School of Physics and Astronomy

There's More than One Way to Ride the Wave: A Multi-Disciplinary Approach to Gravitational Wave Data Analysis

Rhys D Green

Submitted for the degree of Doctor of Philosophy
School of Physics and Astronomy
Cardiff University

06/08/2021

Summary of thesis

Since the first detection in 2015, gravitational-wave astronomy has progressed hugely. Several observing runs have been completed, resulting in many more confirmed detections of compact binary coalescence. As the number of detections grows larger, the potential for exciting science also increases, however, this is not without challenges. Specifically efficiently analyzing growing data will present many computational problems going forward. In order to properly interpret and understand this growing data, we must develop new ways to approach these computational problems.

When seeking to tackle a difficult problem there are broadly two ways to do this. One can tackle the problem using some physical or mathematical insight, this understanding can then be translated into a simpler formulation or good approximation which makes the problem tractable. This has been the standard way to tackle problems since the beginning of science, recently, however, data-driven methods have become hugely popular. These data-driven methods such as machine learning generally do not use physical insight but make use of large amounts of data efficiently to produce solutions to these intractable problems.

This thesis draws on both of these approaches and presents several new methods to analyze gravitational-wave data. In chapters 2 - 3 we derive a way to describe a precessing waveform as a harmonic decomposition, where each harmonic is a simple non-precessing waveform. With this formulation, we are able to obtain a simple picture of precession as the beating of two waveforms. We then use this understanding to answer questions such as when will we observe precessing waveforms? And where in parameter space will we be able to observe precessing waveforms?

The remaining chapters look at data-driven approaches, using machine learning techniques to improve different aspects of gravitational-wave data analysis. Chapter 5uses Gaussian Processes to interpolate posterior samples, this allows us to have a smooth continuous representation of our posterior as opposed to histograms for example. Chapter 6 uses using advances in waveform modeling and GPUs to potentially make parameter estimation more efficient. In chapters 7 and 8we look at how reliable machine learning techniques are, we show that often they do not incorporate uncertainty properly into their predictions. We then present a simple algorithm for both classification and regression pipelines that can be used with any machine learning model to address this.

Finally, in the conclusions, we review the work presented as a whole and discuss ways in which these two approaches can be combined to get the best of both. We suggest that using our physical insights to guide and constrain our data-driven methods will eventually provide the best path forward for gravitational-wave data analysis.

Contents

List of Figures

List of Tables

Acknowledgements

Firstly I would like to thank my supervisor Stephen Fairhurst, I really can't imagine enjoying my time during the PhD as much with anyone else as my supervisor. As well as being brilliant you're a great person and genuine inspiration to me. I would also like to thank Mark Hannam and Vivien Raymond, there was never a dull moment working with you both! I think between us all we managed to produce some pretty good work as well as having a good time.

I was hugely lucky during the PhD to work at Oracle AI Apps, the six months there were some of the best of the entire PhD. I would like to thank the entire AI Apps team, I learnt more from you all in this period than I thought possible. Special thanks must go to Matt Rowe and Alberto Polleri for supervising me during this time. I feel massively privileged to have been given the opportunity to work with you both and come out of the experience with an entirely new outlook on my work and my future. The placement experience was also made much better due to the hard work of Rosemary Granger. I really appreciate everything that you did for me during the four years and especially in helping make the placement go as smoothly as it did.

My time studying in Cardiff was an amazing experience. I can barely remember a day that I didn't enjoy and this is mostly due to the amazing people in the group and the department as a whole. Apologies to anyone I've missed I could honestly name everyone there because I have great memories with all of you. Thanks to Ali, Cam, Charlie, Dave, Ed, Eleanor, James, Iain, Jonathan, Matt Smith, Matt Bates, Phil, Ronaldas, Seb, Vassilis and Virginia I'm glad we became friends rather than just colleagues. Special thanks to Seb for all the help you gave me during the last year, even whilst fighting Tensorflow our team chats were often the highlight of my week! Finally an enormous thank you to Charlie Hoy, you're one of the most talented people I know and I really don't think the PhD would have been the same if I hadn't spent most of it sitting next to you (sometimes virtually)!

The final eighteen months of this PhD were completed during COVID but I can honestly say that they have been some of the best months of my life. This is in a large part thanks to Ali, Josh and most of all Elle. I've said it lots of times but I'm pretty sure that between our house crawls and various escapes, no one else had as much fun as we did.

Finally, I would like to thank my parents, I feel incredibly lucky to have had three amazing parents and I don't think I would have been able to achieve this without each of them.

. . . if it is humanly possible, consider it to be within your reach

Marcus Aurelius

Chapter 1

Introduction

1.1 General Relativity and Gravitational Waves

In 1915 Einstein completed his theory of General Relativity. This work unified space and time into the four-dimensional quantity known as spacetime [4]. One of the insights from this work was that spacetime is curved by all matter and this curvature can be described using the mathematical framework of differential geometry developed by Reimann and others. Shortly after this theory was developed, many consequences were discovered by probing these equations.

Gravitational waves were initially proposed by Poincaré in 1905 [5] he suggested analogously to electromagnetic waves that gravitational waves would be produced by accelerating masses. When the framework of General Relativity was developed, this description could be formalized as gravitational waves producing ripples in the 4d spacetime. The controversy and skepticism (even from Einstein himself) surrounding the acceptance of gravitational waves is well documented [6] but eventually, they became an accepted consequence of General Relativity [7]. It is now accepted that gravitational waves are produced by the aspherical acceleration of mass. However, the measurable effect of them on spacetime is generally very small and therefore the only feasible way to currently detect them is to observe gravitational waves from very massive, dense objects such as black holes and neutron stars.

The gravitational wave emission from binary mergers can be obtained by solving Einstein's equations. It's generally not possible to do this analytically however so we must resort to either approximate methods such as Post-Newtonian [8] or computational ones such as Numerical relativity [9]. These methods allowed us to understand the types of signals that could be produced and therefore allowed us to understand that it would be possible to detect gravitational-waves on Earth. This then led to the development on gravitational-wave detectors that would in theory be sensitive enough to observe the signals emitted by these objects.

1.2 Gravitational Wave Detectors

One of the first investigations into observing gravitational-waves carried out was by Joseph Weber, he developed an instrument known as Weber bars [10]. These weber bars were used in an experiment to probe gravitational waves produced by the Galactic Centre. Weber claimed to have discovered gravitational waves in the 1970s however these claims were never reproduced by other experimental efforts and have therefore never been accepted by the scientific community [11]. Despite this, these pioneering investigations began the era of searching experimentally for gravitational waves.

In the same decade that Weber was attempting to find gravitational waves experimentally, Russell Hulse and Joseph Taylor were gathering observational evidence of their existence. They discovered the first binary pulsar an observation which would eventually earn the Nobel prize. They observed that the orbital decay for the binary pulsar [12, 13] matched the energy that would be emitted as gravitational waves for the system. This observational evidence generated more interest and work in this field and experimental work, pioneered by Rainer Weiss, Ron Drever, and many others [14, 15, 16, 17], began focusing on the idea of using laser interferometers as a means to detect GWs.

A gravitational wave effectively stretches and squeezes spacetime. These interferometers work by detecting the very small changes in spacetime caused by a passing gravitational wave. This small change is detected by splitting a laser beam in perpendicular directions, these beams travel along the interferometer arms and are reflected back along the interferometer arms to the origin of the split and are then re-combined. If the length of the arms are exactly the same, the beams should interfere with one another destructively. If there is a difference between the arm lengths caused by a passing gravitational wave then there will be an interference pattern that will be dependent on the source of the gravitational wave.

Though the fundamental principle of detecting the differential arm length with an interference pattern is the same as this simple description, modern detectors have several enhancements such as “Fabry Pérot cavities” which effectively increase the distance the laser travels which then has a similar effect to having longer arms, power recycling to produce a much more powerful laser and many more [18]. As well as this, to ensure these detectors are sensitive enough many sources of noise must be reduced. At low frequencies (0-20Hz) the dominant source of noise comes from Seismic activity, at high frequencies ($>$ 100Hz) quantum noise at due to the uncertainty in *photon counting* dominates and between these thermal noise dominates [19]. The combination of these noise sources produce a noise budget which limits our sensitivity, more formally these noise sources determine our PSD, see section 1.3a, and our strain sensitivity is the square root of this, see figure 1.1 for an example PSD.

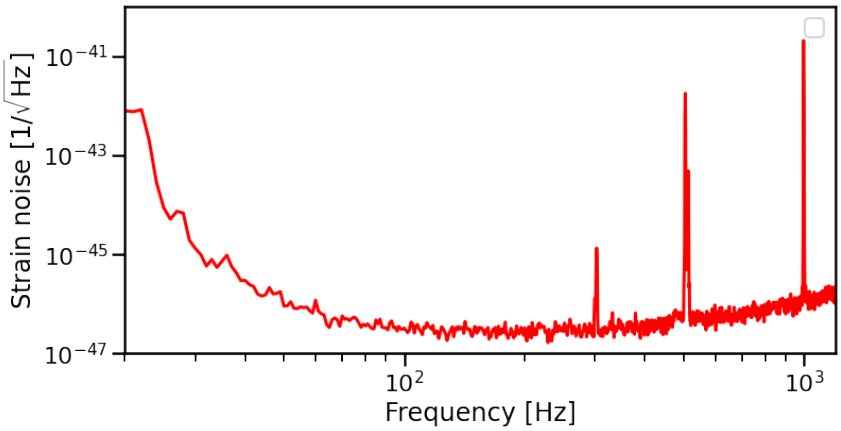

Figure 1.1: An example PSD. This is the welch average of the data taken the Hanford detector using 32s around the time of the detection [1]

The principles outlined above have been developed and refined over many years, culminating in a (growing) network of incredibly sophisticated interferometers spread across the globe [18, 20, 21, 22]. As well as the current detectors, several *next generation* detectors have been suggested such as the Einstein Telescope (ET) and the Laser Interferometer Space Antenna (LISA) [23, 24]. This growing network as well as the future generation of detectors, means that gravitational-wave astronomy is very much still in its infancy and points to an exciting future where it will continue to progress for decades to come.

1.3 Data Analysis

Now that theoretical and experimental foundations had been laid for gravitational-wave astronomy, the final problem to be addressed is that of efficient data analysis; A GW observed at a detector will be a weak signal buried in noise. We need to confidently extract it from the data and extract the properties of the source from the observed signal

a The Power Spectrum Density and the natural inner product

The data collected at a gravitational wave observatory is a time series, if there is a signal present in the data it is likely to be relatively weak compared to the noise in the detector therefore one must use statistical methods to extract these signals from the noise. Following [25, 26, 27, 28] we will now derive the optimal method to extract these signals from the time series given the assumed properties of our detector.

We first assume the output of our detector (in the absence of a gravitational-wave) follows a stationary random process i.e.:

$$\langle x \rangle = \lim_{T \to \infty} \frac{1}{T} \int_{\frac{T}{2}}^{\frac{T}{2}} x(t) dt \tag{1.1}$$

If we further assume that our time series is windowed and zero-mean, such that the expectation of x is zero, then the average power for a stationary process such as this can then be written as :

$$\langle x^2 \rangle = \lim_{T \to \infty} \frac{1}{T} \int_{-\infty}^{\infty} |x(t)^2| dt \tag{1.2}$$

Then by invoking Parseval's Theorem, we can describe the average power in the frequency domain as:

$$\langle x^2 \rangle = \lim_{T \to \infty} \frac{1}{T} \int_{-\infty}^{\infty} |\tilde{x}(f)^2| df \tag{1.3}$$

Where $\tilde{x}(f)$ is the Fourier transform of our time series, $x(t)$, using the following convention for our transform.

$$\tilde{x}(f) = \int_{-\infty}^{\infty} x(t) e^{-2\pi i f t} dt \tag{1.4}$$

Then with the knowledge that our signal is real, we know that the postive and negative frequencies will be symmetric we than have:

$$\langle x^2 \rangle = \frac{2}{T} \int_{0}^{\infty} |\tilde{x}(f)^2| df \tag{1.5}$$

$$= \int_{0}^{\infty} S_x(f) df \tag{1.6}$$

The integrand $S_x(f) = \frac{2}{T} |\tilde{x}(f)^2|$ is referred to as the power spectral density as it reflects the distribution of power across the range of frequencies. Equivalently by noting that the power spectrum density is twice the auto-correlation function of a stationary process we have that

$$S_x(f) = 2 \int_{-\infty}^{\infty} R_x(\tau) e^{-2\pi f \tau} d\tau \,, \tag{1.7}$$

where $R_x(\tau) = \langle x(t)(x(t-\tau) \rangle$. The probability of observing a time series, $x(t)$, from this process is therefore:

$$p_x(x(t)) \propto \exp \left[-\frac{1}{2} \, 4 \int_{0}^{\infty} \frac{|\tilde{x}(f)|^2}{S_x(f)} df \right] \tag{1.8}$$

This then leads to a natural inner product to determine how similar any two time series are in the space defined by our power spectrum density, we define this inner product as:

$$(a,b) = 4\mathrm{Re}\int_0^\infty \frac{\tilde{a}(f)\tilde{b}^*(f)}{S(f)}df, \tag{1.9}$$

such that we now now define the probability of a given time series as:

$$p_x(x(t)) \propto \exp\left[-\frac{1}{2}(x,x)\right] \tag{1.10}$$

As the inner product defines the similarity between two time series, it can be shown that it is the optimal detection statistic when trying to determine whether the data contains a signal.

In order to compute this detection statistic, we must however have very accurate models of a gravitational-wave signal. To produce exact waveforms one must solve the Einstein equations in full, this is not possible analytically and to solve them numerically is computationally prohibitive for most data analysis applications. One must therefore produce approximate waveform models often called approximants [29, 30, 31, 32]. These approximants all take a given set of physical parameters and then either produce a time or frequency domain representation of the waveform that would be produced.

The detection statistic defined above shows us the natural way to compare the data to our waveform models. This detection statistic will only be large if our waveform model parameters match the true source parameters that produced the signal in the data. This means that we must compare the data to a very large sample of models across the eight dimensional mass and spin space to ensure that we are not missing possible detections. This leads to the notion of *template banks*, which are banks of waveform models which cover the entire space densely enough to ensure that any point in this hyper-surface has a sensitivity above some minimum threshold [33, 34]. The data recorded at the detectors is then continuously compared to each of the models in these template banks, in principle if any of the templates has a signal to noise ratio (SNR) higher than some minimum detection threshold we can confidently claim there is a signal in the data.

This theoretical detection pipeline is generally an idealised picture of what must be done when working with the real detector data however. Due to effects such as glitches and other non-Gaussian noise characteristics several additional data quality steps such as detector characterisation, signal consistency etc, [35, 36, 19] are required to confidently detection GWs in practise. These additional checks are incorporated into sophisticated detection pipelines such as PyCBC and GstLal [37, 38] which go well beyond simple matched filtering.

b The Inverse Problem

Once we have detected a signal in the data, for most of the science we would like to do we need to estimate the parameters of the source. Estimating the source

parameters for a given signal observed at a detector is, therefore, a key objective for gravitational-wave data analysis. This is a classic case of an inverse, problem i.e. given an observation of a signal in the data one must solve the inverse mapping to estimate parameters that produced the signal

$$h(\cdot) = F(\vec{\theta}) \tag{1.11}$$

Where $h(\cdot)$ is the waveform, the dot signifies that this can be a function of either time or frequency, and $F(\vec{\theta})$ is some function of the latent physical parameters. At its most simple this problem, assuming F can be thought of as a linear operator, reduces to :

$$h(\cdot) = F\vec{\theta} \tag{1.12}$$

$$\Rightarrow \vec{\theta} = F^{-1}h(\cdot) \tag{1.13}$$

An important but challenging aspect of inverse problems is that often (and certainly the case here in gravitational-wave analysis) the map is not simply bijective, different combinations of parameters may generate the same (or extremely similar) signal with our models. When you also consider additional noise in the signals that we observe, it is clear that this becomes even more difficult as any small differences between signals may be masked by different noise realisations. The problem, therefore, becomes to estimate a range of plausible parameters, along with their associated probabilities, that produced the signal seen in the data.

This range of plausible parameters must therefore reflect our uncertainty around any estimated parameters. The most natural framework to describe uncertainty for this type of problem is Bayesian Inference.

Bayesian inference is one framework in which to tackle inverse problems. Philosophically, from a Bayesian perspective, one starts with a set of prior assumptions about the problem, which are then updated once data has been observed. The introduction of the data (via the likelihood) transforms these prior assumptions into a posterior probability.

$$\begin{aligned} p(\vec{\theta}|d) &= \frac{p(d|\vec{\theta})p(\vec{\theta})}{p(d)} \\ &\propto p(d|\vec{\theta})p(\vec{\theta}) \end{aligned} \tag{1.14}$$

In equation 1.14 we formalize this and can consider each of these parts individually, the left-hand term on the numerator is the likelihood $p(d|\theta)$, this term is defined by your assumptions about the noise. It generally can be thought of as how well a particular model of a system fits the data. We will describe this in more detail when we consider the gravitational wave likelihood specifically. The second term in

the numerator are the priors $p(\theta)$, these priors quantify your assumptions about a system before observing any data. Finally, we have the denominator, this term is known as the evidence $p(d)$. The evidence quantifies the probability of observing the data under all possible configurations of your model. In the gravitational-wave example, this means that for a given model it is the probability of observing the source under all possible parameter configurations. Generally, this term is intractable but as it is a constant we can still estimate the relative probabilities between different points in parameter space and therefore can still perform estimate the posterior up to a normalization constant.

From equation 1.14 one can then marginalize out other parameters to obtain posterior distributions for individual parameters

$$p(\theta_i|d) = \int p(\vec{\theta}|d) d\theta_1 ... d\theta_{i-1} d\theta_{i+1} ... d\theta_N. \tag{1.15}$$

c The GW likelihood and GW priors

Here we make the assumption of gaussian noise, for a frequency domain signal (FFT of the original time series observed at a detector) we then have data which contains a signal plus noise;

$$d(f) = h(\vec{\theta}, f) + \epsilon(f) \tag{1.16}$$

Where $h(\vec{\theta}, f)$ is a frequency domain template waveform for a given set of parameters, θ, and epsilon is our noise distribution, which we assume to be a zero mean Gaussian distribution, where the noise level at a given frequency is determined by the PSD, i.e. $\epsilon \sim \mathcal{N}(0, \vec{\sigma})$, where $\vec{\sigma}$ is the noise level. We then rearrange as follows to obtain our likelihood [39, 40]:

$$\begin{aligned} \epsilon(f) &= d(f) - h(\vec{\theta}, f) \\ \mathcal{N}(0, \vec{\sigma}) &= d(f) - h(\vec{\theta}, f) \end{aligned} \tag{1.17}$$

We see here then that the assumptions about the noise distribution then defines our likelihood. Using the assumptions about the PSD when considering a single frequency bin we then have.

$$\mathcal{L}(d_i|\vec{\theta}) = \frac{1}{2\pi\sigma_i} \exp\left(-2\Delta f \frac{|d_i - h_i(\vec{\theta})|^2}{\sigma_i}\right) \tag{1.18}$$

Where σ_i is the expected noise level as defined by the PSD. We generally want to consider many frequency bins. If we assume stationary noise then the covariance matrix across our range of frequencies is diagonal. Therefore the likelihood for the entire signal across n frequency bins is:

$$\mathcal{L}(d|\vec{\theta}) = \prod_{i=1}^{n} \mathcal{L}(d_i|\vec{\theta}) \tag{1.19}$$

For simplicity we generally work with the log-likelihood. Relating this to matched filtering. We see that our log likelihood resembles the inner product and therefore can be most naturally written in these terms i.e.

$$\begin{aligned} \ln \mathcal{L}(d_i|\vec{\theta}) &= -\frac{1}{2}\sum_{i=1}^{n} \ln(2\pi\sigma_i) - 2\Delta f \sum_{i=1}^{n} \frac{|d_i - h(\vec{\theta})|^2}{\sigma_i} \\ &= \Phi - \frac{1}{2}(d - h(\vec{\theta}), d - h(\vec{\theta})) \end{aligned} \tag{1.20}$$

Where $\Phi = \frac{1}{2}\sum_{i=1}^{n} \ln(2\pi\sigma_i)$ and the right hand term is the inner product (defined in section a) between the data and a template waveform.

As well as the likelihood, in order to to compute the posterior probability we need priors. These should define our expectations about a parameter in the absence of data. In the strong signal limit where signals have a very high SNR then the likelihood dominates our posterior probability, however for ground-based gravitational-wave detectors we generally have relatively low SNR signals and therefore our priors can have significant effects on our posteriors.

There is no such thing as a *correct* prior and it is completely acceptable in a Bayesian philosophy to define your priors differently to others, conduct an analysis and obtain different results without either of the results technically being incorrect. In practise however many of our priors are determined naturally by geometric relationships such as those for sky localisation and orientation or by physical bounds such as the limits of extremal spins(though you could of course define different valid priors within these ranges e.g. [41, 42]). For details of the priors generally used in gravitational-wave analysis see Appendix B.1 of [2].

d Stochastic Tools for Bayesian inference

We now have the theoretical framework to estimate our source parameters, $\vec{\theta}$, however, the integrals required to evaluate equation 1.14 are generally intractable. The problem then becomes a computational problem of how to approximate this equation numerically. In simple cases, it may be possible to evaluate the integrals by sampling on a grid, here one would try many points on a very dense grid to estimate the best-fit parameters. If the spaces between points on the grid are small then in principle one could estimate the source parameters this way, this becomes exponentially inefficient as the number of dimensions increase however. In gravitational wave parameter estimation, we are usually dealing with $n > 15$ dimensions, so we need to use alternative methods.

The common solution to this problem has been to use stochastic samplers, these fall very broadly (and not exclusively as often nested samplers use MCMC in the algorithm) into two techniques:

1. Markov Chain Monte Carlo (MCMC) [43, 44]
2. Nested sampling [45]

Monte Carlo methods generally refer to methods that obtain approximate numerical results (such as integrals and expectations) by randomly simulating a system many times and exploiting the convergence properties to obtain good estimates to the underlying system. A Markov chain is a stochastic process where a chain moves through possible states of a system, for a chain to be Markovian it must be *memoryless*. This property means that the chain can have no knowledge of its past. More formally conditional on the current state of a chain, the future states are independent of the past states [46]. Combining Markow chains with Monte Carlo simulation allows one to draw samples from a posterior distribution in a way that guarantees that the chain will asymptotically converge to the true distribution [43, 44].

The most common MCMC algorithm is known as the Metropolis-Hastings algorithm [47, 48]. This algorithm works by proposing new positions in parameter space according to some proposal distribution, one then simulates the system at this new state and calculates the ratio of probabilities between the current state and the proposed state. If the proposed state has a higher probability then this position is accepted and the chain continues from this new position. If the proposed state has a lower probability then the point is only accepted if the ratio of probabilities is larger than a draw from a random uniform distribution, i.e. the acceptance probability is:

$$\alpha(\theta, \theta*) = min\left(1, \frac{Q(\theta^*|\theta)p(\theta_*|d)}{Q(\theta|\theta_*)p(\theta|d)}\right) \tag{1.21}$$

Where Q is our transition kernel that determines how we move from one position in parameter space to another and p is the probability density there.

As mentioned above, MCMC methods have asymptotic guarantees. These are by definition only then true with infinite samples, often a practical problem when performing parameter estimation with MCMC methods is how many samples are *enough* to provide reasonable approximations. This problem is particularly difficult due to the geometry of high-dimensional surfaces. I will briefly outline problems associated with sampling from the typical set [49] for a more in-depth explanation.

The typical set can be thought of as the region in parameter space that has a non-negligible contribution to any expectations. In high dimensional surfaces, the neighborhood around the mode of the distribution contains very large densities relative to areas outside of it. This means that the areas outside of this neighborhood will not have a significant contribution to any expectations, however the neighborhood around the mode has a very low volume relative to the area outside of it.

This results in the mode itself having increasingly negligible contribution to any expectations as the dimensionality of the problem increases. This means that there is generally a very small region of parameter space where these two quantities (the volume and the density) are balanced such that there will be any significant contribution. This region is known as the typical set and is the region of parameter space where one should focus the computation resources and draw samples from. In practice, this means designing samplers that can find and explore the typical set efficiently.

Exploring the typical set is generally difficult when the parameter space is complicated, e.g. if there are non-trivial correlations between parameters or there are many distant local optima (known as a multi-modal problem). In the case of non-trivial correlations, the sampler may propose many either *bad* points, i.e. ones that are unlikely to be from the typical set if it does not walk along these correlation, or propose points very near to the current position which makes the positions correlated and hence breaks the *memoryless* condition. This manifests in producing a large integrated auto-correlation time is longer, and hence, it takes the chain longer to move to a statistically independent point. In this case, one must accept many of these points before one independent sample has been generated. In the case of multi-modal of these cases, the chains may spend a long time *stuck* in a local optimum before moving on to find the global minima.

Therefore the key to designing a good sampler is to create a proposal distribution that is able to both explore the typical set, (with multiple chains this is often referred to as mixing) and collect many *independent* samples as quickly as possible.

Nested sampling [45] has also been a popular method for parameter estimation in the field of gravitational-waves. Nested sampling is primarily an algorithm to estimate the evidence term, i.e. the denominator in 1.14. It can however be utilised to draw samples from the posterior. The algorithm works by distributing a set of points across the posterior surface according to the prior, an evidence is then computed by the contour created with these initial points. This first estimate is gradually refined by removing the point on the contour edge with the lowest likelihood value and drawing a new point subject to the constraint that the likelihood is larger than the previous. This procedure is carried out literately until the evidence estimate converges to a stable contour which should encapsulate the typical set as long as the initial points are sufficient to contain the typical set. Posterior distributions can be created by storing the discarded points, these will be drawn according to the posterior as long as they are weighted by the relative stages at which they are discarded.

The gravitational wave parameter space is generally a difficult parameter space to sample from the typical set has non-trivial correlations, multi-modal distributions and moderately high dimensionality ($\sim$ 15 dimensions depending upon the approximant). Because of this difficulty, much work has been done on designing

software that is able to reliably produce source parameter estimates for the detections [50, 51]. Bayesian inference packages such as LALinference, Bilby and PyCBCInference [52, 53, 54] have been developed to address this problem and implement techniques such as gravitational-wave specific proposal distributions and parallel tempering [55, 56]. These tools have proven to be incredibly powerful and have been vital components of the gravitational-wave data analysis efforts in the previous decade as well as much of the analysis presented in this Thesis.

1.4 Summary of Thesis

At the time of writing, gravitational-wave astronomy is progressing rapidly, there have been three successful observing runs which have produces several catalogs of GW events [2, 57]. The number of detections means that gravitational-wave astronomy is now moving towards an era where it can provide answers to many open questions such as the validity/limits of General Relativity [58], astrophysical population properties [59], and many others in physics and astrophysics [60, 61]. Looking forward to the fourth observing run (O4) and beyond, and as the number of events grows, these questions are most likely to be answered by combining the information from many events. This however brings challenges, in particular, the computational and data analysis techniques will have to become more efficient to accommodate the growing number of detections.

This thesis attempts to tackle some of these computational problems, the work can be broadly broken down into two sections which look at two different ways of making a difficult computational challenge tractable; we look at specific examples but we address these problems using broader themes that may have applications beyond the examples shown here.

If a problem is intractable then one way to address this is to re-frame the problem as an easier one using approximations that lose very little information, in the first section we use traditional analytical insights into the mathematics of the problem and re-frame the problem of measuring precession into a considerably easier one the two harmonic approximation.

If it is not clear how to make a problem simpler then one must resort to data-driven methods, this could include using more efficient algorithms or using optimized hardware. In the later sections, we exploit state-of-the-art machine learning methods and apply these to gravitation-wave data analysis. These methods generally provide tools that can either give us new information and insight from the data or provide the same information considerably more efficiently.

First, we look at the problem of measuring a phenomenon known as precession in compact binary coalescence (CBCs), this is a direct prediction of GR which has not clearly been detected in any of the events thus far. Precession is well understood from a theoretical perspective [62] however prior to this work it was not well understood

exactly where in the parameter space we were likely to see precession when we were likely to see precession and more importantly there was no simple metric to determine whether a signal was precessing. The standard way to do this prior to this work was to use a Bayes factor between a precessing and non-precessing model. This significantly increases the computational cost of carrying out any large-scale studies into precession.

In chapter 2, we first discuss a method to describe a precessing waveform as a harmonic decomposition, where each harmonic is a simple non-precessing waveform. We then show that for the vast majority of signals we can well approximate a processing waveform using only the first two leading order harmonics. This then motivates the two harmonic approximations. With this approximation we are able to re-frame much of the existing analysis in a simpler way, if there is measurable power in the sub-dominant harmonic then the waveform is not well described by a non-precessing waveform and therefore this suggests that there is measurable precession in the system.

In chapters 3 and 4 we then use the two harmonic decomposition to carry out in-depth studies into the *measurability* of precession, we look at where we are likely to measure precession and also carry out a population analysis which predicts how often we are likely to measure precession. Both of these studies would not have been practical prior to the two harmonic formulation of the problem.

Machine learning has advanced rapidly in the previous decades, these methods have revolutionized many fields and are now starting to make significant advances in the natural sciences and astronomy [63, 64, 65]. The remaining chapters in this thesis apply several different machine learning methods to tackle data analysis problems in gravitational-wave astronomy.

As mentioned above, much of the exciting science in gravitational-wave astronomy will be derived from combining the information from many detections. The sampling routines produce this information in the form of discrete samples from the posterior surface. These samples can be combined to produce uncertainty estimates about source parameters however they are often not suitable for population analysis. For analysis like that, we need to derive a continuous density surface from these discrete samples. In chapter 5, we propose a method that uses Gaussian Processes to interpolate these samples and obtain a continuous density estimate across this surface. As well as producing a point estimate for the density, we show that we are able to incorporate uncertainty into our analysis. This can then be incorporated into downstream analysis such as population studies.

In chapter 6 we exploit recent developments in waveform surrogate modeling [66, 67] to make Bayesian inference more efficient. We show that using advances in waveform modelling and GPUs, we can potentially perform parameter estimation much more quickly. We present two methods to do this, firstly we use simple random walk MCMC but run this in large batches as a vectorised operation. This batching

can produce huge speed ups in efficiency, meaning we are able to produce many more samples per second than is possible otherwise. We then present a method to perform Hamiltonian Monte Carlo, [68, 49], which uses Automatic Differentiation [69, 70] to compute the gradient of the likelihood surface using no approximate or numerical gradient calculations. This provides a slower (in terms of samples per second) but more efficient (in terms of effective samples per likelihood evaluation) sampler. We briefly compare these methods and then point to future directions that would likely improve this further.

Finally in chapters 7 and 8 we look at very important questions when using machine learning algorithms; can we trust them? And can we trust the confidence and uncertainty estimates that they provide? Having reliable uncertainty estimates is essential if these data analysis techniques are to be adopted into the field of gravitational-wave astronomy. A closely related problem is the question of whether a model understands its *domain of validity* and it is able to understand situations where it is not able to produce reliable predictions. In this section I show that often machine learning methods do not incorporate uncertainty properly and therefore are not able to produce reliable and trustworthy predictions. I then present a novel algorithm which can be applied (with some slight differences) for both classification and regression problems and allows one to properly account for uncertainty when using ML methods for prediction.

Chapter 2

Precession and the Harmonic Decomposition

2.1 Introduction

For completeness and in order to cover the relevant background material for later chapters, this chapter reviews the theory of precession, in particular focusing on the harmonic decomposition that was presented in [3].

When the spins of black holes in a binary system are mis-aligned with the binary's orbital angular momentum, both the spins and orbital angular momentum will precess [62, 71, 72, 73]. We therefore expect that most astrophysical binaries will undergo precession, but to date there has been no evidence of precession in gravitational-wave (GW) observations from the Advanced LIGO and Virgo detectors [2, 74]. This is not necessarily surprising, because precession often leaves only a weak imprint on the observable signal, particularly when the black holes are of comparable mass and the binary's orbit is face-on to the detector, which are the most likely configurations that have been observed so far. Despite this heuristic picture, there is no simple means to estimate the measurability of precession of a given binary configuration, and as such it is difficult to predict when precession effects will be conclusively observed in GW events.

Detailed parameter estimation techniques have been developed, which enable the reconstruction of the parameters of observed signals [50, 51, 52, 54, 75], in addition to approximate Fisher-matrix methods [76, 77]. In parallel, techniques have been developed that provide an intuitive understanding of the measurement accuracy of certain parameters (or parameter combinations) [78, 79, 80, 81, 82, 83]. These have typically involved either approximations (such as leading order, Fisher Matrix type calculations), restriction to a subset of system parameters (for example masses and spins; timing and sky location; binary orientation). Combined, these give an understanding of the accuracy of parameter estimation for non-precessing systems.

In parallel, there have been significant developments in understanding the impli-

cations of precession, starting with the early work in Refs. [62, 72, 73] which provided insights into the impact of precession on the gravitational waveform emitted during the inspiral of compact binaries. Subsequently, black hole binary waveforms which incorporate precession through merger have been developed [84, 85, 86, 87, 88]; large scale parameter estimation studies of precession have been performed to identify the regions of parameter space where precession will be observable [89, 90, 91, 92, 93, 94, 95, 96]; and new theoretical insights into the impact of precession on both detection and parameter estimation have been obtained [97, 98, 99]. Complementary to this, there have been several efforts to understand the impact of precession on searches [100, 97], and to implement searches for precessing signals [72, 101, 102, 103, 73]. This has led to an increasingly clear picture of the impact of precession: it is most significant for binaries with large mass ratios, where the in-plane spin components are large and for systems where the total angular momentum is mis-aligned with the line of sight.

At leading order, the gravitational waveform emitted by a precessing binary is composed of five harmonics, which are offset by multiples of the precession frequency [84, 99]. We show that these harmonics form a natural hierarchy with the amplitude of the sub-leading harmonics suppressed by a factor that depends upon the opening angle (the angle between the orbital and total angular momenta). Using this approximation, and restricting to the two leading harmonics, we are able to obtain relatively simple expressions for the precession waveform. Each harmonic takes the form of a non-precessing-binary waveform (i.e., with monotonic amplitude and frequency evolution during the inspiral of non-eccentric systems), and the amplitude and phase modulations of the complete precessing-binary waveform arise as beating between the two harmonics.

The purpose of this chapter is to introduce this decomposition (Sec. 2.3), with an alternative derivation given in the Appendix, and the two-harmonic approximation (Sec. 2.4), and to identify its range of validity and accuracy (Sec. 2.5). Then a proposed search for precessing binaries is discussed using the two-harmonic approximation (Sec. 2.6) and finally introduce the notion of a "precession SNR" that can be used to determine whether precession effects are observable in a given system (Sec. 2.7). We begin in the next section with a summary of precession in black-hole binaries.

2.2 Black hole Spin Induced Precession

In the general theory of relativity a binary consisting of two objects of masses, m_1 and m_2 (where we choose $m_1 \geq m_2$ and denote $q = m_1/m_2$, so that $q \geq 1$), with spin angular momenta $\mathbf{S_1}$ and $\mathbf{S_2}$, orbiting each other with angular momentum $\mathbf{L}$, will slowly inspiral due to the loss of energy and momentum through the emission of gravitational waves. If $\mathbf{S}_1 \parallel \mathbf{S}_2 \parallel \mathbf{L}$, then the plane of the orbit remains fixed and

in non-eccentric binaries the amplitude and frequency of the emitted gravitational wave increases as the orbital separation decreases. The system eventually merges and forms a single perturbed black hole that emits gravitational radiation as a superposition of quasinormal ringdown multipoles, until the system settles down to its final state [104].

For the case where the total spin is not aligned with the total orbital angular momentum, $(\mathbf{S}_1 + \mathbf{S}_2) \times \mathbf{L} \neq 0$, in most cases the orbital plane of the binary will precess around the approximately constant total angular momentum $\mathbf{J} = \mathbf{S_1}+\mathbf{S_2}+\mathbf{L}$, i.e., $\mathbf{L}$ precesses around $\mathbf{J}$, and the spins precess such that $\dot{\mathbf{S}} = -\dot{\mathbf{L}}$ [62]. For configurations where $\mathbf{J} \approx 0$, the system undergoes "transitional precession", this results in the binary losing its gyroscopic axis and tumbling chaotically through space [62, 71], but these dynamics are generally short lived (it is a transitional phase between simple precession phases) and are expected to be rare in LIGO-Virgo detections. The angle between $\mathbf{L}$ and $\mathbf{J}$ is denoted by β. In simple precession cases β slowly increases during inspiral as L decreases (recall that in the Newtonian limit $L \propto \sqrt{r}$, where r is the orbital separation), but the spin magnitudes S_1 and S_2 remain fixed, and, to a good approximation, so do their orbit-averaged components parallel and perpendicular to L, $S_{i||}$ and $S_{i\perp}$. The opening angle β typically varies very little over the portion of a binary's inspiral that is visible in a GW detector, and so it is often possible to make the approximation that β is constant. This approximation has been used to good effect in Ref. [97], and we will also use it in some of the discussion in this chapter.

Adopting the notation that the inclination angle of the binary as seen by an observer, ι, is the angle between the orbital angular momentum and the line of sight (see Fig.2.1), $\cos\iota = \hat{\mathbf{L}} \cdot \hat{\mathbf{N}}$, where a caret denotes a unit vector (e.g. $\hat{\mathbf{a}} = \mathbf{a}/|\mathbf{a}|$), the binary's orbital inclination becomes times dependent. As a result the energy emitted in GWs in the $\hat{\mathbf{N}}$ direction will also be time dependent, where the maximum instantaneous energy emission is approximately in the direction of $\hat{\mathbf{L}}$. If $\hat{\mathbf{N}}$ is aligned with $\hat{\mathbf{J}}$, then $\iota \approx \beta$ and varies slowly and with minimal oscillations due only to orbital nutation. If $\hat{\mathbf{N}}$ is in some other direction, then the energy emission will be modulated on the precession timescale. In the following we will not use the inclination ι, but rather combinations of β and the angle between $\mathbf{J}$ and $\hat{\mathbf{N}}$, denoted by $\theta_{\rm JN}$. As noted previously, $\hat{\mathbf{J}}$ is approximately constant for simple precession cases, and we will treat it as a constant in the analysis in Sec. 2.3.

The signal measured in a detector will exhibit modulations in phase and amplitude that depend on β, $\theta_{\rm JN}$, the precession angle of $\mathbf{L}$ around $\mathbf{J}$, denoted by α, and the polarisation ψ of the observed signal. These angles are illustrated in Fig. 2.1, and discussed further in Sec. 2.3. For now we note several well-known features of precession waveforms [62, 71], which will be further sharpened in the discussion later in the chapter. The strength of precession in a system is characterised by the degree of tilt of the binary's orbit, given by β, and by the precession frequency Ω_P of $\mathbf{L}$

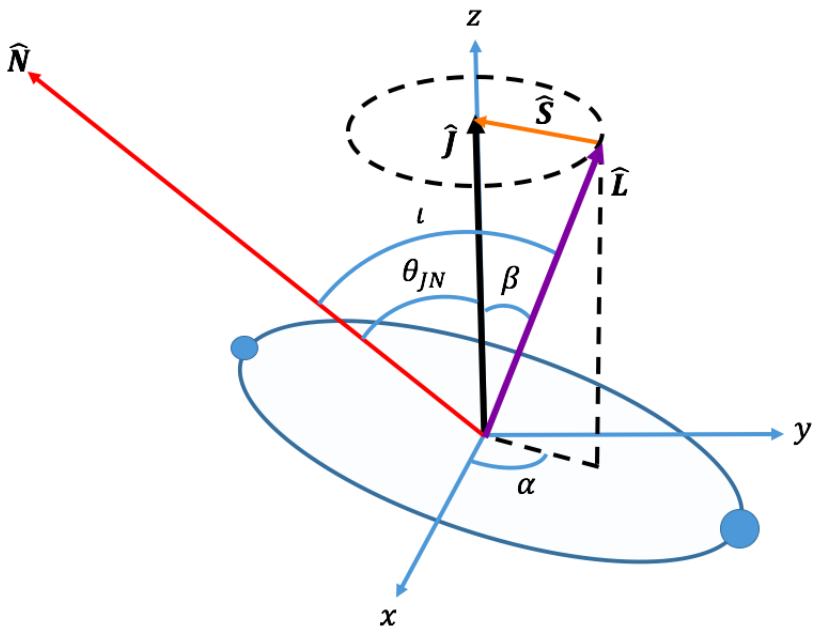

Figure 2.1: Plot showing how the precession angles used in this study are defined in the J-aligned frame. The normal vector here indicates the line of sight of the observer, $\hat{\mathbf{L}}$ and $\hat{\mathbf{J}}$ are the orbital angular momentum and total angular momentum vectors respectively, S_{1x}, S_{1y} and S_{1z} are the x, y and z components of the spin on the larger black hole.

around $\mathbf{J}$, which is given by

$$\Omega_p = \dot{\alpha}\,. \tag{2.1}$$

The angle β is determined primarily by the total spin in the plane, and binary's mass ratio and separation. At leading order we can write the orbital angular momentum of the system as $L = \mu\sqrt{Mr}$, where μ is the reduced mass, $\mu = m_1 m_2/M = qM/(1+q)^2$, and so to first approximation,

$$\tan\beta = \frac{S_\perp}{\mu\sqrt{Mr} + S_\parallel}, \tag{2.2}$$

which provides us with the basic dependence of β on the binary configuration. At leading order the precession frequency can be written as,

$$\Omega_p \approx \left(2 + \frac{3}{2q}\right)\frac{J}{r^3}, \tag{2.3}$$

meaning that to first approximation it *does not* depend on the spins (or therefore the opening angle β), but only on the binary's total mass, mass-ratio, and separation (or equivalently orbital frequency). The number of precession cycles over a certain time or frequency range (e.g., over the course of an observation), depends on the total mass and mass-ratio of the binary. In a GW observation there is a partial

degeneracy between the mass ratio and the aligned spin $S_{||}$ [105, 76, 79], meaning that one of the chief effects of a measurement of precession will be to improve the measurement of these two physical properties [92].

In the remainder of this chapter we choose to describe the gravitational wave signal, precessing or non-precessing, with the `IMRPhenomPv2` phenomenological model presented in Ref. [84]. This model exploits the phenomenology of simply precessing binaries described earlier, with the additional approximation that a precessing-binary waveform can be factorised into an underlying non-precessing waveform, and the precessional dynamics [106]. The underlying non-precessing-binary model is `IMRPhenomD` [107, 108], using only the spin components aligned with $\mathbf{L}$. In `IMRPhenomD` both aligned spin components are used to generate an approximate post-Newtonian phasing and amplitude, with corrections provided by fits to numerical-relativity waveforms, that are parameterised by two different combinations of the two spin components. Although `IMRPhenomD` has been found to model well two-spin systems [109], its dominant spin dependence can be characterised well by the effective spin,

$$\chi_{\rm eff} = \frac{1}{M}\left(\frac{\mathbf{S_1}}{m_1} + \frac{\mathbf{S_2}}{m_2}\right)\cdot\hat{\mathbf{L}}, \tag{2.4}$$

which takes values between -1 (both maximal anti-aligned spins) and $+1$ (both maximal aligned spins) to describe the magnitude of spin aligned with the total angular momentum. For a given configuration `IMRPhenomPv2` uses the corresponding `IMRPhenomD` waveform, but with the final spin modified to take into account the in-plane spin components. A frequency-dependent rotation is then applied to the non-precessing waveform to introduce the precession dynamics, which are modelled by frequency-domain post-Newonian expressions for the precession angles for an approximately equivalent single-spin system [110, 84], where the large black hole has spin,

$$\chi_{\rm p} = \frac{1}{A_1 m_1^2}\max\left(A_1 S_{1\perp}, A_2 S_{2\perp}\right), \tag{2.5}$$

where $A_1 = 2 + 3q/2$ and $A_2 = 2 + 3/(2q)$ and $\mathbf{S}_{i\perp}$ is the component of the spin perpendicular to $\mathbf{L}$. The effective precession spin parameter is obtained by averaging the relative in-plane spin orientation over a precession cycle, and so more accurate for a system that undergoes many precession cycles. There are several important features which are not incorporated in the `IMRPhenomPv2` waveform. These include two-spin effects [85, 111, 112], gravitational wave multipoles other than the leading 22 mode [86], significant precession during merger [113], and spin alignment due to spin-orbit resonances during inspiral [114, 115]. Some of these effects will have an impact upon the distributions of black hole spin orientations when the binaries enter the LIGO or Virgo sensitivity band while others can leave imprints on the waveform which may be observable, particularly close to the merger. Nonetheless, the `IMRPhenomPv2` has been used in the analysis of all LIGO-Virgo observations

during the first two observing runs [116, 117, 74, 2], and it captures much of the dominant phenomenology of precessing-binary waveforms. In addition, the decomposition presented in the next section is in no way tied to the particular waveform used and could be equally well applied to other waveform models for precessing binaries which, for example, incorporate two-spin effects and precession during merger. The current formalism does not include additional gravitational wave multipoles, and we will investigate this in a future work. We expect the broad features of many of the results presented in the remainder of the chapter to be relatively unaffected by the specific waveform choice, but the details for any specific signal could change.

