Wendland polynomial covariance functions. All these functions have compact support, $k(\tau) = 0$ for $\tau > 1$. As the value of q increases the functions become smoother near the origin.

compact support,

$$\begin{aligned} k(\tau) &> 0 \quad \forall\, \tau \in [0, T]\,, \\ k(\tau) &= 0 \quad \forall\, \tau \in (T, \infty)\,, \end{aligned} \tag{5.40}$$

where T is some threshold distance beyond which we assume that the waveform differences become uncorrelated. This leads to a sparse, band-diagonal covariance matrix, which is much easier to invert. Care must be taken when specifying the covariance function to ensure that the function is still positive definite (which is required of a GP): if the SE covariance function is truncated, then the matrix formed from the new covariance function is not guaranteed to be positive definite.

Nevertheless, it is possible to construct covariance functions which have the requisite properties and satisfy the compact support condition in Eq. (5.40). These are typically based on polynomials. We consider a series of polynomials proposed by Wendland (2004), which we will refer to as the *Wendland* polynomials. These have the property that they are positive definite in $\mathbb{R}^D$ and are $2q$-time differentiable at the origin. Therefore the discrete parameter q is in some sense analogous to the η hyperparameter of the Matérn covariance function in that it controls the smoothness of the GP. Defining β to be

$$\beta = \left\lfloor \frac{D}{2} \right\rfloor + q + 1 \tag{5.41}$$

and using $\Theta(x)$ to denote the Heaviside step function, the first few Wendland polynomials

$k_{D,q}(\tau)$ are given by,

$$
\begin{aligned}
k_{D,0}(\tau) &= \sigma_f^2 \Theta(1-\tau)(1-\tau)^{\beta}\,, && (5.42)\\
k_{D,1}(\tau) &= \sigma_f^2 \Theta(1-\tau)(1-\tau)^{\beta+1}\left[1+(\beta+1)\,\tau\right]\,, && (5.43)\\
k_{D,2}(\tau) &= \frac{\sigma_f^2}{3}\Theta(1-\tau)(1-\tau)^{\beta+2}\left[\,3+(3\beta+6)\,\tau\right.\\
&\quad \left.+\left(\beta^2+4\beta+3\right)\tau^2\right]\,, && (5.44)\\
k_{D,3}(\tau) &= \frac{\sigma_f^2}{15}\Theta(1-\tau)(1-\tau)^{\beta+3}\left[15+(15\beta+45)\,\tau\right.\\
&\quad +\left(6\beta^2+36\beta+45\right)\tau^2\\
&\quad \left.+\left(\beta^3+9\beta^2+23\beta+15\right)\tau^3\right]\,. && (5.45)
\end{aligned}
$$

These are plotted in Fig. 5.2. Other types of covariance function with compact support have also been proposed and explored in the literature (e.g., Gneiting (2002); Liang and Marcotte (2015); Melkumyan and Ramos (2009)), but we do not consider them in this thesis.

Chapter 6

Properties of the method

In this chapter proofs of several useful features of the marginalised likelihood are presented. In Sec. 6.1 we derive the PE error in a linearised formalism, recovering results of Cutler and Vallisneri (2007) as well as new results for our marginalised likelihood; in Sec. 6.2 we use these results to show that the marginalised likelihood should not exclude the true parameter values even at large SNR, and in Sec. 6.3, we derive other limits of the marginalised likelihood at specific points in parameter space.

6.1 The error at linear order

A more detailed understanding of the theoretical error problem, and the solution offered by the marginalised likelihood can be gained by examining the behaviour of the likelihoods in the vicinity of a maximum.

The *exact likelihood*, from Eq. (5.4), is given by

$$L'(\vec{\lambda}) \propto \exp\left(-\frac{1}{2}\left\|s - h(\vec{\lambda})\right\|^2\right) , \tag{6.1}$$

and has a maximum at the best fit parameters, $\vec{\lambda}_{\rm bf}$, which satisfy the equation

$$\left\langle \partial_a h(\vec{\lambda}_{\rm bf}) \middle| s - h(\vec{\lambda}_{\rm bf}) \right\rangle = 0 \, . \tag{6.2}$$

The measured data consist of noise and the physical signal with the true parameters, $\vec{\lambda}_{\rm tr}$, that is $s = n + h(\vec{\lambda}_{\rm tr})$. Therefore Eq. (6.2) becomes

$$\left\langle \partial_a h(\vec{\lambda}_{\rm bf}) \middle| n + h(\vec{\lambda}_{\rm tr}) - h(\vec{\lambda}_{\rm bf}) \right\rangle = 0 \, . \tag{6.3}$$

Expanding the difference in the signals to leading order in $\Delta\vec{\lambda} = \vec{\lambda}_{\rm bf} - \vec{\lambda}_{\rm tr}$ gives

$$\left\langle \partial_a h(\vec{\lambda}_{\rm bf}) \middle| n - \Delta\vec{\lambda}^b \partial_b h(\vec{\lambda}_{\rm bf}) \right\rangle = 0\,, \tag{6.4}$$

whence

$$\Delta\vec{\lambda}^a = \left(\Sigma^{-1}\right)^{ab} \left\langle n \middle| \partial_b h(\vec{\lambda}_{\rm bf}) \right\rangle\,, \tag{6.5}$$

where $\Sigma_{ab} = \langle \partial_a h(\vec{\lambda}_{\rm bf}) | \partial_b h(\vec{\lambda}_{\rm bf}) \rangle$. Therefore, at leading order, the shift between the best fit and true parameters for the exact likelihood consists of one term proportional to n; we call this the noise error. The matrix Σ_{ab} is the usual Fisher information matrix (FIM) which characterises the random errors at leading order (Vallisneri, 2008).

The *approximate likelihood*, from Eq. (5.7), is given by

$$L(\vec{\lambda}) \propto \exp\left(-\frac{1}{2}\left\| s - H(\vec{\lambda}) \right\|^2\right)\,, \tag{6.6}$$

and has a maximum at the best fit parameters which satisfy the equation

$$\left\langle \partial_a H(\vec{\lambda}_{\rm bf}) \middle| s - H(\vec{\lambda}_{\rm bf}) \right\rangle = 0\,. \tag{6.7}$$

Using $s = n + h(\vec{\lambda}_{\rm tr})$ in Eq. (6.7) and expanding to leading order in $\Delta\vec{\lambda}$ gives

$$\left\langle \partial_a H(\vec{\lambda}_{\rm bf}) \middle| n - \delta h(\vec{\lambda}_{\rm tr}) - \Delta\vec{\lambda}^b \partial_b H(\vec{\lambda}_{\rm bf}) \right\rangle = 0\,, \tag{6.8}$$

thus

$$\Delta\vec{\lambda}^a = \left(\Gamma^{-1}\right)^{ab} \left\langle n \middle| \partial_b H(\vec{\lambda}_{\rm bf}) \right\rangle - \left(\Gamma^{-1}\right)^{ab} \left\langle \delta h(\vec{\lambda}_{\rm tr}) \middle| \partial_b H(\vec{\lambda}_{\rm bf}) \right\rangle\,, \tag{6.9}$$

where $\Gamma_{ab} = \langle \partial_a H(\vec{\lambda}_{\rm bf}) | \partial_b H(\vec{\lambda}_{\rm bf}) \rangle$. Therefore, at leading order the shift between the best fit and true parameters for the approximate likelihood consists of two terms: the noise error as before (except with the FIM evaluated with the approximate model) and what we call the model error,

$$\Delta_{\rm model}\vec{\lambda}^a = -\left(\Gamma^{-1}\right)^{ab} \left\langle \delta h(\vec{\lambda}_{\rm tr}) \middle| \partial_b H(\vec{\lambda}_{\rm bf}) \right\rangle\,. \tag{6.10}$$

The model error is independent of the noise realisation, and hence represents a systematic error in the PE associated with using inaccurate models.

The above treatment of the exact and approximate likelihoods is a brief summary of part of the analysis done by Cutler and Vallisneri (2007). We now apply the same type of analysis to the new marginalised likelihood to see how this reduces or removes the model error.

The *marginalised likelihood* is given in Eq. (5.9). From Eq. (5.25) it can be seen that the interpolated waveform difference $\mu(\vec{\lambda})$ is a linear combination of $\delta h(\vec{\lambda}_i)$ from the training set. We will assume, for this calculation only, that the waveform difference is also a small quantity in the sense that $\|\delta h\| \ll \|h\|$ with the norm from Eq. (5.6). Therefore, $\mu = \mathcal{O}(\delta h)$ and $\sigma_f = \mathcal{O}(\delta h)$. We shall keep contributions up to $\mathcal{O}(\delta h)$.

Under the twin assumptions that $\Delta\vec{\lambda}$ and $\|\delta h\|$ are small, the marginalised likelihood is approximately given by

$$\mathcal{L}(\vec{\lambda}) \approx \exp\left(-\frac{1}{2}\left\|s - H(\vec{\lambda}) + \mu(\vec{\lambda})\right\|^2\right) . \tag{6.11}$$

This has a maximum at the best fit parameters $\vec{\lambda}_{\rm bf}$ which satisfy the equation

$$\left\langle \partial_a \left(\mu(\vec{\lambda}_{\rm bf}) - H(\vec{\lambda}_{\rm bf})\right) \middle| s - H(\vec{\lambda}_{\rm bf}) + \mu(\vec{\lambda}_{\rm bf}) \right\rangle = 0 . \tag{6.12}$$

Using $s = n + h(\vec{\lambda}_{\rm tr})$, and expanding to leading order in $\Delta\vec{\lambda}$ and δh gives

$$\begin{aligned}
\left\langle -\partial_a \left(\mu(\vec{\lambda}_{\rm bf}) - H(\vec{\lambda}_{\rm bf})\right) \middle| n + h(\vec{\lambda}_{\rm tr}) - H(\vec{\lambda}_{\rm bf}) + \mu(\vec{\lambda}_{\rm bf}) \right\rangle &= 0 , \quad (6.13)\\
\left\langle -\partial_a \left(\mu(\vec{\lambda}_{\rm bf}) - H(\vec{\lambda}_{\rm bf})\right) \middle| n - \delta h(\vec{\lambda}_{\rm tr}) - \Delta\vec{\lambda}^b \partial_b H(\vec{\lambda}_{\rm bf}) + \mu(\vec{\lambda}_{\rm bf}) \right\rangle &= 0 . \quad (6.14)
\end{aligned}$$

This expression can be rearranged to find $\Delta\vec{\lambda}$, dropping all terms second order in small quantities,

$$\begin{aligned}
\Delta\vec{\lambda}^a &= \left(\Gamma^{-1}\right)^{ab} \left\langle n \middle| \partial_b \left(H(\vec{\lambda}_{\rm bf}) - \mu(\vec{\lambda}_{\rm bf})\right) \right\rangle \\
&\quad - \left(\Gamma^{-1}\right)^{ab} \left\langle \delta h(\vec{\lambda}_{\rm tr}) \middle| \partial_b H(\vec{\lambda}_{\rm bf}) \right\rangle \\
&\quad + \left(\Gamma^{-1}\right)^{ab} \left\langle \mu(\vec{\lambda}_{\rm bf}) \middle| \partial_b H(\vec{\lambda}_{\rm bf}) \right\rangle .
\end{aligned} \tag{6.15}$$

Therefore, at leading order, the shift between the best fit and true parameters for the marginalised likelihood consists of three terms: the noise and model errors from before, and a new shift arising from the marginalisation, $\Delta_{\rm marg}\vec{\lambda}^a = \left(\Gamma^{-1}\right)^{ab} \langle \mu(\vec{\lambda}_{\rm bf}) | \partial_b H(\vec{\lambda}_{\rm bf}) \rangle$. The expression for the model and marginalisation errors are similar and appear with opposite signs (as would be hoped since the maginalised likelihood was designed to remove the model error) so the remaining model error is proportional to $\delta h(\vec{\lambda}_{\rm bf}) - \mu(\vec{\lambda}_{\rm bf})$ (integrated inside the inner product).

If the training set is dense in the region of the peak, and the hyperparameters have been correctly estimated, it is reasonable to assume that the GPR interpolant of the waveform difference performs well, and we have $\langle \delta h(\vec{\lambda}) - \mu(\vec{\lambda}) | \cdot \rangle \approx 0$. Under these

conditions the marginalised likelihood removes the systematic model error from the parameter estimates. In reality the interpolation is not perfect, and the method is limited by the available information in the training set, so that a residual model error proportional to $\langle \delta h(\vec{\lambda}_{\rm bf}) - \mu(\vec{\lambda}_{\rm bf})|\cdot\rangle$ remains.

6.2 The limit of large signal-to-noise

As first pointed out by Cutler and Vallisneri (2007), the systematic error associated with the inaccurate model used in the approximate likelihood is independent of the SNR, whereas the random error associated with the noise realisation decreases with increasing SNR. Therefore, there exists a critical SNR for the approximate likelihood above which the systematic model error dominates the random noise error. If the approximate likelihood is used to infer the parameters of a source with an SNR close to or above this critical value then the inferred parameters are significantly and systematically biased. In this section we examine the behaviour of all three likelihood functions for large SNR and show that the marginalised likelihood does *not* suffer from this problem even in the limit of infinite SNR. Therefore, parameter estimates obtained using the marginalised likelihood can always be trusted.

In this section in order to ease the process of taking the limit of large SNR all waveforms are understood to be normalised such that $\|h(\vec{\lambda})\| = 1$, and the amplitude is taken out as a prefactor, so the full signal is $Ah(\vec{\lambda})$. In addition we will assume for simplicity that the measured value of A is equal to the true value for the signal; this has no effect our final result.

The *exact likelihood* Eq. (5.4) is given by

$$L'(\vec{\lambda}) \propto \exp\left(-\frac{1}{2}\left\|s - Ah(\vec{\lambda})\right\|^2\right) . \tag{6.16}$$

The measured data is given by $s = n + Ah(\vec{\lambda}_{\rm tr})$, and the exact likelihood is peaked at $\vec{\lambda}_{\rm bf} = \vec{\lambda}_{\rm tr} + \Delta\vec{\lambda}$, where (see Eq. (6.5))

$$\Delta\vec{\lambda}^a = \frac{1}{A}\left(\Sigma^{-1}\right)^{ab}\left\langle n\middle|\partial_b h(\vec{\lambda}_{\rm bf})\right\rangle . \tag{6.17}$$

In this section, the FIM Σ_{ab} is defined in terms of the normalised waveforms, i.e. Σ_{ab} is independent of A; this is done so that all of the dependence on A remains explicit. The exact likelihood evaluated on the true parameters is given by

$$L'(\vec{\lambda}_{\rm tr}) \propto \exp\left(-\frac{1}{2}\|n\|^2\right) . \tag{6.18}$$

The exact likelihood evaluated on the best-fit parameters is given by

$$L'(\vec{\lambda}_{\rm bf}) \propto \exp\left[-\frac{1}{2}\left\|n + A\left(h(\vec{\lambda}_{\rm tr}) - h(\vec{\lambda}_{\rm bf})\right)\right\|^2\right] . \tag{6.19}$$

The ratio of these two likelihood values is denoted $R_{\rm exact} = L'(\vec{\lambda}_{\rm tr})/L'(\vec{\lambda}_{\rm bf})$. Expanding the difference $h(\vec{\lambda}_{\rm tr}) - h(\vec{\lambda}_{\rm bf})$ in the above equation to leading order in $\Delta\vec{\lambda}$ gives

$$\ln R_{\rm exact} = -\frac{1}{2}\left(\Sigma^{-1}\right)^{ab}\left\langle n\middle|\partial_a h(\vec{\lambda}_{\rm bf})\right\rangle\left\langle n\middle|\partial_b h(\vec{\lambda}_{\rm bf})\right\rangle . \tag{6.20}$$

The quantity $R_{\rm exact}$ is the factor by which the likelihood of the true parameters is suppressed with respect to the peak likelihood. From Eq. (6.20) it can be seen that this factor is a random variable dependent on the noise realisation n; the expectation of this random variable is given by Cutler and Flanagan (1994*b*)

$$\overline{\ln R_{\rm exact}} = -\frac{1}{2} . \tag{6.21}$$

Both Eqs. (6.21) and Eq. (6.20) are independent of the signal amplitude A, and hence are unchanged by taking the limit of large SNR, $A \to \infty$. Therefore (as one might have expected) the exact likelihood evaluated at $\vec{\lambda}_{\rm tr}$ remains finite in this limit and the true parameters are never completely excluded from the posterior at any value of the SNR.

The *approximate likelihood* Eq. (5.7) is given by

$$L(\vec{\lambda}) \propto \exp\left(-\frac{1}{2}\left\|s - AH(\vec{\lambda})\right\|^2\right) , \tag{6.22}$$

The approximate likelihood is peaked at $\vec{\lambda}_{\rm bf} = \vec{\lambda}_{\rm tr} + \Delta\vec{\lambda}$, where (see Eq. (6.9))

$$\Delta\vec{\lambda}^a = \frac{1}{A}\left(\Gamma^{-1}\right)^{ab}\left\langle n\middle|\partial_b H(\vec{\lambda}_{\rm bf})\right\rangle - \left(\Gamma^{-1}\right)^{ab}\left\langle \delta h(\vec{\lambda}_{\rm tr})\middle|\partial_b H(\vec{\lambda}_{\rm bf})\right\rangle . \tag{6.23}$$

The FIM Γ_{ab} is also here defined to be independent of A. The approximate likelihood evaluated on the true parameters is given by

$$\begin{aligned} L(\vec{\lambda}_{\rm tr}) &\propto \exp\left[-\frac{1}{2}\left\|n + A\left(h(\vec{\lambda}_{\rm tr}) - H(\vec{\lambda}_{\rm tr})\right)\right\|^2\right] \\ &\propto \exp\left(-\frac{1}{2}\left\|n - A\delta h(\vec{\lambda}_{\rm tr})\right\|^2\right) . \end{aligned} \tag{6.24}$$

The approximate likelihood evaluated on the best fit parameters is given by

$$\begin{aligned} L(\vec{\lambda}_{\rm bf}) &\propto \exp\left[-\frac{1}{2}\left\|n - A\left(h(\vec{\lambda}_{\rm tr}) - H(\vec{\lambda}_{\rm bf})\right)\right\|^2\right] \\ &\propto \exp\left[-\frac{1}{2}\left\|n + A\left(\delta h(\vec{\lambda}_{\rm tr}) - \Delta\vec{\lambda}^a\partial_a H(\vec{\lambda}_{\rm bf})\right)\right\|^2\right] , \end{aligned} \tag{6.25}$$

where, as before, the waveform difference has been expanded to leading order in $\Delta\vec{\lambda}$. The ratio of the two likelihood $R_{\rm approx} = L(\vec{\lambda}_{\rm tr})/L(\vec{\lambda}_{\rm bf})$ can be evaluated from Eq. (6.24) and Eq. (6.25), and taking the limit of large SNR gives

$$\begin{aligned}\lim_{A\to\infty} \ln R_{\rm approx} &= -\frac{A^2}{2}\left(\Gamma^{-1}\right)^{ab}\left\langle \delta h(\vec{\lambda}_{\rm tr})\middle|\partial_a H(\vec{\lambda}_{\rm bf})\right\rangle \\ &\quad\times\left\langle \delta h(\vec{\lambda}_{\rm tr})\middle|\partial_b H(\vec{\lambda}_{\rm bf})\right\rangle . \end{aligned} \tag{6.26}$$

Unlike $R_{\rm exact}$, this ratio does not depend on n. This is because in the limit of large SNR, only the terms from the exponents of Eq. (6.24) and Eq. (6.25) proportional to A^2 contribute, and the noise-dependent terms are all proportional to A. Also unlike $R_{\rm exact}$, this ratio does depend on the amplitude and $R_{\rm approx} \to 0$ in the limit of large SNR. Therefore, as anticipated above, the approximate likelihood excludes the true source parameters with complete certainty in the limit of large SNR (unless $\langle \delta h(\vec{\lambda}_{\rm tr})|\cdot\rangle = 0$, in which case the exact likelihood is recovered).

The *marginalised likelihood* Eq. (5.9) is given by

$$\mathcal{L}(\vec{\lambda}) \propto \exp\left(-\frac{1}{2}\frac{\left\|s - AH(\vec{\lambda}) + A\mu(\vec{\lambda})\right\|^2}{1 + A^2\sigma^2(\vec{\lambda})}\right) . \tag{6.27}$$

The marginalised likelihood is peaked at $\vec{\lambda}_{\rm bf} = \vec{\lambda}_{\rm tr} + \Delta\vec{\lambda}$, where, by comparison with Eq. (6.15),

$$\begin{aligned}\Delta\vec{\lambda}^a &= \frac{1}{A}\left(\Gamma^{-1}\right)^{ab}\left\langle n\middle|\partial_b\left(H(\vec{\lambda}_{\rm bf}) - \mu(\vec{\lambda}_{\rm bf})\right)\right\rangle \\ &\quad - \left(\Gamma^{-1}\right)^{ab}\left\langle \delta h(\vec{\lambda}_{\rm tr})\middle|\partial_b H(\vec{\lambda}_{\rm bf})\right\rangle \\ &\quad + \left(\Gamma^{-1}\right)^{ab}\left\langle \mu(\vec{\lambda}_{\rm bf})\middle|\partial_b H(\vec{\lambda}_{\rm bf})\right\rangle . \end{aligned} \tag{6.28}$$

The values of the marginalised likelihood evaluated on the true and best-fit parameters are given by Eq. (6.29) and Eq. (6.30), and the ratio of these two likelihoods is denoted $R_{\rm marg}$,

$$\mathcal{L}(\vec{\lambda}_{\rm tr}) \propto \exp\left(-\frac{1}{2}\frac{\left\|n - A\delta h(\vec{\lambda}_{\rm tr}) + A\mu(\vec{\lambda}_{\rm tr})\right\|^2}{1 + A^2\sigma^2(\vec{\lambda}_{\rm tr})}\right) ; \tag{6.29}$$

$$\mathcal{L}(\vec{\lambda}_{\rm bf}) \propto \exp\left(-\frac{1}{2}\frac{\left\|n - A\delta h(\vec{\lambda}_{\rm tr}) - A\Delta\vec{\lambda}^a\partial_a H(\vec{\lambda}_{\rm bf}) + A\mu(\vec{\lambda}_{\rm bf})\right\|^2}{1 + A^2\sigma^2(\vec{\lambda}_{\rm bf})}\right) ; \tag{6.30}$$

$$\lim_{A\to\infty} \ln R_{\rm marg} = -\frac{1}{2\sigma^2(\vec{\lambda}_{\rm bf})}\left[\left(\Gamma^{-1}\right)^{ab}\left\langle \delta h(\vec{\lambda}_{\rm tr}) - \mu(\vec{\lambda}_{\rm bf})\middle|\partial_a H(\vec{\lambda}_{\rm bf})\right\rangle\left\langle \delta h(\vec{\lambda}_{\rm tr}) - \mu(\vec{\lambda}_{\rm bf})\middle|\partial_b H(\vec{\lambda}_{\rm bf})\right\rangle\right.$$
$$\left. - \left\|\mu(\vec{\lambda}_{\rm bf})\right\|^2 + \left\|\mu(\vec{\lambda}_{\rm tr})\right\|^2 + 2\left\langle \delta h(\vec{\lambda}_{\rm tr})\middle|\mu(\vec{\lambda}_{\rm tr}) - \mu(\vec{\lambda}_{\rm bf})\right\rangle\right] \tag{6.31}$$
$$\approx -\frac{1}{2\sigma^2(\vec{\lambda}_{\rm bf})}\left(\Gamma^{-1}\right)^{ab} \times$$
$$\left\langle \delta h(\vec{\lambda}_{\rm tr}) - \mu(\vec{\lambda}_{\rm bf})\middle|\partial_a H(\vec{\lambda}_{\rm bf})\right\rangle\left\langle \delta h(\vec{\lambda}_{\rm tr}) - \mu(\vec{\lambda}_{\rm bf})\middle|\partial_b H(\vec{\lambda}_{\rm bf})\right\rangle . \tag{6.32}$$

The approximation made in going from Eq. (6.31) to Eq. (6.32) involves dropping terms which are products of small quantities. Because the FIM is a symmetric, positive-definite matrix, the numerator in Eq. (6.32) is a negative number, and hence $R_{\rm marg} < 1$ is required to ensure $\vec{\lambda}_{\rm bf}$ is the peak of the likelihood.

As was the case with $R_{\rm approx}$, this expression for $R_{\rm marg}$ does not depend on the noise. However, unlike $R_{\rm approx}$ the expression for $R_{\rm marg}$ also does not depend on the amplitude A. Therefore, in the limit that the SNR becomes large $R_{\rm marg}$ tends to a constant value which depends quadratically on $\langle \delta h(\vec{\lambda}_{\rm tr}) - \mu(\vec{\lambda}_{\rm bf})|\cdot\rangle$. As the SNR increases, the true parameters are not excluded from the marginalised likelihood, instead the likelihood distribution tends to a constant distribution (i.e. no dependence on n), and the ratio by which the true parameters are disfavoured compared to the best fit parameters is set by the ability of the GPR to recover the true waveform difference.

Intuitively, the reason the marginalised likelihood is able to achieve this useful behaviour, even if the true waveform difference is not perfectly recovered by the GPR interpolation (i.e. $\langle \delta h(\vec{\lambda}_{\rm tr}) - \mu(\vec{\lambda}_{\rm bf})|\cdot\rangle \neq 0$), is due to the way the hyperparameters in the covariance function are chosen. The hyperparameters were fixed to their optimum values by maximising the hyperlikelihood for the training set (as described in Sec. 5.3). During this process the overall scale hyperparameter σ_f gains a dependence on the amplitude proportional to A^2. Hence the GPR uncertainty $\sigma^2(\vec{\lambda})$ is also proportional to A^2. As can be seen from Eq. (6.27), in the limit of large SNR the amplitude dependence cancels in the exponential and the marginalised likelihood tends to a constant distribution. Therefore, the marginalised likelihood never excludes the true source parameters from the final posterior with complete certainty.

6.3 Limits of the marginalised likelihood across parameter space

In this section we examine the behaviour of the marginalised likelihood in the limit of being far from any training points and being at a training point.

First we examine the behaviour of the marginalised likelihood in the former case, at a large distance ($\tau^2 \gg 1$) from any of the points in the training set. From Eq. (5.25) it can be seen that well outside of the training set $\mu(\vec{\lambda}) \to 0$ and $\sigma^2(\vec{\lambda}) \to \sigma_f^2$. Therefore, from Eq. (5.28), the log marginalised likelihood tends to

$$\ln \mathcal{L}(\vec{\lambda}) \to \frac{\ln L(\vec{\lambda})}{1 + \sum_{i,j} \boldsymbol{K}_{ij} \left\langle \delta h(\vec{\lambda}_i) \middle| \delta h(\vec{\lambda}_j) \right\rangle} \,. \tag{6.33}$$

Well outside of the training set the marginalised likelihood $\ln \mathcal{L}(\vec{\lambda})$ recovers the standard, approximate likelihood $L(\vec{\lambda})$ up to a constant factor. This constant factor is one plus a linear combination of the overlap integrals of all the waveform differences in the training set. Since the denominator in Eq. (6.33) is always greater than unity (this is ensured by the positive-definite property of the covariance matrix), it broadens any peak in the likelihood outside of the training set. The amount of the broadening is set by the magnitude of the waveform differences in the training set via the overlap matrix $\langle \delta h(\vec{\lambda}_i) | \delta h(\vec{\lambda}_j) \rangle$. This is the behaviour that would be expected; in the absence of any accurate waveforms the parameter uncertainties obtained from the approximate waveforms should be multiplied by a constant factor depending upon our level of belief in the accuracy of the approximate waveform model. In turn, our level of belief in the accuracy of the approximate waveform is learnt from the training set in the process of training the GP.

We now consider the behaviour of the marginalised likelihood evaluated at one of the training set points $\vec{\lambda}_\ell$. First, consider the case where $\sigma_n = 0$. In this case, the interpolated waveform difference, from Eq. (5.24), at $\vec{\lambda}_\ell$ recovers the true waveform difference, and the GPR uncertainty, from Eq. (5.25), vanishes at $\vec{\lambda}_\ell$;

$$\mu(\vec{\lambda}_\ell) = \delta h(\vec{\lambda}_\ell) \,, \tag{6.34}$$

$$\sigma^2(\vec{\lambda}_\ell) = 0 \,. \tag{6.35}$$

Therefore the marginalised likelihood in Eq. (5.28) recovers the exact likelihood with no additional broadening.

$$\mathcal{L}(\vec{\lambda}_\ell) = L'(\vec{\lambda}_\ell) \,. \tag{6.36}$$

This is also the behaviour that would be expected; at a point in parameter space where the accurate waveform is known, the accurate likelihood is recovered.

If $\sigma_n \neq 0$, then Eq. (6.34) and Eq. (6.35) become

$$\begin{aligned} \mu(\vec{\lambda}_\ell) &= \delta h(\vec{\lambda}_\ell) - \sigma_n^2 \sum_i k(\vec{\lambda}_i, \vec{\lambda}_\ell)\delta h(\vec{\lambda}_i) + \mathcal{O}(\sigma_n^4)\,, \\ \sigma^2(\vec{\lambda}_\ell) &= \sigma_n^2 \sum_i k(\vec{\lambda}_i, \vec{\lambda}_\ell)k(\vec{\lambda}_i, \vec{\lambda}_\ell) + \mathcal{O}(\sigma_n^4)\,. \end{aligned} \tag{6.37}$$

In this case any peak in the marginalised likelihood will be slightly shifted and broadened relative to the peak in the accurate likelihood by an amount consistent with the uncertainty σ_n in the accurate waveform model.

6.4 The bias in the likelihood

A common way to quantify the performance of parameter estimation is via the probability-probability (P-P) plot. The P-P plot shows the probability that the true source parameters will lie in a given confidence interval estimated from the detector data, against the value of the confidence interval. In the ideal, unbiased, case the P-P should be a diagonal line; i.e., $x\%$ of the time the true source parameters should lie within the $x\%$ confidence interval. However, there are a variety of effects that can cause the P-P plot to deviate from this ideal. For example, use of a greedy algorithm to build a multi-dimensional confidence interval from a kD-tree constructed from a random sample of points from a distribution (this problem was discussed in the context of sky-localisation by Sidery et al. (2014)), deviations between the waveform model and the true signal due to a breakdown of general relativity in the strong field (the case of undetectable deviations from general relativity, the so-called "stealth-bias", was considered in Vallisneri and Yunes (2013)), and mis-estimating the noise properties of the detector can all cause the P-P plot to deviate from a ideal diagonal line. However, the cause of biased parameter estimation that we will consider in this work is the presence of inaccuracies in the waveform model used to analyse the data (Cutler and Vallisneri, 2007). If such a systematic error is present the returned confidence intervals from a parameter estimation study will be shifted away from the true parameters making it less likely that the confidence interval contains the true parameters. Therefore the P-P plot will "sag" below the ideal diagonal line.

The marginalised likelihood discussed above uses Gaussian processes (GPs) to fold in extra information from a small *training set* of accurate waveforms, e.g. numerical

relativity (NR) waveforms. Accurate here refers to how well these waveforms represent solutions of the Einstein field equations. Numerical relativity waveforms are not perfectly accurate, but they are the best solutions currently available and inaccuracies in them can be folded into the GP analysis. If astrophysical gravitational waves are governed by a theory other than general relativity, these waveforms will not be accurate representations of reality. This will also lead to a bias, but one that is harder to quantify without knowing the true theory of gravity. Here we proceed assuming general relativity is correct and look only at biases from model uncertainties. Once observations are made this assumption could be revisited if evidence arises for departures from general relativity.

The GP marginalised likelihood in general shifts the best fit parameters closer to the true parameters and broadens the peak in the posterior, making it more likely that a given confidence contour contains the true parameters. Therefore, it would be expected that parameter estimates obtained using the marginalised likelihood would exhibit less of a bias, and the P-P plots would exhibit less of a "sag". However, the Gaussian process regression (GPR) which underlies the marginalised likelihood makes some assumptions about how the error in the waveform model varies over parameter space. In this section, we investigate the P-P plots both in the case where these assumptions turn out to be correct, and, more importantly, when they are incorrect.

There are two main results presented in this section. The first is a derivation of an analytic expression for the expected sag in a P-P plot arising from waveform uncertainties. This is derived under the assumption that the waveform error is small so that we can use the linear signal approximation. The second is that the use of the marginalised likelihood constructed via Gaussian process regression to analyse data leads to a reduction in the size of the deviation from the diagonal line. The sag is removed completely if the true waveform errors are drawn from the same model used to construct the marginalised likelihood. However, even when the errors follow a different distribution, the marginalised likelihood leads to a reduction in the sag.

This section is organised as follows. Sec. 6.4.2 presents derivations of analytic expressions for the P-P plots for both the standard and marginalised likelihoods for a variety of possible waveform errors. Sec. 6.4.3 describes the numerical simulations that were performed to back-up the analytic results in Sec.6.4.2. Finally Sec. 6.4.4 contains a discussion of the results and concluding remarks.

6.4.1 The posterior maximum estimator

The aim of a parameter estimation study given measured data s, is to estimate the posterior probability on the parameters, $P(\vec{\lambda}|s)$. In the case of a signal with true parameter $\vec{\lambda}_0$, and stationary, Gaussian, additive noise n the measured data is given by $s = h(\vec{\lambda}_0) + n$ and the likelihood is given by Eq. 5.4.

The *exact likelihood* (defined in Eq. 5.4) will have a posterior peaked at a position given by Eq. 6.4; the shift from the true parameters depends only on the noise realisations. The *approximate likelihood* (defined in Eq. 5.7) will not in general agree with posterior distributions obtained from the exact likelihood. Denote by $\vec{\lambda}_{\rm exact}$ the best fit parameters obtained from Eq. (5.4) and $\vec{\lambda}_{\rm approx}$ the best fit parameters obtained from Eq. (5.7). If both the waveform difference and the parameter shift $\Delta\vec{\lambda} \equiv \vec{\lambda}_{\rm approx} - \vec{\lambda}_{\rm exact}$ are small quantities, $\mathcal{O}(\epsilon)$, then by expanding in ϵ an expression for the shift was found in Eq. 6.10. The shift in best-fit parameters to linear order in ϵ was obtained in Cutler and Vallisneri (2007) as $\Delta\vec{\lambda} \equiv \Delta\vec{\lambda}_1$ where

$$\Delta\lambda_1^a = -\left(\Sigma^{-1}\right)^{ab} \langle \delta h(\vec{\lambda}_0) | \partial_b H(\vec{\lambda}_0) \rangle \,, \tag{6.38}$$

$\Sigma_{ab} = \left\langle \partial_a H(\vec{\lambda}) | \partial_b H(\vec{\lambda}) \right\rangle$, and $\partial_a = \partial/\partial\lambda^a|_{\vec{\lambda}=\vec{\lambda}_0}$. For completeness we include a derivation of this result, and an extension of it to quadratic order, in Appendix D.

From Eq. 6.38 (and Eq. 19 in Appendix D) it can be seen that the systematic shift in parameters caused by using the approximate likelihood is independent of the signal-to-noise ratio. This fact was observed by Cutler and Vallisneri (2007), and since the statistical errors that arise from detector noise decrease with increasing signal-to-noise this means that the systematic shift is most important for the loudest sources.

When using the approximate likelihood in Eq. (5.7) to characterise a single source one would usually use the condition that the systematic error due to the model uncertainty is less than the random error arising from noise to determine if the model is "good enough". This condition ensures that the true parameters will be consistent with the posterior — the amount by which the systematic error shifts the peak of the posterior is less than the typical posterior width.

However, whilst this condition ensures that the true parameters will always be consistent with the posterior, on average they will be further from the centre of the posterior and hence lie at a lower significance than they should. This starts to become important when observing a population of sources (as we hope will be the case for

Advanced LIGO). Even small systematic shifts may lead one to make incorrect inferences about the properties of the population. This can be understood by imagining that we observe a NS-NS binary with identical astrophysical parameters n independent times with Advanced LIGO. The error in the combined estimate for the mean mass of the population is the error in each measurement divided by $\sqrt{n}$. Therefore even if the systematic model error is insignificant for making inferences regarding a single binary it becomes increasingly significant for inferences regarding populations as new sources are added.

The importance of the model errors for LIGO observations of NS-NS binaries was considered by Canitrot (2001). Model error effects could also be seen in the parameter estimation analysis of the "big-dog" blind injection (Abadie et al., 2012*c*). In that case, the recovered masses for the compact binary injection were significantly biased (in part) by the fact that different signal models were used for the injection and parameter estimation. This indicates the importance of considering how to incorporate model uncertainties in parameter estimation before the advanced detector era begins. A detailed investigation of parameter estimation on various injections into data from the LIGO/Virgo interferometers and employing a range of different models for the analysis was carried out in Aasi et al. (2013*b*). These results clearly show how the analysis of the same data using two different models can give mutually inconsistent results.

The marginalised likelihood, discussed in Sec. 5.3.1, attempted to account for the systematic error in the posterior, and hence remove the bias. The expression for the marginalised likelihood is Eq. 5.28. In the following sections we will compare the bias, quantified via the P-P plot, for the exact, approximate and marginalised likelihoods.

6.4.2 Analytic calculation of the P-P plot

In the limit of high signal-to-noise the posterior probability distribution obtained in the analysis of data from a detector will be strongly peaked in the vicinity of the true parameters. Within the vicinity of this peak it is reasonable to expand both the exact and approximate signal models in the usual linear signal approximation (LSA), i.e.

$$\begin{aligned} h(\vec{\lambda}) &= h(\vec{\lambda}_0) + \Delta\vec{\lambda}^a \partial_a h(\vec{\lambda}_0)\,, \\ H(\vec{\lambda}) &= H(\vec{\lambda}_0) + \Delta\vec{\lambda}^a \partial_a H(\vec{\lambda}_0)\,. \end{aligned} \tag{6.39}$$

where $\vec{\lambda}_0$ denotes the parameter values of the true signal, $\vec{\lambda}$ denotes the parameter values at which we want to evaluate the signal or likelihood and $\Delta\vec{\lambda} = \vec{\lambda} - \vec{\lambda}_0$. This LSA is the usual approximation made in the derivation of the Fisher Matrix and the approximation used in the derivation of Eqs. (6.38) and (19).

We are interested in predicting the "sag" that would be expected in a P-P plot. If we use an approximate waveform model to compute the posterior, then we would expect some bias in the recovered parameters and a sag in the P-P plot - on average the true parameters would be further away from the peak of the posterior than we would expect, and so fewer injections would be recovered at a given significance level.

6.4.2.1 The exact likelihood

The exact likelihood, by definition, will give a diagonal unbiased P-P plot. However we will re-derive this obvious result to shed light on the calculations that follow.

The *Exact Likelihood* is given by Eq. (5.4). The measured data is assumed to consist of a signal with true parameters $\vec{\lambda}_0$ and additive Gaussian noise; $s = h(\vec{\lambda}_0) + n$. In the limit of high signal-to-noise, the difference between two nearby signals in parameter space may be expanded using the LSA,

$$\begin{aligned} L'(\vec{\lambda}) &\propto \exp\left(-\frac{1}{2}\left\langle n - \Delta\lambda^a \partial_a h | n - \Delta\lambda^a \partial_a h\right\rangle\right), \\ &= \exp\left(-\frac{1}{2}\left[\langle n|n\rangle - 2\Delta\lambda^a \langle n|\partial_a h\rangle + \Delta\lambda^a \Delta\lambda^b S_{ab}\right]\right), \end{aligned} \tag{6.40}$$

where the exact Fisher matrix is $S_{ab} = \langle \partial_a h | \partial_b h \rangle$. Since the Fisher matrix is symmetric by construction, we may adopt new coordinates in parameter space $\tilde{\Delta\lambda}^a = Q^a_b \Delta\lambda^b$ such that the Fisher matrix in these coordinates becomes diagonal, $S_{ab} = Q^p_a Q^q_b \delta_{pq}$. This amounts to rescaling the coordinate axes such that the iso-probability contour, which originally was an n-ellipsoid, becomes an n-sphere. Derivatives with respect to the new coordinates will be denoted with a tilde, $\partial_a h = Q^b_a \tilde{\partial}_b h$. In these new coordinates the likelihood separates to become

$$L'(\vec{\lambda}) \propto \prod_x \exp\left(-\frac{1}{2}\left(\tilde{\Delta\lambda}^x - \left\langle n|\tilde{\partial}_x h\right\rangle\right)^2\right). \tag{6.41}$$

In order to exploit the spherical symmetry about the peak in the rescaled parameters we adopt (n-dimensional) spherical coordinates centred on the peak; the radial coordinate given by $r^2 = \sum_x (\tilde{\Delta\lambda}^x - < n|\tilde{\partial}_x h >)^2$. The significance of the true parameters is

given by the volume of the posterior that is "closer to the peak", i.e., that has higher posterior weight than the true parameters,

$$\begin{aligned} \text{sig} &= \frac{\int_0^R \mathrm{d}r\; r^{N-1}\exp(-r^2/2)}{\int_0^\infty \mathrm{d}r\; r^{N-1}\exp(-r^2/2)} \\ &= 1 - \frac{\Gamma\left(\frac{N}{2},\frac{R^2}{2}\right)}{\Gamma\left(\frac{N}{2}\right)} = 1 - \bar{\Gamma}\left(\frac{N}{2},\frac{R^2}{2}\right), \end{aligned} \tag{6.42}$$

where $\Gamma(x,y)$ is the incomplete Gamma function,

$$\Gamma(x,y) = \int_y^\infty t^{x-1}\mathrm{e}^{-t}\mathrm{d}t\,, \tag{6.43}$$

$\Gamma(x) = \Gamma(x,0)$ is the complete Gamma function, and $\bar{\Gamma}(x,y)$ is the regularised incomplete gamma function defined via the last equality in Eq. (6.42). In Eq. (6.42) the assumption has been made that the prior distribution on the parameters may be approximated as a constant across the width of the peak; this is reasonable in the high signal-to-noise limit when the posterior is narrow. The quantity R^2 is given by

$$R^2 = \sum_x \left\langle n|\tilde{\partial}_x h\right\rangle^2 = \left(S^{-1}\right)^{ab}\left\langle n|\partial_a h\right\rangle\left\langle n|\partial_b h\right\rangle\,, \tag{6.44}$$

and is distributed as a χ^2 random variable with $N = \dim(\vec{\lambda})$ degrees of freedom. The inverse regularised incomplete gamma function is defined via $y = \bar{\Gamma}(x, \bar{\Gamma}^{-1}(x,y))$. The quantity on the ordinate axis of a a standard P-P plot is the probability that the true parameters lie within a given significance, $P(\text{sig} < X)$. From Eq. (6.44) it may be seen that this can be rewritten as a cumulative probability of the random variable R^2;

$$P(\text{sig} < X) = 1 - P\left(R^2 < 2\bar{\Gamma}^{-1}\left(\frac{N}{2}, 1-X\right)\right)\,. \tag{6.45}$$

The cumulative distribution function of the χ^2 distribution is the regularised Gamma function, $P(R^2 < y) = \bar{\Gamma}(N/2, y/2)$. Using this to evaluate Eq. 6.45 gives the expected, unbiased diagonal form of the P-P plot for the exact likelihood;

$$P(\text{sig} < X) = 1 - (1 - X) = X\,. \tag{6.46}$$

This diagonal P-P plot is shown in the dotted black curve in the left-hand panel of Fig. 6.1. The fact that the PP plot for the exact likelihood is always diagonal follows from the definition of the likelihood, and this remains true even if the LSA fails. The derivation just presented assumes the LSA in order to make it resemble as closely as possible the upcoming derivation for the approximate likelihood.

6.4.2.2 The approximate likelihood

We now move on to the more interesting case when we have biased parameter estimation from using the approximate likelihood. As mentioned in the introduction we expect to obtain a P-P plot that is "sagging" below the diagonal indicating the bias. We first treat the simple case where the waveform model depends on just a single parameter, $\vec{\lambda} = \theta$, where the expression for the P-P plot is given in terms of the inverse error function, $\mathrm{erf}^{-1}(x)$. A treatment will then be given for the general N dimensional case in which the expression for the P-P plot is given in terms of the MarcumQ function, $Q_N(x, y)$, along with a illustration of how this reduces to the 1D result.

The *Approximate Likelihood* is given by Eq. (5.7). We assume the approximate model is "nearly" correct and use the LSA to expand signals that are nearby in parameter space. As before, denoting the waveform difference by $\delta h(\vec{\lambda}) = H(\vec{\lambda}) - h(\vec{\lambda})$, we have

$$\begin{aligned} L(\vec{\lambda}) &\propto \exp\left(-\frac{1}{2}\left\langle n - \delta h(\vec{\lambda}_0) - \Delta\lambda^a \partial_a H \middle| \ldots \right\rangle\right) \\ &= \exp\left(-\frac{1}{2}\left[\left\langle n - \delta h(\vec{\lambda}_0)\middle| \ldots \right\rangle\right.\right. \\ &\qquad \left.\left. -2\Delta\lambda^a \left\langle n - \delta h(\vec{\lambda}_0) | \partial_a H \right\rangle + \Delta\lambda^a \Delta\lambda^b \Sigma_{ab}\right]\right) , \end{aligned} \tag{6.47}$$

where the ellipsis in the right hand entry in the inner product denotes a repeat of the left hand entry and the approximate Fisher matrix is $\Sigma_{ab} = \langle \partial_a H | \partial_b H \rangle$. As before coordinates which diagonalise the Fisher matrix Σ_{ab} may be adopted, which give the following separated expression for the approximate likelihood,

$$L(\vec{\lambda}) \propto \prod_x \exp\left(-\frac{1}{2}\left(\tilde{\Delta\lambda}^x - \left\langle n - \delta h(\vec{\lambda}_0) | \tilde{\partial}_x H \right\rangle\right)^2\right) . \tag{6.48}$$

Example for one dimensional parameter space:

If the waveform depends on only one unknown parameter, $\vec{\lambda} = \lambda$, Eq. (6.48) becomes

$$L(\theta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}\left(\Delta\theta - \mu\right)^2\right] , \tag{6.49}$$

where

$$\frac{1}{\sigma^2} = \left\langle \frac{\mathrm{d}H}{\mathrm{d}\lambda}\Big|_{\lambda=\lambda_0} \middle| \frac{\mathrm{d}H}{\mathrm{d}\lambda}\Big|_{\lambda=\lambda_0} \right\rangle , \tag{6.50}$$

$$\mu \;=\; \sigma^2 \left\langle n - \delta h(\lambda_0) \Big| \frac{\mathrm{d}H}{\mathrm{d}\lambda}\Big|_{\lambda=\lambda_0} \right\rangle , \tag{6.51}$$

and we have included the correct normalisation of the posterior. The true parameter value is at $\Delta\theta = 0$ and the points with larger posterior weight than the true parameters lie in the range $0 < \Delta\theta < 2\mu$ when $\mu > 0$ or in the range $2\mu < \Delta\theta < 0$ when $\mu < 0$. The significance at which the true parameters lie is therefore

$$\int_0^{2\mu} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}\left(\Delta\theta - \mu\right)^2\right] \mathrm{d}\Delta\theta = \mathrm{erf}\left(\frac{|\mu|}{\sqrt{2}\sigma}\right) , \tag{6.52}$$

where

$$\mathrm{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z \mathrm{e}^{-t^2} \mathrm{d}t \tag{6.53}$$

is the usual error function. The quantity μ defined above depends on the particular realisation of the noise. We want to know the fraction of times, over many realisations of the noise, that the true parameters will lie within a certain significance contour. This is just

$$P(\mathrm{sig} < X) = P\left(\frac{|\mu|}{\sqrt{2}\sigma} < \mathrm{erf}^{-1}(X)\right) . \tag{6.54}$$

The quantity $\mu/(\sqrt{2}\sigma)$ is distributed as a Gaussian with mean $\tilde{\mu} = \sigma \left\langle \Delta h(\lambda_0) | \mathrm{d}H/\mathrm{d}\lambda \right\rangle / \sqrt{2}$ and variance $1/2$ and so

$$P(\mathrm{sig} < X) = \frac{1}{2}\mathrm{erf}\left(\mathrm{erf}^{-1}(X) - \tilde{\mu}\right) + \frac{1}{2}\mathrm{erf}\left(\mathrm{erf}^{-1}(X) + \tilde{\mu}\right) . \tag{6.55}$$

In the special case where the approximate waveform model and the exact waveform model are the same, we have $\tilde{\mu} = 0$ and recover the expected unbiased result from Sec. 6.4.2.1;

$$P(\mathrm{sig} < X) = X \, . \tag{6.56}$$

This derivation assumed that $\tilde{\mu}$ was constant, but in practice this will vary from signal to signal. If we denote the probability distribution function for $\tilde{\mu}$ over the astrophysical population by $f(\tilde{\mu})$, the generalisation of Eq. (6.55) can be seen straightforwardly to be

$$\begin{aligned} P(\mathrm{sig} < X) = \int \Bigg[& \frac{1}{2}\mathrm{erf}\left(\mathrm{erf}^{-1}(X) - \tilde{\mu}\right) \\ & + \frac{1}{2}\mathrm{erf}\left(\mathrm{erf}^{-1}(X) + \tilde{\mu}\right) \Bigg] f(\tilde{\mu}) \mathrm{d}\tilde{\mu} \, . \end{aligned} \tag{6.57}$$

Parameter space of arbitrary dimension

We will now generalise the expression for the P-P plot sag in a one-dimensional parameter space, given in Eq. (6.55), to arbitrary numbers of parameters. Identical manipulations to those performed on the exact likelihood yields the same expression for the significance obtained in Eq. (6.42),

$$\mathrm{sig} = 1 - \bar{\Gamma}\left(\frac{N}{2}, \frac{R^2}{2}\right) , \tag{6.58}$$

except this time R^2 is a random variable given by

$$\begin{aligned} R^2 &= \sum_x \left\langle n - \delta h(\vec{\lambda}_0) | \tilde{\partial}_x H \right\rangle^2 \\ &= \left(\Sigma^{-1}\right)^{ab} \left\langle n - \delta h(\vec{\lambda}_0) | \partial_a H \right\rangle \left\langle n - \delta h(\vec{\lambda}_0) | \partial_b H \right\rangle . \end{aligned} \tag{6.59}$$

If $\delta h(\vec{\lambda})$ is constant across parameter space, R^2 is now a non-central χ^2 random variable with N degrees of freedom and non-centrality parameter

$$\Lambda = \left(\Sigma^{-1}\right)^{ab} \left\langle \delta h(\vec{\lambda}_0) | \partial_a H \right\rangle \left\langle \delta h(\vec{\lambda}_0) | \partial_b H \right\rangle . \tag{6.60}$$

As before the expression for the P-P plot is given in terms of the CDF of the distribution of the random variable R^2. The CDF of the non-central χ^2 distribution is the Marcum-Q function, $P(R^2 < y) = Q_{N/2}(\sqrt{\Lambda}, \sqrt{y})$,

$$P(\mathrm{sig} < X) = 1 - P\left(R^2 < 2\bar{\Gamma}^{-1}\left(\frac{N}{2}, 1 - X\right)\right) \tag{6.61}$$

$$= 1 - Q_{\frac{N}{2}}\left(\sqrt{\Lambda}, \sqrt{2\bar{\Gamma}^{-1}\left(\frac{N}{2}, 1 - X\right)}\right) . \tag{6.62}$$

This is an analytic approximation to the P-P plot in the LSA and in the case of a constant waveform difference over parameter space; this function is plotted as a dotted black line in Fig. 6.1. In this case the P-P plot always sags below the diagonal indicating biased parameter recovery.