2.3 Harmonic decomposition of the waveform from a precessing binary

The gravitational waveform emitted by a precessing system, as observed at a gravitational wave detector, can be expressed approximately as [73, 97]

$$\begin{aligned} h(t) &= -\left(\frac{d_o}{d_L}\right) A_o(t)\,\mathrm{Re}\Big[e^{2i\Phi_S(t)} \\ &\qquad \big(F_+(C_+ - iS_+) + F_\times(C_\times - iS_\times)\big)\Big]\,. \end{aligned} \tag{2.6}$$

Here, $A_o(t)$ denotes the amplitude of the gravitational wave signal in a (time-varying) frame aligned with the orbital angular momentum of the binary, and depends upon the masses and spins of the binary. Since the amplitude scales linearly with the luminosity distance, we have chosen to introduce a fiducial normalization $A_o(t)$ for a waveform at a distance d_o and explicitly extract the distance dependence.[1] Φ_S is the phase evolution in the frame aligned with the total angular momentum $\mathbf{J}$ of the binary. The phase evolution, Φ_S, is related to the orbital phase, ϕ_{orb}, as

$$\Phi_S(t) = \phi_{orb}(t) - \epsilon(t) \tag{2.7}$$

where [118]

$$\dot{\epsilon}(t) := \dot{\alpha}(t)\cos\beta(t) \tag{2.8}$$

and, as before, β is the opening angle and α gives the phase of the precession of $\mathbf{L}$ around $\mathbf{J}$ as shown in Fig. 2.1. F_+ and $F_\times$ give the detector response relative to the $\mathbf{J}$-aligned frame and $C_{+,\times}$, $S_{+,\times}$ encode the time-varying response to the gravitational wave due to the evolution of the binary's orbit relative to the detector. They depend upon the three angles introduced previously: the precession opening angle β and phase α and the angle between the total orbital angular momentum

[1] Of course, the observed waveform is also affected by the redshifting of frequencies. For the calculation discussed here, we work in the detector frame and consider the *observed* masses, which are $(1+z)$ times the source frame masses.

and the line of sight $\theta_{\rm JN}$. In terms of these angles, we can express $C_{+,\times}$ and $S_{+,\times}$ as[2]

$$\begin{aligned}
C_+ &= -\left(\frac{1+\cos^2\theta_{\rm JN}}{2}\right)\left(\frac{1+\cos^2\beta}{2}\right)\cos 2\alpha \\
&\quad -\frac{\sin 2\theta_{\rm JN}}{2}\frac{\sin 2\beta}{2}\cos\alpha - \frac{3}{4}\sin^2\theta_{\rm JN}\sin^2\beta, \\
S_+ &= \left(\frac{1+\cos^2\theta_{\rm JN}}{2}\right)\cos\beta\sin 2\alpha + \frac{\sin 2\theta_{\rm JN}}{2}\sin\beta\sin\alpha, \\
C_\times &= -\cos\theta_{\rm JN}\left(\frac{1+\cos^2\beta}{2}\right)\sin 2\alpha - \sin\theta_{\rm JN}\frac{\sin 2\beta}{2}\sin\alpha, \\
S_\times &= -\cos\theta_{\rm JN}\cos\beta\cos 2\alpha - \sin\theta_{\rm JN}\sin\beta\cos\alpha.
\end{aligned} \tag{2.9}$$

The non-precessing expressions can be recovered in the limit of $\beta \to 0$ and $\alpha \to$ constant (which is then degenerate with the polarization of the system). When β is non-zero, the effect of precession is to modulate the detector response at frequencies Ω_P and $2\Omega_P$. To make the harmonic content of $C_{+,\times}$ and $S_{+,\times}$ more explicit, we first introduce the parameter,

$$b = \tan\left(\beta/2\right), \tag{2.10}$$

and write the response functions in terms of it. The terms involving β can be expressed as

$$\begin{aligned}
\frac{1+\cos^2\beta}{2} &= \frac{1+b^4}{(1+b^2)^2}, \\
\cos\beta &= \frac{1-b^4}{(1+b^2)^2}, \\
\frac{\sin 2\beta}{2} &= \frac{2b(1-b^2)}{(1+b^2)^2}, \\
\sin\beta &= \frac{2b(1+b^2)}{(1+b^2)^2}, \\
\sin^2\beta &= \frac{4b^2}{(1+b^2)^2}\,.
\end{aligned} \tag{2.11}$$

Substituting the trigonometric identities from Eq. (2.11) into the expressions for C_+ and S_+ in Eq. (2.9) we obtain,

$$\begin{aligned}
\left(\frac{d_o}{d_L}\right)(C_+ - iS_+) &= -e^{2i\alpha}\sum_{k=0}^{4}\mathcal{A}_k^+\left[\frac{b^k e^{-ik\alpha}}{(1+b^2)^2}\right], \\
\left(\frac{d_o}{d_L}\right)(C_\times - iS_\times) &= ie^{2i\alpha}\sum_{k=0}^{4}\mathcal{A}_k^\times\left[\frac{b^k e^{-ik\alpha}}{(1+b^2)^2}\right],
\end{aligned} \tag{2.12}$$

[2] We have re-written the C_+ term relative to what is normally given in the literature, e.g. [73, 97], to group terms with the same α dependence.

where we have introduced $\mathcal{A}_k^+$ and $\mathcal{A}_k^\times$ as

$$\begin{aligned}
\mathcal{A}_0^+ = \mathcal{A}_4^+ &= \frac{d_o}{d_L}\left(\frac{1+\cos^2\theta_{\rm JN}}{2}\right), \\
\mathcal{A}_0^\times = -\mathcal{A}_4^\times &= \frac{d_o}{d_L}\cos\theta_{\rm JN}, \\
\mathcal{A}_1^+ = -\mathcal{A}_3^+ &= 2\frac{d_o}{d_L}\sin\theta_{\rm JN}\cos\theta_{\rm JN}, \\
\mathcal{A}_1^\times = \mathcal{A}_3^\times &= 2\frac{d_o}{d_L}\sin\theta_{\rm JN}, \\
\mathcal{A}_2^+ &= 3\frac{d_o}{d_L}\sin^2\theta_{\rm JN}, \\
\mathcal{A}_2^\times &= 0\,.
\end{aligned} \tag{2.13}$$

In the approximation where the direction of total angular momentum is constant, the $\mathcal{A}_k^{+,\times}$ are time independent amplitudes, and the time dependence of the amplitude functions is captured as a power series in the parameter $b = \tan(\beta/2)$.

Finally, we can use the harmonic decomposition in Eq. (2.12) to obtain a decomposition of the waveform, Eq. (2.6),

$$\begin{aligned}
h(t) \;=\; {\rm Re}\Bigg[&\left(\frac{A_o(t)e^{2i(\Phi_S+\alpha)}}{(1+b^2)^2}\right) \\
&\sum_{k=0}^{4}(be^{-i\alpha})^k(F_+\mathcal{A}_k^+ - iF_\times\mathcal{A}_k^\times)\Bigg]\,.
\end{aligned} \tag{2.14}$$

This allows us to clearly identify the impact of precession on the waveform. First leads to an additional phase evolution, 2α (which is related to the frequency Ω_P using equation 2.1) and a decrease in the amplitude by a factor $(1+b^2)^2$. The precessing waveform contains five harmonics that form a power series in b, whose amplitude depends upon the detector response, distance and viewing angle of the binary. The frequency of each harmonic is offset from the next by the precession frequency Ω_P. Similar results have been obtained previously, by manipulating the spin-weighted spherical harmonic decomposition of the waveform, e.g. [98, 99]. However, it was not previously observed that the relative amplitudes of the harmonics were related in a straightforward manner. In the Appendix, we present an alternative derivation of the result in Eq. (2.14) in terms of this spin-weighted spherical harmonic decomposition of the waveform, as is customary when producing waveform models for precessing binaries [84].

As a final step, we would like to explicitly extract three more time-independent angles that characterize the waveform, namely the polarization angle ψ, the initial phase ϕ_o and the initial polarization phase α_o.[3]

The unknown polarization ψ is currently folded into the detector response func-

[3]The initial polarization phase α_o is sometimes denoted in the literature as ϕ_{JL}.

tions $F_{+,\times}$. It is more useful to extract ψ and then consider the detector response to be a *known* quantity dependent upon only the details of the detector and the direction to the source. Thus, we write the detector response as,

$$\begin{aligned} F_+ &= w_+ \cos 2\psi + w_\times \sin 2\psi, \\ F_\times &= -w_+ \sin 2\psi + w_\times \cos 2\psi, \end{aligned} \tag{2.15}$$

where w_+ and $w_\times$ are the detector response functions in a fixed frame — for a single detector it is natural to choose $w_\times = 0$ and for a network to work in the dominant polarization for which w_+ is maximized [119]. The unknown polarization of the source relative to this preferred frame is denoted ψ.

To isolate the initial orbital and precession phases, we explicitly extract them from the binary's phase evolution by introducing,

$$\begin{aligned} \Phi(t) &:= \Phi_S(t) - \phi_o + \alpha(t) - \alpha_o \\ &= \phi_{\rm orb}(t) - \phi_o + \int_{\alpha_o}^{\alpha(t)} \frac{2b^2}{1+b^2}\, d\alpha\,. \end{aligned} \tag{2.16}$$

Thus $\Phi(t)$ vanishes at $t = 0$ and evolves as the sum of the orbital phase and an additional, precession dependent, contribution.

We then substitute the expressions for $F_{+,\times}$, Eq. (2.15), and Φ, Eq. (2.16), into the expression for $h(t)$ given in Eq. (2.14), and isolate the time-varying terms from the constant, orientation dependent angles. The waveform can be written as the sum of five precessing harmonics, the amplitudes of which are constants that depend upon the binary's sky location, distance and orientation:

$$h = \sum_{k=0}^{4} w_+ (h_0^k \mathcal{A}_k^1 + h_{\frac{\pi}{2}}^k \mathcal{A}_k^3) + w_\times (h_0^k \mathcal{A}_k^2 + h_{\frac{\pi}{2}}^k \mathcal{A}_k^4), \tag{2.17}$$

where $h^k_{0,\frac{\pi}{2}}$ are the waveform harmonics and $\mathcal{A}_k^\mu$ are constants. The waveform harmonics are

$$\begin{aligned} h_0^k(t) &= {\rm Re}\left[A_o(t) e^{2i\Phi} \left(\frac{b^k e^{-ik(\alpha-\alpha_o)}}{(1+b^2)^2} \right) \right], \\ h_{\frac{\pi}{2}}^k(t) &= {\rm Im}\left[A_o(t) e^{2i\Phi} \left(\frac{b^k e^{-ik(\alpha-\alpha_o)}}{(1+b^2)^2} \right) \right]. \end{aligned} \tag{2.18}$$

The harmonics form a simple power series in $be^{-i\alpha}$, so the amplitude of each successive harmonic is reduced by a factor of b, and the frequency is reduced by Ω_P.

The amplitudes for the harmonics are given by

$$\begin{aligned}
\mathcal{A}^1_k &= \mathcal{A}^+_k \cos\phi_k \cos 2\psi - \mathcal{A}^\times_k \sin\phi_k \sin 2\psi, \\
\mathcal{A}^2_k &= \mathcal{A}^+_k \cos\phi_k \sin 2\psi + \mathcal{A}_\times \sin\phi_k \cos 2\psi, \\
\mathcal{A}^3_k &= -\mathcal{A}^+_k \sin\phi_k \cos 2\psi - \mathcal{A}^\times_k \cos\phi_k \sin 2\psi, \\
\mathcal{A}^4_k &= -\mathcal{A}^+_k \sin\phi_k \sin 2\psi + \mathcal{A}^\times_k \cos\phi_k \cos 2\psi,
\end{aligned} \tag{2.19}$$

where the $\mathcal{A}^{+,\times}$ were introduced in Eq. (2.13), ψ is the polarization and the phase angle for each harmonic is,

$$\phi_k = 2\phi_o + (2-k)\alpha_o\,. \tag{2.20}$$

These amplitudes form a generalization of the $\mathcal{F}$-statistic decomposition of the non-precessing binary waveform (see e.g. [119]). In the limit that $b \to 0$, the precessing decomposition reduces to the standard expression for the non-precessing waveform as the amplitude for all harmonics other than $k = 0$ vanish.

The precessing waveform can equally well be written in the frequency domain by performing a Fourier transform of the time-domain expressions given above [120]. In this case, Eq. (2.17) is unchanged, as are the constant amplitude terms in Eq. (2.19). The frequency dependent harmonics are simply the Fourier transform of the time-domain modes given in Eq. (2.18), and naturally satisfy $h^k_{\frac{\pi}{2}} = i h^k_0$.

The expansion above is most natural when $b < 1$, which corresponds to opening angles of $\beta < 90^\circ$. In cases where the opening angle is greater than 90° it is natural to re-express the waveform in terms of $c = b^{-1} = \cot(\beta/2)$ in which case the waveform can be expressed as a power series in c. We will not discuss the large opening angle calculation further in this chapter, but note that many of the arguments presented below would extend in a straightforward manner to this case.

a Obtaining the harmonics

Here, we give an explicit prescription to obtain the five harmonics for the waveform, introduced in Eq. (2.17). To do so, we generate waveforms for orientations that contain only a subset of the harmonics, and combine them to isolate a single harmonic. For simplicity, we restrict attention to the + polarization by fixing $w_+ = 1$, $w_\times = 0$ and consider a binary at a distance $d_L = d_o$.

Harmonics $k = 0$ and $k = 4$. When the viewing angle of the signal is aligned with the total angular momentum, $\theta_{\rm JN} = 0$, the observed waveform contains only the zeroth and fourth harmonics as $\mathcal{A}^{+,\times}_{1,2,3}$ vanish for $\theta_{\rm JN} = 0$. We also fix $\alpha_o = 0$,

to obtain,

$$\begin{aligned} h_{\phi_o=0;\psi=0} &= h_0^0 + h_0^4, \\ h_{\phi_o=\frac{\pi}{4};\psi=\frac{\pi}{4}} &= -h_0^0 + h_0^4. \end{aligned} \tag{2.21}$$

From these, we can extract the $k = 0$ and 4 harmonics,

$$\begin{aligned} h_0^0 &= \tfrac{1}{2}\left(h_{\phi_o=0,\psi=0} - h_{\phi_o=\frac{\pi}{4},\psi=\frac{\pi}{4}}\right), \\ h_0^4 &= \tfrac{1}{2}\left(h_{\phi_o=0,\psi=0} + h_{\phi_o=\frac{\pi}{4},\psi=\frac{\pi}{4}}\right). \end{aligned} \tag{2.22}$$

The $\frac{\pi}{2}$ phases of the harmonics can be obtained in an identical way.

Harmonics $k = 1$ and $k = 3$. When the signal is edge on, the $\times$ polarization contains only the first and third harmonics. Then, fixing $\theta_{\rm JN} = \frac{\pi}{2}$ and $\psi = \frac{\pi}{4}$, we have,

$$\begin{aligned} h_{\alpha_o=0;\phi_o=\frac{\pi}{4}} &= -2\left(h_0^1 + h_0^3\right), \\ h_{\alpha_o=\frac{\pi}{2};\phi_o=0} &= -2\left(h_0^1 - h_0^3\right), \end{aligned} \tag{2.23}$$

so that,

$$\begin{aligned} h_0^1 &= -\tfrac{1}{4}\left(h_{\alpha_o=0;\phi_o=\frac{\pi}{4}} + h_{\alpha_o=\frac{\pi}{2},\phi_o=0}\right), \\ h_0^3 &= -\tfrac{1}{4}\left(h_{\alpha_o=0;\phi_o=\frac{\pi}{4}} - h_{\alpha_o=\frac{\pi}{2},\phi_o=0}\right). \end{aligned} \tag{2.24}$$

Harmonic $k = 2$. Finally, from the + polarization of the edge-on waveform, we can extract the second harmonic — in principle we could also get $k = 0$ and $k = 4$, but we have already described a method to obtain them. Fixing $\theta_{\rm JN} = \frac{\pi}{2}$ and $\psi = 0$ we have,

$$\begin{aligned} h_{\alpha_o=0,\phi_o=0} &= \tfrac{1}{2}h_0^0 + 3h_0^2 + \tfrac{1}{2}h_0^4, \\ h_{\alpha_o=\frac{\pi}{2},\phi_o=0} &= -\tfrac{1}{2}h_0^0 + 3h_0^2 - \tfrac{1}{2}h_0^4, \end{aligned} \tag{2.25}$$

so that,

$$h_0^2 = \tfrac{1}{6}\left(h_{\alpha_o=0,\phi_o=0} + h_{\alpha_o=\frac{\pi}{2},\phi_o=0}\right). \tag{2.26}$$

b Precession with varying orientation

The observable effect of precession will vary significantly with the binary orientation, as has been discussed in many previous works, for example [62, 97]. Interestingly, both the amplitude and frequency of the observed precession depends upon the viewing angle. The harmonic decomposition derived above provides a straightforward way to understand this effect. The observed amplitude and phase modulations can

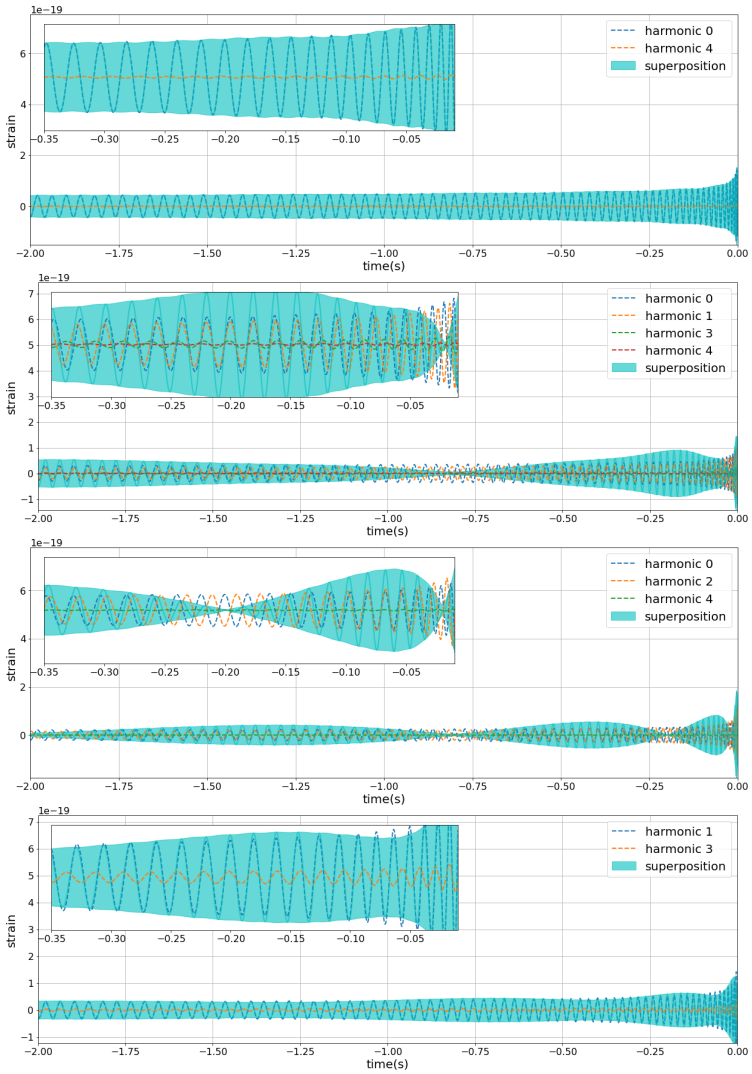

Figure 2.2: The observed waveform from a $40M_\odot$ binary with mass ratio $q = 6$, $\chi_{\rm eff} = 0$ and $\chi_p = 0.6$. The waveform is shown for four different binary orientations: $\theta_{JN} = 0$ (top; $\theta_{JN} = 45°$, $\times$ polarization (upper middle); $\theta_{JN} = 90°$, $+$ polarization (lower middle); $\theta_{JN} = 90°$, $\times$ polarization (bottom). For each waveform, the harmonics that contribute to the signal, their sum and the envelope of the full precessing waveform are shown. The insets show a zoom of a portion of the waveform to more clearly demonstrate that precession arises as a beating between the different harmonics.

be understood as the beating of the different harmonics against each other, with the amplitude of the composite waveform being maximum when the harmonics are in phase and minimum when they are out of phase.

In Fig. 2.2, we show the waveform for four different orientations: a) along $\mathbf{J}$, b) $\times$ polarization at 45° to $\mathbf{J}$, c/d) $+/\times$ polarization orthogonal to $\mathbf{J}$. In all cases, we show the last two seconds of the waveform (from around 25 Hz) for a $40M_\odot$ binary, with $q = 6$, and in-plane spin on the larger black hole of $\chi_P = 0.6$. This configuration

gives an opening angle of $\beta \approx 45^\circ$ (and $b \approx 0.4$) which leads to significant precession effects in the waveform.

When viewed along $\mathbf{J}$, there is minimal precession as only the $k = 0$ and 4 harmonics are present in the system and the $k = 4$ harmonic is down-weighted by a factor of $b^4 \approx 0.03$ relative to the leading harmonic. Furthermore, the modulation comes from the beating of the $k = 0$ and $k = 4$ harmonics and occurs at four times the precession frequency. When the line of sight is orthogonal to the total angular momentum, the $k = 0, 2, 4$ harmonics are present in the + polarized waveform and $k = 1, 3$ in the $\times$ polarization. The $k = 0$ and 2 harmonics have close to equal amplitude (although $k = 2$ is down-weighted by $b^2 \approx 0.17$, the amplitude as given in Eq. (2.13) is maximal). Consequently the observed waveform has maximal amplitude and phase modulation due to precession. For the $\times$ polarized signal, it is the $k = 1, 3$ harmonics that contribute, with $k = 3$ a factor of $b^2 \approx 0.17$ smaller than $k = 1$. Consequently, precession effects are less significant. In both cases, precession occurs at twice the precession frequency as it is from the beating of the the $k = 0$ and $k = 2$ (+ polarization) or $k = 1$ and $k = 3$ ($\times$ polarization). For the $\times$ polarized signal with $\theta_{\rm JN} = 45^\circ$, the $k = 0, 1, 3, 4$ harmonics are present, with $k = 0, 1$ dominating and having approximately equal amplitude. For this signal, the binary precesses from a face-on orientation, $\iota = 0$ to edge-on, $\iota = 90^\circ$, and the waveform amplitude oscillates from the maximum to zero. Here, modulations occur at the precession frequency.

c Importance of precession over parameter space

From the intuitive discussion of precession presented in [62, 73, 97] and summarized in Section 2.2, it is straightforward to identify regions of parameter space where precession is most likely to have a significant impact upon the binary dynamics and, consequently, the observed waveform. Specifically, we expect that higher mass ratios, larger in-plane spins and negative aligned spin components will all lead to a larger opening angle and more significant precession [97]. Here we briefly revisit this discussion, framing our results in terms of the parameter b introduced earlier. Explicitly, we introduce the waveform-averaged value of b as,

$$\overline{b} := \frac{|h^1|}{|h^0|} = \sqrt{\frac{\int df \frac{|h_1|^2}{S_n(f)}}{\int df \frac{|h_0|^2}{S_n(f)}}}\,, \tag{2.27}$$

where $h^{0,1}$ are the harmonics of the waveform introduced in Eq. (2.18) and $S_n(f)$ is the noise power spectrum of the detector. For this work, we choose $S_n(f)$ to be the design-sensitivity Advanced LIGO noise curve [2] and evaluate the integral over

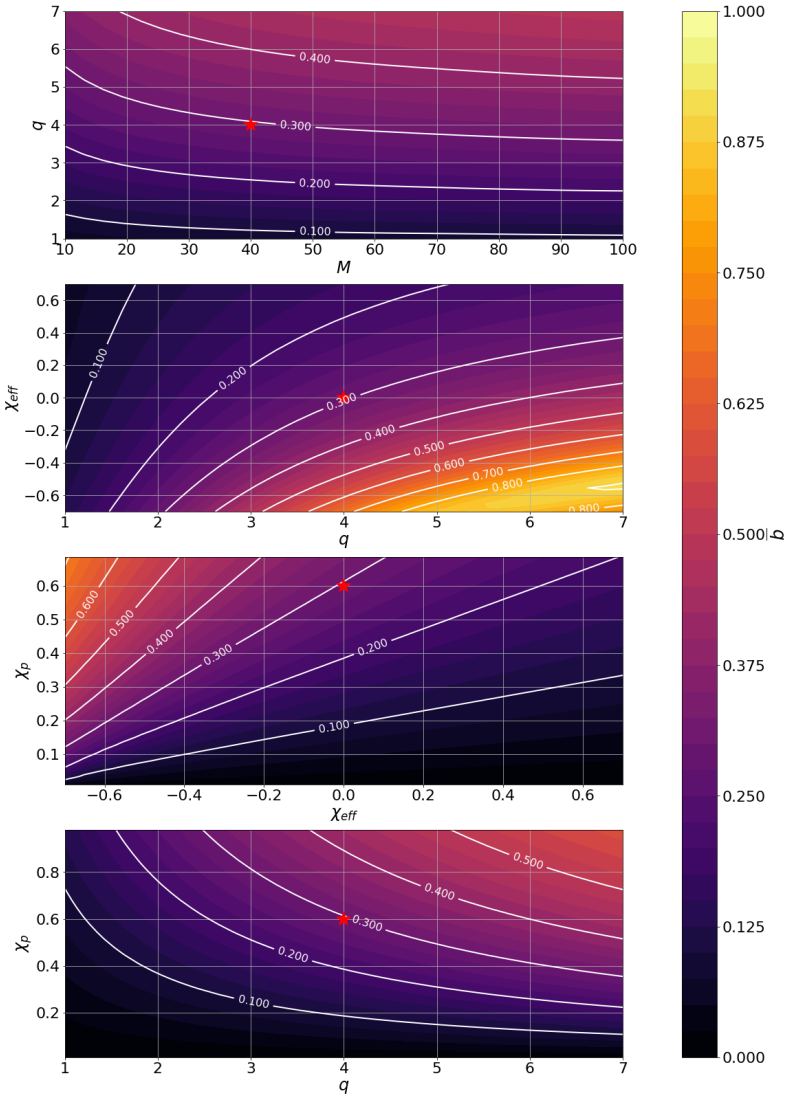

Figure 2.3: The value of $\bar{b}$ across the parameter space of total mass, mass ratio, $\chi_{\rm eff}$ and χ_p. In each figure, two of the parameters are varied while the other two are fixed to their fiducial values of $M = 40M_\odot$, $q = 4$, $\chi_{\rm eff} = 0$, $\chi_p = 0.6$ (this point is marked with a $\star$ in all the plots). The total mass has a limited impact on the value of $\bar{b}$, for masses over $M \approx 40M_\odot$; below this the $\bar{b}$ increases with mass, as the later parts of the merger are brought into the most sensitive band of the detector. The value of $\bar{b}$ is seen to increase as the mass ratio or precessing spin χ_p are increased and decrease as the aligned component of the spin $\chi_{\rm eff}$ increases. Thus, the value of b is largest for a binary with unequal masses, a large spin on the more massive component which has significant components both in the plane of the orbit and anti-aligned with the orbital angular momentum.

the frequency range $f \in [20, 1024]\,\mathrm{Hz}$ [4]. For binaries where the opening angle β is approximately constant, $\bar{b} \approx \tan(\overline{\beta}/2)$.

[4]Using a realistic noise curve similar to the observed curves during 01 and O2 would change the reported values slightly, as these noise curves are less sensitive than design, particularly at low frequencies. The qualitative patterns seen in the figure would remain the same however

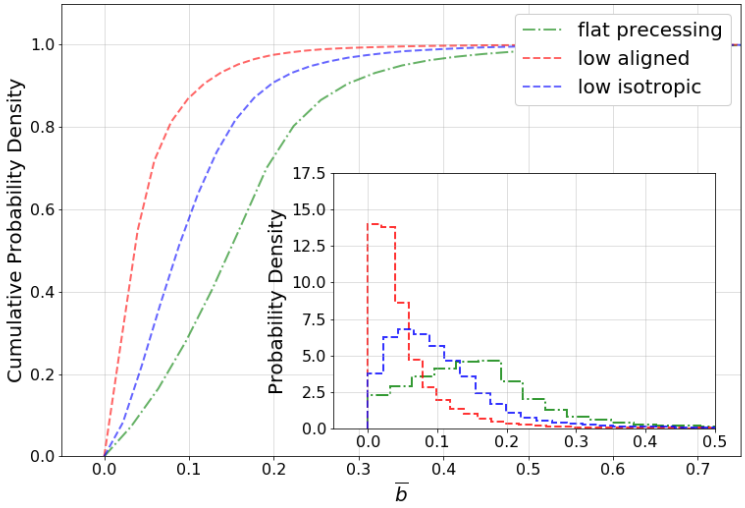

Figure 2.4: The distribution of $\overline{b}$ for a 3 different populations of binary black holes. Each population assumes either a low-isotropic, low-aligned or a flat precessing spin distribution. A power-law distribution in masses is assumed in all cases (see text for details).

Fig. 2.3 shows the value of $\overline{b}$ on several two-dimensional slices through the four dimensional parameter space of total mass M, mass ratio q, effective spin $\chi_{\rm eff}$ and precessing spin χ_p. Keeping other quantities fixed, the value of $\overline{b}$ increases with total mass. For higher masses, the late inspiral and merger occur in the sensitive band of the detectors and, close to merger, the opening angle increases as orbital angular momentum is radiated. For masses above $40M_\odot$ the mass dependence of $\overline{b}$ is small, with only a 10% decrease from $40M_\odot$ to $100M_\odot$. Thus, for the other figures, we fix $M = 40M_\odot$ and investigate the dependence of $\overline{b}$ on $q, \chi_{\rm eff}$ and χ_p. The dependence of $\overline{b}$ follows directly from Eq. (2.2). The opening angle will increase with mass ratio, as the orbital angular momentum decreases. The opening angle, and also $\overline{b}$, increase with χ_p. It follows directly from the definition that $\tan\beta$ scales linearly with χ_p, and hence approximately linearly for $b = \tan(\beta/2)$. Finally, the opening angle decreases as the effective spin $\chi_{\rm eff}$ increases, so that the largest value of $\overline{b}$ is obtained with significant spin anti-aligned with $\mathbf{J}$.

Over much of the parameter space we have explored, $\overline{b} \lesssim 0.3$. This includes binaries with mass ratio up to 4:1, with precessing spin $\chi_p \lesssim 0.6$, and zero or positive aligned spin, $\chi_{\rm eff} \geq 0$. Only a small part of parameter space has $\overline{b} > 0.4$, the value used in generating the waveforms in Figure 2.2, and $b > 0.5$ is only achieved with at least two of: a) close to maximal χ_p, b) high mass ratio, $q \gtrsim 5$ or c) significant spin anti-aligned with the orbital angular momentum $\chi_{\rm eff} \lesssim -0.4$.

Next, we consider the importance of precession for an astrophysically motivated population. In Fig. 2.4, we show the distribution of $\overline{b}$ for three distributions of black hole masses and spins. For each population, we generate 100,000 binaries

uniformly in co-moving distance, with masses drawn from a power law distribution — $p(m_1) \propto m_1^{-\alpha}$, with $\alpha = 2.35$ — and different spin distributions, which are the same as those used in Refs. [121, 122, 123]. We consider populations where the spins are preferentially low and aligned with the binary orbit; low and isotropically aligned or drawn from a flat distribution and preferentially leading to precession. A low spin distribution is a triangular distribution peaked at zero spin and dropping to zero at maximal spin while a flat distribution is a uniform between zero and maximal spin. The *aligned* distribution is strongly peaked towards aligned spins, while the *isotropic* distribution assumes a uniform distribution of spin orientations over the sphere. The *precessing* distribution is strongly peaked towards spins orthogonal to the orbital angular momentum, i.e., with significant orbital precession [124, 125]. To account for observational biases, we keep only those signals that would be observable above a fixed threshold in a gravitational wave detector. We find that even for the most extreme precessing population considered, the mean value of $\overline{b}$ is 0.15 with over 90% of binaries having $\overline{b} < 0.3$. This result is obviously sensitive to the assumptions on the mass and spin distribution. In Ref. [123] we investigate a larger number of spin distributions, including ones which allow for large spin magnitudes, and we find that the peak of the $\overline{b}$ distribution is below 0.2 and that over 90% of binaries have $\overline{b} < 0.4$ in all cases.

Fig. 2.5 shows $\overline{b}$ for a range of neutron star–neutron star and neutron star–black hole binaries. For neutron star–black hole binaries, the picture is similar to that for black hole binaries, with large values of $\overline{b}$ observed for high mass ratios and large χ_P. However, as an earlier part of the waveform is in the detector's sensitive band, the impact of precession is less observable at fixed mass ratio than for higher mass black hole binaries. For neutron star binaries, the value of $\overline{b}$ remains below 0.15 across the parameter space, and is less than 0.05 for reasonable neutron star spins, $\chi \lesssim 0.4$.

2.4 The two-harmonic approximation

The precessing waveform can be expressed as the sum of five harmonics whose amplitudes form a power series in $b = \tan(\beta/2)$. Furthermore, over the majority of the space of binary mergers, the value of b is less than 0.3. In addition, for $\overline{b} \leq 0.4$ the dominant harmonic — the one containing the most power — must be either $k = 0$ or 1. Thus, for the vast majority of binary mergers, we expect that these two harmonics will be the most significant.

This motivates us to introduce the *two-harmonic approximation*, in which we generate a waveform containing only the $k = 0$ and $k = 1$ harmonics, i.e.,

$$h = \sum_{k=0,1} w_+(h_0^k \mathcal{A}_k^1 + h_{\frac{\pi}{2}}^k \mathcal{A}_k^3) + w_\times(h_0^k \mathcal{A}_k^2 + h_{\frac{\pi}{2}}^k \mathcal{A}_k^4)\,. \tag{2.28}$$

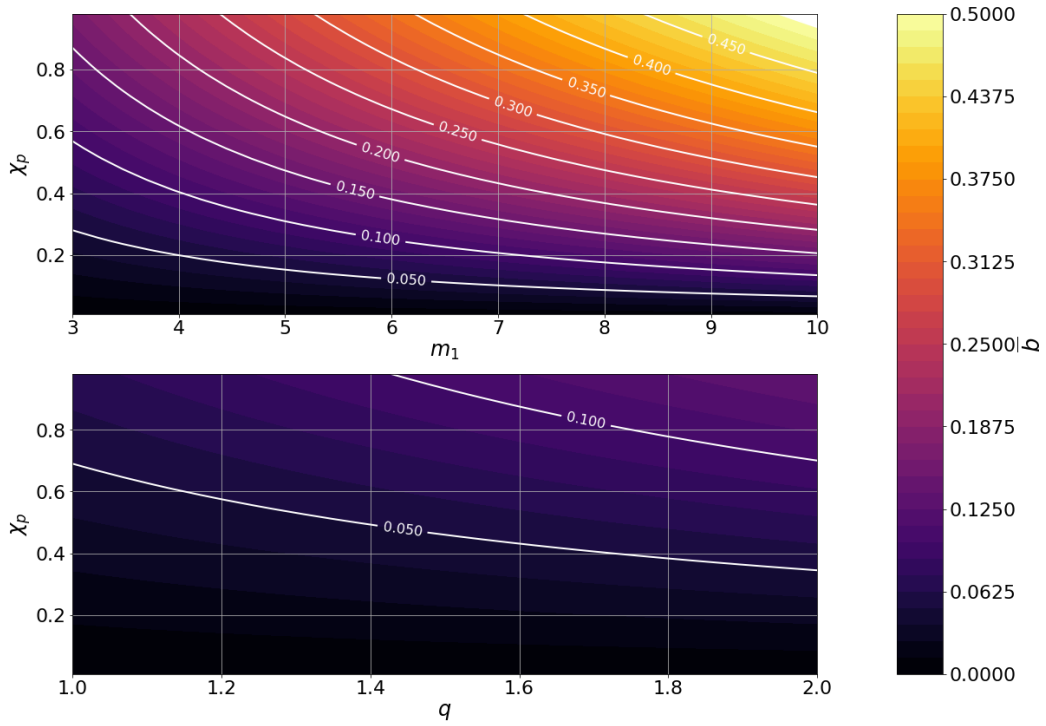

Figure 2.5: The value of $\bar{b}$ across the binary neutron star and neutron-star–black-hole space. The left figure shows the variation of $\bar{b}$ for an NSBH system with a $1.4M_\odot$ neutron star, $\chi_{\rm eff} = 0$ and varying black hole mass and χ_p. The right figure shows the variation of $\bar{b}$ against mass ratio and χ_p for a binary neutron star system of total mass $2.7M_\odot$ and $\chi_{\rm eff} = 0$.

The expression for the two-harmonic waveform can be simplified by restricting to the single detector case (i.e., setting $w_+ = 1$ and $w_\times = 0$), explicitly working with the waveform in the frequency domain, for which $h^k_{\frac{\pi}{2}}(f) = ih^k_0(f)$, and dropping the subscript 0 on the zero-phase waveform, so that $h^k(f) := h^k_0(f)$. The two harmonics of interest are,

$$h^0(f) = A_o(f)e^{2i\Phi(f)}\left(\frac{1}{(1+b(f)^2)^2}\right), \tag{2.29}$$

$$h^1(f) = A_o(f)e^{2i\Phi(f)}\left(\frac{b(f)e^{-i(\alpha(f)-\alpha_o)}}{(1+b(f)^2)^2}\right), \tag{2.30}$$

and the two-harmonic waveform then becomes,

$$h_{\rm 2harm} = \mathcal{A}_0 h^0 + \mathcal{A}_1 h^1\,, \tag{2.31}$$

where,

$$\begin{aligned}\mathcal{A}_0 =& \frac{d_0}{d_L}\left(\frac{1+\cos^2\theta_{\rm JN}}{2}\cos 2\psi - i\cos\theta_{\rm JN}\sin 2\psi\right)\times \\ & e^{-i(2\phi_o+2\alpha_o)}, \\ \mathcal{A}_1 =& \frac{d_0}{d_L}\left(\sin 2\theta_{\rm JN}\cos 2\psi - 2i\sin\theta_{\rm JN}\sin 2\psi\right)\times \\ & e^{-i(2\phi_o+\alpha_o)}.\end{aligned} \tag{2.32}$$

Thus, the two-harmonic waveform is composed of two components that have frequencies offset by Ω_P, and any observed amplitude and phase modulation of the waveform is caused by the beating of one waveform against the other. The relative amplitude and phase of the two harmonics is encoded by

$$\begin{aligned}\zeta &:= \frac{\bar{b}\mathcal{A}_1}{\mathcal{A}_0} \\ &= \bar{b}e^{i\alpha_o}\left(\frac{\sin 2\theta_{\rm JN}\cos 2\psi - 2i\sin\theta_{\rm JN}\sin 2\psi}{\frac{1}{2}(1+\cos^2\theta_{\rm JN})\cos 2\psi - i\cos\theta_{\rm JN}\sin 2\psi}\right).\end{aligned} \tag{2.33}$$

The value of ζ depends upon the viewing angle, encoded in $\theta_{\rm JN}$ and ψ, and the initial precession phase α_o. It is not difficult to show that ζ can take any value as the parameters $\theta_{\rm JN}$, ψ, α_o are varied. For example, at $\theta_{\rm JN} = 0$, $\mathcal{A}_1$ vanishes and so does ζ, while at $\theta_{\rm JN} = \pi/2$ and $\psi = \pi/4$, $\mathcal{A}_0$ vanishes and $\zeta \to \infty$. Since the initial precession phase α_o is a free parameter, the phase of ζ also can take any value. The overall amplitude and phase of the signal also depends upon the distance and coalescence phase so that any values of the amplitude and phase of the signal in the two harmonics are consistent with a signal.

2.5 Validity of the two-harmonic waveform

To investigate the validity of the two-harmonic approximation, we compare the approximate waveform with the full, five-harmonic, precessing waveform across the parameter space. The error will be of order b^2, which is small over much of the parameter space, and for the majority of orientations.

Fig. 2.6 shows the overlap between the full waveform and a subset of the harmonics for a binary with $M = 40M_\odot$, $q = 4$ and $\chi_{\rm eff} = 0$, while varying the orientation and value of χ_P. In each case, we calculate,

$$O(h,h') = \frac{\max_{\phi_o}(h|h')}{|h||h'|}, \tag{2.34}$$

where,

$$(a|b) = 4\,{\rm Re}\int_{f_o}^{\infty}\frac{a^\star(f)b(f)}{S(f)}df, \tag{2.35}$$

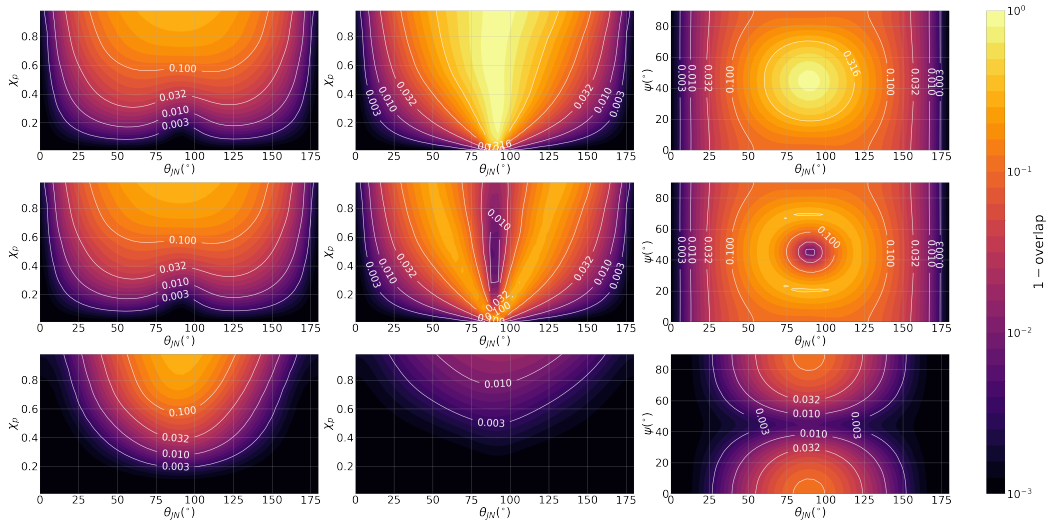

Figure 2.6: The overlap between a precessing waveform and a subset of the harmonics, as a function of the precessing spin and binary orientation for a $40M_\odot$ binary with mass ratio $q = 4$ and $\chi_{\rm eff} = 0$. The top row shows the overlap between the leading, $k = 0$, harmonic and the full waveform; the second row shows the overlap between the dominant harmonic and the full waveform; the bottom row shows the overlap between our two-harmonic precessing waveform and the full waveform. The first column is for the + polarization, second for $\times$ and third for fixed $\chi_P = 0.6$ and varying polarization.

and $S(f)$ is the power spectral density of the detector data. Thus the overlap is maximized over the phase, but not over time or any of the mass and spin parameters. An overlap of close to unity shows that the two waveforms are very similar, while a lower value of overlap implies significant deviations between the waveforms. As a rule of thumb, an overlap $O(h, h') \lesssim 1 - 3/\rho^2$ will be observable at a signal to noise ratio ρ [126, 127, 79].

We calculate the overlap of the full waveform, h, against

1. the leading order waveform in the precession expansion, h^0;
2. the dominant harmonic, i.e. the harmonic of h^0 and h^1 which contains the largest fraction of the power in the full waveform;
3. the two-harmonic waveform with the appropriate values of $\mathcal{A}_0$ and $\mathcal{A}_1$.

For the + polarized waveform (left column), the $k = 0$ harmonic is dominant for all values of $\theta_{\rm JN}$ and χ_P, so that the observed overlap with the full waveform is above 0.8 across the parameter space. For $\theta_{\rm JN} \approx 0$ or small values of χ_P, the other harmonics make a minimal contribution and the overlap is close to unity. For larger values of $\theta_{\rm JN}$ and χ_P the other harmonics are more significant and the overlap drops to 0.9 or less. The two-harmonic waveform is a significantly better match to the full waveform, with an overlap greater than 0.99 for much of the parameter space,

and only below 0.9 for edge-on systems with high χ_P where the $k = 2$ harmonic contributes most strongly (and the $k = 1$ contribution vanishes).

For the $\times$ polarized waveform (center column), the effect of incorporating the $k = 1$ harmonic is dramatic. For $\theta_{\rm JN} = 90^\circ$ the $k = 0$ contribution vanishes and only the $k = 1, 3$ harmonics are present. Thus, the overlap with harmonic $k = 0$ is essentially zero. Using the best of $k = 0, 1$ provides a good overlap with the edge-on waveform, but there is still a poor overlap at $\theta_{\rm JN} \approx 60^\circ$ where both the $k = 0$ and 1 harmonics contribute significantly to the waveform. This effect has been observed previously, for example in [97, 98] and a geometric understanding of its origin provided. The two-harmonic waveform matches remarkably well to the full waveform, with the largest differences for $\theta_{\rm JN} = 90^\circ$ and $\chi_P \approx 1$ where the overlap drops to 0.99 due to the contribution from the $k = 3$ harmonic.

The right column shows the overlap as the orientation of the binary changes. As expected, at points where the $k = 0$ harmonic vanishes ($\theta_{\rm JN} = 90^\circ$ and $\psi = 45^\circ$), the overlap with this harmonic drops to zero. The dominant harmonic is a good match to the waveform, except for orientations where two harmonics contribute significantly. As discussed in detail in Ref. [97], this corresponds to configurations where the binary orientation passes through the null of the detector response (i.e. the signal goes to zero) once per precession cycle. Thus, the radius of the circle with poor overlaps is approximately equal to the opening angle of the binary. The two-harmonic approximation provides an excellent fit to the full waveform over the majority of orientations, only dropping below 0.95 for orientations where $\theta_{\rm JN} \to 90^\circ$ and $\psi \approx 0, 90^\circ$, where the $k = 2$ harmonic is most significant.

Next, we investigate the validity of the two-harmonic approximation for a population of binaries. To begin with, let us fix the masses and spins and just consider the effect of binary orientation. As before, we choose $M = 40 M_\odot$, $q = 4$, $\chi_{\rm eff} = 0$ and $\chi_P = 0.6$, corresponding to $\bar{b} \approx 0.3$, with the binary orientation distributed uniformly over $\cos(\theta_{\rm JN}), \phi_o, \alpha_o, \psi$. Fig. 2.7 shows the distribution of the overlap between the full waveform and 1) the $k = 0$ harmonic, 2) the dominant harmonic and 3) the two-harmonic approximation. The results are shown for both a uniformly distributed population, and a population of signals observable above a fixed threshold in the detector — thereby favoring orientations that produce the largest amplitude gravitational wave. The median overlap with either the $k = 0$ or dominant harmonic is $\lesssim 0.9$, while the two-harmonic approximation improves the median overlap to 0.99. Using the dominant harmonic, there are a small fraction of signals with overlaps of 0.7 or lower (and for the $k = 0$ harmonic, this tail extends to overlaps of 0.2), while for the two harmonic approximation, the worst overlap is 0.88.