If $\delta h(\vec{\lambda})$ is not constant over parameter space, the generalisation of this result takes the same form as Eq. (6.57), but with the term in square brackets replaced by Eq. (6.62) and with $f(\tilde{\mu})$ replaced by the corresponding probability distribution function for Λ. For example, in the case that $\delta h(\vec{\lambda})$ is distributed at different times and at different points in parameter space as an uncorrelated, zero-mean Gaussian with variance in each

component of ϵ^2 (i.e. $\delta h(\vec{\lambda}_0) \sim \mathcal{N}(0, \epsilon^2)$) then the quantities $\langle \delta h(\vec{\lambda}_0) | \partial_a H \rangle$ are distributed as $N(0, \Sigma)$ and we see that Λ is distributed as ϵ^2 times a χ^2 distribution with N degrees of freedom with probability distribution function

$$f(\Lambda) = \frac{1}{2^{\frac{N}{2}}\Gamma(N/2)\epsilon^N} \Lambda^{\frac{N}{2}-1} \mathrm{e}^{-\frac{\Lambda}{2\epsilon^2}}. \tag{6.63}$$

Writing $x_u^2 = 2\bar{\Gamma}^{-1}\left(\frac{N}{2}, 1-X\right)$ we must evaluate

$$\begin{aligned}
P(\mathrm{sig} < X) =& \int_0^\infty \left[\frac{\Lambda^{\frac{N}{2}-1}\mathrm{e}^{-\frac{\Lambda}{2\epsilon^2}}}{2^{\frac{N}{2}}\Gamma\left(\frac{N}{2}\right)\epsilon^N} \int_0^{x_u} x \left(\frac{x}{\sqrt{\Lambda}}\right)^{\frac{N}{2}-1} \mathrm{e}^{-\frac{1}{2}(x^2+\Lambda)} I_{\frac{N}{2}-1}(\sqrt{\Lambda}x)\mathrm{dx} \right] \mathrm{d}\Lambda \\
=& \frac{1}{(2\epsilon)^{\frac{N}{2}-1}\Gamma\left(\frac{N}{2}\right)} \int_0^{x_u} \left[x^{\frac{N}{2}} \mathrm{e}^{-\frac{x^2}{2}} \int_0^\infty y^{\frac{N}{2}} \mathrm{e}^{-\frac{1}{2}(1+\epsilon^2)y^2} I_{\frac{N}{2}-1}(\epsilon x y)\mathrm{d}y \right] \mathrm{d}x \\
=& \frac{1}{(2\epsilon)^{\frac{N}{2}-1}\Gamma\left(\frac{N}{2}\right)(1+\epsilon^2)^{\frac{1}{2}+\frac{N}{4}}} \int_0^{x_u} \left[x^{\frac{N}{2}} \mathrm{e}^{-\frac{x^2}{2}} \right. \\
& \left. \times \int_0^\infty \tilde{y}^{\frac{N}{2}} \mathrm{e}^{-\frac{1}{2}\tilde{y}^2} I_{\frac{N}{2}-1}\left(\frac{\epsilon}{\sqrt{1+\epsilon^2}} x\tilde{y}\right) \mathrm{d}\tilde{y} \right] \mathrm{d}x \\
=& \frac{1}{2^{\frac{N}{2}-1}\Gamma\left(\frac{N}{2}\right)(1+\epsilon^2)^{\frac{N}{2}}} \int_0^{x_u} x^{N-1} \mathrm{e}^{-\frac{x^2}{2(1+\epsilon^2)}} \mathrm{d}x \\
=& \frac{1}{2^{\frac{N}{2}}\Gamma\left(\frac{N}{2}\right)} \int_0^{\frac{x_u^2}{1+\epsilon^2}} u^{\frac{N}{2}-1} \mathrm{e}^{-\frac{u}{2}} \mathrm{d}u \\
=& 1 - \bar{\Gamma}\left(\frac{N}{2}, \frac{\bar{\Gamma}^{-1}\left(\frac{N}{2}, 1-X\right)}{1+\epsilon^2}\right),
\end{aligned} \tag{6.64}$$

where the second line follows by a change of variable and a change in the order of integration, the fourth line follows from the fact that $Q_m(a, 0) = 1$ and the final lines follow from another change of variable. We have also made use of the integral expression for the Marcum-Q function given below. This result can also be obtained directly by noticing that the random variable R^2, which depends on both n and $\delta h(\vec{\lambda})$, is distributed as $1+\epsilon^2$ times a χ^2 random variable with N degrees of freedom. The analytic expression for the P-P plot is therefore very similar to the case of the exact likelihood, but with the argument of the regularised Gamma function scaled appropriately to give the same result as in the final line of Eq.(6.64). This P-P plot also exhibits a sag below the diagonal; see the orange dotted curve in Fig. 6.1.

The N dimensional result in Eq. (6.62) can be shown to reduce to the 1 dimensional result in Eq. (6.55) using the standard properties of the Marcum-Q function. The

Marcum-Q function is defined by the integral

$$\begin{aligned} Q_m(a,b) &= \int_b^\infty x\left(\frac{x}{a}\right)^{m-1}\exp\left[-\frac{1}{2}(x^2+a^2)\right]I_{m-1}(ax)\mathrm{d}x \\ &= \exp\left[-\frac{1}{2}(a^2+b^2)\right]\sum_{k=1-m}^{\infty}\left(\frac{a}{b}\right)^k I_k(ab)\,, \end{aligned} \tag{6.65}$$

in which $I_n(x)$ is the modified Bessel function of the first kind. For $N = 1$, the Marcum-Q function, Eq. (6.65), can also be simplified

$$\begin{aligned} Q_{\frac{1}{2}}(a,b) &= \sqrt{a}\int_b^\infty \sqrt{x}\exp\left[-\frac{(x^2+a^2)}{2}\right]I_{-1/2}(ax)\mathrm{d}x \\ &= \sqrt{\frac{1}{2\pi}}\int_b^\infty\left(\exp\left[-\frac{(x+a)^2}{2}\right]+\exp\left[-\frac{(x-a)^2}{2}\right]\right)\mathrm{d}x \\ &= 1-\frac{1}{2}\left(\mathrm{erf}\left(\frac{b-a}{\sqrt{2}}\right)+\mathrm{erf}\left(\frac{b+a}{\sqrt{2}}\right)\right), \end{aligned} \tag{6.66}$$

which follows from $I_{-\frac{1}{2}}(x) = \sqrt{2/\pi}\cosh(x)/\sqrt{x}$. When $N = 1$ the regularised Gamma function becomes

$$\begin{aligned} \frac{\Gamma\left(1/2, R^2/2\right)}{\Gamma\left(1/2\right)} &= \frac{\int_{R^2/2}^\infty \mathrm{e}^{-t}/\sqrt{t}\ \mathrm{d}t}{\int_0^\infty \mathrm{e}^{-t}/\sqrt{t}\ \mathrm{d}t} \\ &= \frac{\int_{R/\sqrt{2}}^\infty \mathrm{e}^{-u^2}\ \mathrm{d}u}{\int_0^\infty \mathrm{e}^{-u^2}\ \mathrm{d}u} \\ &= 1-\mathrm{erf}(R/\sqrt{2})\,. \end{aligned} \tag{6.67}$$

Eq. (6.61) therefore becomes

$$\begin{aligned} P(\mathrm{sig}<X) &= \frac{1}{2}\left(\mathrm{erf}\left(\mathrm{erf}^{-1}(\mathrm{X})-\sqrt{\frac{\Lambda}{2}}\right)\right. \\ &\qquad \left.+\mathrm{erf}\left(\mathrm{erf}^{-1}(\mathrm{X})+\sqrt{\frac{\Lambda}{2}}\right)\right), \end{aligned} \tag{6.68}$$

as we can identify $\tilde{\mu} = \Lambda/2$, we recover Eq. (6.55) as expected.

6.4.2.3 The marginalised likelihood

The *Marginalised Likelihood* is given by Eq. (5.9), as before this may be expanded in the LSA. In the high signal-to-noise limit the posterior is narrow compared to the length scale over which the waveform changes. The waveform difference changes over the same

length scale as the waveform. The quantity $\sigma^2(\vec{\lambda})$ also changes over this length scale, as it is "learnt" by the GP in the procedure of maximising the evidence. Therefore in the high signal-to-noise limit $\sigma^2(\vec{\lambda})$ may be approximated as a constant. As before coordinates which diagonalise the Fisher matrix may be adopted, which give the following separated expression for the approximate likelihood,

$$\mathcal{L}(\vec{\lambda}) \propto \prod_x \exp\left(-\frac{1}{2}\frac{\left(\tilde{\Delta\lambda}^x - \left\langle n + \mu(\vec{\lambda}_0) - \delta h(\vec{\lambda}_0)|\tilde{\partial}_x(H-\mu)\right\rangle\right)^2}{1+\sigma^2}\right) . \tag{6.69}$$

The waveform difference is assumed to be a small quantity, therefore in Eq. (6.69) the derivative $\tilde{\partial}_x(H-\mu)$ may be replaced by $\tilde{\partial}_x(H)$, as the difference is the product of small quantities. Identical manipulations to those performed on the exact and approximate likelihoods give the same expression for the significance obtained in Eqs. (6.42) and (6.58),

$$\text{sig} = 1 - \bar{\Gamma}\left(\frac{N}{2}, \frac{R^2}{2}\right) , \tag{6.70}$$

except this time the random variable R^2 is given by

$$R^2 = \frac{1}{1+\sigma^2}\sum_x \left\langle n + \mu(\vec{\lambda}_0) - \delta h(\vec{\lambda}_0)|\tilde{\partial}_x H\right\rangle^2 , \tag{6.71}$$

$$= \frac{1}{1+\sigma^2}\left(\Sigma^{-1}\right)^{ab}\left\langle n + \mu(\vec{\lambda}_0) - \delta h(\vec{\lambda}_0)|\partial_a H\right\rangle\left\langle n + \mu(\vec{\lambda}_0) - \delta h(\vec{\lambda}_0)|\partial_b H\right\rangle \tag{6.72}$$

The GPR technique assumes that the $\delta h(\vec{\lambda})$ are distributed as a Gaussian process across parameter space, with zero mean and a covariance estimated from a training set and any prior knowledge. If this assumption is in fact true, and the covariance has been correctly estimated, then the quantity $\mu(\vec{\lambda}_0) - \delta h(\vec{\lambda}_0)$ is distributed as a zero mean Gaussian with variance σ^2. In this case (perhaps unsurprisingly) the marginalised likelihood completely fixes the sag. The new R^2 random variable is distributed as a χ^2 random variable with N degrees of freedom and using the regularised Gamma function as the CDF of this distribution we recover the diagonal P-P plot;

$$P(\text{sig} < X) = 1 - P\left(R^2 < 2\bar{\Gamma}^{-1}\left(\frac{N}{2}, 1-X\right)\right) , \tag{6.73}$$

$$P(\text{sig} < X) = X . \tag{6.74}$$

This case is shown, both analytically and numerically, in orange in Fig. 6.1.

More interesting is the behaviour in the realistic case when $\delta h(\vec{\lambda})$ is not distributed exactly as the GPR has predicted. This case is more complicated because the different

components that make up the R^2 random variable are no longer independent random variables and a simple expression for the distribution of $\delta h(\vec{\lambda})$ cannot be found. In particular, from Eq. (6.71) it can be seen that R^2 is the sum of the squares of a noise term, $\langle n|\tilde{\partial}_x H\rangle$, a GPR term, $\langle \mu(\vec{\lambda}_0)|\tilde{\partial}_x H\rangle$, and (minus) a physical term, $\langle \delta h(\vec{\lambda}_0)|\tilde{\partial}_x H\rangle$. In particular the GPR and physical terms are now related because the expression for $\mu(\vec{\lambda}_0)$ in Eq. (5.24) is a linear combination of the realisations of $\delta h(\vec{\lambda})$ in the training set, $\mathcal{D}$. The sag will still be given by the analogue of Eq. (6.57), but this integral will not in general be analytically tractable. Instead, we will consider such cases numerically in Sec. 6.4.3.

As we have seen, in the particular case considered above where the waveform difference is distributed as assumed by the GPR, the marginalised likelihood completely removes the systematic bias present in the standard, approximate, likelihood. In addition, as we will see in Sec. 6.4.3, even in unfavourable situations the marginalised likelihood is often able to remove significant portions of the bias. We conclude this section with a discussion of why it is expected that the bias in parameter estimates obtained using the marginalised likelihood will usually be less than those obtained using the standard likelihood.

From Eqs. (6.59) and (6.71) it can be seen that the condition for the marginalised likelihood to yield *more* biased parameter estimates than the approximate likelihood, *for a particular event*, is $R^2_{\text{approx}} < R^2_{\text{GPR}}/(1+\sigma^2(\vec{\lambda}_0))$, where

$$\begin{aligned} R^2_{\text{approx}} &= \sum_x \left\langle n - \delta h(\vec{\lambda}_0)|\tilde{\partial}_x H\right\rangle^2 \\ R^2_{\text{GPR}} &= \sum_x \left\langle n - \delta h(\vec{\lambda}_0) + \mu(\vec{\lambda}_0)|\tilde{\partial}_x H\right\rangle^2 . \end{aligned} \tag{6.75}$$

These terms both involve a projection onto the space spanned by the derivatives, $\tilde{\partial}_x H$, at the point $\vec{\lambda}_0$. Since these "tilde" derivatives were constructed to be an orthonormal basis, the condition for the marginalised likelihood to give worse parameter estimates than the approximate likelihood can therefore be written as

$$\left|n - \delta h(\vec{\lambda}_0)\right|^2_{\mathcal{D}} < \frac{\left|n - \delta h(\vec{\lambda}_0) + \mu(\vec{\lambda}_0)\right|^2_{\mathcal{D}}}{1+\sigma^2(\vec{\lambda}_0)} \tag{6.76}$$

where the modulus is taken with respect to the function inner product in Eq. (5.5), projected into the space, $\mathcal{D}$, spanned by the derivatives. For this to be satisfied, it would be necessary not only for the interpolation to have the wrong sign when expressed in the

basis $\tilde{\partial}_x H$ (i.e. $0 > \sum_x \langle\, h(\vec{\lambda}_0)|\tilde{\partial}_x H \,\rangle \langle\, \mu(\vec{\lambda}_0)|\tilde{\partial}_x H \,\rangle$), but also for it to be large enough in magnitude to overcome the GPR uncertainty $\sigma^2(\vec{\lambda}_0)$ in the denominator. Moreover, this is just for one particular realisation of the noise and true waveform parameters. We are really interested in the sag that arises when considering a population of events. In that case, we would need Eq. (6.76) to be true in some average sense and so the interpolation would have to have the wrong sign and be too large for the majority of choices of waveform parameters. Although this is technically possible, it is clear that any reasonable interpolation algorithm with decent coverage of the parameter space in the training set and a reasonable covariance function should violate the above bound on average and therefore yield better parameter estimates on average and show a smaller sag in the P-P plot than the approximate likelihood.

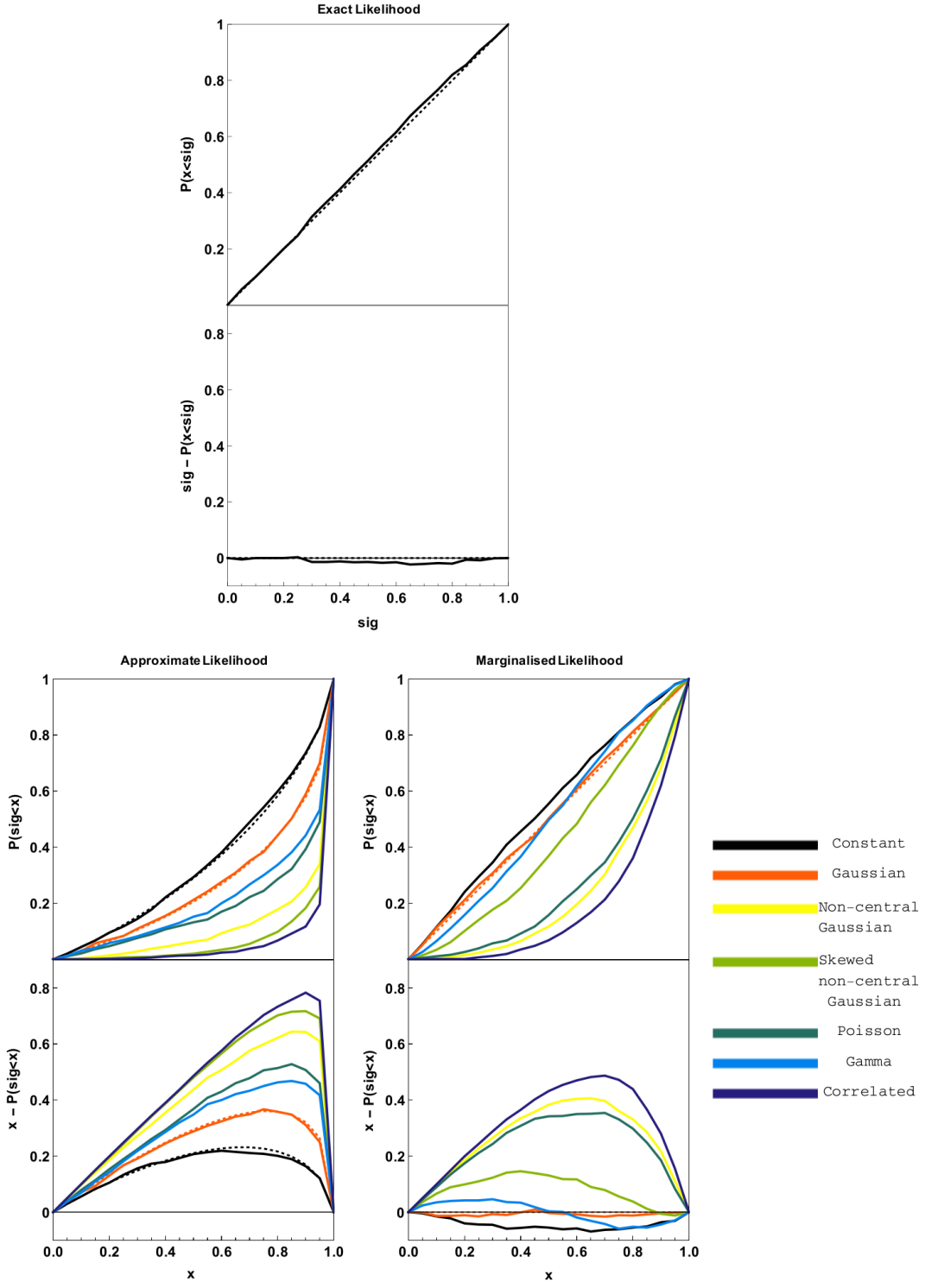


Figure 6.1: *See following page for figure caption.*

P-P plots for parameter estimation using the three likelihoods $L'(\vec{\lambda})$, $L(\vec{\lambda})$, and $\mathcal{L}(\vec{\lambda})$ shown in the three columns respectively. In each column the top panel shows a P-P plot whilst the bottom panel shows the "sag"; i.e. the difference between the ideal diagonal line and the actual P-P plot. In each panel curves drawn as dotted lines correspond to analytic results whilst solid curves are numerical results. The top-most column shows ideal, unbiased parameter recovery for the exact likelihood. In the bottom-left and bottom-right columns different colour curves correspond to different distributions of $\langle \delta h(\lambda_0) | \tilde{\partial}_x Hangle$. The curves in black are for a constant distribution giving a non-centrality $\Lambda = 2$ (Λ defined in Eq. (6.60)). The curves in orange are for a zero mean Gaussian distribution with variance 1. The curves in yellow are for a non-central Gaussian distribution with mean 4/3 and variance 1. The curves in light-green are for a non-central, skewed normal distribution with location parameter 1, scale parameter 1, and skew parameter 1. The curves in dark-green are for a Poisson distribution with mean and variance 1. The curves in blue are for a Gamma distribution with shape parameter 1 and scale parameter 1. And finally, the curves in purple are for a correlated random walk distribution with mean Gaussian step size 1. In all cases the number of parameter dimensions is $N = 4$, and the number of points used for the numerical simulations was $n = 10^3$. The top-most panel clearly shows the exact likelihood does not suffer from any bias, as expected. The bottom-left panel shows that in all cases the approximate likelihood suffers from a bias. The bottom-right column shows that in all cases the marginalised likelihood reduces the bias relative to the approximate likelihood. In the ideal case (shown in orange) of a zero mean Gaussian distribution for $\langle \delta h(\lambda_0) | \tilde{\partial}_x H \rangle$ the bias is completely removed.

6.4.3 Numerical calculation of the P-P plot

In all of the above calculations the expression for the P-P plot was written in terms of the CDF of the distribution of the R^2 random variable. This random variable is written in terms of a signal inner product of the model derivatives, it therefore depends both on the properties of the gravitational wave source and of the detector. By expressing our results in terms of R^2 we ensure that they remain valid for any detector and any source (assuming the LSA holds). In the cases considered above where analytic expression for the P-P plots could be found these can also be verified numerically by drawing n values of R^2 from the relevant distribution and numerically estimating the CDF. In cases where an analytic expression for the P-P plot can not be found the same procedure can be used to investigate the P-P plot numerically.

First consider the unbiased, diagonal P-P plot obtained for the exact likelihood. The analytic expression for this P-P plot is given in Eq. (6.46). A numerical validation of this result may be performed by drawing random realisations of the R^2 value in Eq. (6.44). It can be seen that R^2 is the sum of the squares of N standard Gaussian random variables $\langle n|\tilde{\partial}_x H\rangle$; i.e. a χ^2 distribution with N degrees of freedom. We drew n realisations of R^2 from this distribution, numerically estimated the CDF and plotted the P-P plot using Eq. (6.45). The results for $n = 10^3$ and $N = 4$ are shown in the left panels of Fig. 6.1 (analytic results shown as a dotted line, numerical results as a solid line[1]). Within the scale of fluctuations the numerical results agree well with the analytic results. The bottom left panel of the same figure shows the sag of the P-P plot below the diagonal, i.e. $\mathrm{sig} - P(x < \mathrm{sig})$. The values $n = 10^3$ and $N = 4$ will also be used for all subsequent numerical calculations in this section.

P-P plots for the approximate likelihood are shown in the centre panels of Fig. 6.1 for a variety of different distributions of the waveform difference projected into the model derivatives; $\langle \delta h(\lambda_0)|\tilde{\partial}_x H\rangle$. In the case of a constant distribution, or a zero-mean Gaussian distribution the analytic expressions in Eqs. (6.62) and (6.64) respectively are shown as dotted lines. For the numerical calculations the procedure followed was first to specify the distribution for the $\langle \delta h(\lambda_0)|\tilde{\partial}_x H\rangle$ random variables (for example the black curves show results when this is a constant). The quantity R^2 was then calculated using Eq. (6.59) by drawing a random value from this distribution and a random value for $\langle n|\tilde{\partial}_x H\rangle$ from

[1]The PDF of a skew Gaussian distribution with location parameter μ, scale parameter σ and skew parameter α is given by $\left[1 + \mathrm{erf}(\alpha(x-\mu)/\sqrt{2}\sigma)\right] \exp(-(x-\mu)^2/2\sigma^2)$

a standard Gaussian distribution. The R^2 variable was calculated n times, the CDF of this variable estimated, and the P-P plot calculated from Eq. (6.61). Different colours in Fig. 6.1 indicate different distributions for $\langle \delta h(\lambda_0) | \tilde{\partial}_x H \rangle$, the specification of these distributions are given in the figure caption.

	Approximate	Marginalised
Constant	0.158	-0.044
Gaussian	0.237	0.000
non-central Gaussian	0.385	0.263
Skew non-central Gaussian	0.426	0.079
Poisson	0.317	0.235
Gamma	0.293	-0.001
Correlated	0.441	0.308

Table 6.1: Table of the total integrated biases for the curves shown in Fig. 6.1. The integrated bias is defined as the total area in the sag, i.e. $\int_0^1 \mathrm{d}(\mathrm{sig})\ (\mathrm{sig} - P(x < \mathrm{sig}))$.

P-P plots for the marginalised likelihood are shown in the right panels of Fig. 6.1 for a variety of different distributions of $\langle \delta h(\lambda_0) | \tilde{\partial}_x H \rangle$. In the case of a zero-mean Gaussian the analytic expressions in Eq. (6.74) is shown as a dotted line. For the numerical calculations it is necessary to construct a training set for the GPR interpolation. Instead of using GPR to interpolate the waveform differences, $\delta h(\vec{\lambda})$, it is simpler for our present purpose to instead interpolate the projections of the waveform differences onto the waveform derivatives, i.e., $\langle \delta h(\vec{\lambda}) | \tilde{\partial}_x H \rangle$, as these are what appear in Eq. (6.71), . The training set was taken to consist of points at $\lambda = 1, 2, \ldots, 20$ and the actual experimental realisation at a value $\lambda_0 = 21$. For the majority of distributions (constant, Gaussian, non-central Gaussian, skew non-central Gaussian, Poisson, and Gamma distributions) shown in Fig. 6.1 the random variables in the training set were drawn independently and interpolated using an uncorrelated Gaussian process, i.e. $K_{ij} = \sigma_f \delta_{ij}$. The R^2 value was calculated from Eq. 6.71 (with $\mu = 0$ from Eq. (5.24), because of the assumption of an uncorrelated process), the CDF estimated and the P-P plot calculated from Eq. (6.73).

The assumption of an uncorrelated Gaussian process is a conservative assumption. In the absence of correlations the Gaussian process regression assumes a "worst-case" scenario and returns a mean waveform difference of zero (see Eq. (5.24)). If correlations

were present then the GPR would return a non-zero estimate for μ and shift the position of the posterior peak into better agreement with the true value, thus improving the P-P plot. To investigate the effect of correlations the final numerical calculation (labelled as "correlated" in Fig. 6.1) was performed using a random walk distribution. The values of $\langle \delta h(\vec{\lambda}) | \tilde{\partial}_x H \rangle$ at the points $\lambda = 1, 2, \ldots, 21$ were taken to be a realisation of a random walk with Gaussian step width $a = 1/3$. The first 20 of these values were taken as the training set and used to extrapolate the final value. For the GPR interpolation a squared exponential covariance function $k(x, y) = \exp((-1/2)(x - y)^2)$ was used. The squared exponential covariance function is not able to accurately capture the covariance of the random walk distribution, so this again represents a conservative choice to examine how the marginalised likelihood performs in the presence of un-modelled correlations. However, even in this unfavourable case the marginalised likelihood still significantly reduces the bias in the P-P plot.

The purpose of considering such a wide variety of different distributions for the waveform difference is to test whether the marginalised likelihood is robust against different types of errors in the waveform models, which are not correctly modelled by the Gaussian process. For example, the marginalised likelihood assumes the waveform difference is a zero mean Gaussian process across parameter space, therefore it is perhaps not surprising that it performs well in the case of a zero mean Gaussian distribution. However the list of distributions used here also test the robustness of the method against non-central distributions (e.g non-central Gaussian), skewed distributions (e.g. skewed Gaussian), one-sided and non-Gaussian distributions (e.g. Poisson or Gamma distributions), and the presence of un-modelled correlations in the waveform difference (the random walk distribution).

By comparing the curves of the same colour between the centre and right-hand panels of Fig. 6.1 it can be seen that in all cases the P-P plot for the marginalised likelihood exhibits less of a bias, i.e., less of a "sag", than the approximate likelihood. In the ideal case where the distribution of the waveform differences is precisely that assumed by the GPR, a diagonal, unbiased P-P plot is recovered; however the bias is also almost completely removed for several of the other distributions considered. In all cases a significant improvement in performance can be seen when using the marginalised likelihood in place of the standard approximate likelihood. These results are summarised in Table 6.1, which lists the *total bias* (defined as the area between the sagging curve

and the ideal diagonal) for all the curves shown in Fig. 6.1.

6.4.4 Discussion

The P-P plot provides a way to quantify the bias that results when using inaccurate models to perform gravitational wave parameter estimation. For individual sources the systematic error in the parameters is independent of the signal-to-noise (SNR), whilst the random errors scale as 1/SNR, and hence the bias is most significant for the loudest sources. Even in cases where, for each individual source, the systematic error is small compared to the random error, the bias can still be significant when observing populations of sources, since the statistical error in a parameter estimated from combining a population of sources reduces as $1/\sqrt{N}$ as more sources are added, while the systematic errors remain fixed.

In this section several analytic expressions have been obtained that predict the sag of the P-P plots that results from different distributions of the model error. These results have been derived within the linear signal approximation, and are valid to $\mathcal{O}(1/\mathrm{SNR})$. These analytic expressions for the P-P plots may be viewed in the same spirit as Fisher matrix estimates for the random errors, or Cutler and Vallisneri (2007) expression for the systematic error in a single measurement. This latter result has also here been generalised (in Appendix D) to include terms of $\mathcal{O}(1/\mathrm{SNR}^2)$.

It is now well established that model errors will present significant problems for a range of sources. In this section the performance of the marginalised likelihood was examined by comparing the P-P plots (obtained both analytically and numerically) with those obtained from the standard likelihood. In particular, it was found that in favourable conditions the marginalised likelihood was able to completely remove the parameter estimation bias. More importantly, it was found that the marginalised likelihood was robust against a wide range of un-modelled features in the distribution of waveform differences, and in all cases considered outperformed the standard likelihood. These results provide further illustration of the need to account for model uncertainties (using GPR or other techniques) when drawing inferences from near future gravitational wave observations.

Chapter 7

Implementation

In this chapter we present an illustrative implementation of our GPR approach. Our focus will be on the novel features of the method, for example the use of Gaussian processes and the choice of the covariance function, and not on the aspects the method has in common with the standard approach, for example the need to use a stochastic sampler to draw independent points from an unknown, high dimensional distribution. To this end we will consider a relatively simple parameter estimation problem involving simply estimating a single parameter; a full multidimensional application that would be appropriate for actual gravitational wave data analysis will be investigated in future work. We then focus our attention on how the performance of the standard likelihood compares to the marginalised likelihood, and how the latter is affected by changing various features of the GPR.

We begin in Sec. 7.1 by introducing the waveforms we use for this study. In Sec. 7.2 we describe the placement of the training set points for the GPR; in order to investigate the effect of training set on the GPR interpolant two sets were constructed with different numbers of points and grid spacings. In Sec. 7.3 we present results for maximising the hyperlikelihood to find the optimum hyperparameters, $\vec{\theta}_{\rm op}$, for the interpolation; this is done for a range of different covariance functions on each of the training sets described in Sec. 7.2. In Sec. 7.4 we interpolate the waveforms across parameter space for the different training sets and different covariance functions described and compare the interpolated waveforms $H(\vec{\lambda}) - \mu(\vec{\lambda})$ to the accurate waveforms $h(\vec{\lambda})$. In Sec. 7.5 we present results for the GPR uncertainty, $\sigma^2(\vec{\lambda})$, for the different training sets and different covariance functions considered. Finally in Sec. 7.6 we present results for the marginalised likelihood

$\mathcal{L}(\vec{\lambda})$, and compare with the results obtained using the approximate likelihood $L(\vec{\lambda})$, and the exact likelihood $L'(\vec{\lambda})$.

7.1 Model waveforms

In order to implement the GPR, a choice has to be made regarding which waveform models to use. The method uses two waveform approximants; the accurate $h(\vec{\lambda})$ and the approximate $H(\vec{\lambda})$ waveforms. The accurate waveform should be the most accurate available at a computational cost that permits the offline construction of the training set $\mathcal{D}$. The criteria for choosing the approximate waveform is less clear, a balance needs to be struck between accuracy and speed. If the model is computationally cheap but not accurate enough the waveform difference, $\delta h(\vec{\lambda}) = H(\vec{\lambda}) - h(\vec{\lambda})$, will be large and vary on short length scales over parameter space; these are the situations which will cause the GPR to perform worst. On the other hand an accurate model which is too computationally expensive could slow down any PE to such an extent that there ceases to be any benefit in using the marginalised likelihood instead of the accurate likelihood.

We used two waveform models implemented in the LIGO Scientific Collaboration Algorithm Library (LAL).[1] As our intention here is to provide a proof of principle, we choose the IMRPhenomC approximant (Santamaría et al., 2010) as the accurate waveform and the widely used TaylorF2 approximant (Buonanno et al., 2009; Damour et al., 2001, 2002) as the approximate waveform; both of these models are sufficiently fast to evaluate that we can compute and then compare the three likelihoods (accurate $L'(\vec{\lambda})$, approximate $L(\vec{\lambda})$, and marginalised $\mathcal{L}(\vec{\lambda})$) and directly assess the performance of the GPR.

Both of the approximants we have chosen to use here are frequency-domain models, i.e. they naturally return the waveform in the Fourier domain $\tilde{h}(f)$.[2] The IMRPhenomC waveform includes inspiral, merger and ringdown, while the TaylorF2 waveform only includes the inspiral.

[1] http://www.lsc-group.phys.uwm.edu/lal

[2] In previous published work by Moore and Gair (2014) the marginalised likelihood has been implemented with time-domain approximants. The method works equally with frequency-domain, which are used throughout this thesis. In the offline stage the waveforms enter only via the overlap matrix $\langle \delta h(\vec{\lambda}_i) | \delta h(\vec{\lambda}_j) \rangle$, and in the online stage the waveforms enter only in the linear combination for $\mu(\vec{\lambda})$ in Eq. (5.24), which commutes with the operation of taking the Fourier transform.

We investigate the merger of non-spinning circular binaries. This limits the number of intrinsic parameters describing the system to two, the masses of the two component objects, $\vec{\lambda} = \{m_1, m_2\}$. To further simplify the problem we place training set points only along a one-dimensional subspace, which we choose to be a surface of constant mass ratio, $Q = m_2/m_1$ (with $m_1 \geq m_2$), parametrised by the value of the chirp mass $\mathcal{M}_\mathrm{c} = (m_1 m_2)^{3/5}/(m_1 + m_2)^{1/5}$. This keeps the size of the training set small, and hence the computational complexity of the GPR to a minimum. This allows us to instead focus our attention on the novel features of the marginalised likelihood, and explore the effect of changing various features of the method.

7.2 The training set

For simplicity we restrict the range of the coordinates which we search over to reduce the computational complexity. This is again to allow us to focus our attention on the novel features of the method. The training sets cover the chirp mass in the range $\mathcal{M}_\mathrm{c} \in [5, 5.6]\, M_\odot$ and the mass ratio is fixed to the (nearly equal mass) $Q = 0.75$. The placement of training set points was done as a regular grid in chirp mass with a step size between points of $\Delta\mathcal{M}_\mathrm{c}$.

The chirp-mass range has been chosen to demonstrate the properties of the method. For lower masses, the signal is dominated by the inspiral where both approximants agree well. Therefore, interpolating these small differences would not be a robust test. At higher masses, where the signal is just merger and ringdown, the two approximants are completely different; we get no useful information from the TaylorF2 waveform and may as well interpolate IMRPhenomC directly. We do not anticipate that in practice we would consider waveform differences as significant as the complete absence of merger and ringdown; hence, this example, although only one-dimensional, should be a rigorous test of what waveform uncertainties can be successfully interpolated. Understanding if this continues to be true for the interpolation of a more intricate waveform difference across a higher-dimensional parameter space, must wait for further studies to be completed.

To allow us to explore the effect that the density of points in the training set has on the marginalised likelihood two different values for $\Delta\mathcal{M}_\mathrm{c}$ were considered. This leads to two different training sets whose total number n of points are different; the properties of these two training sets are summarised in Tab. 7.1. It is expected that the GPR interpolation, and hence the marginalised likelihood, will perform better when using the

	$\Delta\mathcal{M}_\mathrm{c}$	N
$\mathcal{D}_0$	$1.0 \times 10^{-2} M_\odot$	60
$\mathcal{D}_1$	$5.0 \times 10^{-3} M_\odot$	120

Table 7.1: The properties describing the positions of the template waveforms for each of the three training sets used.

denser set $\mathcal{D}_1$.

Once the training set points $\{\vec{\lambda}_i\}$ were specified, both the approximate $H(\vec{\lambda}_i)$ and accurate $h(\vec{\lambda}_i)$ waveforms discussed in Sec. 7.1 were evaluated at each point, and the waveform differences $\{\delta h(\vec{\lambda}_i)\}$ stored for use during the GPR interpolation. The matrix of waveform difference overlaps $M_{ij} = \langle \delta h(\vec{\lambda}_i) | \delta h(\vec{\lambda}_j) \rangle$ was also evaluated and stored for use during the hyperlikelihood maximisation procedure.

7.3 The hyperparameters

Initially the training sets described in Sec. 7.2 were interpolated using the SE covariance function in Eq. (5.33). This covariance function has just two hyperparameters, $\vec{\theta} = \{\sigma_f, g_{\mathcal{M}_\mathrm{c}\mathcal{M}_\mathrm{c}}\}$. The one-dimensional metric $g_{\mathcal{M}_\mathrm{c}\mathcal{M}_\mathrm{c}}$ can be exchanged for a length scale in the chirp mass parameter $\delta\mathcal{M}_\mathrm{c} \equiv 1/\sqrt{g_{\mathcal{M}_\mathrm{c}\mathcal{M}_\mathrm{c}}}$. A fixed noise term with $\sigma_n^2 = 10^{-4}$ was used for all the covariance functions in this section, to make the inverse of the covariance function numerically stable as discussed in Sec. 5.4.3. The hyperlikelihood for the training set $\mathcal{D}_0$ was maximised with respect to these two hyperparameters. The optimum values for the hyperparameters were found to be

$$\sigma_f = 3.49 \times 10^4 \,, \tag{7.1}$$

$$\delta\mathcal{M}_\mathrm{c} = 1.11 \times 10^{-2} M_\odot \,. \tag{7.2}$$

The hyperlikelihood is shown in Fig. 7.1. The hyperlikelihood was also maximised for the training set $\mathcal{D}_1$ using the same SE covariance function and those results are also shown in Fig. 7.1. For the denser training set $\mathcal{D}_1$ the optimum length scale was found to be smaller, $\delta\mathcal{M}_\mathrm{c} = 6.31 \times 10^{-3} M_\odot$. For both training sets, in the limit that the length scale becomes much larger than the total width of the training set ($0.6 M_\odot$) or much smaller than the grid point spacing ($\Delta\mathcal{M}_\mathrm{c}$), the hyperlikelihood tends to a constant value. This behaviour can be understood by examining the expression for the hyperlikelihood in

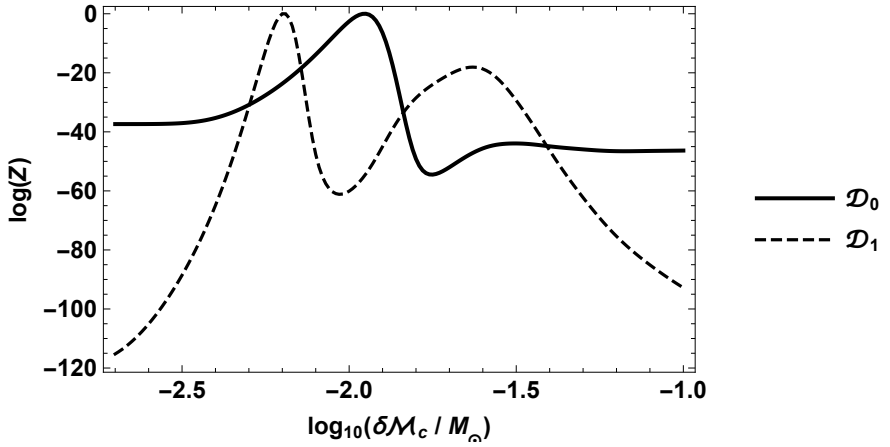


Figure 7.1: The hyperlikelihood, from Eq. (5.18), for the SE covariance function, maximised over the scale hyperparameter σ_f, plotted against the chirp mass length scale $\delta\mathcal{M}_{\rm c}$. The hyperlikelihood is shown for both of the training sets (normalised to a peak value of 1). The denser training set $\mathcal{D}_1$ was found to favour smaller length scales.

Eq. (5.18).

In order to explore the effect that the choice of covariance function has on the marginalised likelihood, the training sets were also interpolated using the Matérn covariance function in Eq. (5.36). This covariance function has an additional hyperparameter, $\vec{\theta} = \{\sigma_f, g_{\mathcal{M}_{\rm c}\mathcal{M}_{\rm c}}, \eta\}$. The hyperlikelihood for training set $\mathcal{D}_0$ was maximised for this covariance function. It was found that the hyperlikelihood surface did not possess a peak, instead a ridge was found tending to a maximum at a value $\eta \to \infty$, and values of σ_f and $g_{\mathcal{M}_{\rm c}\mathcal{M}_{\rm c}}$ were found to be the same as for the SE covariance function. In Fig. 7.2 we plot the log-hyperlikelihood (maximised over σ_f) against chirp-mass length scale and the additional hyperparameter η.

As the Matérn covariance function recovers the SE function in the limit $\eta \to \infty$, there will be no difference in the performance of the interpolants for this training set when using the Matérn or SE covariance functions. If the volume under the hyperlikelihood surface (the hyperevidence) is used as a figure-of-merit for which covariance function the data favours, then in this case the data is equally well described by either covariance function, but the SE covariance function is favoured over the Matérn due to the smaller prior volume (the Occam penalty).

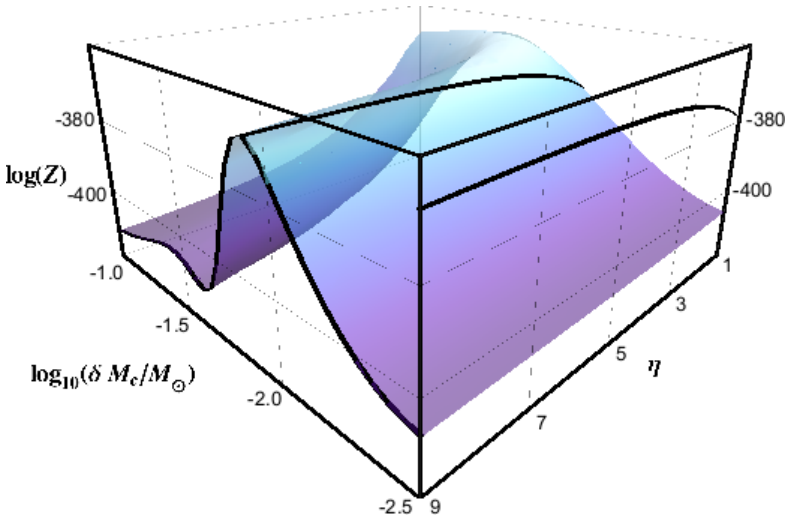


Figure 7.2: The hyperlikelihood, from Eq. (5.18), surface for the training set $\mathcal{D}_0$ using the Matérn covariance, maximised over the hyperparameter σ_f, plotted against the chirp mass length scale $\delta\mathcal{M}_\mathrm{c}$ and the hyperparameter, η. The hyperlikelihood does not show a clear peak, instead a ridge in the hyperparameter space favours the limiting case $\eta \to \infty$, in which limit the Matérn covariance function is equal to the SE covariance function. On the near-side faces of the plot box we show the hyperlikelihood sliced parallel to the coordinate axes though the point $(\delta\mathcal{M}_\mathrm{c} = 10^{-1.9} M_\odot, \eta = 10)$. The solid black line on the near, left-hand face of the box very closely matches the solid black curve in Fig. 7.1 (up to an arbitrary additive constant).

The hyperlikelihood was also calculated for both training sets $\mathcal{D}_0$ and $\mathcal{D}_1$ using the PLE covariance, see Eq. (5.34), and the Cauchy covariance, see Eq. (5.35), considered in Sec. 5.4. In both cases a similar behaviour was observed. For the PLE covariance, a peak in the hyperlikelihood was found at $\eta = 2$, where the PLE covariance equals the SE covariance. For the Cauchy covariance, a ridge in the hyperlikelihood was found tending to a maximum for $\eta \to \infty$ (similar to the Matérn case shown in Fig. 7.2), in which limit the Cauchy covariance also recovers the SE covariance. As with the Matérn covariance, if the hyperlikelihood is used as a figure-of-merit for selecting the covariance function then the SE covariance is favoured over both the PLE and Cauchy functions due to the Occam penalty.

It is clear that interpolations of the training sets $\mathcal{D}_0$ and $\mathcal{D}_1$ using any of the PLE,

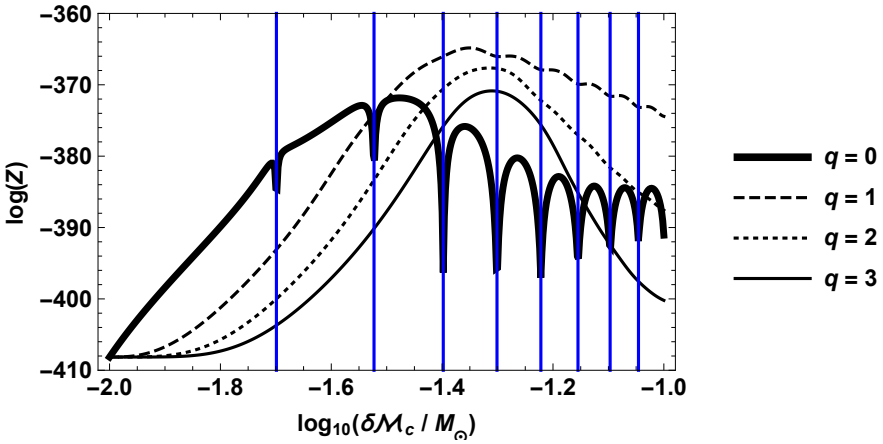


Figure 7.3: The hyperlikelihood, from Eq. (5.18), for the training set $\mathcal{D}_0$ using the Wendland polynomial covariance functions, maximised over the scale hyperparameter σ_f, plotted against the chirp-mass length scale $\delta\mathcal{M}_\mathrm{c}$. The vertical blue lines indicate multiples of the training-set grid spacing $\Delta\mathcal{M}_\mathrm{c}$.

Cauchy, or Matérn covariance functions, evaluated at the hyperlikelihood-maximising hyperparameters, would yield identical results to an interpolation using the simpler SE covariance. For this reason, in the following sections we do not use the PLE, Cauchy, or Matérn functions further and instead focus on the SE covariance function. We will, however, also consider using the Wendland polynomial function in the following sections as it reduces the computational cost.