We can use these results to obtain a *rough* sense of the benefits of performing a search using the two-harmonic approximation. Previous, more detailed, investigations of this question have been carried out in, e.g. [103, 97, 128]. Current gravitational wave searches make use of spin-aligned waveforms [129, 130], and a

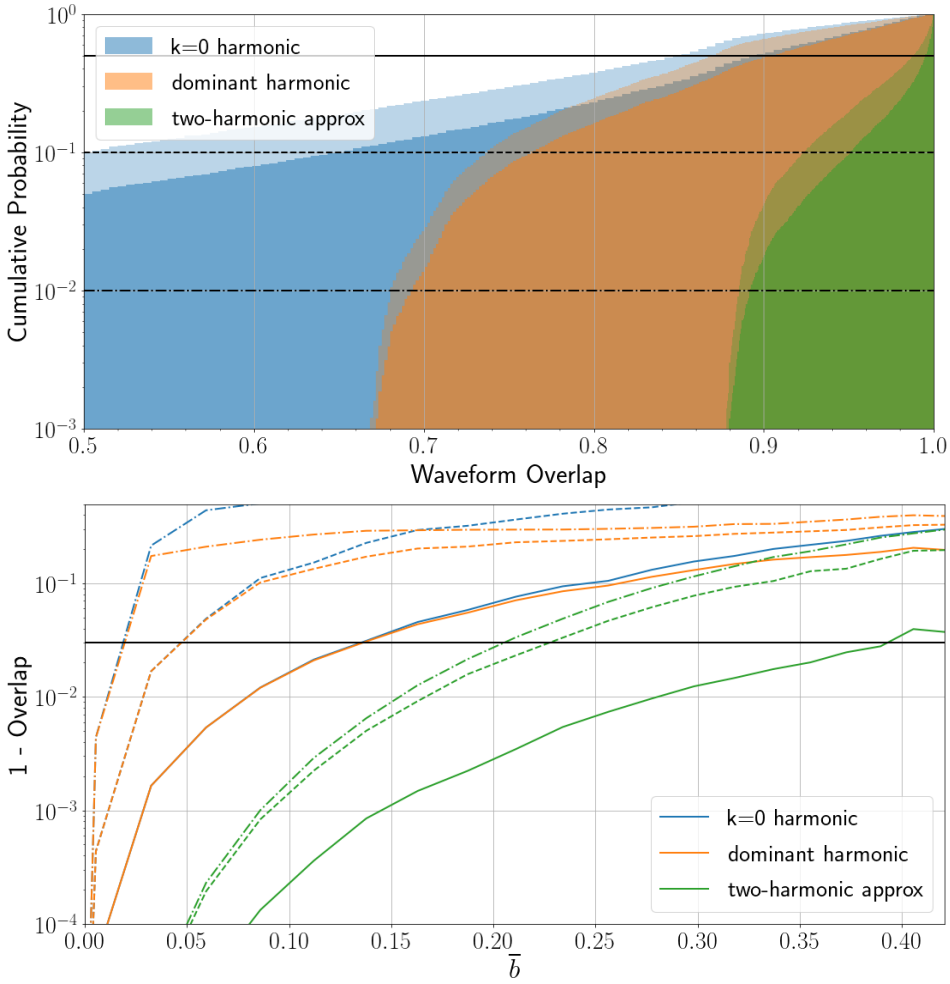

Figure 2.7: The distribution of the overlap of the precessing waveform with the $k = 0$, dominant and two-harmonic waveforms for a population of signals with $M = 40M_\odot$, $q = 4$, $\chi_{\rm eff} = 0$. The top plot shows the overlap distribution for $\chi_P = 0.6$, with random orientation of the signal. The lighter shaded regions give the distribution for a randomly oriented population of sources and the darker regions for the expected observed distribution (for a uniform-in-volume source). The lower plot shows the overlap between full and approximate waveforms as a function of $\bar{b}$. The lines on the plot show the value of the overlap for the median (solid line), worst 10% (dashed) and worst 1% (dot-dashed) of signals.

precessing waveform will naturally be identified by a spin-aligned waveform which matches well the dominant harmonic. Thus, we can use the overlaps between the precessing waveforms and dominant harmonics as a proxy for the performance of an aligned spin search. Since the median overlap is 0.9 we would expect to recover approximately 70% as many signals ($\approx 0.9^3$ for a population uniform in volume) as with a full precessing search, above a fixed threshold. A search based upon the two-harmonic approximation would recover around 97% of these signals, indicating an improvement of over 30% in sensitivity to such systems.

We also show how the distribution of overlaps varies across the mass and spin parameter space, as encoded by the parameter $\overline{b}$ and plotted for three choices of spin distribution in Figure 2.4.[5] For $\overline{b} \lesssim 0.13$ — accounting for three quarters of signals in the low-isotropic population — the median overlap between the dominant harmonic and the full waveform is above 0.97. Thus, for the majority of expected signals, the spin-aligned search will have good sensitivity. However, even for low values of $\overline{b}$ there will be some orientations of signals where two dominant harmonics will not match the waveform well, while the two-harmonic waveform still provides an essentially perfect representation of the waveform for all orientations. At $\overline{b} \approx 0.25$ the median overlap with the dominant harmonic waveform drops to 0.9, and it is here that a search with the two-harmonic approximation could provide a 30% improvement. We note, however, that for the low-isotropic distribution this accounts for only 5% of systems. While systems with such significant precession may be rare they would come from interesting areas of parameter space, with high mass ratios and spins. It is only at $\overline{b} = 0.4$ that the median overlap for the two harmonic waveform drops to 0.97, indicating a 10% loss relative to an ideal search, but also 70% improvement over a spin-aligned search.

2.6 Searching for precessing binaries

The two-harmonic approximation provides an ideal basis to develop a search for binaries with precession. The typical approach to searching for binary coalescences has been to generate a template-bank of waveforms that covers the parameter space [131, 132, 133]. These templates comprise discrete points in the mass and spin space chosen so that the waveform produced by a binary anywhere in the parameter space of interest has a match of at least 97% with one of the templates. The waveform for each template is then match-filtered against the data to identify peaks of high SNR, and various signal consistency and coincidence tests are used to differentiate signals from non-stationary noise transients [134, 135, 136, 129, 130]. Current searches make use of a template bank covering the four dimensional mass and aligned-spin space [137, 138].[6] The search takes advantage of the fact that changing the sky location, distance and orientation of the binary only changes the overall amplitude and phase

[5] While these plots were made with fixed masses and $\chi_{\rm eff}$, they should give a reasonable indication of the accuracy of the two-harmonic waveform across the mass and spin parameter space, as a function of $\overline{b}$. For different masses and spins, the evolution of the precession angle during the coalescence can have a slight impact upon the relative importance of the modes but, as b typically does not change significantly over the observable waveform, this effect is likely to be small. Furthermore, as different modes are not perfectly orthogonal, the degree to which they are not will also have a small effect upon the results. As shown in Section 2.7, the harmonics are close to orthogonal for $M \lesssim 40 M_\odot$ so that the results shown here will be representative, at least at lower masses.

[6] As we have discussed, the most significant effect on the observed waveform arises due to the effective spin $\chi_{\rm eff}$, which is a combination of the aligned spin components of the two waveforms. Thus, although the template space is four dimensional, one of the spin directions provides limited variation to the waveforms, and thus is relatively straightforward to cover.

of the signal, and these quantities can be maximized over in a simple manner.

When developing a search for precessing binaries, the search becomes more challenging due to the increasing number of parameters. In principle, it is necessary to search over two masses and six spin components, although, in practice it will probably be sufficient to restrict to the masses, $\chi_{\rm eff}$ and χ_P. The second complication is that the observed morphology of the waveform varies as the orientation of the binary changes, and it becomes necessary to search over binary orientation $\theta_{\rm JN}$, polarization ψ and precession phase α_o, although methods have been developed to straightforwardly handle a subset of these parameters [139, 103].

The two-harmonic waveform can be used to maximize the SNR over the binary orientation in a simple way. The two complex amplitudes $\mathcal{A}_0$ and $\mathcal{A}_1$, defined in Eq. (2.32), are dependent upon five variables: the distance, d_L, binary orientation, $\theta_{\rm JN}$, ψ, and the initial orbital and precession phases, ϕ_o, α_o. Since $\mathcal{A}_0$ and $\mathcal{A}_1$ can take any value in the complex plane, it is possible to construct the two-harmonic SNR by filtering the two harmonics h_0 and h_1 against the data and then freely maximizing the amplitudes so that,

$$\rho^2_{\rm 2harm} = \rho^2_0 + \rho^2_1 \,. \tag{2.36}$$

If the harmonics are not orthogonal, the two-harmonic SNR should be calculated using h^0 and $h^1_\perp$ — the $k = 1$ harmonic with any component proportional to h^0 removed. The extrinsic parameters of the binary (distance, sky location, orientation, orbital and precession phase) can be searched over through maximization over the amplitudes of the two harmonics, leaving only the masses and spins as dimensions to search using a bank of waveforms.

We must still construct a bank of waveforms to cover the four-dimensional parameter space of masses, the effective aligned $\chi_{\rm eff}$ and precessing χ_P components of the spins. The amplitude and phase evolution of a single harmonic does not carry the tell-tale amplitude and phase modulation caused by precession, but does have a different phase evolution due to precession [99, 84]. Since the phase evolution of each precessing harmonic is degenerate with a non-precessing waveform with different mass-ratio or effective spin, the bank of templates will essentially be a bank of non-precessing waveforms. This may allow us to reduce the size of the template bank.

The k=0 harmonic of the precessing waveform has an additional phase (see Eq. (2.16)) of,

$$\delta\phi_0(t) = \int_{t_o}^{t} \frac{2b^2}{1+b^2}\,\dot{\alpha}\,dt' \,. \tag{2.37}$$

For systems in which orbital angular momentum dominates over spin angular momentum, the precession frequency is inversely proportional to orbital frequency, $\Omega_P = \dot{\alpha} \propto f^{-1}$ [62, 73, 97]. This is the same frequency dependence as the 1PN

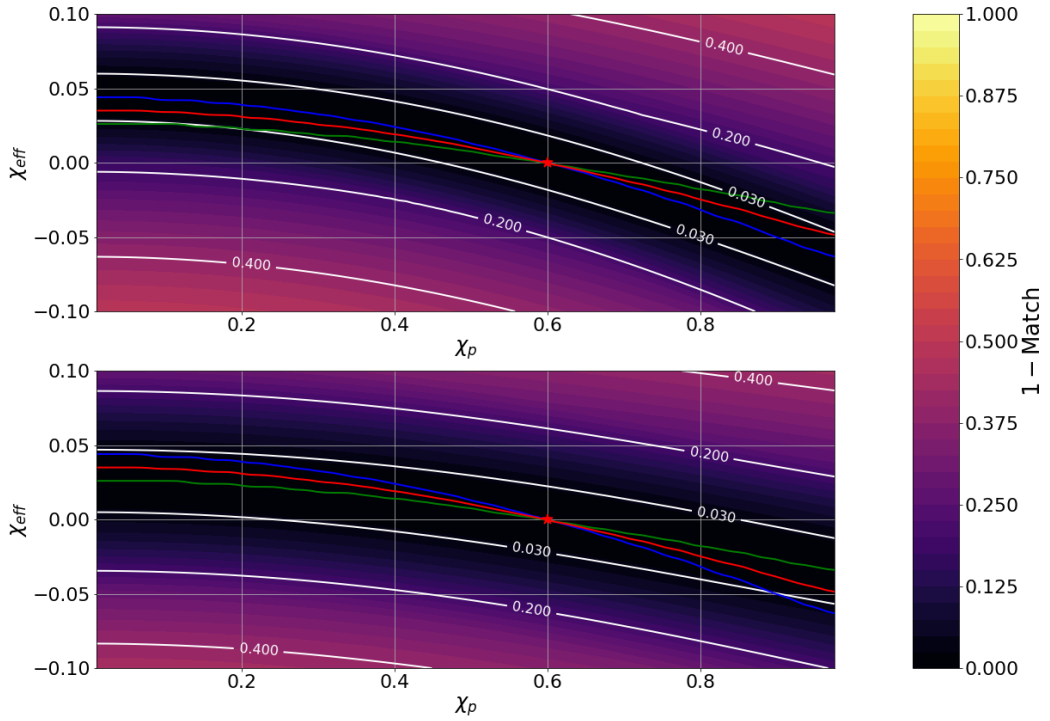

Figure 2.8: The mismatch between the $k = 0$ (left) and $k = 1$ (right) harmonic of two precessing signals as the effective spin $\chi_{\rm eff}$ and precessing spin χ_P are varied. For all waveforms, the total mass is fixed to $40M_\odot$ and the mass ratio to 4. One waveform has $\chi_{\rm eff} = 0$ and $\chi_P = 0.6$ (the point marked by a star), while the spins of the second waveform are varied. The blue and green lines show the value of $\chi_{\rm eff}$, for the $k = 0$ and $k = 1$ harmonics respectively, which gives the largest match with the fiducial waveform; the red line is the average of these values.

contribution to the waveform, whose amplitude depends upon the mass ratio. Consequently, it is reasonable to expect that the precession-induced phase will be indistinguishable from a systematic offset in the binary mass ratio, or the effective spin [80]. Similarly, the $k = 1$ harmonic has essentially the same amplitude evolution as the non-precessing waveform, but with a phase difference of,

$$\delta\phi_1(t) = -\int_{t_o}^{t} \frac{1-b^2}{1+b^2}\,\dot{\alpha}\,dt'\,, \tag{2.38}$$

which will also, in many cases, be degenerate with a change in the mass ratio or aligned spin.

In Figure 2.8, we investigate the degeneracy in the spin ($\chi_{\rm eff}$–χ_P) space of the two leading precession harmonics. We consider a system with masses, $M = 40M_\odot$ and $q = 4$, and spins $\chi_{\rm eff} = 0$, $\chi_p = 0.6$ and investigate how the two waveform harmonics vary as the spins are changed. The figure shows the match — the overlap maximized over time-offsets — between our fiducial waveform and one with the same masses but different spins. For both harmonics, there is a band in the $\chi_{\rm eff}$–χ_P plane where

the is mismatch is small — the different phase evolution of each harmonic caused by varying χ_P can be offset by a suitable change in $\chi_{\rm eff}$. The relation is approximately quadratic, $\Delta\chi_{\rm eff} \propto (\Delta\chi_P)^2$, which is to be expected. Recall, from Eq. (2.37), that the the change in phase due to precession is quadratic in b, and therefore also in χ_P at least for small values of b. Meanwhile the phasing of the waveform varies, at leading order, linearly with $\chi_{\rm eff}$.

This degeneracy in the $\chi_{\rm eff}$–χ_P plane suggests that a single template waveform could be used to search over an extended region corresponding, for example, to the region of mismatch < 0.03 in Figure 2.8. However, this will only work if the degenerate region for the $k = 0$ and $k = 1$ harmonics is the same. It is clear from Equations (2.37) and (2.38) and Figure 2.8 that they are not identical. Nonetheless,[7] for the example we have considered, the two degenerate regions are similar, and along the line that traces the mid-point between the the best fit values of $\chi_{\rm eff}$ for the two harmonics, both harmonics have a match above 0.97 with the initial point. Thus, to an accuracy appropriate for generating a template bank, we can use the two harmonics from a single waveform to cover a band in the $\chi_{\rm eff}$–χ_P plane which spans all values of χ_P. This effectively reduces the dimensionality of the parameter space to three dimensions: mass, mass ratio and one spin parameter.

Our proposal to develop a precessing search is as follows. First, generate a bank of templates to cover the space of non-precessing binaries. At each M, q, $\chi_{\rm eff}$ point in the template bank, construct the two-harmonic waveform for a fixed value of χ_P. Then, filter the data against the two harmonics and calculate the two-harmonic SNR, as defined in Eq. (2.36) to identify candidate events in a single detector. It will be necessary to extend the existing χ^2 signal consistency test [135] to each harmonic, taking into account the presence of the other harmonics, to reduce the impact of non-stationarity in the data. Next, perform coincidence between detectors by requiring a signal in the same template at the same time, up to the allowed time delays based upon speed of propagation. For a non-precessing signal observed in two detectors, the relative amplitude and phase of the SNR in each detector can take any value, even though some are astrophysically more likely [140] (and this can be used to increase search sensitivity). However, for the two-harmonic waveform not every signal observed in two detectors will be compatible with an astrophysical source. This can be seen through simple parameter counting: there are ten measured quantities (two complex amplitudes and a time of arrival in each detector), which depend upon eight parameters, the five orientation parameters (d_L, $\theta_{\rm JN}$, ψ, ϕ_o, α_o), sky location and merger time. An additional coincidence test to check for consistency between parameters will likely be necessary to reduce the search background. A similar problem arises already in extending the amplitude and phase consistency of

[7] Strictly, when doing this comparison, we must use the same time offset for the two harmonics, whereas the figure allows for an independent maximization of the time delay for each harmonic. Fixing a single time delay does slightly decrease the matches, but not significantly enough to change the conclusions.

[140] to three or more detectors and methods developed for that purpose may be helpful for the precessing search.

We can estimate the likely sensitivity improvement from a precessing search, as we have briefly discussed in Section 2.5. A non-precessing search will typically find the dominant harmonic of the waveform. Thus, for signals where two harmonics provide a significant contribution, a search based on the two-harmonic waveform has the potential to out-perform the non-precessing search. The two-harmonic waveform has four degrees of freedom, encoded in $\mathcal{A}_0$ and $\mathcal{A}_1$, compared to two for the non-precessing search. Thus, the noise background is higher for the two-harmonic search and, based upon a comparison of the tails of the χ^2 distribution with 2 and 4 degrees of freedom, an increase of around 5% in SNR is required to obtain the same false alarm rate (see e.g., Ref. [103] for a discussion of this issue). Thus, a signal will be observed as more significant in the two-harmonic search than a non-precessing search if the SNR can be increased by 5% or more. Fig. 2.7 shows that this occurs for $\overline{b} \gtrsim 0.15$, and for binaries with $\overline{b}$ above this value the two-harmonic search has the potential to outperform a non-precessing search. We note, however, that a given template will cover a range of spin values and consequently a range of $\overline{b}$, so it may be more appropriate to deploy the two-harmonic search for templates with an *average* of $\overline{b}$ which is greater than 0.15.

Another challenge of searches for precessing systems is the associated computational cost [103], which can be prohibitive. The maximum computational cost for the two-harmonic search would be double that of a comparable non-precessing search: it becomes necessary to filter both the $k = 0$ and 1 harmonics, and computational time is dominated by this matched filtering. However, since both the $k = 0$ and $k = 1$ harmonics are essentially non-precessing waveforms, there may be waveforms associated with the $k = 1$ harmonics are *already* in the set of $k = 0$ waveforms, but associated with different parameters. If so, this could further reduce the computational cost.

2.7 Observability of precession

The two-harmonic approximation allows us to easily identify regions of the binary merger parameter space for which precession will leave an observable imprint on the waveform. Since the amplitude and phase evolution of a single harmonic is generally consistent with that of a non-precessing waveform (see above and [99, 98]), it is only when two harmonics can be observed that we are able to clearly identify precession in the system. We are therefore interested in deriving an expression for the *precession SNR*, ρ_p, defined as the SNR in the second most significant harmonic, and determining when it will be observable. If the two harmonics h^0 and h^1 in

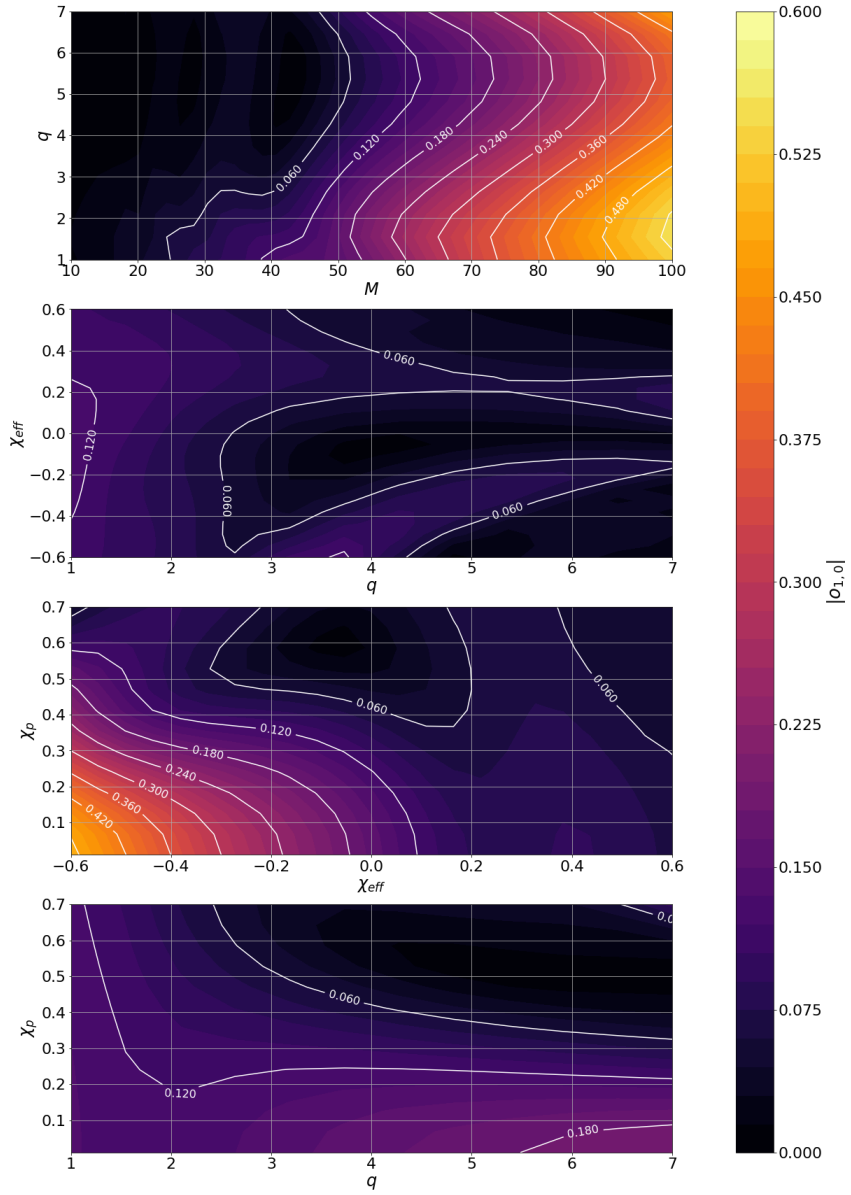

Figure 2.9: The overlap $O(h_0, h_1)$ between the $k = 0$ and $k = 1$ harmonics across two-dimensional slices in the parameter space of total mass, mass ratio, $\chi_{\rm eff}$ and χ_p. In each plot, two of the parameters are varied while the other two are fixed to their fiducial values of $M = 40M_\odot$, $q = 4$, $\chi_{\rm eff} = 0$, $\chi_p = 0.6$.

Eq. (2.31) are orthogonal, then the precession SNR is simply,

$$\begin{aligned} \rho_p &= \min(|\mathcal{A}_0 h^0|, |\mathcal{A}_1 h^1|), \\ &= \rho_{\rm 2harm} \left(\frac{\min(1, |\zeta|)}{\sqrt{1 + |\zeta|^2}} \right), \end{aligned} \tag{2.39}$$

where ζ, defined in Eq. (2.33), gives the ratio of the SNR in the $k = 1$ and $k = 0$ harmonics and $\rho_{\rm 2harm}$ is the total SNR in the two-harmonic waveform.

Let us briefly examine where in parameter space the two harmonics are close to orthogonal. Where there are sufficient precession cycles we expect the two harmonics, h^0 and h^1, will be close to orthogonal, and the overlap to be close to zero [99]. The overlap between the two harmonics for various two-dimensional slices through the parameter space is shown in Fig. 2.9. At higher masses, where the binary completes one, or fewer, precession cycles in the detector's sensitive band, there is a larger overlap between the harmonics. At negative $\chi_{\rm eff}$ and minimal χ_p, the overlap is also significant. However, providing the mass of the system is below $50M_\odot$, for the much of the parameter space the overlap is less than 0.1 and simple expression in Eq. (2.39) will be applicable.

Taking into account the overlap between harmonics, the total power in the two-harmonic waveform is,

$$\rho^2_{\rm 2harm} = |\mathcal{A}_0 h^0|^2 \left(1 + 2{\rm Re}[\zeta\, o_{1,0}] + |\zeta|^2\right) . \tag{2.40}$$

where $o_{1,0}$ is complex overlap between the two harmonics:

$$o_{1,0} = \frac{(h^1|h^0) + i(h^1|ih^0)}{|h^1||h^0|} . \tag{2.41}$$

We can project the SNR onto directions parallel and perpendicular to the h^0 waveform to obtain the SNR in these two directions as,

$$\begin{aligned} \rho^2_0 &= |\mathcal{A}_0 h^0|^2 \left(1 + 2{\rm Re}[\zeta\, o_{1,0}] + |\zeta\, o_{1,0}|^2\right) , \\ \rho^2_{\perp,0} &= |\mathcal{A}_0 h^0|^2 |\zeta|^2 \left(1 - |o_{1,0}|^2\right) . \end{aligned} \tag{2.42}$$

Similarly, the power parallel to and perpendicular to the $k = 1$ harmonic is,

$$\begin{aligned} \rho^2_1 &= |\mathcal{A}_0 h^0|^2 \left(|o_{1,0}|^2 + 2{\rm Re}[\zeta\, o_{1,0}] + |\zeta|^2\right) , \\ \rho^2_{\perp,1} &= |\mathcal{A}_0 h^0|^2 \left(1 - |o_{1,0}|^2\right) . \end{aligned} \tag{2.43}$$

The precession SNR is defined as the power orthogonal to the dominant harmonic,[8]

$$\begin{aligned} \rho_p &:= \min(\rho_{\perp,0}, \rho_{\perp,1}), \\ &= \rho_{\rm 2harm} \min(1, |\zeta|) \left(\frac{1 - |o_{1,0}|^2}{1 + 2{\rm Re}[\zeta\, o_{1,0}] + |\zeta|^2}\right)^{\frac{1}{2}} . \end{aligned} \tag{2.44}$$

As expected, the precession SNR scales with the total SNR of the signal, so that precession will be more easily observed for louder events. If there is significant de-

[8] In exceptional circumstances, where the overlap is large and $\zeta o_{1,0}$ is close to -1, there can be more power in $\rho_{\perp,i}$ than ρ_i. In such cases, it is natural to use ρ_i to determine if precession is present, although this is not ideal as $\rho_{\perp,i}$ need not resemble a non-precessing waveform.

generacy between the harmonics, the numerator will be reduced, making the observation of precession more difficult. Finally, in the limit that $o_{1,0} \to 0$, the expression simplifies to the one given earlier for orthogonal harmonics in (2.39), as expected.

What value of ρ_p will be required to observe precession? This will happen if the evidence for a signal with $\chi_p \neq 0$ in the data is greater than that for a non-precessing source. This can be evaluated through Bayesian model selection, by considering the Bayes factor between the hypotheses. However, such a calculation requires a full exploration of the parameter space. We can, instead, obtain an approximate answer by considering the maximum likelihood. Since the two-harmonic waveform is more general than the non-precessing waveform, it will always give a larger maximum likelihood *even in the absence of precession* due to its ability to fit the detector noise. Thus, we are interested in examining the expected increase in SNR due to the inclusion of the second harmonic, in the absence of any power in it.

The two-harmonic SNR can be written as

$$\rho^2_{\rm 2harm} = \rho^2_{np} + \rho^2_p \,. \tag{2.45}$$

where ρ_{np} is the non-precessing SNR or, equivalently, the SNR in the dominant harmonic. In the absence of precession, ρ_p will be χ^2 distributed with 2 degrees of freedom, as we are able to freely maximize over the amplitude and phase of the two harmonics independently. Consequently, in 90% of cases, noise alone will give a value of $\rho_p < 2.1$. Therefore, as a simple criterion, we require that,

$$\rho_p \geq 2.1, \tag{2.46}$$

for precession to be observable. In Ref. [141] we use this definition to investigate in detail the observability of precession over the parameter space.

2.8 Discussion

We have presented a new, intuitive way to understand the observability of precession in GW observations. By keeping only the leading precession term, we have derived a precession SNR and argued that this can be used to determine when precession will be observable. Before discussing applications we point out the main limitations of this analysis. As is clear from the formulation, this analysis works best for binaries where $b = \tan(\beta/2)$ is small. This typically corresponds to situations where the masses are comparable, the precessing spin is small and any aligned component of the spin is aligned (rather than anti-aligned) with the orbital angular momentum. We have shown above that this assumption is valid for a reasonable population.

We now point to several advantages and applications of this formulation:
First, it gives new understanding of the observability of precession, and also of the origin of precession as the beating of two waveform components with slightly differing

frequencies (also discussed in [99]). It is difficult to identify the presence of precession in a GW observation directly from χ_P, since the prior astrophysical expectation disfavours $\chi_P = 0$. While the deviation from the prior can be determined through the Bayes factor, the precession SNR ρ_P has the advantage of providing a direct measure of whether precession has been measured in a signal. This will be illustrated in more detail in chapter 4, where we probe the measurability of precession across the gravitational wave parameter space. The precession SNR has also been used for just this purpose in the interpretation of the recent GW observation GW190412 [1].

There exist a number of detailed population analyses which extract the features of the underlying population of gravitational waves from the set of observed gravitational wave events, for example [142, 143, 144, 74]. These typically use the full posterior distributions recovered from the gravitational wave signal [52, 145] to infer the population and, as such, naturally account for precession effects in the observed signals when inferring the black hole mass and spin populations. Nonetheless, there have been a number of studies performed which investigate the population properties using a subset of the recovered parameters, see e.g. [146, 121, 147, 148, 122, 74], and have been successfully used to infer interesting properties of the mass and spin distributions. The majority of these studies have restricted attention to the aligned components of the spins. The precession SNR provides a straightforward method to determine the significance of precession, and provides away to probe observability of precession in populations of binaries. In using this method we have been able to derive constraints on the preferred spin distribution including precession effects [123].

Both of the applications highlighted above are currently possible using other more sophisticated but computationally expensive methods such as Bayesian model comparison. This is, of course, a more general method that makes fewer assumptions than we do in computing ρ_p, however the computational costs associated with calculating the marginal likelihood over multiple, e.g. precessing and non-precessing, models per binary are not feasible for a large number of binaries. For example the analysis in [123] involved calculating ρ_p for 1 million binaries, and computing the Bayes factor for 1 million binaries would certainly not be practical. Similar, lightweight analyses, could also be developed using the formalism introduced in, e.g. [97], and if this is done,it would be interesting to compare them with the results from the two harmonic analysis.

Finally, we have outlined a method by which the two-harmonic approximation could be used to develop a search for precessing binaries. We have shown that in principle that this approach could result in a significant increase in sensitivity without the computational overheads associated with other precessing search methods. In addition, the formalism should provide a way to identify the parts of parameter space where a precessing search is likely to increase sensitivity. We plan a detailed investigation into the feasibility of a precessing search based upon the two-harmonic approximation in future work.

Appendix: Derivation using spin-weighted spherical harmonics

In this appendix, we provide an alternative derivation of the power series decomposition of the precessing waveform, given in Section 2.3, based upon the spin-weighted spherical harmonic decomposition of the waveform [149] and its application to precession as described in [84, 118]. Specifically, we wish to obtain the result in Eq. (2.14). Throughout, we follow the notation used in [86].

The gravitational waveform emitted during a binary merger,

$$h := h_+ - ih_\times \tag{2.47}$$

can naturally be decomposed into a set of spin-weighted spherical harmonics as

$$h(t, \vec{\lambda}, \theta, \alpha_o) = \sum_{\ell \geq 2} \sum_{-\ell \leq m \leq \ell} h_{\ell,m}(t, \vec{\lambda})\, {}^{-2}Y_{\ell,m}(\theta, \phi) \tag{2.48}$$

where θ and ϕ give the orientation of the observer relative to a co-ordinate system used to identify the spherical harmonics, $\vec{\lambda}$ encodes the physical parameters of the system (masses, spins, etc) and t is the time.

The multipoles for a precessing system are approximated by "twisting up" [84, 118] the multipoles of the non-precessing counterpart based upon the orientation of the orbital angular momentum given by the opening angle β, precession angle α and the third Euler angle ϵ defined via

$$\dot{\epsilon} = \dot{\alpha} \cos\beta \,. \tag{2.49}$$

Then, the precessing multipoles are given by

$$h^{\rm prec}_{\ell,m}(t) = \sum_{-\ell \leq n \leq \ell} h^{\rm NP}_{\ell,n} D^{\ell}_{n,m}(\alpha(t), \beta(t), \epsilon(t)) \tag{2.50}$$

where the Wigner D-matrix is

$$D^{\ell}_{n,m}(\alpha, \beta, \epsilon) = e^{im\alpha} d^{\ell}_{n,m}(-\beta) e^{-in\epsilon} \tag{2.51}$$

and the Wigner d-matrix given, for example, in [150].

Combining these decompositions gives the waveform for a precessing binary as

$$h = \sum_{\ell,m,n} {}^{-2}Y_{\ell,m}(\theta, \phi) D^{\ell}_{n,m}(\alpha, \beta, \epsilon) h_{\ell,n}(t, \vec{\lambda}) \,. \tag{2.52}$$

In performing the twisting, it's natural that the precessing waveform is described in a coordinate system aligned with the orbital angular momentum, so that $\theta = \theta_{\rm JN}$.

Furthermore, the orientation of the x-axis will be specified relative to the (initial) precession phase so that $\phi = -\alpha_o$.

In this work, we restrict attention to the case where the non-precessing model contains only the $\ell = 2$ and $n = \pm 2$ modes, and require symmetry in gravitational wave emission above and below the plane of the binary so that $h_{\ell,n} = (-1)^\ell h^\star_{\ell,-n}$. This eliminates the sum over ℓ and m from Eq. (2.52). Furthermore, we can expand the spherical harmonics using

$$ {}^{-2}Y_{2,m}(\theta_{\rm JN}, -\alpha_o) = \sqrt{\frac{5}{4\pi}} d^2_{m,2}(\theta_{\rm JN}) e^{-im\alpha_o} \tag{2.53} $$

to obtain

$$ \begin{aligned} h^{\rm prec} \;=\; & \sum_{-2\le m\le 2} \sqrt{\frac{5}{4\pi}} d^2_{2,m}(\theta_{\rm JN}) e^{im(\alpha-\alpha_o)} \times \\ & \left[h^{\rm NP}_{22} d^2_{2,m}(-\beta) e^{-2i\epsilon} + (h^{\rm NP}_{22})^\star d^2_{-2,m}(-\beta) e^{2i\epsilon}\right] \end{aligned} \tag{2.54} $$

We now wish to re-write the above to show that the waveform can be decomposed in modes whose amplitudes form a power series in $b = \tan(\beta/2)$. To do so, we note that the Wigner d-matrices can be evaluated as powers of $\sin(\beta/2)$ and $\cos(\beta/2)$, so that if we are able to group terms with the same indices we will arrive at the desired expression. We do this by using the d-matrix identities:

$$ d^\ell_{n,m} = (-1)^{m-n} d^\ell_{m,n} = (-1)^{m-n} d^\ell_{-n,-m} \tag{2.55} $$

and relabelling the dummy index $m \to -m$ in the second term of Eq. (2.54) to obtain:

$$ \begin{aligned} h^{\rm prec} \;=\; & \sum_{-2\le m\le 2} \sqrt{\frac{5}{4\pi}} d^2_{2,m}(-\beta) \times \\ & \Big[(-1)^m d^2_{2,m}(\theta) \left(h^{\rm NP}_{22}(t) e^{-2i\epsilon} e^{im(\alpha-\alpha_o)}\right) \\ & + d^2_{2,-m}(\theta) \left(h^{\rm NP}_{22}(t) e^{-2i\epsilon} e^{im(\alpha-\alpha_o)}\right)^\star\Big] \end{aligned} \tag{2.56} $$

Finally, we can evaluate the Wigner d-matrices as

$$ \begin{aligned} d^2_{2,m}(-\beta) \;&:=\; C_m \cos^{2+m}(\beta/2) \sin^{2-m}(\beta/2) \\ &=\; \frac{C_m b^{2-m}}{(1+b^2)^2} \end{aligned} \tag{2.57} $$

where $C_{\pm 2} = 1$, $C_{\pm 1} = 2$, $C_0 = \sqrt{6}$ and, as before, $b = \tan(\beta/2)$. Similarly, we

introduce $\tau = \tan\theta_{JN}/2$, and evaluate the d-matrices for the angle $\theta_{\rm JN}$. This gives

$$\begin{aligned} h^{\rm prec} &= \sum_{-2\le m\le 2} \sqrt{\frac{5}{4\pi}} \frac{(C_m)^2 b^{2-m}}{(1+b^2)^2} \times \\ &\quad \left[\frac{\tau^{2-m}}{(1+\tau^2)^2} \left(h^{\rm NP}_{22}(t) e^{-2i\epsilon} e^{im(\alpha-\alpha_o)} \right) \right. \\ &\quad \left. + \frac{(-\tau)^{2+m}}{(1+\tau^2)^2} \left(h^{\rm NP}_{22}(t) e^{-2i\epsilon} e^{im(\alpha-\alpha_o)} \right)^\star \right] \end{aligned} \tag{2.58}$$

This is close to the desired form and, in particular, we have obtained an decomposition where the relative strength of each mode is decreased by a factor of b. To obtain an expression comparable to Eq. (2.14) we must evaluate the waveform observed at a detector with response F_+ and $F_\times$ to the two gravitational polarizations.

$$\begin{aligned} h(t) &= {\rm Re}\left[(F_+ + iF_\times) h^{\rm prec}\right] \\ &= {\rm Re}\left[\left(\sqrt{\frac{5}{4\pi}} \frac{(h^{\rm NP}_{22})^\star e^{2i(\epsilon+\alpha-\alpha_o)}}{(1+b^2)^2} \right) \right. \\ &\quad \sum_{m=-2}^{2} \frac{(C_m)^2}{(1+\tau^2)^2} (b e^{-i(\alpha-\alpha_o)})^{2-m} \\ &\quad \left. \left(F_+ [\tau^{2-m} + (-\tau)^{2+m}] - iF_\times [\tau^{2-m} - (-\tau)^{2+m}] \right) \right] \end{aligned} \tag{2.59}$$

Then, to finally equate this with the desired expression, we must make the identification

$$\sqrt{\frac{5}{4\pi}} (h^{\rm NP}_{22})^\star e^{2i\epsilon} = \frac{d_o}{d_L} A_o(t) e^{2i\Phi_S} \,, \tag{2.60}$$

where Φ_S is defined in Eq. (2.7). Thus the amplitude of the waveform, $A_o(t)$ is the same as the scaled 22 mode while the phase of the 22 mode is the (negative) of the orbital phase. Furthermore, it is straightforward to show that the $\mathcal{A}^{+,\times}_k$ coefficients are given by

$$\begin{aligned} \mathcal{A}^{+}_{(2-m)} &= \frac{d_o}{d_L} (C_m)^2 \left(\frac{\tau^{2-m} + (-\tau)^{2+m}}{(1+\tau^2)^2} \right) , \\ \mathcal{A}^{\times}_{(2-m)} &= \frac{d_o}{d_L} (C_m)^2 \left(\frac{\tau^{2-m} - (-\tau)^{2+m}}{(1+\tau^2)^2} \right) . \end{aligned} \tag{2.61}$$

Substituting these identifications, we obtain the desired expression for the waveform

observed at a detector,

$$h(t) = \mathrm{Re}\left[\left(\frac{A_o(t)e^{2i(\Phi_S+\alpha)}}{(1+b^2)^2}\right) \sum_{k=0}^{4}(be^{-i\alpha})^k(F_+\mathcal{A}_k^+ - iF_\times\mathcal{A}_k^\times)\right] . \quad (2.62)$$

Chapter 3

Population Analysis using the Two Hamornic Approximation

The Advanced Laser Interferometer Gravitational-Wave Observatory (aLIGO) [151] and Advanced Virgo (AdV) [152], provide a unique method of observing mergers of black holes and/or neutron stars. Observations to date already provide insights into the mass and spin distributions of black holes [2, 74]. This chapter applies the methods presented in chapter 2 to carry out a population study looking at the relative probabilities of different spin population models.

One important general relativistic effect that has *not yet* been observed is orbital precession. This arises when the black-hole spins are not aligned with the binary's orbital angular momentum. In contrast to Newtonian mechanics, where all angular momenta are individually conserved, in general relativity the binary's total angular momentum is (approximately) conserved, and the orbital angular momentum (and hence the orbital plane) and spins precess around it [62, 72]. This leads to modulations in the amplitude and phase of the gravitational wave (GW) signal. These are in general small effects and, in addition, whether they can be measured depends not only on the black-hole masses and spin magnitudes and directions, but also on the binary's orientation relative to the detector, and the observed GW polarization. For this reason, until the introduction of the precession SNR presented in chapter 2 there was no straightforward way to determine how significantly precession would be imprinted onto a given waveform. The usual approach is to perform computationally expensive Bayesian analyses (see e.g. [2, 52]), but even then, the misaligned spin components (which signify whether the binary is precessing) are degenerate with other parameters, and do not provide a direct measure of precession features in the signal. This makes it difficult to infer the impact of precession measurements on the properties of astrophysical binary populations and their formation mechanisms.

3.1 Observability of precession:

Since the individual harmonics are indistinguishable from non-precessing waveforms, it is only when two precession harmonics can be independently observed that precession can be unambiguously identified. For precession to be observable, we therefore require that the expected signal-to-noise ratio (SNR) in *both* of the harmonics is above some threshold.

It remains to determine a threshold above which ρ_p can be considered as evidence for precession, this question is discussed in detail in chapter 4. For a simple approximation for this threshold, consider the situation where an event has been observed, so there is significant SNR in at least one harmonic. In the absence of measurable precession, and assuming well-modelled Gaussian noise, the SNR in the second harmonic will be χ^2 distributed with two degrees of freedom, where the two degrees of freedom correspond to the real and imaginary parts of the complex amplitude. Therefore, in the absence of precession, $\rho_p > 2.1$ is expected in less than 10% of cases, and $\rho_p > 3$ in approximately 1% of cases. We therefore use these simple thresholds as a measure of the strength of evidence for observable precession.

In Fig. 3.1 we show the recovered distribution of χ_p and ρ_p for a number of signals, both real and simulated. For each signal, we use a nested sampling routine within the LALInference code [51, 52] to obtain posterior probability distributions for the parameters. This is the same infrastructure that was used to measure the properties of the LIGO-Virgo observations, and we present our results in the same form as in, for example, the GWTC-1 catalogue [2], by using the PESummary library [153]. The new feature is our calculation of ρ_p.

First, we show the recovered χ_p and ρ_p distributions for a set of simulated signals, generated using the `IMRPhenomPv2` model [84], each with the same choices of masses and spins — total mass $M = 40M_\odot$, mass ratio 2:1, and an in-plane spin of $\chi_p = 0.4$ on the large black hole only — but varying orientation, encoded by the angle θ between the total angular momentum and the line of sight. The distance to each signal is chosen to ensure a *fixed* expected SNR of 20 in the aLIGO detectors at the sensitivity of the second observing run (O2) [2], resulting in a distance variation by a factor of ≈ 3.5 between the least and most inclined systems.

For binaries with total angular momentum closely aligned with the line of sight, $\theta < 45°$, the precessing SNR is consistent with no power in the h^1. The posterior on χ_p is consistent with the prior at low χ_p but excludes $\chi_p \gtrsim 0.7$. When $\theta > 45°$, the angular momentum is significantly mis-aligned with the line of sight and there is significant power in both harmonics, leading to a value of ρ_p inconsistent with noise alone and little support for values of $\chi_p \lesssim 0.1$. However, using χ_p alone, *even after performing the parameter recovery* there are no simple criteria to determine when precession is observed. A natural choice might require that the 90% confidence interval for χ_p exclude zero, but this will *always* be the case, primarily due to the

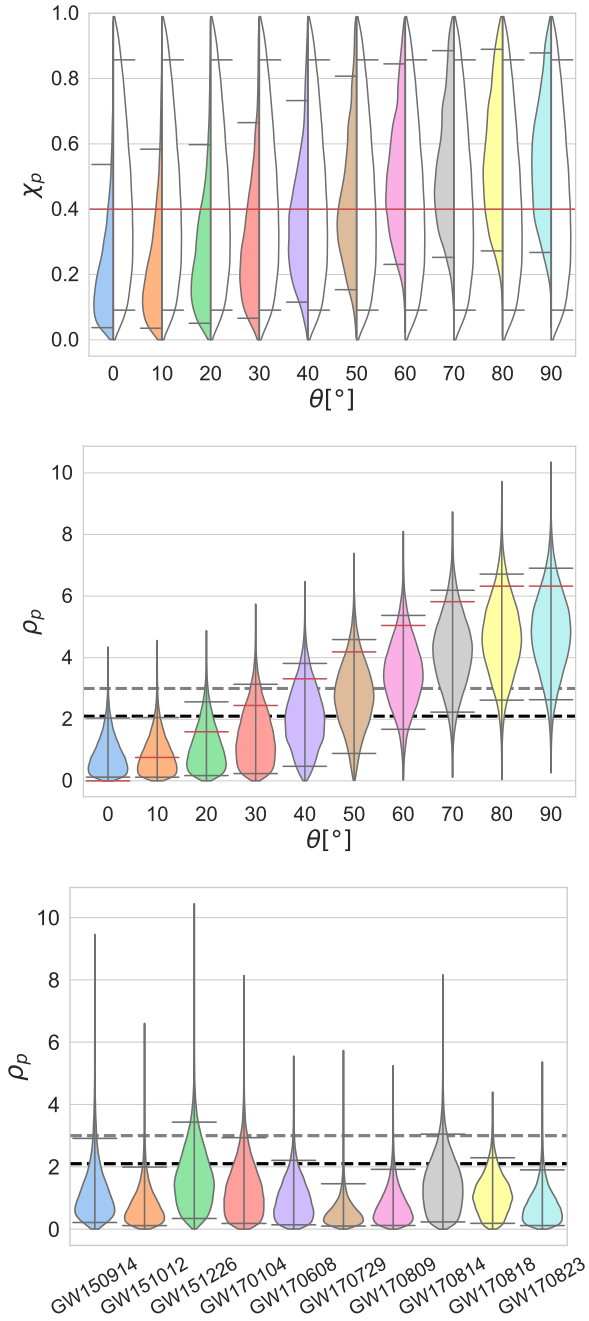

Figure 3.1: For a set of simulated signals with fixed masses and spins (see text), we show the posterior and prior (white) distributions for χ_p (top), and posterior distributions for ρ_p (middle) for a range of different binary orientations, θ. The grey lines show the 90% confidence regions, the solid red lines show the *true* values of χ_p and ρ_p respectively and the dashed black and grey lines indicates the thresholds for observable precession at $\rho_p = 2.1$ and $\rho_p = 3$. The bottom panel shows the ρ_p distribution for the ten binary-black-hole observations in O1 and O2 [2].

shape of the prior. Furthermore, even though we know all of the parameters *a priori*, it is impossible to determine whether precession will be observable without

generating the waveform and performing the parameter recovery.