The hyperlikelihood for the compact support Wendland polynomial covariance functions are shown in Fig. 7.2, for the cases $q = 0$, 1, 2, 3. The compact-support functions can develop multiple-peaks in the hyperlikelihood surface associated with the length-scale of the training set: multiples of the training-set grid spacing are indicated with vertical blue lines in Fig. 7.2. These subsidiary peaks occur in the $\delta\mathcal{M}_\mathrm{c}$ hyperparameter because as the size of the compact-support region grows, the (integer) number of training set points it contains changes discontinuously.

From Fig. 7.3 it can be seen that for the training set $\mathcal{D}_0$, a value of $q = 1$ is favoured with a length-scale $\delta\mathcal{M}_\mathrm{c} = 4.37 \times 10^{-2} M_\odot$. In the following sections we will use Wendland covariance function with all values of q (and their associated peak hyperlikelihood length scales) to interpolate $\mathcal{D}_0$.

The optimum hyperparameters depend on the detector noise power spectral density via the overlap matrix $\langle \delta h(\vec{\lambda}_i) | \delta h(\vec{\lambda}_j) \rangle$. In Appendix B, an investigation of the sensitivity of the optimum hyperparameters to small changes in the detector noise properties is described. It was found that for any realistic changes to the noise curve, the optimum hyperparameters were changed by an amount too small to have any noticeable effect on the interpolant.

7.4 The interpolated waveforms

The GPR waveform $H(\vec{\lambda}) - \mu(\vec{\lambda})$ could be viewed as a new waveform approximant formed from the approximant waveforms and the use of GPR on the training set of accurate waveforms. It is then natural to ask how this new approximant compares to the original ones. This can be assessed by calculating the overlap between the different waveforms, where the overlap is defined by

$$\text{overlap}(a, b) = \frac{\langle a|b \rangle}{\|a\| \, \|b\|} \,, \tag{7.3}$$

using the inner product defined in Eq. (5.5).

Only considering the overlap misses the important extra benefit which the marginalised likelihood approach brings. Our method is not just supplying a new waveform approximant, but also providing a way of modifying the posterior to account for the uncertainties known to be in the approximant. This extra information which modifies the likelihood surface is included through $\sigma(\vec{\lambda})$. Nonetheless, it is still informative to temporarily treat $H(\vec{\lambda}) - \mu(\vec{\lambda})$ as if it were a new waveform approximant and see how it compares with the approximants $h(\vec{\lambda})$ and $H(\vec{\lambda})$ from which it was built. Fig. 7.4 shows the waveform overlap between the interpolated waveform $H(\vec{\lambda}) - \mu(\vec{\lambda})$ and the accurate waveform $h(\vec{\lambda})$ as a function of chirp mass near the edge of the training set. Also shown in the dotted curve is the overlap between the approximate waveform $H(\vec{\lambda})$ and the accurate waveform $h(\vec{\lambda})$. The interpolated waveforms have a much higher overlap than the approximate waveforms, as would be expected. Within the training set the overlap is increased from ~ 0.35 to no less than ~ 0.985 even for the sparser training set D_0. For the denser training set $\mathcal{D}_1$ overlaps no worse than ~ 0.999 were found inside the range of the training set. Outside the training set the interpolated waveform tends rapidly to the approximate waveform $H(\vec{\lambda})$.

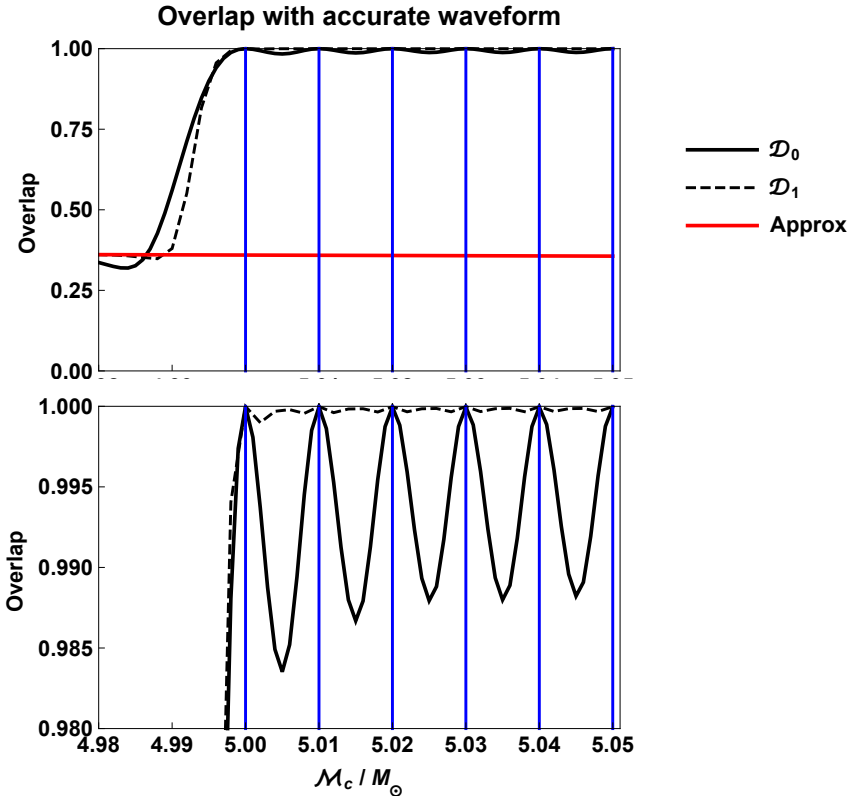


Figure 7.4: A plot of the overlap between the interpolated waveform $H(\vec{\lambda}) - \mu(\vec{\lambda})$ and the accurate waveform $h(\vec{\lambda})$ as a function of the chirp mass $\mathcal{M}_c$. The bottom panel is the same plot with a different ordinate axis scale. The two black lines show the overlap using both training sets, $\mathcal{D}_0$ and $\mathcal{D}_1$, interpolated using the SE covariance function. The red line shows the overlap between the approximate waveform $H(\vec{\lambda})$ and the accurate waveform $h(\vec{\lambda})$ for comparison. The vertical blue lines show the position of the training set points for $\mathcal{D}_0$. In the bottom panel, it can be seen that, for either interpolant, the overlap becomes one when evaluated at the training set points.

The training set waveforms were also interpolated using the Wendland compact support covariance functions discussed in Sec. 5.4.4. The cases $q = 0, 1, 2, 3$ were considered separately. The waveform overlap using these interpolants is plotted in Fig. 7.5. The performance of these interpolants should be compared with the results

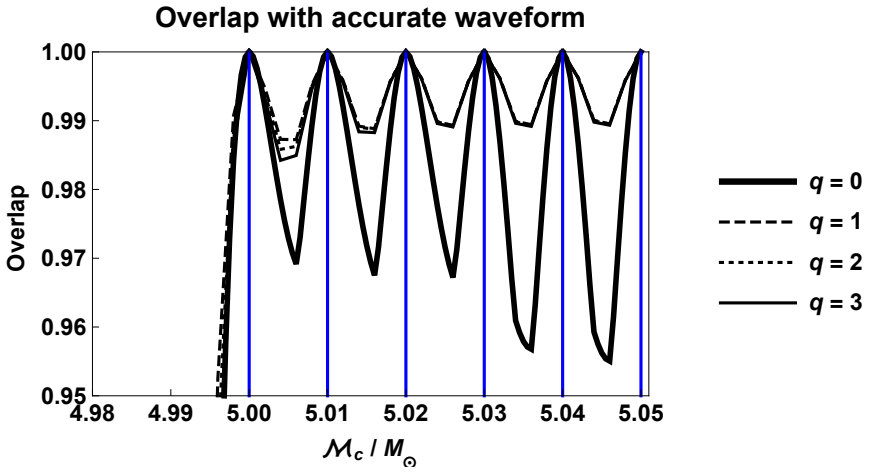


Figure 7.5: A plot of the overlap (or overlap) between the interpolated waveform $H(\vec{\lambda}) - \mu(\vec{\lambda})$ and the accurate waveform $h(\vec{\lambda})$ as a function of the chirp mass $\mathcal{M}_c$. The different curves correspond to using the Wendland polynomial covariance functions with different values of q to interpolate the training set $\mathcal{D}_0$. The vertical blue lines show the position of the training set points for $\mathcal{D}_0$.

using the SE covariance function in Fig. 7.4.

The least smooth of the Wendland polynomials, the $q = 0$ case, performs noticeably worse than the SE covariance; inside the training set the overlap drops as low as ~ 0.955 compared to ~ 0.985 for the SE. However, even a overlap of ~ 0.955 is still a great improvement over the overlap of ~ 0.35 for the approximate waveform alone. For the $q = 0$ Wendland polynomial the interpolant has a discontinuous first derivative, which can be seen in Fig. 7.5 (this is expected and was discussed in Sec. 5.4 and in detail in Appendix C). The higher values of q have discontinuities in the higher ordered derivatives, but these curves look smooth to the eye. The smoother Wendland polynomials, with $q > 0$, all perform very similarly to the SE covariance function; inside the training set the overlap drops as low as ~ 0.985 for the $q = 2$ interpolant.

7.5 The GPR uncertainty

The GPR performs an interpolation of the points in the training set and naturally returns a Gaussian error $\sigma(\vec{\lambda})$, see Eq. (5.25), for each interpolated point. In our present

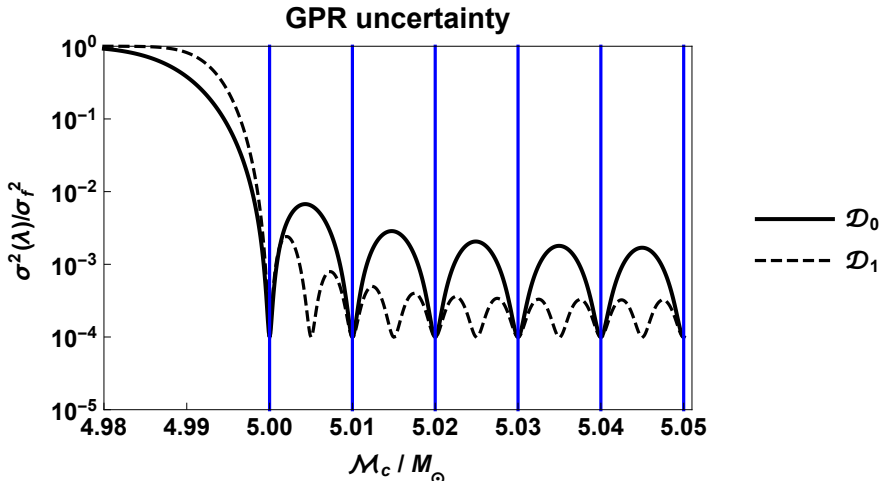


Figure 7.6: A plot of the GPR uncertainty $\sigma^2(\vec{\lambda})$ as a function of the chirp mass parameter for both of the training sets, using the SE covariance function. The vertical blue lines show the position of the training set points for $\mathcal{D}_0$. Outside of the training set the uncertainty tends to a constant σ_f^2. Inside the training sets the error is approximately periodic with minima at the training set points. The maximum uncertainty inside the training set is smaller for the denser training sets.

one-dimensional interpolation this is simply a function of $\mathcal{M}_c$. A small section of this curve taken from the edge of the training set is shown in Fig. 7.6. Inside the training set, the error surface has a regular, periodic pattern with minima at the training set points and maxima in between. This regularity is because the GP used for the interpolation is stationary, the training-set points used are regularly spaced, and each point has an identical error (a jitter $J = 10^{-4}$). If these conditions were to be relaxed, then the error surface would become more complicated. In general, a larger $\sigma(\vec{\lambda})$ indicates greater theoretical uncertainty and highlights regions where we would benefit from additional accurate waveforms (e.g., where it would be beneficial to perform more NR simulations).

Near the edge of the training set the behaviour becomes less regular and well outside of the training set the error tends to a constant value, $\sigma^2(\vec{\lambda}) \to \sigma_f$ as $\vec{\lambda} \to \infty$. This behaviour is seen in Fig. 7.6 for all three training sets. The training sets with smaller grid spacings have smaller uncertainties everywhere in parameter space.

The GPR uncertainty was also calculated using the Wendland polynomial covariance

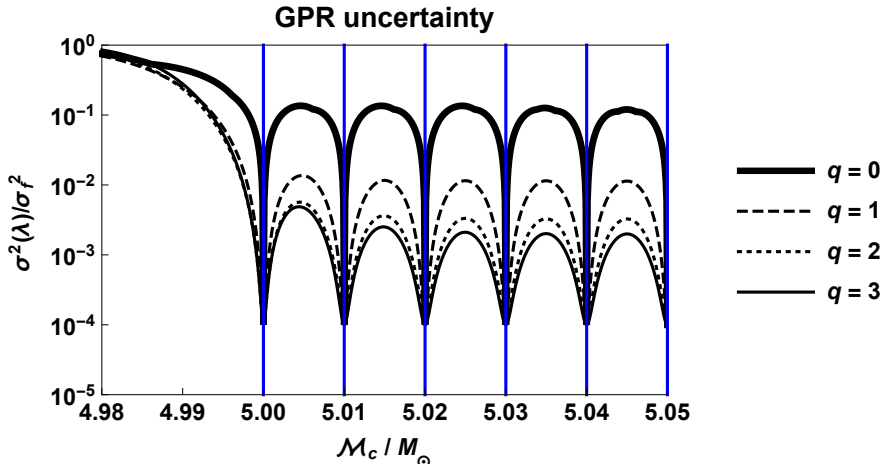


Figure 7.7: A plot of the GPR uncertainty $\sigma^2(\vec{\lambda})$ as a function of the chirp mass parameter for the training set $\mathcal{D}_0$, using the Wendland polynomial covariance functions. The vertical blue lines show the position of the training set points.

functions to interpolate the training set $\mathcal{D}_0$; these are shown in Fig. 7.7. The GPR uncertainty, expressed as a fraction of σ_f^2, is largest for the smallest values of q; this can be traced back to the optimum length scale for the Wendland polynomials increasing with q (see Fig. 7.3). This means that the uncertainty grows more slowly as the interpolating point moves away from the training set points, and hence reaches a smaller maximum value between training set points. The smoother ($q > 0$) Wendland polynomials perform similarly to the SE covariance function, in the sense that both the GPR interpolants (which we quantify via the overlap) and the GPR uncertainties are almost identical. Hence, in the following sections we will only consider using the SE covariance function; the high q Wendland polynomials would yield identical results.

7.6 The likelihood

Finally we put together the interpolated waveform $H(\vec{\lambda}) - \mu(\vec{\lambda})$ and the GPR uncertainty $\sigma^2(\vec{\lambda})$ to give the marginalised likelihood in Eq. (5.28). We compare the performance of the marginalised likelihood $\mathcal{L}(\vec{\lambda})$ to the approximate likelihood $L(\vec{\lambda})$ and the accurate likelihood $L'(\vec{\lambda})$. For the injected signal we use the accurate waveform $h(\vec{\lambda})$. We also consider the case where the noise realisation is zero (the most likely realisation), this

makes comparisons easier.

We injected a signal at a chirp mass of $\mathcal{M}_c = 5.045 M_\odot$; this is inside the training set $\mathcal{D}_0$ and midway between training set points. Injecting the signal midway between the points is conservative as this is the point at which we would expect the marginalised likelihood to perform worst. The three different likelihoods were evaluated as a function of chirp mass (all other parameters set to the injected values). This was done at a range of signal-to-noise ratios and the results are shown in Fig. 7.8. The top row of panels in Fig. 7.8 show the likelihoods renormalised to a peak value of one, this makes the relative positions of the peaks clear and easy to compare. The bottom row of panels shows the log-likelihood without any renormalisation, this illustrates how the approximate likelihood is suppressed relative to the true likelihood (the detection problem discussed in Sec. 5.1).

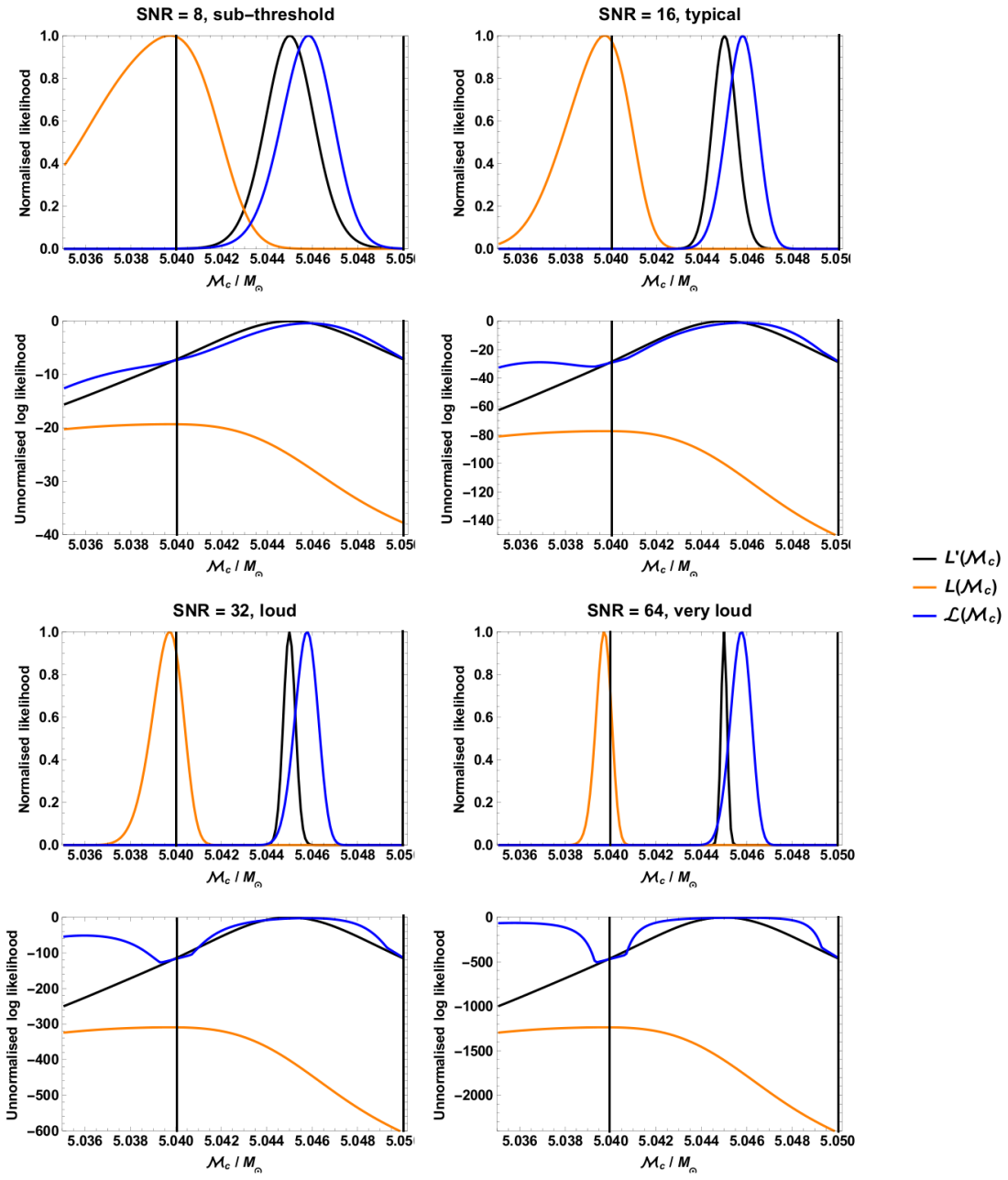


Figure 7.8: A plot of the different likelihoods for a variety of signal-to-noise ratios. Vertical lines indicate the position of training set points. The upper-most panel in each pair of plots shows the likelihood normalised to the same peak value; this makes the peak positions clear and shows how the marginalised likelihood tackles the parameter estimation problem associated with the inaccurate models. The lower-most panel in each pair of plots shows the log-likelihood; this makes the suppression of the peak value of the approximate likelihood clear and shows how the marginalised likelihood could be used to tackle the detection problem.

The exact likelihood $L'(\vec{\lambda})$ is always peaked at the injected value of the chirp mass (because the injected noise realisation is zero) and the width of the peak decreases with increasing signal-to-noise ratio (SNR). The approximate likelihood $L(\vec{\lambda})$ is peaked away from the true value, indicating a systematic error of $\Delta_{\rm sys}\mathcal{M}_{\rm c} = 5.2 \times 10^{-3} M_{\odot}$. The width of the approximate likelihood peak also decreases with increasing SNR and for SNR $\gtrsim 12$ (which is also roughly the detection threshold (Abbott et al., 2013; Berry et al., 2015)) the true parameters are excluded at increasing significance. The bottom row of panels in Fig. 7.8 shows that the approximate likelihood is suppressed by a significant amount, for a typical SNR of 16 it is supressed by 80 in log relative to the exact likelihood; this reduces the Bayesian evidence for a detection. The factor by which the approximate likelihood is suppressed increases exponentially with SNR. Finally, the marginalised likelihood is peaked much closer to the exact likelihood: the systematic error is reduced to $\Delta_{\rm sys}\mathcal{M}_{\rm c} = 9.0 \times 10^{-4} M_{\odot}$. However, as discussed in Sec. 6.2, the peak in the marginalised likelihood does not continually narrow as the SNR increases; for SNR $\gtrsim 30$ the width becomes constant. Consequently, the true parameters are never excluded at high significance; in the limit of infinite SNR the true parameters lie at the $\sim 1\sigma$ level. The bottom panel of Fig. 7.8 shows that the marginalised likelihood is not suppressed relative to the exact likelihood in the vicinity of the peak.

Comparing the properly normalised likelihoods, we see that the marginalised and exact likelihoods roughly agree at low SNR. As the SNR is increased, the marginalised likelihood deviates from the exact likelihood and develops oscillatory behaviour with period equal to the training set point spacing. In the limit of low SNR, all of the parameter estimation uncertainty comes from the noise, but as the SNR increases, the relative size of this statistical uncertainty becomes smaller and at high SNR we are dominated by model uncertainty. The marginalised likelihood correctly encapsulates this behaviour, as can be seen in the sequence from left to right in Fig. 7.8.

7.7 Discussion

The marginalised likelihood has been suggested as a means of incorporating theoretical uncertainty into gravitational wave data analysis. We have now thoroughly investigated the properties of this method, elucidating considerations for a practical implementation. A detailed derivation of the marginalised likelihood, and the use of GPR to interpolate model error was presented in Sec. 5.3. GPR is non-parametric, in the sense that only the

functional form of the covariance function is specified by hand, with its hyperparameters then learnt from the training set, making it well suited to modelling theoretical uncertainty. The expression for the marginalised likelihood derived in Sec. 5.3 made some assumptions about the frequency covariance of the waveforms in the training set; in particular it was assumed that at a particular point in parameter space, the model error is highly correlated in frequency. These assumptions may prove to be too restrictive in the future, and could be relaxed by simultaneously performing GPR interpolation in frequency (as well as in parameter space) on the training set waveforms; this would have the added advantage of allowing the inclusion of waveforms with different frequency samplings, which may be beneficial when using multiple waveform approximants and NR waveforms from multiple sources.

The choice of covariance function is central to GPR as it encodes our prior beliefs about the function space that we are interpolating. We discussed various choices of covariance function in Sec. 5.4. We have found that the simple SE covariance function (as used in Moore and Gair (2014)) performs as well as more complicated alternatives, at least for the relatively small one dimensional training sets considered here. The compact-support Wendland covariance functions with large q were found to perform comparably to the SE, but offer the additional advantage of reduced computational cost. This makes them appealing for future work involving larger training sets.

We proved a number of properties for our marginalised likelihood in Sec. 6, in particular its limiting behaviour for large signal amplitude (where the theoretical errors are known to be most significant (Cutler and Vallisneri, 2007)) and its limiting behaviour both far from and near a point in the training set. In the discussion of the latter, the linearised results previously obtained in Cutler and Vallisneri (2007) were recovered. All of these properties demonstrate the suitability of GPR for making robust inferences. The marginalised likelihood successfully describes our belief in our inferences, including our uncertainty in waveform templates.

In Sec. 7, we presented a one-dimensional implementation of our marginalised likelihood and demonstrated that it offers an improvement in PE accuracy. For this, we chose two inexpensive waveforms to aid computation; in real data analysis situations, we expect more accurate waveforms (including those calculated using NR) to be used in the training. However, this choice of waveforms does illustrate the efficacy of the new marginalised likelihood. In particular we find that even when using qualitatively different waveforms

(inspiral-only TaylorF2 compared to inspiral–merger–ringdown IMRPhenomC), waveform matches as high as $\sim 98.5\%$ can be obtained in the mass range we considered. Restricting ourselves to the simple case of one-dimensional PE, we explored various possibilities for GPR. In particular, the effect of different training set sizes was examined; as expected, the performance of the marginalised likelihood is improved by using denser training sets. Additionally, the impact of varying the SNR of the injected waveform was studied. In the standard likelihood model, errors become more severe as the SNR is increased, but we confirmed that even in the limit of large SNR, the marginalised likelihood remained consistent with the injected parameters. We expect these results to carry over to a full multidimensional analysis, which is the next step in developing this technique.

A possibly complementary use for the GPR approximant is as a less expensive alternative to the accurate waveform for more expedient PE; however, there exist other means of constructing computationally inexpensive waveforms, such as reduced-order modelling (Canizares et al., 2015; Pürrer, 2014). The advantage of GPR is that it not only supplies an interpolant, but also gives an uncertainty, which can be used to gauge accuracy away from training points. A second, related feature is that GPR naturally allows for uncertainty to be included in the accurate models that the interpolant is calibrated against. It is the ability of GPR to include theoretical uncertainty that makes it attractive for gravitational wave astronomy; that this can be done without significant on-line cost is a welcome bonus.

In conclusion, marginalising over waveform uncertainty is a robust and effective method of accounting for theoretical error in both PE and detection problems. GPR is a natural and effective means of performing this marginalisation. The marginalised likelihood is naturally inferior to a likelihood calculated with more accurate (but inevitably more computationally expensive) waveforms, but it offers significantly improved performance over the standard likelihood calculated with cheap waveforms. In addition, the marginalised likelihood is almost as quick to evaluate online as the standard likelihood, although there is additional offline computation required to construct the training set and train the Gaussian process.

Chapter 8

Future generalisations

In the previous chapter the promise of the marginalised likelihood was demonstrated on a toy problem. In order to be used for routine binary black hole parameter estimation in the new era of gravitational wave astronomy the technique will need to be efficiently implemented for large scale use in the LIGO Scientific Collaboration Algorithm Library (LAL).[1] Therefore the main avenue for future work is in implementing this technique. However, there are also a number of other theoretical avenues where the marginalised likelihood can be developed; we discuss one such issue in this section.

An objection to the marginalised likelihood as presented above may be raised on the grounds that the GPR error, $\sigma^2(\vec{\lambda})$, does not depend on the gravitational wave frequency. It is known that the approximate waveform models (for example, those built on the post Newtonian approximation) are more accurate at early times (or low frequencies) when the orbiting bodies are well separated than at late times (or high frequencies) when the bodies merge. It would be desirable to be able to include this additional prior knowledge into the expression for the likelihood because this may lead to more accurate parameter estimates.

The fact that the above method does not adequately include all available prior information can be seen from the following simple thought experiment[2]. Let the true signal, $h(\vec{\lambda})$ be a simple sine wave with amplitude A, observed for a large number of cycles over a time T, with observed signal-to-noise $\rho \approx \sqrt{AT/2}$ which we assume is much greater than 1. If we attempt to measure the frequency of the sine wave using the accurate

[1] http://www.lsc-group.phys.uwm.edu/lal

[2] Considered at the suggestion of Will Farr

model of a sine wave and the exact likelihood (Eq. 5.4) we will have some uncertainty Δf in this measurement. Now suppose that we only have access to an approximate model, $H(\vec{\lambda})$, which is equal to the true model for the first half the signal, but equal to zero for the second half. Now we attempt to measure the frequency using the approximate likelihood (Eq. 5.7); the approximate model only provides information about half the signal, the signal-to-noise is $\rho/\sqrt{2}$ and the uncertainty on the measured frequency will be $\sim \sqrt{2}\Delta f$. (In the high signal-to-noise ratio (SNR) limit the parameter estimation errors scale as 1/SNR.) The toy waveform models are shown in Fig. 8.1.

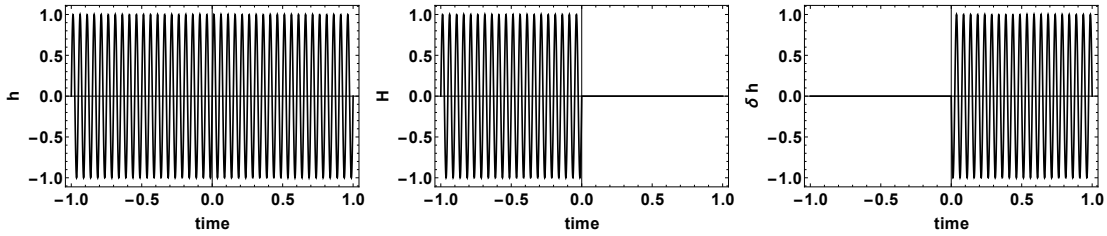


Figure 8.1: A very simple waveform model. The accurate model shown on the left is a sine wave with fixed frequency and amplitude. The approximate model shown in the centre is the same sine wave for negative x, and zero for positive x. The waveform difference shown on the right is defined, as before, as $\delta h = H - h$.

However, if we now analyse the entire signal using the marginalised likelihood in Eq. 5.28 we get a different result. For simplicity we will assume the worst case scenario where the injected signal is far away in parameter space from any training set points, and so (from Eqs. 5.24 and 5.25) we have $\mu = 0$ and $\sigma^2 \sim \rho^2/2$. The error obtained for the frequency from the marginalised likelihood will be $(1 + \sigma^2)\Delta f$. This is obviously not the desired behaviour; for large signal amplitudes the error estimate obtained from the marginalised likelihood are much larger than those obtained from the far simpler approach of just considering half the data. The reason for this behaviour stems from our original assumption about the waveform difference; namely that the waveform difference is a realisation of a GP with a particular covariance function. In particular, the entire waveform difference function at all times has been assumed to have the same error. Therefore, when training the GP, large errors at certain times (or frequencies) contribute to the GPR uncertainty which is then assumed to apply at all times (or frequencies).

This leads to an incorrect posterior distribution.

The solution to this problem is to generalise the GPR so that different portions of the data are assumed to come from their own, independent GPs. In the toy example we are currently considering we can construct a modified, marginalised likelihood in the following way. Split the data in half, on the first half construct a training set and use the marginalised likelihood from Eq. 5.28. In this half of the data the waveform difference is zero, and so the marginalised likelihood recovers the exact likelihood evaluated on this part of the data. In the second half of the data construct another training set and again use the marginalised likelihood. In the second half of the data we have $\mu = 0$ and $\sigma^2 \sim \rho^2/2$. (To be consistent, we again assume the worst case scenario where the injected signal is far away in parameter space from any training set points.) Now we form a new total likelihood by taking the product of the likelihoods for the first and second halves. This new marginalised likelihood, evaluated for this simple toy problem, has the property that it recovers exactly the approximate likelihood in the limit of large signal-to-noise.

Therefore, in the worst case scenario the performance from the newly constructed marginalised likelihood matches that from the approximate likelihood. In a better scenario where there are training set waveforms nearby the performance will improve. Fig. 8.2 shows the four different likelihoods for this toy problem; as expected, the exact likelihood has the smallest width, the approximate likelihood has a width wider by a factor of $\sqrt{2}$, the marginalised likelihood from Eq. 5.28 has a much larger width, and the new marginalised likelihood has the same width as the approximate likelihood.

8.1 The new marginalised likelihood

In the simple example considered above the signal was monochromatic, and the likelihood was computed in the time domain as the product of the likelihoods for the first and second halves of the signal separately. However, for general broadband signals with coloured detector noise the expression for the standard likelihood only separates in this way in the frequency domain:

$$L(\vec{\lambda}) = \exp\left(-\frac{1}{2}\left\langle s - H(\vec{\lambda}) | s - H(\vec{\lambda})\right\rangle\right) \tag{8.1}$$

$$= \exp\left(-2\sum_i \delta f \frac{|\tilde{s}(f_i) - \tilde{H}(f_i)|^2}{S_n(f_i)}\right) \tag{8.2}$$

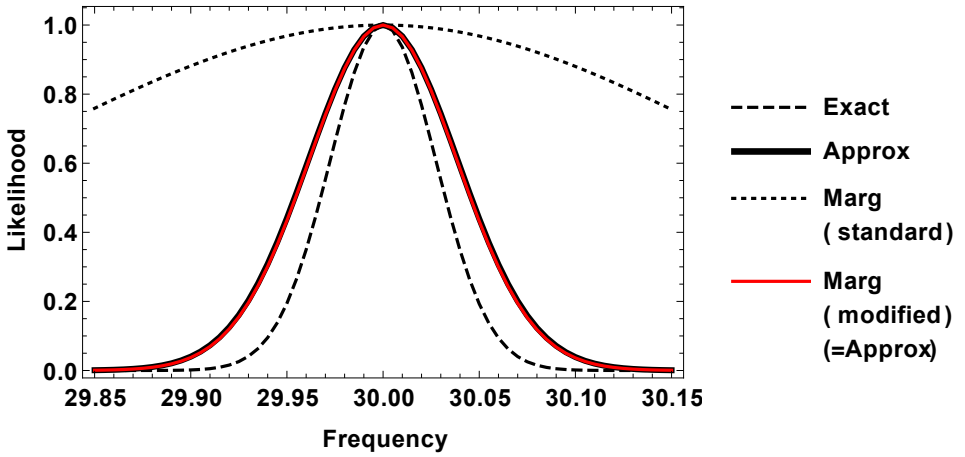


Figure 8.2: The different likelihoods for the simple thought experiment, with a SNR for the accurate model of $\rho = 10$. The SNR is large, so the exact likelihood is almost a Gaussian. The approximate likelihood is also almost a Gaussian, with width wider by a factor of $\sim \sqrt{2}$. The standard marginalised likelihood has a much wider posterior indicating the poor performance in this case. The modified marginalised likelihood curve lies almost exactly on top of the approximate likelihood curve in this case; in the limit of infinite signal-to-noise the agreement would be perfect.

$$= \prod_i \exp\left(-2\delta f \frac{|\tilde{s}(f_i) - \tilde{H}(f_i)|^2}{S_n(f_i)}\right) , \tag{8.3}$$

where i indexes the different frequency bins. We will now construct a prior to marginalise over the waveform difference by assuming that each frequency component is drawn from an independent Gaussian process. In particular, training the GPs separately, this allows for different GPR errors in each frequency bin.

At the i^{th} frequency bin we assume that the waveform difference is drawn from a GP and we construct the training set $\mathcal{D}_i$. After training, the GP provides a Gaussian distribution for the waveform difference at some new parameter point, $\vec{\lambda}$, with mean $\tilde{\mu}(\vec{\lambda}, f_i)$, and variance $\sigma^2(\vec{\lambda}, f_i)$.

$$\mathcal{D}_i = \left\{ \left(\vec{\lambda}_x, \delta\tilde{h}(\vec{\lambda}_x, f_i) \right) \middle| \, x = 1, 2, \ldots, n \right\} \tag{8.4}$$

$$\tilde{\mu}(\vec{\lambda}, f_i) = k_x^* K_{xy}^{-1} \delta\tilde{h}(\vec{\lambda}_y, f_i) \tag{8.5}$$

$$\sigma^2(\vec{\lambda}, f_i) = k^{**} - k_x^* K_{xy}^{-1} k_y^* \tag{8.6}$$

The k-quantities in the above equations are defined in terms of a covariance function in an analogous way to the previous chapter, except now the hyperparameters of the covariance function can vary with frequency. For example, one possible choice is the SE covariance,

$$k(\vec{\lambda}_1, \vec{\lambda}_2) = \sigma_f(f) \exp\left(-\frac{1}{2} g_{ab}(f)(\vec{\lambda}_1 - \vec{\lambda}_2)^a(\vec{\lambda}_1 - \vec{\lambda}_2)^b\right) . \tag{8.7}$$

The probability distribution for the component of the waveform difference is given by

$$P(\tilde{\delta h}(f_i)) = \frac{\exp(-\frac{1}{2}\frac{|\tilde{\mu}(\vec{\lambda},f_i)-\tilde{\delta h}(\vec{\lambda},f_i)|^2}{\sigma^2(\vec{\lambda},f_i)})}{2\pi\sigma^2(\vec{\lambda}, f_i)} . \tag{8.8}$$

The unusual looking normalisation in the denominator of Eq. 8.8 comes from the normalisation condition $\int \mathrm{d}(\tilde{\delta h}(f_i))\, P(\tilde{\delta h}(f_i)) = 1$ where it is understood that $\mathrm{d}(\tilde{\delta h}(f_i)) = \mathrm{d}(\mathrm{Re}\{\tilde{\delta h}(f_i)\}) \times \mathrm{d}(\mathrm{Im}\{\tilde{\delta h}(f_i)\})$.

The GPs in different frequency bins are assumed to be uncorrelated, leading to the following probability distribution for the total waveform difference; this takes the place of Eq. 5.23,

$$P(\tilde{\delta h}) = \prod_i P(\tilde{\delta h}(f_i)) . \tag{8.9}$$

The marginalised likelihood is defined in the same way as the previous chapter, except now with Eq. 8.9 forming the prior on the waveform difference. We will denote by $\Lambda(\vec{\lambda})$ the marginalised likelihood defined in this frequency dependent way, to distinguish it from the frequency independent $\mathcal{L}(\vec{\lambda})$. Performing the integrals gives,

$$\begin{aligned}\Lambda(\vec{\lambda}) = &\prod_i \frac{1}{2\pi\sigma^2(\vec{\lambda}, f_i)} \int \mathrm{d}\left(\delta h(\tilde{f}_i)\right) \exp\left(-2\delta f \frac{|\tilde{s}(f_i) - \tilde{H}(\vec{\lambda}, f_i) + \tilde{\delta h}(f_i)|^2}{S_n(f_i)}\right) \\ &\times \exp\left(-\frac{1}{2}\frac{|\tilde{\mu}(\vec{\lambda}, f_i) - \tilde{\delta h}(f_i)|^2}{\sigma^2(\vec{\lambda}, f_i)}\right)\end{aligned} \tag{8.10}$$

$$\Lambda(\vec{\lambda}) = \prod_i \left\{ \frac{S_n(f_i)}{S_n(f_i) + 4\delta f \sigma^2(\vec{\lambda}, f_i)} \exp\left(-2\delta f \frac{|\tilde{s}(f_i) - \tilde{H}(\vec{\lambda}, f_i) + \tilde{\mu}(\vec{\lambda}, f_i)|^2}{S_n(f_i) + 4\delta f \sigma^2(\vec{\lambda}, f_i))}\right)\right\} . \tag{8.11}$$

The exponential factor in the expression for the modified, marginalised likelihood looks very much like $\sim \exp(-(1/2)\langle s - H + \mu | s - H + \mu\rangle)$, except that the inner product is taken with respect to a modified noise curve,

$$S'_n(f) = S_n(f) + 4\sigma(\vec{\lambda}, f)/T , \tag{8.12}$$

where T is the total observation time. We can write Eq. 8.11 in the more compact form

$$\Lambda(\vec{\lambda}) = \left(\prod_i \frac{S_n(f_i)}{S'_n(f_i)}\right) \exp\left(-\frac{1}{2}\left[s - H(\vec{\lambda}) + \mu(\vec{\lambda})|\ldots\right]\right) \tag{8.13}$$

where square brackets now indicate the inner product defined in Eq. 5.5, except with $S_n(f) \to S'_n(f)$. From Eq. 8.13 it can be seen that (up to a normalisation) $\Lambda(\vec{\lambda})$ is equal to the standard likelihood, $L(\vec{\lambda})$, just with the waveform model replaced with the interpolation, $H(\vec{\lambda}) \to H(\vec{\lambda}) - \mu(\vec{\lambda})$, and with the inner product taken with respect to the modified noise S'_n. The factor of 4 in the modified noise curve is related to our convention for the signal inner product in Eq. 5.5.

8.2 Smoothness Power Spectra Vs Power Spectral Density

The quantity $|\sigma^2(f)|$ which appears in the effective noise curve is the usual power spectra of the waveform difference. The quantity $S_n(f)$ is the power spectral *density* of the detector noise. Consider a segment of detector data of length T, and take the limit $T \to \infty$; the total energy in the detector noise in this segment diverges in the limit. However the waveform difference, $|\sigma^2(f)|$, remains fixed. In order to compare the effective "noise" due to the model error with the physical noise in the detector, a choice for T has to be made and the power in the waveform difference divided by T to form a density. This explains the factor $\delta f \equiv 1/T$ in the effective noise.

Other than the condition that T should be greater than the length of the waveform, there is no natural choice for T. However, it is already standard practice in the gravitational wave parameter estimation community to fix a segment length of data before any parameter estimation study, and it is this choice that fixes the frequency resolution δf. It should be noted however, that different choices of T will lead to different effective noise PSD as defined in Eq. 8.12. This peculiarity simply reflects the fact that one source of noise is naturally a density, while another source is naturally a power, and an arbitrary choice has to be made to compare the two. Choosing a larger value of T means that the data has a higher frequency resolution (there are more frequency bins).

Our choice of covariance function assumes that the error in each frequency bin in independent. A possible concern with interpolating each frequency bin of the waveform difference separately is that the resulting interpolant may not be a sufficiently smooth function of frequency. From Eq. 8.5 it can be seen that the interpolant, $\mu(\vec{\lambda}, f)$, is formed from a combination of training set waveforms and the covariance function. Provided the covariance function is chosen to be a smooth (i.e. infinitely differentiable) function of frequency then the interpolant will have the same level of differentiability as the waveform differences in the training set.

8.3 Conclusions

The accuracy with which we can measure gravitational wave source parameters and make inferences regarding populations of sources is limited not only by the noise levels in our detectors, but also by the fidelity of our models. There are many important sources of noise in an interferometer, and a great effort over many years has gone into minimising these over as wide a frequency range as possible. Whilst our model errors don't physically shake the mirrors in the interferometer we have shown here how they may also be interpreted as an effective source of noise. It is the sum of the physical noise and the effective noise which limits our ability to draw useful astrophysical inferences, and so overcoming the model errors is equally important as reducing the physical detector noise.

Part III

Testing the Kerr Hypothesis

Chapter 9

Accretion Disks

> Macavity, Macavity, there's no one like Macavity,
> He's broken every human law, he breaks the law of gravity.
>
> T. S. Eliot

General Relativity (GR) has been extensively tested in the weak field regime. The famous classical tests of GR began before the theory was even fully formulated with the explanation of the anomalous precession of the perihelion of mercury by Einstein (1915*b*). Only a few years later further confirmation came with the first measurements of the gravitational deflection of light during the 1919 solar eclipse by Dyson et al. (1920). More recently, but no less famously, the confirmation by Taylor et al. (1979) that the loss of energy from the binary pulsar PSR1913+16 was in the manner predicted by the theory won the authors the 1993 Nobel prize in physics [1]. Interspersed among these landmark experiments GR has been the subject of a constant stream of increasingly stringent tests; to date no deviations from the predictions of GR have been detected. For a review of experimental tests of GR see Will (2006).

All of the tests mentioned above concern phenomena in weak gravitational fields; it has proved much more difficult to devise similarly stringent tests of GR in the strong gravitational field. This is due in part to the immense difficulty associated with measuring the motion of bodies on the small gravitational length scale associated with black holes (a few kilometers for a solar mass black hole) at astrophysical distances. If

[1] `http://www.nobelprize.org/nobel_prizes/physics/laureates/1993/index.html`

such measurements could be made and it were possible to accurately track the motion of a test particle around an astrophysical black hole it would become possible to place constraints on the strong field spacetime metric. It is a prediction of GR, together with the no-hair theorem, that the spacetime metric around an astrophysical black holes is described by the famous Kerr solution (Kerr, 1963). Alternative theories of gravity predict the existence of different solutions (examples of known solutions considered here include the Kehagias and Sfetsos (2009) solution to Hořava gravity (Hořava, 2009*a*,*b*) and the slowly rotating Yunes and Pretorius (2009) solution to dynamical Chern-Simons gravity (Jackiw and Pi, 2003)). If a deviation from the Kerr solution was detected it could indicate a failure of GR in the strong field regime. Alternatively, it could indicate a failure of the no-hair theorem and raise the possibility of exotic compact objects within GR. It should be noted that there are other ways in which alternative theories of gravity may differ from GR besides changes to the metric around a black hole; for example, scalar-tensor theories of gravity (the theory of (Brans and Dicke, 1961) is perhaps the best known example) generically predict the existence of addition gravitational wave polarisation states (Will, 2006). In this chapter however we restrict our attention to testing the spacetime metric around a black hole.

Accretion disks provide an ideal candidate for such tests, because X-ray irradiation of matter in the inner regions of the disk ($r \lesssim 20M$) imprints characteristic features upon the X-ray spectra, in particular the fluorescent Kα Iron line (rest energy 6.38keV). As this emission originates from so close to the Black Hole (BH) it is strongly distorted by gravitational and Doppler shifting producing the characteristic 'two-horned' profiles observed in nature (Tanaka et al., 1995). Furthermore the light, once emitted, is strongly gravitationally lensed as it propagates through the spacetime, altering the observed spectra and imprinting upon it extra information about the strong field metric. Indeed the large width of these lines is evidence for the existence of highly spinning black holes (Iwasawa et al., 1996); for a recent review of black hole spin measurements using X-ray emission see McClintock et al. (2011). In addition, viscous torques present in the disk dissipate energy causing the material to gradually spiral inwards, this energy is radiated locally in the form of thermal emission. This thermal radiation is subjected to the same strong-field gravitational effects as the line emission, and so may also be used as a probe of the metric.

The approach adopted in this chapter is to consider Iron line and thermal emission

in a large class of parametrically deformed Kerr black holes. The deformed black hole spacetimes are referred to as "bumpy black holes" and the individual deformation parameters as "bumps". The spacetimes all have the property that if all the bump parameters are simultaneously set to zero then the Kerr solution is recovered. The advantage of this approach, compared with considering the emission in a small number of known alternative black hole solutions, is that one is able to consider a wide range of different deformations simultaneously and identify particular "bumps" which are especially easy or difficult to constrain. Furthermore, even if a particular alternative black hole is not contained within this class it may be hoped that the disk spectrum in this new metric would have a significant overlap with the spectrum of a black hole in the class; and therefore a deviation from GR would still be detectable. The family of bumpy black holes used for this purpose was the class of metrics constructed by Vigeland et al. (2011); these have the property that the perturbed spacetimes possess a fourth constant of motion (analogous to the Carter constant (Carter, 1968) in Kerr spacetime). In addition to this very general family of bumpy black holes the results are put into context by comparing them with the bounds it is possible to place on some known black hole solutions; in particular both the linear (Yunes and Pretorius, 2009) and quadratic (Yagi et al., 2012) in spin solutions to weakly coupled dynamical Chern-Simons gravity, and the slowly rotating Kehagias-Sfetsos black hole (Lee et al., 2010) were considered. There have been several previous efforts to systematically constrain deviations from the Kerr metric using accretion disk observations; for example see Johannsen (2013) for a study using regular black hole metric parameterised by four free functions. In this work we use a very general family of bumpy black holes, and we also compare our results against known black hole solutions in several specific alternative theories of gravity.

We begin this chapter by describing the family of bumpy black holes that were considered in this work, we also describe a couple of specific examples of known black hole solutions in alternative theories of gravity which are also considered here. In Sec. 9.2 the necessary theory for calculating both the thermal and Iron line emission from an accretion disk in a general stationary, axisymmetric spacetime is described. Sec. 9.3 describes the methods used for data analysis and the method used for estimating the bounds it will be possible to place on the different deformation, or "bump", parameters. The results are presented in Sec. 9.5 and finally a discussion and concluding remarks are given in Sec. 9.6. Throughout this chapter natural units, where $G = c = k_B = 1$,

are used. The work in the current chapter and Chapter 10 are partly based on the following published paper Moore and Gair (2015), and the following paper is currently in preparation Moore and Gair (in prep).

9.1 Bumpy black hole spacetimes

A metric that represents a rotating black hole is required to be stationary and axisymmetric (i.e. there exists a timelike Killing vector $\partial/\partial t$ and a spacelike Killing vector $\partial/\partial\phi$), invariant under simultaneous inversion of the ϕ and t coordinates, and reflection symmetric about the equatorial plane. A sufficiently general metric which captures all these properties is given by Chandrasekhar (2002)

$$\mathrm{d}s^2 = g_{tt}\,\mathrm{d}t^2 + g_{t\phi}\,\mathrm{d}t\,\mathrm{d}\phi + g_{rr}\,\mathrm{d}r^2 + g_{\theta\theta}\,\mathrm{d}\theta^2 + g_{\phi\phi}\,\mathrm{d}\phi^2\,, \tag{9.1}$$

where the metric coefficients depend only on the radial and polar coordinates r and θ. In fact Eq. 9.1 retains a considerable degree of gauge freedom, however, it is sufficient for our present purpose.

As our interest here is in testing the hypothesis that the metric is the Kerr solution, and it is known that the Kerr solution is an excellent description in the weak field, it is natural to expand the metric as

$$g_{\mu\nu} = g^{\mathrm{Kerr}}_{\mu\nu} + \epsilon h_{\mu\nu}\,, \tag{9.2}$$

where $\epsilon \ll 1$ and $h_{\mu\nu} \to 0$ as $r \to \infty$. The background Kerr solution is given by,

$$\begin{aligned} g^{\mathrm{Kerr}}_{tt} &= -\left(1 - \frac{2Mr}{\rho^2}\right)\,,\; g^{\mathrm{Kerr}}_{t\phi} = \frac{-2M^2ar}{\rho^2}\sin^2\theta\,, \\ g^{\mathrm{Kerr}}_{rr} &= \frac{\rho^2}{\Delta}\,,\; g^{\mathrm{Kerr}}_{\theta\theta} = \rho^2\,,\; g^{\mathrm{Kerr}}_{\phi\phi} = \frac{\Sigma^2}{\rho^2}\sin^2\theta\,, \end{aligned} \tag{9.3}$$

for a black hole with a mass M and dimensionless spin a, in which

$$\begin{aligned} &\rho^2 = r^2 + a^2M^2\cos^2\theta\,,\; \Delta = r^2\left(1 - \frac{2M}{r}\right) + M^2a^2\,, \\ &\text{and}\quad \Sigma^2 = \left(r^2 + M^2a^2\right)^2 - M^2a^2\Delta\sin^2\theta\,. \end{aligned} \tag{9.4}$$

From here on dimensionless units with $M = 1$ are used.

Geodesic motion with four-momentum p^μ in a metric with the symmetries outlined thus far would posses three constants of motion; energy $E = -p^\mu(\partial/\partial t)_\mu$, z-component

of angular momentum $L = p^\mu(\partial/\partial\phi)_\mu$, and particle rest mass $m^2 = -p^\mu p_\mu$. However the Kerr solution also possess an additional constant of motion related to the existence of a second rank Killing tensor, $C = p^\mu p^\nu \xi_{\mu\nu}$. In place of C the Carter constant is often defined as $Q = C - (L - aE)^2$ (Carter, 1968). The existence of the addition constant of motion ensures that geodesic motion in the spacetime is separable and tri-periodic (i.e. with a well defined frequency associated with motion in each of the r, θ and ϕ coordinates). Although it is not clear that it should be required of any alternative to the Kerr solution to possess these properties, they are sufficiently appealing that it is worth considering the possibility carefully.

The most general form of metric perturbation, $h_{\mu\nu}$, that could be added to the Kerr solution such that the resultant metric also admitted a second rank Killing tensor and hence a Carter-like constant (at least to $\mathcal{O}(\epsilon^2)$ in the perturbation) was considered by Vigeland et al. (2011). The metric perturbation was also required to tend to zero in the limit $r \to \infty$ faster than r^{-2} so that all weak field tests are still satisfied. A similar solution was also found earlier by Benenti and Francaviglia (1979). Working in Boyer-Linquist-like coordinates Vigeland et al. (2011) found a series of differential relations that must be satisfied by the metric perturbation components. The only non zero components of the perturbation are $\{h_{tt}, h_{t\phi}, h_{rr}, h_{\phi\phi}\}$; in particular $h_{\theta\theta}$ vanishes. In subsequent work Gair and Yunes (2011) expansions for the metric perturbation components in powers of $1/r$ were found,

$$h_{\mu\nu} = \sum_n h_{\mu\nu,n} \left(\frac{1}{r}\right)^n \tag{9.5}$$

where the coefficients $h_{\mu\nu,n}$ are functions of θ only. The leading order non-zero coeffeicients are given in Eq. 9.6 and all the coefficients up to $\mathcal{O}(1/r^5)$ are reproduced in Appendix E.1;

$$\begin{aligned}
h_{tt,2} &= \gamma_{1,2} + 2\gamma_{4,2} - 2a\gamma_{3,1}\sin^2\theta\,, \\
h_{rr,2} &= -\gamma_{1,2}\,, \\
h_{t\phi,2} &= -M\sin^2\theta\left[\gamma_{3,3} + a\left(\gamma_{1,2} + \gamma_{4,2}\right) + a^2\gamma_{3,1}\right]\,, \\
h_{\phi\phi,0} &= 2M^2 a\gamma_{3,1}\sin^4\theta\,.
\end{aligned} \tag{9.6}$$

Note that the leading order correction enters at a lower order in the $h_{\phi\phi}$ component. The components h_{tt} and h_{rr} are dimensionless, whilst $h_{t\phi}$ has units of length and $h_{\phi\phi}$ has units of length squared; reflecting the dimensions of the components of $g_{\mu\nu}$.

At low order in the expansion the metric perturbation is fully characterised by a small number of coefficients, $\gamma_{i,j}$. Adopting the notation of Gair and Yunes (2011) the metric perturbation up to $\mathcal{O}(1/r^2)$ is given by the four constants $\mathcal{B}_2 = (\gamma_{1,2}, \gamma_{3,1}, \gamma_{3,3}, \gamma_{4,2})$ (see Eq. 9.6); up to $\mathcal{O}(1/r^3)$ it is given by the 7 constants $\mathcal{B}_2 \cup \mathcal{B}_3$, where $\mathcal{B}_3 = (\gamma_{1,3}, \gamma_{3,4}, \gamma_{4,3})$; up to $\mathcal{O}(1/r^4)$ it is given by the 10 constants $\mathcal{B}_2 \cup \mathcal{B}_3 \cup \mathcal{B}_4$, where $\mathcal{B}_4 = (\gamma_{1,4}, \gamma_{3,5}, \gamma_{4,4})$; up to $\mathcal{O}(1/r^5)$ it is given by the 13 constants $\mathcal{B}_2 \cup \mathcal{B}_3 \cup \mathcal{B}_4 \cup \mathcal{B}_5$, where $\mathcal{B}_5 = (\gamma_{1,5}, \gamma_{4,5}, \gamma_{3,6})$.

Throughout this chapter various $\mathcal{B}_N$ limits will be referred to, these correspond to setting all the parameters $\gamma_{i,j}$ to zero except for those in the set $\mathcal{B}_N$, which are treated as independent free parameters. This greatly reduces the number of free parameters and allows for a systematic way of examining perturbations to the Kerr metric. Because the Kerr metric is known to be an excellent approximation at large radii it is natural to expect any deviation from the Kerr solution to show up, initially at least, at lowest order in $1/r$. This is the justification for examining each of the $\mathcal{B}_N$ limits separately. In general, it would be possible for any combination of the $\gamma_{i,j}$ to be be none-zero. In this chapter results are presented for the $\mathcal{B}_2$, $\mathcal{B}_3$, $\mathcal{B}_4$ and $\mathcal{B}_5$ metrics.

9.1.1 Known black hole solutions

In addition to the very general family of perturbed metrics described above, it is also useful to consider some examples of known black hole solutions in alternative theories of gravity. In this section a few such solutions are listed. One of these spacetimes (the linear in spin solution to dynamical Chern-Simons gravity, Sec. 9.1.1.2) is a member of the family of solutions described in Sec. 9.1; it can be obtained by a particular choice of the constants $\gamma_{i,j}$.

9.1.1.1 The Kehagias Sfetsos metric

A spherically symmetric black hole solution to Hořava gravity (Hořava, 2009*a*,*b*) was found by Kehagias and Sfetsos (2009). This was generalised to a slowly rotating solution by Lee et al. (2010). Accretion disk signatures for this type of black hole have been considered previously by Harko et al. (2011*b*). The metric is

$$\mathrm{d}s^2_{\mathrm{KS}} = -f_{\mathrm{KS}}(r)\mathrm{d}t^2 + \frac{\mathrm{d}r^2}{f_{\mathrm{KS}}(r)} + r^2\mathrm{d}\Omega^2 - \frac{4a\sin^2\theta}{r}\mathrm{d}\phi\mathrm{d}t\,, \tag{9.7}$$

where

$$f_{\rm KS}(r) = 1 + \omega r^2 \left(1 - \sqrt{1 + \frac{4}{\omega r^3}}\right) , \tag{9.8}$$

and

$$\mathrm{d}\Omega^2 = \mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2 . \tag{9.9}$$

In the limit $\omega \to \infty$ the slowly rotating limit of the Kerr metric is recovered. In order to avoid a naked singularity at the origin an extra constraint is needed, $\omega M^2 \geq \frac{1}{2}$. For the remainder of this section ω is made dimensionless by multiplying by M^2, and we choose to work with the small parameter $Y \equiv 1/(\omega M^2) \ll 1$. For the remainder of this section the metric in Eq. 9.7 will be refered to as the KS metric.

9.1.1.2 The Chern Simons metric linear in spin

Dynamical Chern Simons (CS) modified gravity (Jackiw and Pi, 2003) is a parity violating theory of gravity constructed by adding a Pontryagin invariant term to the action. As the Schwarzschild solution is spherically symmetric and has even parity it remains a solution to the modified CS field equations; however, the Kerr solution does not have even parity and fails to satisfy these equations. No complete rotating black hole solution is known in the CS theory, however perturbative solutions in the spin and CS coupling constant have been found analytically. The rapidly rotating case was considered numerically by Stein (2014), here we focus on the slowly rotating solutions.

The slowly rotating black hole solution (linear in spin) to dynamical Chern-Simons gravity was found by Yunes and Pretorius (2009);

$$\mathrm{d}s^2_{\rm CS1} = \mathrm{d}s^2_{\rm Kerr} + \frac{5}{4}\zeta\chi\frac{1}{r^4}\left(1 + \frac{12}{7r^2} + \frac{27}{10r^2}\right)\mathrm{d}t\mathrm{d}\phi . \tag{9.10}$$

Accretion disk signatures for this type of black hole have been considered previously by Harko et al. (2010). For the remainder of this section the metric in Eq. 9.10 will be refered to as the CS1 metric.

9.1.1.3 The Chern Simons metric quadratic in spin

A slowly rotating black hole solution (quadratic in spin) to dynamical Chern-Simons gravity was found by Yagi et al. (2012). Disk emission in this metric has not been considered before. The metric is

$$\mathrm{d}s^2_{\rm CS2} = \mathrm{d}s^2_{\rm Kerr} + \delta\left(g^{\rm CS2}_{\mu\nu}\right)\mathrm{d}x^\mu \mathrm{d}x^\nu \tag{9.11}$$

and expressions for the components $\delta\left(g_{\mu\nu}^{\rm CS2}\right)$ are given in Appendix E.2. For the remainder of this section the metric in Eq. 9.11 will be referred to as the CS2 metric.

Unlike the KS metric, both the CS1 and CS2 metrics recover the Schwarzschild solution in the limit $a \to 0$. This is associated with the fact that dynamical Chern Simons modified gravity is a parity violating theory of gravity and hence the spherically symmetric Schwarzschild solution in general relativity, which is parity even, remains a solution of the modified theory.

9.2 Disk emission in bumpy black holes spacetimes

The theory of thin accretion disks in the steady state was developed by Shakura and Sunyaev (1973) in the Newtonian case and generalised by Page and Thorne (1974) to the general relativistic case. It is assumed that the disk lies in the equatorial plane with material moving in (approximately) circular geodesic orbits, with a small radial accretion velocity superposed on the circular motion. In addition, it is assumed that the disk is thin (i.e. at a radius r the thickness $2h$ satisfies $h \ll r$; the disk height may depend on radius, $h = h(r)$, but for notational simplicity this dependence is suppressed here) and is in a steady state with a constant accretion rate $\dot{M}_0$. The disk extends from the innermost stable circular orbit (ISCO) of the spacetime to some finite outer radius (strictly, to be consistent with the steady state assumption, the disk must extend indefinitely; here the more conservative approach of leaving the outer radius as a free parameter is taken). When the accreting material reaches the ISCO it plunges quickly into the black hole without radiating any further or interacting viscously with the material at larger radii. The fact that the material plunges quickly after crossing the ISCO provides the physical motivation for imposing a zero torque boundary condition at the inner edge of the disk. This allows the differential equation governing the radial dependence of the flux to be integrated; see Eq. 9.26.

The model described above is often referred to as the standard relativistic model of accretion disks, and has been extensively studied for the Kerr geometry. Here it is necessary to retain sufficient generality to perform the calculations for any of the bumpy black holes described in Sec. 9.1. Since the disk is thin and assumed to lie in the equatorial plane we may switch from using Boyer-Lidquist polar-like coordinates to

cylindrical-like coordinates. The metric near the equatorial plane is given by

$$\mathrm{d}s^2 = g_{tt}\mathrm{d}t^2 + 2g_{t\phi}\mathrm{d}t\mathrm{d}\phi + g_{rr}\mathrm{d}r^2 + g_{\phi\phi}\mathrm{d}\phi^2 + g_{zz}\mathrm{d}z^2\,, \tag{9.12}$$

where the metric components depend only on r; corrections that depend on z enter at second order and are neglected because the disc is assumed to be thin. By a rescaling of the z coordinate the metric component g_{zz} may be set to unity without any loss of generality.

The metric in Eqs. 9.1 and 9.12 does not depend on the timelike or azimuthal coordinates, so the corresponding covariant components of the four-momentum are conserved, $p_t = -E$ and $p_\phi = L$. Using the metric to raise the indices on these momentum components gives the first two geodesic equations in first order form, where a tilde denotes an orbital quantity per unit mass of the test particle,

$$\begin{aligned}\frac{\mathrm{d}t}{\mathrm{d}\tau} &= \frac{\tilde{E}g_{\phi\phi} + \tilde{L}g_{t\phi}}{g_{t\phi}^2 - g_{tt}g_{\phi\phi}}\,,\\ \frac{\mathrm{d}\phi}{\mathrm{d}\tau} &= -\frac{\tilde{E}g_{t\phi} + \tilde{L}g_{tt}}{g_{t\phi}^2 - g_{tt}g_{\phi\phi}}\,,\end{aligned} \tag{9.13}$$

where τ is the proper time along the worldline of the orbiting test particle.

Because the material in the disk is orbiting in the equatorial plane the vertical component of its four-velocity vanishes, $\mathrm{d}z/\mathrm{d}\tau = 0$. Therefore, the third and final geodesic equation may be conveniently obtained from the normalisation condition of the four velocity;

$$g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\tau}\frac{\mathrm{d}x^\nu}{\mathrm{d}\tau} \equiv g_{\mu\nu}u^\mu u^\nu = -1\,, \tag{9.14}$$

$$\Rightarrow \left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 = \frac{V_{\mathrm{eff}}(r)}{g_{rr}}\,, \tag{9.15}$$

where,

$$V_{\mathrm{eff}}(r) = \frac{\tilde{E}^2 g_{\phi\phi} + 2\tilde{L}\tilde{E}g_{t\phi} + \tilde{L}^2 g_{tt}}{g_{t\phi}^2 - g_{tt}g_{\phi\phi}} - 1\,. \tag{9.16}$$

From Eq. 9.15 it can be seen that the conditions for stable circular orbits are

$$V_{\mathrm{eff}}(r) = V_{\mathrm{eff},r}(r) = 0\,, \tag{9.17}$$

where a comma in the subscript denotes a partial derivative with respect to all subsequent indices. These two conditions yield expressions for the specific energy and angular momentum per unit mass of a particle on a circular, equatorial geodesic;

$$\tilde{E} = -\frac{g_{tt} + g_{t\phi}\Omega}{\sqrt{-\left(g_{tt} + 2g_{t\phi}\Omega + g_{\phi\phi}\Omega^2\right)}} \tag{9.18}$$

$$\tilde{L} = \frac{g_{t\phi} + g_{\phi\phi}\Omega}{\sqrt{-\left(g_{tt} + 2g_{t\phi}\Omega + g_{\phi\phi}\Omega^2\right)}} \,. \tag{9.19}$$

Combining Eqs. 9.13 with Eqs. 9.18 and 9.19 gives an expression for the coordinate angular velocity of the particle,

$$\Omega = \frac{\mathrm{d}\phi}{\mathrm{d}t} = \frac{-g_{t\phi,r} + \sqrt{\left(g_{t\phi,r}\right)^2 - g_{tt,r}g_{\phi\phi,r}}}{g_{\phi\phi,r}} \,. \tag{9.20}$$

In the metrics of interest in this chapter (i.e. the Kerr metric and small perturbations from it) there exists an ISCO. The radius of the ISCO may be found by solving $V_{\mathrm{eff},rr}(r) = 0$ simultaneously with $V_{\mathrm{eff}}(r) = V_{\mathrm{eff},r}(r) = 0$. This equation admits the following analytic solution in the Kerr case (Bardeen et al., 1972);

$$\begin{aligned} Z_1 &= 1 + \left(1 - a^2\right)^{1/3}\left((1+a)^{1/3} + (1-a)^{1/3}\right) \,, \\ Z_2 &= \left(3a^2 + Z_1^2\right)^{1/2} \,, \\ r_{isco} &= 3 + Z_2 - \left((3 - Z_1)(3 + Z_1 + 2Z_2)\right)^{1/2} \,. \end{aligned} \tag{9.21}$$

For the general bumpy black holes discussed in Sec. 9.1 the radius for the ISCO must be found numerically. However, if the spacetime is characterised by a small deformation from Kerr, then the radius of the ISCO may be expanded perturbatively in the bump parameter, ϵ,

$$r_{isco} = r^{\mathrm{Kerr}} - \epsilon \frac{\left.\frac{\mathrm{d}V_{\mathrm{eff},rr}(r,\epsilon)}{\mathrm{d}\epsilon}\right|_{r=r^{\mathrm{Kerr}}_{isco},\epsilon=0}}{\left.\frac{\mathrm{d}V_{\mathrm{eff},rr}(r,\epsilon)}{\mathrm{d}r}\right|_{r=r^{\mathrm{Kerr}}_{isco},\epsilon=0}} + \mathcal{O}\left(\epsilon^2\right) \,. \tag{9.22}$$

If the metric is characterised by more than one bump parameter, ϵ_i, as is the case with the $\mathcal{B}_N$ spacetimes described in Sec. 9.1 then the following replacement should be made in Eq. 9.22,

$$\epsilon \left.\frac{\mathrm{d}V_{\mathrm{eff},rr}(r,\epsilon)}{\mathrm{d}\epsilon}\right|_{r=r^{\mathrm{Kerr}}_{isco},\epsilon=0} \to \sum_i \epsilon_i \left.\frac{\mathrm{d}V_{\mathrm{eff},rr}(r,\epsilon_i)}{\mathrm{d}\epsilon_i}\right|_{r=r^{\mathrm{Kerr}}_{isco},\epsilon_i=0} . \tag{9.23}$$

From Eq. 9.20, and the fact that the orbit is circular and equatorial, the four-velocity of a small fluid element of the disk orbiting the black hole is given by

$$\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}, \frac{\mathrm{d}r}{\mathrm{d}\tau}, \frac{\mathrm{d}\phi}{\mathrm{d}\tau}, \frac{\mathrm{d}z}{\mathrm{d}\tau}\right) = \left(u^t, u^r, u^\phi, u^z\right) = u^t\left(1, 0, \Omega, 0\right) \,, \tag{9.24}$$

where u^t may be found from the normalisation condition on the four velocity in Eq. 9.14,

$$u^t = \frac{1}{\sqrt{-\left(g_{tt} + 2\Omega g_{t\phi} + \Omega^2 g_{\phi\phi}\right)}} \,. \tag{9.25}$$

Viscous forces in the disk cause the orbiting material to dissipate energy and gradually move inwards to smaller radii. An expression for the radial dependence of the flux of energy from the disk was derived by Page and Thorne (1974); for completeness this derivation is reproduced in Appendix F. The resulting expression for the radial dependence of the flux is,

$$F(r) = \frac{-\dot{M}_0 \Omega_{,r}}{4\pi\sqrt{-\mathbf{g}}\left(\tilde{E} - \Omega\tilde{L}\right)^2} \int_{r_{\rm isco}}^{r} \left(\tilde{E} - \Omega\tilde{L}\right) \tilde{L}_{,r}\, \mathrm{d}r \,, \tag{9.26}$$

where $\dot{M}_0$ is the accretion rate and $\mathbf{g}$ is the metric determinant. Since the disk is in thermodynamic equilibrium, and not heating up or cooling down, this flux of energy is radiated away from the disk in the form of a thermal distribution of photons.

9.2.1 Line emission

In addition to the flux of thermal radiation a higher energy power-law component of hard X-ray photons is also observed (Zdziarski et al., 1994). For supermassive black holes with cooler disks this component can in fact dominate over the thermal component. The hard X-ray power-law component is generally believed to be caused by inverse Compton scattering of the thermal photons radiated from the disk by the hot surrounding corona. A certain fraction of the hard X-ray photons are radiated back towards the disk surface where, upon incidence, they produce fluorescent transition lines which may also be observed as a third component of the spectrum. One frequently observed line is that due to the Iron (Fe) Kα transition, which in its rest frame has an energy of 6.38 keV. The combined effects of gravitational redshifting and Doppler boosting broaden this line into the characteristic shapes shown in Fig. 9.1.

The observed line profile can be computed as follows. Let $\mathrm{d}F_0(E_0)$ be the infinitesimal element of flux at energy E_0 observed at infinity due to an infinitesimal element of the disk. If $I_0(E_0)$ is the specific intensity (i.e. the energy per unit time, per unit area, per unit spectral energy (or frequency), per unit solid angle) in the observer's rest frame then

$$\mathrm{d}F_0(E_0) = I_0(E_0)\mathrm{d}\Xi \,, \tag{9.27}$$

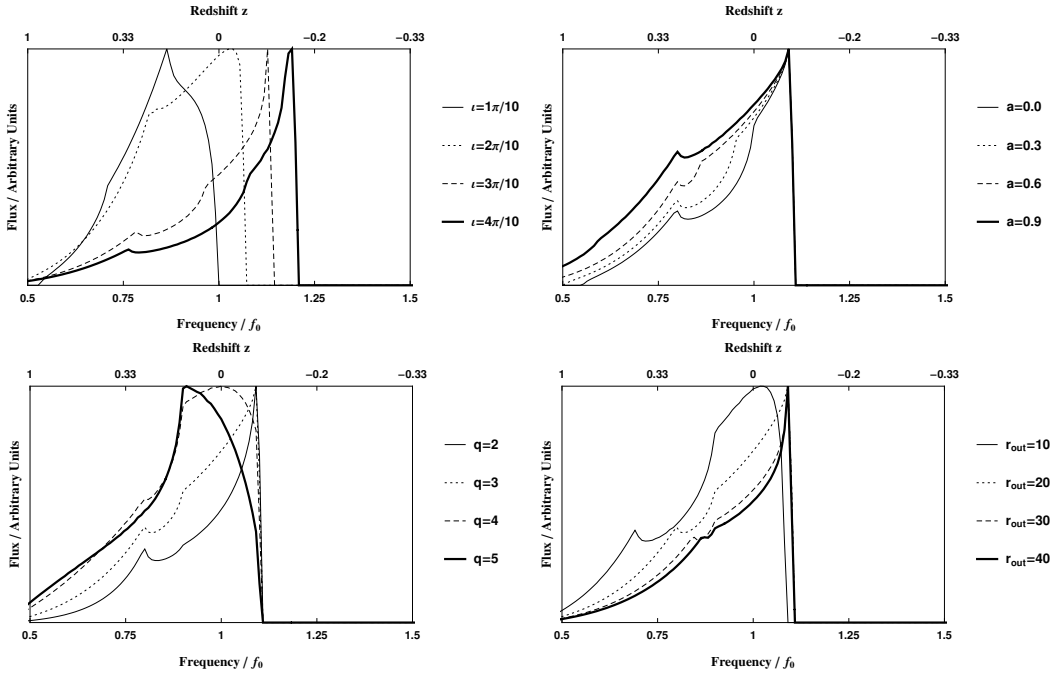


Figure 9.1: The dependence of the Kerr Iron line profile on various parameters. All spectra were normalised to the same peak flux. Unless otherwise stated in the figure all parameters were set to their fiducial values; $a = 0.5$, $\iota = \pi/4$, $q = 3$ and $r_{\rm out} = 30$.

where the element of the disk subtends solid angle element $\mathrm{d}\Xi$. If the disk has a specific intensity $I_e(E_e)$ in its own rest frame, given the invariance of I/E^3 along the photon's worldline (Misner et al., 1973), it follows that

$$\mathrm{d}F_0(E_0) = g^3 I_e(E_e)\mathrm{d}\Xi \,, \tag{9.28}$$

where $g \equiv 1/(1+z) \equiv E_0/E_e$ is the red-shift factor. For line emission the specific intensity may be approximated as a delta-function, $I_e(E_e) = \epsilon(r_e, \mu_e)\delta(E_e - E_{int})$, where E_{int} is the energy of the transition line in its rest frame, ϵ is the emissivity of the material, r_e is the radius of the emitting fluid element and μ_e is the cosine of the angle of the emission with respect to the disk normal. Eq. 9.28 together with the delta-function expression for the line specific intensity gives

$$\mathrm{d}F_0(E_0) = g^4 \epsilon(r_e, \mu_e)\delta(E_0 - gE_{\rm int})\mathrm{d}\Xi \,; \tag{9.29}$$

the extra factor of g in Eq. 9.28 compared to Eq. 9.29 comes from the change of argument

in the delta-function. Now let r_e and ϕ_e be the plane polar coordinates in the disk of the emitting point (where $\phi_e = 0$ is along the line of nodes where the disk intersects the observers plane of the sky). In addition let α and β be the cartesian coordinates in the observer's plane of the sky (where the x axis appears to lie along the lines $\phi_e = 0$ and $\phi_e = \pi$), and r_0 be the distance to the source. It follows that the solid angle element is given by $\mathrm{d}\Xi = \mathrm{d}\alpha\mathrm{d}\beta/r_0^2$, and therefore integrating Eq. 9.29 gives

$$F_0(E_0) = \frac{1}{r_0^2}\iint \mathrm{d}\,\alpha\,\mathrm{d}\beta\; g^4\epsilon(r_e,\mu_e)\delta(E_0 - gE_{\rm int})\;. \tag{9.30}$$

If light-bending is neglected then the polar coordinates of the emitting point in the disk (r_e and ϕ_e) may be related to the cartesian coordinates of the point in the observer's plane of the sky (α_e and β_e) via the transformation,

$$\alpha_e = r_e\cos\phi_e\;,\quad \beta_e = r_e\sin\phi_e\cos\iota\;, \tag{9.31}$$

with Jacobian

$$\frac{\partial(\alpha,\beta)}{\partial(r_e,\phi_e)} = r_e\cos\iota\,, \tag{9.32}$$

where ι is the inclination from which the disk is viewed (i.e. the angle between the observer's line of sight and the spin axis of the BH). For further discussion of this assumption see Sec. 9.2.3. With this in hand we may change the integral in Eq. 9.30 from being over the observer's plane of the sky to being over the disk itself (Dropping the subscript e),

$$F_0(E_0) = \frac{\cos\iota}{r_0^2}\int_0^{2\pi}\int_{r_{isco}}^{r_{out}} r\mathrm{d}r\mathrm{d}\phi\, g^4\epsilon(r,\mu)\delta\left(E_0 - gE_{int}\right)\;. \tag{9.33}$$

Hereafter the emissivity is taken to be a function of r only (no μ_e dependence), and is parameterised as a power law, $\epsilon = r^{-q}$ where q will be referred to as the emissivity index. We have now reduced the problem to finding the red-shift factor as a function of the position in the disk. The red-shift factor is defined by

$$g \equiv \frac{E_0}{E_e} = \frac{p_\mu v^\mu\big|_{\rm observer}}{p_\mu u^\mu\big|_{\rm emitter}}\;, \tag{9.34}$$

where p_μ is the four-momentum of the photon linking the emitter to the observer, v^μ is the four-velocity of the observer, and u^μ is the orbital four-velocity of the emitter (see Eq. 9.24). Since the metric is independent of both the t and ϕ coordinates, the corresponding components of the photon's four-momentum are conserved;

$$[p_\mu] = (p_t, p_r, p_\theta, p_\phi) = (-E, p_r, p_\theta, \Lambda)\,. \tag{9.35}$$

If the observer is at infinity and is at rest with respect to the black hole the numerator of Eq. 9.36 is simply given by $p_\mu v^\mu|_{\rm observer} = -E$. Using the orbital four-velocity from Eq. 9.24 the denominator of Eq. 9.36 is $p_\mu u^\mu|_{\rm emitter} = -Eu^t + \Omega u^t \Lambda$, and hence the red-shift factor is given by

$$g = \frac{1}{u^t(1-\Omega\lambda)} \,, \tag{9.36}$$

where $\lambda = \Lambda/E$. Since λ is conserved along the photon's worldline it may be evaluated at infinity where the indices on the photon four-momentum may be raised and lowered using the usual Minkowski metric. At large distance the ratio of the contravariant components of the photon's four-momentum gives the azimuthal impact parameter α, defined in Eq. 9.31,

$$\alpha = -\frac{rp^\phi}{p^t}\Big|_{r\to\infty} = -\frac{\lambda}{\sin\iota} \,. \tag{9.37}$$

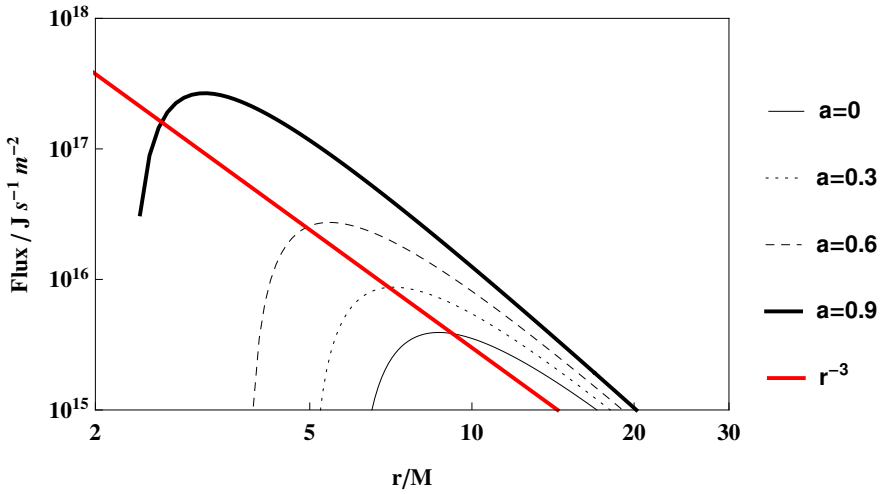


Figure 9.2: The Flux as a function of radius, for various values of the spin parameter. Also shown is the power law, $(r/M)^{-3}$. All other parameters were set to their fiducial values, $\iota = \pi/4$, $M = M_\odot$, $\dot{M}_0 = 10^{-12} M_\odot/{\rm yr}$ and $r_{\rm out} = 30$.

If light-bending is neglected then the photon linking the emitting point in the disk to the observer at infinity travels on a straight line with impact parameter $\alpha = r\cos\phi$ (see Eq. 9.31). Substituting these results into Eq. 9.36 gives the red-shift;

$$g\,(r,\phi) = \frac{\sqrt{-\left(g_{tt} + 2\Omega g_{t\phi} + \Omega^2 g_{\phi\phi}\right)}}{1 + \Omega r\cos\phi\sin\iota} \,. \tag{9.38}$$

The flux integral in Eq. 9.33 may now be simply evaluated numerically. This was done initially for the unperturbed Kerr metric in order to reproduce known results (see, for example, Kojima (1991); Laor (1991)) and examine the dependence of the spectra on various parameters. The results of these calculations are shown in Fig. 9.1. For all calculations in this work, unless otherwise stated, the disk parameters were set to the following fiducial values; $a = 0.5$, $\iota = \pi/4$, $q = 3$ and $r_{\rm out} = 30$.

It can be seen from Fig. 9.1 that the profile is strongly sensitive to disk inclination. Qualitatively this is because the orbiting material is moving at relativistic speeds and for large inclinations material on one side of the disk is moving towards the observer while on the other side it is moving away. This produces both a blue and a red-shifted horn to the profile, the blue-shifted horn is always more intense due to the relativistic beaming effect which enhances the intensity of light emitted in the direction of travel (the headlight effect).

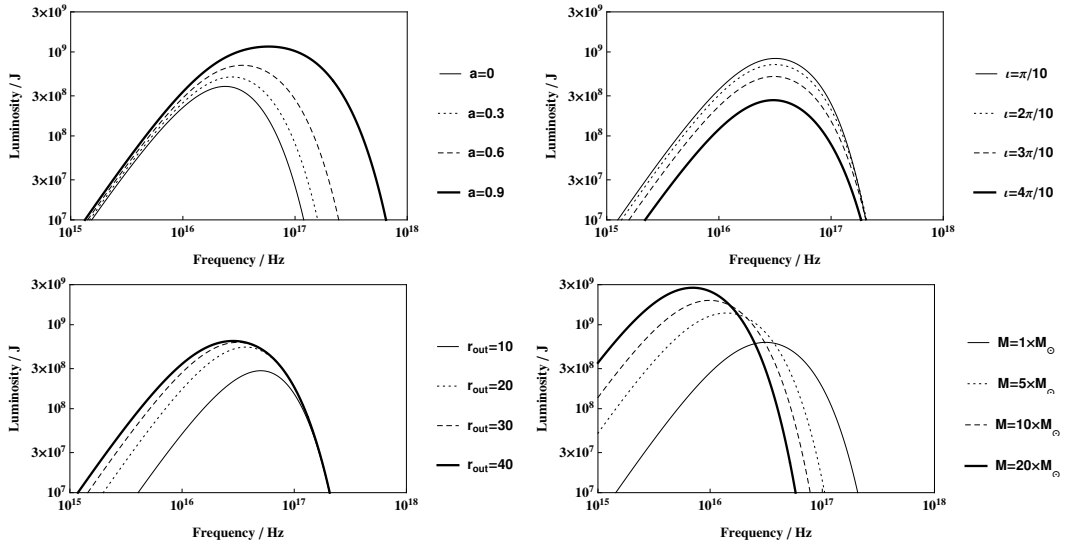


Figure 9.3: The dependence of the Kerr black body spectrum on various parameters. Unless indicated otherwise all parameters were set to their fiducial values, $a = 0.5$, $\iota = \pi/4$, $M = M_\odot$, $\dot{M}_0 = 10^{-12} M_\odot$/yr and $r_{\rm out} = 30$.

The profile is also strongly sensitive to the radius of the inner edge of the disk, which in the Kerr case (under the modelling assumptions made here) is monotonically related to the spin parameter, a (see Eq. 9.21). For higher spins (and smaller values of r_{isco})

it can be seen from Fig. 9.1 that the red-shifted wing of the line profile is much more prominent. Under the assumptions that the black hole is Kerr and the inner edge of the disk is located at the ISCO the line profile can be used as an accurate probe of the black hole spin, indeed such observations are among the best evidence for the existence of near maximally spinning black holes in some active galactic nuclei (see, for example, Iwasawa et al. (1996)).

The line profile also depends strongly on q; this parameter must be fit for simultaneously with all other parameters. The fiducial value $q = 3$ was chosen as this is approximately the form of the gravitational energy release per unit area of the disk (see Fig. 9.2). Likewise the value of r_{out} must also be fit for, but this is less problematic because (at least for $r_{out} \gtrsim 20$) the profile only depends weakly r_{out}.

9.2.2 Black-body spectra

The line emission profiles calculated above encode how the gravitational and Doppler shifting (and lightbending if included) affect a single frequency source in the disk. If instead there is a broadband source the resulting spectrum may be obtained by convolving the spectrum in the source's rest frame with the line profile obtained in the preceding section. This method can be used to calculate the black body spectra by convolving with the Plank distribution of a black body; this approach is described by Cunningham (1975).

Here we take the alternative, and more physically motivated approach of directly integrating the flux over the disk to find the spectra. From Eq. 9.33 the radial flux (shown in Fig. 9.2) is known, and assuming that the radiation is that of a black body the Stefan-Boltzmann law gives the radial temperature distribution in the disk

$$T(r) = \left(\frac{F(r)}{\sigma}\right)^{1/4} , \tag{9.39}$$

where σ is the Stefan-Boltzmann constant. Hence the thermal spectrum may be written as the following integral where the gravitational effects on the spectrum are included through the redshift factor,

$$F_0(E_0) = \frac{8\cos\iota}{\pi} \int_{r_{isco}}^{r_{out}} \int_0^{2\pi} r \mathrm{d}\phi \, \mathrm{d}r \, \frac{E_0^3}{g^3 \left(e^{\frac{E_0}{gT}} - 1 \right)} . \tag{9.40}$$

As was the case for the line emission in Eq. 9.33, the thermal spectrum in Eq. 9.40 is

now in the form of an integral over the disk in r and ϕ coordinates; this was evaluated numerically. Several example thermal spectra are shown for the Kerr metric in Fig. 9.3.

The thermal spectra of the disk is the convolution between the line emission spectra and the Planck distribution. Since the Planck distribution contains no information about the gravitational field the thermal spectra in Fig. 9.3 contain the same information about the black hole as the line spectra in Fig. 9.1, only significantly smoothed out. Because of this smoothing it appears that the thermal spectra are less distinctive, and that therefore it would be harder to measure the parameters of the black hole using thermal emission than line emission. However, the thermal spectra may be observed across a much wider frequency range and at a larger signal-to-noise ratio. Therefore it is not obvious *a priori* which method offers the best opportunity to constrain the metrics in Sec. 9.1. In practice the most suitable technique depends upon the mass of the black hole; Iron line emission is typically most suitable for supermassive black holes whilst thermal emission is used for the hotter disks around stellar mass black holes. In rare instances both techniques may be used simultaneosuly, and they have been demonstrated to give consistent resuts (Steiner et al., 2011). Henceforth we consider only the Iron line emission, and we leave a detailed study of parameter estimation using thermal emission to future work.

9.2.3 The effect of lightbending

The formalism for calculating both the Iron line and thermal emission outlined in the preceding sections assumed that points in the image plane could be related to points in the disk by *straight lines*, in the sense that the Boyer-Lindquist-like coordinates were treated as if they were spherical polar coordinates in flat space (these assumptions are summarised in Eq. 9.31). As the light originates from the strong gravitational field the effects of lightbending (including frame dragging, if $a \neq 0$) may be significant.

On the other hand it may be hoped that the effect of lightbending will vary slowly with changing system parameters, $\vec{\theta}$. If this is the case then while the inclusion of lightbending will have a significant impact on the observed profile it will have only a limited effect on our ability to measure the disk parameters, and hence on our ability to constrain deviations from the Kerr metric. In the language of the Fisher matrix calculations in Sec. 9.3, even if the spectra depends strongly on whether or not lightbending is included if the derivatives of the spectra do not then the Fisher matrix remains unchanged. In order to test the validity of this assumption a small number of calculations were performed

including the effect of lightbending, and the results compared.

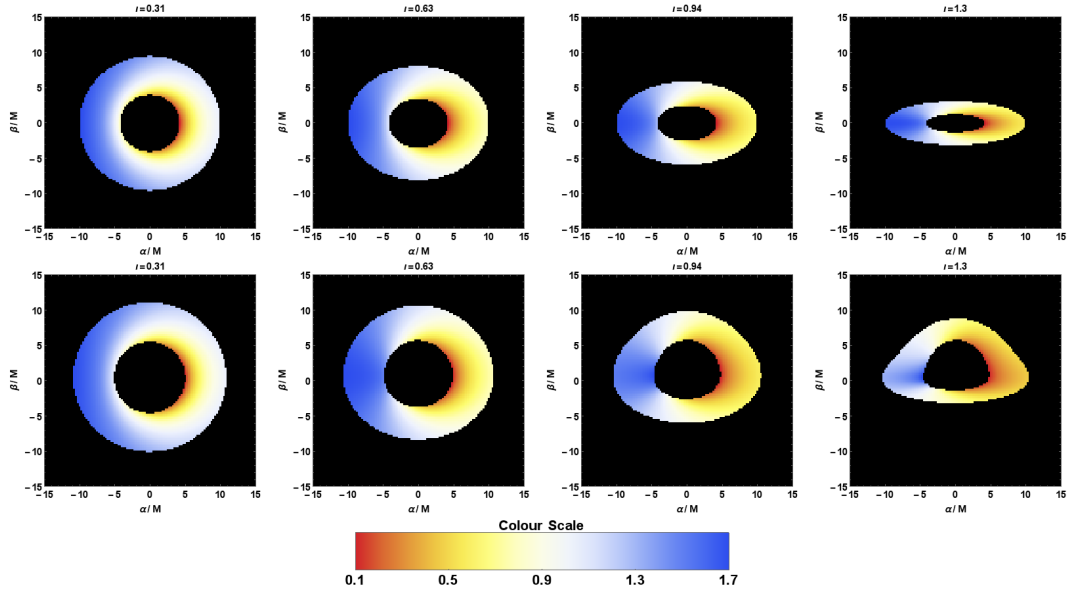


Figure 9.4: A set of images showing the resolved appearance of an accretion disk around a Kerr black hole. The colour indicates the redshift of the light from that portion of the disk; the top row shows the disk with no lightbending and the bottom row shows the disk including lightbending. The inclination angle is varried between plot; from left to right it takes the values $\iota = \{\pi/10, 2\pi/10, 3\pi/10, 4\pi/10\}$. The spin parameter and outer radius of the disk were fixed to $a = 0.5$ and $r_{out} = 10$.

A method for ray-tracing photons from the image plane to the disk was outlined by Psaltis and Johannsen (2012). The method uses the fact that the spacetime is stationary and axisymmetric to write the t and ϕ geodesic equations for the photon in first order form, but integrates the r and θ geodesic equations in second order form. This is ideal for our present purpose as the method does not require the existence of a Carter-like constant, which does not exist in all the metrics discussed in Sec. 9.1. If a fourth and final constant of motion does exist (as is the case for the $\mathcal{B}_N$ metrics) then evaluating it along the resulting trajectory provides a convenient check on the numerical accuracy of the integration.

The effect of lightbending on the appearance of a spatially resolved disk is shown in Fig. 9.4 for varying inclination (the colour scale indicates redshift factor, g). Particularly

at high inclinations the inclusion of lightbending significantly alters the appearance of the disk; the effect is most pronounced for large values of spin and inclination where the light from the far-side, inner edge passes very close to the horizon. However, the disk is not spatially resolved by our telescope, instead we observed the integrated flux across the disk. This shows a much smaller difference. The effect of including lightbending is also more significant for higher values of the spin parameter, because the disk extends closer to the black hole where the gravitational effects are stronger.

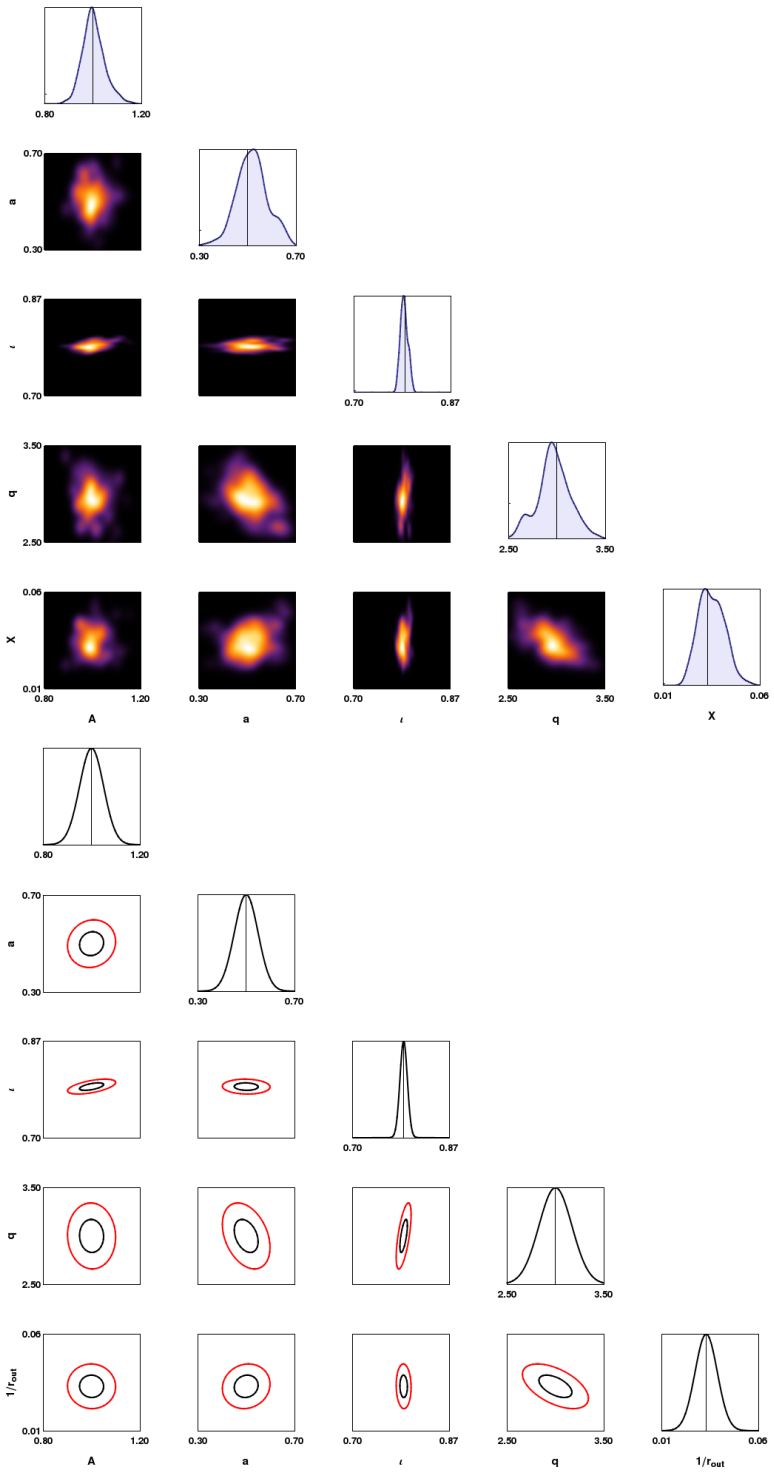


Figure 9.5: A comparison of the likelihood surface predicted by the Fisher matrix (indicated by 1σ (black) and 2σ (red) contours) and the true posterior (shown as a density plot using samples produced from an MCMC algorithm). All parameters were set to their fiducial values; $a = 0.5$, $\iota = \pi/4$, $q = 3$ and $r_{\rm out} = 30$.

For example, for the fiducial black hole parameters the Fisher matrix estimates an error on the dimensionless spin parameter of $\Delta a = 0.06$ if light bending is neglected (see right hand panel of Fig. 9.5). If instead lightbending is included then the same Fisher matrix analysis yields an error estimate for the spin parameter of $\Delta a = 0.03$. Changes in the error estimates of a factor of ~ 2 were observed for the other parameters.

As anticipated above larger differences were obtained for higher values of spin and lower values of inclination, and smaller differences in the opposite limits. These changes are not enough to affect our conclusions in this section.

9.3 The Fisher matrix

For each black hole metric described in Sec. 9.1 the Iron line and thermal spectra are characterised by a small number of parameters $\vec{\theta}$; e.g. for the Iron line spectra $h(\vec{\theta})$, $\vec{\theta} = (A, a, \iota, q, r_{out}, \vec{\epsilon})$, where A is an overall multiplicative factor relating to the unknown total luminosity and distance to the source, a is the spin parameter, ι is the disk inclination, q is the emissivity index described in Sec. 9.2.1, r_{out} is the outer radius of the disk and $\vec{\epsilon}$ is the vector of any metric deformation parameters. Given a measured spectrum, s, the challenge is to infer the posterior probability density on these parameters, $P(\vec{\theta}|s, I)$. The peak of this distribution is positioned at the best estimate of $\vec{\theta}$, and the characteristic width of the peak in each parameter direction indicates the uncertainty. The posterior probability density is related to the likelihood of the data given the parameters, $P(s|\vec{\theta}, I)$, via Bayes theorem,

$$P(\vec{\theta}|s) = \frac{P(s|\vec{\theta})P(\vec{\theta})}{P(s)} \; ; \tag{9.41}$$

where $P(\vec{\theta})$ is the prior probability density of the parameters and $P(s)$ is the normalisation constant known as the evidence,

$$P(s) = \int \mathrm{d}\vec{\theta}\, P(s|\vec{\theta})P(\vec{\theta}) \; . \tag{9.42}$$

For all calculations performed in this chapter flat priors were assumed over all physically allowed regions of parameter space so the posterior is simply proportional to the likelihood within this region.