The precession SNR solves these problems. A value of $\rho_p > 2.1$ tells us immediately that there is evidence of observable precession. Most significantly, the expected precession SNR ρ_p (red lines on the middle plot) can be calculated directly from the signal parameters; no detailed parameter estimation analysis is necessary. Thus, for the first time, we are able to identify immediately whether precession would be measurable in a given configuration. We see in Fig. 3.1 that for each inclination, the true value for ρ_p lies within the recovered 90% credible interval however the posterior is not centred around the true value. This is due to selection and prior effects. In chapter 4, we investigate these selection effects as well as providing a detailed exploration of the observability of precession over the parameter space of masses, spins and binary orientation.

Fig. 3.1 also shows the distribution of ρ_p for the BBH merger signals that were observed in the aLIGO and AdV O1 and O2 runs [2]. No evidence of precession was found in these signals [74], as is made clear from the recovery of ρ_p. There are several cases where the distribution extends to higher values, but the median never exceeds the 2.1 threshold. These results demonstrate the efficacy of ρ_p.

3.2 When will we observe precession?

We can use the observation of precession to distinguish different binary formation scenarios. The precession SNR makes it straightforward to perform an in-depth investigation of various models and identify the fraction of signals for which precession effects will be observable. Such a study was not previously possible, due to the difficulty in classifying observability of precession. Instead, limited investigations of the parameter space have been performed [89], or inferences of the distributions for the spin magnitudes and orientations obtained [144, 142], again with a limited sample size.

We investigate nine astrophysically-motivated populations of black hole binaries, comprised of three distributions of spin magnitude, and three distributions of spin orientation for the individual black holes in the binary. The spin-magnitude distributions are those used in Refs. [121, 146, 122]: *low* and *high* are triangular, peaked either at zero or extremal spin, and *flat* is a uniform distribution between zero and one. The spin-orientation is characterized by the distribution for the angle σ between each black hole's spin and the orbital angular momentum: *aligned* is a triangular distribution in $\cos\sigma$, which peaks at 1 and can take values $0.85 < \cos\sigma < 1.0$, ($\sigma \lesssim 30^\circ$); *precessing* is triangular in $\cos\sigma$ peaked at 0, with values $-0.15 < \cos\sigma < 0.15$, ($80^\circ \lesssim \sigma \lesssim 100^\circ$); *isotropic* is uniform in $\cos\sigma$ between -1 and 1. For each population, we generate 10^5 binaries with masses drawn from a power law mass distribution with $p(m_1) \propto m^{-2.35}$, and $p(m_2)$ uniform in m_2 between $5\,M_\odot$ and m_1 (as in [122]),

	Aligned		Isotropic		Precessing	
Low	0.043	0.644	0.151	0.194	0.173	0.150
Flat	0.077	0.448	0.276	0.040	0.327	0.019
High	0.105	0.331	0.354	0.013	0.412	0.005

Table 3.1: The probability of observing precession, $\rho_p > 3$, for an observed binary (white) from each spin distribution and the probability of *not* observing precession in 10 random draws (grey) from each spin distribution.

and distributed uniformly in volume and binary orientation.

Table 3.1 shows the probability of observing precession in a single event drawn from each of the nine populations, observed with O2 sensitivity while assuming that the PSD corresponds to the special case of the zero realisation from this process, this can also be thought of as the mean PSD of infinite draws from the underlying process. This noise realisation is often referred to as a *zero noise* realisation. For this study, we use the higher threshold of $\rho_p > 3$, corresponding to a 1% false rate, to indicate strong evidence for observed precession. When observing a population of events, the number of events exceeding this threshold when there is no precession in the system remains low.[1]

As expected, we are most likely to observe precession when the black holes have high spins that lie preferentially in the orbital plane [2] (*high-precessing* configurations) and least likely for black holes with low spins, or with spins preferentially aligned with the orbital angular momentum (*low-aligned* configurations). Given that precession has not been observed in GW detections to date, we are able to restrict the spin distribution. Table 3.1 shows the probability of detecting ten signals with no observable precession from each of the nine spin distributions. Based on precession measurements alone, we strongly disfavour all *precessing* distributions. Although these are already considered astrophysically unlikely, there are models that predict preferentially in-plane spins [124, 125]. We also disfavour *isotropic* spins with *flat* or *high* magnitudes. Thus, the lack of observed precession points towards low spins, or spins preferentially aligned with the orbital angular momentum.

Previous constraints on spins have primarily been provided by considering the measurable aligned-spin component [121, 146, 122, 2] and provide strong evidence against all but *low aligned* or *iostropic* distributions, with *low isotropic* spins preferred. Combining the aligned spin and precession results will further restrict the spin distribution consistent with GW observations, and will likely require spin mag-

[1] As our analysis assumes zero noise, the fraction of binaries with observable precession will be slightly underestimated. At a threshold of $\rho_p > 2.1$, the effect would be significant while at a threshold of $\rho_p > 3$, the difference is small.

[2] We note here than under many astrophysical models these systems are thought to be unusual, especially by the time that they would be observable in the current LIGO/VIRG0 frequency ranges. There are however some models such as triples [124] where one might expect these high in-plane spins

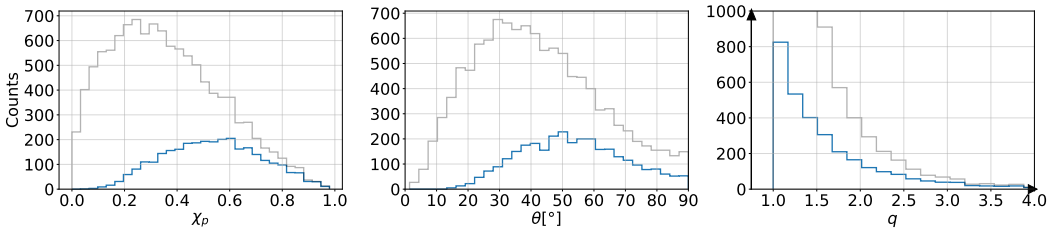

Figure 3.2: The distribution of χ_p, θ and q for observable binaries (grey), and those with measurable precession (blue), assuming a low isotropic spin distribution. θ is the inclination angle folded to $[0, \pi/2]$. The y-axis labels the number of observed events in each bin, out of 10^5 simulated signals with low isotropic spins.

nitudes even smaller than our *low* distribution (see also Ref. [74]).

3.3 Where will we observe precession?

We are able to identify, for the first time, the regions of parameter space that lead to signals with observable precession. In Fig. 3.2, we show the expected distribution of the precessing spin χ_p, binary orientation θ and mass ratio q for observable binaries and binaries with observable precession, $\rho_p > 3$, assuming a *low isotropic* distribution of spins. We identify clear regions of the parameter space where precession is more likely to be observed: large values of χ_p, binaries that are close to edge-on, $\theta > 45°$, and systems with high mass ratio. Regions where the chance of observing precession is close to zero include binaries with $\chi_p < 0.2$ or where the total angular momentum is within 20° of the line of sight. These results are consistent with expectations based upon smaller studies using detailed parameter estimation techniques [89]. We also note that most observations of precession will be in comparable-mass binaries, i.e., $q \leq 2$. This is a surprising, new result. It is well known that precession is more easily measured at higher mass ratios [62], which is confirmed by our study: precession is observed in <12% of detections with $q < 2$, but >35% for $q > 2$. However, with $\sim 90\%$ of observations expected to have $q < 2$, these low mass ratio observations vastly outnumber the higher-mass-ratio observations. This means that despite the probability for any single precession observation being lower for smaller mass ratios, as a population the many low probability observations will in aggregate produce more observations of measurable precession than higher mass ratio events (of which we expect considerably fewer). In this study we find that ∼75% of precession observations will come from detections of binaries with $q < 2$.

3.4 Discussion:

In this letter we have used a simple method to identify when precession is measurable in a compact binary GW signal. The gravitational waveform is well approximated by the first two harmonics in a power series expansion in the tangent of the half-opening angle [3], and the unambiguous observation of precession requires the identification of both of these harmonics in the data. The precession SNR ρ_p is a simple measure of this observability. We have demonstrated the efficacy of ρ_p through parameter estimation studies and also provided the distributions of ρ_p for the aLIGO-AdV observations to date. Using our definition of precession SNR, we have identified how often precession will be observed for a variety of potential astrophysical spin distributions. For the most likely distribution, based on current observations (*low-aligned*) there is a 83% chance that precession will be measured after $\sim$40 observations, and is therefore likely to be observed during the current third aLIGO-AdV observing run (O3). The non-measurement of precession by the end of O3 would place much stronger constraints on spin orientations and magnitudes.

The precession SNR has many applications. Most immediately, it allows us to determine the measurability of precession in a system *without performing computationally expensive parameter estimation.* This allows us to, e.g., easily fold precession information into population analyses of black-hole binaries. In future work, we will explore whether the value of ρ_p can be used to predict the measured χ_p distribution. The precessing SNR also gives us a simple way to identify regions of the parameter space where precession is important, a necessary first step in extending existing GW searches to explicitly use precessing waveforms [103].

Chapter 4

Identifying where precession is Measurable

4.1 Introduction

In September 2015, the first direct detection of gravitational-waves (GWs) marked the beginning of GW astronomy [116]. Another 14 detections have been announced by the LIGO Scientific and Virgo collaborations (LVC), the vast majority of which were due to black-hole (BH) mergers [2, 154, 1, 155, 156, 157]. Additional events have also been reported by independent groups [158, 159, 160, 161]. These GW observations have already provided significant insights into gravitational physics, cosmology, astronomy, nuclear physics and fundamental physics (see e.g. Refs. [162, 163, 164, 165, 166, 167, 168, 74]). With an order of magnitude more observations expected over the next 5-10 years, as the sensitivities of the LIGO [169, 151], Virgo [21] and KAGRA [170] detectors improve and additional detectors come online, GW astronomy from compact-binary mergers has the potential to transform our understanding of gravitational and fundamental physics [171, 172, 173].

Everything we learn from GW binary-black-hole (BBH) observations is a consequence of a detailed parameter estimation analysis that extracts the source parameters of the binary. While some parameters are extracted with good precision, inspiral dominanted signals show strong correlations between certain parameters which means that they cannot be measured so accurately, for example correlations between the binary's distance and inclination [105, 174, 83], the two masses [105, 76], and the mass-ratio and spin components aligned to the binary's orbital angular momentum [76, 79, 175, 176]. As well as studies of the inspiral, work has been done to extract the source properties for high mass signals dominated by the merger ringdown, see e.g. [177, 178, 91, 179].

Spin components misaligned with the binary's orbital angular momentum, leading to a precession of the binary's orbital plane and hence modulations of the amplitude and phase, have not yet been unambiguously measured in GW observations [2],

see Fig. 3.1. Precession effects and correlations with other parameters are understood in principle [62, 71] but since theoretical signal models of precessing binaries that include the merger and ringdown date from only shortly before the first detections [84, 180], we have less experience of when precession will be measurable, and what the impact will be on other parameter measurements.

The purpose of this chapter is to explore when precession will be measurable, and its impact on other parameter measurements, in the kind of configurations that are representative of expectations from binary populations based on LIGO-Virgo-KAGRA observations to date [2]. By utilizing the precession signal-to-noise ratio (SNR) ρ_p [3, 123] as a quantifier for the measurability of precession, we also verify that ρ_p is indeed a good metric for the measurability of precession across the vast majority of the parameter space, and relate it to the standard means to identify the presence of precession, the Bayes factor. In doing so, we show that computationally expensive parameter estimation runs can be avoided by simply calculating the precession SNR.

Previous work has explored the general phenomenology of precession effects: its increased measurability with large in-plane spins [181, 77, 182], large mass ratios [181, 77], high inclination [62, 97, 183, 184, 123, 91], and of course high SNR [181, 77, 185]. Beyond these general expectations, the *quantitative* behaviour of parameter measurements in the presence of precession has not been studied in great detail for typical LIGO-Virgo-KAGRA observations. The measurability of precession for high mass ratio LIGO-Virgo-KAGRA observations like GW190814 has been investigated in recent work [186].

In this chapter, we focus on the region of parameter space most likely to yield binaries with observable precession: binaries of comparable mass, with moderate in-plane and aligned-spin components [123]. We perform a series of one-dimensional investigations of the parameter space, in which we vary one parameter at a time: total mass, mass ratio, spins (both in-plane as characterized by χ_p, and the aligned spin combination $\chi_{\rm eff}$), the binary orientation (both the inclination of the orbit and also binary polarization), and the sky location and show the impact of varying each of the binary parameters individually. These investigations serve to confirm that much of the known phenomenology is apparent even at relatively low SNR, while also demonstrating that the precession SNR can be effectively used across a significant fraction of the parameter space to *predict* the observable consequences of precession *without* the need for computationally costly parameter estimation analyses.

This chapter is structured as follows: Sec. 4.2 provides an introduction to the parameter estimation techniques used here, and parameter estimation results and interpretation for our fiducial system. In Sec. 4.3 we perform a series of one-dimensional explorations of the parameter space. In Sec. 4.6 we compare the predicted precession SNR with observations and in Sec. 4.5 we compare precession SNR with the Bayes factors between precessing and non-precessing runs. We conclude with a summary

and discussion of future directions.

a Observability of precession

The strength of the modulations in the GW signal depend primarily on the opening angle, β, and this is reflected in the expansion parameter b in the two-harmonic approximation; the precession frequency $\dot{\alpha}$ also plays a role. The strength of the modulations in the *observed* signal also depend on the binary's inclination to the observer, $\theta_{\rm JN}$, and the detector polarisation ψ, and these are all incorporated into the precession SNR ρ_p, through Eqs. (2.33) and (2.39). From these we can draw immediate conclusions about the scenarios in which precession will be most easily measured. These observations are in general not new (see, as always, the pioneering discussions in Refs. [62, 71]), but we summarise them here and, where salient, present them in terms of the two harmonic formalism, which highlights the insights and intuition that are simplified in this formulation. We then compare these expectations with the quantitative results that we find in our full parameter estimation study.

Our first basic picture of the strength of precession effects comes from Eq. (2.2), which gives the dominant effect on β during the inspiral. If we first consider cases where the spin is entirely in the orbital plane, i.e., $S_{||} = 0$, we see that the opening angle β will be zero if $S_{\perp} = 0$ (as we would expect), and increases linearly for small $S_{\perp}$. The opening angle also increases as μ decreases, i.e., as the mass ratio is increased. Eq. (2.2) is no longer accurate near merger, and for equal-mass systems β does not become large, but for large mass ratios the opening angle can approach $90°$.

If we now consider non-zero $S_{||}$, we see that the level of precession will be reduced for systems with a positive aligned-spin component, and will be increased for systems with a negative aligned-spin component. The importance of this effect will depend on the other terms, but we can see that for a high-mass-ratio system where μ is very small, and close to merger, so rM is also small, the aligned-spin component will have a strong effect on β, and therefore the measurability of precession. A negative $S_{||}$ is necessary to achieve $\beta > 90°$, and for large mass-ratio systems near merger (small μ and rM) and large negative $S_{||}$, β can approach $180°$, but such systems will be rare.

The measurability of precession also depends on the orientation of the binary with respect to the detector, $\theta_{\rm JN}$. As we see in Eq. (2.33), precession effects will be minimal if $\theta_{\rm JN} \sim 0°$ or $180°$, i.e., the observer views the system from the direction of $\hat{\mathbf{J}}$. We expect precession to be strongest in the observed waveform for orientations close to $\theta_{\rm JN} \sim 90°$. Additionally, when the detector, or network is primarily sensitive to the $\times$ polarization, precession effects will be more significant. The amplitude of the $k = 1$ harmonic vanishes in the $+$ polarization for both face on $\theta_{\rm JN} = 0°$ and $180°$ and edge-on $\theta_{\rm JN} = 90°$ systems, while the $\times$ polarization is maximal for edge-

on systems. Additionally, the $\times$ polarization for the $k = 0$ harmonic vanishes for edge on systems, while the $+$ polarization is only reduced by a factor of two. Thus, even when b is small, there can be observable precession when the system is close to edge on and the network is preferentially sensitive to the $\times$ polarization. For a given choice of masses and spins, the maximum precession SNR is $\rho_p = \rho/\sqrt{2}$.

4.2 Parameter Estimation Results

a Standard configuration

We begin by describing the results of the parameter recovery routine for a specific simulated signal. The details of the signal are given in Tab. 4.1. These parameters were chosen so that precession effects would be significant in the observed waveform while still being consistent with the observed population of BBHs. In the following sections, we vary over the parameters of the signal one-by-one to investigate the impact of each parameter on the observability of precession and the accuracy of parameter recovery. For each parameter, we are able to both increase and decrease the significance of precession.

Using the first, second and third observing runs [57] to infer the mass for distribution of the BBHs observed in the it is predicted that 90% of detected binaries will have mass ratios $q < 4$ and $\sim 97\%$ of BHs in these binaries will have masses less than $45M_\odot$ [59]. Our "standard" simulated signal was chosen to have total mass $M = 40M_\odot$ and mass ratio $q = 2$ inclined at an angle of $\theta_{JN} = 60^\circ$. This corresponds to component masses of $26.7M_\odot$ and $13.3M_\odot$. This mass ratio and inclination was chosen to increase the observability of precession.

Of the 50 events reported by the LIGO/Virgo, 13 exclude the aligned-spin measure $\chi_{\rm eff} = 0$ at 90% confidence [187, 2, 57]. The other 37 observations peak at $\chi_{\rm eff} = 0$ [2, 57]. Based on this, studies have shown that it is likely BHs in binaries have low spin magnitudes [74, 121, 122, 123]. For this reason, in our standard configuration the BH spins were chosen such that there is zero spin aligned with the binary's orbital angular momentum, $\chi_{\rm eff} = 0$. We introduce precession by giving the more massive BH a spin of 0.4 in-plane and leaving the second BH with zero spin; two-spin effects are generally far weaker than the dominant precession effect, which exhibits the same phenomenology as a single-spin system [188, 110]. From Eq. (2.5) we see that this gives us a system with $\chi_p = 0.4$. The opening angle for the binary when the signal enters the detector's sensitivity band is 10° and the average value of the parameter $b = \tan(\beta/2)$ is $\overline{b} = 0.11$, from Eq . (2.27). The signal is simulated using the `IMRPhenomPv2` waveform model that incorporates precession effects, but not higher harmonics ($\ell > 2$) in the signal [84, 189].

Our "standard" simulated signal was chosen to be more favourable to precession measurements than typical LIGO-Virgo observations. Assuming systems are

distributed uniformly in binary orientation, masses drawn from a power law distribution and spins drawn from a low isotropic distribution (see Ref. [123] for details), we expect that 4 in every 100 binaries detected by LIGO-Virgo will be inclined at angles greater than 60° and have $\bar{b} > 0.11$.

The sky location of the binary was chosen to have RA $= 1.88\,\mathrm{rad}$, DEC $= 1.19\,\mathrm{rad}$ (we investigate the effect of this choice in 4.4 c). The coalescence time is $t = 1186741861$ GPS (corresponding to the merger time of GW170814 [190]). The polarization angle, defined by the orientation of the orbital plane when entering the sensitive band at 20Hz, is $\psi = 40°$. The two harmonic approximation is calculated in the J-aligned frame ($\hat{z} = \hat{\mathbf{J}}$). In this frame, the polarization angle is $\psi_J = 120°$, which gives antenna factors for H1 of $F_+ = 0.34$ and $F_\times = 0.53$ and for L1 of $F_+ = -0.45$ and $F_\times = -0.30$, thus both detectors are roughly equally sensitive to the two GW polarizations.

We injected the signals into zero noise. The zero-noise analysis results will be similar to those obtained from the average results of multiple identical injections in different Gaussian noise realisations. The simulated signal is recovered using the LIGO Livingston and Hanford detectors with sensitivities matching those achieved in the second observing run (O2) [2]. A low frequency cut-off of 20Hz was used for likelihood evaluations, this frequency is also used as the reference frequency when defining all frequency dependent parameters such as θ_{JN}. Both the LIGO Livingston and Hanford sensitivities improved prior to the third observing run (O3) [191] and are expected to improve further prior to the fourth observing run (O4) [192]. The results presented in this work are unlikely to be affected significantly by these changes and therefore we expect the main conclusions to be valid for O4 and beyond.

The SNR of the signal is fixed to be 20, corresponding to a moderately loud signal for aLIGO and AdV observations [192]. This sets the distance to $d_L = 223\,\mathrm{Mpc}$. The simulated SNR in the two detectors is 16.2 in L1 and 11.7 in H1. The simulated precession SNR in each of the detectors is 3.7 and 3.4 respectively, giving a network precession SNR of 5.0. Thus, we expect that precession will be clearly observable in this signal.

b Parameter Estimation Techniques

We will adopt a parameter estimation methodology that uses matched filtering with phenomenological gravitational waveforms and Markov Chain Monte Carlo (MCMC) techniques to sample the posterior.

We begin by introducing the matched filtering formalisation for parameter estimation. We assume that the time series received from the GW detectors can be decomposed as a sum of the GW signal, $h(t)$, plus noise, $n(t)$, which is assumed stationary and Gaussian with zero mean,

$$d(t) = h(t) + n(t). \tag{4.1}$$

Under the assumption of Gaussian noise, the probability of observing data d given a signal $h(\boldsymbol{\lambda})$ parameterised by $\boldsymbol{\lambda} = \{\lambda_1, \lambda_2, ..., \lambda_N\}$, otherwise known as the likelihood, is [27],

$$p(d|\boldsymbol{\lambda}) \propto / \exp\left(-\frac{1}{2}\langle d - h(\boldsymbol{\lambda})|d - h(\boldsymbol{\lambda})\rangle\right), \tag{4.2}$$

where $\langle a|b\rangle$ denotes the inner product between two waveforms a and b and is defined as in Eq. 1.9

The posterior probability density function (PDF) can then be computed through a simple application of Bayes' theorem,

$$\begin{aligned} p(\boldsymbol{\lambda}|d) &= \frac{p(\boldsymbol{\lambda})p(d|\boldsymbol{\lambda})}{p(d)}, \\ &\propto p(\boldsymbol{\lambda}) \exp\left(-\frac{1}{2}\langle d - h(\boldsymbol{\lambda})|d - h(\boldsymbol{\lambda})\rangle\right), \end{aligned} \tag{4.3}$$

where $p(\boldsymbol{\lambda}|d)$ is the posterior distribution for the parameters λ, $p(\boldsymbol{\lambda})$ is the prior probability distribution where $\int p(\boldsymbol{\lambda})d\boldsymbol{\lambda} = 1$, and $p(d)$ is the marginalised likelihood where $p(d) = \int p(\lambda_i)p(d|\lambda_i)d\lambda_i$. Posterior distributions for specific parameters can then be found by marginalising over all other parameters,

$$p(\lambda_i|d) = \int p(\boldsymbol{\lambda}|d)d\lambda_1...d\lambda_{i-1}d\lambda_{i+1}...d\lambda_N. \tag{4.4}$$

In the idealised situation of zero noise, Eq. (4.2) has a maximum at $h(\boldsymbol{\lambda}) = h(\boldsymbol{\lambda}_0)$. However, as can be seen in Eq. (4.3) the posterior also includes priors, this means that, as well as effects due to noise, certain priors may cause the maxima to be deflected away from $h(\boldsymbol{\lambda}) = h(\boldsymbol{\lambda}_0)$. This would then lead to Eq. (4.4) recovering a biased posterior. In this work, we consider the effect of three closely related priors,

- *Global*: the prior used during the parameter estimation analysis. This reflects our prior belief before observing any data,
- *Conditioned*: the global prior conditioned upon the posterior distributions of other parameters from the same analysis. For example since $\chi_{\rm eff}$ and $\chi_{\rm p}$ are correlated, any informative measurement of $\chi_{\rm eff}$ restricts the range of plausible values for χ_p and therefore modifies our prior beliefs about $\chi_{\rm p}$. The most simple case of this would be if the z components of spin for both black boles were 1 (the maximum value for spin in a black hole) then we know there can be no in-plane spin components for those black holes and as such our prior belief conditioned on a measurement of χ_{eff} is that $\chi_p = 0$.This prior has been used in previous LVC publications, see e.g. [2],
- *Informed*: the global prior conditioned upon the posterior distributions from a different analysis. Here, we use this to inform our expectations of the degree

	Simulated	Median		maxL	
	Precessing	Non-Precessing	Precessing	Non-Precessing	Precessing
Total mass $M/M_\odot$	40.0	40^{+3}_{-2}	40^{+4}_{-2}	40.161	40.507
Chirp mass $\mathcal{M}/M_\odot$	16.22	$16.5^{+0.3}_{-0.2}$	$16.3^{+0.3}_{-0.3}$	16.459	16.113
Mass ratio q	2.0	$1.8^{+0.8}_{-0.7}$	$1.9^{+0.9}_{-0.7}$	1.895	2.191
Inclination angle $\theta_{JN}/^\circ$	60.0	110^{+50}_{-100}	120^{+40}_{-90}	30.0	40.0
Precession phase $\phi_{JL}/^\circ$	45.0	–	200^{+100}_{-200}	–	80.0
Effective aligned spin, $\chi_{\rm eff}$	0.0	$0.044^{+0.099}_{-0.084}$	$-0.005^{+0.098}_{-0.092}$	0.06	−0.011
Effective precessing spin, χ_p	0.4	–	$0.5^{+0.4}_{-0.3}$	–	0.554
Right ascension RA/rad	1.88	3^{+3}_{-3}	3^{+3}_{-3}	1.418	1.325
Declination DEC/rad	1.19	$0.2^{+1.0}_{-1.2}$	$0.2^{+1.0}_{-1.2}$	1.229	1.221
Luminosity distance d_L/Mpc	223	500^{+200}_{-200}	400^{+200}_{-200}	451.834	372.706
Network SNR ρ	20.0	$19.3^{+0.1}_{-0.2}$	$19.7^{+0.2}_{-0.2}$	19.52	19.936
Precessing SNR ρ_p	5.05	–	4^{+2}_{-2}	–	4.649

Table 4.1: Table showing the simulated and inferred parameters for the "standard" injection when recovered by a non-precessing (`IMRPhenomD`) and a precessing (`IMRPhenomPv2`) waveform model. We report the median values along with the 90% symmetric credible intervals and the maximum likelihood (maxL) value.

of precession given the results from a non-precessing analysis. See Section 4.6 for details.

c Parameter recovery

We performed parameter estimation on the signal using the `LALInference` [52] and `LALSimulation` libraries within `LALSuite` [193]. Parameter recovery was performed with the `IMRPhenomPv2` model [189, 84], which matches the simulated signal to remove any systematic error caused by waveform uncertainty, and the corresponding `IMRPhenomD` aligned-spin waveform model [108, 107], which does not include any precession effects. Additionally, all analyses used exactly the same priors as those used in the LIGO-Virgo discovery papers, for details, see Appendix B.1 of [2]. All post-processing was handled by the `PESummary` python package[153].

Tab. 4.1 summarises the key results for the standard configuration. All uncertainties are the 90% symmetric credible intervals.

We begin by comparing the overall differences between parameter recovery with the precessing, `IMRPhenomPv2`, and non-precessing, `IMRPhenomD`, runs. From the table, we see that the maximum likelihood SNR for the non-precessing model is, as expected, lower than for the precessing waveforms. This can be easily understood from the two-harmonic approximation. Since the precessing waveform is well approximated by the sum of two non-precessing harmonics, we would expect the non-precessing recovery to accurately recover the more significant of these two. If that were the case, the we would expect that,

$$\rho_D^2 \approx \rho^2 - \rho_p^2, \tag{4.5}$$

and this is indeed the case, as $\rho_D = 19.52$, $\rho = 19.94$ and the recovered power in the second harmonic is $\rho_p = 4.6$. Furthermore, we see that the recovered waveforms con-

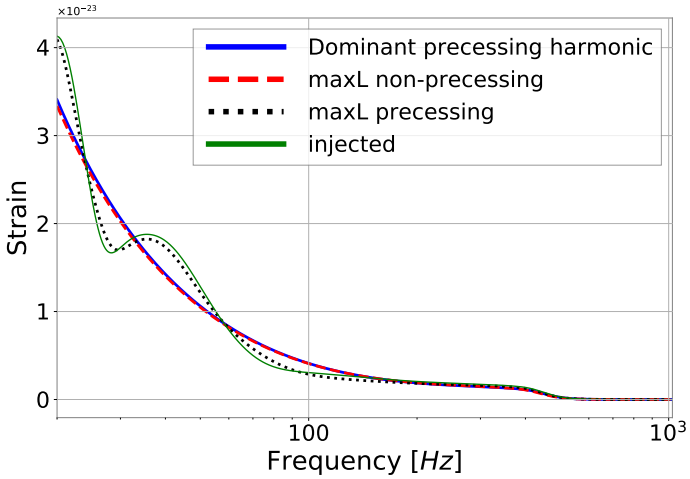

Figure 4.1: Comparison of the simulated precessing (green), non-precessing maximum likelihood (red), precessing maximum likelihood (black) and dominant precessing harmonic (blue) waveforms as a function of frequency. Waveforms are projected onto the LIGO Hanford detector.

firm this expectation: the recovered waveform when we include precession matches well with the simulated signal, while the non-precessing run recovers a waveform that matches the dominant harmonic, as show in Fig. 4.1.

We first consider the accuracy with which the masses and (aligned) spins are recovered. As expected, the chirp mass of the system is well recovered, in that it matches the simulated value with only a 2% uncertainty, which remains constant for both precessing and non-precessing runs. As is well known, there is a degeneracy between mass-ratio and spin, particularly during the inspiral part of the waveform [76, 79, 175, 176], which leads to significant uncertainty in both parameters. In Fig. 4.2 we show the recovery of the mass ratio and spin, for both precessing and non-precessing runs. When the model used to recover includes precession effects, the peak of the posteriors is located close to the simulated value ($\chi_{\rm eff} = 0$ and $q = 2.0$) and, while the degeneracy leads to significant uncertainty in both parameters, the mass-ratio distribution is clearly peaked away from $q = 1$. Interestingly, when we recover with a non-precessing waveform model, the inferred *aligned* spin component is systematically offset, with a peak at $\chi_{\rm eff} \approx 0.05$. This can be understood by recalling that precession induces a secular drift in the phase evolution of the binary, and this can be mimicked by a change in the value of the aligned spin [62, 3]. This discrepancy has not been seen in LIGO/Virgo observations [2] as we have not observed any systems with significant ρ_p (see Fig. 3.1). We investigate this further in Sec. b, where we study the effect of varying the mass ratio.

For non-precessing binaries, it is generally not possible to accurately recover the distance and orientation of the source, due to a well known degeneracy (see e.g., Ref. [83] for details), although the observation of higher signal harmonics can break

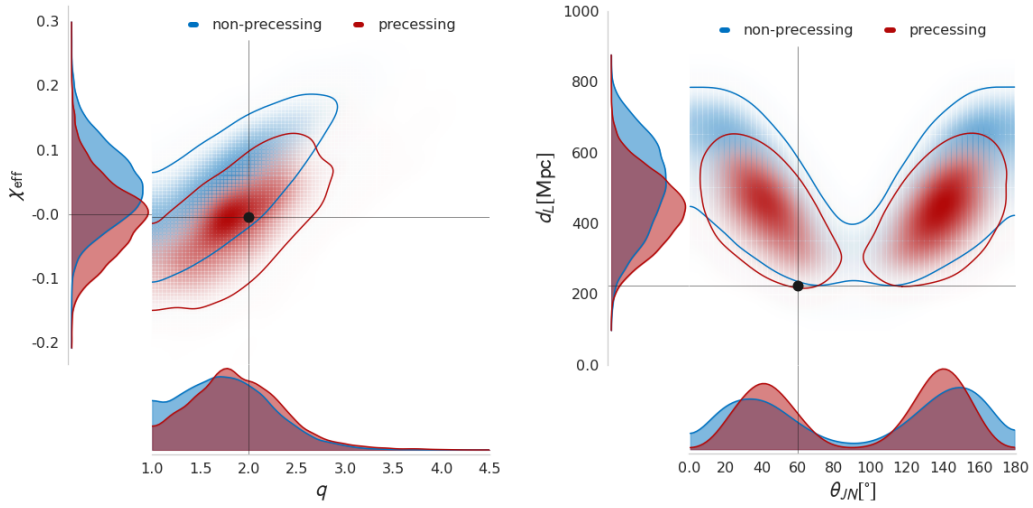

Figure 4.2: 2d contour comparing q–$\chi_{\rm eff}$ (left) and distance–inclination (right) degeneracies when precession effects are included. Contours show the 90% confidence interval. Bounded two-dimensional kernel density estimates (KDEs) are used for estimating the joint probability density. The black circle indicates the simulated values.

this degeneracy through an independent measurement of the source inclination [105, 83, 194]. Similarly, the observation of precession can break this degeneracy [195]. Precession causes an oscillation of the orbital plane leading to a time-dependence of the orientation of the orbital plane relative to the line of sight. Equivalently, in the two-harmonic picture, precession leads to the observation of a second harmonic and, consequently, additional constraints on the binary orientation as the amplitudes of the harmonics depend upon the viewing angle. In Fig. 4.2, we show the inferred two-dimensional distance and inclination posteriors for the precessing and non-precessing runs. As expected, the precessing run constrains the source to be away from face-on, while the non-precessing run simply returns the prior. However, even with observable precession, the simulated distance and orientation are not accurately recovered — a significant fraction of the posterior support is for a system at a greater distance and oriented closer to face-on. We will see how these measurements improve with stronger precession in Sec. a.

The sky location of the source is not well recovered. The analysis was performed with only the two LIGO detectors, and therefore we expect to recover the source restricted to a ring on the sky, which corresponds to a fixed time delay between the detectors [78, 81]. The location along the ring cannot be well constrained and, as expected the inferred location is preferentially associated with sky positions where the detector network is more sensitive. Thus, while the simulated sky location is within the 90% region, it is not at or close to the peak. This impacts the recovery of the distance, with the signal being recovered at larger distances, although the simulated distance remains within the 90% range. In Section c, we show results

from a set of runs with varying sky location, and verify that at sky locations where the network is more sensitive, the distance posterior is more consistent with the simulated value.

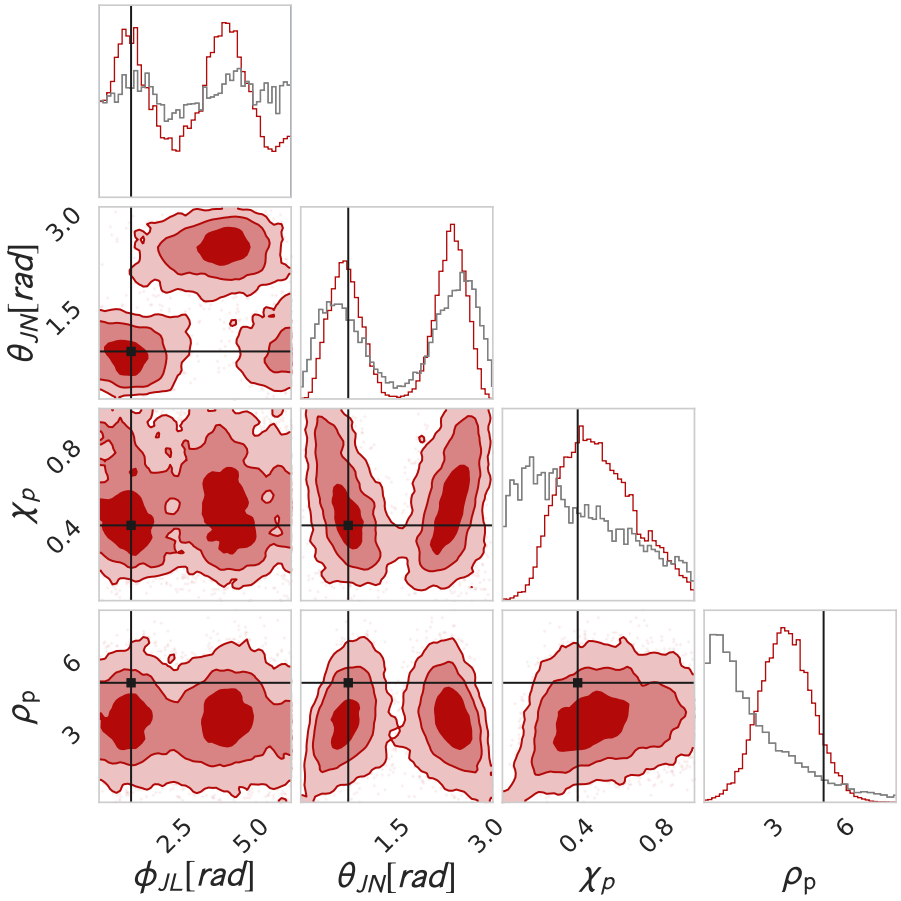

Figure 4.3: A corner plot showing the recovered values of binary orientation θ_{JN}, precessing spin χ_p, precession phase ϕ_{JL} and precession SNR ρ_p. Shading shows the 1σ, 3σ and 5σ confidence intervals. Black dots show the simulated values. The grey histograms show the *informed* prior, see Sec. 4.6. There is a clear correlation between the binary orientation and inferred precession spin, with signals which are close to face on ($\cos\theta \approx \pm 1$) having larger values of precessing spin, while those which are more inclined having less precessing spin. The precession SNR only weakly correlated with χ_p.

Lastly, we turn to measurement of precession. In Fig. 4.3 we show the recovered distributions for binary orientation, θ_{JN}, precessing spin χ_p, initial precession phase, ϕ_{JL}, and precession SNR, ρ_p. There is a clear correlation between the inferred orientation and χ_p, with binaries that are more inclined having lower values of χ_p. Neither of these quantities are directly observable, it is only the amount of observable precession in the system, encoded by ρ_p, that can be measured. Thus the orientation and spin must combine to give the right amount of power in precession, and we see that this is the case — there is little correlation between the recovered values of

ρ_p and the precessing spin χ_p. The inferred value of the precessing spin χ_p and precession SNR ρ_p are both consistent with the simulated values. Specifically, the signal has $\chi_p = 0.4$ and this is consistent with the recovered value, although the posterior distribution is broad, with support over essentially the entire range from 0 to 1. The precession SNR peaks well away from zero, giving clear indication of precession in the system. However, the peak of the distribution occurs at 3.5, while the simulated value is 5.0. We have deliberately chosen an event with significant observable precession. Only a small fraction of the parameter-space volume leads to such significant precession as shown by the *informed* prior on Fig. 4.3. This is calculated by estimating the allowed values of ρ_p conditioned on the measurements from a non-precessing analysis. See Sec. 4.6 for further details.

The precession phase, ϕ_{JL}, while not measured with great accuracy, does show two peaks, which are consistent with the simulated value of 45° (0.8 rad). The precession phase can be inferred from the relative phase of the two precessing harmonics using Eq. (2.33), provided the binary orientation is well measured. There is a clear dependence with the binary orientation: if $\theta_{JN} < 90°$ then the peak is in ϕ_{JL} at the simulated value and if it is greater then ϕ_{JL} is offset by 180°, to compensate for the change in sign of the $\cos\theta_{JN}$ terms in Eq. (2.33).

4.3 Impact of Varying Parameters

We now look at the effect of varying individual parameters one at a time on the recovered posteriors, in particular focusing on the measurement of precession as described by the posterior distributions of ρ_p and χ_p. All subsequent one-dimensional investigations of the parameter space maintain a constant SNR (except for Sec.a where the effect of the SNR is investigated). This is achieved by varying the distance to the source.

Primary results presented in this section will be displayed in the form of violin plots. We show the χ_p posterior distribution (left hand side, colored) compared to the global prior (right hand side, white) unless otherwise stated. We show the ρ_p posterior distribution as a single violin. Horizontal grey lines show the 90% symmetric credible interval. Horizontal red lines show the simulated value. A solid black line corresponds to the $\rho_p = 2.1$ threshold. Bounded kernel density estimates (KDEs)are used for estimating the probability density. We use the same 2d contour plots and multi-dimensional corner plots as described in Sec. c. Plots were generated with the PESummary [153] python package.

a In-plane spin components

We first look at the effect of varying the amount of precession in the system, varying χ_p from 0 to 1 in steps of 0.25. At $\chi_p = 1$ we have maximal spin, all in the plane

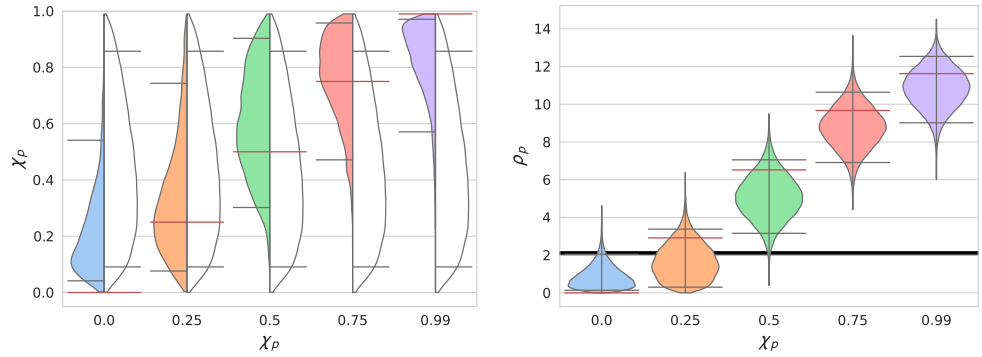

Figure 4.4: Violin plots showing the recovered posterior distributions distributions for χ_p compared to its prior (left) and ρ_p (right). Distributions are plotted for varying χ_p. Parameters other than χ_p match the "standard injection" (see Table 4.1)

of the binary. The inferred values of precessing spin and precession SNR are shown in Fig. 4.4. We observe, as expected, that increasing the in-plane spin leads to an increase in the magnitude of precession effects observable in the system. With zero precessing spin, there is no evidence for precession in the system; the recovered χ_p is consistent with zero[1]. Similarly, there is no support for significant precession SNR, with ρ_p constrained near zero. As χ_p increases, the amount of precession in the system grows and the measurement of χ_p becomes both more accurate and more precise. Fig. 4.4 shows the relationship between ρ_p and χ_p, and a larger value for ρ_p enables a better measurement for χ_p.

Fig. 4.5 shows how the inferred mass ratio–aligned spin and distance–orientation contours change as the magnitude of the in-plane spins change. When there is no observable precession in the system, there is a clear degeneracy in both cases. However, as precession effects become stronger the degeneracy between both pairs of parameters is broken. If ρ_p is small then this can be explained by both a small amount of precession observed at almost any inclination angle, or a large χ_p observed close to face on, as seen in Fig. 4.3. Since precession effects are not strong enough to provide an accurate measurement of the orientation, the degeneracy between distance and θ_{JN} persists. When ρ_p clearly excludes small values, there is *no support* for close to face-on signals, allowing a more precise measurement of the inclination angle θ_{JN}, breaking the degeneracy with distance.

Stronger precession also allows for improved measurement of the mass ratio. The opening angle β, and consequently the precession parameter $\bar{b}$, increases as the mass-ratio is increased, as can be seen from Eq. (2.2). Thus, when strong precession effects are observed, the signal is inconsistent with an equal mass system. In addition, the difference in frequency between the two leading precession harmonics depends upon the mass-ratio [3], and this may also improve our measurement of q. This can also

[1] We do not expect the χ_p posterior to contain $\chi_p = 0$ as there is no prior support there, however the posterior is relatively well constrained at low precession.

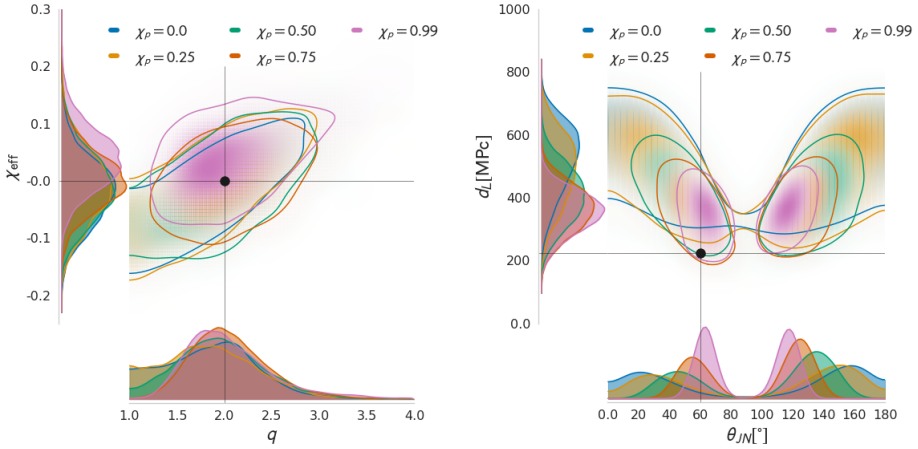

Figure 4.5: Two dimensional posteriors for (left) mass ratio and aligned spin, $\chi_{\rm eff}$, (right) binary orientation and distance. Contours show the 90% confidence interval. Bounded two-dimensional KDEs are used for estimating the joint probability density. The black circle with corresponding horizontal and vertical lines indicates the simulated values. For the simulated distance, a solid horizontal band indicates the maximum and minimum simulated values.

be seen from the precession dynamics, where the precession rate of L around J, $\dot{\alpha}$, depends the mass ratio, and the number of observable precession cycles corresponds to improved accuracy in the measurement of the mass ratio [92].