In order to calculate the likelihood it is first necessary to make some assumptions about the performance of our detector. For simplicity we assume that across the entire energy (frequency) range the detector has a constant energy (frequency) resolution ΔE

(Δf), and that in each energy (frequency) bin there is an independent Gaussian error of size σ. The values of $\Delta E = 100\,\mathrm{eV}$ ($\Delta f = 2.4 \times 10^{14}\,\mathrm{Hz}$) and $\sigma = 0.05$ were choosen to make the hypothetical instrument broadly equivalent to the best current observations. For instance, the error on the spin measurement for the fiducial disk parameters under these assumption is $\Delta a = 0.06$ (see, for example, the results in Fig. 9.5).

These simplifying assumptions neglect some important effects. For example, the frequency resolution of most X-ray detectors changes substantially across observable bandwidth, the random errors in each frequency bin are due to variations in the arrival rate of photons which is a Poisson not a Gaussian process (although in the limit of high signal-to-noise the assumption of Gaussian errors becomes correct), and in addition to the uncorrelated random errors there will be systematic errors which may be correlated between frequency bins. Nevertheless, under these simplifying assumptions the likelihood is given by the following expression,

$$\begin{aligned}\mathcal{L}(\vec{\theta}) &= \frac{\exp\left(\frac{-1}{2}\sum_{i=1}^{N}\frac{\left(s_i - h_i(\vec{\theta})\right)^2}{\sigma^2}\right)}{\sqrt{2\pi}\sigma^N} \\ &= \frac{\exp\left(\frac{-1}{2}\left\langle s - h(\vec{\theta}) | s - h(\vec{\theta})\right\rangle\right)}{\sqrt{2\pi}\sigma^N}\,,\end{aligned} \tag{9.43}$$

where the inner product has been defined as

$$\langle a|b\rangle = \sum_{i=1}^{N}\frac{a_i b_i}{\sigma^2}\,. \tag{9.44}$$

In general the likelihood (Eq. 9.43) is a complicated function of the parameters, $\vec{\theta}$. However, expanding the signal about the true parameter values, $\vec{\theta_0}$, (using the Einstein summation convention)

$$h(\vec{\theta}) = h(\vec{\theta_0}) + \frac{\partial h}{\partial\theta_i}\Big|_{\vec{\theta}=\vec{\theta}_0}\delta\theta_i + \mathcal{O}\left((\delta\theta_i)^2\right)\,, \tag{9.45}$$

and using $s = n + h(\vec{\theta_0})$ (where n is the particular realisation of the noise observed in the detector) gives

$$\begin{aligned}\mathcal{L}(\vec{\theta}) &\approx \frac{\exp\left(\frac{-1}{2}\left[(n|n) - 2\left(n|\frac{\partial h}{\partial\theta_i}|_{\vec{\theta}=\vec{\theta}_0}\delta\theta_i\right) + \Sigma_{ij}\delta\theta_i\delta\theta_j\right]\right)}{\sqrt{2\pi}\sigma^N}\,, \\ &\text{where } \Sigma_{ij} = \left\langle \frac{\partial h(\vec{\theta})}{\partial\theta_i}\Big|_{\vec{\theta}=\vec{\theta}_0}\Big|\,\frac{\partial h(\vec{\theta})}{\partial\theta_j}\Big|_{\vec{\theta}=\vec{\theta}_0}\right\rangle\,.\end{aligned} \tag{9.46}$$

Therefore, within the linear signal approximation used in Eq. 9.45, the likelihood is a multivariate Gaussian, peaked a noise-realisation-dependent distance away from the true parameters, and with a covariance matrix given by the inverse of the Fisher information matrix, Σ_{ij}. An estimate for the error in each parameter may be read off from the corresponding component of the covariance matrix,

$$\Delta\theta_i = \sqrt{(\Sigma^{-1})_{ii}} \quad \text{(no sum on } i\text{)}. \tag{9.47}$$

The Fisher matrix formalism was used first to estimate the parameter estimation accuracy for a disk in the Kerr metric (Sec. 9.5.1 below) and subsequently for all of the bumpy black hole spacetimes discussed in Sec. 9.1 (Secs. 9.5.2 to 9.5.4 below). In the case of the bumpy black hole spacetimes the true value of the bump parameters were set to zero, i.e. the Kerr metric was used, and the error estimate obtained for the bump is reported. The error on the bump parameter(s) are then interpreted as an estimate of the bound it may be possible to place on the size of the bump; i.e. if the true value of the bump parameter took this value then it would be marginally detectable with these observations. This bound should be interpreted as a lower limit, i.e. a best case scenario; in reality even if a non-zero value of a given deformation parameter was returned in a particular experiment it would still be a non-trivial task to rule out more mundane explanations. For example, before claiming a detection of a deviation from the Kerr solution it would presumably be necessary to consider more complicated forms for the radial emissivity, $\epsilon(r)$, than a simple power law. The free parameters in this new emissivity law would then have to be marginalised over and this would have the effect of increasing the errors on the other parameters. Other possibilities must also be considered; for example thick accretion disks, emitting material within the ISCO, reprocessing of the light by surrounding material, etc.

The applicability of the Fisher matrix rests on the validity of the linear signal approximation in Eq. 9.45; this must hold at least within a few standard deviations from the peak in all directions in parameter space. In general it is impossible to know from the Fisher matrix alone whether one is within the region where the linear signal approximation may be safely applied. This question of the applicability of the Fisher matrix is addressed in Sec. 9.4.

9.4 Verification of the applicability of the Fisher matrix formalism

The gold standard for parameter estimation is to numerically calculate the likelihood (or in general the posterior) surface over the region of parameter space of interest. This may be achieved in low dimensional problems by simply evaluating the likelihood function on a grid of parameter points. Alternatively, and more efficiently in high dimensional problems, there exist a variety of Markov chain Monte Carlo (MCMC) algorithms designed to sample points from the target probability distribution; the simplest of these is the Metropolis-Hastings algorithm. In Sec. 9.4.1 a MCMC analysis is performed on a sample Iron line spectra for a typical case; the resulting likelihood surfaces are compared with those predicted using the Fisher matrix.

MCMC type algorithms quickly become (prohibitivly) expensive as the dimension of the parameter space increases. A consistency check on the Fisher matrix was proposed by Vallisneri (2008) which is relativly quick to implement. The consistency check involves testing the validity of the linear signal approximation. For some central parameter values, $\vec{\theta_0}$, a point $\vec{\theta}$ is picked at random from the 1σ surface estimated by the Fisher matrix. The likelihood is then evaluated exactly, and approximated with the linear signal approximation, the ratio of the two likelihoods is denoted $r(\vec{\theta})$. The logarithm of this ratio is given by

$$\left|\log r(\vec{\theta})\right| = \frac{1}{2}\left(\Delta\theta_i h_i - \Delta h(\vec{\theta})\right|\Delta\theta_i h_i - \Delta h(\vec{\theta})\Big)\ , \tag{9.48}$$

where

$$\Delta\vec{\theta} = \vec{\theta} - \vec{\theta}_0\ ,\ \Delta h(\vec{\theta}) = h(\vec{\theta}) - h(\vec{\theta}_0)\ ,\text{and}\ \ h_i = \left.\frac{\partial h(\vec{\theta})}{\partial\theta^i}\right|_{\vec{\theta}=\vec{\theta}_0}\ . \tag{9.49}$$

Small values of $|\log r(\vec{\theta})|$ incate that the linear signal approximation is holding out as far as the 1σ surface in that particular parameter direction. This procedure may then be repeated for many points drawn randomly from the 1σ surface to assess whether the approximation holds in all directions. It should be stressed that this only checks the internal consistency of the linear signal approximation and does not guarantee the accuracy of the Fisher matrix. In Sec. 9.4.2 this consistency check was performed for the Iron line emission likelihood surface, as the consistency check is faster than a full MCMC it was performed for a range of spin and inclination parameter values.

These two checks on the applicability of the Fisher matrix give us increased confidence in the results for the bounds on the various metric deformations found in Sec. 9.5.

9.4.1 MCMC

A simple Metropolis-Hastings MCMC was used to sample from the likelihood distribution in Eq. 9.43 with the Iron line spectra described in Sec. 9.2.1. The resulting chain, plotted as a density histogram, is shown in Fig. 9.5 alongside the Gaussian contours from the Fisher matrix analysis. From Fig. 9.5 it can be seen that there is unquestionably additional structure in the true posterior which is not captured by the Fisher matrix; however the widths of the Fisher matrix Gaussians give a good indiction of the scale of the true posterior.

9.4.2 Mismatch ratio

For a range of values of a and ι the Fisher matrix was evaluated, and used to choose 100 points, randomly distributed, on the 1σ surface. The mismatch ratio was evaluated for all of these points and the results are summarised in Tab. 9.1; all values are less than unity, indicating that the Fisher matrix performs well for this problem.

	$\iota = \frac{\pi}{10}$	$\iota = \frac{2\pi}{10}$	$\iota = \frac{3\pi}{10}$	$\iota = \frac{4\pi}{10}$
$a = 0.1$	0.0044/0.013	0.0088/0.032	0.025/0.15	0.10/0.50
$a = 0.3$	0.0046/0.016	0.0091/0.030	0.022/0.064	0.032/0.14
$a = 0.5$	0.0037/0.0069	0.011/0.026	0.019/0.083	0.028/0.077
$a = 0.7$	0.0039/0.0095	0.012/0.036	0.017/0.061	0.028/0.088
$a = 0.9$	0.0078/0.015	0.014/0.028	0.027/0.053	0.031/0.10

Table 9.1: All entries are of the form μ/M, where μ is the mean value of the logarithm of mismatch ratio calculated on 100 points selected uniformly from the 1σ surface and M is the maximum value of the mismatch ratio on the same set of points. All other parameters were set to their fiducial values of $q = 3$ and $r_{\rm out} = 30$. (The forward slash between entries does *not* indicate division.)

9.5 Results

9.5.1 Kerr errors: Iron line

Before attempting calculations in the bumpy black hole spacetimes the Fisher matrix formalism was used to assess the accuracy with which it is possible to measure the standard parameters $\{A, a, \iota, q, X\}$ with the Iron line observations described in Sec. 9.2.1. The results are shown in Fig. 9.6. In particular we note that the spin parameter can usually be measured with an error of $\Delta a \sim 0.1$. The spin parameter can be measured most accurately for highly spinning black holes; this is because the location of the inner edge of the disk varies most strongly with black hole spin for near-extremal black holes. The inner edge of the disk is also the most strongly radiating region of the disk. In Fig.9.6 the error estimate on the spin parameter decreases by a factor of ~ 2 over the range of spin parameters. However these calculations were performed holding the total emission, A, constant. If instead the emission per unit area of the disk were held constant, a much stronger dependence of the the spin parameter would be observed as most radiation would come from the inner edge.

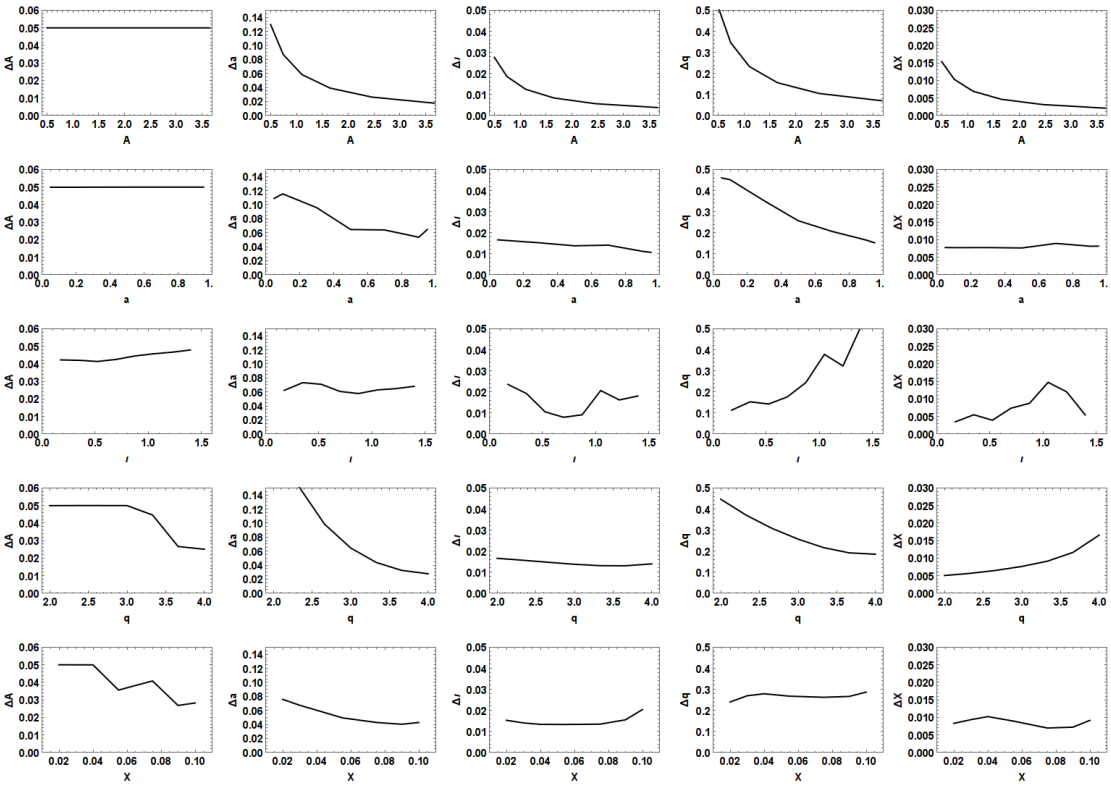


Figure 9.6: Each panel of this figure gives the error in the column parameter versus the value of the row parameter for fixed fiducial values of the other parameters; $a = 1$, $\iota = \pi/4$, $q = 3$ and $r_{\rm out} = 30$.

9.5.2 KS metric

The formalism described in Sec. 9.2 was used to calculate the Iron line profile in the KS metric, Eq. 9.7. As described in Sec. 9.2 the shift in the position of the ISCO, relative to the Kerr values, to first order in the small parameter $1/\omega$ may be calculated; $r_{isco} = r_{isco}^{\rm Kerr} + \Delta r_{isco}$, where

$$\Delta r_{isco} = \frac{-11}{36\omega} - \frac{59a}{54\sqrt{6}\omega}\,. \tag{9.50}$$

The ISCO moves inwards for increasing deformation, therefore the effect of a large value of $1/\omega$ is to boost the redshifted red wing of the line profile relative to the blue-shifted peak, similar to the effect of increasing spin; see Fig. 9.7.

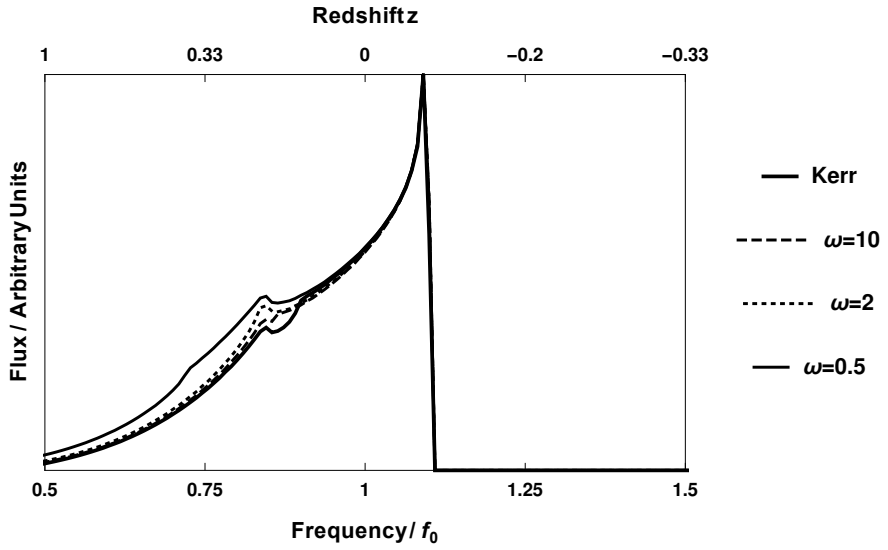


Figure 9.7: A series of Iron line spectra for accretion disks around KS black holes varying the value of ω. All other parameters were set to fiducial values, $a = 0.5$, $\iota = \pi/4$, $q = 3$ and $r_{\rm out} = 30$.

The 1σ and 2σ contours from a Fisher matrix analysis are plotted in Fig. 9.8, it can be seen that there is a rather stark degeneracy between the spin and deformation parameter. By comparing Fig. 9.8 to Fig. 9.5 it can be seen that the errors in all other parameters are virtually unaffected by the inclusion of the deformation parameter. The effect of the degeneracy, apart from inhibiting any measurement of the spin parameter, is to make it very difficult to place any bound on the deformation; Tab. 9.2 gives the

bound it is possible to place on ω for different values of spin. It should be remembered that the KS metric is valid only to linear order in a, and in addition the metric exhibits unphysical properties (e.g. naked singularities, closed timelike curves, etc) for $\omega < 1/2$. Therefore, bearing in mind the fact that the bounds derived from the method discussed in Sec. 9.3 should be treated as lower limits, the conclusion to be drawn from Tab. 9.2 is that it would be extremely difficult to place any constraint on the KS deformation parameter using observations of this type.

There have been several attempts to place bounds on the KS deformation parameter using solar system tests of gravity such as perihelion precession, deflection of light by the Sun and radar echo delay observations (Harko et al., 2011*a*; Iorio and Ruggiero, 2011; Liu et al., 2011). The bounds obtained from these studies are typically on the length scale $\omega \gtrsim 5 \times 10^{-28}\,\mathrm{m}$, corresponding to a dimensionless bound $\omega M^2 \gtrsim 3 \times 10^{-16}$, with $M = M_\odot$. It should be noted that these bounds are much less stringent than the requirement imposed here that $\omega > 1/2$, which is necessary to ensure that there is an event horizon. However, as these tests were conducted in the weak field around a material object whose radius is much larger than the gravitational radius the vacuum solution is not valid down near the event horizon and the constraint $\omega > 1/2$ need not apply.

	$\Delta(1/\omega)$
$a = 0.1$	15.2
$a = 0.3$	9.23
$a = 0.5$	5.47
$a = 0.7$	4.76
$a = 0.9$	3.44

Table 9.2: The bounds it is possible to place on the small KS deformation parameter $(1/\omega)$ for different values of the spin parameter, a. All other parameters were set to their fiducial values; $\iota = \pi/4$, $q = 3$ and $r_{\mathrm{out}} = 30$ and $Y \equiv 1/\omega = 0$.

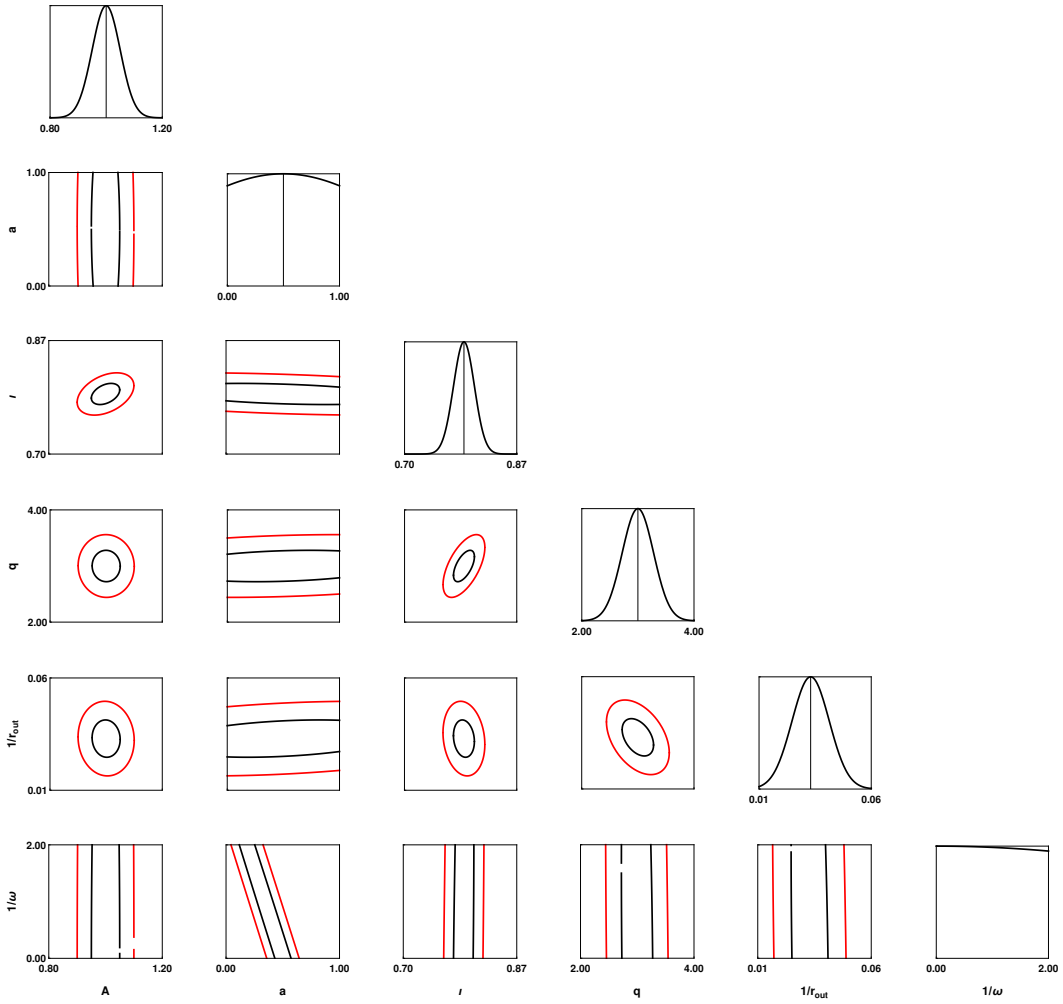


Figure 9.8: The 1σ and 2σ contours from a Fisher matrix analysis on the KS metric. It can be seen that introducing the extra degree of freedom in the KS deformation parameter has introduced a clear degeneracy with the spin parameter. From the bottom righthand plot it can be seen that the data is almost equally consistent with any value of ω in the range $(1/2, \infty)$, so no bound may be placed in the deformation. All parameters were set to their fiducial values; $a = 0.5$, $\iota = \pi/4$, $q = 3$, $r_{\rm out} = 30$ and $Y \equiv 1/\omega = 0$.

9.5.3 CS metric

Shown in Fig. 9.9 is the Iron line profile for a Kerr black hole and a CS2 black hole with deformation $\zeta = 1$. It should be remembered that the CS1 and CS2 metrics in Eqs. 9.10 and 9.1.1.3 are only valid in the limit $\zeta \ll 1$. However, even with the large deformation used in Fig. 9.9 the change in the line profile is only visible in the residual plot shown in the right-hand panel. The Iron line profile for the CS1 metric shows very similar behaviour to that in Fig. 9.9.

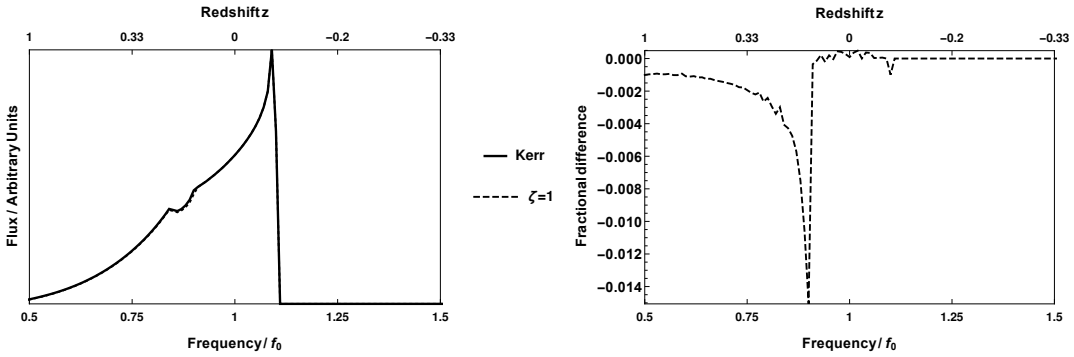


Figure 9.9: The left-hand panel shows a pair of Iron line spectra for accretion disks around a Kerr black hole and a quadratic in spin CS black hole (CS2 metric) with a large deformation, $\zeta = 1$. Also shown in the right-hand panel is the residual plot. All parameters were set to their fiducial values; $a = 0.5$, $\iota = \pi/4$, $q = 3$ and $r_{\rm out} = 30$.

Bearing in mind the results in Fig. 9.9, it is clear that it will be extremely difficult to bound the CS1 deformation parameter using this technique. Tab. 9.3 shows the bounds it is possible to place on both the CS1 and CS2 deformation parameters for a range of values of a; it was found that no bounds less than unity were possible with Iron line observations, however, the best results were obtained for the CS1 metric and high values of spin. Both the CS1 and CS2 metrics are expansions in the a parameter, therefore the bounds for the higher values of spin (particularly $a = 0.9$) should be treated with more caution than the low spins. However, this does not effect our main conclusion that no bounds less than unity were possible.

For comparison, weak field tests using the frame-dragging effect around the Earth measured by the Gravity Probe B and the LAGEOS satellites places a bound $\zeta^{1/4} <$

10^8 km (Ali-Haïmoud and Chen, 2011). Tighter bounds will be possible using strong field tests, for example it was found that bounds of $\zeta^{1/4} < 10^4$ km corresponding to a dimensionless bound of $\zeta < 1.0 \times 10^{-7}$ would be possible with eLISA observations of EMRIs (Canizares et al., 2012).

	$\Delta\zeta_{\rm CS1}$	$\Delta\zeta_{\rm CS2}$
$a = 0.1$	101	122
$a = 0.3$	97.0	62.6
$a = 0.5$	24.0	25.4
$a = 0.7$	23.9	9.98
$a = 0.9$	1.32	5.34

Table 9.3: The bounds it is possible to place on the CS1 and CS2 deformation parameters for different values of the spin parameter, a. It should be remembered that the CS parameter is constrained to be $\zeta \ll 1$, therefore no meaningful constraint may be placed with these observations. All other parameters were set to fiducial values, $\iota = \pi/4$, $q = 3$, $r_{\rm out} = 30$ and $\zeta = 0$.

9.5.4 $\mathcal{B}_N$ metrics

Shown in Fig. 9.10 are a series of Iron line profiles for the metrics defined in Sec. 9.1 with the constants $\mathcal{B}_2$ varied. The deformation parameters were varied between 0 and 0.5. Shown in Fig. 9.11 is the bounds it is possible to place on the different deformations using Iron line observations for different values of a.

The results in Fig. 9.11 show that the tightest constraints on the $\mathcal{B}_N$ bumpy black holes can be placed on the lowest values of N for the highest values of spin. The lower values of N have deformations entering at lower powers of $1/r$, since all the emission originates from outside the horizon (where $r > 1$) it is to be expected that deformations at lower N are easier to constrain. It is also to be expected that higher values of spin make placing constraints easier, because the ISCO moves to smaller values of r for larger spins, and most of the emission comes from close to the ISCO, the deformation has a greater effect on the spectra for high spins.

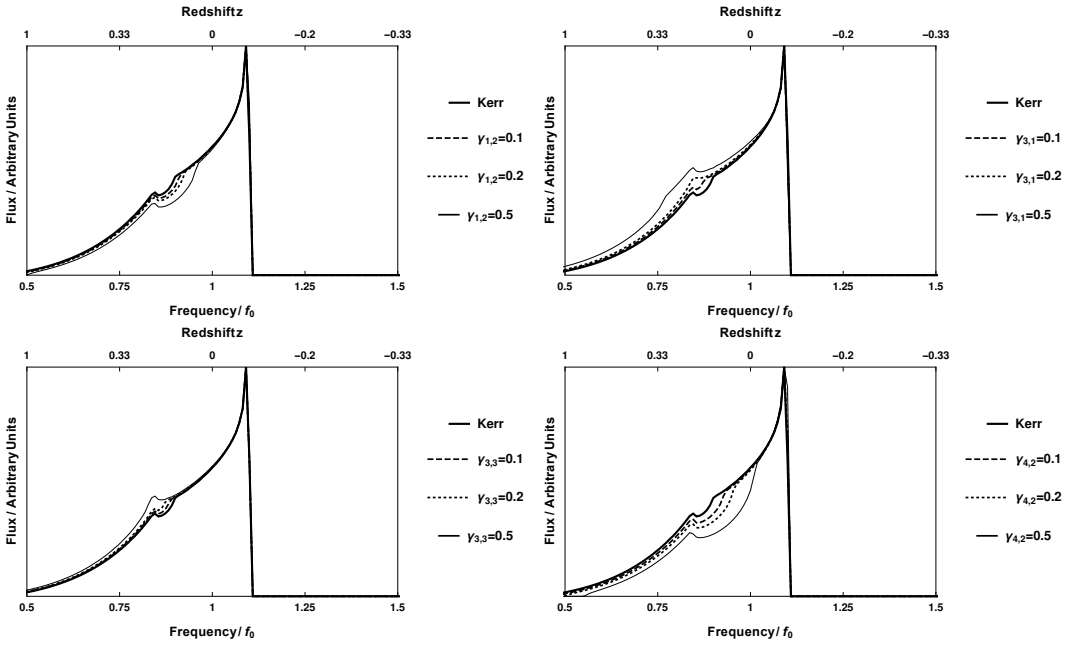


Figure 9.10: A series of Iron line spectra for accretion disks around B2 bumpy black holes with varying values of the bump parameters. All other parameters were set to fiducial values; $a = 0.5$, $\iota = \pi/4$, $q = 3$, $r_{\rm out} = 30$, and $\gamma_{i,j} = 0$ unless otherwise indicated.

9.6 Discussion

Observations of accretion disks offer an enticing opportunity to probe strong gravitational fields. Such observations have already provided some of the best evidence for the existence of black holes (in particular the existance of an event horizon, and a separate innermost stable circular orbit) and are routinely used to measure the spin of black holes. In this chapter we have considered the possibility of using X-ray emission from a thin accretion disk to distinguish between the Kerr black hole predicted by general relativity and alternative black hole metrics. For this purpose a series of parametrically deformed Kerr metrics, referred to as bumpy black holes, expanded in powers of M/r, were used (Gair and Yunes, 2011; Vigeland et al., 2011). In this chapter both the transition line and the thermal emission around these general bumpy black holes were calculated. The Fisher matrix formalism was used for the Iron line emission to estimate the accuracy to which the various disk parameters could be measured.

In addition to this general family of bumpy black holes the disk emission in several

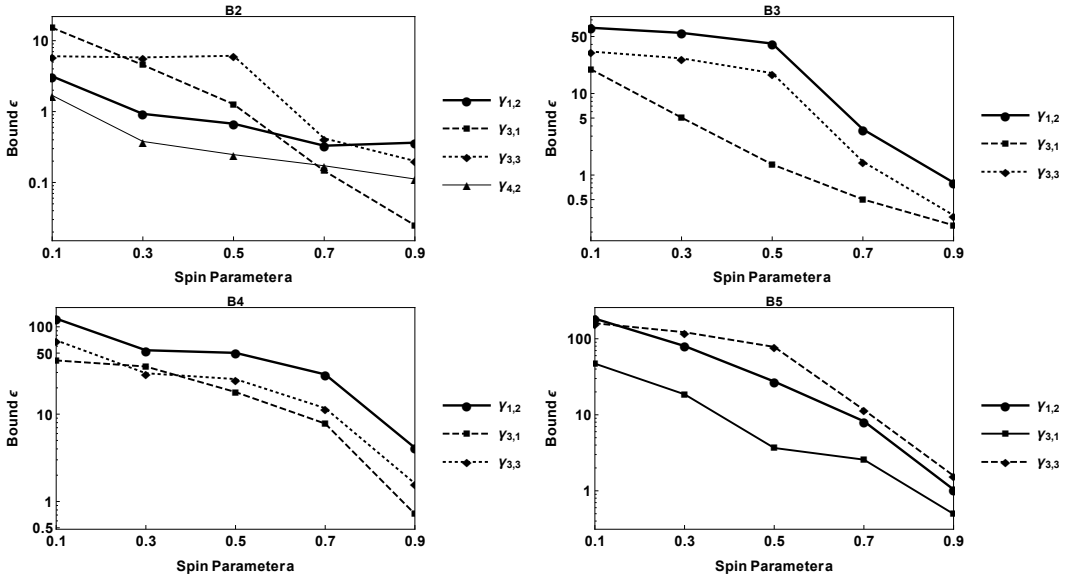


Figure 9.11: The bounds it is possible to place on the different $\mathcal{B}_N$ bump parameters for different values of the spin parameter, a. The tightest bounds can typically be placed for low values of N (i.e. $\mathcal{B}_2$) and for the most highly spinning black holes. All other parameters were set to fiducial values; $\iota = \pi/4$, $q = 3$, $r_{\rm out} = 30$ and $\epsilon = 0$.

specific, known black hole spacetimes were also considered. In particular the Iron line emission in the metric CS2 due to Yagi et al. (2012) was considered for the first time in the literature. For the CS2 metric it was found that it is impossible to place constraints on the small, dimensionless coupling parameter less than unity.

The technique of using accretion disk observations to constrain alternative black hole solutions has limitations. Qualitatively, the method works well for measuring the spin of a Kerr black hole, but adding extra deformation parameters typically introduces degeneracies which make constraining the bump parameters very difficult. This was very clearly seen in the case of the Kehagias Sfetsos metric, see Fig. 9.8. For the BN metrics it was found that bounds less than unity were only possible in a limited number of cases, typically for the leading order deformations around highly spinning black holes. It should also be remembered that the bounds presented here are optimistic, "best case" scenarios for the reasons discussed in Sec. 9.3.

In the near future it will become possible to perform this sort of test of general

relativity using gravitational wave observations of the inspiral and coalescence of a binary containing neutron stars or black holes. In particular, the possibility of using eLISA observations of extreme mass ratio inspirals to constrain the CS1 metric was recently considered by Canizares et al. (2012). It was found that bounds on the dimensionless deformation parameter $\epsilon < 10^{-6}$ may be possible, several orders of magnitude better than the bounds estimated here. As the CS1 metric is a special case of the general $\mathcal{B}_N$ metrics considered here, it would be interesting to try to constrain the $\mathcal{B}_N$ metrics here using the same techniques and see if similar improvements are possible.

Chapter 10

Extreme Mass-Ratio Inspirals

> We can lick gravity, but sometimes the paperwork is overwhelming.
>
> Wernher von Braun

The previous chapter considered the possibility of using X-ray observations of accretion disks to test the hypothesis that the gravitational field around an astrophysical black hole is described by the Kerr solution. It was found that while such tests were possible, their sensitivity was generally severely limited by a strong degeneracy between the black hole spin and any deformation to the Kerr metric. The cause of this degeneracy was the fact that the majority of the observed signal came from material that was moving on a single trajectory; the innermost stable circular orbit. In this chapter an alternative possibility for testing the Kerr metric, using gravitational wave observations, which does not suffer from this limitation will be considered.

The new era of gravitational wave astronomy opens up many new possibilities for testing general relativity. The announcement of the first direct detection of gravitational waves from a binary black hole merger (The LIGO Scientific Collaboration and the Virgo Collaboration, 2016*c*) was accompanied by the publication of a paper dedicated to checking whether the signal was consistent with the expectations of general relativity (The LIGO Scientific Collaboration and the Virgo Collaboration, 2016*f*); no statistically significant deviation from general relativity was found. The possibility of using observations of binary neutron star mergers to constrain the deviations from general relativity has also been considered (see, for example. Agathos et al. (2014)).

In the near future space-based detectors (such as the European eLISA mission, for

details see eLISA Consortium (2013)) will detect low-frequency gravitational waves. Such detectors will be sensitive to gravitational waves from a source with a greater total mass; it is expected that these will include a number of extreme mass-ratio inspiral (EMRI) sources. For a full discussion of the range of gravitational wave sources anticipated for space-based detectors see Sec. 3.2. An EMRI is a binary system comprising a small compact object (usually a neutron star or stellar mass black hole) which inspirals and eventually mergers with a much larger central black hole.

To detect the gravitational waves from an EMRI system with central black hole in the mass range $M = (10^5 - 10^7)\, M_\odot$ a space-based detector similar to eLISA sensitive to GWs in the frequency range $f = (10^{-4} - 10^{-1})\,\mathrm{Hz}$ is needed; the European space agency is planning such a mission for launch in 2034 (eLISA Consortium, 2013). Current estimates for the astrophysical rate of EMRI events are highly uncertain, but most estimates centre around the range $10^1 - 10^2$ detections per year with a LISA or eLISA-like detector. These sources can be observed out to distances of redshift $z \approx 1$ (see, for example, Amaro-Seoane et al. (2007, 2013)).

The instantaneous amplitude of the gravitational radiation from an EMRI can be more than an order of magnitude smaller than the noise in the detectors; however, the observed signal-to-noise ratio scales as $N^{1/2}$, where N is the number of wave cycles observed. A typical EMRI may be observed for $\sim 10^5$ orbits, and signal-to-noise ratios as high as $\sim 10^2$ can be achieved. Because of this large signal-to-noise and the fact that the waveform shows frequency modulation due to several different relativistic effects (see Sec. 10.1) it will be possible to extract precise measurements for the system parameters; typical fractional errors on the two object masses of $\sim 10^{-4}$ are possible, along with an absolute error of $\sim 10^{-4}$ for the central black hole dimensionless spin (see Barack and Cutler (2004)). This should also mean that it will be possible to put constraints on any "non-Kerr-ness" of the metric to a comparably strong level.

Observations of EMRIs are in many ways the ideal way to probe the gravitational field around a supermassive black hole; the compact object is generally on an inclined and eccentric orbit which gradually shrinks until merger, so the particle radiates information about much of the gravitational field, information far more rich than merely the location of the innermost stable circular orbit.

Given an EMRI observation one approach to performing a test of general relativity is to compare the observations with two sets of predictions; one calculated assuming

the Kerr metric of general relativity and a second assuming an alternative metric; for example, the slowly rotating solution to dynamical Chern-Simons gravity discussed in Sec. 9.1.1.2 (Canizares et al., 2012). The obvious drawback of such an approach is that it is limited to metrics which have already been considered theoretically. A preferable approach is to construct a large set of possible metrics that are continuously parametrised deformations of the Kerr metric and to use observations to place constraints on the sizes of the various different deformations. In this way a systematic and model independent test of the Kerr metric is possible. Although it should be noted that such a scheme can never hope to be a complete test of all possible deviations.

Several such schemes have been proposed, for example spacetimes that satisfy Einstein's equations but possess arbitrary and independent multipole moments have been considered. The metric in the vaccum region outside of a stationary, axisymmetric body may be written in terms of a series of multipole moments (Hansen, 1974); i.e. a power series in $1/r$ (where r is a radial coordinate) where each term is multiplied by an angular function. Such spacetimes contain naked singularities, closed time-like curves and other non-physical matter distributions near the origin which is why the multipole moments are usually not taken to be independent but their values fixed by just two parameters (the "no hair theorem") which may be taken as the black hole mass and spin, hence reducing to the Kerr metric. One approach to testing the Kerr hypothesis, is to measure the first $n > 2$ multipole moments and verify that they are consistent with the Kerr values. Ryan (1997) calculated the scalar radiation for a charge on a circular equatorial orbit in a stationary, axisymmetric spacetime with general multipole moments. Collins and Hughes (2004) considered orbits in spherical symmetric spacetimes with arbitrary multipole moments. Glampedakis and Babak (2006) computed "kludge" EMRI waveforms in stationary, axisymmetric spacetimes with an arbitrary quadrupole moment. Barack and Cutler (2007) also computed "kludge" EMRI waveforms in similar spacetimes, and then used the Fisher matrix to estimate the bound it might be possible to place on an anomalous contribution to the quadrupole moment.

In this thesis an alternative appoach to the multipole moments is considered. The same set of parametrically deformed Kerr metrics, otherwise known as "bumpy spacetimes", originally proposed by Vigeland et al. (2011) and previously used in Chapter 9, are considered. These bumpy spacetimes retain the symmetry properties of original Kerr metric; namely, they remain stationary, axisymmetric and admit a second rank Killing

tensor (at least to leading order in the deformation). The requirement that the bumpy spacetimes should retain this additional symmetry familiar to the Kerr metric is desireable, although not strictly necessary, because it ensures that geodesic motion in the bumpy spacetime has a fourth constant of motion (the other three being the particle rest mass, the energy and the z-component of the angular momentum). This in turn ensures that geodesic motion is integrable and, in the absence of radiation reaction, the motion of a test particle will be regular and tri-periodic (in the radial, polar and azimuthal directions). Integrable motion is strongly desireable (and perhaps even necessary) for matched filtering to be able detect the signal in the noisy detector data.

The plan of this chapter is as follows. In Sec. 10.1 the formalism used for calculating the EMRI gravitational waveform signal is described. In Sec. 10.2 we described how this formalism is to be modified to compute EMRI signals in the bumpy black holes previously described in Sec. 9.1. The necessary formalisms of signal analysis for extracting estimates of the errors on the system parameters are discussed in Sec. 10.3. The results of the analysis are presented in Sec. 10.4 and finally discussions and concluding remarks are given in Sec. 10.5.

10.1 Gravitational Waveforms

In this section the *analytic kludge* formalism developed by Barack and Cutler (2004) for computing an approximate gravitational wave signal (or waveform) from an EMRI around a standard Kerr black hole is briefly summarised. The idea is to approximate the compact object (CO) supermassive black hole (SMBH) binary at each instant as a non-relativistic Keplerian binary. The gravitational radiation from such a binary, at lowest order, is described by the famous result of Peters and Mathews (1963). The effects of general relativity are then approximately superimposed on the system by hand by artificially rotating the system to account for the perihelion precession, and the precession of the orbital plane (Lens-Thirring precession). The higher post Newtonian order effects of radiation reaction are also approximately included by forcing the frequency and eccentricity of the orbit to evolve with time. In all cases the equations governing the evolution come from post-Newtonian theory. The orbital motion of the detector also modulates the observed signal, both via the rotation of the detector with respect to the source direction and the Doppler effect of the motion in the direction of the source (the velocity of the Earth is sufficiently small that transverse Doppler shifts may be neglected).

The resultant waveforms include qualitativly all of the key features of the true EMRI signal; but in a simplified framework, and at a greatly reduced computational cost. In Sec. 10.2 the extension of this analytic kludge approach of to the bumpy spacetimes described in Sec. 9.1 is discussed.

The analytic kludge models have been used previously to assess parameter estimation accuracy, for example in the mock LISA data challenges (Babak et al., 2010). However, the kludge waveform models are not especially accurate (these waveforms drift out of phase with waveforms calculated using more accurate methods on a timescale of a few hours (Babak et al., 2007)) and results obtained from them should be treated only as estimates of results which would be obtained using more accurate methods.

Let M and μ be the masses of the SMBH and the CO respectively, where the SMBH is at a luminosity distance D from the solar system barycentre. The SMBH has a dimensionless spin parameter S and the unit 3-vector $\hat{S}$ points in the direction of this spin. The CO is on a Keplerian orbit around the SMBH with spatial position vector $\vec{r}(t)$, eccentricity e and orbital angular momentum $\vec{L}$ which makes a constant inclination angle λ with the direction of the SMBH spin. The orbital frequency is denoted by ν and $\Phi(t)$ is the mean anomaly from the pericentre. The azimuthal angle α on the angular momentum vector is defined by

$$\hat{L} = \hat{S}\cos\lambda + \frac{\hat{z} - \hat{S}\cos\theta_K}{\sin\theta_K}\sin\lambda\cos\alpha + \frac{\hat{S}\times\hat{z}}{\sin\theta_K}\sin\lambda\sin\alpha\,. \tag{10.1}$$

It is also necessary to specify the direction of the pericentre; this is done by defining the angle $\tilde{\gamma}$ between the direction $\vec{L}\times\vec{S}$ and the pericentre. See Fig. 10.1 for an illustration of these parameters.

For a true Keplerian orbit the quantities e, α, γ and ν are all constant, however here we allow them to vary with time. Qualitatively, the evolution of the angle γ is principally due to the precession of the perihelion in general relativity, the evolution of the angle α is due to the Lens-Thirring prscession of the orbital plane, and the evolution of ν and e is due to the radiation reaction effect of the emitted gravitational waves. The radiation reaction has the effect of shrinking and circularising the orbit over time. Quantative equations governing the evolution of the quantities $e(t)$, $\alpha(t)$, and $\nu(t)$ based on a post-Newtonian approach are described, along with the evolution of $\Phi(t)$ and $\gamma(t)$, in Sec. 10.1.1. For a true inspiral the inclination angle λ also evolves with time but typically only by a small amount (Hughes, 2000), here it is approximated as a constant.

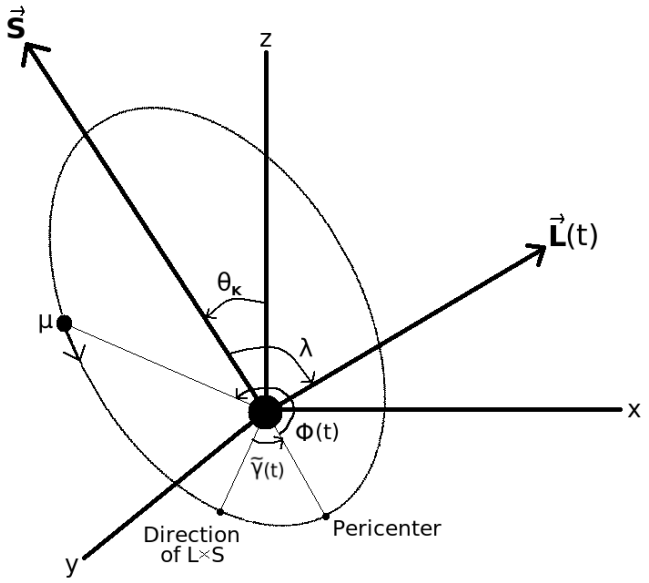


Figure 10.1: The CO-SMBH system. The vector $\vec{S}$ is the spin of the SMBH which is at an angle θ_K to the z axis of the eccliptic based coordinate system. The vector $\vec{L}(t)$ is the orbital angular momentum which makes a constant angle λ with $\vec{S}$ and precesses around it at a rate determined by $\alpha(t)$. The angle $\tilde{\gamma}(t)$ is the intrinsic direction of the pericentre from the direction of $\vec{L} \times \vec{S}$, and $\Phi(t)$ is the mean orbital anomaly from pericentre.

The position on the sky of the SMBH-CO system is specified by a pair of angular coordinates defined in the fixed ecliptic based coordinate system, $\{\theta_S, \phi_S\}$. The direction of the SMBH spin vector $\hat{S}$ is also specified by a pair of angular coordinates defined with respect to the same axis, $\{\theta_K, \phi_K\}$.