As χ_p is increased, the peak of the recovered ρ_p distribution is closer to the simulated value. This is likely due to a better measurement of the binary orientation as shown in Fig. 4.5.

b Inclination

It is well known that the inclination angle will affect our ability to measure precession. In particular, from Eq. (2.33) we see that in the two-harmonic approximation the second harmonic vanishes when $\theta_{\rm JN} = 0°$ or $180°$. In this section we consider the effect of changing the orientation of our standard configuration, which allows us to quantify how it will manifest in realistic LIGO-Virgo signals. A related study has looked at the effect at higher mass ratios [186].

The effect of varying $\theta_{\rm JN}$ is shown in Fig. 4.6. For binaries where the total angular momentum is nearly aligned with the line of sight, precession effects are not observable, as is clear from both the ρ_p and χ_p posteriors. It is not until $\theta_{JN} \geq 40°$ that we begin to be able to see that the median value for ρ_p is larger than 2.1 which is the 90% value for the χ^2 distribution. For 90% of the probability density to be larger than 2.1, we need inclinations around 70°, this is then however a very stringent threshold for detection. Although the accuracy of the measurement clearly improves as we increase θ_{JN}, the uncertainty in the measurement of χ_p remains large and even at $\theta_{JN} = 90°$ the posterior is very broad. This can be understood

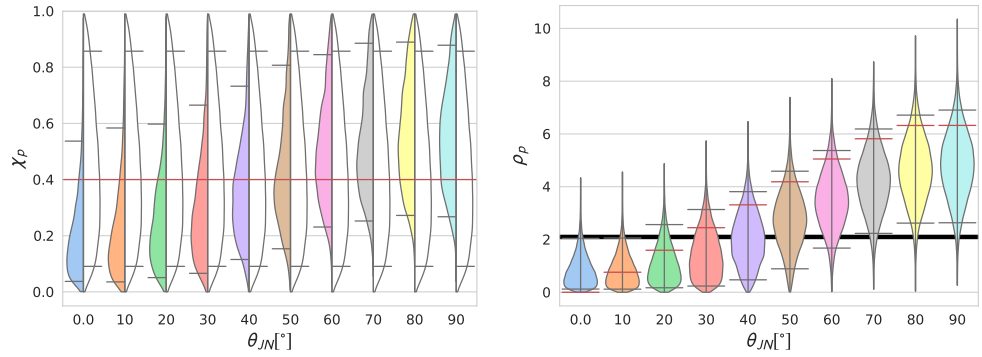

Figure 4.6: Violin plots showing the recovered posterior distributions distributions for χ_p compared to its prior (left) and ρ_p (right). Distributions are plotted for varying θ_{JN}. Parameters other than θ_{JN} match the "standard injection" (see Table 4.1)

by considering the degeneracies shown in Fig. 4.3 for the standard signal and in Fig. 4.7 for the $\theta_{JN} = 90°$ signal. In both cases, the measured quantity, ρ_p, is relatively well constrained but neither the binary orientation nor χ_p are accurately measured. The observed precession is consistent with both a highly inclined system with lower precessing spin (i.e., low χ_p and large θ_{JN}) or by a less inclined system with higher precessing spin (i.e., high χ_p and small θ_{JN}). Both of these will produce similar observable effects in the waveform.

This allows us to explain the measured posterior for χ_p. At low inclination the posterior is consistent with small values of χ_p. While we are unable to rule out large χ_p, there is limited support as it would require the system to be observed very close to face-on, otherwise precession effects become significant. At large values of θ_{JN}, when precession is clearly observable in the signal, $\chi_p = 0$ is excluded but the distribution remains broad and extends to $\chi_p = 1$.

c Total mass

We now vary the total mass of the system, keeping all other parameters including mass ratio fixed, in steps of $20\,M_\odot$. As before, we keep the SNR of the system constant at 20, so the higher mass systems are generated at a greater distance. The inferred distributions for χ_p and ρ_p are shown in Fig. 4.8.

As the total mass of the source increases, the length of the waveform decreases, as does the number of precession cycles, with the number scaling approximately inversely to the total mass (see Eq. (45) of [62]). From the two-harmonic perspective, a small number of precession cycles leads to a large overlap between the harmonics. Specifically, for the $M = 100M_\odot$ system the overlap between the normalised harmonics is $\langle \hat{h}_0 | \hat{h}_1 \rangle = 0.77$, where $\hat{h} = h/|h|$ and the inner product is defined in Eq. (1.9). At $M = 20M_\odot$, the harmonics are close to orthogonal with $\langle \hat{h}_0 | \hat{h}_1 \rangle = 0.15$. The opening angle doesn't change significantly, with $\bar{b} = 0.14$ at $M = 20M_\odot$ and $\bar{b} = 0.21$ at $M = 100M_\odot$.

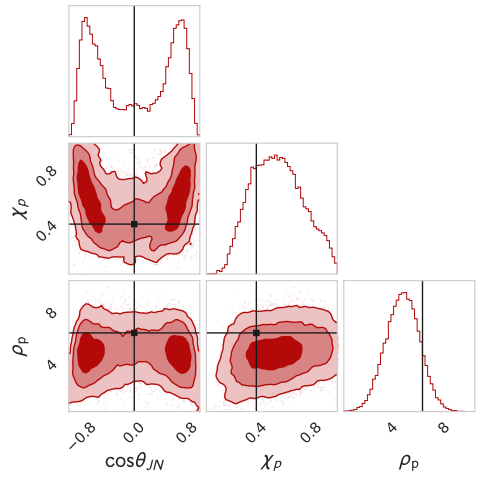

Figure 4.7: A corner plot showing the recovered values of binary orientation θ_{JN}, precessing spin χ_p and precession SNR ρ_p for a system simulated at edge on. Shading shows the 1σ, 3σ and 5σ confidence intervals. Black dots show the simulated values, We see the strong correlation between θ_{JN} and χ_p reflecting the measurement of a certain ρ_p

At lower masses, $M \leq 40M_\odot$, while the precessing spin is not tightly constrained, it is clearly restricted to be non-zero and the precession SNR has essentially no support for $\rho_p = 0$. For the $60M_\odot$ and $80M_\odot$ mergers, the precessing spin is still peaked close to the simulated value while ρ_p peaks above 2.1 showing evidence for observable precession, although both ρ_p and χ_p distributions do extend to zero.

For the high-mass system, $M = 100M_\odot$, the χ_p posterior more closely matches the prior and we are unable to exclude $\chi_p = 0$. The inferred ρ_p distribution peaks close to zero, and is consistent with no precession, even though the precession SNR in the simulated signal is similar to the lower mass signals. This is likely due to the breakdown of the two-harmonic approximation for this short signal. In particular, for a high-mass system, the power orthogonal to the leading harmonic will depend sensitively upon the initial precession phase ϕ_{JL}. The fact that the recovered value of ρ_p is inconsistent with the simulated value may be due to this fact: the value of $\phi_{JL} = 45°$ used in the simulation leads to maximal observable precession. Across the full parameter space there are very few configurations with significant precession, so this observation is dis-favoured by our priors. We explore the prior effects such as this in detail in Sec. b.

d Polarization

The effect of changing the relative sensitivity to the two GW polarizations is clear from Eq. (2.33). Recalling that $\bar{b} = 0.11$ and $\theta_{JN} = 60°$, we can express ζ (the ratio

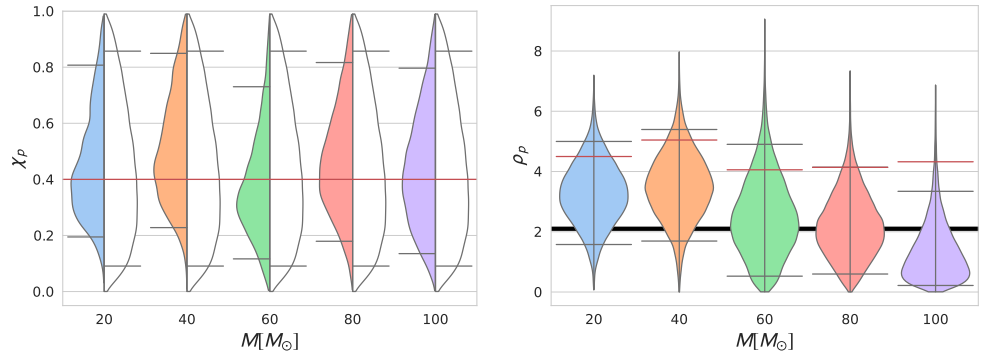

Figure 4.8: Violin plots showing the recovered posterior distributions distributions for χ_p compared to its prior (left) and ρ_p (right). Distributions are plotted for varying total mass. Parameters other than the total mass of the signal match the "standard injection" (see Table 4.1)

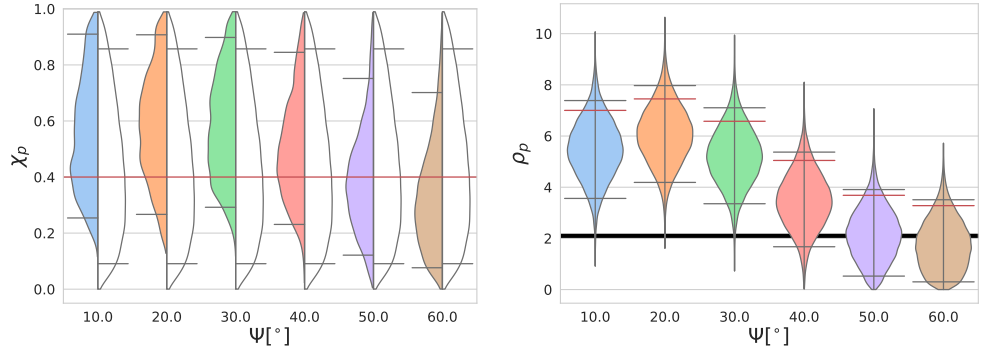

Figure 4.9: Violin plots showing the recovered posterior distributions distributions for χ_p compared to its prior (left) and ρ_p (right). Distributions are plotted for varying ψ_J. Parameters other than ψ_J match the "standard injection" (see Table 4.1)

of the amplitudes of the two harmonics) as

$$|\zeta| = 0.15 \left| \frac{F_+ + 2iF_\times}{1F_+ + 0.8iF_\times} \right| ,$$

Thus, ζ, and consequently the imprint of precession on the waveform, will be maximized when the detector network is primarily sensitive to the $\times$ polarization and minimized when the network is sensitive to the $+$ polarization. We can investigate this by varying the polarization angle of the simulated signal, in steps of 10° from the "standard" value of 40°. At $\psi = 40°$, the sensitivity to the two polarizations is approximately equal, $|F_\times|/|F_+| = 0.9$. It is largest for $\psi = 20°$ where $|F_\times|/|F_+| = 25$ and smallest for $\psi = 60°$ where $|F_\times|/|F_+| = 0.04$. This leads to a variation in the precession SNR from $\rho_p \approx 3$ to $\rho_p \approx 7$.

In Fig. 4.9 we show the recovered posteriors for χ_p and ρ_p for a set of runs where the precession is varied. The precession SNR varies in accordance with expectation — it is largest at $\psi = 20°$, where the median of the posterior is at $\rho_p = 6$ and there

is no support for non-precessing systems, and smallest at 60° where the posterior extends down to $\rho_p = 0$. The amount of observable precesssion directly impacts the inferred distribution for ρ_p. For the $\psi = 60°$ signal, the posterior for χ_p is consistent with zero, or small in-plane spins, and large values are excluded. Meanwhile for $\psi = 20°$, $\chi_p < 0.1$ is excluded while extremal in-plane spins are consistent with the observation.

It is well known that precession leaves a stronger imprint upon the × polarization. However, we are not aware of previous results showing how simply changing the polarization of the system can so dramatically change the observable consequences of precession — from being barely observable when the observed signal is primarily the + polarization to being strongly observed in ×. Using the two-harmonic approximation, we are able to straightforwardly predict this effect and then verify it with detailed parameter estimation studies.

4.4 Additional Results

In this section I present results from [141] which are relevant to the work but were produced by other authors on the paper

a SNR

We now start with the fiducial run configuration described above and vary the SNR of the simulated signal.

In the strong-signal limit, where the likelihood surface can be well approximated by a multivariate gaussian, it is well known that the accuracy with which parameters can be measured is generally inversely proportional to the SNR [105, 76]. However, this is not always the case due to, for example, degeneracies between parameters (see Ref. [196] for a discussion of the limits of this approximation).

Fig. 4.10 shows that as the SNR of the simulated signal increases, the accuracy and precision of the inferred χ_p posterior distribution improves. As expected the width of the 90% credible interval decreases approximately linearly with increasing SNR. The improvement in the χ_p posterior distribution can be mapped to a linear increase in ρ_p.

When the simulated signal has low SNR ($\rho = 10$), the recovered χ_p posterior distribution resembles the prior, implying that there is no information about precession in the data. For this case, ρ_p matches the expected distribution in the absence of any measurable precession — a χ distribution with 2 degrees of freedom. As the SNR increases ($\rho = 20$-30), the 5th percentile of the the ρ_p distribution is comparable or greater than the $\rho_p = 2.1$ threshold. This maps to the χ_p posterior distribution removing all support for near-zero χ_p ($\chi_p \lesssim 0.1$). For larger SNRs ($\rho > 40$), the entire ρ_p distribution is greater than the 2.1 threshold. This implies significant power from

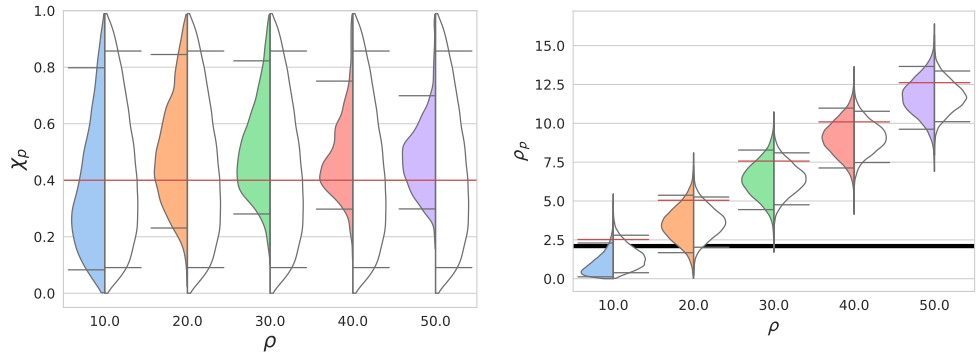

Figure 4.10: Violin plots showing the recovered posterior distributions distributions for χ_p compared to its prior (left) and ρ_p compared to a non-central χ distribution with 2 degrees of freedom and non-centrality equal to the median of the ρ_p distribution (right). Distributions are plotted for varying SNR. Parameters other than the SNR of the signal match the "standard injection" (see Table 4.1).

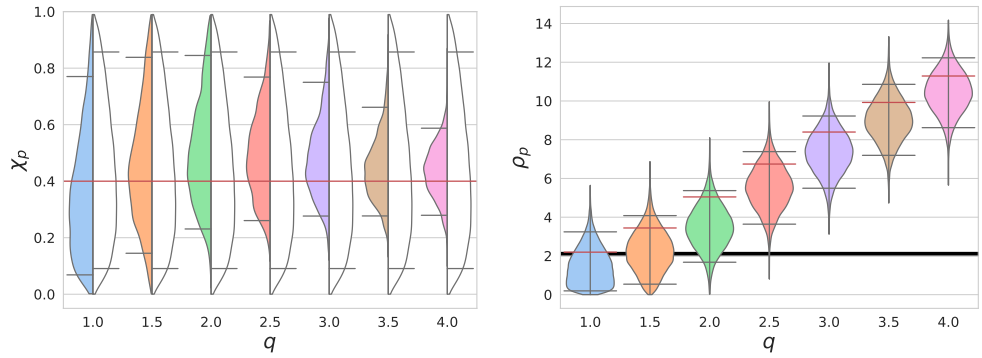

Figure 4.11: Violin plots showing the recovered posterior distributions distributions for χ_p compared to its prior (left) and ρ_p (right). Distributions are plotted for varying mass ratio. Parameters other than the mass ratio of the signal match the "standard injection" (see Table 4.1).

precession. For these cases, we remove support for maximal precession $\chi_p \sim 1$.

As expected we find good agreement between ρ_p and a non-central χ distribution with 2 degrees of freedom and non-centrality equal to the inferred power in the second harmonic (median of the ρ_p distribution).

b Mass ratio and aligned spin

Fig. 4.11 shows how the inferred precessing spin and precession SNR varies with the mass ratio of the system. As expected the mass ratio increases, an in-plane spin on the larger BH leads to a larger opening angle and more significant precession effects. For near equal-mass systems ($q \lesssim 1.5$), the inferred χ_p posterior distribution resembles its prior, and there is not significant power in precession, as shown by the value of ρ_p. As the mass ratio increases, the inferred power in precession also increases and for $q \gtrsim 2.5$, the 90% credible interval of the inferred ρ_p distribution is

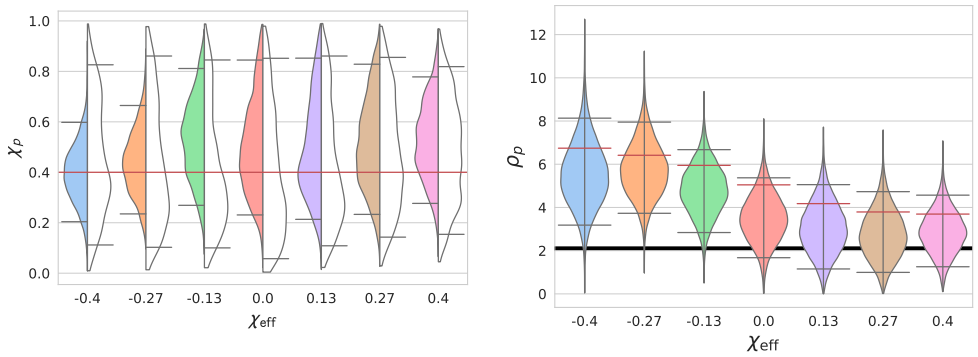

Figure 4.12: Violin plots showing the recovered posterior distributions distributions for χ_p compared to its prior conditioned on the $\chi_{\rm eff}$ and mass ratio posterior distributions (left) and ρ_p (right). Distributions are plotted for varying $\chi_{\rm eff}$. Parameters other than the $\chi_{\rm eff}$ of the signal match the "standard injection" (see Table 4.1).

entirely above $\rho_p = 2.1$. At this stage, precession is clearly identified and $\chi_p \approx 0$ is clearly excluded. In addition, the maximum value of χ_p is also bounded away from maximal.

Fig. 4.12 shows how varying $\chi_{\rm eff}$ affects our ability to measure precession. A system with a large negative $\chi_{\rm eff}$ results in a larger opening angle compared to an equivalent system with a large positive $\chi_{\rm eff}$. Thus, based upon Eq. (2.2), we expect the observable impact of precession to be greater for negative values of $\chi_{\rm eff}$ and smaller for positive values. The results are consistent with this expectation, in that the precession SNR decreases with increasing $\chi_{\rm eff}$ and the width of the recovered χ_p distribution increases. However, for the $\chi_{\rm eff} = 0.4$ analysis, we find that the range of χ_p is restricted, with both $\chi_p = 0$ and $\chi_p = 1$ excluded. This is *not* due to the measurement of precession, but is actually due to the measured non-zero aligned-spin component.

A non-zero measurement of $\chi_{\rm eff}$ forces $\chi_p < 1$ as the primary and secondary spin magnitudes must be less than unity. For example, in the $\chi_{\rm eff} = 0.4$ analysis, we measure $\chi_{\rm eff} = 0.38^{+0.07}_{-0.07}$. Under the single spin assumption, this limits $\chi_p < 0.95$. Similarly, since we are using prior distributions that are uniform in spin magnitude and orientation, the observation of a large aligned spin component leads to greater support for a large in-plane spin component. This is shown in Fig. 4.13, where we plot both the uninformed prior on the primary spin as well as the prior conditioned on $\chi_{\rm eff} = 0.4$, which removes all support for $\chi_p \approx 0$.

The χ_p measurement for the $\chi_{\rm eff} = 0.27$ and 0.4 analyses are similar to the conditional prior but do restrict the lower χ_p bound beyond prior effects. Although the distribution for ρ_p does extend to zero, it still peaks above $\rho_p = 2.1$ indicating some evidence, although not particularly strong, for precession.

As we vary the mass ratio and aligned spin, the length of the waveform will change. In particular, the aligned spin and high mass ratio configurations produce

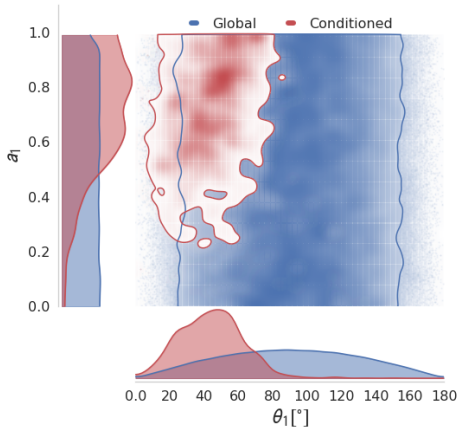

Figure 4.13: 2d contours showing the prior 90% credible interval over the primary spin magnitude and spin direction parameter space. Blue shows the global prior and red shows the global prior conditioned on the $\chi_{\rm eff} = 0.4$ mass ratio and $\chi_{\rm eff}$ posterior distributions

longer waveforms than those with anti-aligned spins and equal masses [197]. In principle, this will impact the measurability of precession, as longer waveforms allow for a greater number of precession cycles in the detectors' sensitive band. For very short signals, with less than one precession cycle in band, the two leading harmonics are no longer orthogonal (or even approximately so), which make it more challenging to unambiguously identify the second harmonic. This is not an issue for the signals considered here, but does become important when we vary the mass of the binary in Section c. With a greater number of precession cycles, we will also be able to more accurately measure the precession frequency (the frequency difference between the harmonics), which may improve the measurement of mass ratio [92]. However, it is still the precession SNR that determines the observability of precession. Finally, we note that changing the mass ratio and aligned spin will change the overall amplitude of the waveform. Since our study is performed at a *fixed SNR*, this simply leads to the signals being placed at a larger or smaller distance and therefore doesn't impact the results presented here.

c Sky Location

We performed a series of runs where we altered the sky location of the signal, keeping the masses and spins of the components fixed. We also maintained the binary orientation $\theta_{\rm JN} = 60°$, but varied the distance and polarization of the source to ensure that the SNR remained constant and that the relative contribution of the $+$ and $\times$ polarizations was consistent with the standard run. Furthermore, sky locations were restricted to those for which the relative time of arrival between the Hanford and Livingston detectors remains the same (i.e., we were sampling from the

Label	RA/rad	DEC/rad	$\psi/^\circ$	d_L/MPc	ρ_p	d_L/MPc
A	0.31	0.92	320	370	5.02	480^{+130}_{-180}
B	0.80	1.15	345	320	5.09	470^{+140}_{-160}
C	1.31	1.22	10	280	5.11	450^{+150}_{-160}
D	1.88	1.19	40	220	5.05	430^{+160}_{-160}
E	6.11	0.21	40	310	5.09	440^{+150}_{-170}

Table 4.2: Table showing the simulated parameters for the sky location set (see Sec. c). All other parameters match the "standard injection" (see Table 4.1). The recovered luminosity distance (far right column) is also shown.

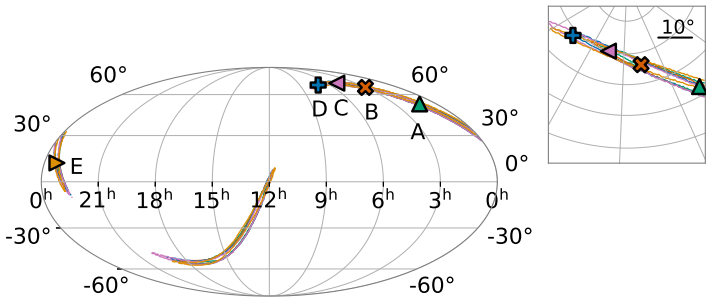

Figure 4.14: Skymap showing the different simulated sky positions, see Table 4.2. The solid lines show the 90% credible intervals and the markers show the simulated sky position. Their respective colors matches their corresponding credible intervals. We vary the distance and polarization of the source to ensure that the SNR remains consistent with the standard injection in Table 4.1.

nearly degenerate ring on the sky of constant time delays). Details of the runs are given in Tab. 4.2.

Table 4.2 shows that the inferred luminosity distance remains approximately constant despite the simulated luminosity distance varying by almost a factor of two. In addition, the recovered ρ_p distribution remains consistent with the "standard" injection. Fig. 4.14 shows that the inferred sky position of the source remains essentially unchanged, and consistent with locations of the detectors' greatest sensitivity. We note here that for this study we only considered the two detector LIGO network. Including VIRGO would likely have considerably improved the precision of the inferred sky location. We do not expect that this would affect any of the inferred physical parameters or any of the main conclusions in this work.

4.5 Relating the precessing SNR to Bayes Factors

An alternative method for identifying evidence for precession can be calculated within the Bayesian framework. We can calculate the Bayes factor, $\mathcal{B}$, by comparing the marginalized likelihoods (see Eq. (4.3)) from two competing hypotheses (A, B) [198],

$$\ln \mathcal{B} = \ln p(d_A) - \ln p(d_B). \tag{4.6}$$

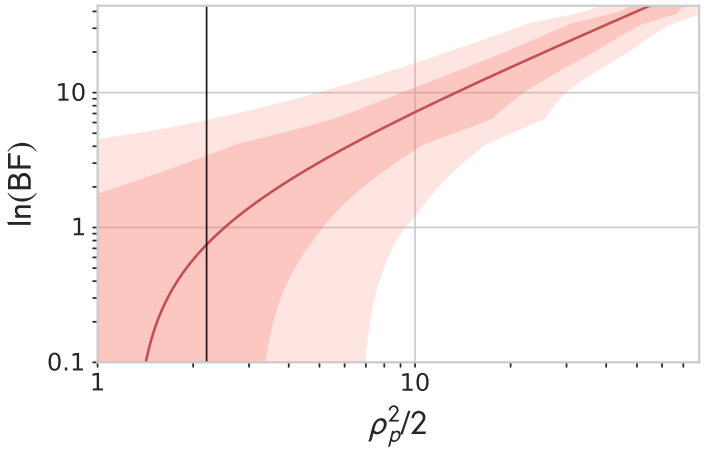

Figure 4.15: Plot comparing the Bayes factor in favour of precession to the inferred ρ_p distribution. Bayes factors were calculated by comparing the evidences for a precessing analysis and a non-precessing analysis. The uncertainties on the Bayes factors are calculated by taking the 90% confidence interval across multiple LALInferenceNest chains. The solid line uses the median of the ρ_p distribution. The shading gives the 1σ and 2σ uncertainties on the ρ_p measurement. The solid black lines shows the $\rho_p = 2.1$ threshold.

Bayes factors have thus far been the gold standard for identifying evidence for precession within the GW community and have been used extensively in previous works, see e.g., Ref. [186].

In the same way that Bayes factors can be used to quantify evidence for precession, it is also possible to quantify the significance of a GW signal by calculating the Bayes factor for signal verses noise [51]. It has been shown that the log Bayes factor for signal versus noise scales approximately with ρ^2 [199]. Here, we investigate the relationship between the Bayes factor in favour of precession and the precession SNR ρ_p. Both of these quantities have been used together in recent works when assessing the evidence for observable precession [1, 155, 186]

For a subset of the runs described in Section. b, we reran the analysis using the aligned-spin waveform model `IMRPhenomD`. Bayes factors in favour of precession could then be calculated and compared to the derived ρ_p posterior distributions.

Fig. 4.15 shows an approximately linear relationship between the log Bayes factor ($\ln \mathrm{BF}$) and the square of the precession SNR (ρ_p^2). This is expected given that the likelihoods recovered from the precessing waveform model will be larger than the likelihoods recovered from the aligned-spin waveform model by a factor of $\exp(\rho_p^2/2)$.

The commonly used heuristic when assessing the strength of evidence using Bayes factors is that $1 \leq \ln \mathrm{BF} \geq 3$ is marginal evidence and $\ln \mathrm{BF} > 3$ is strong evidence in favour of a hypothesis. From the line of best fit in the plots above we conclude that if 90% (50%) of the ρ_p posterior distribution is above the $\rho_p = 2.1$ threshold, this corresponds to a $\ln \mathrm{BF} \approx 3.5$ ($\ln \mathrm{BF} \approx 0.8$) and is therefore very strong (marginal)

evidence for precession. The posterior distribution on ρ_p can therefore be approximately mapped to the commonly used $\ln \mathrm{BF}$. We note here that the mapping, although clearly linear, is relatively uncertain at the moment and all quantitative maps contain a large amount of uncertainty. This uncertainty could be reduced by including more runs in the analysis, however for computational reasons we avoided multiple Bayes factor calculations. If one wantes to map them precisely then it would be simple enough to increase the sample size, this would allow for precisely assessing the strength of evidence for precession using ρ_p and would reduce the need for additional parameter estimation runs using non-precessing models, which are *necessary* to compute the Bayes factor. This reduction in computational cost will not be significant for a single event, but for population analyses and large scale PE studies this alternative metric could be extremely useful.

4.6 Predicting the Precession SNR Posterior

For the majority of simulations presented in this chapter, the distribution for the precession SNR, ρ_p, has been peaked significantly below the simulated value, although in nearly every case the simulated value does lie within the 90% confidence region. While the naive expectation is that the recovered posterior will peak at the simulated value, for complex parameter recovery where there are dependencies and degeneracies between the different parameters, this is often not the case. We have already seen that the distance is typically over-estimated in the simulations we have performed — this is a well-known effect and arises for two reasons, first that the network is less sensitive to sources from the chosen sky location than from other locations consistent with the observed signal (as discussed in Sec. c), and second that the signal was simulated significantly inclined from face-on, yet preferentially recovered close to face-on (as discussed in Sec. b). Similarly, it seems likely that the signals we have simulated have more significant precession effects (deliberately, as we wish to understand the observability of precession) than the vast majority of possible sources. Thus, our conjecture is that the likelihood peaks at the simulated value of ρ_p but the posterior distribution will be biased to recover a smaller value owing to the much larger volume of parameter space consistent with low ρ_p. To demonstrate this, we calculate a prior distribution for ρ_p which uses the information gleaned from a non-precessing analysis to take into consideration the much larger volume of parameter space consistent with low ρ_p. We then show that when multiplying the likelihood by the prior, the *predicted* posterior for ρ_p agrees well with the *inferred* posterior from a fully precessing parameter estimation analysis.

Let us first show that the likelihood peaks at the simulated value of ρ_p. The two-harmonic approximation allows us to factorize the likelihood in Eq. (4.7) into two terms: a non-precessing component (dependent on h_0) $\Lambda_{\mathrm{np}}(\boldsymbol{\lambda})$ and precessing component (dependent on h_1) $\Lambda_{\mathrm{p}}(\boldsymbol{\lambda})$,

$$
\begin{aligned}
p(d|\boldsymbol{\lambda}) &\propto \exp\left(-\frac{1}{2}\langle d-(\mathcal{A}_0(\boldsymbol{\lambda})h^0(\boldsymbol{\lambda})+\mathcal{A}_1(\boldsymbol{\lambda})h^1(\boldsymbol{\lambda}))|d-(\mathcal{A}_0(\boldsymbol{\lambda})h^0(\boldsymbol{\lambda})+\mathcal{A}_1(\boldsymbol{\lambda})h^1(\boldsymbol{\lambda}))\rangle\right) \\
&\propto \exp\left(\langle d|\mathcal{A}_0(\boldsymbol{\lambda})h^0(\boldsymbol{\lambda})\rangle-\frac{|\mathcal{A}_0(\boldsymbol{\lambda})|^2}{2}\langle h^0(\boldsymbol{\lambda})|h^0(\boldsymbol{\lambda})\rangle\right)\times \\
&\quad \exp\left(\langle d|\mathcal{A}_1(\boldsymbol{\lambda})h^1(\boldsymbol{\lambda})\rangle-\frac{|\mathcal{A}_1(\boldsymbol{\lambda})|^2}{2}\langle h^1(\boldsymbol{\lambda})|h^1(\boldsymbol{\lambda})\rangle\right) \\
&\propto \Lambda_{\rm np}(\boldsymbol{\lambda})\times\Lambda_{\rm p}(\boldsymbol{\lambda}),
\end{aligned}
\tag{4.7}
$$

For simplicity we use the approximations that $\langle h^0|h^1\rangle = 0$ and that h^0 is the dominant harmonic, i.e., that the SNR in the h^0 harmonic is larger than in h^1. The calculation proceeds analogously when h^1 is dominant, and can be extended to the general case by replacing h^1 by its projection onto the space orthogonal to h^0.

We can re-express the precessing contribution to the likelihood Λ_p in terms of the precession SNR using Eq. (2.39). To do so, we introduce $\hat{\rho}_p$ which is the simulated value of ρ_p, and $\rho_p(\boldsymbol{\lambda})$ which is the precession SNR for the set of parameters $\boldsymbol{\lambda}$. Furthermore, we define the simulated phase (as given in Eq. (2.32)) of the precession harmonic as $\hat{\phi}_1$ and the phase associated with the parameters $\boldsymbol{\lambda}$ as $\phi_1(\boldsymbol{\lambda})$. Following the procedure described in, e.g. Ref. [200], we can rewrite the precession likelihood as

$$
\Lambda_{\rm p}(\rho_p,\phi_1)\propto\exp\left(-\tfrac{1}{2}\left(\rho_p^2(\boldsymbol{\lambda})-2\hat{\rho}_p\rho_p(\boldsymbol{\lambda})\cos(\hat{\phi}_1-\phi_1)+\hat{\rho}_p^2\right)\right). \tag{4.8}
$$

In general, we have no prior knowledge of the precession phase, so it is natural to assume a uniform prior on ϕ_1. We may then analytically marginalise $\Lambda_{\rm p}(\rho_p,\phi_1)$ over ϕ_1 to obtain,

$$
\begin{aligned}
\Lambda_{\rm p}(\rho_p) &\propto \int_0^{2\pi}\Lambda_{\rm p}(\rho_p,\phi_1)\,p(\phi_1)\,d\phi_1 \\
&\propto I_0(\hat{\rho}_p\,\rho_p)\exp\left(-\frac{\hat{\rho}_p^2+\rho_p^2}{2}\right).
\end{aligned}
\tag{4.9}
$$

We therefore see that the precession likelihood peaks at $\hat{\rho}_p$. We may then calculate the posterior distribution for ρ_p using Bayes' Theorem,

$$
p(\rho_p|d)\propto p(\rho_p)\Lambda_{\rm p}(\rho_p)\,, \tag{4.10}
$$

where $p(\rho_p)$ is the prior for the precession SNR.

Previously, in Ref. [3], we obtained a distribution for $p(\rho_p|d)$ by maximising the likelihood over $\mathcal{A}_1$. This is equivalent to assuming uniform priors for the real and imaginary components of $\mathcal{A}_1$, and leads to a prior $p(\rho_p)\propto\rho_p$. It follows from Eq. 4.10 that this results in a χ^2 distribution with 2 degrees of freedom. Here, we instead use a prior for ρ_p which is informed by the information obtained from

a non-precessing analysis, we refer to this as the *informed* prior. This *informed* prior better represents our prior knowledge about ρ_p before explicitly accounting for precession in our analysis.

The majority of parameters required to calculate the *informed* prior are already given in the non-precessing results. The two exceptions are the amplitude of the precessing spin χ_p and the initial precession phase ϕ_{JL}. As discussed in Section b, we can obtain a prior for χ_p *conditioned* upon the other parameters, specifically the mass ratio and aligned spin $\chi_{\rm eff}$, and this can be used to generate the informed prior on ρ_p. The initial precession phase is unconstrained by the non-precessing parameter recovery, this then allows us to assume it to be uniformly distributed. By calculating the *predicted* posterior distribution for ρ_p based upon a set of non-precessing samples, we may examine the effect of other measured parameters on the final ρ_p distribution. For example, if the aligned-spin run favours a binary that is close to equal mass and an orientation consistent with a face-on system, then our prior belief will be that the precessing SNR will be low — it is only with unequal masses and systems misaligned with the line of sight that there are significant precession effects in the observed waveform. A prior belief of ρ_p peaking at low values will cause the predicted ρ_p to peak at values lower than the simulated one and consequently so too will the inferred posterior distribution for ρ_p inferred from a full 15-dimensional parameter estimation analysis.

a Precessing signal

We now apply this conjecture to a precessing signal by attempting to *predict* the posterior distributions for ρ_p. This allows us to investigate how much our recovered posterior distributions may differ from the idealised case of a precession likelihood function distributed about the simulated (true) value. In Fig. 4.16 we show the results of this for the $q = 4$ simulation presented in Sec. b. This specific simulation was chosen since this case has the largest ρ_p and corresponds to a simulation where a non-precessing analysis is less justified. It is therefore a good case to show how the combination of the informed prior and the additional likelihood from precession $\Lambda_{\rm p}$ correctly estimates the large ρ_p. In Fig. 4.17, we show how the predicted posterior distribution compares to the inferred distribution over the full range of mass ratio simulations presented in Sec. b.

In Fig. 4.16 we show this predicted distribution, the informed prior, the χ^2 likelihood function and the posterior distribution obtained from a full parameter estimation analysis. By explicitly calculating the informed prior and likelihood terms separately for ρ_p, we can see the effect of the prior on the ρ_p posterior. The prior strongly disfavours large observable precession and therefore pulls the posterior towards *smaller* values than the simulated value i.e. where the likelihood function peaks.

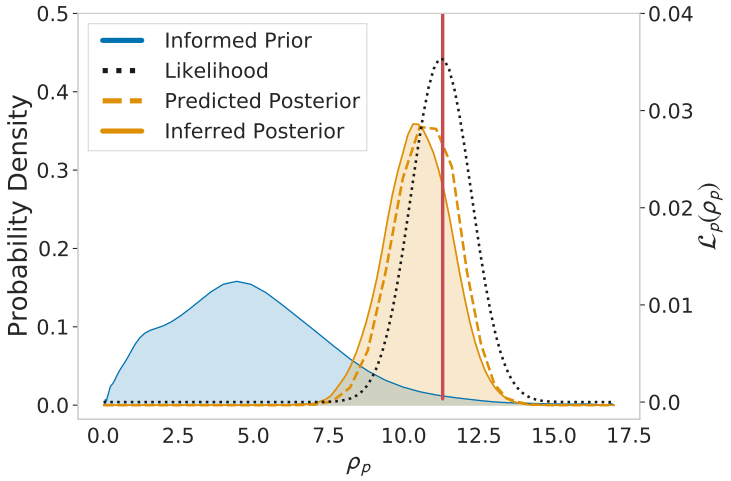

Figure 4.16: The predicted distribution for the precession SNR ρ_p (dashed orange) calculated as the product of the precessing contribution to the likelihood (black dotted line) and the informed prior of ρ_p (blue) for the $q = 4$ simulation presented in Sec. b. For comparison, we show the inferred ρ_p posterior distribution from the full 15 dimensional parameter estimation analysis (solid orange) and ρ_p for the injection (red line). The informed prior is peaked at low values of ρ_p causing the peak of the posterior to be smaller than the maximum likelihood value.

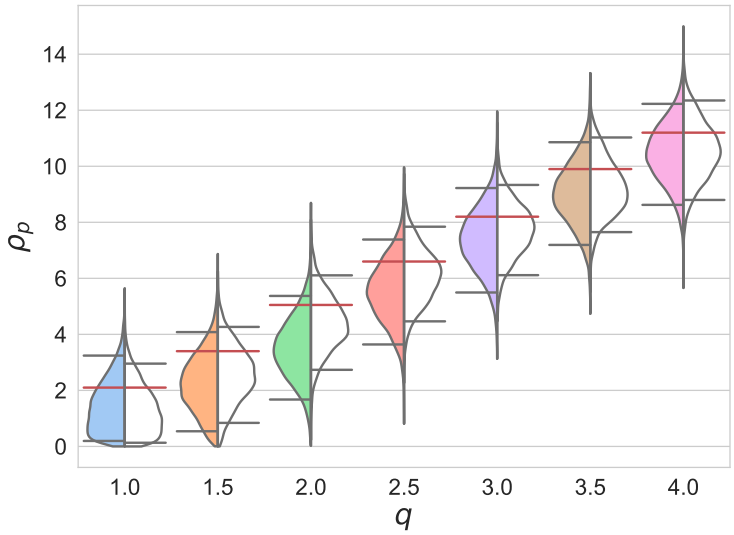

Figure 4.17: Violin plot comparing the observed ρ_p distribution (colored) from a precessing analysis, and the predicted distribution (white) based on the aligned-spin results and simulated value of ρ_p for the set of varying mass ratio simulations presented in Sec. b. The predicted and observed distributions for precession SNR are in good agreement, even though the ρ_p in the simulated signal (red lines) lies above the peak of either distribution.

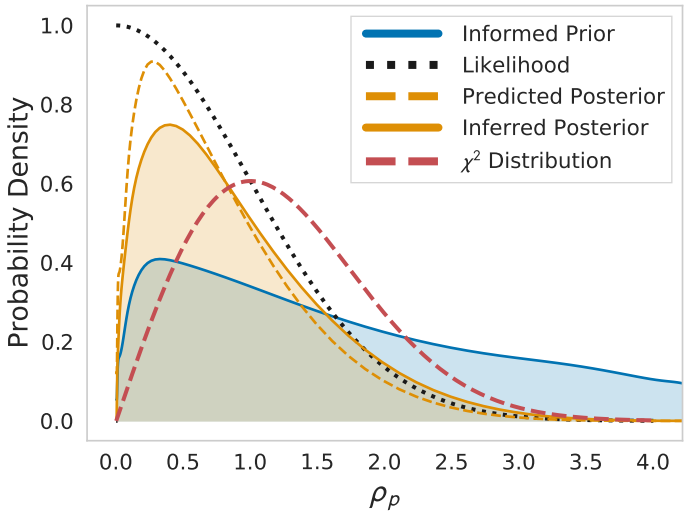

Figure 4.18: Distribution of ρ_p in the absence of precession for the "standard injection". The inferred ρ_p distribution using the IMRPhenomPv2 approximant for recovery is shown by the solid orange line. The dashed orange line shows the predicted distribution using samples collected from an aligned-spin analysis and setting the simulated precession SNR to be 0. We also shows the χ^2 distribution used previously ([3]) as a red dashed line

In Fig. 4.17, we show a comparison between the predicted and measured ρ_p distributions for the set of runs with varying mass ratio presented in Sec. b. When we calculate the posterior, explicitly accounting for the parameter space weighting encoded in the informed prior on ρ_p, we find good agreement between the predicted and the inferred ρ_p distributions and note that neither predicted nor inferred are centred around the *true* value for the set of signals that we have simulated. Of course, if we were to draw signals uniformly from the prior distribution, we would expect to observe the inferred distributions of ρ_p matching with the simulated values.

b Non-precessing signal

We now look at the expected posterior distribution for ρ_p when there is no precession in the the signal. As explained in Sec.4.6, previously a χ^2 distribution with two degrees of freedom was used to model the ρ_p distribution in the absence of any precession (see Ref. [3]). This then led to the natural heuristic that $\rho_p = 2.1$ should be the threshold for observable precession. Using Eq. (4.10) we can now use a more informative prior on $\tilde{\rho}_p$ and obtain a more accurate estimate of the expected posterior distribution in the absence of precession. We do this by using parameter estimation samples from an aligned-spin model and setting the simulated precession SNR to be 0, this then allows us to account for the effects of priors and different noise realisations.

In Fig. 4.18 we show the predicted and observed distributions for the precession SNR for a non-precessing signal. We use a non-precessing equivalent of the "standard" injection as our simulated signal (i.e., we set $\chi_p = 0$ while ensuring all other parameters match those in Tab. 4.1). We inject with zero noise and use the `IMRPhenomPv2` model for parameter recovery.

The inferred ρ_p distribution is peaked at lower values that the χ^2 distribution as shown in Fig. 4.18. However using the prediction from the likelihood (Eq.4.7) and the *informed prior* we are able to obtain a better estimate of the posterior in the absence of precession. This estimate can be obtained without performing parameter estimation incorporating precession, this therefore allows for a better metric for determining whether or not there is measurable precession in the system.

The distribution for the informed prior on precession SNR will depend upon the details of the signal. In particular, it will be strongly peaked near zero for events that are likely to have small opening angle (eqivalently $\bar{b}$), i.e., events that are close to equal mass and have significant spin aligned with the orbital angular momentum, while high mass-ratio events and those with large anti-aligned spins will lead to greater support for large values of ρ_p. Furthermore, for binaries where the orientation can be well measured, *without* precession information, for example where higher modes are important, those that are close to face-on will lead to predictions of smaller ρ_p while those that are edge-on will give larger values. Given that the majority of signals observed to date are consistent with equal mass binaries, in most cases the prior on ρ_p will tend to be peaked at low values. Consequently, the simple threshold of $\rho_p \gtrsim 2.1$ as evidence for precession, remains appropriate and is likely more stringent than suggested by the simple likelihood calculation.

4.7 Discussion

In most candidate astrophysical binary distributions, precession is likely to be first measured in a comparable-mass binary [123]. We have considered a fiducial example of such a possible signal (mass-ratio $q = 2$, SNR $\rho = 20$, and in-plane spin $\chi_p = 0.4$, such that the precession contribution to the total SNR is $\rho_p = 5$), and performed an extensive parameter-estimation study that has systematically explored the impact on parameter measurements of changes in each of the key source parameters: the SNR, the in-plane spin magnitude, binary inclination, the binary mass ratio and aligned-spin contribution, the binary's total mass, the polarisation, and sky location. These examples illustrate well-known features of precession signals [181, 77, 62, 97, 183, 184, 123, 185], and quantify their effect on both the measurement of precession, and their impact on the measurement accuracy and precision of other parameters.