Following Barack and Cutler (2004) the spin of the CO has been neglected, leaving 14 parameters which completely specify the inspiral event. Defining some arbitrary fiducial frequency ν_0, the time of the event, t_0, is defined as the instant the orbital frequency equals ν_0. Hereafter a subscript 0 indicates the value a time dependent quantity takes at time t_0, for example $\alpha_0 = \alpha(t = t_0)$. It is useful to distinguish between parameters which are intrinsic to the SMBH-CO system (i.e. are independent of the position and orientation of the system with respect to Earth) and those which are extrinsic. The 14 parameters are chosen to be the following (logarithms of dimensionfull parameters have

been used), where the first 7 are intrinsic;

$$\{\log M, \log \mu, S, e_0, \tilde{\gamma}, \Phi_0, \lambda, \log D, \theta_S, \phi_S, \theta_K, \phi_K, \alpha_0, t_0\} \ . \tag{10.2}$$

At any instant the compact object is assumed to be on a Keplerian orbit around the central SMBH, the leading order emission from such a binary was described in the famous paper by Peters and Mathews (1963), the necessary results from which will now be briefly summarised. The emitted GWs are described by the metric perturbation, h_{ij}, which at leading order is given by the quadrupole formula(Misner et al., 1973)

$$h_{ij} = \frac{2}{D}(P_{ik}P_{jl} - \frac{1}{2}P_{ij}P_{kl})\ddot{I}^{kl}(t) \tag{10.3}$$

where the projection tensor is defined $P_{ij} = \eta_{ij} - n_i n_j$, η_{ij} is the flat space 3-metric, and $\hat{n}$ is a unit spacial vector pointing from the solar system barycentre to the source. The inertia tensor is defined as

$$I^{ij}(t) = \mu\, r^i(t)\, r^j(t)\,. \tag{10.4}$$

At large distance from the source the metric perturbation is a plane wave which may be written as the sum of two basis polarisations,

$$h_{ij}(t) = A^+(t)\, H^+_{ij}(t) + A^\times(t)\, H^\times_{ij}(t) \tag{10.5}$$

where A^+ and $A^\times$ are the amplitudes, and the basis tensors are defined as

$$H^+_{ij}(t) = \hat{p}_i\, \hat{p}_j - \hat{q}_i\, \hat{q}_j\,, \tag{10.6}$$

$$H^\times_{ij}(t) = \hat{p}_i\, \hat{q}_j + \hat{q}_i\, \hat{p}_j\,. \tag{10.7}$$

The unit vectors $\hat{p}$ and $\hat{q}$ have been defined as

$$\hat{p} = \frac{\hat{n} \times \hat{L}}{\left|\hat{n} \times \hat{L}\right|}\,, \tag{10.8}$$

$$\hat{q} = \hat{p} \times \hat{n}\,. \tag{10.9}$$

Because the orbit is planar, working in Cartesian coordinates adapted to the binary with $\hat{L}$ along the z-axis, the only non-zero components of the inertia tensor are $\{I^1, I^{12} = I^{21}, I^{22}\}$. Since the motion is periodic these may be expressed as the sum of harmonics of the orbital frequency $I^{ij} = \sum_n I^{ij}_n$. Defining,

$$a_n = \frac{1}{2}\left(\ddot{I}^{11}_n - \ddot{I}^{22}_n\right)\,, \tag{10.10}$$

$$b_n = \ddot{I}_n^{12} \,, \tag{10.11}$$

$$c_n = \frac{1}{2}\left(\ddot{I}_n^{11} + \ddot{I}_n^{22}\right) \,, \tag{10.12}$$

Peters and Mathews (1963) showed that

$$\begin{aligned} a_n = &- n\mathcal{A}\Big[J_{n-2}(ne) - 2eJ_{n-1}(ne) + \frac{2}{n}J_n(ne) \\ &+ 2eJ_{n+1}(ne) - J_{n+2}(ne)\Big]\cos\left(n\Phi(t)\right) \,, \end{aligned} \tag{10.13}$$

$$\begin{aligned} b_n = &- n\mathcal{A}\left(1-e^2\right)^{1/2}\Big[J_{n-2}(ne) - 2J_n(ne) \\ &+ J_{n+2}(ne)\Big]\sin\left(n\Phi(t)\right) \,, \end{aligned} \tag{10.14}$$

$$c_n = 2\mathcal{A}J_n(ne)\cos\left(n\Phi(t)\right) \,, \tag{10.15}$$

where $\mathcal{A} = \mu\left(2\pi M\nu\right)^{2/3}$. Bessel functions of the first kind are denoted by J_n and the mean anomaly simply increases linearly with time for a Newtonian binary,

$$\Phi(t) = 2\pi\nu(t - t_0) + \Phi_0 \,. \tag{10.16}$$

The amplitudes of the two polarisation states in Eq. 10.5 may also be decomposed into harmonics of the orbital frequency;

$$A^+ = \sum_n A_n^+ \,, \tag{10.17}$$

$$A^\times = \sum_n A_n^\times \,, \tag{10.18}$$

where,

$$A_n^+ = -\left[1 + \left(\hat{L}\cdot\hat{n}\right)^2\right]\left[a_n\cos\left(2\gamma\right) - b_n\sin\left(2\gamma\right)\right] + \left[1 - \left(\hat{L}\cdot\hat{n}\right)^2\right]c_n \tag{10.19}$$

$$A_n^\times = 2\left(\hat{L}\cdot\hat{n}\right)\left[b_n\cos\left(2\gamma\right) + a_n\sin\left(2\gamma\right)\right] \,. \tag{10.20}$$

The angle γ is related to the previously defined $\tilde{\gamma}$, the exact relation is described in Appendix G.

Now that we have a description of the gravitational readiation from the source it is necessary to describe the detectors' response to this radiation. A LISA-like detector has three arms which function as two crossed Michelson interferometers, see Tinto et al. (2002). The output from each of these effective interferometers is determined by the metric perturbation and the geometric response functions $F_I^+(t)$, $F_{II}^+(t)$, $F_I^\times(t)$, and

$F^{\times}_{II}(t)$ (where I and II denote the two output chanels of the detector). The response functions are similar to those given for a ground based interferometer in Sec. 2.4 and depend upon the relative orientations of the detector and source. Explicit expressions for these functions are given in Appendix H. The two detector outputs are denoted $h_\alpha(t)$, where $\alpha = I, II$, and are given by

$$h_\alpha(t) = \sum_n \left[F^+_\alpha(t) A^+_n(t) + F^\times_\alpha(t) A^\times_n(t) \right] \ . \tag{10.21}$$

The expression in Eq. 10.21 gives the measured gravitational wave signal for an instantenous Keplerian binary; it is now necessary to specify how the binary evolves in time.

10.1.1 Orbital Evolution

The orbital quantities $\Phi(t)$, $\nu(t)$, $\tilde{\gamma}(t)$, $e(t)$ and $\alpha(t)$ are evolved using the following post Newtonian formulae given in Eqs.10.22 to 10.26. These equations are valid to PN order 3.5, i.e. one order higher than when radiation reaction effects first appear. These are the same PN formulae as used in Barack and Cutler (2004), and the terms are mostly due to Junker and Schaefer (1992), with some additional terms from Brumberg (1991), Ryan (1996), and Barker and O'Connell (1975).

$$\frac{d\Phi}{dt} = 2\pi\nu \tag{10.22}$$

$$\begin{aligned}\frac{d\nu}{dt} =& \frac{96}{10\pi}\left(\frac{\mu}{M^3}\right)(2\pi M\nu)^{11/3}\left(1-e^2\right)^{-9/2}\left\{\left[1+\frac{73e^2}{24}+\frac{37e^4}{96}\right]\left(1-e^2\right)\right. \\ &+ (2\pi M\nu)^{2/3}\left[\frac{1273}{336}-\frac{2561e^2}{224}-\frac{3885e^4}{128}-\frac{13147e^6}{5376}\right] \\ &\left.- (2\pi M\nu)\left(\frac{S}{M^2}\right)\cos\lambda\left(1-e^2\right)^{-1/2}\left[\frac{73}{12}+\frac{1211e^2}{24}+\frac{3143e^4}{96}+\frac{65e^6}{64}\right]\right\}\end{aligned} \tag{10.23}$$

$$\begin{aligned}\frac{d\gamma}{dt} =& 6\pi\nu\left(2\pi M\nu\right)^{2/3}\left(1-e^2\right)^{-1}\left[1+\frac{1}{4}\left(2\pi M\nu\right)^{2/3}\left(1-e^2\right)^{-1}\left(26-15e^2\right)\right] \\ &- 12\pi\nu\cos\lambda\left(\frac{S}{M^2}\right)\left(2\pi M\nu\right)\left(1-e^2\right)^{-3/2}\end{aligned} \tag{10.24}$$

$$\begin{aligned}\frac{de}{dt} =& -\frac{e}{15}\left(\frac{\mu}{M^2}\right)\left(1-e^2\right)^{-7/2}\left(2\pi M\nu\right)^{8/3}\left[\left(304+121e^2\right)\left(1-e^2\right)\right. \\ &\left.\left(1+12\left(2\pi M\nu\right)^{2/3}\right)-\frac{1}{56}\left(2\pi M\nu\right)^{2/3}\left(8\times 16705+12\times 9082e^2-25211e^4\right)\right]\end{aligned}$$

$$+ e\left(\frac{\mu}{M^2}\right)\left(\frac{S}{M^2}\right)\cos\lambda\,(2\pi M\nu)^{11/3}\left(1-e^2\right)^{-4}\left[\frac{1364}{5}+\frac{5032e^2}{15}+\frac{263e^4}{10}\right] \tag{10.25}$$

$$\frac{d\alpha}{dt} = 4\pi\nu\left(\frac{S}{M^2}\right)(2\pi M\nu)\left(1-e^2\right)^{-3/2} \tag{10.26}$$

10.1.2 Doppler Modulation

The final ingredient in the analytic kludge model waveform is the Doppler effect on the observed signal due to the orbital motion of the detector around the sun. This becomes significant for observations lasting more than a few weeks, which is typically the case for EMRIs. The main effect on the parameter estimation of the Doppler shifting due to the space-craft motion is improved estimation of the source orientation and sky position; if the signal frequency is higher in a given month, then the direction to the source may be partially infered to be in direction of motion of the Earth in that month (Cornish and Rubbo, 2003). The effect of the motion of the centre of mass of the detector may be accounted for by making the transformation $\Phi(t) \to \Phi(t) + \Phi_D(t)$ in all the above expression;

$$\Phi_D(t) = 2\pi\nu(t) R \sin\Theta_S \cos\left(\frac{2\pi t}{T} - \phi_S\right). \tag{10.27}$$

The orbital radius of the LISA spacecraft is $R = 1\,\mathrm{AU}$, the orbital period is $T = 1\,\mathrm{year}$, and the angles $\{\Theta_S, \phi_S\}$ define the sky position of the source.

There is an additional, smaller, Doppler effect on the signal due to the rotational motion of the triangular constellation of spacecraft about the centre of mass of the detector which is neglected here; for a discussion see Cutler (1998).

10.2 Bumpy black hole spacetimes

In this section the necessary formalism for generating EMRI waveforms in the family of perturbed black hole spacetimes considered in Sec. 9.1 is presented. This family of "bumpy black holes" are continuously parametrised deformations of the Kerr metric and were first proposed by Vigeland et al. (2011). These bumpy black hole spacetimes are constructed such that they retain the symmetries of the background Kerr metric; namely the properties of stationarity, axisymmetry, and the existence of a second rank Killing tensor (at least to leading order in the deformation). These symmetries ensure that geodesic motion in the perturbed spacetime possesses four constants of motion: rest

mass, energy, z-component of angular momentum, and a Carter-like constant (again, at least to leading order in the perturbation).

The deformation to the metric tensor will translate, through the geodesic equations, to a deformation in the instantaneous orbital trajectory which will further translate, through the emission of gravitational waves (which is here represented with the quadrupole approximation), into a deformation in the gravitational waveforms described in Sec. 10.1. In this section the changes proposed by Vigeland et al. (2011) to the gravitational waveforms of the previous section are discussed.

It should be noted that in this approach only the metric tensor $g_{\mu\nu}$ has been modified from the standard general relativity prediction. It would also be possible to consider alternative theories of gravity where any combination of the metric tensor, the sourcing of gravitational waves by matter, the polarisation and propagation of the gravitational waves, or the radiation reaction effect is modified from the general relativity prediction. For example, the LISA response functions to alternative polarisation states were considered in Tinto and Alves (2010). Futhermore, the possibility of constraining differences in gravitational wave propagation (specifically placing bound on the mass of the graviton) was considered by Berti et al. (2011), using the dispersion effects in the chirping signals of SMBH mergers. For a comprehensive review of possible tests of general relativity using space-based gravitational wave detectors see Gair et al. (2013).

In Gair and Yunes (2011) it was shown how the metric perturbations in Eq. 9.6 and Appendix E.1 lead to corrections in Eqs. 10.22 through 10.26. The corrections at order $\mathcal{B}_2$ are given in Eqs. 10.28 to 10.31 and corrections up to order $\mathcal{B}_5$ are given in Appendix I.

$$2\pi M^2\delta\left(\frac{d\nu}{dt}\right) = \frac{16}{5}\eta\frac{(2\pi M\nu)^{13/3}}{(1-e^2)^{9/2}}\left(18+78e^2+\frac{99}{4}e^4\right)(\gamma_{1,2}+2\gamma_{4,2}) \quad (10.28)$$

$$M\delta\left(\frac{d\gamma}{dt}\right) = \frac{(2\pi M\nu)^{5,3}}{2\,(1-e^2)}(\gamma_{1,2}+2\gamma_{4,2}) \quad (10.29)$$

$$\delta\left(\frac{de}{dt}\right) = -\frac{16}{5}\eta\frac{(2\pi M\nu)^{10/3}}{(1-e^2)^{7/2}}\left(\frac{93}{4}e+\frac{67}{4}e^3+\frac{1}{4}e^5\right)(\gamma_{1,2}+2\gamma_{4,2}) \quad (10.30)$$

$$M\delta\left(\frac{d\alpha}{dt}\right) = -\frac{a\,(2\pi M\nu)^2}{(1-e^2)^{3/2}}(\gamma_{1,2}+2\gamma_{4,2}) \quad (10.31)$$

It should be noted that all these corrections depend upon similar combinations of parameters, at the $\mathcal{B}_2$ order this is $(\gamma_{1,2}+2\gamma_{4,2})$. This is a dimensionless quantity of order unity which gets multiplied by the small quantity ϵ in the expression for the metric.

Here $(\gamma_{1,2} + 2\gamma_{4,2})$ is set to unity, then ϵ is treated as an additional free parameter to add to those in Eq. 10.2. Including these corrections in the prescription described in Sec. 10.1 allows gravitational waveforms to be generated for the perturbed spacetimes.

As discussed in Sec. 9.1, one special case of the general family of bumpy black holes is of particular interest; the linear in spin Chern Simons metric (Yunes and Pretorius, 2009). This metric has leading order terms of $\mathcal{O}((M/r)^4)$, but as it also contains additional higher order terms as well, and so is not purely a $\mathcal{B}_4$ metric.

At quadratic order in spin, $\mathcal{O}(S^2)$, Yagi et al. (2012) found an analytic black hole solution to the equations of dynamical Chern-Simons gravity. This metric is not in the class of spacetimes discussed in Sec. 10.2 and geodesic motion in this spacetime does not posses a Carter-like constant. The quadratic in spin solution has bumps which enter at $\mathcal{O}((M/r)^3)$, one order lower than was the case for the linear in spin metric. Therefore it would be expected that bounds placed on the Chern-Simons coupling parameter using the linear in spin solution will only be valid for very small values of S, because as S increases the bumps will become dominated by the quadratic in spin solution. At higher order in spin no analytic black hole solutions are known; however, there has been some numerical progress studying highly spinning Chern-Simons black holes (Stein, 2014).

10.2.1 A Sample Waveform

In Secs. 10.1 and 10.2 the necessary theory was introduced to allow approximate *analytic kuldge* EMRI waveforms to be calculated in the Kerr black hole solution of Einstein's equations, as well as the parametrised bumpy black hole solutions of Vigeland et al. (2011). Shown in Fig. 10.2 are short segments of sample waveforms in both the Kerr metric and a $\mathcal{B}_2$ metric with bump parameter $\epsilon = 1.0$. These signals were generated for the same values of the other system parameters and are arranged such that they are in phase at the end of the plotted time window. Over a short time scale (around $\sim 10^3\,\mathrm{s}$) the $\mathcal{B}_2$ waveform drifts out of phase relative to the Kerr waveform. EMRI waveforms can be observed over a period of several years, and phase differences of $< 1\,\mathrm{rad}$ over the entire signal can generally be detected (assuming the cause of the phase difference is not degenerate with the other system parameters). The two waveforms shown in Fig. 10.2 could be easily distinguished by a LISA-like detector.

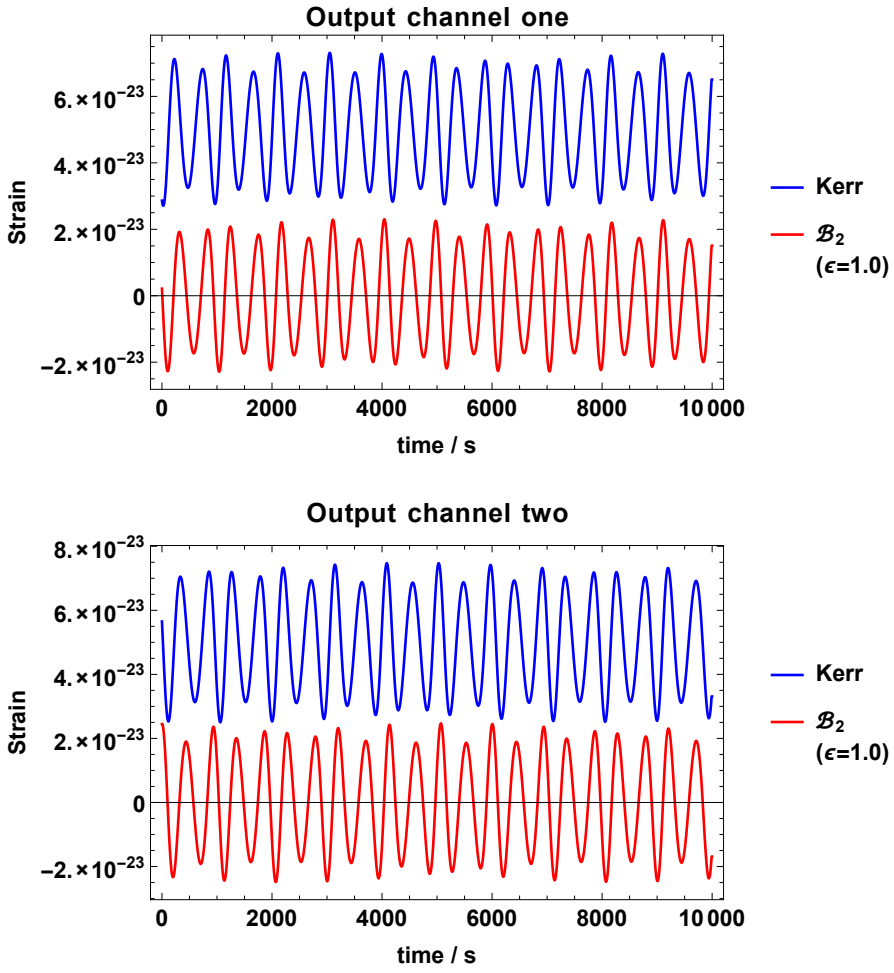


Figure 10.2: A short segment (~ 2 hours) of a sample waveform for the Kerr spacetime and one for a $\mathcal{B}_2$ with $\epsilon = 1.0$. The two detector outputs, $h_I(t)$ and $h_{II}(t)$, are shown in the top and bottom panels respectively. The values of the parameters were $M = 10^5 M_\odot$, $\mu = 10 M_\odot$, $S = 0.1$, $e_0 = 0.1$, and $\lambda = \pi/4$; the source is at a luminosity distance of $D_L = 1\,\mathrm{Gpc}$, and the remaining angles were randomly drawn from the relevent ranges. The Kerr waveform, shown in blue, has been vertically shifted for clarity. The waveforms are plotted for a short segment of $10^4\,\mathrm{s}$ around a central frequency of $\nu_0 = 10^{-3}\,\mathrm{Hz}$. The $\mathcal{B}_2$ waveform, shown in red, is in phase with the Kerr waveform at the end of the segment shown, but dephases from the Kerr waveform over a timescale of a few times $10^3\,\mathrm{s}$.

10.3 Signal Analysis

A LISA-like detector has three arms which have the equivalent response to two crossed Michelson detectors here labelled by the subscript index $\alpha = I, II$. The output of each detectors is a time series denoted $s_\alpha(t)$, with Fourier transform $\tilde{s}(f)$ defined in Eq. 2.3. The measured signal is the sum of the true gravitational waveform, $h_\alpha(\vec{\theta}_0, t)$, and the noise in the experiment, $n_\alpha(t)$,

$$s_\alpha(t) = h_\alpha(\vec{\theta}_0, t) + n_\alpha(t)\,. \tag{10.32}$$

The noise in each detector is assumed to have identical statistical properties; i.e. is stationary, Gaussian and uncorrelated between the two detector chanels. Under these assumptions the noise is fully specified by the power spectral density (see Sec. 2.1.2) $S_n(f)$,

$$\langle \tilde{n}_\alpha(f)\tilde{n}_\beta(f')\rangle = \frac{1}{2}\delta(f-f')S_h(f)\delta_{\alpha\beta}\,. \tag{10.33}$$

The functional form of the power spectral density (which is assumed to be known *a priori*) for a LISA-like detector is discussed in Appendix J. Under these assumptions the likelihood, $L(\vec{\theta}) \equiv P(s|\vec{\theta})$, of the measured data s containing the gravitational wave signal $h(\vec{\theta})$ (where $\vec{\theta}$ is a vector containing the source parameters) is given by

$$L(\vec{\theta}) \propto \exp\left(\frac{-1}{2}\langle s - h(\vec{\theta})|s - h(\vec{\theta})\rangle\right)\,, \tag{10.34}$$

where angled brackets denote the inner product defined in Eq. 2.15. For simplicity, it is assumed in this chapter that the priors on the source parameters are flat over the region with significant posterior support.

This expression for the likelihood should be compared with that in Eq. 9.43 used in Chapter 9. As in the previous chapter, the Fisher matrix is used to estimate the accuracy with which the source parameters may be measured. Expanding the difference in the signals to linear order in the difference in the source parameters gives

$$h_\alpha(\vec{\theta}) = h_\alpha(\vec{\theta}_0) + \delta\theta_i \frac{\partial h_\alpha(\vec{\theta})}{\partial\theta_i}\Big|_{\vec{\theta}=\vec{\theta}_0} + \mathcal{O}\left((\delta\theta_i)^2\right)\,, \tag{10.35}$$

where $\delta\vec{\theta} = \vec{\theta} - \vec{\theta}_0$. Substituting this into Eq. 10.34 gives a Gaussian approximation to the likelihood valid in the limit of high SNR;

$$L(\vec{\theta}) \propto \exp\left(\frac{-1}{2}\Sigma_{ij}\delta\theta^i\delta\theta^j\right) + \mathcal{O}(1/\rho)\,, \tag{10.36}$$

where the *Fisher matrix* has been defined as

$$\Sigma_{ij} = \left\langle \frac{\partial h(\vec{\theta})}{\partial \theta_i}\Big|_{\vec{\theta}=\vec{\theta}_0} \middle| \frac{\partial h(\vec{\theta})}{\partial \theta_j}\Big|_{\vec{\theta}=\vec{\theta}_0} \right\rangle , \tag{10.37}$$

and the signal-to-noise ratio is $\rho = \langle h|h\rangle^{1/2}$. Therefore, for large SNR, the maximum likelihood estimator of the soure parameters is Gaussian distributed with mean at the true parameters and with covariance matrix given by the inverse of the Fisher matrix. This inverse matrix can be used to estimate the parameter estimation errors (c.f. Eq. 9.47),

$$\Delta\theta_i = \sqrt{(\Sigma^{-1})_{ii}} \quad \text{(no sum on } i\text{)}. \tag{10.38}$$

The Fisher matrix in Eq. 10.37 is given in terms of an inner product which is evaluated in the frequency domain. For computational speed however, and following the approach of Barack and Cutler (2004); Barack and Cutler (2007), it is convenient to adopt an approximate version of the inner product defined purely in the time domain. This avoids the need for performing multiple Fourier transforms of the (typically very long) signals. Firstly, define the noise weighted waveform as

$$\hat{h}_\alpha(t) = \sum_n \frac{h_{\alpha,n}(t)}{S_h^{1/2}\left(f_n(t)\right)} , \tag{10.39}$$

where f_n is a combination of the radial orbital frequency, $\nu(t)$, and the azimuthal orbital frequency,

$$f_n(t) = n\nu(t) + \frac{\dot{\tilde{\gamma}}}{\pi} . \tag{10.40}$$

The inner product can then be approximated as an integral of the noise weighted waveforms in the time domain.

$$\langle a(t), b(t)\rangle \approx 2\sum_\alpha \int_0^T dt\; \hat{a}_\alpha(t)\hat{b}_\alpha(t) . \tag{10.41}$$

This time domain inner product is motivated by considering Parseval's theorem in the special case of a flat noise PSD, $S_h(f)$. This approximation is exact for circular orbits.

Mirroring the calculations performed in Chapter 9 for the Iron line observations, the validity of the Fisher matrix was checked using the mismatch ratio proposed by Vallisneri (2008). As described in Sec.9.4, this checks the internal consistency of the linear signal approximation which is central to accuracy of the Fisher matrix approximation. N points

were selected at random from the 1σ surface returned from the Fisher matrix analysis. At each of these points the *mismatch ratio*, $|\log r|$, was calculated,

$$|\log r(\theta)| = \frac{1}{2}\left\langle \Delta\theta_i \frac{\partial h}{\partial \theta_i}\Big|_{\vec{\theta}=\vec{\theta}_0} - \Delta h(\theta) \middle| \Delta\theta_j \frac{\partial h}{\partial \theta_j}\Big|_{\vec{\theta}=\vec{\theta}_0} - \Delta h(\theta) \right\rangle \tag{10.42}$$

where $\Delta h(\theta) = h(\vec{\theta}_0 + \Delta\vec{\theta}) - h(\vec{\theta}_0)$ is the difference in the waveforms between the true point and the point selected on the 1σ surface. Summation convention over repeated indices is used.

This procedure was followed for the waveforms described in Sec. 10.1 for the EMRI system shown in Fig. 10.2, observed for 1 year ending at plunge at a distance such that the observed signal-to-noise ratio was 30. The mismatch ratio was calculated at 100 points on the 1σ surface, producing an average value of $|\log r| = 0.06$ and a maximum value of $|\log r| = 0.07$, implying that the linear signal approximation holds to better than 10% out to the 1σ surface. This suggests that the Fisher matrix provides a reasonable estimate of the accuracy with which the system parameters can be measured.

10.4 Results

The Fisher matrix was used to assess the bounds that it would be possible to place on the $\mathcal{B}_N$ metric deformation parameters. This was done for the EMRI system with parameters the same as those used to generate the data plotted in Fig. 10.2. The source was assumed to be observed for 1 year with the observation period ending when the CO plunges into the SMBH. The EMRI system was allowed to have a single $\mathcal{B}_N$ deformation parameter at a time to deviating from the zero Kerr value. Deformation parameters for $\mathcal{B}_2$, $\mathcal{B}_3$, $\mathcal{B}_4$, and $\mathcal{B}_5$ were considered. The source was placed at a luminosity distance such that the observed signal-to-noise was 30 (the required distance was 850 Mpc). The deformation parameter was set to zero, and the accuracy with which the deformation parameter can be measured was assessed using the Fisher matrix and the result was interpreted as the bound that is would be possible to place on that particular deformation. This mimics the situation where the true gravitational field around the black hole is in fact Kerr, and the challenge is to place the most stringent bounds on any deformations. The results of this analysis for $N = 2, 3, 4, 5$ are shown in Fig. 10.3.

As was the case with the Iron line observations discussed in Chapter 9 it can be seen that the most stringent bounds can be placed on the deformations with the lowest values

of N; this is because all of the information is coming from outside of the horizon in the region where $r > 1$, and the size of the metric deformation is proportional to $1/r^N$. It can also be seen, by comparing the results from Fig. 9.11 to those in Fig. 10.3, that EMRI observations offer the opportunity to place much more stringent bound (more than 4 orders of magnitude) than x-ray observations of accretion disks. These huge improvements were anticipated, and the reasons for them have already been discussed in great detail at the beginning of the current chapter.

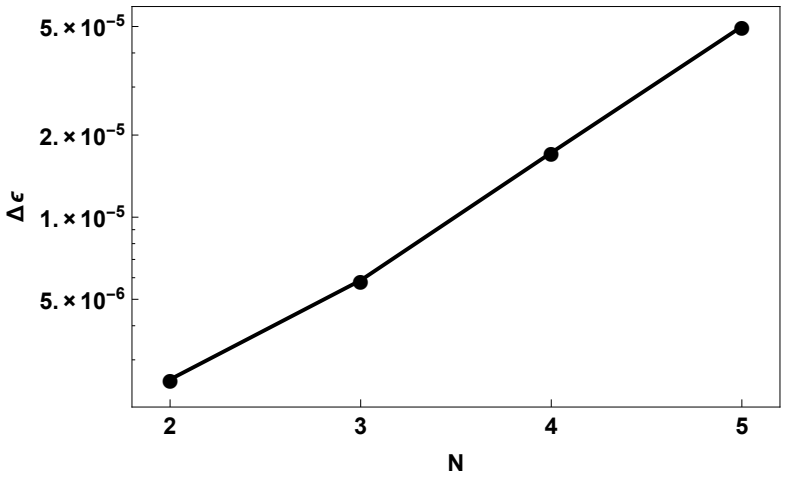


Figure 10.3: A plot of the bound it is possible to place on the $\mathcal{B}_N$ deformation parameter as a function of the order, N. As expected, and as was the case for the X-ray observations considered in Chapter 9, the bounds are strongest for the lowest order deformations. These bounds were computed for the EMRI system plotted in Fig. 10.2, observed for 1 year (ending at plunge) and at a distance such that the system has an observed signal-to-noise ratio of 30.

10.4.1 Chern-Simons metric

The linear in spin black hole solution to dynamical Chern-Simons gravity (see Sec. 9.1.1.2) is a special case of the general family of bumpy black holes considered so far. It is therefore of interest to see what bounds it might be possible to place on the Chern-Simons coupling parameter using EMRI observations and compare the results to those

obtained for Iron-line observations in Chapter 9 and previous published studies using EMRIs (Canizares et al., 2012).

For the EMRI system shown in Fig. 10.2 (but with the spin parameter set to $S = 0.1$, so that the linear in spin approximation remains reasonably valid) the bound that it is possible to place on the Chern-Simons deformation parameter was assessed using the Fisher matrix to be $\xi^{1/4} < 5 \times 10^7\,\mathrm{m}$. This is similar to bounds previously obtained by Canizares et al. (2012).

However, it is now known that the quadratic in spin solution (see Sec. 9.1.1.3) contains metric deformation terms at order $1/r^3$, whereas the linear in spin solution (see Sec. 9.1.1.2) only contains terms at orders $1/r^4$ and higher. As has been observed previously deformations that enter at lower order in $1/r$ can typically have tighter constraints placed upon them. Therefore, observations of EMRIs around central black holes with intermediate or high spins may actually allow more stringent limits than the calculations reported here suggest due to the presence of metric deformation terms $\propto S^2/r^3$.

The quadratic in spin Chern-Simons metric is not part of the family of perturbed spacetimes described in Sec. 9.1 (geodesic motion in this spacetime does not posses a Carter-like constant of motion) and therefore EMRI waveforms cannot be generated for these systems using the formalism presented here. Therefore all that can currently be said about bounding the Chern-Simons coupling parameter using EMRI observations is that all the available estimates are rough, preliminary estimates. A more detailed assessment would require model EMRI waveforms more general than those presented here, and this must necessarily be left to future work.

10.5 Discussion and concluding remarks

This chapter along with the one that preceded it have considered some possibilities for performing strong-field tests of general relativity.

Observations of accretion disks offer an enticing opportunity to probe strong gravitational fields. Such observations have already provided some of the best evidence for the existence of black holes (in particular the existence of an event horizon, and a separate innermost stable circular orbit) and are routinely used to measure the spin of black holes. In the previous chapter we have considered the possibility of using X-ray emission from a thin accretion disk to distinguish between the Kerr black hole predicted by general

relativity and alternative black hole metrics. For this purpose a series of parametrically deformed Kerr metrics, referred to as bumpy black holes, expanded in powers of M/r, were used (Gair and Yunes, 2011; Vigeland et al., 2011). Both the transition line and the thermal emission from accretion disks around these general bumpy black holes were calculated. The Fisher matrix formalism was used for the Iron line emission to estimate the accuracy to which the various disk parameters could be measured.

The parametrised family of bumpy black holes is able to recover some known black hole solutions in particular alternative theories of gravity; for example, the linear in spin solution to dynamical Chern-Simons gravity (Yunes and Pretorius, 2009). The possibility of using EMRIs to probe this particular metric has been considered previously, in Sec. 10.4.1 it was shown that the results of this special case can be recovered from the general parameterised framework.

The use of EMRI observations to probe the metric around an astrophysical black hole is a technique that shows a lot of promise. The phase of the gravitational waves tracks the orbital phase of the compact object as it orbits and precesses about the central black holes, and the phase of the gravitational wave can typically be tracked to better than 1 radian over hundreds of thousands of orbits. And because the compact object is generally on an inclined and eccentric orbit the EMRI waveform contains information about the entire spacetime, not limited to just the equatorial plane, or the location of the innermost stable circular orbit.

Part IV

Conclusions

Chapter 11

Retrospect

I have been very fortunate to have had the opportunity to conduct my PhD research during an extremely exciting time for the field of gravitational wave astronomy. Midway through my PhD excitement was stoked by the surprise announcement by the BICEP2 collaboration of the possible detection of ultra-low-frequency gravitational waves via the B-mode polarisation in the cosmic microwave background, although subsequent research has cast some doubt on the cosmological origin of this polarisation. Furthermore, around 18 months on from the BICEP2 announcement, a huge step was taken towards the detection of mHz gravitational waves in space with the successful launch of the LISA Pathfinder mission. In addition, exciting progress has been made towards the first detection of nHz gravitational waves; during the course of my PhD each of the three main pulsar timing arrays have published improved limits on the gravitational wave stochastic background. Finally, only a few weeks before this thesis was submitted, the first direct detection of gravitational waves was announced! The advanced LIGO detectors detected the telltale signal of the merger of two black holes, each around 30 times the mass of our own sun. I feel very privileged to have had the opportunity of being involved in the field during such a historic period.

Against this backdrop it was impossible not to become excited by the science which would soon become possible with gravitational wave observations. In part I of this thesis I reviewed a wide range of gravitational wave detectors and gravitational wave sources. This work culminated with the figures shown in Chapter 4, which attempt to condense as much of this information as possible into a single plot. The over-riding impression that I was left with after producing these plots is of the breadth of the field. The figures

manage to capture fairly well the huge range of scales probed by gravitational wave astronomy (frequencies, and hence masses, range over at least 14 orders of magnitude, even when not including cosmological and early universe sources). However, the figures almost entirely fail to capture the range of different signal morphologies which occur, the diversity of the data analysis techniques needed to extract these diverse signals, and the incredible range of technologies used in the construction of all of the different types of detector.

In part II of this thesis I considered a particular problem which can occur when using gravitational wave observations to measure the properties of the source. This analysis must necessarily be performed using imperfect models; imperfections can arise for a wide variety of reasons discussed in Chapter 5 and these lead to systematic errors in the recovered source parameters. The problems associated with imperfect models affect gravitational wave science in particular as it generally relies quite heavily on detailed signal models (or templates) for both detection and subsequent characterisation of sources. The technique that I have developed in this thesis for overcoming this problem involves using a pair of models for the same system but which have differing computational costs and accuracies (and in general including different physical effects). The technique of Gaussian process regression was then used to marginalise over the model error, treating it as any nuisance parameter would normally be treated in a Bayesian analysis. In this way I obtained a new likelihood, called the *marginalised likelihood*, which may be evaluated in reasonable computational time, but which has reduced systematic errors and properly accounts for any remaining model uncertainty. In Chapter 6 I demonstrated that the marginalised likelihood has a number of very appealing properties; in particular, parameter estimates obtained from this likelihood suffer from a *greatly* reduced bias compared to standard techniques.

In part III of this thesis I considered some possibilities for testing our current best theory of gravity, the general theory of relativity. This theory is now over 100 years old (another important milestone passed during the course of my PhD). In that time it has been tested in numerous, and ever more precise and stringent ways, and no defect which stands up to scrutiny has been found. A very brief summary of some of the most historically important tests was given in Chapter 9. These tests have, until recently, been limited to observations carried out in weak and slowly varying gravitational fields. Gravitational waves, in contrast, are typically generated in strong

and highly dynamic gravitational fields; these waves then propagate to us with almost no distortion or absorption by any intervening material. The specific tests I considered in this thesis concern the gravitational field around an astrophysical black hole, which general relativity (together with cosmic censorship and the "no-hair" theorems) predicts should be described by the famous Kerr solution. In Chapter 9 I considered a way of testing this hypothesis which utilised X-ray observations of an accretion disk around a black hole. I found that although such tests were possible, their sensitivity was generally limited by a strong degeneracy between the black hole spin and any deviation from the Kerr metric. In Chapter 10 I considered an alternative which utilised observations of gravitational waves from extreme mass-ratio inspirals. These systems are ideally suited for the purpose because the gravitational waves directly track the phase of a small object whose trajectory around the black hole is governed entirely by the gravitational field. I found that more stringent tests were possible via this approach. Testing general relativity in the strong field regime is one of many areas which stand to be revolutionised in the era of gravitational wave astronomy.

Chapter 12

Prospect

The future is undoubtedly bright for gravitational wave astronomy. Indeed, the aspect of the recent announcement by the LIGO and VIRGO collaborations which I personally found most exciting was not the detection of gravitational waves (it seemed inevitable that this would be achieved at some point soon), rather it was the fact that just two detectors, operating for only a couple of weeks at a sensitivity substantially below their full design sensitivity, were able to make such a clear and convincing detection. Over the next few years several more detectors, including Advanced Virgo, LIGO India and KAGRA, will join in the hunt for gravitational waves. The sensitivity of all these detectors can be expected to continually improve over time. As this ever-improving network of ground-based detectors is jointly operated for months, and then years, at a time, we can expect numerous more detections; it is likely that the number of confidently detected black hole binaries will exceed 100 by the end of the decade (The LIGO Scientific Collaboration and the Virgo Collaboration, 2016*g*), and this number does not include all of the other sources of gravitational waves that ground-based detectors may observe. It is the implications of the recent announcement for the future prospects of ground-based laser interferometers that excite me most of all.

Of course, the timely and accurate analysis of this wealth of observational data poses as many challenges as it does opportunities. The precise analysis of the loudest detected events will always test the fidelity of our best models. In addition, as the number of detected sources grows, the accuracy of our models will also limit our ability to infer unbiased parameters for the underlying population of sources. The Gaussian process marginalised likelihood developed in this thesis is one technique designed to quantify

and overcome these problems. A posterior probability distribution obtained using the marginalised likelihood should be more robust than one obtained using the standard likelihood. The routine use of the Gaussian process marginalised likelihood will help to ensure that any inferences drawn from gravitational wave observations are robust against the inaccuracies present in our models. The basic theoretical aspects of the marginalised likelihood have been thoroughly developed in this thesis and the papers referenced within. The technique has been demonstrated using a realistic, but significantly simplified toy example. The next stage is of course to implement the marginalised likelihood in the LIGO algorithm library and apply it to real gravitational wave data analysis.

Additionally, there are some outstanding theoretical aspects of the marginalised likelihood which can be developed further. Some of these have been discussed in detail in Chapter 8. These developments centre around relaxing some of the stringent conditions which have been placed on the the covariance function which governs the correlation structure of the Gaussian process across parameter space. Allowing for more general covariance functions may enable the interpolant to better model the true underlying model errors and this will in turn lead to narrower peaks in the posterior distribution, and hence more precise parameter estimates. However, overly complicated covariance functions run the risk of overfitting the model, and this may lead to erroneously narrow posterior peaks and possibly introduce systematic errors into the parameter estimates; this is precisely the problem that the marginalised likelihood was originally introduced to solve. Without more work using a range of different covariance function it is difficult to know how close to being optimal the covariance functions considered up until this point are.

The first detection of gravitational waves has opened the door on a new era of possibilities for testing general relativity. Indeed, the initial announcement of the detection of the first binary black hole was accompanied by the publication of a paper dedicated to checking whether the observed signal was consistent with the expectations of general relativity (The LIGO Scientific Collaboration and the Virgo Collaboration, 2016*f*); no statistically significant inconsistencies were found. This thesis has considered the possibility of using observations of extreme mass-ratio inspirals to test the nature of the gravitational field around an astrophysical black hole. We are still some years away from being able to make such observations, and the work described in this thesis is at a correspondingly early stage. In particular, as the eLISA mission design is finalised

over the coming years the work here should be extended in several ways: by using a more detailed model of the spacecraft and its noise spectrum; by using more realistic gravitational waveform models, in place of the analytic kludge models used here; and by performing full Bayesian parameter estimation using MCMC methods (or an alternative stochastic sampling technique) in place of the Fisher matrices used here.

The first detection of gravitational waves also provides encouragement for the efforts to detect lower frequency gravitational waves, both nHz waves by pulsar timing and mHz waves by space-based detectors. At the beginning of my PhD, when I made the first version of the sensitivity curve plots presented in Chapter 4, no gravitational wave of any frequency had been detected. I hope that over the next few years we will be able to fill in these plots with detections spanning the entire range of frequencies shown, and possibly more besides. I look forward to playing a small part in those discoveries.

Part V

Appendices

A Bayesian monochromatic sensitivity

Rearranging Eq. 2.94 in Sec. 2.4.4 gives the Bayesian expression for the sensitivity of a PTA to a monochromatic source.

$$
\begin{aligned}
h_c =& 4\chi\sqrt{\log(\mathcal{B})}\left(-\frac{90N_p\cos^2(\phi)}{\pi^6\delta t f^8\sigma^2T^5}-\frac{90N_p\cos^2(2\pi fT+\phi)}{\pi^6\delta t f^8\sigma^2T^5}+\frac{180N_p\cos(\phi)\cos(2\pi fT+\phi)}{\pi^6\delta t f^8\sigma^2T^5}\right.\\
&+\frac{180N_p\sin(\phi)\cos(\phi)}{\pi^5\delta t f^7\sigma^2T^4}+\frac{180N_p\cos(\phi)\sin(2\pi fT+\phi)}{\pi^5\delta t f^7\sigma^2T^4}-\frac{180N_p\sin(\phi)\cos(2\pi fT+\phi)}{\pi^5\delta t f^7\sigma^2T^4}\\
&-\frac{180N_p\sin(2\pi fT+\phi)\cos(2\pi fT+\phi)}{\pi^5\delta t f^7\sigma^2T^4}-\frac{96N_p\sin^2(\phi)}{\pi^4\delta t f^6\sigma^2T^3}-\frac{96N_p\sin^2(2\pi fT+\phi)}{\pi^4\delta t f^6\sigma^2T^3}\\
&-\frac{168N_p\sin(\phi)\sin(2\pi fT+\phi)}{\pi^4\delta t f^6\sigma^2T^3}+\frac{120N_p\cos(\phi)\sin(\pi fT)\sin(\pi fT+\phi)}{\pi^4\delta t f^6\sigma^2T^3}\\
&-\frac{120N_p\sin(\pi fT)\sin(\pi fT+\phi)\cos(2\pi fT+\phi)}{\pi^4\delta t f^6\sigma^2T^3}-\frac{144N_p\sin(\phi)\sin(\pi fT)\sin(\pi fT+\phi)}{\pi^3\delta t f^5\sigma^2T^2}\\
&-\frac{96N_p\sin(\pi fT)\sin(\pi fT+\phi)\sin(2\pi fT+\phi)}{\pi^3\delta t f^5\sigma^2T^2}-\frac{24N_p\sin(\phi)\cos(2\pi fT+\phi)}{\pi^3\delta t f^5\sigma^2T^2}\\
&+\frac{24N_p\sin(2\pi fT+\phi)\cos(2\pi fT+\phi)}{\pi^3\delta t f^5\sigma^2T^2}-\frac{72N_p\sin^2(\pi fT)\sin^2(\pi fT+\phi)}{\pi^2\delta t f^4\sigma^2T}\\
&-\frac{24N_p\cos^2(2\pi fT+\phi)}{\pi^2\delta t f^4\sigma^2T}-\frac{48N_p\sin(\pi fT)\sin(\pi fT+\phi)\cos(2\pi fT+\phi)}{\pi^2\delta t f^4\sigma^2T}\\
&\left.+\frac{N_p\sin(2\phi)}{\pi\delta t f^3\sigma^2}-\frac{N_p\sin(2(2\pi fT+\phi))}{\pi\delta t f^3\sigma^2}+\frac{4N_pT}{\delta t f^2\sigma^2}\right)^{-1/2}
\end{aligned}
\tag{1}
$$

B The effect of small changes in the noise PSD on the GPR interpolant

In the offline stage of the method, the GP was trained using the hyperlikelihood in Eq. 5.18. The result of this process was an interpolant which enabled fast online PE. However, this splitting into offline and online stages also has a potential problem, because the training process makes use of the overlap matrix $M_{ij} = \langle \delta h(\vec{\lambda}_i) | \delta h(\vec{\lambda}_j) \rangle$ which, in turn, depends upon the detector noise PSD, $S_n(f)$. The noise PSD is not constant; it changes on short timescales as the noise drifts in the instrument (e.g., Aasi et al. (2015*b*)), on longer timescales it changes more dramatically as the instrument is gradually upgraded (Abbott et al., 2013). There are also differences between different detectors, for example between the aLIGO and AdV instruments (or even between the two aLIGO interferometers). It would be a significant drawback if the offline training stage of the process had to be repeated for every single candidate signal because of small differences in the detector PSD.

We do not expect small changes in the noise curve to have a significant effect on the resulting interpolant. First, the noise can be rescaled by an overall constant and have no effect on the position of the peak in the hyperlikelihood; this can be seen from Eq. 5.18. Second, the peak in the hyperlikelihood is typically wide, and using the hyperparameters from anywhere in the vicinity of the peak still gives reasonable, if not perfect, interpolation. Accordingly, when the PSD changes, some of the difference can be absorbed by an overall scaling, which has no effect on the results, and the remaining change shifts the peak of the hyperlikelihood away from the previously optimised values, but not enough to limit their applicability. If this is the case, then GPs trained on slightly different noise PSDs perform nearly identically to each other and there is no need to retrain for the new PSD.

To assess the sensitivity of our results to changes in the noise curve, we considered three different noise curves chosen to represent the range of possibilities in the advanced-detector era. These are: an estimate of the observing run 1 (O1) aLIGO sensitivity (the early curve of Barsotti and Fritschel (2012)); the zero-detuned high-power (ZDHP) design sensitivity of aLIGO (Aasi et al., 2015*a*; Shoemaker, 2010), and the design sensitivity of AdV (Accadia, Acernese, Agathos et al., 2012; Acernese et al., 2009). As an additional check, we also considered an inverted top-hat noise curve. All of these noise curves are

plotted in the left-hand panel of Fig. 1. We then took the training set $\mathcal{D}_0$ and trained the SE GP to find the optimum hyperparameters. Shown in the centre-panel of Fig. 1 is the hyperlikelihood surface as a function of chirp mass length scale for the different noise curves. As expected, for the range of realistic noise curves the peak in the hyperlikelihood only shifts by a small amount. Finally we used the optimum hyperparameters from each of these hyperlikelihood surfaces to interpolate the training set and calculated the overlap to the accurate waveforms using the ZDHP noise curve; the results of this are shown in the right-hand panel of Fig. 1. For the range of realistic noise curves, the overlap is equally good; cf. Stein (1999). Although the inverted top-hat noise curve gives noticeably lower overlaps, even in that case the drop in the overlap is still less than 0.1%, which is smaller than the difference between the approximate and GPR likelihoods.

This suggests it is safe to train a GP with a fixed noise curve (typical for the instruments considered). The resulting interpolants can be used to analyse all signals without worrying about small drifts in the noise.

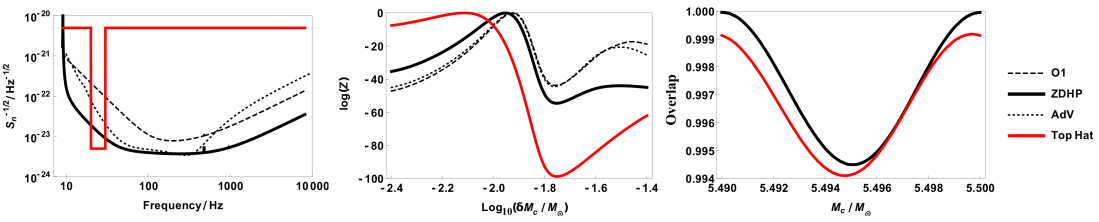


Figure 1: The left-hand panel shows three different noise curves for ground-based detectors in the advanced era. Also shown in red is an unrealistic noise which we use for comparison. The centre-panel shows the hyper-likelihood surface for the training set $\mathcal{D}_0$ using the SE covariance function and with the overlap matrix calculated using each of the noise curves in the left-hand panel. The right-hand panel shows the waveform overlap between the accurate and the interpolated waveform evaluated for parameter values between two training set points. The interpolants based on the realistic noise curves perform equally well (the curves lie on top of each other). The unrealistic noise curve performs worst, but still gives overlaps greater than 0.994.