We have also verified that ρ_p provides a suitable and intuitive metric for determining whether or not we have measured precession, and shown that there is an approximate mapping between ρ_p and the use of the Bayes factor to assess the ev-

idence of precession. We suggest that given these results, future large scale studies of precession can be made considerably computationally cheaper by computing ρ_p, rather than a full Bayesian analysis.

We note that as ρ_p captures precession by identifying additional power beyond a simple non-precessing waveform model, it could therefore be effected by phenomena such as eccentricity and higher order multipoles. As BFs simply compare the evidence for two models, one precessing and one non-precessing, using BFs as the sole metric would also be biased by properties like eccentricity and higher order multipoles.

However, a similar approach to the 2-harmonic decomposition for precessing signals has recently been applied to GWs including the effects of higher harmonics [200]. In future work, we will combine these approaches and explore the measurability of precession in systems with significant evidence for higher harmonics, and the impact of the combination of higher modes and precession upon parameter accuracy. It may also be possible to account for eccentricity through a similar decomposition.

As highlighted in section 4.6 these decompositions provide powerful insights into how the addition of physical phenomena introduce information into the analysis. Here we show that the likelihood can be simply factored into precessing and non-precessing contributions. This then allows us quantify the extra information that can be gained from a precessing analysis and even predict the recovered ρ_p distribution with or without these effects taken into consideration in the analysis.

The current study does not include higher harmonics, and uses a signal model (`IMRPhenomPv2`) that neglects two-spin precession effects, mode asymmetries that lead to out-of-plane recoil [201], and detailed modelling of precession effects through merger and ringdown. Although these effects are typically small, so is the imprint of precession on the signal, and it would be interesting in future to investigate the impact of these additional features on our results. We also emphasize that, although we consider it to be extremely useful to provide quantitative examples of the effects of each of the binary parameters, these will necessarily depend on the location in parameter space of our fiducial example. However, having chosen a configuration from amongst what we expect to be the most likely signals, we hope that these examples will act as a useful guide in interpreting precession measurements when they arise in future gravitational-wave observations.

Chapter 5

Density Estimation with Gaussian Process

5.1 Introduction

The first detection of gravitational waves (GW) in 2015 [202] sparked a new era of Astronomy. Several years on from that event the number of detected gravitational waves keeps increasing and within this decade we expect to observe $O(10^3)$ signals [203] from compact binaries coalescences (CBCs). This huge progress brings with it the challenge of efficiently analysing a large number of events. To address these computational challenges, machine learning techniques are being increasingly investigated within the field of gravitational-wave physics [63]. Many studies have focused on speeding parameter estimation of the source parameters of the signals with various techniques, such as deep learning [204], variational autoencoders [205] and autoregressive neural flows [206]. Other work has focused on combining detection and parameter estimation with deep neural networks [207] as well as using neural networks to perform the interpolation step in reduced order modeling to rapidly generate surrogate waveforms [67, 66].

While the research efforts to speed up or completely revolutionise parameter estimation are ongoing, the issue of how to effectively deal with a large number of results from different events remains. In particular, how to streamline the analysis of the results, while maintaining accuracy. In this work, we demonstrate the efficiency and usefulness of using Gaussian processes (GP) for post-processing parameter estimation results of CBCs. Applications of GPs in the field of gravitational waves span a wide range of use-cases, such as marginalising waveform errors [208], regression of analytical waveforms [209], predictions of population synthesis simulations [210], hierarchical population inference [211] and Equation of State (EOS) calculations [212]. They have also been exploited for the development of fast parameter estimation with `RIFT` sampler [213].

Here we exploit GPs to estimate probability density functions (PDFs) from pa-

rameter estimation of gravitational-wave signals. Non-parametric density estimation from a finite set of samples is an active research field in machine learning and statistics [214, 215, 216].

For most GW analysis, histograms are usually the preferred estimators to visualise the marginal posterior PDFs and to avoid over-smoothing sharp features, but often are not convenient for post-processing analyses such as population inference. These sorts of analyses either re-weight the posterior samples directly [217] or need to estimate a continuous representation of the gravitational-wave posterior density surface. Several density estimation methods such as Dirichlet processes [218], Gaussian Mixture Models [219] and others have been employed to address this problem specifically for gravitational-waves. As well as these, A closely related method to GPs [220], Gaussian Kernel Density Estimators (KDEs) are sometimes employed in gravitational waves' analyses [221, 222, 223].

These KDEs are often effective but they assume correlations between parameters to be linear and smooth, making this method sometimes limited in flexibility. There exist many variations of the KDE algorithm to take into account specific interpolations problems, but there isn't a single implementation that is guaranteed to be robust against all possibilities. A specific KDE implementation might solve an issue in one case and be the cause of some inaccuracies in another [224].

We implement a single technique that can interpolate arbitrary multi-dimensional slices in parameter space, which can handle both simple and difficult space morphology, such as sharp bounds and non-Gaussian correlations. Our modelling tool is based on the histogram density estimate, combining the histogram's accurate treatment of the samples' features with the predictive capabilities of GPs. An additional advantage of this technique is that it can provide a Bayesian measure of uncertainty from the finite (and sometimes small) number of samples for post-processing analysis. This measure of model uncertainty could then be incorporated into any analysis where the marginalised posterior density is used.

In Sec. 5.2 we describe our density estimation technique in the context of gravitational-wave parameter estimation and machine learning. We propose a series of example applications in Sec. 5.3, which allows us to discuss the advantageous features of our method. Finally, in Sec. 5.4 we summarise our findings and suggest future extensions of this work.

5.2 Methods

In this section, we introduce the mathematical framework of the techniques discussed. In subsection a we discuss the Bayesian inference problem for gravitational waves and the density estimation techniques currently employed in the field. In subsection b, we outline the fundamentals of GPs and their interpretation for interpolating a posterior density surface. We then describe the details of our GP

implementation and how to model probability densities from parameter estimation.

a Bayesian inference and density estimators

The posterior probability we consider is generally a 15-dimensional surface for a circular binary black hole (BBH) merger but can be 17-dimensional in the case of a binary neutron star (BNS) merger due to the inclusion of parameters that describe the physical structure of the neutron stars. The dimensionality depends on the physical parameters describing the signals. Generally, these are distinguished between extrinsic parameters, such as sky localisation, and intrinsic parameters, such as the masses of the sources.

The posterior is however generally intractable and therefore must be evaluated via stochastic methods such as Markov Chain Monte Carlo (MCMC) and nested sampling, these are implemented (and specifically tuned for the gravitational wave problem) in Bayesian inference packages such as LALInference [225] and `bilby` [53].

b Density estimation with Gaussian Processes

Definition and interpretation

GPs are interpolation methods with a probabilistic interpretation, they are built on a Bayesian philosophy, which allows you to update your beliefs based on new observations. The process can be understood as an infinite-dimensional generalization of multivariate normal distributions, such that any finite collection of points within the domain of the process are related by a multivariate Gaussian distribution. As data is observed, the GP is *conditioned* and the range of possible functions that can explain the observations is constrained. As such a GP is defined by a mean, which represents the expectation value for the best fitting function, and by a covariance matrix, called a kernel, which measures the correlations between observations [226]. In the absence of observations, the GP predictions will revert to a prior mean function, which is usually chosen to be zero, and which properties are determined by the kernel architecture. Mathematically this is written as:

$$f(\vec{x}) \sim \mathcal{GP}(m(\vec{x}), \kappa(\vec{x}, \vec{x}')) \tag{5.1}$$

where the mean and covariance are denoted as:

$$\begin{aligned} m(\vec{x}) &= \mathbb{E}\left[f(\vec{x})\right] \\ \kappa(\vec{x}, \vec{x}') &= \mathbb{E}\left[(f(\vec{x}) - m(\vec{x}))(f(\vec{x}') - m(\vec{x}'))\right] \end{aligned} \tag{5.2}$$

We can then model any point on the surface as a Normal distribution where the mean and standard deviation are defined by our process conditional on previous

observations:

$$y_*|f, x \sim \mathcal{N}(\mu(x_*), \sigma_*^2) \tag{5.3}$$

where we have that the predictive mean, $\mu(x*)$, and variance σ_*^2 are defined as:

$$\mu(x_*) = \kappa(\vec{x_*}, \vec{x}^{)}\kappa(\vec{x}, \vec{x}^{)}f \tag{5.4}$$

$$\sigma_*^2 = \kappa(\vec{x_*}, \vec{x}_*) - \kappa(\vec{x_*}, \vec{x})\kappa(\vec{x}, \vec{x})^{-1}\kappa(\vec{x_*}, \vec{x}^{)} \tag{5.5}$$

These equations have a simple interpretation, the predictive mean is given as a weighted combination of the previous function values. Where the weightings are given by the similarity, as defined by the kernel response, between the new point x_* and the previously seen values x. The uncertainty again has a similar interpretation, it has a maximum at $\kappa(\vec{x_*}, \vec{x}_*)$, which is the prior uncertainty. This uncertainty is then reduced when there is a kernel response e.g. the more similar the training points are to the point you are predicting then the more information you have about the point and therefore the uncertainty is reduced.

The non-parametric nature of GPs makes this technique flexible, but it can be computationally expensive as the whole training set needs to be taken into account at each prediction. The standard implementation has $\mathcal{O}(N^3)$ computations and $\mathcal{O}(N^2)$ storage, this then becomes prohibitive for $\sim$ 10k data observations or more. To tackle this issue it is common to use sparse inference methods, which approximate the conditioning of the GP over a set of $M \ll N$ 'inducing' points. The evaluation over the inducing points M is then much cheaper than for an 'exact' GP resulting in $\mathcal{O}(NM^2)$ computations rather than $\mathcal{O}(N^3)$ [227, 228]. As well as sparse methods one can exploit multi-GPU parallelization and methods like linear conjugate gradients to distribute the kernel matrix evaluations which then allows for exact inference to be performed on a short time scale [229]. In this work, however, we find that sparse approximations were accurate enough to effectively model the marginalised posterior surfaces that we were interested in. Moreover, once a GP has been 'trained' over the data, it is possible to draw infinitely many function realisations from it without recomputing the expensive covariance matrix.

A recognised advantage of GPs is reliable uncertainty estimate when making predictions over unseen data. In this application, we are not interested in predicting the value of the posterior in unexplored regions of the parameter space, but only in generating a faithful model where we have posterior samples. In regions within the space of parameters, the GP variance depends on our choice of training points, which is useful to assess the accuracy of our density estimation. In terms of uncertainty estimation this can be explained as our model having very low *epistemic* uncertainty everywhere, we then seek to estimate the *aleatoric* uncertainty due to our model fit

around the random fluctuations in the histogram densities which are used to train the GP.

Model construction

In this application, we want to use a GP to estimate the marginalised posterior density for any subset of parameters. We train our GP using the normalised histogram counts over a grid of points, i.e. the centroids of the histogram bins, that cover the marginalised parameter space. This gives us a non-smooth, noisy (due to Poisson errors) estimate of the density, we then use the GP to fit this this discrete set of points to generate a smooth,continuous representation of the density surface.

An important choice when modeling a system using GPs is the choice of kernel, this encodes your assumptions about the relationship or covariance between data points. For all examples presented in this work we used a combination of the RBF and Matern($\frac{1}{2}$ or $\frac{5}{2}$) kernels. Further technical details regarding this choice and our data pre-processing scheme (which also had a significant impact on our model accuracy) are included in Appendix 5.5 b.

We employ `TensorFlow` and `GPFlow` to implement our GP training infrastructure, which includes two inference schemes: exact inference for 1-2 dimensional problems ($\mathcal{O} \sim 1000$ samples) and sparse inference for higher dimensionality due to computational costs. As well as a difference in the inference scheme, when extending this method to higher dimensions, our choice of training data changes. When creating the grid over four dimensions, due to the sparsity of the parameter space, we find that the typical set has a volume of $O(1\%)$ relative to the total prior volume (this is a common problem associated with the curse of dimensionality [49]). We, therefore, choose to discard the empty bins and encode our knowledge of these points through the choice of prior over our GP.

Since the model is constructed with converged posterior samples, there is no probability support where the histogram bins are empty. To encode this, we set the mean of the GP to be equal to zero, such that far away from the training data the model will have a high variance but a mean of zero.

To estimate the density for a given region of parameter space we then simply evaluate the GP at those parameters, i.e.

$$\begin{aligned} p(\vec{\theta} = \vec{x_*}|d) &\approx y_*|f, x \\ &\sim \mathcal{N}(f(\vec{x_*}), \sigma_*^2) \end{aligned} \tag{5.6}$$

The choice to set the GP prior to zero means that we would be allowing for negative probaility densities, to avoid this we apply the ReLU function [230] as a layer on top of the density evaluation. This sets all negative values to zero meaning that some points in parameter space will be distributed as a truncated-Gaussian.

Due to bounded priors (e.g at mass ratio $m_2/m_1 \coloneqq q = 1$), the posterior sur-

face often presents sharp discontinuities and therefore the surface is only *piece-wise continuous*. GPs are in principle flexible enough to model any surface including piece-wise continuous ones, however, we found in practice that it is more favourable to decompose our density function into two components, one smooth, continuous function, and one step function. We do this by multiplying the density and our GP estimate by a step function, which is zero at any discontinuities and 1 otherwise.

$$\pi(\vec{x_*}) = \begin{cases} 1 & \text{if } x_{min} < \vec{x_*} < x_{max} \\ 0 & \text{otherwise} \end{cases}$$

Multiplying by this step function is then analogous to imposing a prior over our posterior surface, i.e. it allows us to rewrite the equation 5.6 as

$$\begin{aligned} p(\vec{\theta} = \vec{x_*}|d)\pi(\vec{x_*}) &\approx (y_*|f, x)\pi(\vec{x_*}) \\ p(\vec{\theta} = \vec{x_*}|d) \qquad &\sim \mathcal{N}(f(\vec{x_*}), \sigma_*^2)\pi(\vec{x_*}) \end{aligned} \tag{5.7}$$

We are free to encode our knowledge in this way and perform the decomposition as we do not change the original posterior surface that we would like to model in any way. This enhances the robustness of the model against all discontinuities, including artificial cuts in parameter space that might be required for post-processing analysis.

The variance of the GP depends on the kernel, but also on the noise variance parameter of the likelihood. Usually, the noise variance is given by a single number, i.e. homoskedastic noise, which reflects the random fluctuations of the posterior samples. In low-dimensional examples, where we employ an exact inference scheme, we can assign multiple values to the noise variance, i.e. heteroskedastic noise [231]. In such instances, we are then able to propagate the error from the histogram on the density estimate, which is simply given by the Poisson noise in each bin $\sigma_{\text{bin}} \sim \sqrt{N_{\text{counts}}}$. Incorporating heteroskedastic errors within a sparse inference scheme is an area of current research in the field of machine learning [232].

It is common practice to build an interpolation of a posterior surface in order to draw more samples from it. As our model is implemented in `TensorFlow` we can quickly draw more samples from the marginalised posteriors using the many samplers available in the package library, such as Hamiltonian Monte Carlo (HMC) [49].

5.3 Results

In this section, we present our model and a series of example applications for gravitational waves. In Sec. a we illustrate the method on a simple 1-dimensional analytical example. In Sec. b we show examples of common post-processing applications for our density estimation tool. Finally, we discuss our treatment of GP model uncertainty and how we propagate it to produce uncertainty on the marginalised posterior

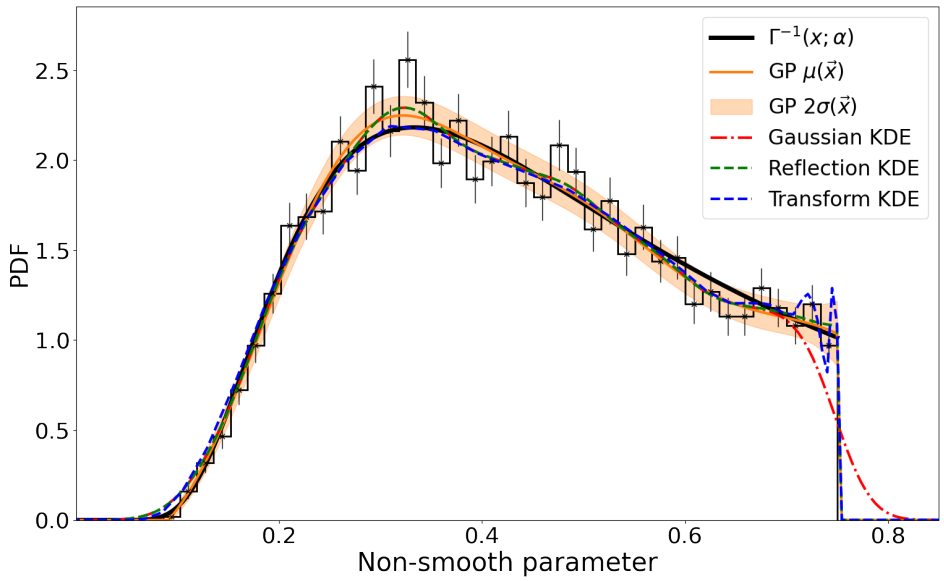

Figure 5.1: Interpolation of a bounded one-dimensional inverse gamma density function (in solid black) with our GP-based method (in solid orange). The histogram points used to generate the model and its uncertainty are shown as black points with error bars. Alternative KDE methods are shown for comparison as coloured dashed lines.

distributions.

a Analytical 1D example

Our proposed GP modelling technique is by construction flexible and robust against all distribution morphologies. To illustrate this, we construct a non-trivial 1-dimensional example: an inverse gamma function $f(x,\alpha) = \frac{x^{-\alpha-1}}{\Gamma(a)}\exp(-\frac{1}{x})$, with $\alpha = 2$ and a sharp bound at $x = 0.75$.

In Figure 5.1 we show our GP model mean prediction and uncertainty, compared to a Gaussian KDE from `scipy.stats` [233] and two KDE transformations implemented in `PESummary` [234], a commonly used post-processing package in gravitational-wave astronomy. The *reflection* and *transform* KDEs, are examples of augmentations on the standard (Gaussian) KDE, and are generally used to model difficult features introduced at the boundaries of posterior distributions. Both of these improvements to the standard KDE apply a transformation at the boundary which implicitly assumes some distributional features (see [234] for more details). A Gaussian Process on the other hand makes no assumptions about the distributional shape and can in principle fit any distribution.

We show an example in Fig 5.1 where our GP is able to well model the posterior and the *reflection* KDE provides a better fit than the other KDE methods. The GP

is slightly too high, relative to the underlying function, this is seen with both the KDEs and is due to random poisson errors in the samples from the function that we use to fit the models. This highlights a useful feature of the GP however as the uncertainty region captures the true model despite the mean being too high.

The *transform* KDE is more sensitive to noisy features in the samples and can present artifacts, while the Gaussian KDE over-smooths the sharp cut at 0.75. Following this illustrative example, there are others where the *reflection* KDE is less appropriate. This example was chosen to highlight a case where the choice of KDE is important to fit the distribution well. While synthetic and not representative, it does illustrate features that can and do happen in gravitational-wave astronomy when analysing posteriors. In examples such as this our GP model provides an alternative method to KDEs, requires less hand-tuning, and also provides a Bayesian estimate of the error on the density estimate, as propagated from the histogram errors.

b GW Applications

We now look at a few important post-processing problems in gravitational-wave astrophysics. The training time required to generate the models presented in this section is of the order $\mathcal{O}(2mins)$, with variations due to the dimensionality of the surface and to the inference scheme employed. To assess the quality of the model in more than one dimension we decide to re-sample the surrogate surface and compare the new samples to the original set, part of which has been used for training. All samples used in the following sections are taken from the Bilby GWTC-1 catalog [235].

Catalogue of gravitational-wave properties

Gravitational-wave detection parameters can be distinguished between those intrinsic to the sources, such as the component masses, and those extrinsic to them, such as the sky location. Interpolating the marginal posteriors of combinations of these parameters is often necessary for post-processing. The following example illustrates a simple case where one can use a GP to interpolate the intrinsic parameters for a given detection. In practice, this could then be repeated for entire GW catalogues so that these interpolated posterior surfaces are then combined for population inferences on the sources of GWs.

For this example, we interpolate the marginal posterior distribution of the intrinsic parameters of the first BBH detection GW150914 [236], parametrised as follows: chirp mass $\mathcal{M}$, mass ratio $q = m_2/m_1$ (where $m_1 > m_2$), effective inspiral spin component $\chi_{\rm eff}$ and effective precession spin $\chi_{\rm p}$, defined by the spin components that lie in the orbital plane [237]. In Figure 5.2 we compare the marginal distributions sampled from our GP model to the original PE samples. We can visually assess that

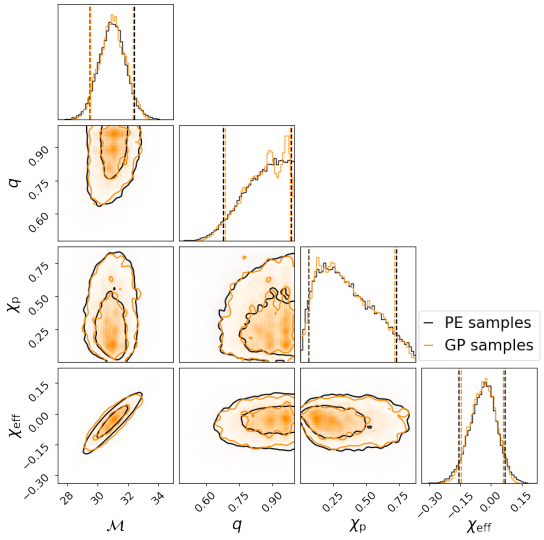

Figure 5.2: Corner plot of the intrinsic parameters of GW150914, drawn from our GP surrogate (in orange) compared to the original PE samples (in black).

	GP samples	PE samples
Chirp mass $\mathcal{M}/M_\odot$	$30.95^{+0.93}_{-0.97}$	$30.96^{+0.86}_{-0.89}$
Mass ratio q	$0.87^{+0.09}_{-0.12}$	$0.87^{+0.09}_{-0.12}$
Effective precession spin component χ_p	$0.33^{+0.26}_{-0.19}$	$0.32^{+0.27}_{-0.19}$
Effective inspiral spin component χ_eff	$-0.04^{+0.07}_{-0.07}$	$-0.04^{+0.06}_{-0.07}$

Table 5.1: Source properties of the intrinsic parameters of GW150914, original samples and samples from the GP interpolation.

the correlations between parameters are accurately reconstructed as the 50% and 90% contour lines overlap for each pair of parameters. In Table 5.1 we report the mean and 90% confidence intervals of the samples drawn from our model and which we find in agreement to the values from the original samples within the expected uncertainty.

Accurate interpolation for conditional integrals

Many astrophysical inquiries in gravitational-wave astronomy require evaluating conditional integrals across parameter space, which in turns require sampling additional posterior points constrained to a hyperplane. This is for instance the case when estimating the Equation of State (EOS) from BNS collisions, an important post-processing analysis that allows us to probe extreme conditions of matter [165]. This is possible because the compactness of the objects is imprinted in the gravitational waveform and can be measured by the tidal deformability parameters. The EOS integral involves evaluating the marginal posterior distribution over the masses $(\mathcal{M}, \eta)$

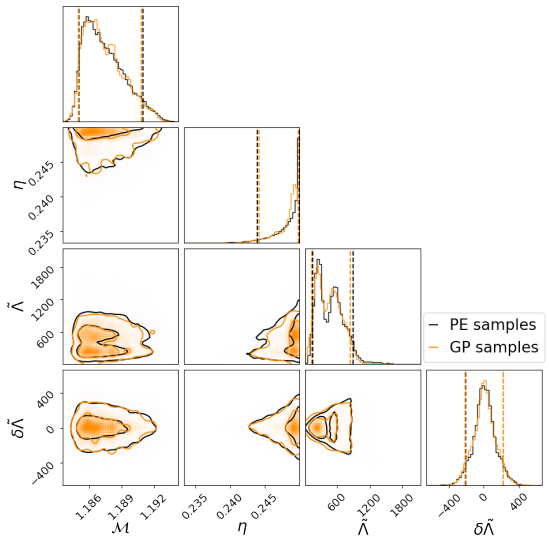

Figure 5.3: Corner plot of the mass and tidal parameters of GW170817, drawn from our GP model (in orange), compared to original PE samples in black.

and tidal parameters $(\tilde{\Lambda}, \delta\tilde{\Lambda})$, subject to constraints between those parameters as parametrised by the EOS.

There are instances where the marginal posterior for these parameters contain non-linear correlations, as is the case for the first BNS event GW170817 [238]. We test our interpolation model over this 4-dimensional surface. In Figure 5.3 we compare the marginal distributions sampled from our GP model to the original PE samples. We see that our GP is able to faithfully represent the marginalised posterior surface, in particular, we see that there is good agreement between the 90% credible intervals. When looking at the 2d contours see that the 50% and 90% levels agree very well and that the GP model is able to capture degenerate features and bi-modalities. Finally, our interpolation of the surface can be re-sampled efficiently and for this example, we obtained 750k samples in a few $\mathcal{O}(5mins)$ (depending on hardware) using an HMC sampler. Hence this method can be advantageous over traditional methods, where the interpolation is generally performed with a Gaussian KDE by transforming the symmetric mass ratio parameter to be $log(0.25 - \eta)$ [222] and there is no measure of uncertainty over the fit.

Propagating GP uncertainty

GPs provide a fully Bayesian estimation of the uncertainty over model predictions, as the full covariance matrix between posterior samples is computed. In each of the GW applications shown so far we have utilised the mean prediction of the GP function. This uncertainty measurement can be very important in many cases, however here we illustrate with a single example how one can extract the uncertainty from the modelling. Accurate localisation of a gravitational signal can be of fundamental

importance for multi-messenger astronomy [239, 240] and for measurements of cosmological parameters with dark sirens [241]. As the localisation accuracy decreases, the marginal posteriors for the sky location parameters can look degenerate and non-Gaussian. We build an interpolation of the sky location parameters, right ascension (ra) and declination (dec), of GW150914. This event was observed by only two detectors, so albeit its high SNR, its sky location presents a typical ring-like shape.

The uncertainty measure produced by the GP is a Gaussian distribution about any given point on the surface, when considering the entire surface the combination of these Gaussians can be interpreted as a range of plausible density surfaces for any given confidence level (e.g. 2σ). The uncertainty on the 1D marginal distributions can then be obtained from an upper and lower bound for each point in the surface (given by the GP error σ, equation 5.3) and then marginalising these across one of the dimensions to obtain an uncertainty estimate about the mean 1D predicted posterior density. For brevity let $ra = \alpha$, $dec = \delta$.

$$\begin{aligned} p(\alpha|d) &= \textstyle\int_\delta p(\alpha, \delta|d)\, \delta \\ p(\alpha|d) \pm \sigma(\alpha) &= \textstyle\int_\delta (p(\alpha) \pm \sigma(\alpha, \delta))\, \mathrm{d}\delta \end{aligned} \tag{5.8}$$

In 2D and especially when considering sky localisation, we are also interested in the contours that enclose a given volume of probability density to plan optimal observation strategies in the search for electromagnetic counterparts. We propagate the uncertainty estimate produced by the GP (in the space of all realisations from the GP) to the physical parameter space on credible interval contour levels. We define a function, f_q, which truncates the posterior density function as follows:

$$f_q(\alpha) = \begin{cases} p((\alpha, \delta|d) & \text{if } p((\alpha, \delta|d) \geq q \\ 0 & \text{otherwise} \end{cases}$$

Such that the integral of f_q contains a given proportion of the total probability mass determined by the desired confidence level i.e.

$$\int_{\alpha,\delta} f_q(\alpha, \delta|d)\, \mathrm{d}\delta\, \mathrm{d}\alpha = cl \tag{5.9}$$

For a given a confidence level cl (usually the 50% and 90% levels), solving equation 5.9 for q gives q_{cl}, the value of the posterior density of the relevant contour. We obtain the contour, and the error on the contour, by plotting the (ra,dec) values for which:

$$\begin{aligned} p(\alpha, \delta|d) &= q_{cl} \\ p(\alpha, \delta|d) \pm \sigma(\alpha, \delta) &= q_{cl} \end{aligned} \tag{5.10}$$

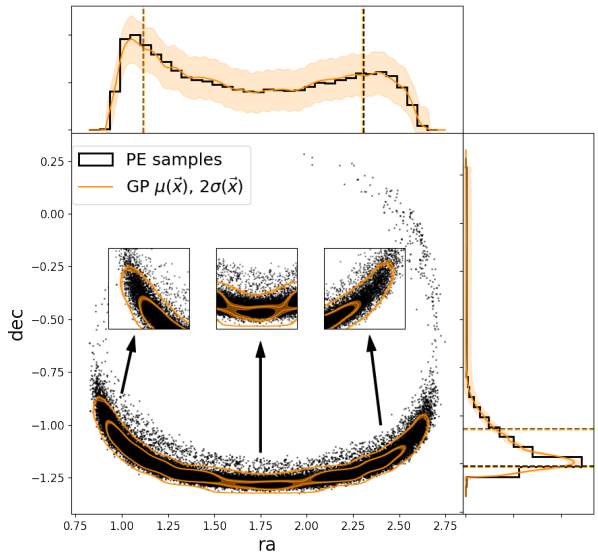

Figure 5.4: Central panel: contours of the 2D sky-location of GW150914, the GP model mean prediction and uncertainty (in orange) is compared to the points used to construct the fit (black crosses). Top and left panels show the GP model projections in 1D, compared to the original PE samples. All plots show the 2σ uncertainty around the density estimate as a shaded band

Geometrically, we are building an uncertainty envelope around the mean GP prediction. The uncertainty on the contour is then given by the location in physical parameter space where the edges of that envelope intersect the plane defined by the mean prediction contour.

In the central panel of Fig 5.4 we show the samples used to construct the model as well as the 50% and 90%, contours of the GP interpolation in 2D with their respective 2σ uncertainty (the shaded regions). The top and left panels of Fig. 5.4 show the mean prediction and its 2σ uncertainty marginalised over each parameter by a simple integration of the density over its projection.

The inclusion of the uncertainty highlights several features. On the central inset in Fig 5.4 we see that the lower bound on the 50% contour is composed of three islands which correspond to peaks, while for both the mean and the upper bound these islands are connected to obtain a smooth surface at this contour level. For the outer 90% contour we see that the differences mainly manifest in the tails, where as expected the upper bound follows the well known *ring* around the sky slightly further. This matches our intuition that there is possibly more density around the ring than around the edges of the contour in the middle of the plot.

The *ra* and *dec* parameter space highlights several interesting features which are in principle difficult to model, such as their highly curved correlation. For this

particular example the simple kernels used throughout the study were appropriate. However it is important to note that in general this may not be the case, and possible enhancements include simply encoding the 2π wrapping formaly using a periodic kernel, or in order to account for the non-trivial correlations between parameters a non-stationary kernels such as deep kernels [242] which would effectively allow for position dependent length scales."

5.4 Conclusions

We have presented an alternative method for density estimation of marginal PDFs for gravitational-wave parameters. Our method combines the desirable features of histograms to the extrapolation capabilities of KDEs, within a Bayesian framework. The choice of histogram binning determines the resolution of the PDF, while the kernel of the GP allows the interpolation to be flexible over non-Gaussian correlations and yet smooth. The noise variance parameter of the GP ensures that sources of stochastic noise from the histogram density estimation are taken into account. In cases where we employ an exact inference scheme, this noise variance can be evaluated for each histogram bin and it is equivalent to heteroskedastic errors over the density estimation. This allows to fully propagate the uncertainty from the PE samples. We plan to extend this method and fully incorporate uncertainties, as we showed in this work for the sky localisation example, over higher-dimensional posterior surfaces in future work.

This method may be preferable to other methods such as KDEs, a closely related method which is sometimes adopted in the field, depending upon the use-case requirements. It comes with three main advantages: it is suitable for most interpolation problems commonly encountered for gravitational-wave marginal posteriors; it provides a Bayesian measure of uncertainty over the model predictions; it allows to quickly re-sample the interpolation using HMC and other samplers available in `TensorFlow` . We presented a series of examples where we know the accuracy of the interpolation is important, such as EOS calculations and sky localisation. As the number of events will increase in the next observing run ($O4$), we need reliable tools to post-process the large volume of results.

This work has highlighted the power of GPs to fit a gravitational-wave posterior surface, a natural extension of this work is to generate a surrogate for the entire likelihood surface, similar to what was done by the authors of [243] using a random forest regressor. Such use of GPs has been already investigated in the field of cosmology to model the Planck18 posterior distribution [244]. This work has laid the foundation for us to apply a similar methodology to the gravitational-wave problem in a future work which is currently in preparation [245]. This has applications such as Bayesian quadrature [246], efficient jump proposals [247, 248] and more general use of the GP variance to guide the sampling process. The surface learned by the GP can

be evaluated directly for a given set of parameters, therefore, avoiding the need to compute expensive waveforms. An example where such likelihood surrogates could be exploited is fast re-sampling with new astrophysical priors. This could replace an often difficult re-weighting procedure, especially when a prior assumption limits the number of available samples in a region of interest [249].

5.5 Appendix: Technical details of the GP model

a Data pre-processing

Data pre-processing, often referred to as data-set standardisation, is a common practice within the realm of machine learning and it can have a very high impact on the accuracy of the model. Our posterior samples have a wide range of values, some having bounds $[-1, 1]$ and some reaching $\mathcal{O}(10^3)$. We re-scale our posterior samples such that each parameter ranges between $[0, 1]$ by using the following transformation:

$$\vec{\tilde{\theta}}_d = \frac{(\vec{\theta}_d - min(\vec{\theta}_d))}{(max(\vec{\theta}_d) - min(\vec{\theta}_d))} \tag{5.11}$$

where $\vec{\tilde{\theta}}_d$ is the vector of transformed samples and the *min* and *max* are evaluated for each parameter (i.e. each dimension of the posterior samples vector). The approximate marginalised posterior is scaled according to the *z-score*, such that it has zero mean and unit variance:

$$\widetilde{p}(\theta_i|d) = \frac{p(\theta_i|d) - \mu_{p(\theta_i|d)}}{\sigma_{p(\theta_i|d)}} \tag{5.12}$$

where $\widetilde{p}(\theta_i|d)$ is the transformed marginalised posterior, $\mu_{p(\theta_i|d)}$ and $\sigma_{p(\theta_i|d)}$ are respectively the mean and standard deviation of the marginalised posterior points. All pre-processing in this work is performed using *Scikit-Learn* [250].

b Kernel design

The kernel is defined as the prior covariance between any two function values. Our prior knowledge about the morphology of the posterior can be encoded via this covariance, as it determines the space of functions that the GP sample paths live in. The radial basis function (RBF) or squared exponential kernel is the most basic kernel and it's given as:

$$\kappa_{\rm RBF}(x, x') = \sigma^2 \exp(-\frac{1(x - x')^2}{2\ell^2}) \tag{5.13}$$

where the Euclidian distance between (x, x') is scaled by the length-scale parameter ℓ (measure of deviations between points) and the overall variance is denoted by σ^2 (average distance of the function away from its mean). Functions drawn from a GP with this kernel are infinitely differentiable.

For our application, a more complex kernel architecture that can capture the correlations between parameters is needed. We need smoothness over small scale features, such that we don't model random noise fluctuations of samples, and flexibility over the large scale characteristics of the posterior. For this purpose we

employ a combination of RBF and Matern, which is a generalisation of the RBF kernel with an additional smoothness parameter ν. The smaller ν, the less smooth the approximated function is:

$$\kappa_{\mathrm{M}\nu}(x, x') = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu} \frac{(x - x')}{\ell} \right)^{\nu} K_\nu \left(\sqrt{2\nu} \frac{(x - x')}{\ell} \right) \tag{5.14}$$

We choose $\nu = (\frac{1}{2}, \frac{5}{2})$ depending on the specific morphology of the posterior, as this kernel is responsible for encoding its overall shape such as sharp boundary features. The resulting kernel equation is given by:

$$\kappa_{GP}(\vec{\theta_d}, \vec{\theta_d}') = \kappa_{\mathrm{RBF}} \times \kappa_{\mathrm{M52}}$$

The kernel multiplication is equivalent to an AND operation, as it corresponds to an element-wise multiplication of their corresponding covariance matrices. This means that the resulting covariance matrix will only have a high value if both covariances have a high value. We also apply automatic relevance determination (ARD), which modifies the kernel such that for each dimension an appropriate length scale is chosen [251].

Chapter 6

Exploiting GPU and Autodiff capabilities to rapidly sample gravitational-wave posteriors distributions

6.1 Introduction

Gravitational-wave astronomy is progressing rapidly. There have currently been three successful observing runs, resulting in around fifty confirmed observations (depending upon the choice of threshold for a detection) of Compact Binary Coalescence (CBCs)[2, 57, 252, 158, 159]. With a planned fourth observing run using upgraded detectors on the horizon, as well as further future improvements, we expect the number of detections to grow rapidly over the next decade, potentially reaching around event a day [203]. This increase in detections will provide new insights and allow us to understand the universe as never before, however the computational challenges associated with analysing the growing number of detections will become a major data analysis problem.

One of the major bottlenecks in the data analysis pipeline is estimating the source parameters of detections which pass the detection threshold, to do this one must run computationally expensive sampling routines which can often take O(days-weeks) to converge depending on the source parameters [52, 253]

These parameter estimation routines involve running stochastic samplers which numerically approximate the posterior by drawing many samples from it . This process is computationally expensive for two reasons, the first is that calculating a single likelihood evaluation requires generating a template waveform which is relatively expensive. A waveform which does include precession can take between 1-2000 ms to generate a single template [67], adding in more physics such as precesion or higher

modes can increase this considerably. The second reason is that gravitational-wave posterior surfaces are generally complex and are therefore difficult to sample from. Where difficult to sample from means that to draw a single independent sample from the posterior may require taking many likelihood evaluations [253]. This inefficiency in the sampling effectively means that we must do a relatively slow operation, waveform generation, a huge number of times resulting in large computational costs .

In this work we show two approaches to addressing this problem, either of which can in principle reduce the wall time required to complete a parameter estimation analysis. Both of these methods make use of recent developments in hardware and in waveform surrogate modelling[67], The first approach is effectively a brute force method in which we run random-walk metropolis MCMC to draw samples. This method is inefficient in terms of samples per likelihood evaluations however we are able to rapidly evaluate the likelihood in large batches, meaning that it is efficient in terms of wall time. The second approach we show in this work is to use the automatic differentiation (*Autodiff*) [70, 69] capabilities of *Tensorflow* to compute the gradient of the likelihood which then allows us to perform gradient based sampling such as Hamiltonian Monte Carlo (HMC) [68, 49].

HMC has been proposed in the gravitational-wave community as a way to reduce the computational cost of parameter estimation, however previously it required performing some approximation or fit to the gradient which adds additional costs [254, 255]. The method presented here requires no additional steps, we use a neural network surrogate model[1],implemented in *Tensorflow* and are therefore able to trivially compute the gradients using *Autodiff*

Recently a lot of work has been carried out which looks at replacing sampling using a variety of machine learning methods such as, [206, 256]. These methods generally involve learning the inverse likelihood function in such a way that allows one to produce samples very efficiently. Currently the adoption of these methods in production analysis has been slow, mostly due to the lack of asymptotic guarantees that are implicitly present when using stochastic methods. These asymptotic guarantees, along with years of experience which have resulted in useful sampling diagnostics ensure that the results produced by them are reliable and well understood. In this work we argue that sampling should *not be abandoned* as the primary method for performing parameter estimation and show that using modern hardware and techniques, we may be able to produce a parameter estimation analysis using stochastic sampling in a practical timescale.

We show an example of this where using these methods we are able to produce a converged parameter estimation analysis in tens of minutes. We find that for the

[1]We are not necessarily restricted to using surrogate models, if the waveform model can be evaluated in Tensorflow (or any other language such as PyTorch that allows for automatic differentiation) then it is trivial to use gradient based sampling.

examples presented here the brute force method is faster than gradient based sampling, however it is likely that as the dimensionality and difficulty of the parameter space increases gradient based methods will be more important and therefore this may be an important tool going forward.

6.2 MCMC methods

The most common MCMC algorithm is known as the Metropolis-Hastings algorithm. This algorithm works by proposing new positions in parameter space according to some proposal distribution, Q. One then simulates the system at this new state and calculates the ratio of probabilities between the current state and the proposed state. The point is accepted according to the acceptance probability,

$$\alpha(\theta, \theta_*) = min\left(1, \frac{Q(\theta^*|\theta)p(\theta_*|d)}{Q(\theta|\theta_*)p(\theta|d)}\right) \tag{6.1}$$

Where Q is our transition kernel that determines how we move from one position in parameter space to another and p is the probability density there.

Heuristically this algorithm ensures that in general we are mostly moving toward the regions of larger probability but allowing some probability of exploration. If this chain is run for infinite time then the chain will explore all regions of non-zero probability within our parameter space and will draw samples exactly according to the relative probabilities.

As mentioned above, MCMC methods have asymptotic guarantees. These are by definition only then true with infinite samples, often a practical problem when performing parameter estimation with MCMC methods is how many samples are *enough* to provide reasonable approximations. This problem is particularly difficult due to the geometry of high-dimensional surfaces. I will briefly outline problems associated with sampling from the typical set but see [49] for a more in-depth explanation.

The typical set can be thought of as the region in parameter space that has a non-negligible contribution to any expectations. In high dimensional surfaces, the neighborhood around the mode of the distribution contains very large densities relative to areas outside of it. This means that the areas outside of this neighborhood will not have a significant contribution to any expectations, however the neighborhood around the mode has a very low volume relative to the area outside of it. As we increase the dimensionality of the problem, the neighborhood around the mode tends to a singular point therefore the mode itself has an increasingly negligible contribution to any expectations. This means that there is generally a very small region of parameter space where these two quantities (the volume and the density) are balanced such that there will be any significant contribution to expectations. This region is known as the typical set and is the region of parameter space where

one should focus the computation resources. This balance means that efficient computational Bayesian Inference generally involves designing samplers that can find and explore the typical set efficiently.

a *Vectorised* MCMC

Often when performing MCMC in practise one uses several chains in parallel, this in principle makes it easier to diagnose behaviour such as finding local minima as well as allowing for diagnostics such as the split $\hat{R}$ which checks for global convergence between chains. As well as these diagnostics, using multiple chains in parallel allows one to produce samples quickly by distributing the computation. This increase in efficiency is not guaranteed as by using multiple chains the burn in costs increase relative to a single long chain, however using multiple chains is usually the standard practise. Traditionally this could only be parallelised by computing multiple chains separately on multiple CPUs, this allows for almost perfect parallelisation however this can be computationally expensive if you want to run lots of chains. As well as the cost of using many CPUs, paralelising in this way fails to exploit "single instruction, multiple data" (SIMD) parallelisation which effectively allows for efficiency savings when doing calculations in batches [257, 258]. The SIMD speed ups can be particularly significant when trying to perform linear algebra heavy compuations such as matrix calculations on GPUs

In [257], they showed how to vectorise MCMC by treating the entire operation with matrix algebra. We have an initial matrix which is $(n_{chains} \times n_{dimensions})$ all transition operations are applied to the matrix as a whole which allows for sampling to be carried out in large batches on GPUs with a much less than linear increase in wall time. The results of doing this are equivalent to running n_{chains} independently as is often done by doing n cpus however the computational costs are reduced massively due to SIMD savings.

b HMC

Hamiltonian Monte Carlo involves redefining the problem of sampling from a posterior in terms of classical physics, where the exploration of the typical set is analogous to traversing a potential about the mode. By framing the problem in this way, we see that the natural way to efficiently explore the space is to incorporate the gradient of the posterior surface into our calculation and for our chains follow trajectories that are dictated by this gradient. To do this we first introduce some momentum variables, ρ and define a joint density over our parameters θ and our momenta ρ;

$$p(\rho, \theta) = p(\rho|\theta)p(\theta) \tag{6.2}$$

Generally we take ρ to be draws from a multivariate Gaussian, i.e. $\rho \sim \mathcal{N}(0, \Sigma)$ where the covariance, Σ should reflect your assumption about the covariance be-

tween the physical parameters. Often this is assumed to be a diagonal matrix or is estimated using the initial warm up or burn in phase at the start of the analysis. These draws introduce the stochasticity into our draws and ensure that we will traverse all regions of the parameter space.

The density defined in 6.2 then defines the Hamiltonian of our system.

$$\begin{aligned}\mathcal{H}(\rho,\theta) &= -\log p(\rho,\theta) \\ &= -\log p(\rho|\theta) - logp(\theta) \\ &= T(\rho,\theta) + V(\theta)\end{aligned} \tag{6.3}$$

Where $T(\rho,\theta)$ is the kinetic energy and $V(\theta)$ is the potential energy of the system. Using the standard Hamiltonian equations we can then draw some momentum variable from our pre-defined distribution and evolve the system as:

$$\frac{\partial\theta}{\partial t} = \frac{\partial T}{\partial\rho} \tag{6.4}$$

$$\begin{aligned}\frac{\partial\rho}{\partial t} &= -\frac{\partial T}{\partial\theta} - \frac{\partial V}{\partial\theta} \\ &= -\frac{\partial V}{\partial\theta}\end{aligned} \tag{6.5}$$

Where the derivative of T with respect to θ is zero because the kinetic energy term is independent of the parameters θ, it depends only on the draws for the momenta ρ.

We now have two differential equations which can be solved numerically, in most software packages such as [259, 257], this is done using the Leaprfrog integrator. This integrator takes a step size δ and a number of steps, L, and then moves Lδ steps along the path derived by the integrator to a new position (ρ^*,θ^*).The step size,δ and number of leapfrog steps, L, are hyper-parameters that can significantly affect the efficiency when sampling using HMC. We discuss this further in c.

As with all numerical integrators there is a small error introduced however this can be corrected using the Metropolis acceptance condition, this then guarantees that our chains will asymptocially converge to the true posterior. The Metropolis condition here is $min(1,\exp(\mathcal{H}((\rho,\theta) - \mathcal{H}(\rho^*,\theta^*))$.