C Continuity and differentiability of GPs

In this appendix we give proofs of the results stated in Sec. 5.4 concerning the continuity and differentiability of GPs, following the approach of Adler (1981). Let $\vec{\lambda}_1, \vec{\lambda}_2, \vec{\lambda}_3 \ldots$ be a sequence of points in parameter space which converges to a point $\vec{\lambda}_*$, in the sense $\lim_{\ell\to\infty} |\vec{\lambda}_\ell - \vec{\lambda}_*| = 0$, where, as in Sec. 5.4, $|\vec{x}|$ denotes the norm with respect to the metric on parameter space, as discussed in Sec. 5.4.1. The GP $Y(\vec{\lambda})$ is said to be MS continuous at $\vec{\lambda}_*$ if

$$\lim_{\ell\to\infty} \mathbb{E}\left[\left(Y(\vec{\lambda}_\ell) - Y(\vec{\lambda}_*)\middle| Y(\vec{\lambda}_\ell) - Y(\vec{\lambda}_*)\right)\right] = 0\,, \tag{2}$$

where $\mathbb{E}[\ldots]$ denotes the expectation of the enclosed quantity over realisations of the GP. For notational convenience, we denote this MS limit as

$$Y(\vec{\lambda}_*) = \underset{\ell\to\infty}{\text{l.i.m.}}\, Y(\vec{\lambda}_\ell)\,, \tag{3}$$

where l.i.m. stands for limit in mean (Wiener, 1949*a*). MS continuity implies continuity in the mean,

$$\lim_{\ell\to\infty} \mathbb{E}\left[Y(\vec{\lambda}_\ell) - Y(\vec{\lambda}_*)\right] = 0\,. \tag{4}$$

This follows from considering the variance of the quantity $Y(\vec{\lambda}_\ell) - Y(\vec{\lambda}_*)$, and the fact that variance is non-negative. There are other notions of continuity of GPs used in the literature, but the notion of MS continuity relates most easily to the covariance. The mean and the covariance of a GP are defined as

$$\begin{aligned} m(\vec{\lambda}) &= \mathbb{E}\left[Y(\vec{\lambda})\right]\,, \\ k(\vec{\lambda}_1, \vec{\lambda}_2) &= \mathbb{E}\left[\left(Y(\vec{\lambda}_1) - m(\vec{\lambda}_1)\middle| Y(\vec{\lambda}_2) - m(\vec{\lambda}_2)\right)\right]. \end{aligned} \tag{5}$$

Using these, Eq. 2 can be written as

$$\begin{aligned} \lim_{\ell\to\infty} \Big\{ & k(\vec{\lambda}_*, \vec{\lambda}_*) - 2k(\vec{\lambda}_\ell, \vec{\lambda}_*) + k(\vec{\lambda}_\ell, \vec{\lambda}_\ell) \\ & + \left(m(\vec{\lambda}_*) - m(\vec{\lambda}_\ell)\middle| m(\vec{\lambda}_*) - m(\vec{\lambda}_\ell)\right)\Big\} \;=\; 0\,, \end{aligned} \tag{6}$$

and using the continuity of the mean in Eq. 4 gives

$$\lim_{\ell\to\infty} \left[k(\vec{\lambda}_*, \vec{\lambda}_*) - 2k(\vec{\lambda}_\ell, \vec{\lambda}_*) + k(\vec{\lambda}_\ell, \vec{\lambda}_\ell)\right] = 0\,. \tag{7}$$

This condition is satisfied if the covariance function is continuous at the point $\vec{\lambda}_1 = \vec{\lambda}_2 = \vec{\lambda}_*$. Therefore, we arrive at the result that if the covariance function is continuous in the usual

sense at some point $\vec{\lambda}_*$, then the corresponding GP is MS continuous at this point. In fact, a GP is continuous in MS if *and only if* the covariance function is continuous (Adler, 1981), although this is not proved here. In the special case of stationary covariance this reduces to checking continuity of $k(\vec{\tau})$ at $\vec{\tau} = 0$, and in the special case of isotropic covariance, continuity of $k(\tau)$ at $\tau = 0$.

We now move on from continuity to consider differentiability. In the spirit of Eq. 2, the notion of taking the MS derivative of a GP is defined as

$$\frac{\partial Y(\vec{\lambda})}{\partial \vec{\lambda}^a} = \underset{\epsilon \to 0}{\text{l.i.m.}}\, X_a(\vec{\lambda}, \epsilon)\,, \tag{8}$$

$$\text{where } X_a(\vec{\lambda}, \epsilon) = \frac{Y(\vec{\lambda} + \epsilon\, \hat{e}_a) - Y(\vec{\lambda})}{\epsilon} \tag{9}$$

with parameter-space unit vector $\hat{e}_a$. This definition can be easily extended to higher-order derivatives (Adler, 1981).The MS derivative of a GP is also a GP; this follows simply from the fact that the sum of Gaussians is also distributed as a Gaussian. The covariance of $X_a(\vec{\lambda}, \epsilon)$ is given by

$$K_\epsilon(\vec{\lambda}_1, \vec{\lambda}_2) = \mathbb{E}\Big[\Big(X_a(\vec{\lambda}_1, \epsilon) - \Xi(\vec{\lambda}_1, \epsilon)\Big| \\ X_a(\vec{\lambda}_2, \epsilon) - \Xi(\vec{\lambda}_2, \epsilon)\Big)\Big] \tag{10}$$

where $\Xi_a(\vec{\lambda}, \epsilon) = \mathbb{E}[X_a(\vec{\lambda}, \epsilon)]$. It then follows that

$$K_\epsilon(\vec{\lambda}_1, \vec{\lambda}_2) = \frac{k(\vec{\lambda}_1 + \epsilon, \vec{\lambda}_2 + \epsilon) - k(\vec{\lambda}_1, \vec{\lambda}_2 + \epsilon)}{\epsilon^2} \\ - \frac{k(\vec{\lambda}_1 + \epsilon, \vec{\lambda}_2) - k(\vec{\lambda}_1, \vec{\lambda}_2)}{\epsilon^2}\,. \tag{11}$$

Substituting this into Eq. 9, the limit in MS becomes a normal limit, and the result is obtained that the MS derivative of a MS continuous GP with covariance function $k(\vec{\lambda}_1, \vec{\lambda}_2)$ is a GP with covariance function $\partial^2 k(\vec{\lambda}_1, \vec{\lambda}_2)/\partial \vec{\lambda}_1^a \partial \vec{\lambda}_2^a$. In general the covariance function of the n_d-times MS differentiated GP

$$\frac{\partial^{n_d} Y(\vec{\lambda})}{\partial \vec{\lambda}^{a_1} \partial \vec{\lambda}^{a_2} \ldots \partial \vec{\lambda}^{a_{n_d}}}\,, \tag{12}$$

is given by the $2n_d$-times differentiated function

$$\frac{\partial^{2n_d} k(\vec{\lambda}_1, \vec{\lambda}_2)}{\partial \vec{\lambda}_1^{a_1} \partial \vec{\lambda}_2^{a_1} \partial \vec{\lambda}_1^{a_2} \partial \vec{\lambda}_2^{a_2} \ldots \partial \vec{\lambda}_1^{a_{n_d}} \partial \vec{\lambda}_2^{a_{n_d}}}\,. \tag{13}$$

From the above results relating the MS continuity of GPs to the continuity of the covariance function at $\vec{\lambda}_1 = \vec{\lambda}_2 = \vec{\lambda}_*$, it follows that the n_d-times MS derivative of the GP is MS continuous (the GP is said to be n_d-times MS differentiable) if the $2n_d$-times derivative of the covariance function is continuous at $\vec{\lambda}_1 = \vec{\lambda}_2 = \vec{\lambda}_*$ (Stein, 1999). So it is the smoothness properties of the covariance function along the diagonal points that determines the differentiability of the GP. (It can also be shown that if a covariance function is continuous at all diagonal points $\vec{\lambda}_1 = \vec{\lambda}_2$ then it's everywhere continuous.)

D Systematic bias due to waveform errors

We assume that an approximate model $H(\vec{\lambda})$ is used to recover the parameters of a gravitational wave signal that is described by the true model $h(\vec{\lambda})$ with parameters $\vec{\lambda}_0$. The best fit parameters of the approximate model are $\vec{\lambda}_{\rm bf} = \vec{\lambda}_0 + \Delta\vec{\lambda}$. These parameters minimise the squared distance between the true and approximate model spaces,

$$\left\langle \delta h(\vec{\lambda}_0) + H(\vec{\lambda}_{\rm bf}) - H(\vec{\lambda}_0) | \delta h(\vec{\lambda}_0) + H(\vec{\lambda}_{\rm bf}) - H(\vec{\lambda}_0) \right\rangle \tag{14}$$

where $\delta h(\vec{\lambda}_0) = H(\vec{\lambda}_0) - h(\vec{\lambda}_0)$ (note the different sign convention from Cutler and Vallisneri (2007)). Differentiating with respect to each of the parameters in turn, we find that the best-fit parameters must satisfy the equations

$$\langle \delta h(\vec{\lambda}_0) + H(\vec{\lambda}_{\rm bf}) - H(\vec{\lambda}_0) | \partial_a \left(H(\vec{\lambda}_{\rm bf}) - H(\vec{\lambda}_0) \right) \rangle = 0 \,. \tag{15}$$

We use the notation $\partial_a x \equiv \partial x / \partial \lambda^a$ and subsequently will use $\partial_{ab} x \equiv \partial^2 x / \partial\lambda^a \partial\lambda^b$. If we now assume that the approximation is good, we can write $\delta h(\vec{\lambda}) \sim \mathcal{O}(\epsilon)$, a small parameter, and $\vec{\lambda}_{\rm bf} = \vec{\lambda}_0 + \Delta\vec{\lambda}$ with $\Delta\lambda^i \sim \mathcal{O}(\epsilon)\ \forall i$. We can then expand the difference between the approximate waveforms as a Taylor series

$$\begin{aligned} H(\vec{\lambda}_{\rm bf}) - H(\vec{\lambda}_0) &= \partial_a H(\vec{\lambda}_0)\Delta\lambda^a + \\ &\quad \frac{1}{2}\partial_{ab} H(\vec{\lambda}_0)\Delta\lambda^a\Delta\lambda^b + \cdots \,. \end{aligned} \tag{16}$$

Eq. 15 becomes

$$\left\langle \delta h(\vec{\lambda}_0) + \partial_b H(\vec{\lambda}_0)\Delta\lambda^b + \frac{1}{2}\partial_{bc} H(\vec{\lambda}_0)\Delta\lambda^b\Delta\lambda^c | \partial_a H(\vec{\lambda}_0) + \partial_{ad} H(\vec{\lambda}_0)\Delta\lambda^d \right\rangle = 0 \,. \tag{17}$$

where all derivatives are now evaluated at $\vec{\lambda}_0$. Keeping only terms of order ϵ we find the Cutler and Vallisneri result

$$\Delta\lambda_1^a = -\left(\Sigma^{-1}\right)^{ab} \langle \delta h(\vec{\lambda}_0) | \partial_b H(\vec{\lambda}_0) \rangle \tag{18}$$

where $\Sigma^{ij} \equiv \langle \partial_a H(\vec{\lambda}_0) | \partial_b H(\vec{\lambda}_0) \rangle$ is the Fisher Matrix.

We now extend to the next order in ϵ by writing $\Delta\lambda^a = \Delta\lambda_1^a + \Delta\lambda_2^a$, where $\Delta\lambda_1^i$ is the previous solution, Eq. 18. Keeping terms to $\mathcal{O}(\epsilon^2)$ we obtain

$$\begin{aligned} \Delta\lambda_2^a = -\left(\Sigma^{-1}\right)^{ab} \Big[&\langle \delta h(\vec{\lambda}_0) | \partial_{ab} H(\vec{\lambda}_0) \rangle \Delta\lambda_1^b + \langle \partial_{ac} H(\vec{\lambda}_0) | \partial_b H(\vec{\lambda}_0) \rangle \Delta\lambda_1^b \Delta\lambda_1^c \\ &+ \frac{1}{2} \langle \partial_a H(\vec{\lambda}_0) | \partial_{ab} H(\vec{\lambda}_0) \rangle \Delta\lambda_1^b \Delta\lambda_1^c \Big] \,. \end{aligned} \tag{19}$$

A suitable validity criterion for the Cutler and Vallisneri formula, Eq. 18, is

$$\max_a \{|\Delta\lambda_2^a/\Delta\lambda_1^a|\} \ll 1. \tag{20}$$

We note also that Eq. 19 provides an improved estimate of the systematic bias and that we can readily extend this method to higher order in ϵ by including further terms in the expansion in Eq. 16.

E Metric components

In this appendix explicit coordinate expressions are given for some of the metric components of the bumpy black hole spacetimes used in chapters 9 and 10.

E.1 BN

Here expressions are given for all of the bumpy $\mathcal{B}_N$ metrics discussed in Sec. 9.1; all of the metric coefficients up to $\mathcal{O}(1/r^5)$ are reproduced here.

$$\begin{aligned}
h_{tt,2} = &\ \gamma_{1,2} + 2\gamma_{4,2} - 2a\gamma_{3,1}\sin^2\theta \\
h_{tt,3} = &\ \gamma_{1,3} - 8\gamma_{4,2} - 2\gamma_{1,2} + 2\gamma_{4,3} + 8a\gamma_{3,1}\sin^2\theta \\
h_{tt,4} = &\ -8\gamma_{4,3} - 2\gamma_{1,3} + 2\gamma_{4,4} + 8\gamma_{4,2} + \gamma_{1,4} - 8a\gamma_{3,1}\sin^2\theta \\
&+a^2\left(\gamma_{1,2} + 2\gamma_{4,2}\right)\sin^2\theta + 2a^3\gamma_{3,1}\cos^2\theta\sin^2\theta \\
h_{tt,5} = &\ 16a^3\gamma_{3,1}\sin^4\theta + \sin^2\theta\left[4a\gamma_{3,3} + a^2\left(\gamma_{1,3} - 2\gamma_{1,2}\right.\right. \\
&\left.\left.-12\gamma_{4,2} + 2\gamma_{4,3}\right) - 12a^3\gamma_{3,1}\right] + a^2\left(8\gamma_{4,2} + 2\gamma_{1,2}\right) \\
&+\gamma_{1,5} + 2\gamma_{4,5} - 2\gamma_{1,4} + 8\gamma_{4,3} - 8\gamma_{4,4}
\end{aligned} \tag{21}$$

$$\begin{aligned}
h_{t\phi,2} = &\ -M\sin^2\theta\left[\gamma_{3,3} + a\left(\gamma_{1,2} + \gamma_{4,2}\right) + a^2\gamma_{3,1}\right] \\
h_{t\phi,3} = &\ -8Ma^2\gamma_{3,1}\sin^4\theta \\
&+M\sin^2\theta\left[\left(2\gamma_{3,3} - \gamma_{3,4}\right) + a\left(6\gamma_{4,2} - \gamma_{4,3}\right.\right. \\
&\left.\left.+2\gamma_{1,2} - \gamma_{1,3}\right) + 2\gamma_{3,1}a^2\right] \\
h_{t\phi,4} = &\ M\sin^4\theta\left[a^2\left(8\gamma_{3,1} - \gamma_{3,3}\right) + a^3\left(-\gamma_{1,3} - \gamma_{4,2}\right)\right. \\
&\left.-a^4\gamma_{3,1}\right] + \sin^2\theta\left[\left(2\gamma_{3,4} - \gamma_{3,5}\right) + a\left(-\gamma_{4,4}\right.\right. \\
&\left.\left.-8\gamma_{4,2} + 6\gamma_{4,3} - \gamma_{1,4} + 2\gamma_{1,3}\right) - a^2\gamma_{3,3}\right] \\
h_{t\phi,5} = &\ -16Ma^4\gamma_{3,1}\sin^6\theta \\
&+M\sin^4\theta\left[a^2\left(-2\gamma_{3,3} - \gamma_{3,4}\right)\right. \\
&\left.+a^3\left(\gamma_{4,3} + 10\gamma_{4,2} + 2\gamma_{1,2} - \gamma_{1,3}\right) + 14a^4\gamma_{3,1}\right] \\
&+\sin^2\theta\left[\left(2\gamma_{3,5} - \gamma_{3,6}\right)\right. \\
&+a\left(-\gamma_{1,5} - 8\gamma_{4,3} - \gamma_{4,5} + 2\gamma_{1,4} + 6\gamma_{4,4}\right) \\
&\left.-\gamma_{3,4}a^2 + a^3\left(-2\gamma_{1,2} - 6\gamma_{4,2}\right) - 2a^4\gamma_{3,1}\right]
\end{aligned} \tag{22}$$

$$
\begin{aligned}
h_{rr,2} &= -\gamma_{1,2} \\
h_{rr,3} &= -\gamma_{1,3} - 2\gamma_{1,2}, \\
h_{rr,4} &= -\gamma_{1,4} - 2\gamma_{1,3} - 4\gamma_{1,2} + (1/2)\gamma_{1,2}a^2\left(1-\cos 2\theta\right) \\
h_{rr,5} &= a^2\sin^2\theta\left(\gamma_{1,3}+2\gamma_{1,2}\right) \\
&\quad -\gamma_{1,5} - 2\gamma_{1,4} - 4\gamma_{1,3} - 8\gamma_{1,2} + 2a^2\gamma_{1,2}
\end{aligned} \tag{23}
$$

$$
\begin{aligned}
h_{\phi\phi,-2} &= 0 \\
h_{\phi\phi,-1} &= 0 \\
h_{\phi\phi,0} &= 2M^2 a\gamma_{3,1}\sin^4\theta \\
h_{\phi\phi,1} &= 0 \\
h_{\phi\phi,2} &= M^2\sin^4\theta\left[2a\gamma_{3,3} + a^2\gamma_{1,2} + a^3\gamma_{3,1}\left(4-2\cos^2\theta\right)\right] \\
h_{\phi\phi,3} &= 8M^2a^3\gamma_{3,1}\sin^6\theta + M^2\sin^4\theta\left[a\left(-4\gamma_{3,3}+2\gamma_{3,4}\right)\right. \\
&\quad \left. +a^2\left(-2\gamma_{1,2}-4\gamma_{4,2}+\gamma_{1,3}\right) - 4a^3\gamma_{3,1}\right]
\end{aligned} \tag{24}
$$

E.2 CS2

Here expressions are given for the CS2 metric discussed in Sec. 9.1.1.3. The following expressions use $f(r) = 1-(2/r)$.

$$
\begin{aligned}
\delta\left(g_{tt}^{\mathrm{CS2}}\right) &= \zeta a^2\frac{1^3}{r^3}\left[\frac{201}{1792}\left(1+\frac{1}{r}+\frac{4474}{4221}\frac{1^2}{r^2}\right.\right. \\
&\quad -\frac{2060}{469}\frac{1^3}{r^3}+\frac{1500}{469}\frac{1^4}{r^4}-\frac{2140}{201}\frac{1^5}{r^5} \\
&\quad \left.+\frac{9256}{201}\frac{1^6}{r^6}-\frac{5376}{67}\frac{1^7}{r^7}\right)\left(3\cos^2\theta-1\right) \\
&\quad -\frac{5}{384}\frac{1^2}{r^2}\left(1+100\frac{1}{r}\right. \\
&\quad \left.\left.+194\frac{1^2}{r^2}+\frac{2220}{7}\frac{1^3}{r^3}-\frac{1512}{5}\frac{1^4}{r^4}\right)\right],
\end{aligned} \tag{25}
$$

$$
\delta\left(g_{t\phi}^{\mathrm{CS2}}\right) = \frac{5}{4}\zeta\chi\frac{1}{r^4}\left(1+\frac{12}{7r^2}+\frac{27}{10r^2}\right) \tag{26}
$$

$$
\begin{aligned}
\delta\left(g_{rr}^{\mathrm{CS2}}\right) &= \zeta a^2\frac{1^3}{r^3 f(r)^2}\left[\frac{201}{1792}f(r)\left(1+\frac{1459}{603}\frac{1}{r}\right.\right. \\
&\quad \left.\left.+\frac{20000}{4221}\frac{1^2}{r^2}+\frac{51580}{1407}\frac{1^3}{r^3}-\frac{7580}{201}\frac{1^4}{r^4}\right.\right.
\end{aligned}
$$

$$
\begin{aligned}
& -\frac{22492}{201}\frac{1^5}{r^5} - \frac{40320}{67}\frac{1^6}{r^6}\Bigg)(3\cos^2\theta - 1) \\
& -\frac{25}{384}\frac{1}{r}\Bigg(1 + 3\frac{1}{r} + \frac{322}{5}\frac{1^2}{r^2} + \frac{198}{5}\frac{1^3}{r^3} \\
& +\frac{6276}{175}\frac{1^4}{r^4} - \frac{17496}{25}\frac{1^5}{r^5}\Bigg)\Bigg], \qquad (27)
\end{aligned}
$$

$$
\begin{aligned}
\delta\left(g_{\theta\theta}^{\mathrm{CS2}}\right) &= \frac{201}{1792}\zeta a^2 1^2 \frac{1}{r}\Bigg(1 + \frac{1420}{603}\frac{1}{r} + \frac{18908}{4221}\frac{1^2}{r^2} \\
& +\frac{1480}{603}\frac{1^3}{r^3} + \frac{22460}{1407}\frac{1^4}{r^4} \\
& +\frac{3848}{201}\frac{1^5}{r^5} + \frac{5376}{67}\frac{1^6}{r^6}\Bigg)(3\cos^2\theta - 1) \qquad (28)
\end{aligned}
$$

$$
\delta\left(g_{\phi\phi}^{\mathrm{CS2}}\right) = \sin^2\theta\, g_{\theta\theta}^{\mathrm{CS2}}\,. \qquad (29)
$$

F Radial dependence of the flux

In this appendix the expression for the radial dependence of the flux from an accretion disk, used in Sec. 9.2 of the main text, is derived. This derivation follows that of Page and Thorne (1974).

The thin disk is assumed to be axisymmetric, stationary, and lying in the equatorial place; therefore all quantities in the disk depend only on the radial coordinate. It is assumed that the material in the disk moves (very nearly) on circular geodesics. The four-velocity of individual fluid elements, u_0^μ, when mass averaged over the disk structure must therefore be the four-velocity of the geodesic orbit u^μ in Eq. 9.25.

$$u^\mu = \frac{1}{\Sigma(r)} \int_{-h}^{+h} \mathrm{d}z \; \rho_0 u_0^\mu \,, \tag{30}$$

where

$$\Sigma(r) = \int_{-h}^{+h} \mathrm{d}z \; \rho_0 \,. \tag{31}$$

Where ρ_0 mass density in the rest frame of the orbiting material. Without loss of generality the stress-energy tensor may be decomposed by writing

$$T^{\mu\nu} = \rho_0 \left(1 + \Pi\right) u^\mu u^\nu + t^{\mu\nu} + 2u^{(\mu} q^{\nu)} \tag{32}$$

where the physical interpretation of each term becomes clear in the rest frame of the orbiting material; Π is the specific internal energy, $t^{\mu\nu}$ is the stress tensor in the averaged rest frame of the material and q^μ is the energy flow vector. The round brackets in the superscript of Eq. 32 denote symmeterisation with respect to the enclosed indices. The tensors $t^{\mu\nu}$ and q^μ obey the orthogonality relations $u_\mu q^\mu = 0$ and $u_\mu t^{\mu\nu} = u_\nu t^{\mu\nu} = 0$. By analogy with $\Sigma(r)$ in Eq. 30 the average stresses in the disk are defined as

$$W_\mu^{\ \nu} = \int_{-h}^{+h} \mathrm{d}z \; t_\mu^{\ \nu} \,. \tag{33}$$

It is also assumed that heat flow within the disk is negligible except for in the vertical direction, which is reasonable as the disk is thin. The non-local heating effects due to light emitted by one portion of the disk being re-absorbed by another portion are also neglected.

$$q^t = q^r = q^\phi = 0 \,, \quad \text{at} \quad z = \pm h. \tag{34}$$

Since the only time-averaged stress that reaches out of the disk to infinity is carried by photons (neglecting gravitational radiation and any coherent superposition of long

wavelength radiation), and using Eq. 34, on the upper and lower edges of the disk the following terms of the stress-tensor disappear;

$$t_{\phi}^{\ z} = t_{r}^{\ z} = t_{t}^{\ z} = 0 \tag{35}$$

and

$$|q^{z}| = F(r)\,,\ \text{at } z = \pm h\,. \tag{36}$$

Since we are assuming the particles in the disk are, very nearly, on circular, geodesic orbits it follows that the acceleration due to pressure gradients in the disk must be much less than the acceleration due to gravity otherwise the material would be pushed off its geodesic trajectory. Using the approximate relation $t_{rr} \approx \rho_0 \Pi$ (which is valid for any astrophysically reasonable matter (Page and Thorne, 1974)) leads directly to the condition of negligible specific heat;

$$\begin{aligned}
&\text{radial pressure acceleration} \approx |\partial_r t_{rr}/\rho_0| \approx |\partial_r\,(t_{rr}/\rho_0)| \\
&\text{gravitational acceleration} \approx |\partial_r \tilde{E}| \approx |\partial_r (1-\tilde{E})| \\
&|\partial_r (t_{rr}/\rho_0)| \ll |\partial_r (1-\tilde{E})| \quad \Rightarrow \quad \Pi \ll 1 - \tilde{E}\,.
\end{aligned} \tag{37}$$

If the internal energy is small compared to the gravitational potential energy, this means that as the material spirals in towards the black hole all of the gravitational potential energy is radiated away.

With these simplifying assumptions in place the equations governing the structure of the disk follow from the conservation of stress-energy ($\nabla_\mu T^{\mu\nu} = 0$), and the conservation of rest mass of the fluid (Misner et al., 1973),

$$\nabla_\mu\,(\rho_0 u^\mu) = 0\,. \tag{38}$$

This is integrated over the spacetime volume $\{\mathcal{V} : t \in (t_0, t_0+T)\,,\ r \in (r, r+\Delta r)\,,\ \phi \in (0, 2\pi)\,,\ z \in (-h, +h)\}$. Gauss's theorem is then used to convert the volume integral into a surface integral over the boundary $\partial\mathcal{V}$ with area element $|d^3A|$.

$$\begin{aligned}
0 &= \int_{\partial\mathcal{V}} \rho_0 u^\mu n_\mu \left| d^3A \right| \qquad (39)\\
0 &= \left[\int_r^{r+\Delta r} \int_0^{2\pi} \int_{-h}^{+h} \mathrm{d}r\mathrm{d}\phi\mathrm{d}z\ \sqrt{-\mathbf{g}}\rho_0 u^t \right]_{t=t_0}^{t=t_0+T} \\
&\quad + \left[\int_t^{t+T} \int_0^{2\pi} \int_{-h}^{+h} \mathrm{d}t\mathrm{d}\phi\mathrm{d}z\ \sqrt{-\mathbf{g}}\rho_0 u^r \right]_{r'=r}^{r'=r+\Delta r}
\end{aligned}$$

$$+\left[\int_t^{t_0+T}\int_r^{r+\Delta r}\int_{-h}^{+h}\mathrm{d}t\mathrm{d}r\mathrm{d}z\ \sqrt{-\mathbf{g}}\rho_0 u^\phi\right]_{\phi=0}^{\phi=2\pi}$$
$$+\left[\int_{t_0}^{t_0+T}\int_r^{r+\Delta r}\int_0^{2\pi}\mathrm{d}t\mathrm{d}r\mathrm{d}\phi\ \sqrt{-\mathbf{g}}\rho_0 u^z\right]_{z=-h}^{z=+h} \tag{40}$$

The first and third terms in the above expression are zero by the assumed stationarity and axisymmetry of the system. The final term is also zero because there is no motion in the vertical direction, $u^z = 0$. Therefore Eq. 40 simplifies to

$$0 = 2\pi T\Delta r\left(\sqrt{-\mathbf{g}}\Sigma(r)u^r\right)_{,r}$$
$$\Rightarrow \dot{M}_0 = -2\pi\sqrt{-\mathbf{g}}\Sigma(r)u^r = \text{constant}\,, \tag{41}$$

where $\dot{M}_0$ is the accretion rate.

The second conservation law is that of angular momentum. Again the differential form of the conservation law is integrated over the volume $\mathcal{V}$ and Gauss's law used to turn this into a surface integral over $\partial\mathcal{V}$.

$$0 = \nabla_\mu J^\mu \quad \text{where} \quad J^\mu = T^{\mu\nu}\left(\frac{\partial}{\partial\phi}\right)_\nu \tag{42}$$
$$0 = \int_{\partial\mathcal{V}} J^\mu n_\mu \left|d^3A\right| \tag{43}$$

$$0 = \left[\int_{t_0}^{t_0+T}\int_0^{2\pi}\int_{-h}^{+h}\mathrm{d}t\mathrm{d}\phi\mathrm{d}z\ \left[\rho_0(1+\Pi)u_\phi u^r + t_\phi^{\ r} + u_\phi q^r + q_\phi u^r\right]\sqrt{-\mathbf{g}}\right]_{r'=r}^{r'=r+\Delta r}$$
$$+\left[\int_{t_0}^{t_0+\Delta t}\int_r^{r+\Delta r}\int_0^{2\pi}\mathrm{d}t\mathrm{d}r\mathrm{d}\phi\ \left[\rho_0(1+\Pi)u_\phi u^z + t_\phi^{\ z} + u_\phi q^z + q_\phi u^z\right]\sqrt{-\mathbf{g}}\right]_{z=-h}^{z=+h} \tag{44}$$

The ϕ index has been lowered and the t and ϕ integral terms have been set to zero due to the assumed stationarity and axisymmetry of the system. Using the negligible internal energy condition derived above, Eqs. 34 and 35, and the fact that $u^z = 0$ this becomes

$$0 = \left[2\pi T\int_{-h}^{+h}\mathrm{d}z\ \left[\rho_0 u_\phi u^r + t_\phi^{\ r}\right]\sqrt{-\mathbf{g}}\right]_{r'=r}^{r'=r+\Delta r}$$
$$+\left[2\pi T\int_r^{r+\Delta r}\mathrm{d}r\ u_\phi q^z\sqrt{-\mathbf{g}}\right]_{z=-h}^{z=+h} \tag{45}$$
$$\Rightarrow 4\pi\sqrt{-\mathbf{g}}F(r)\tilde{L} = \left[\dot{M}_0\tilde{L} - 2\pi\sqrt{-\mathbf{g}}W_\phi^{\ r}\right]_{,r}\,. \tag{46}$$

An extra factor of two has appeared on the left-hand side of Eq. 46 from the fact that a flux $F(r)$ is radiated from both sides of the disk.

The third and final conservation law is that of conservation of energy. By performing the same type of manipulations to this equation as was done for Eq. 44 we obtain,

$$0 = \nabla_\mu E^\mu \tag{47}$$

where

$$E^\mu = -T^{\mu\nu}\left(\frac{\partial}{\partial t}\right)_\nu , \tag{48}$$

$$\left[\dot{M}_0\tilde{E} + 2\pi\sqrt{-\mathbf{g}}W_t{}^r\right]_{,r} = 4\pi\sqrt{-\mathbf{g}}F(r)\tilde{E} . \tag{49}$$

Making use of the orthogonality $u^\mu t_\mu{}^\nu = 0$ which implies that $u^\mu W_\mu{}^\nu = 0 \;\Rightarrow\; W_t{}^r + \Omega W_\phi{}^r = 0$, this equation may be rewritten in terms of $W_\phi{}^r$ as was the case with the angular momentum equation.

$$\left[\dot{M}_0\tilde{E} - 2\pi\sqrt{-\mathbf{g}}W_\phi{}^r\Omega\right]_{,r} = 4\pi\sqrt{-\mathbf{g}}F(r)\tilde{E} \tag{50}$$

From Eqs. 9.18, 9.19 and 9.20 it can be seen that the energy, angular momentum and angular velocity satisfy the following energy angular momentum relation,

$$\tilde{E}_{,r} = \Omega\tilde{L}_{,r} . \tag{51}$$

Eqs. 46 and 50 may now be integrated to find the radial dependence of the flux. This is done by multiplying Eq. 46 by Ω and subtracting the result from Eq. 50 to obtain an expression for the torque;

$$W_\phi{}^r = 2F(r)\frac{\Omega\tilde{L} - \tilde{E}}{\Omega_{,r}} . \tag{52}$$

Substituting this back into Eq. 46 and rearranging and using the energy angular momentum relation in Eq. 51 gives a differential equation for $F(r)$,

$$\left[4\pi\sqrt{-\mathbf{g}}\frac{\left(\tilde{E} - \Omega\tilde{L}\right)^2}{\Omega_{,r}}F(r)\right]_{,r} = \dot{M}_0\left(\tilde{E} - \Omega\tilde{L}\right)\tilde{L}_{,r} , \tag{53}$$

which may be readily integrated. To fix the constant of integration we use the zero torque boundary condition at the inner edge of the disk, $F(r_{\rm isco}) = 0$. Therefore we have an expression for the radial flux from the disk,

$$F(r) = \frac{-\dot{M}_0\Omega_{,r}}{4\pi\sqrt{-\mathbf{g}}\left(\tilde{E} - \Omega\tilde{L}\right)^2}\int_{r_{\rm isco}}^{r}\left(\tilde{E} - \Omega\tilde{L}\right)L_{,r}{\rm d}r . \tag{54}$$

For completeness we also present the final radial structure expression derived in Page and Thorne (1974), the expression for the torque per unit circumference as a function of radius, W^r_ϕ,

$$W^r_\phi(r) = \frac{-\dot{M}_0 \Omega_{,r}}{2\pi\sqrt{-\mathbf{g}}\left(\tilde{E} - \Omega\tilde{L}\right)^2} \frac{\tilde{E} - \Omega\tilde{L}}{-\Omega_{,r}} \int_{r_{\rm isco}}^{r} \left(\tilde{E} - \Omega\tilde{L}\right) L_{,r} \mathrm{d}r \,. \tag{55}$$

G The pericentre Direction

In Sec. 10.1 two angles were defined relating to the direction of the pericentre; γ measured the angle between pericentre and $\hat{x}$ and $\tilde{\gamma}$ measured the angle between pericentre and $\hat{L}\times\hat{S}$. The angle $\tilde{\gamma}$ has the advantage of being a purely intrinsic quantity to the gravitational wave source, however the angle γ, with mixed intrinsic and extrinsic qualities, made the expressions for the amplitude coefficients A_n^+ and $A_n^\times$ is Eqs. 10.19 and 10.20 simpler.

In this appendix we present the relation between the angles, as derived by Barack and Cutler (2004), necessary in order to use the results of Sec. 10.1 to calculate and EMRI waveform. The angle β is defined as the difference between the two different γ's,

$$\gamma = \tilde{\gamma} + \beta\,. \tag{56}$$

With this definition β is the angle between $\hat{x}$ and $\hat{L}\times\hat{S}$;

$$\sin\beta = \frac{\cos\lambda\hat{L}\cdot\hat{n} - \hat{S}\cdot\hat{n}}{\sin\lambda\left[1-\left(\hat{L}\cdot\hat{n}\right)^2\right]^{1/2}} \tag{57}$$

$$\cos\beta = \frac{\hat{n}\cdot\left(\hat{S}\times\hat{L}\right)}{\sin\lambda\left[1-\left(\hat{L}\cdot\hat{n}\right)^2\right]^{1/2}} \tag{58}$$

In order to evaluate β it is necessary to use,

$$\begin{aligned}\hat{S}\cdot\hat{n} &= \cos\theta_S\cos\theta_K \\ &\quad + \sin\theta_S\sin\theta_K\cos\left(\phi_S-\phi_K\right)\,,\end{aligned} \tag{59}$$

$$\begin{aligned}\hat{n}\cdot\left(\hat{S}\times\hat{L}\right) &= \sin\theta_S\sin\left(\phi_K-\phi_S\right)\sin\lambda\cos\alpha \\ &\quad + \frac{\hat{S}\cdot\hat{N}\cos\theta_K-\cos\theta_S}{\sin\theta_K}\sin\lambda\sin\alpha\,.\end{aligned} \tag{60}$$

To an excellent approximation the quantity $\hat{S}\cdot\hat{n}$ may be considered constant throughout the inspiral.

H LISA Antenna Response Functions and the polarisation angle

The classic LISA mission has three arms (6 laser links), which produce two independent outputs equivilent to two crossed Michelson interferometes. Some more recent mission proposals have two arms (4 laser links) in which case only one of the outputs is present. In this appendix we discuss the response of a three arm detector with output chanels labelled by $\alpha = I, II$ (the response of a two arm detector is the same for the I chanel, and zero for the II second chanel).

We use the notation of Cutler (1998); let l_1^i, l_2^i, and $l_3^i = l_2^i - l_1^i$ be unit vectors along the three detector arms all with equal length L. We choose to work in a coordinate system attached to the rotating detector, specfically $l_1^i = \cos(\pi/12)x^i + \sin(\pi/12)y^i$, and $l_2^i = \cos(5\pi/12)x^i + \sin(5\pi/12)y^i$. The strain in the first output is given by

$$\begin{aligned} h_I(t) &= \left(\delta L_1(t) - \delta L_2(t)\right)/L & (61)\\ &= \frac{1}{2}h_{ij}(t)(l_1^i l_1^j - l_2^i l_2^j) & (62)\end{aligned}$$

The gravitational wave strain is written in terms of the polarisation basis tensors introduced in Sec. 10.1,

$$h_{ij}(t) = A_+(t)H_{ij}^+ + A_\times(t)H_{ij}^\times\,. \tag{63}$$

Putting these expressions into Eq. 62 gives

$$h_I(t) = \frac{\sqrt{3}}{2}\left(A_+(t)F_I^+(\theta,\phi,\psi) + A_+(t)F_I^+(\theta,\phi,\psi)\right)\,, \tag{64}$$

where the response functions have been defined as,

$$F_I^+ = \frac{1}{2}\left(1+\cos^2\theta\right)\cos(2\phi)\cos(2\psi) - \cos\theta\sin(2\phi)\sin(2\psi) \tag{65}$$

$$F_I^\times = \frac{1}{2}\left(1+\cos^2\theta\right)\cos(2\phi)\sin(2\psi) + \cos\theta\sin(2\phi)\cos(2\psi) \tag{66}$$

The factor of $\sqrt{3}/2$ is included to make the response functions equal to those for ground based detectors with arms at 90^0, see Sec. 2.4.1. The angles θ, ϕ are the spherical polar angles describe the direction of the incoming gravitational wave relative to the detector; θ is the angle between the wave vector and $l_1 \times l_2$, while ϕ is the angle between the component of the wave vector perpendicular to $l_1 \times l_2$ and $\hat{x}$. The angle ψ is the polarisation angle of the gravitational wave and is defined as

$$\psi = \frac{z^i q_i}{z^j p_j}\,. \tag{67}$$

The response functions of the second chanel, F^{+}_{II} and $F^{\times}_{II}$, may be obtained in a similar manner to that done for the first chanel, or (more simply) by noting that the second detector is indentical to the first with a rotation in the ϕ coordinate;

$$\begin{aligned} F^{+}_{II}(\theta,\phi,\psi) =& F^{+}_{I}(\theta,\phi-\pi/4,\psi) \\ =& \frac{1}{2}\left(1+\cos^2\theta\right)\sin(2\phi)\cos(2\psi)+\cos\theta\cos(2\phi)\sin(2\psi)\ , \qquad (68) \\ F^{\times}_{II}(\theta,\phi,\psi) =& F^{\times}_{I}(\theta,\phi-\pi/4,\psi) \\ =& \frac{1}{2}\left(1+\cos^2\theta\right)\sin(2\phi)\sin(2\psi)-\cos\theta\cos(2\phi)\cos(2\psi)\ . \qquad (69) \end{aligned}$$

In order to make use of these equations in the Sec. 10.1 it is necessary to relate the angles θ, ϕ, and ψ to the angles already introduced. This is simply a geometric relation and the relevent expressions are derived in Barack and Cutler (2007). The angles θ and ϕ are related to the sources sky position via

$$\begin{aligned} \cos\theta(t) =& \frac{1}{2}\cos\theta_S-\frac{\sqrt{3}}{2}\sin\theta_S\cos\left[\bar{\phi}_0+2\pi\frac{t}{T}-\phi_S\right]\ , \qquad (70) \\ \phi(t) =& \bar{\alpha}_0+2\pi\frac{t}{T}+\tan^{-1}\left[\frac{\sqrt{3}\cos\theta_S+\sin\theta_S\cos\left[\phi_0+2\pi\frac{t}{T}-\phi_S\right]}{2\sin\theta_S\sin\left[\phi_0+2\pi\frac{t}{T}-\phi_S\right]}\right]\ , \end{aligned}$$

where the angle $\bar{\phi}_0$ specifies the intial orbital phase of the detector and the angle $\bar{\alpha}_0$ specifies the inital rotation phase. The period, T, of the detector motion is one year. The polarisation angle, ψ, is related to the sky position and the orientation of the binary via

$$\begin{aligned} \tan\psi =& \left\{\frac{1}{2}\cos\theta_L-\frac{\sqrt{3}}{2}\sin\theta_L\cos\left[\bar{\phi}_0+2\pi\left(\frac{t}{T}\right)-\phi_L\right]\right. \\ & \left.-\cos\theta(t)\Big[\cos\theta_L\cos\theta_S+\sin\theta_L\sin\theta_S\cos(\phi_L-\phi_S)\Big]\right\}\Big/ \\ & \left[\frac{1}{2}\sin\theta_L\sin\theta_S\sin(\phi_L-\phi_S)-\frac{\sqrt{3}}{2}\cos\left(\bar{\phi}_0+2\pi\frac{t}{T}\right)\right. \\ & \times(\cos\theta_L\sin\theta_S\sin\phi_S-\cos\theta_S\sin\theta_L\sin\phi_L) \\ & \left.-\frac{\sqrt{3}}{2}\sin\left(\bar{\phi}_0+2\pi\frac{t}{T}\right)(\cos\theta_S\sin\theta_L\cos\phi_L-\cos\theta_L\sin\theta_S\cos\phi_S)\right]\ . \qquad (71) \end{aligned}$$

The angles $\{\theta_L,\phi_L\}$ measure the direction of the CO's orbital angular momentum, $\hat{L}(t)$ in an ecliptic based coordinate system. These angles are not constant as the orbital angular momentum processes about the SMBH's spin axis. $\{\theta_L,\phi_L\}$ can be related the other angles involved as follows.

$$\cos\dot{\theta}_L(t)=\cos\theta_K\cos\lambda+\sin\theta_K\sin\lambda\cos\alpha(t) \qquad (72)$$

$$\begin{aligned}\sin\theta_L(t)\cos\phi_L(t) = &\sin\theta_K\cos\phi_K\cos\lambda \\ &- \cos\phi_K\cos\theta_K\sin\lambda\cos\alpha(t) \\ &+ \sin\phi_K\sin\lambda\sin\alpha(t) \end{aligned} \tag{73}$$

$$\begin{aligned}\sin\theta_L(t)\sin\phi_L(t) = &\sin\theta_K\sin\phi_K\cos\lambda \\ &- \sin\phi_K\cos\theta_K\sin\lambda\cos\alpha(t) \\ &- \cos\phi_K\sin\lambda\sin\alpha(t) \end{aligned} \tag{74}$$

I Perturbed evolution equation

As described in Sec. 10.2 the perturbations to the metric translate into perturbations to the post Newtonian evolution equations for the EMRI orbital quantities. The additive corrections to Equations (10.22) to (10.26) for the $\mathcal{B}_N$ metrics, where N is in the range 3 to 5, are given by:

$$M\delta\dot{e}_{\mathcal{B}_N} = \frac{-16}{5}\eta\frac{(2\pi M\nu)^{2N/3+2}}{(1-e^2)^{N+3/2}}g_{e,N}(e)\left(\gamma_{1,N}+2\gamma_{4,N}\right)\ , \tag{75}$$

$$2\pi M^2\delta\dot{\nu}_{\mathcal{B}_N} = \frac{16}{5}\eta\frac{(2\pi M\nu)^{2N/3+2}}{(1-e^2)^{N+5/2}}g_{\nu,N}(e)\left(\gamma_{1,N}+2\gamma_{4,N}\right)\ , \tag{76}$$

$$M\delta\dot{\tilde{\gamma}}_{\mathcal{B}_N} = \frac{(2\pi M\nu)^{(2N+1)/3}}{(1-e^2)^{N-1}}g_{\gamma,N}(e)\left(\gamma_{1,N}+2\gamma_{4,N}\right)\ , \tag{77}$$

$$M\delta\dot{\alpha}_{\mathcal{B}_N} = -S\frac{(2\pi M\nu)^{2(N+1)/3}}{(1-e^2)^{N-1/2}}g_{\alpha,N}(e)\left(\gamma_{1,N}+2\gamma_{4,N}\right)\ . \tag{78}$$

The eccentricity dependent factors are defined as follows:

$$g_{e,3}(e) = \frac{725}{12}e+\frac{383}{8}e^3+\frac{143}{48}e^5\,, \tag{79}$$

$$g_{\nu,3}(e) = 42+\frac{407}{2}e^2+\frac{345}{4}e^4+\frac{33}{16}e^6\,, \tag{80}$$

$$g_{\gamma,3}(e) = \frac{3}{2}\,, \tag{81}$$

$$g_{\alpha,3}(e) = \frac{3}{4}+\frac{1}{4}e^2\,, \tag{82}$$

$$\begin{aligned} g_{e,4}(e) = {} & \frac{316}{3}e+\frac{12713}{96}e^3+\frac{2467}{64}e^5+\frac{3}{8}e^7 \\ & +(1-e^2)^{1/2}e\left(\frac{95}{18}-\frac{305}{96}e^2-\frac{605}{288}e^4\right)\,, \end{aligned} \tag{83}$$

$$\begin{aligned} g_{\nu,4}(e) = {} & 69+\frac{3191}{8}e^2+\frac{5035}{16}e^4+\frac{3039}{64}e^6 \\ & -\frac{1}{48}\left(1-e^2\right)^{1/2}\left(48-358e^2+147e^4+163e^6\right)\,, \end{aligned} \tag{84}$$

$$g_{\gamma,4}(e) = 3+\frac{3}{4}e^2\,, \tag{85}$$

$$g_{\alpha,4}(e) = 1+e^2\,, \tag{86}$$

$$\begin{aligned} g_{e,5}(e) = {} & 155e+\frac{31565}{96}e^3+\frac{30749}{192}e^5+\frac{5305}{768}e^7 \\ & +\frac{5}{96}e\left(1-e^2\right)^{1/2}\left(304-183e^2-121e^4\right)\,, \end{aligned} \tag{87}$$

$$g_{\nu,5}(e) = 99 + \frac{5601}{8}e^2 + 905e^4 + \frac{15617}{64}e^6 + \frac{1137}{256}e^8 - \frac{1}{16}\left(1-e^2\right)^{1/2}\left(144 - 466e^2 + 75e^4 + 247e^6\right) \tag{88}$$

$$g_{\gamma,5}(e) = 5 + \frac{15}{4}e^2\,, \tag{89}$$

$$g_{\alpha,5}(e) = \frac{5}{4} + \frac{5}{2}e^2 + \frac{1}{4}e^4\,. \tag{90}$$

J The LISA noise curve

There are three significant sources of noise for a LISA-like mission which we consider here: instrumental noise (which is itself comprised of several sources), confusion noise from galactic binaries and confusion noise from extragalactic binaries (Hughes, 2002).

For the instrumental noise the simple analytic fit to the noise budget in the LISA pre-phase A report (LISA study team, 1998) was used.

$$S_h^{\rm inst}(f) = \left[9.18\times10^{-52}\left(\frac{f}{1{\rm Hz}}\right)^{-4} + 1.59\times10^{-41} + 9.18\times10^{-38}\left(\frac{f}{1{\rm Hz}}\right)^{2}\right]{\rm Hz}^{-1} \quad (91)$$

For the purpose of analysing the confusion noise it is convenient to use the energy density in gravitational waves (see Sec. 2.1.3) as a fraction of the energy density needed to close the universe (assuming that it is matter dominated) per logarithmic frequency interval

$$\Omega_{\rm GW} \equiv \frac{1}{\rho_c}\frac{d\rho_{\rm GW}}{d\,(\log f)}\,, \quad (92)$$

where

$$\rho_c = \frac{3c^2H_0^2}{8\pi G}\,. \quad (93)$$

Assuming that the background is isotropic, stationary and unpolarised the energy density is related to the noise spectral density, $S_h(f)$, as described in Sec. 2.1.4 (where we have now included the detector the response function, see Appendix H, which modifies Eq. 2.30 only by a geometric factor)

$$S_h(f) = \frac{3\rho_c}{5\pi}f^{-3}\Omega_{\rm GW}(f)\,. \quad (94)$$

For the extragalactic background Farmer and Phinney (2003) estimate that for a value of the Hubble constant, $H_0 = 70{\rm km\ s^{-1}Mpc^{-1}}$ and at frequencies near 1 mHz the energy density in gravitational waves due to white dwarf white dwarf mergers is given by

$$\Omega_{\rm GW}(f) = 3.6\times10^{-12}\left(\frac{f}{10^{-3}{\rm Hz}}\right)^{2/3}\,, \quad (95)$$

and hence

$$S_h^{\text{ex-gal}}(f) = 4.2\times10^{-47}\left(\frac{f}{1{\rm Hz}}\right)^{-7/3}{\rm Hz}^{-1}\,. \quad (96)$$

For the galactic confusion noise, Nelemans et al. (2001) estimate a value to be approximately 50 times larger,

$$S_h^{\rm gal}(f) = 2.1\times10^{-45}\left(\frac{f}{1{\rm Hz}}\right)^{-7/3}{\rm Hz}^{-1}\,. \quad (97)$$

For frequencies in the range $f \sim (10^{-4} - 10^{-2})\,\mathrm{Hz}$ the galactic confusion noise is larger than the instrumental noise. However, for higher frequencies the galactic confusion is sufficiently small that it can be "fit out" from the data, the resulting noise spectra is estimated as,

$$S_h^{\mathrm{inst+gal}}(f) = \min\left\{ \frac{S_h^{\mathrm{inst}}(f)}{\exp\left(\frac{-\kappa}{T}\frac{dN}{df}\right)}, S_h^{\mathrm{inst}}(f) + S_h^{\mathrm{gal}}(f) \right\} \tag{98}$$

where $\mathrm{d}N/\mathrm{d}f$ is the number density of galactic white dwarf binaries per unit frequency, T is the mission lifetime and κ is the average number of frequency bins lost when each galactic binary is "fitted out". It is estimated that $\kappa T = 1.5/\mathrm{year}$.