In this work we use the ANN-SUR waveform model presented in [67]. This is a reduced order surrogate model for SEOBNRv4 [260] which uses neural networks to perform the difficult interpolation step required for fast and accurate reduced order surrogates. In [67] it was shown that this model can produce very accurate surrogate models and also provide huge speed-ups in waveform generation time relative to the original model. As well as this and more importantly for this work

using ANN-SUR waveform approximants can be evaluated efficiently in very large batches. In the ANN-SUR model the expensive computation is carried out using a neural network which means that the gradient of this computation can be calculated using *Autodiff* [70, 69]. This allows us to compute derivatives relatively cheaply compared to numerical methods and means that this waveform model is compatible with gradient-based samplers such as HMC. Using *Autodiff* allows us to overcome the computational challenges that have meant that HMC has not been widely used in gravitational-wave parameter estimation.

c Implementation

To make use of the methods described above, we designed a gravitational wave sampling code that is compatible with these techniques. This leads to several differences when compared to codes such as *LALinference* or `bilby` . Firstly we had to ensure that all calculations were compatible with Batching, this means that the waveform generation, likelihood calculation and sampling must all be coded up using vectorised operations which preserve both batch and shape semantics. The waveform generation can be trivially batched if we use the ANN-SUR model, we just have to ensure that we pass the physical parameters in batches of shape $(n_{chains}, n_{dimensions})$. Currently ANN-SUR is the only waveform model which is compatible with this methodology, however due to the potential computational benefits it is hoped that other waveforms may adopt the batching philosophy going forward.

Next we must ensure that the likelihood calculation can also be computed in large batches, the implementations in *LALinference* or `bilby` are not compatible with this. To allow us to do this we implemented the gravitational wave likelihood function using *Tensorflow* [261], this means that all calculations such as FFTs, multiplications, etc are handled as matrix operations which can be efficiently batched to produce a vectorised likelihood function. Finally we had to ensure that the sampling itself could be batched, we again make use of recent software developments and use *Tensorflow probability* [258, 257]. This allows for sampling in batches of n_{chains}, as well as this we can make use of their off the shelf implementations of gradient based sampling algorithms which are again all trivially batched.

In this work we use the standard HMC algorithm and Gaussian Random-walk mcmc, for both algorithms we have per-parameter step sizes which are adapted in the burn-in phase to ensure a pre-defined target acceptance rate(0.7 for HMC, 0.25 for MCMC). These optimum target acceptance rates are derived by minimising the *cost* in terms of rejected steps for a given proposal distribution, see e.g. [262]. However for complicated posteriors it may be preferable to optimise explicitly for the effective sample size (and therefore the integrated auto-correlation time indirectly) which would likely mean that our target acceptance rates would be smaller. When using HMC we manually tune the number of leapfrog steps based on several runs.

If the number of leapfrog steps is too low then we are less likely to generate effective samples because the new location is not independent of the previous, this then becomes similar to the expected random-walk behaviour. If it is too large then we are wasting computation because we could have generated a sample without computing extra integration steps. In the most extreme case a very long trajectory could result in a closed orbit in phase space, meaning the end point could be close to the start point, giving a random walk chain as in the case where the step size is too small. The current hand-tuning procedure is clearly not optimal and in future we would move towards algorithms which do this automatically such as NUTS [263] or ChEES-HMC [264].

The other hyper-parameter in HMC is the mass matrix, we use the simple implementation which assumes a diagonal covariance matrix, where the relative scalings are estimated using a Fisher matrix approximation and then hand-tuned over serval runs. This assumption does not therefore incorporate any correlations between parameters, this is again clearly a sub-optimal assumption for gravitational-wave data, therefore is likely a major source of inefficiency for our implementation. For example the distance inclination degeneracy that we see in figure 6.1, is well understood from a physical perspective, the amplitude of the waveform is proportional to both the distance and the inclination angle, therefore a face on binary which is further away could have the same amplitude as a binary which is much closer but more inclined. If we could encode this knowledge into our samplers we could move along the degenerate lines and therefore traverse the space much more efficiently. As well as encoding physical knowledge we could estimate the covariance matrix as is done in the STAN [259] HMC implementation. There the mass matrix is adapted on the fly during the warm-up phase, this is also something we would like to implement in future.

To summarise the implementation, we have for the first time constructed vectorized MCMC and HMC analyses for GW parameter estimation. These are both very simple, generic versions of these algorithms which lack many features such as an appropriate estimate for the covariance matrix. We also only currently use a single detector and ignore some physics such as precession and higher modes, however despite these simplifications and limitations, we are still able to produce sensible parameter estimates very quickly. We are also able to run initial tests and identify potential performance improvements that could be incorporated into a full parameter estimation analysis in the future.

6.3 Results

As an example of this method we injected a fiducial signal into Gaussian noise and perform a single detector parameter estimation analysis. We inject a non-spinning M=70 (solar mass, defined in the detector frame), q = 1.2 ANN-SUR

signal into Gaussian Noise at SNR=18, we use the H1 design sensitivity PSD and use a start frequency of 30 Hz. Other parameters are shown in table 6.1. ANN-SUR waveforms ignores precession and higher modes, this leaves us with the two masses (m_1, m_2), two spins (χ_1, χ_2), the binary inclination angle (θ), the orbital phase (ϕ), the coalescence phase, (ϕ_c), the luminosity distance (D_L), the right ascension ra and declination (dec), and the binary polarization angle (ψ) and the time of arrival t_c.

This results in an eleven dimensional posterior surface that we would like to sample from, when sampling we parameterise the masses as $M = m_1 + m_2$ and $q = \frac{m1}{m2}$. We also analytically marginalise over time and phase (see the appendix of [40]). This means we are in effect only sampling over nine dimensions. We will use this fiducial example to verify that both samplers converge to the expected parameters, we then move on to look at the increased efficiency due to batching and finally we highlight key differences between the results obtained using our MCMC and HMC implementations.

a Comparison between sampling methods

In table 6.1 see that using both methods we are able to produce a set of samples which converge to distributions which contain the true parameters at the 90% level. Quantitatively we defined our convergence threshold to to be at least 1000 effective samples for each parameter and an $\hat{R} < 1.01$ The parameters ψ, θ, ϕ, ra and dec recover uninformative priors because we are only using a single detector and a non-precessing signal therefore there is no information about these parameters in the data.

When using the random walk MCMC the analysis takes around 15-30 minutes, if using HMC this is around 60-90 minutes. The corner plot shown in figure 6.1 shows a six dimensional corner plot for the injection (we also sample in ra, dec, ψ and ϕ but exclude them from the plot). We see that the sampler was able to find the true values as well as the expected non-trivial correlations such as in distance and inclination.

We now look at the benefits that are gained by our vectorised sampling approach using GPUs and then look differences in terms of speed and efficiency between the methods, highlighting the key differences between the two approaches.

b Computational Efficiency

In figure 6.2, we show the number of samples per second generated by each of the sampling methods as a function of batch size. Here we ignore the effective sample size as we are purely looking at the computational benefits of batching the calculation. We use an NVIDIA Tesla V100-SXM2-16GB GPU and run the samplers for 2000 iterations per chain. We see MCMC produces many more samples per second, this

	HMC	MCMC	Injected
M	$68.8^{+5.9}_{-6.4}$	$69.2^{+5.8}_{-4.4}$	70
q	$1.5^{+0.6}_{-0.5}$	$1.39^{+0.4}_{-0.4}$	1.2
χ_1	$-0.08^{+0.6}_{-0.8}$	$0.04^{+0.7}_{-0.7}$	0
χ_2	$-0.06^{+0.7}_{-1.0}$	$-0.15^{+0.7}_{-1.0}$	0
θ	$3.21^{+3.1}_{-2.7}$	$3.20^{+3.1}_{-2.7}$	π
ϕ	$3.10^{+2.5}_{-3.1}$	$3.18^{+3.1}_{-2.5}$	π
D_L	349^{+210}_{-140}	373^{+221}_{-239}	500
ra	$3.20^{+2.5}_{-3.1}$	$3.17^{+2.6}_{-3.1}$	1
dec	$0.04^{+1.5}_{-1.2}$	$-0.04^{+1.4}_{-1.3}$	1
ψ	$3.20^{+3.1}_{-2.6}$	$3.15^{+3.1}_{-2.5}$	1

Table 6.1: Summary of the posterior distributions obtained by both MCMC and HMC, numbers shown are the median and then the upper and lower bounds for the 90% highest density interval.

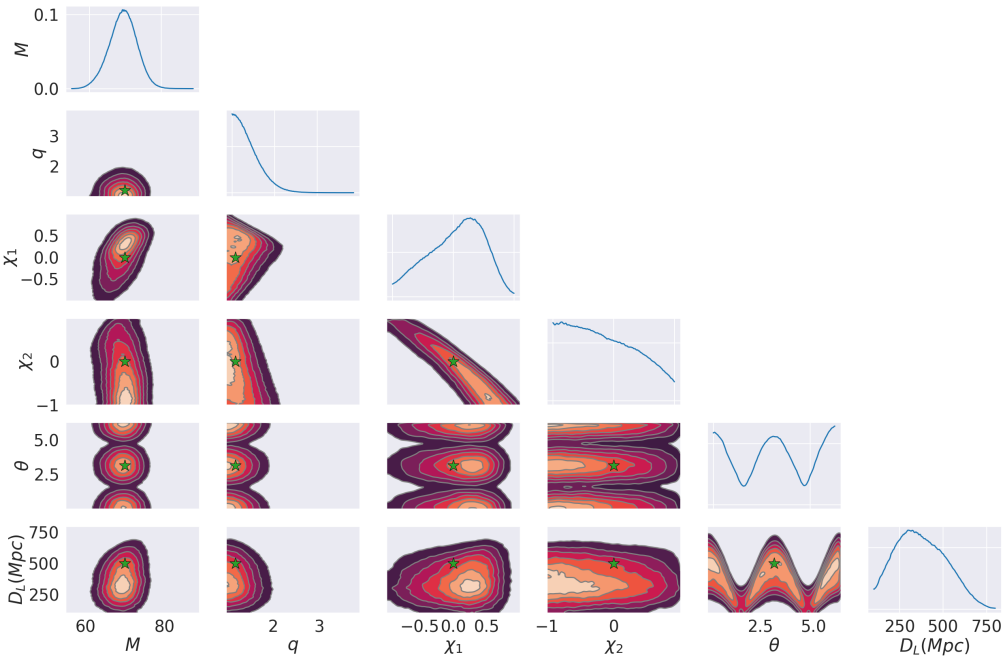

Figure 6.1: Results from a single detector MCMC analysis, the injected value is shown as the red star

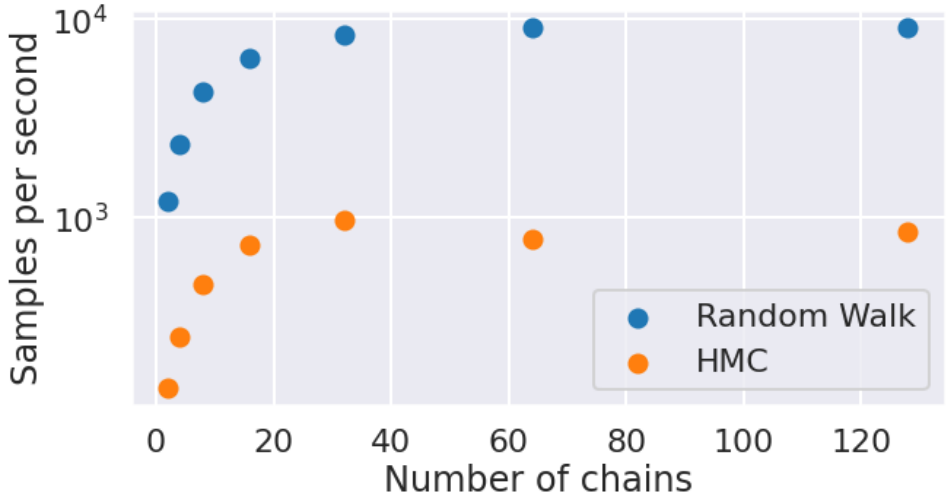

Figure 6.2: Comparing the number of likelihood evaluations per second as a function of the batch size, we see that up to 32 chains we get a better than linear scaling, a linear scaling would be shown here as a horizontal line.

is expected because HMC does more computation before taking a step, this extra computation means that samples are less likely to be rejected and/or correlated compared to MCMC. HMC is then in principle a a slower but more efficient sampler. For both methods we see that the effect of batching is huge, the gradient of the curves show the increase when adding a new chain, going from 2-32 chains in both cases gives an increase in samples per second of around 10x for both sampling methods, this speed-up can be translated exactly into an overall speed up of a parameter estimation analysis.

If we were to reproduce this plot without batching and used a single CPU we would expect to see a linear scaling which would correspond to a roughly horizontal line on the plot, if we double the chains we could get double the samples but it would take approximately double the time. This is why multiple chains are generally parallelised across CPUs. If we were batching but using a CPU we would expect to see similar pattern, i.e. some some SIMD efficiencies, but these would likely not be as large.

The numbers shown above are for the raw numbers of samples produces per second however, due to the often correlated sequences of samples produced an MCMC analysis, the number of samples generated per second is not a particularly good metric when comparing samplers. One must also consider how many samples are needed to generate an independent or effective sample. Where we define the number

of effective samples, n_{eff} as:

$$\begin{aligned}\vec{n}_{\text{eff}} &= \frac{Mn}{\sum_{t=-\infty}^{\infty} \vec{P}_t} \\ &= \frac{Mn}{1 + 2\sum_{t=1}^{\infty} \vec{P}_t}\end{aligned} \tag{6.6}$$

Where M is the number of chains, n is the number of samples per chain, P_t is the estimated auto-correlation[2] at lag t. In practise we do not sum to infinity we sum up to some k, where k is the is the largest integer for which $P_k = P_{2k'} + P_{2k+1}$ is positive. The number of effective samples corresponds to the number of *independent* samples in our chains. n_{eff} is a vector because the autocorrelation length will vary between parameters, we therefore have an effective sample size per parameter.

In figure 6.3, we show the number of samples required to generate an *effective sample*[3], we calculate this as $\frac{n}{n_{eff}}$. The error-bars on the barplot indicate the per parameter spread in this metric, We can see that for this particular example HMC is around a factor of 2-5 more efficient that our random walk MCMC. We would expect the efficiency improvement to be considerably larger than this, for a well tuned HMC sampler we would expect the efficiency to be around D times better for a D dimensional problem (we sample in nine dimensions in this case) [267, 254]. This suggests that our HMC implementation is not optimally tuned. We have mentioned several reasons why this might be the case throughout the paper, also see section 6.4 for more discussion on this.

Combining the sampling efficiency we produce here this information with the figure 6.2, we can estimate the relative time difference between the two samplers for this particular problem. To produce the same number of effective samples, then the random walk sampler around 10 times as fast, but needs 3 times as many samples for an independent sample, so it should be roughly 3 times as fast overall. This agrees with our experience when running the code and also accounting for effects such as burn-in and variance between runs. The relative benefits between efficiency in terms of effective samples and samples produced per second may become more significant as we move to higher dimensional problems and include more complicated parameter spaces, for example if we include precession.In these examples we expect random walk samplers to become considerably less efficient, it would then be likely that a slower but more efficient sampler would actually produce effective samples more quickly.

[2]Following STAN [259] and ArviZ [265] we use the definition of autocorrelation from [266]
[3]note we are ignoring the burn-in efficiency in this section, see [253] for a discussion on this

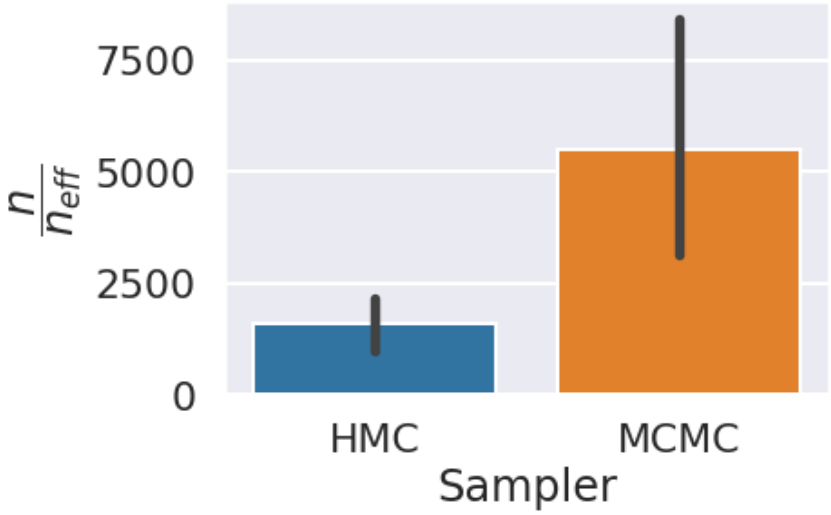

Figure 6.3: Comparing number of samples per effective samples generated, we see that HMC is generally more efficient, the error bars are due to the spread in effective sample size across the different parameters.

6.4 Discussion

This work has developed a framework in which gravitational-wave astronomy can exploit developments in software (batched sampling), hardware (GPUs) and waveform modelling to perform parameter estimation efficiently. We've presented two methods which both show promising results however the current implementation astrophysically. We have only used a single detector which means we are unable to recover informative sky localisation. We also do not make use of the gravitational-wave specific jump proposals which have been shown to make samplers considerably more efficient. Combining these with either strategy would likely make the samplers considerably more efficient. Both of these limitations are currently just practical, there is no reason why any of the arguments presented in this chapter would be effected by them, it is in fact likely that the sampling would become even faster.

The other major limitation of this work is that we also ignore higher modes and precession, this is because currently there are no waveform models which contain these and are compatible with the batching methodology which is essential here. Including these effects generally makes the parameter space more complicated and would therefore increase the amount of time required to fully explore the typical set. This would likely make HMC more efficient relative to random walk MCMC as the guess and check strategy would rapidly produce less effective samples per likelihood evaluation.

Outside of methods that are currently used in gravitational-wave data analysis, there are several improvements that could easily be implemented to improve this work further. Firstly we currently use a diagonal mass matrix, this implicitly

assumes that parameters are not correlated, this is clearly a bad assumption assumption for gravitational-wave data. In the Stan HMC implementation this can be estimated on the fly using a warm-up phase, doing something similar here would likely lead to large improvements in efficiency. As well as estimating this on the fly we could provide good estimates analytically based upon the Fisher matrix and/or physical insights about the parameter space. An interesting enhancement when including precession would be to use ρ_p and encode our understanding about the precessing parameter space into our Mass matrix.

Another improvement that would now be feasible would be to follow the approach presented in [268] and use Neural Networks to learn a bijective map between some simple parameter space and the true complicated gravitational-wave parameter space. A similar method was proposed in [269], where they used normalising flows to learn the contour geometry in nested sampling. This step could be added on at the beginning of an analysis and in principle would be able to remedy some of the sampling difficulties due to the non-trivial geometry of the gravitational wave parameter space.

It is hard to draw conclusions when comparing these methods to existing packages such as [52, 253] because of the difficulty in obtaining an apples to apples comparison between the implementations. These packages benefit from years of gravitational-wave specific knowledge that has resulted in samplers which are very well tuned to the problem. This work has presented simple generic methods but make extensive use of modern hardware and developments in waveform modeling, it does however highlight promising techniques such as batching and more efficient gradient based sampling which could be incorporated into these software packages and provide major benefits in the future. By combining the methods presented with the current knowledge already in the parameter estimation community, it is feasible that we will bring the analysis down to a practical timescale and will not have to abandon stochastic sampling as the primary method for parameter estimation.

Chapter 7

Model Agnostic Confidence Estimation - Classification

7.1 Introduction

Over recent decades a huge amount of progress have been made on improving the global accuracy of machine learning models, however calculating how *likely* a single prediction is to be correct has seen considerably less attention. In some fields, where making a single bad prediction can have major consequences, having trustworthy confidence estimates may be the limiting factor before introducing AI. It is important in these situations that a model is able to understand how likely any prediction it makes is to be correct before acting upon it; being able to do this well requires satisfying two closely related conditions:

1. Confidence estimates must approximate the true frequency of being correct, i.e. if a model estimates a confidence of 80% it should be correct roughly 80% of the time.

2. Confidence estimates should also indicate *ignorance*, i.e. the model must know what it doesn't know so that it will not blindly make bad predictions.

A similar way to think about confidence estimation is to say that any estimate must account for any *uncertainty* that is present. Uncertainty can be split into two forms ([270, 271]):

1. Aleatoric Uncertainty: this refers to the intrinsic variance or randomness inherent in any process, i.e. even with unlimited data there will always be errors in any modelling and therefore aleatoric uncertainty cannot be reduced by collecting more data.

2. Epistemic Uncertainty: this refers to the uncertainty due to the lack of knowledge of a model, the knowledge of any model comes from data, so does the

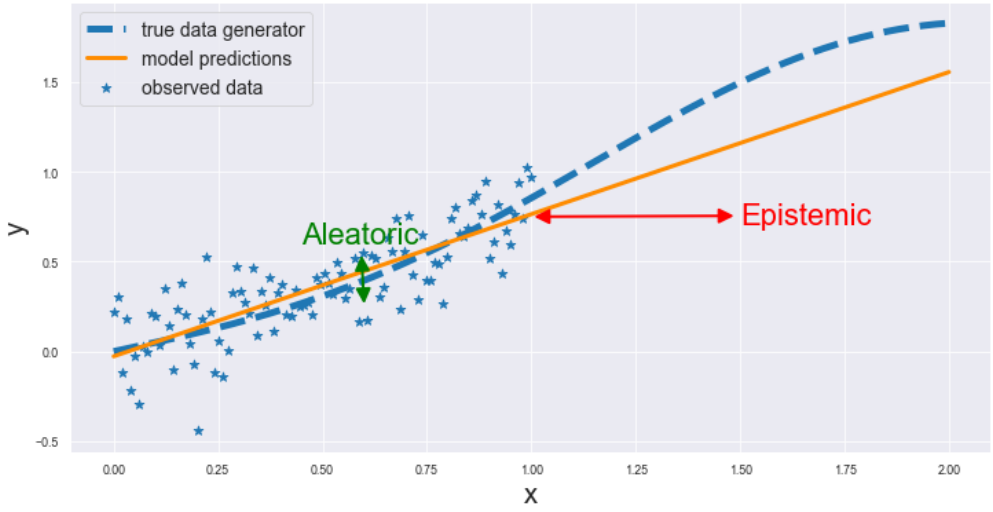

Figure 7.1: A simple example illustrating the two types of uncertainty: Aleatoric refers to the uncertainty due to the intrinsic randomness inherent in any system. This is shown by the spread of observed data about the learnt model which is a good approximation to the true model in the region we have observed data. Epistemic uncertainty is the uncertainty due to a lack of knowledge (i.e data) to inform the model prediction. This is shown above when the model becomes a worse approximation to the true function as we move further away from the observed data. (e.g x> 1)

model have the *relevant* data to predict something? Epistemic Uncertainty can always be reduced by collecting more or different data.

Understanding and accounting for uncertainty is crucial to understanding the capabilities and limitations of any machine learning model but it is especially crucial if one wants to produce a confidence estimate. Therefore confidence estimation for a prediction is most naturally expressed in a Bayesian language where uncertainty can be explicitly accounted for in both the data and the model. Once we have accounted for this uncertainty we can then reason about the effect it has upon a prediction; this is in general done by marginalising out our uncertainty to produce a distribution of outputs for any prediction which will then reflect our confidence ([272]).

Despite the Bayesian framework being the natural way to tackle the problem of confidence estimation, it is often prohibitively computationally expensive in its application. It also does not naturally lend itself to being combined with popular machine learning algorithms which result in a large number of parameters or non-linearities such as Decision trees, Random Forest, support vector machine etc. ([273, 274, 275])

Most popular machine learning algorithms adopt methods to produce an estimate of how likely a prediction is to be correct, for example a random forest model ([274]) estimates the confidence score by the fraction of trees that predict a certain class. Each of these techniques are in effect *heuristics* which correspond to ordering

predictions that are more or less likely, however they certainly cannot be interpreted as true probability or confidence estimates.

Although these heuristics can be useful, and when calibrated can estimate the aleatoric uncertainty well (see a), we argue that they will always be fundamentally flawed when estimating confidence due to using the classification model estimate as a starting point. See Section 7.5 for a more detailed analysis of this claim.

In [276] the authors present a method which provides Trust scores that attempt to quantify how much one can trust a classifier, this is very closely related to the notion of our confidence in being correct. These scores are computed explicitly by calculating the relative distance to a set of k nearest neighbours from each class. This can be seen as quantifying the epistemic uncertainty for a given prediction.

In this work we highlight the fundamental problem with using point prediction algorithms (even if well calibrated to data) and present an alternative model MACE (Model Agnostic Confidence Estimator) that seeks to bridge the gap between trust scores and the current state of the art for confidence calibration. We show that for many tasks our method is competitive with state of the art and, as it can directly account for both aleatoric and epistemic uncertainty, is more robust to problems such as extrapolation and out of sample bias.

The original contributions of this work are as follows:

- We compare trust scores and confidence calibration highlighting key similarities and differences between these methods.
- We show that generally calibration methods do not properly account for epistemic uncertainty and therefore can return high confidence but a low trust score.
- We present a novel algorithm called MACE, which is agnostic to the model which produces point predictions and which explicitly models both aleatoric and epistemic uncertainty.
- We then demonstrate that MACE is competitive with popular calibration methods across several metrics, and also addresses the problems other models have when epistemic uncertainty is large.

7.2 Confidence Estimation

The general approach of classification involves using labelled pairs of training data $(\mathbf{X}, \mathbf{y})$ where each $\mathbf{x_i} \in \mathbf{X}$ is a vector corresponding to the features, and each $\mathbf{y_i} \in \mathbf{Y}$ is an associated class for the element i: the goal being to learn a transformation function $\psi(\mathbf{X})$ that maps any new unseen data point $\mathbf{x}_*$ to an estimated label $\hat{y}_*$.

The problem can then be interpreted as wanting to estimate the class label y_* of $\mathbf{x}_*$ given a learnt model $\psi(\mathbf{X})$ from training data $\mathbf{X}$. It is important to highlight that, although somewhat obvious, any prediction explicitly depends upon the specific model which has been trained using a specific set of training data $\mathbf{X}$. The importance of this statement becomes more obvious when one looks at confidence estimation. For a given point $\mathbf{x}_*$ we estimate the confidence as:

$$C(\hat{y_*} = y_* | \mathbf{x}_*, \psi(\mathbf{x}_*), \mathbf{X}) = \hat{p_*} \tag{7.1}$$

Where y_* is the true class label, $\hat{y_*}$ is the predicted class label for a point $\mathbf{x}_*$ and $\hat{p_*}$ is the estimated confidence associated with that prediction. This then corresponds to estimating the probability that our prediction $\hat{y}$ is correct given a model, ψ learnt from the training data $(\mathbf{X}, \mathbf{y})$.

Note that in this work we will be looking at *Confidence estimation*, therefore we only look at the confidence of a point prediction being correct rather than estimating the full distribution across all classes (these are equivalent for binary classification but not multi-class) - this is similar to the definitions in [277]. We argue this is the more natural question to be answered by a confidence estimator given an already trained point prediction model, it is also more similar to the notion of Trust Scores defined in [276] to which we would like to draw parallels.

The confidence estimator C defined above must also explicitly account for the training data and, more importantly, how *informative* the training data is when trying to classify a specific point, x_*. This notion is the foundation of Gaussian Process modelling ([278, 226]) where the confidence on any prediction is calculated by explicitly considering the kernel (or co-variance) function $K(\mathbf{x}_*, \mathbf{X})$, the kernel can also be interpreted as a measure of similarity between a point $\mathbf{x}_*$ and the training data $\mathbf{X}$. Here, the confidence of any prediction is conditioned upon the similarity between the predicted point and the training data: if a point is not similar to what we have seen during training we therefore cannot be confident when predicting it.

Having set out what we mean by a confidence estimate (equation 7.1), for brevity and ease of comparison of notation with other literature we will also henceforth adopt the convention of (whilst of course not ignoring) leaving the dependence upon training data and model implicit. We therefore define the confidence of a given point prediction being correct as:

$$C(\hat{y_*} = y_*) = \hat{p_*} \tag{7.2}$$

where again we define p as the confidence of being correct, y_* as the true label, and $\hat{y_*}i$ as the predicted label.

7.3 Performance Metrics

a Calibration Metrics

The calibration of a confidence estimator is the degree to which the estimates match the empirical accuracy of a classifier, i.e. if a confidence estimator estimates a probability of 75%, this should correct approximately 75% of the time. We say that it is *perfectly calibrated* if this is true for all estimated probabilities, $p \in [0, 1]$. If this is not the case the case then we say that the confidence estimator is either over or under confident: if it is correct less often than the estimated probability an estimator is said to be over-confident and vice versa.

Formally this can be defined as needing to satisfy the following condition:

$$P(\hat{y}_i = y_i | \hat{p}_i = p_i) = p \quad \forall p[0, 1] \tag{7.3}$$

As the true probability of an event is an unknown random variable, it is generally not possible to calculate this exactly, so the standard technique is to convert the continuous confidence space into a set of n discrete bins, B_n (e.g $[0.5, 0.6, 0.7, 0.8, 0.9, 1.]$) and compare the average confidence estimate $\bar{p}$ for each bin to the empirical accuracy in each bin.

That is:

$$Acc(B_n) = \frac{1}{|B_n|} \sum_{i \in B_n} \delta_{\hat{y}_i, y_i} \tag{7.4}$$

Where,

$$\delta_{\hat{y},y} = \begin{cases} 1, & \text{if } \hat{y} = y \\ 0, & \text{otherwise} \end{cases} \tag{7.5}$$

And,

$$C(B_n) = \frac{1}{|B_n|} \sum_{i \in B_n} \bar{p_i} \tag{7.6}$$

A perfectly calibrated estimator will then have $C(B_n) = Acc(B_n)$ for each bin.

The standard metric when looking to evaluate the calibration error directly is Expected Calibration Error (ECE) [279]. This metric measures the difference between the predicted confidence estimates and the empirical accuracy of each bin. These residuals are then combined in a sum weighted by the number of points in each bin:

$$ECE = \frac{1}{|n|} \sum_{i=1}^{n} |C(B_i) - Acc(B_i)| \ |B_i| \tag{7.7}$$

In [280] the authors evaluate this metric and highlight potential problems when using it. In particular it is shown that many somewhat arbitrary choices in the metric such as the choice of norm, binning strategy, and class weighting can produce metrics which are measuring slightly different definitions of calibration error. They

produce a total of 32 possible variations of this calibration error and when comparing calibration methods their the ranking is generally not consistent across these 32 metrics.

It is likely that the optimal calibration metric will depend upon the use case in mind and therefore that general comparisons between calibration methods are likely to be a difficult problem.

In this work we will evaluate calibration methods using the probability that the classifier is correct together with adaptive binning schemes and weighted by the total number of points in each bin rather than also conditioning on the class - this corresponds to the general calibration metric 20 defined in the appendix of [280]. We chose this metric for ease of comparison with other studies (such as [277, 279]) and because the metric is sufficient to highlight the utility of MACE for confidence estimation.

b From Scoring Rules to Metrics

As noted previously, in particular in [280], it has been shown that calibration errors can sometimes be misleading. It is therefore important to have additional criteria against which to measure estimators. The quality of a set of probabilistic predictions has been a long-standing issue in many fields such as Meteorology where, as well as calibration metrics, the notion of *proper* scoring rules have developed.

Proper scoring rules are metrics which, when minimised, correspond to approximating the ground truth probability distribution [281]. This means that a *biased* forecaster will perform worse in these metrics than an honest forecaster - evaluation by proper scoring rules was advocated by [270]. The two most popular scoring rules are the Brier Score ([282]) and the Negative log loss ([283]).

As we are looking to draw parallels between trust scores and calibration methods we are trying to estimate the probability our point prediction is correct, therefore these scoring rules are also defined with respect to a prediction being correct, i.e. $O_i = \delta_{\hat{y}_i, y_i}$. This is equivalent to giving class probabilities in the binary case but equates to a one vs. rest strategy in multi-class problems. We therefore use the Brier Loss as follows:

$$BL = \sum_{i=1}^{n} (p_i - O_i)^2 \tag{7.8}$$

And the Negative Log Likelihood (NLL) as:

$$NLL = -\sum_{i=1}^{n} \left(O_i \log(p_i) + (1 - O_i) \log(1 - p_i) \right) \tag{7.9}$$

Intuitively both of these metrics penalise (via squared loss for Brier and logarithmic loss for NLL) incorrect predictions with high confidence but also correct predictions with low confidence. Therefore performing well in these metrics gener-

ally corresponds to predicting correctly with high confidence and predicting incorrect with low confidence. This differs slightly with respect to calibration metrics where the only measured quantity is how well the predicted confidence scores approximate the empirical accuracy of the predictions.

We will estimate all evaluation metrics using K-fold cross-validation, this involves splitting the data into K sets uniformly, you then use K-1 of these sets for training (both the point prediction model and the mace parameters see 7.6) and use the Kth set for evaluating your metrics. You repeat this K times so you have K evaluations of your model on different subsets of your data. This in principle estimates the variation in model performance that one might expect to see using the model on unseen data. We report this error as twice the standard deviation of the K folds; thereby approximating roughly the 95% error interval (under the assumption of Gaussian errors). This provides a useful way to evaluate if there are any *significant* differences between calibration methods.

7.4 Related Work

a Confidence Calibration

Confidence calibration was first studied in [284], since then there have been many more calibration techniques developed [285, 286, 279, 277, 287]. Each of these techniques aim to transform the raw scores from the point prediction algorithms into a confidence score that approximates the empirically derived accuracy. This is generally by some distributional transformation or scaling, where the parameters required to do the transformation are learnt using a single hold out set of data.

For example if the point prediction model predicts a confidence score of 80% one of the techniques above may learn a transformation that down-scales that score to a 70%. It should be noted here that these transformations are dependent upon the point prediction algorithm score only, not they are not explicitly dependent upon the data point, for which they are asked to give a confidence estimate.

These methods have been shown to work very well across a range of machine learning problems. However, as we will show in section 7.5, they learn a global representation of uncertainty and are therefore often only actually learning to model aleatoric uncertainty. This means that they are therefore vulnerable to giving overly confident predictions due to ignoring epistemic uncertainty.

b Trust Scores

In ([276]) the authors introduce a method which produces a *trust score* for model predictions, i.e. a high trust score means that a model is likely to be correct and vice versa. This notion is very closely related to the confidence one has that a prediction is correct. They approach the problem of trust by looking at the similarity between

the point you want to predict and the training data. The relative distances between a set of k nearest neighbours from the predicted class and from all other classes are then compared. For example a trust score of 1 corresponds to the prediction being equally as close to the k nearest neighbours from the predicted class to those of all other classes; a trust score of 2 means it is twice as close to the predicted class, etc. The intuition here is that if a data point is similar to previously seen points of the same class then the prediction is more likely to be correct and therefore more *trustworthy*. This can be interpreted as estimating the epistemic uncertainty: when the trust score is large, the distance to similar points is small and therefore the epistemic uncertainty is small. One would then expect that there is a greater chance of the prediction being correct, i.e. more likely to be trustworthy.

c Other related works

Quantifying the uncertainty in a prediction is most naturally considered within a Bayesian framework. Parametric methods such as Bayesian parametric modelling ([288] [289]) and non-parametric methods such as Gaussian Process (GP) Regression ([278, 226]) and Bayesian Neural Networks ([251]) have proven to be effective for providing prediction intervals. These methods however suffer from some drawbacks that MACE attempts to address, generally they are considerably more computationally expensive at both training and inference time than algorithms that do not include confidence estimates: this is potentially a major barrier for many applications.

There is in general no method that the authors are aware of to explicitly combine Bayesian uncertainty estimates with an arbitrary point prediction algorithm. MACE is not a Bayesian algorithm however, and, as described in the introduction, it is motivated by some of the underlying Bayesian principles. It seeks to bridge the gap by combining some of the benefits of Bayesian modelling with less computational costs and, by being compatible with any point prediction algorithm, considerably more flexibility.

Outside of the explicitly Bayesian framework there are several methods that have been utilised for the problem of confidence and uncertainty estimation. Dropouts method introduced in [290] could potentially be applied to other methods by perturbing or dropping model parameters and generating a range of predictions, however such Monte Carlo simulation methods used to get a good coverage of the parameter space where the number of model parameters is large are computationally expensive. Ensemble methods similar to those introduced for neural networks in [270] could also be applied for any model, however the cost of training a large enough ensemble to is often prohibitive for many applications.

7.5 Calibrated Models Are Not Necessarily Trustworthy

Above we have seen that there are two different approaches to answering very similar questions, i.e. how confident are we in a prediction (calibration methods)? And how much can we trust a prediction (Trust scores)? Intuitively these seem to be very related, yet the two approaches estimate their score very differently. This section asks the question: can we trust well calibrated estimators?

To answer this question we perform classification on the classic MNIST data set ([291]), which consists of $60,000$ examples of 28x28 pixel images of handwritten digits with the task of classifying the digits. As trust scores use a distance metric the raw pixels are not suitable so we take 50 principal components and use those as our features. We then train a simple random forest model as the point prediction model and use a selection of the cited calibration methods described above to compute confidence estimates. We now look at the correlation between trust scores and these confidence estimates.

Figure 7.2 shows the joint distributions for the confidence estimates and trust scores evaluated on an unseen test data set. The point prediction model is generally very accurate ($\approx 95\%$) and thus, as expected, for each confidence estimator we have a distribution of confidence which is very high, thereby corresponding to high (≈ 2) trust scores. We calculate the Spearman rank correlation coefficient between the confidence predictions and trust scores, and find that for each calibration model the correlation between them is very strong (Isotonic = 0.797, Platt = 0.803, Temp = 0.814, Dirichlet= 0.773). This is generally what we should expect: the models are reliable and predicting well and so both the confidence and trust scores are high. In this situation we see that the standard calibration methods are working well, and the confidence estimates are reliable.

Next, we simulate 28x28 pixel images using $10,000$ different uniform noise realisations and ask our model to make predictions on this data (See figure 7.3 for an example of this) - after having trained the models on the above MNIST dataset of labelled samples. As this does not look like any digit, we would expect the confidence estimates also be very low for any prediction and should, ideally, return something like a uniform distribution across all classes that would then reflect the minimum confidence. We can see from Figure 7.4 that although the confidence is considerably lower than most of the in-sample confidence estimates, it is still often high. Most predictions are greater than 0.5 and there are a considerable number ($\sim 30\%$) which are greater than 0.8.

What about the trust scores? We see in Figure 7.5 the trust scores are now much lower. 95% of trust scores are now ≤ 1.05, this reflects our intuition that noise should be roughly as close to the predicted class as any other. When comparing to

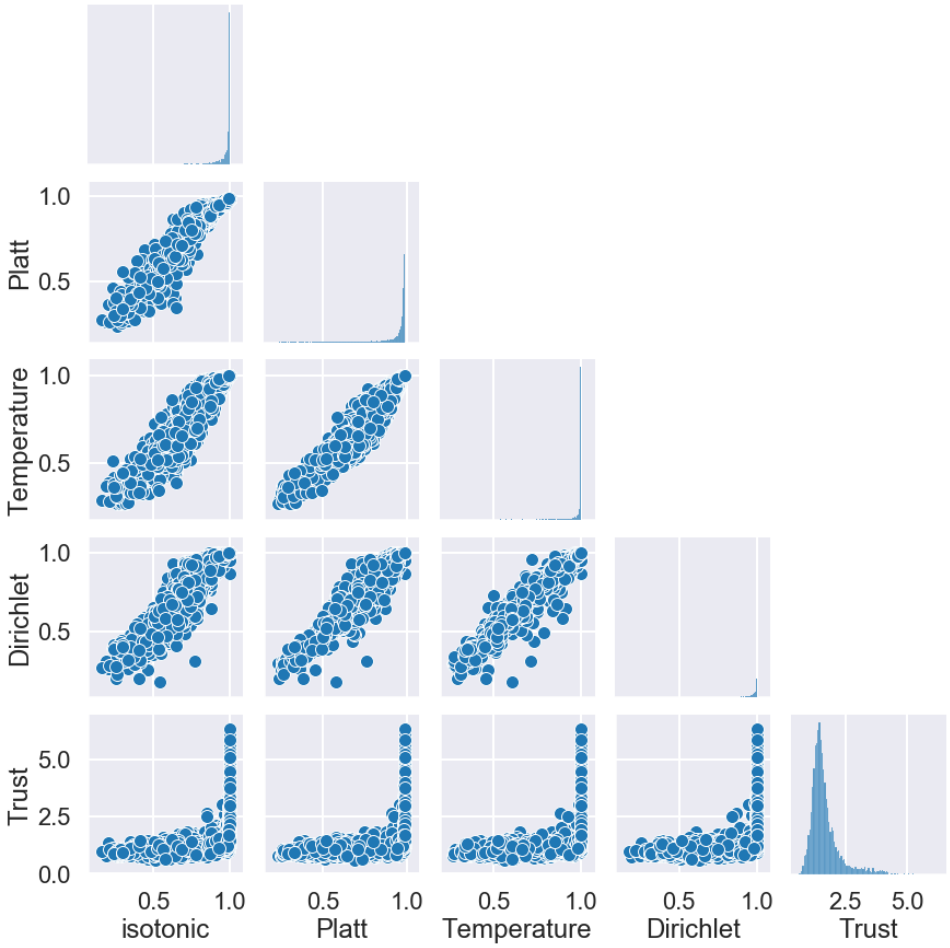

Figure 7.2: Pair plots comparing trust scores and confidence predictions when predicting on an unseen test set for the mnist dataset. Trust scores are correlated strongly with confidence predictions here

the confidence scores however we see that the correlation between trust scores and confidence estimates has effectively disappeared (Isotonic = 0.056 , Platt = 0.085, Temp = 0.108, Dirichlet= 0.089). This example highlights that despite confidence estimators being calibrated to the data they are not necessarily trustworthy, as posited by [276].

We argue that this effect is fundamental to the way classification models work in practice. The classification paradigm is effectively to learn a global decision boundary and split the features into distinct regions of classes. This is a natural and incredibly effective way to perform classification, however we argue it is not a good way to estimate confidence. This is because it does not take into account the *similarity* of a given point to the data that the model used to train, resulting in epistemic uncertainty never being accounted for. If a point falls far away from the decision boundary it is confidently predicting a label even if it is completely different to any of the examples that trained a model ([290]). This can be seen clearly in the simple example shown in Fig. 7.6: here we have an example with a set

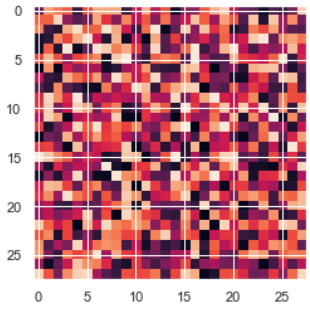

Figure 7.3: An example of the uniform random noise which we asked our model (which was trained on MNIST) to classify. Confidence estimates calibrated using Dirichlet, Temperature, Isotonic and Platt all reported $> 85\%$ confidence.

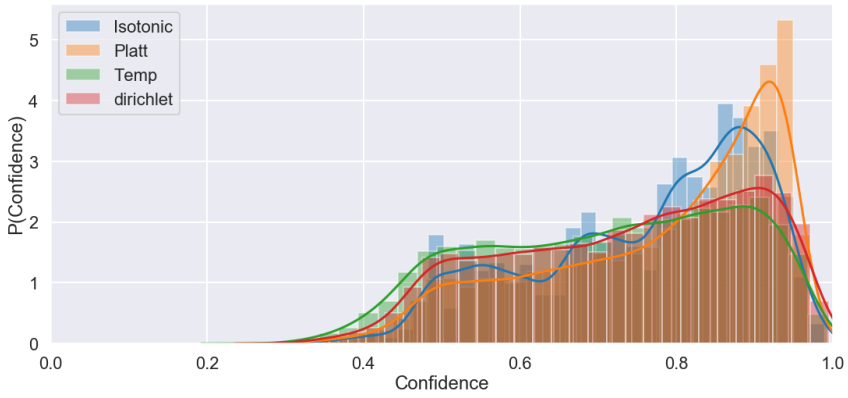

Figure 7.4: Example of 10^4 confidence estimates on random noise similar to Figure 7.3: we show the probability distribution function (top) and the probability density function (bottom). These distributions *should* be shifted towards 0.1 thereby indicating the model's lack of confidence given the presence of pure noise: instead, we see for all calibration methods that there are a large number of high confidence predictions despite the input.

of simulated points in two classes which follow a spiral. There is a good distinction between the classes and therefore the model is able to clearly split the feature space into regions where it will predict either red and blue. In the regions close to the data the confidence is very high because the model is generally predicting very accurately. As we move away from the training data, we see that the predictions are extrapolated according to the global boundary which splits the feature space. We also see however that the confidence estimates are extrapolated, this means that in regions of the feature space very far away from the training data the model is still returning very high confidence results. We argue that the confidence estimate should *decrease* as we move away from the training data because the epistemic uncertainty is considerably higher in these regions.