The total LISA noise curve is then given by

$$S_h^{\mathrm{total}}(f) = S_h^{\mathrm{inst+gal}}(f) + S_h^{\mathrm{ex\text{-}gal}}(f)\ . \tag{99}$$

The relative importance of these noise sources of noise as a function of frequency is shown in Fig. 2. The quantity plotted is the power spectral density (see Sec. 2.1.2)

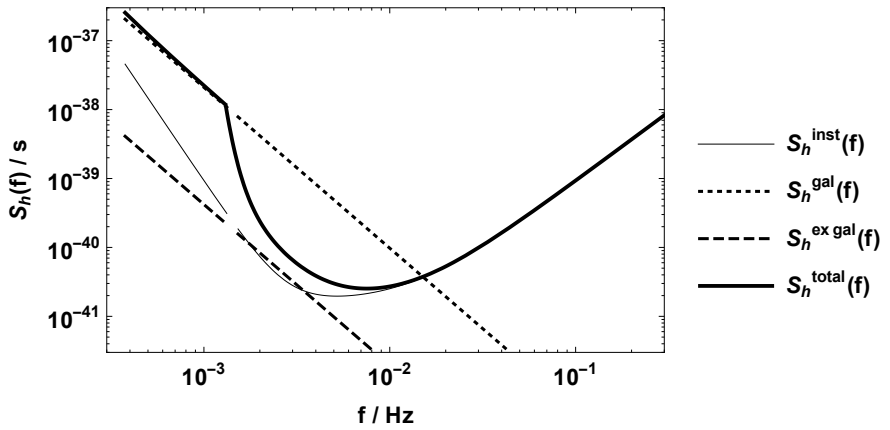


Figure 2: The contributions to the total noise as a function of frequency from the three sources: galactic binaries, extragalactic binaries and instrumental noise.

Bibliography

Aasi J et al. 2013*a*. Einstein@Home all-sky search for periodic gravitational waves in LIGO S5 data. *Physcial Review D* **87**(4), 042001. arXiv:1207.7176.

Aasi J et al. 2013*b*. Parameter estimation for compact binary coalescence signals with the firs t generation gravitational-wave detector network. *Physical Review D* **88**(6), 062001. arXiv:1304.1775.

Aasi J et al. 2014*a*. Constraints on Cosmic Strings from the LIGO-Virgo Gravitational-Wave Detectors. *Physical Review Letters* **112**(13), 131101. arXiv:1310.2384.

Aasi J et al. 2014*b*. Gravitational waves from known pulsars: results from the initial detector era. *The Astrophysical Journal* **785**(2), 119. arXiv:1309.4027.

Aasi J et al. 2014*c*. The NINJA-2 project: detecting and characterizing gravitational waveforms modelled using numerical binary black hole simulations. *Classical and Quantum Gravity* **31**(11), 115004. arXiv:1401.0939.

Aasi J et al. 2015*a*. Advanced LIGO. *Classical and Quantum Gravity* **32**(7), 074001. arXiv:1411.4547.

Aasi J et al. 2015*b*. Characterization of the LIGO detectors during their sixth science run. *Classical and Quantum Gravity* **32**(11), 115012. arXiv:1410.7764.

Abadie J et al. 2010*a*. All-sky search for gravitational-wave bursts in the first joint LIGO-GEO-Virgo run. *Physical Review D* **81**(10), 102001. arXiv:1002.1036.

Abadie J et al. 2010*b*. TOPICAL REVIEW: Predictions for the rates of compact binary coalescences observable by ground-based gravitational-wave detectors. *Classical and Quantum Gravity* **27**(17), 173001. arXiv:1003.2480.

Abadie J et al. 2012*a*. All-sky search for gravitational-wave bursts in the second joint LIGO-Virgo run. *Physical Review D* **85**(12), 122007. arXiv:1202.2788.

Abadie J et al. 2012*b*. Search for Gravitational Waves Associated with Gamma-Ray Bursts during LIGO Science Run 6 and Virgo Science Runs 2 and 3. *The Astrophysical Journal* **760**(1), 12. arXiv:1205.2216.

Abadie J et al. 2012*c*. Search for gravitational waves from low mass compact binary coalescence in LIGO's sixth science run and Virgo's science runs 2 and 3. *Physcial Review D* **85**(8), 082002. arXiv:1111.7314.

Abbott B P et al. 2009. LIGO: the Laser Interferometer Gravitational-Wave Observatory. *Reports on Progress in Physics* **72**(7), 076901. arXiv:0711.3041.

Abbott B P et al. 2013. Prospects for Observing and Localizing Gravitational-Wave Transients with Advanced LIGO and Advanced Virgo. *ArXiv e-prints* . arXiv:1304.0670.

Abbott B et al. 2007. Searches for periodic gravitational waves from unknown isolated sources and Scorpius X-1: Results from the second LIGO science run. *Physical Review D* **76**(8), 082001. arXiv:gr-qc/0605028.

Accadia T, Acernese F, Agathos M et al. 2012 Advanced virgo technical design report Virgo Technical Report VIR-0128A-12. https://tds.ego-gw.it/ql/?c=8940

Accadia T et al. 2012. Virgo: a laser interferometer to detect gravitational waves. *Journal of Instrumentation* **7**(03), 03012.

Acernese F et al. 2009 Advanced virgo baseline design Virgo Technical Report VIR-0027A-09. http://tds.ego-gw.it/itf/tds/file.php?callFile=VIR-0027A-09.pdf

Adler R J 1981 *The Geometry of Random Fields* Wiley Series in Probability and Mathematical Statistics Wiley Chichester.

Agathos M et al. 2014. TIGER: A data analysis pipeline for testing the strong-field dynamics of general relativity with gravitational wave signals from coalescing compact binaries. *Physical Review D* **89**(8), 082001. arXiv:1311.0420.

Ali-Haïmoud Y and Chen Y 2011. Slowly rotating stars and black holes in dynamical Chern-Simons gravity. *Physical Review D* **84**(12), 124033. arXiv:1110.5329.

Allen B and Romano J D 1999. Detecting a stochastic background of gravitational radiation: Signal processing strategies and sensitivities. *Physical Review D* **59**(10), 102001. arXiv:gr-qc/9710117.

Amaro-Seoane P et al. 2007. TOPICAL REVIEW: Intermediate and extreme mass-ratio inspirals - astrophysics, science applications and detection using LISA. *Classical and Quantum Gravity* **24**, 113. arXiv:astro-ph/0703495.

Amaro-Seoane P et al. 2013. eLISA: Astrophysics and cosmology in the millihertz regime. *GW Notes* **6**, 4–110. arXiv:1201.3621.

Ando M 2002. Current status of TAMA. *Classical and Quantum Gravity* **19**, 1409–1419.

Anholm M, Ballmer S, Creighton J D E, Price L R and Siemens X 2009. Optimal strategies for gravitational wave stochastic background searches in pulsar timing data. *Physcial Review D* **79**(8), 084030. arXiv:0809.0701.

Astone P et al. 2010. IGEC2: A 17-month search for gravitational wave bursts in 2005-2007. *Physical Review D* **82**(2), 022003. arXiv:1002.3515.

Baade W and Zwicky F 1934*a*. On Super-novae. *Proceedings of the National Academy of Science* **20**, 254–259.

Baade W and Zwicky F 1934*b*. Remarks on Super-Novae and Cosmic Rays. *Physical Review* **46**, 76–77.

Babak S, Fang H, Gair J R, Glampedakis K and Hughes S A 2007. "Kludge" gravitational waveforms for a test-body orbiting a Kerr black hole. *Physical Review D* **75**(2), 024005. arXiv:gr-qc/0607007.

Babak S et al. 2010. The Mock LISA Data Challenges: from challenge 3 to challenge 4. *Classical and Quantum Gravity* **27**(8), 084009. arXiv:0912.0548.

Baker J G, Centrella J, Choi D I, Koppitz M and van Meter J 2006. Gravitational-Wave Extraction from an Inspiraling Configuration of Merging Black Holes. *Physical Review Letters* **96**(11), 111102. arXiv:gr-qc/0511103.

Barack L and Cutler C 2004. Lisa capture sources: Approximate waveforms, signal-to-noise ratios, and parameter estimation accuracy. *Physical Review D* **69**, 082005.

Barack L and Cutler C 2007. Using LISA extreme-mass-ratio inspiral sources to test off-Kerr deviations in the geometry of massive black holes. *Physical Review D* **75**(4), 042003. arXiv:gr-qc/0612029.

Bardeen J M, Press W H and Teukolsky S A 1972. Rotating Black Holes: Locally Nonrotating Frames, Energy Extraction, and Scalar Synchrotron Radiation. *Astrophysical Journal* **178**, 347–370.

Barker B M and O'Connell R F 1975. Gravitational two-body problem with arbitrary masses, spins, and quadrupole moments. *Physical Review D* **12**, 329–335.

Barsotti L and Fritschel P 2012 Early aligo configurations: example scenarios toward design sensitivity Technical Report LIGO-T1200307-v4. https://dcc.ligo.org/LIGO-T1200307/public

Bender P L, Begelman M C and Gair J R 2013. Possible LISA follow-on mission scientific objectives. *Classical and Quantum Gravity* **30**(16), 165017.

Benenti S and Francaviglia M 1979. Remarks on certain separability structures and their applications to general relativity. *General Relativity and Gravitation* **10**, 79–92.

Berry C P L and Gair J R 2013. Observing the Galaxy's massive black hole with gravitational wave bursts. *Monthly Notices of the Royal Astronomical Society* **429**(1), 589–612. arXiv:1210.2778.

Berry C P L et al. 2015. Parameter Estimation for Binary Neutron-star Coalescences with Realistic Noise during the Advanced LIGO Era. *Astrophysical Journal* **804**, 114. arXiv:1411.6934.

Berti E, Gair J and Sesana A 2011. Graviton mass bounds from space-based gravitational-wave observations of massive black hole populations. *Physical Review D* **84**(10), 101501. arXiv:1107.3528.

BICEP2 Collaboration 2014. Detection of B-Mode Polarization at Degree Angular Scales by BICEP2. *Physical Review Letters* **112**(24), 241101. arXiv:1403.3985.

BICEP2/Keck and Planck Collaborations 2015. Joint Analysis of BICEP2/Keck Array and Planck Data. *Physical Review Letters* **114**(10), 101301. arXiv:1502.00612.

Binétruy P, Bohé A, Caprini C and Dufaux J F 2012. Cosmological backgrounds of gravitational waves and eLISA/NGO: phase transitions, cosmic strings and other sources. *Journal of Cosmology and Astroparticle Physics* **2012**(06), 027. arXiv:1201.0983.

Blanchet L 2014. Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries. *Living Reviews in Relativity* **17**, 2. arXiv:1310.1528.

Brans C and Dicke R H 1961. Mach's principle and a relativistic theory of gravitation. *Physical Review* **124**, 925–935.

Brumberg V A 1991 *Essential relativistic celestial mechanics.*

Buonanno A and Damour T 1999. Effective one-body approach to general relativistic two-body dynamics. *Physical Review D* **59**(8), 084006. arXiv:gr-qc/9811091.

Buonanno A and Damour T 2000. Transition from inspiral to plunge in binary black hole coalescences. *Physical Review D* **62**(6), 064015. arXiv:gr-qc/0001013.

Buonanno A, Iyer B R, Ochsner E, Pan Y and Sathyaprakash B S 2009. Comparison of post-newtonian templates for compact binary inspiral signals in gravitational-wave detectors. *Physical Review D* **80**, 084043.

Burke W L 1975. Large-Scale Random Gravitational Waves. *The Astrophysical Journal* **196**, 329–334.

Campanelli M, Lousto C O, Marronetti P and Zlochower Y 2006. Accurate Evolutions of Orbiting Black-Hole Binaries without Excision. *Physical Review Letters* **96**(11), 111101. arXiv:gr-qc/0511048.

Canitrot P 2001. Systematic errors for matched filtering of gravitational waves from inspiraling compact binaries. *Physical Review D* **63**(8), 082005.

Canizares P, Gair J R and Sopuerta C F 2012. Testing Chern-Simons modified gravity with gravitational-wave detections of extreme-mass-ratio binaries. *Physical Review D* **86**(4), 044010. arXiv:1205.1253.

Canizares P et al. 2015. Accelerated Gravitational Wave Parameter Estimation with Reduced Order Modeling. *Physical Review Letters* **114**(7), 071104. arXiv:1404.6284.

Cannon K C 2008. A Bayesian coincidence test for noise rejection in a gravitational-wave burst search. *Classical and Quantum Gravity* **25**(10), 105024.

Carroll S 2004 *Spacetime and Geometry: An Introduction to General Relativity* Addison Wesley. `https://books.google.co.uk/books?id=1SKFQgAACAAJ`

Carter B 1968. Global structure of the kerr family of gravitational fields. *Physical Review* **174**, 1559–1571.

Chamel N and Haensel P 2008. Physics of Neutron Star Crusts. *Living Reviews in Relativity* **11**(10). `arXiv:0812.3955`.

Chandrasekhar S 2002 *The mathematical theory of black holes* Oxford classic texts in the physical sciences Oxford Univ. Press Oxford.

Chatterji S, Blackburn L, Martin G and Katsavounidis E 2004. Multiresolution techniques for the detection of gravitational-wave bursts. *Classical and Quantum Gravity* **21**(20), S1809. `arXiv:gr-qc/0412119`.

Collins N A and Hughes S A 2004. Towards a formalism for mapping the spacetimes of massive compact objects: Bumpy black holes and their orbits. *Physical Review D* **69**(12), 124022. `arXiv:gr-qc/0402063`.

Cornish N J and Littenberg T B 2015. Bayeswave: Bayesian inference for gravitational wave bursts and instrument glitches. *Classical and Quantum Gravity* **32**(13), 135012. `arXiv:1410.3835`.

Cornish N J and Rubbo L J 2003. Publisher's Note: LISA response function [Phys. Rev. D 67, 022001 (2003)]. *Physical Review D* **67**(2), 029905. `arXiv:gr-qc/0209011`.

Coughlin M and Harms J 2014*a*. Constraining the gravitational-wave energy density of the Universe in the range 0.1 Hz to 1 Hz using the Apollo Seismic Array. *Physical Review D* **90**(10), 102001. `arXiv:1409.4680`.

Coughlin M and Harms J 2014*b*. Upper limit on a stochastic background of gravitational waves from seismic measurements in the range 0.05–1 hz. *Physical Review Letters* **112**, 101102.

Cunningham C T 1975. The effects of redshifts and focusing on the spectrum of an accretion disk around a Kerr black hole. *Astrophysical Journal* **202**, 788–802.

Cutler C 1998. Angular resolution of the LISA gravitational wave detector. *Physical Review D* **57**, 7089–7102. arXiv:gr-qc/9703068.

Cutler C and Flanagan É E 1994*a*. Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral waveform? *Physical Review D* **49**(6), 2658–2697. arXiv:gr-qc/9402014.

Cutler C and Flanagan É E 1994*b*. Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral waveform\? *Physical Review D* **49**, 2658–2697. arXiv:gr-qc/9402014.

Cutler C and Vallisneri M 2007. Lisa detections of massive black hole inspirals: Parameter extraction errors due to inaccurate template waveforms. *Physical Review D* **76**, 104018.

Damour T, Iyer B R and Sathyaprakash B S 2001. Comparison of search templates for gravitational waves from binary inspiral. *Physical Review D* **63**(4), 044023. arXiv:gr-qc/0010009.

Damour T, Iyer B R and Sathyaprakash B S 2002. Comparison of search templates for gravitational waves from binary inspiral: 3.5PN update. *Physical Review D* **66**(2), 027502. arXiv:gr-qc/0207021.

Damour T and Vilenkin A 2005. Gravitational radiation from cosmic (super)strings: Bursts, stochastic background, and observational windows. *Physical Review D* **71**(6), 063510. arXiv:hep-th/0410222.

de Waard A et al. 2006. Preparing for science run 1 of MiniGRAIL. *Classical and Quantum Gravity* **23**, S79–S84.

Detweiler S 1979. Pulsar timing measurements and the search for gravitationa l waves. *The Astrophysical Journal* **234**, 1100–1104.

Dimmelmeier H, Font J A and Müller E 2002. Relativistic simulations of rotational core collapse II. Collapse dynamics and gravitational radiation. *Astronomy and Astrophysics* **393**, 523–542. arXiv:astro-ph/0204289.

Dimopoulos S, Graham P W, Hogan J M, Kasevich M A and Rajendran S 2009. Gravitational wave detection with atom interferometry. *Physics Letters B* **678**, 37–40. arXiv:0712.1250.

Dyson F W, Eddington A S and Davidson C 1920. A Determination of the Deflection of Light by the Sun's Gravitational Field, from Observations Made at the Total Eclipse of May 29, 1919. *Royal Society of London Philosophical Transactions Series A* **220**, 291–333.

Edwards R T, Hobbs G B and Manchester R N 2006. TEMPO2, a new pulsar timing package - II. The timing model and precision estimates. *Monthly Notices of the Royal Astronomical Society* **372**, 1549–1574. arXiv:arXiv:astro-ph/0607664.

Einstein A 1905. Zur Elektrodynamik bewegter Körper. *Annalen der Physik* **322**, 891–921.

Einstein A 1915*a*. Die Feldgleichungen der Gravitation. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 844-847.* .

Einstein A 1915*b*. Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 831-839.* pp. 831–839.

Einstein A 1916. Näherungsweise Integration der Feldgleichungen der Gravitation. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 688-696.* .

Einstein A 1918. Über Gravitationswellen. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 154-167.* .

Einstein A and Rosen N 1937. On gravitational waves. *Journal of the Franklin Institute* **223**, 43–54.

eLISA Consortium T 2013. The Gravitational Universe. *ArXiv e-prints* . arXiv:1305.5720.

Estabrook F B and Wahlquist H D 1975. Response of Doppler spacecraft tracking to gravitational r adiation. *General Relativity and Gravitation* **6**, 439–447.

Farmer A J and Phinney E S 2003. The gravitational wave background from cosmological compact binaries. *Monthly Notices of the Royal Astronomical Society* **346**, 1197–1214. arXiv:astro-ph/0304393.

Feroz F and Hobson M P 2008. Multimodal nested sampling: an efficient and robust altern ative to Markov Chain Monte Carlo methods for astronomical data analyses . *Monthly Notices of the Royal Astronomical Society* **384**, 449–463. arXiv:0704.3704.

Feroz F, Hobson M P and Bridges M 2009. MULTINEST: an efficient and robust Bayesian inference tool for cosmology and particle physics. *Monthly Notices of the Royal Astronomical Society* **398**, 1601–1614. arXiv:0809.3437.

Feroz F, Hobson M P, Cameron E and Pettitt A N 2013. Importance Nested Sampling and the MultiNest Algorithm. *ArXiv e-prints* . arXiv:1306.2144.

Ferrarese L and Ford H 2005. Supermassive Black Holes in Galactic Nuclei: Past, Present and Future Research. *Space Science Reviews* **116**(3–4), 523–624. arXiv:astro-ph/0411247.

Finn L S 1992. Detection, measurement, and gravitational radiation. *Physical Review D* **46**(12), 5236–5249. arXiv:gr-qc/9209010.

Finn L S and Chernoff D F 1993. Observing binary inspiral in gravitational radiation: One interferometer. *Physcial Review D* **47**, 2198–2219. arXiv:gr-qc/9301003.

Finn L S and Thorne K S 2000. Gravitational waves from a compact star in a circular, inspiral orbit, in the equatorial plane of a massive, spinning black hole, as observed by LISA. *Physical Review D* **62**(12), 124021. arXiv:gr-qc/0007074.

Fitz Gerald G 1889. The Ether and the Earth's Atmosphere. *Science* **13**, 390.

Foster R S and Backer D C 1990. Constructing a pulsar timing array. *Astrophysical Journal* **361**, 300–308.

Gair J R and Moore C J 2015. Quantifying and mitigating bias in inference on gravitational wave source populations. *Physical Review D* **91**(12), 124062. arXiv:1504.02767.

Gair J R, Vallisneri M, Larson S L and Baker J G 2013. Testing General Relativity with Low-Frequency, Space-Based Gravitational-Wave Detectors. *Living Reviews in Relativity* **16**, 7. arXiv:1212.5575.

Gair J and Yunes N 2011. Approximate waveforms for extreme-mass-ratio inspirals in modified gravity spacetimes. *Physical Review D* **84**(6), 064016. arXiv:1106.6313.

Ghosh A, Del Pozzo W and Ajith P 2015. Estimating parameters of binary black holes from gravitational-wave observations of their inspiral, merger and ringdown. *ArXiv e-prints* . arXiv:1505.05607.

Glampedakis K and Babak S 2006. Mapping spacetimes with LISA: inspiral of a test body in a 'quasi-Kerr' field. *Classical and Quantum Gravity* **23**, 4167–4188. arXiv:gr-qc/0510057.

Gneiting T 2002. Compactly supported correlation functions. *Journal Multivariate Analysis* **83**(2), 493–508.

Graff P B, Buonanno A and Sathyaprakash B S 2015. Missing Link: Bayesian detection and measurement of intermediate-mass black-hole binaries. *Physical Review D* **92**(2), 022002. arXiv:1504.04766.

Grishchuk L P 2005. Relic gravitational waves and cosmology. *Physics-Uspekhi* **48**(12), 1235–1247. arXiv:gr-qc/0504018.

Grote H 2010. The GEO 600 status. *Classical and Quantum Gravity* **27**(8), 084003.

Hannam M et al. 2014. Simple Model of Complete Precessing Black-Hole-Binary Gravitational Waveforms. *Physical Review Letters* **113**(15), 151101. arXiv:1308.3271.

Hansen R O 1974. Multipole moments of stationary spacetimes. *Journal of Mathematical Physics* **15**, 46–52.

Harko T, Kovács Z and Lobo F S N 2010. Thin accretion disk signatures in dynamical Chern-Simons-modified gravity. *Classical and Quantum Gravity* **27**(10), 105010. arXiv:0909.1267.

Harko T, Kovacs Z and Lobo F S N 2011*a*. Solar System tests of Horava-Lifshitz gravity. *Royal Society of London Proceedings Series A* **467**, 1390–1407. arXiv:0908.2874.

Harko T, Kovács Z and Lobo F S N 2011*b*. Thin accretion disk signatures of slowly rotating black holes in Hořava gravity. *Classical and Quantum Gravity* **28**(16), 165001. arXiv:1009.1958.

Harry G M 2010. Advanced LIGO: the next generation of gravitational wave detectors. *Classical and Quantum Gravity* **27**(8), 084006.

Haskell B, Samuelsson L, Glampedakis K and Andersson N 2008. Modelling magnetically deformed neutron stars. *Monthly Notices of the Royal Astronomical Society* **385**(1), 531–542. arXiv:0705.1780.

Hellings R W and Downs G S 1983. Upper limits on the isotropic gravitational radiation background from pulsar timing analysis. *Astrophyiscal Journal Letters* **265**, L39–L42.

Hild S et al. 2011. Sensitivity studies for third-generation gravitational wave observatories. *Classical and Quantum Gravity* **28**(9), 094013. arXiv:1012.0908.

Hobbs G 2013. The Parkes Pulsar Timing Array. *Classical and Quantum Gravity* **30**(22), 224007. arXiv:1307.2629.

Hobbs G B, Edwards R T and Manchester R N 2006. TEMPO2, a new pulsar-timing package - I. An overview. *Monthly Notices of the Royal Astronomical Society* **369**, 655–672. arXiv:arXiv:astro-ph/0603381.

Hobbs G et al. 2009. TEMPO2: a new pulsar timing package - III. Gravitational w ave simulation. *Monthly Notices of the Royal Astronomical Society* **394**, 1945–1955. arXiv:0901.0592.

Hořava P 2009*a*. Quantum gravity at a Lifshitz point. *Physical Review D* **79**(8), 084008. arXiv:0901.3775.

Hořava P 2009*b*. Spectral Dimension of the Universe in Quantum Gravity at a Lifshitz Point. *Physical Review Letters* **102**(16), 161301. arXiv:0902.3657.

Hughes S A 2000. Evolution of circular, nonequatorial orbits of Kerr black holes due to gravitational-wave emission. *Physical Review D* **61**(8), 084004. arXiv:gr-qc/9910091.

Hughes S A 2002. Untangling the merger history of massive black holes with LISA. *Monthly Notices of the Royal Astronomical Society* **331**, 805–816. arXiv:astro-ph/0108483.

Hui L, McWilliams S T and Yang I S 2013. Binary systems as resonance detectors for gravitational waves. *Physical Review D* **87**, 084009.

Hulse R A 1994. The discovery of the binary pulsar. *Reviews of Modern Physics* **66**, 699–710.

Hulse R A and Taylor H J 1974 *in* 'Bulletin of the American Astronomical Society' Vol. 6 of *Bulletin of the AAS* p. 453.

Iorio L and Ruggiero M L 2011. HOŘAVA-LIFSHITZ Gravity:. Tighter Constraints for the Kehagias-Sfetsos Solution from New Solar System Data. *International Journal of Modern Physics D* **20**, 1079–1093. arXiv:1012.2822.

Iwasawa K et al. 1996. The broad iron K emission line in the Seyfert 2 galaxy IRAS 18325-5926. *Monthly Notices of the Royal Astronomical Society* **279**, 837–846.

Iyer B et al. 2011 LIGO-India, Proposal of the Consortium for Indian Initiative in Gravitational-wave Observations (IndIGO) Technical Report M1100296. https://dcc.ligo.org/cgi-bin/DocDB/ShowDocument?docid=75988

Jackiw R and Pi S Y 2003. Chern-Simons modification of general relativity. *Physical Review D* **68**(10), 104012. arXiv:gr-qc/0308071.

Jenet F A et al. 2006. Upper Bounds on the Low-Frequency Stochastic Gravitational Wave Background from Pulsar Timing Observations: Current Limits and Future Prospects. *Astrophysical Journal* **653**, 1571–1576. arXiv:astro-ph/0609013.

Jennrich O et al. 2011 NGO Revealing a hidden Universe: opening a new chapter of discovery Technical Report ESA/SRE(2011)19 European Space Agency Noordwijk.

Johannsen T 2013. Regular black hole metric with three constants of motion. *@ARTICLE2013PhRvD..88d4002J, author = Johannsen, T., title = "Regular black hole metric with three constants of motion", journal = Physical Review D, archivePrefix = "arXiv", eprint = 1501.02809, primaryClass = "gr-qc", keywords = Modified theories of gravity, Physics of black holes, year = 2013, month = aug, volume = 88, number = 4, eid = 044002, pages = 044002, doi = 10.1103/PhysRevD.88.044002, adsurl = http://adsabs.harvard.edu/abs/2013PhRvD..88d4002J, adsnote = Provided by the SAO/NASA Astrophysics Data System* **88**(4), 044002. arXiv:1501.02809.

Junker W and Schaefer G 1992. Binary systems - Higher order gravitational radiation damping and wave emission. *Monthly Notices of the Royal Astronomical Society* **254**, 146–164.

Kalaghatgi C, Ajith P and Arun K G 2015. Template-space metric for searches for gravitational waves from the inspiral, merger, and ringdown of binary black holes. *Physical Review D* **91**(12), 124042. arXiv:1501.04418.

Kanner J B et al. 2015. Leveraging waveform complexity for confident detection of gravitational waves. *ArXiv e-prints* . arXiv:1509.06423.

Kehagias A and Sfetsos K 2009. The black hole and FRW geometries of non-relativistic gravity. *Physics Letters B* **678**, 123–126. arXiv:0905.0477.

Kepler J 1609 *Astronomia nova aitiologetos [romanized] : sev physica coelestis, tradita commentariis de motibvs stellae Martis, ex observationibus G. V. Tychonis Brahe.*

Kepler J 1619 *Harmonices Mundi.*

Kerr R P 1963. Gravitational field of a spinning mass as an example of algebraically special metrics. *Physical Review Letters* **11**, 237–238.

Klimenko S, Yakushin I, Mercer A and Mitselmakher G 2008*a*. A coherent method for detection of gravitational wave bursts. *Classical and Quantum Gravity* **25**(11), 114029. arXiv:0802.3232.

Klimenko S, Yakushin I, Mercer A and Mitselmakher G 2008*b*. A coherent method for detection of gravitational wave bursts. *Classical and Quantum Gravity* **25**(11), 114029. arXiv:0802.3232.

Kojima Y 1991. The effects of black hole rotation on line profiles from accretion discs. *Monthly Notices of the Royal Astronomical Society* **250**, 629–632.

Kotake K, Sato K and Takahashi K 2006. Explosion mechanism, neutrino burst and gravitational wave in core-collapse supernovae. *Reports on Progress in Physics* **69**(4), 971–1143. arXiv:astro-ph/0509456.

Kramer M and Champion D J 2013. The European Pulsar Timing Array and the Large European Array for Pulsars. *Classical and Quantum Gravity* **30**(22), 224009.

Laor A 1991. Line profiles from a disk around a rotating black hole. *Astrophysical Journal* **376**, 90–94.

Larmor J 1897. A Dynamical Theory of the Electric and Luminiferous Medium. Part III. Relations with Material Media. *Philosophical Transactions of the Royal Society of London Series A* **190**, 205–300.

Lattimer J M 2012. The Nuclear Equation of State and Neutron Star Masses. *Annual Review of Nuclear and Particle Science* **62**(1), 485–515. arXiv:1305.3510.

Lee H W, Kim Y W and Myung Y S 2010. Slowly rotating black holes in the Hořava-Lifshitz gravity. *European Physical Journal C* **70**, 367–371. arXiv:1008.2243.

Liang M and Marcotte D 2015. A class of non-stationary covariance functions with compact support. *Stochastic Environmental Research and Risk Assessment* pp. 1–15.

LISA study team 1998. Lisa, pre-phase a report.

Littenberg T B, Baker J G, Buonanno A and Kelly B J 2013. Systematic biases in parameter estimation of binary black-hole mergers. *Physical Review D* **87**(10), 104003. arXiv:1210.0893.

Liu M, Lu J, Yu B and Lu J 2011. Solar system constraints on asymptotically flat IR modified Hořava gravity through light deflection. *General Relativity and Gravitation* **43**, 1401–1415. arXiv:1010.6149.

Lopes I and Silk J 2014. Helioseismology and Asteroseismology: Looking for Gravitational Waves in Acoustic Oscillations. *Astrophysical Journal* **794**, 32. arXiv:1405.0292.

Lorentz H A 1892. De relatieve beweging van de aarde en den aether. *Zittingsverlag Akad. v. Wet.* .

MacKay D J C 1999. Comparison of Approximate Methods for Handling Hyperparameters. *Neural Computation* **11**(5), 1035–1068.

MacKay D J C 2003 *Information Theory, Inference and Learning Algorithms* Cambridge University Press Cambridge.

Maggiore M 2000. Gravitational wave experiments and early universe cosmology. *Physics Reports* **331**, 283–367. arXiv:gr-qc/9909001.

Maggiore M 2007 *Gravitational Waves: Volume 1: Theory and Experiments* OUP Oxford. https://books.google.it/books?id=AqVpQgAACAAJ

Manchester R N 2010. Detection of Gravitational Waves using Pulsar Timing. *ArXiv e-prints* . arXiv:1004.3602.

Manchester R N and IPTA 2013. The International Pulsar Timing Array. *Classical and Quantum Gravity* **30**(22), 224010.

Maxwell J C 1865. A dynamical theory of the electromagnetic field. *Philosophical Transactions of the Royal Society of London* **155**, 459–513.

McClintock J E et al. 2011. Measuring the spins of accreting black holes. *Classical and Quantum Gravity* **28**(11), 114009. arXiv:1101.0811.

McLaughlin M A 2013. The North American Nanohertz Observatory for Gravitational Waves. *Classical and Quantum Gravity* **30**(22), 224008. arXiv:1310.0758.

Melkumyan A and Ramos F 2009 *in* 'IJCAI' Vol. 9 pp. 1936–1942.

Merritt D, Alexander T, Mikkola S and Will C M 2011. Stellar dynamics of extreme-mass-ratio inspirals. *Physical Review D* **84**(4), 044024. arXiv:1102.3180.

Michelson A A and Morley E W 1887. On the Relative Motion of the Earth and of the Luminiferous Ether. *Sidereal Messenger, vol. 6, pp.306-310* **6**, 306–310.

Mingarelli C M F and Sidery T 2014. The effect of small inter-pulsar distance variations in stochastic gravitational wave background searches with Pulsar Timing Arrays. arXiv:1408.6840.

Minkowski H 1909 'Space and time, address delivered at cologne'.

Misner C W, Thorne K S and Wheeler J A 1973 *Gravitation* W.H. Freeman and Co. San Fansisco.

Moore C J 2015. The sensitivity of pulsar timing arrays. *Journal of Physics Conference Series* **610**(1), 012018. arXiv:1408.0739.

Moore C J, Berry C P L, Chua A J K and Gair J R 2016. Improving gravitational-wave parameter estimation using Gaussian process regression. *Physical Review D* **93**(6), 064001. arXiv:1509.04066.

Moore C J, Chua A J K, Berry C P L and Gair J R 2016. Fast methods for training gaussian processes on large datasets. *Royal Society Open Science* **3**(5). arXiv:http://rsos.royalsocietypublishing.org/content/3/5/160125.full.pdf.

Moore C J, Cole R H and Berry C P L 2015. Gravitational-wave sensitivity curves. *Classical and Quantum Gravity* **32**(1), 015014. arXiv:1408.0740.

Moore C J and Gair J R 2014. Novel Method for Incorporating Model Uncertainties into Gravitational Wave Parameter Estimates. *Physical Review Letters* **113**(25), 251101. arXiv:1412.3657.

Moore C J and Gair J R 2015. Testing the no-hair property of black holes with x-ray observations of accretion disks. *Physical Review D* **92**(2), 024039. arXiv:1507.02998.

Moore C J and Gair J R in prep. Testing the no-hair property of black holes with gravitational wave observations of extreme-mass-ratio inspirals.

Moore C J, Taylor S R and Gair J R 2015. Estimating the sensitivity of pulsar timing arrays. *Classical and Quantum Gravity* **32**(5), 055004. arXiv:1406.5199.

Mroué A H et al. 2013. Catalog of 174 Binary Black Hole Simulations for Gravitational Wave Astronomy. *Physical Review Letters* **111**(24), 241104. arXiv:1304.6077.

Nelemans G, Yungelson L R and Portegies Zwart S F 2001. The gravitational wave signal from the Galactic disk population of binaries containing two compact objects. *Astronomy and Astrophysics* **375**, 890–898. arXiv:astro-ph/0105221.

Newton I 1687 *Philosophiæ naturalis principia mathematica*.

Nordström G 1912. Relativitätsprinzip und gravitation. *Physikalische Zeitschrift* .

Nordström G 1913. Zur theorie der gravitation vom standpunkt des relativitätsprinzip. *Annalan der Physik* .

Ott C D 2009. The gravitational-wave signature of core-collapse supernovae. *Classical and Quantum Gravity* **26**(6), 063001. arXiv:0809.0695.

Owen B J 1996. Search templates for gravitational waves from inspiraling binaries: Choice of template spacing. *Physical Review D* **53**, 6749–6761. arXiv:gr-qc/9511032.

Paciorek C J and Schervish M J 2004 *in* S Thrun, L Saul and B Schölkopf, eds, 'Advances in Neural Information Processing Systems 16' MIT Press Cambridge, MA.

Page D N and Thorne K S 1974. Disk-Accretion onto a Black Hole. Time-Averaged Structure of Accretion Disk. *Astrophysical Journal* **191**, 499–506.

Pan Y et al. 2011. Inspiral-merger-ringdown multipolar waveforms of nonspinning black-hole binaries using the effective-one-body formalism. *Physical Review D* **84**(12), 124052. arXiv:1106.1021.

Peters P C and Mathews J 1963. Gravitational Radiation from Point Masses in a Keplerian Orbit. *Physical Review* **131**, 435–440.

Planck Collaboration et al. 2014. Planck intermediate results. XXX. The angular power spectrum of polarized dust emission at intermediate and high Galactic latitudes. *ArXiv e-prints* . arXiv:1409.5738.

Poincaré H 1900. La thÃ©orie de Lorentz et le principe de rÃ©action. *Archives nÃ©erlandaises des sciences exactes et naturelles* .

Pretorius F 2005. Evolution of binary black-hole spacetimes. *Physical Review Letters* **95**, 121101.

Prix R 2009 *in* W Becker, ed., 'Neutron Stars and Pulsars' Astrophysics and Space Science Library Springer Berlin chapter 24, pp. 651–685.

Psaltis D and Johannsen T 2012. A Ray-tracing Algorithm for Spinning Compact Object Spacetimes with Arbitrary Quadrupole Moments. I. Quasi-Kerr Black Holes. *Astrophysical Journal* **745**, 1. arXiv:1011.4078.

Pürrer M 2014. Frequency-domain reduced order models for gravitational waves from aligned-spin compact binaries. *Classical and Quantum Gravity* **31**(19), 195010. arXiv:1402.4146.

Quiñonero-Candela J, Rasmussen C E and Williams C K I 2007 *in* L Bottou, O Chapelle, D DeCoste and J Weston, eds, 'Large-scale Kernel Machines' MIT Press Cambridge, MA chapter 9, pp. 203–223.

Rasmussen C E and Williams C K I 2006 *Gaussian Processes for Machine Learning* MIT Press Cambridge, MA.

Ryan F D 1996. Effect of gravitational radiation reaction on nonequatorial orbits around a Kerr black hole. *Physical Review D* **53**, 3064–3069. arXiv:gr-qc/9511062.

Ryan F D 1997. Accuracy of estimating the multipole moments of a massive body from the gravitational waves of a binary inspiral. *Physical Review D* **56**, 1845–1855.

Santamaría L et al. 2010. Matching post-newtonian and numerical relativity waveforms: Systematic errors and a new phenomenological model for nonprecessing black hole binaries. *Physical Review D* **82**, 064016.

Sathyaprakash B S and Dhurandhar S V 1991. Choice of filters for the detection of gravitational waves from coalescing binaries. *Physical Review D* **44**, 3819–3834.

Sathyaprakash B S and Schutz B F 2009. Physics, Astrophysics and Cosmology with Gravitational Waves. *Living Reviews in Relativity* **12**, 2. arXiv:0903.0338.

Sazhin M V 1978. Opportunities for detecting ultralong gravitational waves˙ *Soviet Astronomy* **22**, 36–38.

Schmidt P, Hannam M and Husa S 2012. Towards models of gravitational waveforms from generic binaries: A simple approximate mapping between precessing and nonprecessing inspiral signals. *Physical Review D* **86**(10), 104063. arXiv:1207.3088.

Schmidt P, Ohme F and Hannam M 2015. Towards models of gravitational waveforms from generic binaries: II. Modelling precession effects with a single effective precession parameter. *Physical Review D* **91**(2), 024043. arXiv:1408.1810.

Seitenzahl I R et al. 2015. Neutrino and gravitational wave signal of a delayed-detonation model of type ia supernovae. *Physical Review D* **92**, 124013.

Seljak U and Zaldarriaga M 1997. Signature of Gravity Waves in the Polarization of the Microwave Background. *Physical Review Letters* **78**, 2054–2057. arXiv:astro-ph/9609169.

Sesana A 2013. Systematic investigation of the expected gravitational wave signal from supermassive black hole binaries in the pulsar timing band. *Monthly Notices of the Royal Astronomical Society* **433**, L1–L5. arXiv:1211.5375.

Sesana A, Vecchio A and Colacino C N 2008. The stochastic gravitational-wave background from massive black hole binary systems: implications for observations with Pulsar Timing Arrays. *Monthly Notices of the Royal Astronomical Society* **390**, 192–209. `arXiv:0804.4476`.

Sesana A, Vecchio A and Volonteri M 2009. Gravitational waves from resolvable massive black hole binary systems and observations with Pulsar Timing Arrays. *Monthly Notices of the Royal Astronomical Society* **394**(4), 2255–2265. `arXiv:0809.3412`.

Shakura N I and Sunyaev R A 1973. Black holes in binary systems. Observational appearance. *Astronomy and Astrophysics* **24**, 337–355.

Shannon C E 1948. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **27**(4), 623–656.

Shannon R M et al. 2013. Gravitational-wave limits from pulsar timing constrain supermassive black hole evolution. *Science* **342**(6156), 334–337. `arXiv:1310.4569`.

Shoemaker D 2010 Advanced ligo anticipated sensitivity curves LIGO Document LIGO-T0900288-v3. `https://dcc.ligo.org/LIGO-T0900288/public`

Sidery T et al. 2014. Reconstructing the sky location of gravitational-wave detected compact binary systems: Methodology for testing and comparison. *Physical Review D* **89**(8), 084060. `arXiv:1312.6013`.

Snelson E and Ghahramani Z 2006 *in* Y Weiss, B Schölkopf and J. C Platt, eds, 'Advances in Neural Information Processing Systems' Vol. 18 MIT Press Cambridge, MA pp. 1257–1264.

Somiya K 2012. Detector configuration of KAGRA-the Japanese cryogenic gravitational-wave detector. *Classical and Quantum Gravity* **29**(12), 124007. `arXiv:1111.7185`.

Stein L C 2014. Rapidly rotating black holes in dynamical Chern-Simons gravity: decoupling limit solutions and breakdown. *ArXiv e-prints* . `arXiv:1407.2350`.

Stein M L 1999 *Interpolation of Spatial Data* Springer Series in Statistics Springer–Verlag New York, NY.

Steiner J F et al. 2011. The spin of the black hole microquasar XTE J1550-564 via the continuum-fitting and Fe-line methods. *Monthly Notices of the Royal Astronomical Society* **416**, 941–958. arXiv:1010.1013.

Stroeer A and Vecchio A 2006. The LISA verification binaries. *Classical and Quantum Gravity* **23**, 809. arXiv:astro-ph/0605227.

Szilágyi B et al. 2015. Approaching the post-newtonian regime with numerical relativity: A compact-object binary simulation spanning 350 gravitational-wave cycles. *Physical Review Letters* **115**, 031102.

Tanaka Y et al. 1995. Gravitationally redshifted emission implying an accretion disk and massive black hole in the active galaxy MCG-6-30-15. *Nature* **375**, 659–661.

Taracchini A et al. 2012. Prototype effective-one-body model for nonprecessing spinning inspiral-merger-ringdown waveforms. *Physcial Review D* **86**(2), 024011. arXiv:1202.0790.

Taracchini A et al. 2014. Effective-one-body model for black-hole binaries with generic mass ratios and spins. *Physical Review D* **89**(6), 061502. arXiv:1311.2544.

Taylor J H, Fowler L A and McCulloch P M 1979. Measurements of general relativistic effects in the binary pulsar PSR 1913+16. *Nature* **277**, 437–440.

The LIGO Scientific Collaboration and the Virgo Collaboration 2016*a*. Calibration of the Advanced LIGO detectors for the discovery of the binary black-hole merger GW150914. *ArXiv e-prints* . arXiv:1602.03845.

The LIGO Scientific Collaboration and The Virgo Collaboration 2016*b*. Gw150914: The advanced ligo detectors in the era of first discoveries. *Phys. Rev. Lett.* **116**, 131103.

The LIGO Scientific Collaboration and the Virgo Collaboration 2016*c*. Observation of Gravitational Waves from a Binary Black Hole Merger. *Physical Review Letters* **116**(6), 061102. arXiv:1602.03837.

The LIGO Scientific Collaboration and the Virgo Collaboration 2016*d*. Observing gravitational-wave transient GW150914 with minimal assumptions. *ArXiv e-prints* . arXiv:1602.03843.

The LIGO Scientific Collaboration and the Virgo Collaboration 2016*e*. Properties of the binary black hole merger GW150914. *ArXiv e-prints* . arXiv:1602.03840.

The LIGO Scientific Collaboration and the Virgo Collaboration 2016*f*. Tests of general relativity with GW150914. *ArXiv e-prints* . arXiv:1602.03841.

The LIGO Scientific Collaboration and the Virgo Collaboration 2016*g*. The Rate of Binary Black Hole Mergers Inferred from Advanced LIGO Observations Surrounding GW150914. *ArXiv e-prints* . arXiv:1602.03842.

the Nobel Prize in Physics 1993. *http://www.nobelprize.org/nobel_prizes/physics/laureates/1993/index.html* .

Thorne K S 1987 *in* S. W Hawking and W Israel, eds, 'Three Hundred Years of Gravitation' Cambridge University Press Cambridge chapter 9, pp. 330–458.

Thrane E and Romano J D 2013. Sensitivity curves for searches for gravitational-wave backgrounds. *Physical Review D* **88**(12), 124032. arXiv:1310.5300.

Tinto M and Alves M E D S 2010. LISA sensitivities to gravitational waves from relativistic metric theories of gravity. *Physical Review D* **82**(12), 122003. arXiv:1010.1302.

Tinto M and Dhurandhar S V 2005. Time-delay interferometry. *Living Reviews in Relativity* **8**(4).

Tinto M, Estabrook F B and Armstrong J W 2002. Time-delay interferometry for LISA. *Physical Review D* **65**(8), 082003.

Unnikrishnan C S 2013. IndIGO and Ligo-India Scope and Plans for Gravitational Wave Research and Precision Metrology in India. *International Journal of Modern Physics D* **22**(1), 1341010.

Vallisneri M 2008. Use and abuse of the Fisher information matrix in the assessment of gravitational-wave parameter-estimation prospects. *Physical Review D* **77**(4), 042001. arXiv:gr-qc/0703086.

Vallisneri M and Galley C R 2012. Non-sky-averaged sensitivity curves for space-based gravitational-wave observatories. *Classical and Quantum Gravity* **29**(12), 124015. arXiv:1201.3684.

Vallisneri M and Yunes N 2013. Stealth bias in gravitational-wave parameter estimation. *Physical Review D* **87**(10), 102002. arXiv:1301.2627.

van Haasteren R and Levin Y 2013. Understanding and analysing time-correlated stochastic sig nals in pulsar timing. *Monthly Notices of the Royal Astronomical Society* **428**, 1147–1159. arXiv:1202.5932.

Veitch, J. et al. 2015. Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library. *Physical Review D* **91**(4), 042003. arXiv:1409.7215.

Veitch J, Pürrer M and Mandel I 2015. Measuring Intermediate-Mass Black-Hole Binaries with Advanced Gravitational Wave Detectors. *Physical Review Letters* **115**(14), 141101. arXiv:1503.05953.

Veitch J and Vecchio A 2010. Bayesian coherent analysis of in-spiral gravitational wave signals with a detector network. *Physical Review D* **81**(6), 062003. arXiv:0911.3820.

Vigeland S, Yunes N and Stein L C 2011. Bumpy black holes in alternative theories of gravity. *Physical Review D* **83**(10), 104027. arXiv:1102.3706.

Volonteri M 2010. Formation of supermassive black holes. *The Astronomy and Astrophysics Review* **18**(3), 279–315. arXiv:1003.4404.

Volonteri M, Haardt F and Madau P 2003. The Assembly and Merging History of Supermassive Black Holes in Hierarchical Models of Galaxy Formation. *The Astrophysical Journal* **582**(2), 559–573. arXiv:astro-ph/0207276.

Watson G N 1995 *A Treatise on the Theory of Bessel Functions* Cambridge Mathematical Library second edn Cambridge University Press Cambridge.

Weber J 1960. Detection and generation of gravitational waves. *Physical Review* **117**, 306–313.

Wendland H 2004 *Scattered Data Approximation* Cambridge Monographs on Applied and Computational Mathematics Cambridge University Press Cambridge.

Wheeler J A 2000 *Geons, Black Holes, and Quantum Foam: A Life in Physics* W. W. Norton & C.

Wiener N 1949*a* *Extrapolation, Interpolation and Smoothing of Stationary Time Series* John Wiley & Sons New York, NY.

Wiener N 1949*b* *The extrapolation, interpolation and smoothing of stationary time series* The M.I.T. Press Cambridge, Massachusetts.

Will C M 2006. The confrontation between general relativity and experiment. *Living Reviews in Relativity* **9**(3).

Willke B et al. 2006. The GEO-HF project. *Classical and Quantum Gravity* **23**, 207.

Yagi K and Seto N 2011. Detector configuration of DECIGO/BBO and identification of cosmological neutron-star binaries. *Physcial Review D* **83**(4), 044011. `arXiv:1101.3940`.

Yagi K, Yunes N and Tanaka T 2012. Slowly rotating black holes in dynamical Chern-Simons gravity: Deformation quadratic in the spin. *Physical Review D* **86**(4), 044037. `arXiv:1206.6130`.

Yunes N and Pretorius F 2009. Dynamical Chern-Simons modified gravity: Spinning black holes in the slow-rotation approximation. *Physical Review D* **79**(8), 084043. `arXiv:0902.4669`.

Zdziarski A A et al. 1994. Physical Processes in the X-Ray / Gamma-Ray Source of IC4329A. *Monthly Notices of the Royal Astronomical Society* **269**, L55.