Both of these examples highlight a fundamental problem with using the model estimates as a starting point for a confidence estimate. Confidence estimates derived from the model estimates can generally be calibrated to estimate the aleatoric

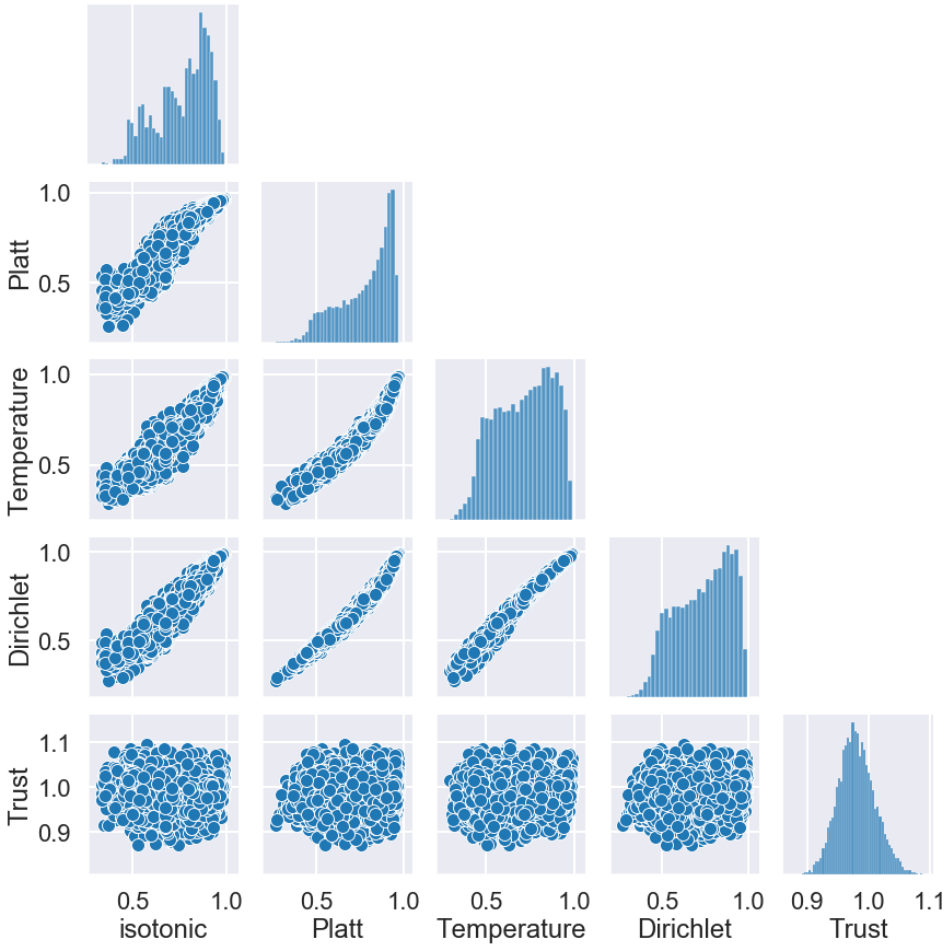

Figure 7.5: Pair plots comparing trust scores and confidence predictions when predicting uniform noise: when epistemic uncertainty is large then trust scores and confidence estimates become uncorrelated. Trust scores are generally low ~ 1 but confidence predictions can be high. Note the correlation between the predictions of the models, this is because they all perform a slightly different transformation on the same set of original predictions: therefore the ranking of points will not change.

uncertainty well, however the confidence estimates are not reliable when being used for data which is significantly different from the training set. This can often result in over-confident predictions, despite a high degree of epistemic uncertainty being present.

Trust scores, as an estimate of the distance to similar data from the training set, are an effective way of highlighting cases where the epistemic uncertainty is high, they however do not have any mechanism to account for aleatoric uncertainty and are also not easy to interpret as a likelihood or probability of being correct. The algorithm we now present seeks to bridge the gap between these two approaches by producing well calibrated confidence estimates which account for both epistemic and aleatoric uncertainty.

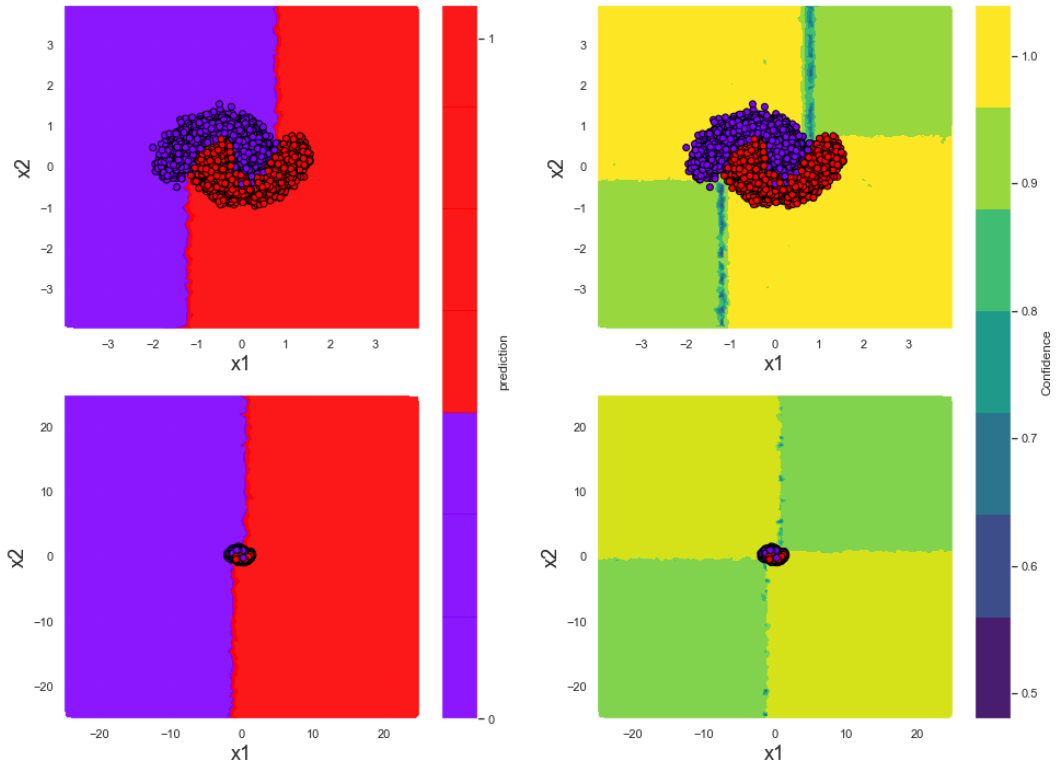

Figure 7.6: Example illustrating the problem with extrapolating confidence: because the confidence estimate does not account for the *similarity* to the training data, the model confidence will not decrease however far we extrapolate from the data from which the model learnt.

7.6 MACE

We start with the heuristic that confidence is a local quantity, e.g. if the model is terrible globally then there are still some predictions for which the model can be very confident. Similarly if the model is very accurate (even perfect) globally on a given test set then for any given new prediction there exists the possibility that the model is not capable of providing a good answer and should therefore have very low confidence for this point. We also argue above that the point prediction model may fail to produce good confidence estimates and therefore the confidence for a prediction should be estimated independently of the model which produced the classification.

The following assumptions then outline our approach to model the local confidence estimate:

1. If we can define a notion of similarity between data points then members of the same class should have a similar representation in this feature space.
2. A point that not similar to any training data is an unusual point and predictions on that point should not have a high confidence value associated with them.

3. The distance to a set of nearest neighbours from the training data is a good proxy for how similar any point is to the training data, this can then be used to estimate the epistemic uncertainty.
4. The accuracy of a set of predictions for a set of k nearest neighbours can be used as a proxy for the local prediction variance, i.e. the aleatoric uncertainty for the prediction can be estimated by the accuracy in these k nearest neighbours.

To implement this for a given prediction on a point $\mathbf{x}_*$, we define the local neighbourhood by finding the k nearest neighbours to $\mathbf{x}_*$. The parameter k is learnt during training when we are seeking to minimise the calibration error. This means we are in effect learning a parameter, for a given dataset, which defines the size of a local neighbourhood in terms of data points. We assume that this is a global property for a given dataset however in future work we would like to drop this assumption and learn the optimum k given the position in feature space. This assumption however has worked reasonably well for the examples shown here.

Now that we have defined how many neighbours define our neighbourhood we can use these nearest neighbours we then attempt to directly estimate proxies for the Epistemic and Aleatoric uncertainty for predictions in that neighbourhood. The epistemic uncertainty can be approximated by the distance to these k closest neighbours, the intuition here is that if relative to other points if a given point is not similar to it's k closest neighbours then the model has not been trained on points like this and therefore we cannot be confident in the prediction. We then estimate the aleatoric uncertainty using the accuracy on the k points, the accuracy on the closest points should then gives us an idea of how likely our model is to be on points similar to the one we want to predict.

Once we have estimated these quantities we define a simple parametric function of them and calibrate the function so that our confidence estimates approximate the empirical accuracy. By modelling these two effects directly, confidence estimates produced by MACE are able to encapsulate the local variance accurately whilst also being aware of when the model is being asked to predict a point that is very different to those it has been trained on. This makes MACE robust to problems such as overconfident extrapolations and bad out of sample predictions. We now describe the algorithm in more detail.

a Algorithm

MACE requires that a given dataset is split into four distinct sets: the standard labelled training set for the point prediction model, the point prediction model then generates the underlying prediction as usual in a machine learning pipeline, MACE then works on top of this as an additional step to work out how confident we are that our model is correct. We then need a set of labelled data from which MACE finds

the nearest neighbours for any given point;[1]and a set of labelled data to optimise the MACE parameters (see equation 7.10 below) by minimising the expected calibration error. Finally we have the unseen hold-out data used to test the model, one can use the same test data for both MACE and the point prediction model.

Next, one must define a notion of similarity, this is equivalent to defining a sensible distance measure between points. Often choices such as a Euclidean distance will be sufficient if the co-ordinates are sensible, i.e. not highly correlated, however any measure of similarity may be suitable depending upon the data one is trying to model. For difficult problems one can generally improve the performance of MACE by exploiting techniques such as PCA, Word2Vec, transformers, etc ([292, 293, 294, 295]) to embed the data into a space where the metric better reflects the similarity between points. In a standard pipeline this would be a modular step, i.e. the downstream algorithm works with any definition of similarity so one can swap in any embedding technique depending upon the problem. The calibration of MACE may then be better or worse depending upon the choice of metric.

Once we have defined a notion of similarity, for any given point $\mathbf{x}_*$, we find the k *most similar* points from the known MACE training data. Finding the k nearest neighbours exactly has $O(n^2)$ complexity which is prohibitively expensive for many situations. Finding Approximate Nearest Neighbours (ANN) can however be considerably less expensive and is an active area of research. For ANN search in this work we use the Hierarchical Navigable Small World algorithm [296], this leads to a complexity of $O(n \log n)$ during training in order to build the graph and then $O(\log n)$ at inference once the graph has been built. Due to the simplicity of MACE these are usually the dominant costs for both training and inference.

We note here that the nearest neighbour search is also a modular step. MACE, in principle, does not rely on any particular nearest neighbour algorithm so other methods for example [297, 298] may be more appropriate depending on the application. See [299] for a comparison of approximate nearest neighbour methods.

Once we have found the k nearest neighbours we use these k points to estimate proxies for both epistemic and aleatoric uncertainty. The local aleatoric uncertainty will be the (distance weighted) error, i.e. the number of incorrect predictions on these k points each weighted by the distance from the point where we would like to predict. The epistemic uncertainty will be approximated by the average (weighted by the rank order of closeness) distance to these k neighbours. We define the error as $\epsilon(\mathbf{x_k})$ and the average distance to the set of neighbours as $\mathcal{K}(\mathbf{x}_*, \mathbf{x_k})$. The unnormalised confidence score is then approximated to be a simple function of these terms as follows:

$$\sigma_* = \alpha\epsilon(\mathbf{x}_*, \mathbf{x_k}) + \beta\mathcal{K}(\mathbf{x}_*, \mathbf{x_k}) \tag{7.10}$$

[1]One could use the point prediction training data for this however this may mean that there is a bias when estimating the local prediction accuracy due to information leakage when training the point prediction model. We therefore use an extra split in our dataset to avoid this

Where we define $\epsilon(\mathbf{x_k})$ and $\mathcal{K}(\mathbf{x}_*, \mathbf{x_k})$ as:

$$\mathcal{K}(\mathbf{x}_*, \mathbf{x_k}) = \sum_{i=i}^{k} \frac{d(\mathbf{x}_*, \mathbf{x_i})}{i} \tag{7.11}$$

And,

$$\epsilon(\mathbf{x}_*, \mathbf{x_k}) = \sum_{i=i}^{k} \frac{\delta_{\hat{y}_i, y_i}}{d(\mathbf{x}_*, \mathbf{x_i})} \tag{7.12}$$

In both definitions, we assume that the k points are ordered by their relative distances to x_*, and we use the Euclidean distance[2] as our distance function:

$$d(\mathbf{p}, \mathbf{q}) = \sqrt{\sum_{j=1}^{D} (p_j - q_j)^2} \tag{7.13}$$

We can then see that if points are very similar to the training set then $\beta\mathcal{K}(\mathbf{x}_*, \mathbf{x_k}) \ll \alpha\epsilon(\mathbf{x_k})$ and the uncertainty will be dominated by the local aleatoric uncertainty but if points are very distant then $\beta\mathcal{K}(\mathbf{x}_*, \mathbf{x_k}) \gg \alpha\epsilon(\mathbf{x_k})$ and the epistemic uncertainty will dominate. The coefficients α and β can then be interpreted by the relative importance of both factors and can be learnt during the MACE training phase.

For the classification problem we repeat these steps for each class and then have a σ_* for each class. These scores are then normalised first by dividing by the average score across all classes for a given point: this can be interpreted as returning a score for each class which is relative to the average uncertainty across all classes. We then apply a negative softmax normalisation to these relative scores in order to return probabilities within the interval $[0, 1]$.

Note that for the purposes of this work we are only concerned with estimating the probability that the point prediction produced by the original model is correct. We therefore we minimise the expected calibration error for the predicted class only when training. This need not necessarily be the largest confidence according to MACE because the point prediction model and MACE are calculating the relative likelihood of each class independently. We find empirically that this rarely happens but in cases where they clearly disagree this suggests that either a point is particularly unusual and therefore the point prediction algorithm is likely to perform badly on them or the distance metric used by MACE may not be optimal.

In principle this algorithm could of course be extended to be made more flexible by adding non-linearities or including higher order effects for each factor however, in principle, we found that a simple linear model was generally sufficient for the experiments shown below.

[2]The euclidean distance may be calculated in an embedded space rather than the euclidean distance between the data points themselves

7.7 Results

We now evaluate MACE with three different experiments: in each case we will compare to some standard calibration methods. The first experiment looks at the performance of MACE compared to other methods on several standard UCI datasets. The models will be compared on a hold out test dataset where this test set will be unseen but from the same dataset which the models were trained on: therefore this experiment will mostly focus on how the models deal with Aleatoric uncertainty. The second experiment simulates a situation where data changes after a point prediction model has been trained: here, standard confidence estimators may fail and highlights a key difference between the way MACE and other methods deal with epistemic uncertainty. Finally we repeat the experiment shown in section 7.5 with MACE to show that MACE is generally robust to problems caused by predicting on out-of-sample data, we also show that the relationship between Trust scores and MACE remains highly correlated with both in-sample and out-of-sample data.

a Aleatoric Experiments

For this set of experiments we used seven datasets from the UCI collection ([300]). As discussed above, the choice of similarity metric will influence the results so for all datasets we first performed a simple embedding strategy: Principal Component Analysis for non-text datasets and TF-IDf followed by Neighbourhood Component analysis [301] for text datasets. Neighbourhood Component Analysis seeks to project the data into a lower dimensional surface which maximises the performance of a K-nearest neighbours classifier. This allows one to reduce the dimensionality between the data whilst still maintaining an effective distance measure between points. Once we had embedded the data we defined similarity to be the Euclidean distance in this space. We then trained a Random Forest classifier [274] and calibrated the Random Forest uncertainty using each of MACE, Platt, Isotonic, Temperature scaling and Dirichlet Calibration methods. Finally we computed the ECE, Log Loss, and Brier score using K-fold cross validation with ten folds. We then reported the 95% error intervals as twice the standard deviation over these ten folds.

Looking first at Tables 7.1 and 7.2, where we are evaluating the models using the proper scoring rules Negative Log Loss and Brier loss, we see that MACE records the lowest mean negative log likelihood for two datasets MNIST and EEG. For these two datasets we see a significant difference between MACE and the other models. The large difference in negative log likelihood is likely because MACE is able to reduce the confidence of *bad* predictions and therefore will have fewer high confidence predictions which are incorrect. This increases the *resolution* of the model as it is able to distinguish between very high and very low confidence predictions to a greater degree. We see that rankings are broadly consistent across the different

	MACE	Platt	Isotonic	Temperature	Dirichlet
EEG	0.147 ± 0.06	0.217 ± 0.05	0.228 ± 0.06	0.225 ± 0.06	0.221 ± 0.06
Text Sentiment	0.525 ± 0.26	0.49 ± 0.17	0.504 ± 0.26	0.520 ± 0.26	0.607 ± 0.61
Letters	0.199 ± 0.03	0.20 ± 0.02	0.187 ± 0.17	0.168 ± 0.02	0.170 ± 0.04
Mnist	0.07 ± 0.01	0.13 ± 0.01	0.125 ± 0.01	0.118 ± 0.01	0.115 ± 0.01
Fashion	0.30 ± 0.01	0.30 ± 0.01	0.28 ± 0.01	0.28 ± 0.01	0.28 ± 0.01
20 news	0.38 ± 0.08	0.29 ± 0.08	0.38 ± 0.26	0.10 ± 0.27	0.28 ± 0.14
Adult	0.43 ± 0.02	0.42 ± 0.02	0.40 ± 0.02	0.49 ± 0.07	0.49 ± 0.07

Table 7.1: Negative log likelihoods calculated for each of the benchmark UCI datasets, error bars are estimated using K-fold validation with 10 folds.

metrics however there are some differences. For example on the Letters dataset when looking at the mean score MACE is ranked fourth on the negative log loss but second on the Brier loss.

Looking now at table 7.3 we see similar results, this is expected as calibration is one part of the proper scoring rules. Again we see some difference between the rankings of various models on each of the datasets. MACE again is clearly the lowest on MACE and EEG with the rankings varying across other datasets. In general across each of the metrics MACE seems to perform less on the NLP tasks, this is not surprising as the induced similarity metric, i.e. the embedding strategy, used in this work is relatively simple. It is likely that by using more sophisticated text embedding methods we would be able to perform better on these datasets.

To summarise the results we find that, as has been seen previously, the differences between various methods is very problem dependent, we find that the ordering of methods changes across the experiments as well as across different metrics. Looking at the error intervals across the three metrics we see that often the distributions are overlapping meaning that it is difficult to draw statistically significant conclusions regarding a *best* model on any of the datasets. We do however see that MACE is generally competitive with other methods commonly used to tackle the problem of confidence estimation.

It is likely that for many problems each of the calibration methods generally performs well and the differences between them are small. We therefore suggest that for problems where the data is unlikely to differ considerably to the training data then each of the calibration methods will likely be *good enough* and that ultimately the choice of method should be based upon other problem-dependent factors such as computational costs, the specific objective function and whether epistemic uncertainty is likely to be a problem.

b Epistemic Experiment

Motivated by use cases where models are trained at a fixed moment in time (e.g. each morning, start of each week) and then used for some period of time after, we split the data into training and testing: a model is then trained on the former while the latter

	MACE	Platt	Isotonic	Temperature	Dirichlet
EEG	0.042 ± 0.02	0.065 ± 0.02	0.068 ± 0.02	0.068 ± 0.02	0.067 ± 0.02
Text Sentiment	0.162 ± 0.07	0.160 ± 0.07	0.156 ± 0.07	0.159 ± 0.08	0.167 ± 0.06
Letters	0.054 ± 0.01	0.055 ± 0.007	0.057 ± 0.01	0.052 ± 0.01	0.055 ± 0.01
Mnist	0.02 ± 0.003	0.037 ± 0.003	0.037 ± 0.002	0.035 ± 0.003	0.035 ± 0.003
Fashion	0.095 ± 0.005	0.09 ± 0.005	0.089 ± 0.004	0.0089 ± 0.03	0.089 ± 0.03
20 news	0.109 ± 0.02	0.083 ± 0.03	0.079 ± 0.03	0.079 ± 0.03	0.082 ± 0.03
Adult	0.13 ± 0.007	0.13 ± 0.006	0.13 ± 0.006	0.13 ± 0.005	0.13 ± 0.005

Table 7.2: Brier score calculated for each of the benchmark UCI datasets, error bars are estimated using K-fold validation with 10 folds

	MACE	Platt	Isotonic	Temperature	Dirichlet
EEG	0.017 ± 0.05	0.022 ± 0.006	0.0222 ± 0.011	0.023 ± 0.010	0.24 ± 0.008
Text Sentiment	0.137 ± 0.05	0.123 ± 0.04	0.104 ± 0.05	0.125 ± 0.05	0.21 ± 0.03
Letters	0.03 ± 0.01	0.05 ± 0.01	0.04 ± 0.01	0.02 ± 0.01	0.02 ± 0.01
Mnist	0.005 ± 0.002	0.020 ± 0.003	0.016 ± 0.005	0.007 ± 0.002	0.007 ± 0.03
Fashion	0.019 ± 0.08	0.40 ± 0.005	0.018 ± 0.003	0.013 ± 0.004	0.011 ± 0.003
20 news	0.051 ± 0.02	0.068 ± 0.022	0.029 ± 0.01	0.032 ± 0.01	0.035 ± 0.02
Adult	0.045 ± 0.01	0.058 ± 0.006	0.022 ± 0.005	0.024 ± 0.06	0.026 ± 0.08

Table 7.3: ECE calculated for each of the benchmark UCI datasets, error bars are estimated using K-fold validation with 10 folds

set has noise artificially added to its features. This simulates heteroskedastic noise, i.e. replicating the situation where data often changes throughout the life cycle of a model. In these cases it is therefore important that a model is able to understand that the data is changing and adjust the confidence in predictions accordingly. We consecutively added Gaussian noise to the features in the test data, increasing the standard deviation of the noise, at each iteration. We then recorded the average point prediction accuracy and the mean confidence for each calibration model.

The results in Figure 7.7 show the average accuracy of the model and the average confidence estimate of each confidence estimator. We see that initially all models respond to the noise in a very similar way, i.e. dropping the mean confidence to follow the trend in mean prediction accuracy. As the amount of noise increases the model accuracy continues to decrease, until eventually it is scoring around 10% which corresponds to a random choice. When we compare the accuracy to the confidence estimates we see that each of the models, apart from MACE, decreases the confidence scores but are bounded at around 50% despite the continuous decrease in model performance. This is likely because the other calibration models, which have learnt a fixed transformation on the random forest confidence score rather than producing their confidence as a function of the data itself, have no way of adapting sufficiently to the extreme changes in the data and therefore have no way of anticipating such a drastic decrease in model performance.

MACE on the other hand continues to follow the trend in accuracy and generally indicates a more appropriate level of confidence relative to the performance of the model. In particular, when the noise becomes large enough, MACE is able to report that the model is returning predictions which are no better than random guesses.

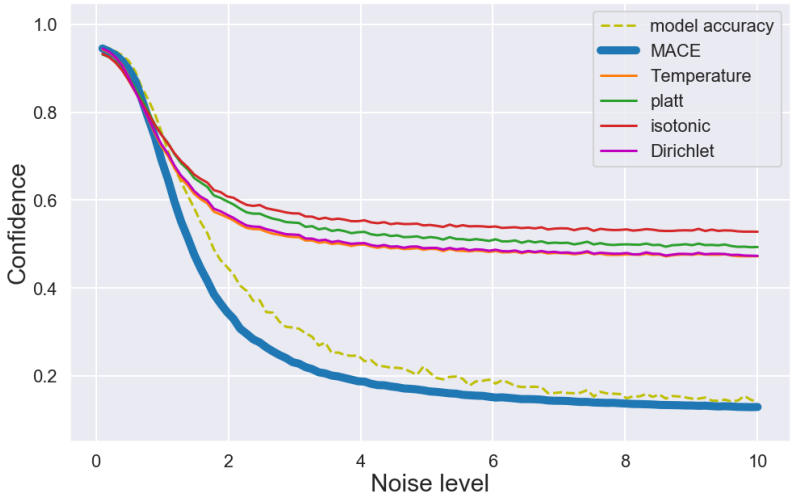

Figure 7.7: Here we add zero mean Gaussian noise with a standard deviation defined by the noise level. This simulates data drift by iteratively making the test set more different to the training data. We show the mean of the test set confidence distribution for each model and see that because MACE explicitly calculates the epistemic uncertainty it is able to track the degradation in model performance considerably better than the other calibration methods.

This is incredibly important when using models in real-world applications as often beyond the training data, it is hard to have immediate feedback regarding the outcome of any given prediction. This means that it is generally non-trivial to monitor the model performance live, so the model confidence may be the only metric available in real-time. If one were to use that metric in this example, beyond the noise level with a standard deviation of 2, then unless MACE was used, there would be no indication that the data had changed and that the model is getting less capable at making good predictions.

c Is MACE trustworthy ?

We now turn to our final critique of traditional calibration models. We have shown that often the correlation between confidence and trust scores can be broken when data is sufficiently different to the training data. This indicates that traditional calibration methods are not accounting for epistemic uncertainty and are therefore not necessarily trustworthy. We now repeat the experiment in section 7.5 with MACE and compare the behaviour between MACE and traditional calibration methods.

As shown in fig 7.8 when MACE is asked to make predictions on random noise, the distribution clearly changes relative to the calibration methods. We quantify this difference with the Kolmogorov–Smirnov (KS) statistic and find that MACE vs. Isotonic, Platt, Temperature, Dirichlet scores are 0.95 ± 0.02. This significant change is because MACE's distribution tends towards very low confidence predictions. We find that often MACE returns probabilities of around 0.1 indicating that the model

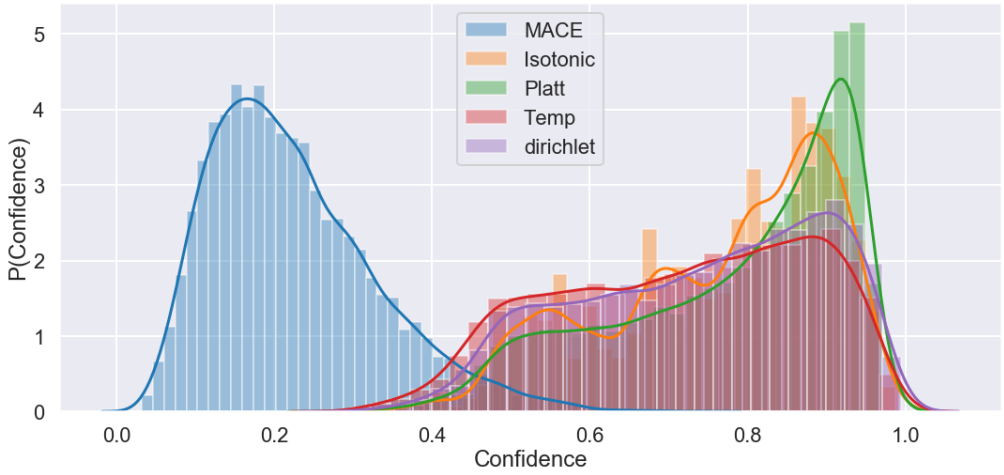

Figure 7.8: Comparison of the distribution of confidence predictions on random noise. We show the smoothed cumulative density function for each model. MACE is shown ton be significantly less confident than the other methods with many predictions returning ~ 0.1 corresponding to no clear information and effectively zero high confident predictions.

	In sample test data	**Out of sample Noise**
MACE	0.964	0.841
Isotonic	0.797	0.056
Platt	0.803	0.085
Temperature	0.814	0.108
Dirichlet	0.773	0.089

Table 7.4: Table comparing the Spearman rank correlation between trust scores and confidence estimates for MNIST data. The first column shows the correlation when estimating both quantities on an unseen test set. The second column shows the correlation when calculating each quantity on random noise

is not able to make a clear prediction.

We now compare MACE and trust scores. As expected when we are making predictions on data that is similar to the training data confidence and trust scores are generally both very high (Spearman rank correlation = 0.964). In Figure 7.9 of the joint distribution between MACE confidence estimates and trust scores when making predictions on random noise we see that MACE and trust scores remain highly correlated (correlation = 0.841). This indicates that MACE is correctly incorporating epistemic uncertainty and therefore remains trustworthy despite the data being entirely different to the data which the point prediction model was trained on. See table 7.4 for a full comparison between each calibration method.

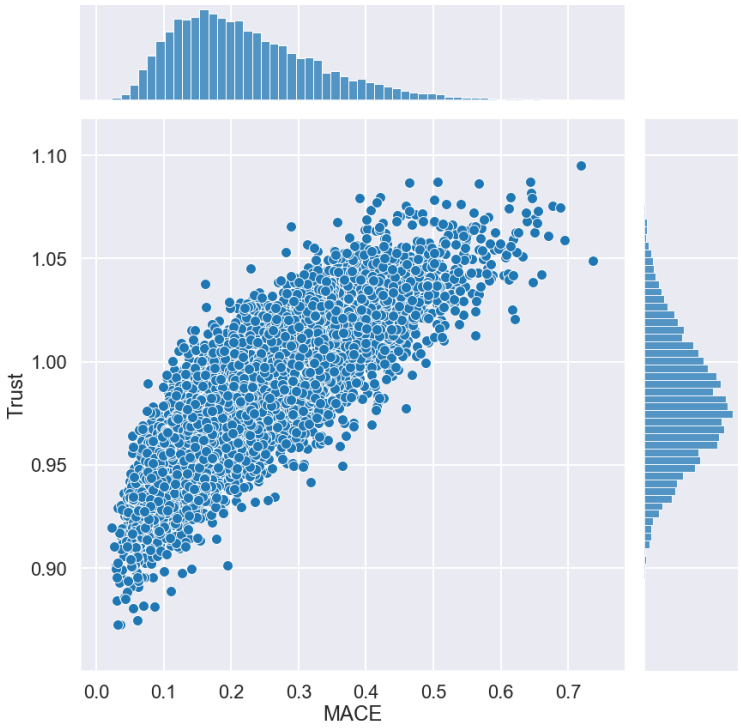

Figure 7.9: The joint distribution between confidence estimates and trust scores for MACE. We see that, unlike other methods, there is a very clear correlation between the two indicating that MACE remains trustworthy under large epistemic uncertainty.

7.8 Discussion

The results from these experiments suggest that despite MACE being an incredibly simple model, a linear sum of two derived quantities (there are likely more sophisticated versions based on this philosophy that may be better), MACE is comparable with the state of the art across a variety of metrics whilst also not suffering from the flaws highlighted in the methods based upon scaling point prediction scores. We will first address some caveats and limitations of this method before pointing to several potential important applications of MACE and looking forward to future work in this area.

a Potential Limitations

Firstly, MACE relies on the assumption that it is possible to define a notion of similarity between data points, however the distance between raw features is not necessarily a good measure of similarity and it is likely that more care will be needed in the data cleaning and feature engineering steps. This can be equivalently expressed as the need to embed the data into a sensible co-ordinate space so that a similarity measure is appropriate. Embedding into a specific distance metric area is a very active area of research (see [302] for a recent survey).

The computational costs associated with MACE are predominantly in the neighbour search, yet exact nearest neighbour search is $O(n^2)$ which is prohibitively expensive for any reasonably large dataset. Despite our use of efficient approximate nearest neighbour search the cost of building an additional nearest neighbour model may be prohibitive for some applications.

In terms of data efficiency, MACE does not have many parameters to train so does not require a large amount of data to tune them; however sufficient data is needed to train the point prediction model, as well as large graph data, and then the calibration and test sets must be sufficiently large to be useful, i.e. to induce the graph space and generate representative accuracy estimations. As a very rough rule of thumb we found that in order to have sufficient data to train both MACE and the point prediction model well, we required several thousand or more data points for the above classification tasks.

b Practical Applications

As well as the intrinsic utility of getting good confidence estimates we believe that the intuitive nature of MACE means that it has other important uses in the data science community. Firstly, the confidence estimates have a clear interpretation. If you have low confidence you either don't have enough relevant data or the model predictions in this region are noisy or error-prone. This can easily be translated into useful insights for human in the loop pipelines, i.e. seeking to understand why the model struggles to predict well and/or collect more data in certain regions.

Anomaly detection is also trivial when using MACE, any anomalous data point will by definition be very different to the training data. This will mean that MACE predicts a large uncertainty, and in particular a large epistemic uncertainty. This could easily be converted into an automated test, for which points with epistemic uncertainty larger than some threshold are automatically flagged.

Another issue that practitioners face is how to effectively monitor a model and when a model will need to be re-trained, this is related to the issue of data-drift over time. This is the notion that over time data will change and models will become less effective. Models must then be re-trained on new data, however re-training models can be costly and when to re-train is a somewhat arbitrary choice because there are generally not good metrics for *goodness of data* making this choice particularly difficult ([303]). MACE again provides a very simple solution to this problem, as shown in section b, as the data becomes more different to the initial training data the confidence estimates will on average decrease. This means there is a very clear metric to track when considering how well a model is performing. This is a better metric than predictive accuracy because the reason for confidence decline will be more clearly related to the data-drift. This could lead to simple automation rules such as re-train if average confidence falls below x%.

c Future work

There are some problems where knowledge of a single feature of many is enough to be very confident about a prediction, this concept of feature importance and feature interaction is not explicitly accounted for in MACE currently. Accounting for feature importance is a common problem in machine learning applications and is an active area of research ([304, 305]). Pre-processing and feature engineering can somewhat mitigate these problems, however in future work we plan to explicitly incorporate feature interactions and feature importance into any similarity metric.

This study has not explicitly looked at applications in deep learning, this is because MACE is a general method which is supposed to be model agnostic. Deep learning is effective when tackling problems where the raw features themselves are not necessarily good predictors, and so the network is able to induce new features. This of course means the notion of a similarity metric between two data points is more challenging.

In future work we plan to explore the applications of MACE to deep learning, in particular we plan to compare methods which will induce a similarity metric such as Auto-encoders, embeddings, etc. In principle as long as there is, or one can induce, a good measure of similarity then MACE should be effective. Combining metric learning and similarity networks [302] with MACE would make this method applicable to a wider range of problems.

The applications to deep learning are important because many of the problems highlighted above, in particular decision boundaries and extrapolation are even bigger issues for deep learning applications. The most extreme version of this being adversarial attacks ([306]). In principle MACE will offer a very robust counter-measure to adversarial attacks for the same reasons that MACE is robust to high confidence predictions on pure noise.

7.9 Conclusion

Modern Machine learning algorithms are incredibly effective at classification tasks, we argue here that despite this, these algorithms are not a good starting point for computing confidence estimates. Our position is that any confidence estimates based upon this paradigm inevitably fail to account for epistemic uncertainty which is crucial to producing reliable confidence estimates. This often leads to *over-confident* predictions on data points that are dissimilar to the data used to train the point prediction model. In this work we present a simple alternative, MACE, which is based upon a fundamentally different paradigm. MACE, using a set of nearest neighbours, estimates uncertainty locally and explicitly accounts for both epistemic and aleatoric uncertainty. MACE can be applied as an ad-hoc step in any machine learning pipeline to provide accurate and robust confidence estimates. We show that

in many situations the confidence estimates produced by MACE will be more reliable and are therefore more suitable than other methods for most practical applications.

Chapter 8

Model Agnostic Confidence Estimation - Regression

As well as applications in classification problems, machine learning techniques have proved to be an incredibly powerful tool for performing regression and forecasting using data. Similarly to the classification case, many of these tools despite their utility fail to properly (if at all) incorporate uncertainty into their predictions. As Machine learning techniques are used more in critical areas such as healthcare, infrastructure, etc the ability for a regression model to be able to indicate both the range of plausible outcomes and whether or not it has the information to make a reliable prediction will become crucial. Both of these problems can be solved by accounting for uncertainty properly.

Incorporating uncertainty into a prediction from a regression results in a *prediction interval* which along with a point prediction predicts an p% upper and lower bound for which should then contain the true value p% of the time.

As is the case of confidence estimation when doing classification, a good prediction interval must encapsulate both epistemic and aleatoric uncertainty, i.e. uncertainty due to the data previously seen by the model and the uncertainy due to the inherent variability in the target variable see figure 7.1. Understanding and accounting for uncertainty is crucial to producing a good prediction interval. Therefore prediction interval estimation is most naturally expressed in a Bayesian language where uncertainty about both the model and the data can be explicitly incorporated into a prediction. However, despite the Bayesian framework it is often prohibitively computationally expensive to be applied. It also does not naturally lend itself to being combined with large non-linear machine learning regression models, (e..g Random Forest regression).

This is because these models generally have a large number of correlated parameters, therefore the sampling techniques that we have discussed through the thesis would take too to converge to be practically useful. As well as this, even if we are able to fit a stochastic model this large, sampling from it to generate an uncertainty

about the prediction is also likely to be prohibitively slow for most applications.

Here we present a simple method (referred to as MACE-PI for creating prediction intervals which can be added ad-hoc onto *any regression pipeline* which makes a point prediction. MACE is not a Bayesian algorithm however it is motivated by the way Bayesian models account for epistemic uncertainty. In the absence of *relevant* data, Bayesian models revert to the prior indicating heuristically that model is saying the training data doesn't provide sufficient information to make this prediction.

MACE estimates Epistemic in a similar way to Gaussian Process regression by computing the *similarity* of a given point to the training data. If epistemic uncertainty is large MACE returns a relatively large prediction interval, this again indicates that the model does not have sufficient information to make a good prediction. MACE then estimates Aleatoric uncertainty by calculating a local error rate on a hold out set indicating the variance one should expect for predictions in that region. See section 8.3 for the details regarding how this is implemented.

This method, for the first time, allows for some of the benefits of Bayesian Modelling, in particular properly accounting for epistemic uncertainty, to be combined with many incredibly powerful regression algorithms which are unable to account for uncertainty. The simplicity of MACE means that uncertainty can now be easily incorporated into many machine learning prediction pipelines where it previously would not have been feasible.

This chapter first introduces some of the metrics that we will use to evaluate the quality of a prediction interval, then after reviewing similar work we present the MACE algorithm and show how it has been adapted to compute prediction intervals instead of confidence estimates for classification problems. We then apply MACE to some simple examples before comparing the prediction intervals produced by MACE and those produced by a Gaussian process on several open source datasets.

8.1 Prediction Interval Estimation

Generally a prediction interval can be defined as an interval within which one expects a future observation to fall given some previously seen data. if we have a model which outputs a point prediction μ then the prediction interval (PI) can be expressed as:

$$PI = \mu^{+\sigma_u}_{-\sigma_l} \tag{8.1}$$

where $\sigma_{u,l}$ are upper and lower errors respectively. Similarly one could define the interval in terms of the bounds i.e. $[L_i, U_i]$ where:

$$\begin{aligned} U_i &= \mu + \sigma_u \\ L_i &= \mu - \sigma_l \end{aligned} \tag{8.2}$$

If our interval is symmetric, then we have

$$PI = \mu \pm \sigma \tag{8.3}$$

Intuitively one can consider the quality of a prediction interval as being the smallest possible interval that is *calibrated*. In order to be *calibrated* we require that the true value y falls within this interval $p\%$ of the time for a specified prediction interval p i.e. a 90% interval should contain the true value 90% of the time. This is often described as a metric known as the Prediction Interval Coverage Probability (PICP). This quantity must be empirically estimated from a hold out data set, i.e. a set of data which the model has not seen during training. We then define the PICP for a given set of data as

$$PICP = \frac{1}{n}\sum_{i=1}^{n} c_i \tag{8.4}$$

Where

$$c_i = \begin{cases} 1, & \text{if } t \in [L_i, U_i] \\ 0, & \text{if } t \notin [L_i, U_i] \end{cases} \tag{8.5}$$

If the PICP matches the desired confidence level, p, for a set of prediction intervals we say that the model satisfies the PICP condition. This is not the only thing to consider however as by increasing the bounds across the entire set uniformly we are able to do this and large uniform intervals are often not particularly useful. Therefore we can characterise out intuitions and say prediction interval should be as narrow as possible given that the PICP condition can be satisfied so one may also consider the mean prediction interval width (MPIW) when considering the quality of a set of predictions.

$$MPIW = \frac{1}{n}\sum_{i=1}^{n} U_i - L_i \tag{8.6}$$

When one uses MPIW as a comparison we will ensure that the coverage probability is the same for the models being compared, this means one may have to increase or decrease the reported confidence to match the true coverage probability. In this work we will ensure compare the MPIW at a coverage probability of 90% (we could of course choose any other interval however the 90% interval is commonly reported). The MPIW effectively reports how precise a set of intervals are, if we can satisfy the PICP condition with a smaller MPIW then we are generating more finely tuned predition intervals.

8.2 Related Work

a Bayesian Methods

Quantifying the uncertainty in a prediction is most naturally considered within a Bayesian framework, parametric methods such as Bayesian parametric modelling ([288] [289]) and non-parametric methods such as Gaussian Process (GP) Regression ([278, 226]) and Bayesian Neural Networks ([251]) have proven to be effective for providing prediction intervals. These methods however suffer from some drawbacks that MACE attempts to address, generally they are considerably more computationally expensive at both training and inference time, this is potentially a major barrier for many applications. For example exact GP regression is $O(n^3)$ to train and $O(n^2$ for prediction. Sparse GPs ([307] use $m \ll n$ inducing variables to reduce these computational costs to $O(nm^2)$ and $O(mn)$ for training and prediction respectively, however this still requires considerable computational resources for many problems.

There is in general no method that the authors are aware of to explicitly combine Bayesian uncertainty estimates with an arbitrary point prediction algorithm. MACE is not a Bayesian algorithm however, as described in the introduction, it is motivated by some of the underlying Bayesian principles. It then seeks to bridge the gap by combining some of the benefits of Bayesian modelling with considerably less computational costs and, by being compatible with any point prediction algorithm, considerably more flexible.

b Non-Bayesian methods

Outside of the explicitly Bayesian framework there are several methods that have been utilised for the problem of prediction intervals. There are methods that specifically address the prediction interval problem for neural networks for example see [308]. Dropouts method introduced in [290] could potentially be applied to other methods by perturbing or dropping model parameters and generating a range of predictions, however monte carlo simulation methods such as this are generally very computationally expensive to get a good coverage of the parameter space if the number of model parameters is large. Ensemble methods similar to those introduced for neural networks in [270] could also be applied for any model however the cost of training a large enough ensemble to is often prohibitive for many applications.

Outside of the Bayesian framework MACE is most similar to Conformal Learning ([309]), here prediction intervals are generated based upon how much a new data point *conforms* to the previously seen data according to some user-defined measure of conformity. This framework often fails to account properly for local variations in model accuracy, and often outputs a nearly constant interval across the variable space. As well as this intervals are set based upon the rank with respect to the previously seen data point, this means that Conformal intervals generally do not

extrapolate well beyond the range of *conformity* seen in the training data. For example they will generally not be able to identify the areas of exceptionally large epistemic uncertainty where the point prediction model does not any data in that region.

c Confidence Calibration

The ah-hoc addition of uncertainty onto an estimate is similar to the notion of confidence calibration for a classification problem, generally here one takes an unreliable confidence estimate from a point prediction model and *calibrates* these confidence estimates with an additional modelling step after the point prediction model has been trained. MACE-PI adopts a similar philosophy here, we take a point prediction and then estimate the prediction interval with one additional modelling step. This approach has generated considerable work in the field of classification ([285, 286, 279, 277, 287]) however it has not previously been applied in the context of regression.

8.3 MACE

a Algorithm - overview

The algorithm for producing prediction intervals follows the same basic principles as the classification algorithm presented in section 7.6. In this section we will briefly again outline the algorithm and highlight areas where this algorithms differ to produce a prediction interval.

The initial steps of the algorithm are the same, one has to split the data into four distinct sets, the point prediction model training set, MACE training set, MACE parameter calibration set, unseen hold-out test set. We then must define a notion of similarity between data points, This definition is an important step that should be guided by knowledge about the data and may require an embedding step to improve performance. As well as defining a notion of similarity, the user must explicitly provide their expectations about the noise distributions, MACE assumes some distributional form for the prediction intervals, e.g. Gaussian, exponential etc. In this work for all examples, we use the euclidean distance between data points as our similarity measure and assume Gaussian errors but in principle any metric and any distribution which can be described by a mode and a scale parameter would be suitable.

Once we have defined these, for any given point x_*, we find the k *most similar* points from the known MACE training data. For computational reasons we use the Hierarchical Navigable Small World algorithm [296], an approximate nearest neighbours algorithm which leads to a complexity of $O(n \log n)$ during training in order to build a graph from the mace training data and then $O(\log n)$ at inference

once the graph has been built. Due to the simplicity of MACE these are usually the dominant costs for both training and inference.

Using these k points we can then estimate proxies for both epistemic and aleatoric uncertainty. The local aleatoric uncertainty will be the error i.e. average absolute error on these k points. The epistemic uncertainty will be approximated by the average distance to these k neighbours. We define the error as $\epsilon(x_k)$ and the distance between points as $K(x_*, x_k)$. The un-normalised confidence scores is then approximated to be a simple function of these terms i.e. :

$$\sigma_* = \alpha\epsilon(x_k) + \beta K(x_*, x_k) \tag{8.7}$$

Where α and β are then co-efficients which weight the relative importance of both factors and can be learnt by optimising the PICP for a given distribution. If points are very similar to the training set then $\beta K(x_*, x_k) \ll \alpha\epsilon(x_k)$ and the uncertainty will be dominated by the local aleatoric uncertainty, if points are very distant then $\beta K(x_*, x_k) \gg \alpha\epsilon(x_k)$ will grow large and eventually the epistemic uncertainty will dominate.

In principle this could of course be extended to be made more flexible by adding non-linearities or learning powers for each factor however in principle we found that a simple linear model sufficient to model the examples data presented in this work.

After learning our hyper-parameters which define the scale parameter σ our prediction interval for a given point x_* will be fined by the distribution:

$$y_* \sim \mathcal{D}(\mu_*, \sigma_*) \tag{8.8}$$

Where μ_* is computed independently of MACE by the model we have chosen for our point predictions and $\mathcal{D}$ is our assumed error distribution about our prediction. We can then generate our prediction interval by cutting this distribution at our desired confidence level.

8.4 Experiments

a Simple Examples

In figure 8.1 we see a simple example where the noise is heteroskedastic, i.e. after 0.5 the data becomes considerably less noisy. MACE is able to capture this locality and the prediction intervals rapidly shrink to reflect the fact that any prediction between 0.5-1 is considerable more confident than between 0 - 0.5. This reflects that the model has been able to effectively capture the local aleatoric uncertainty. At values greater than 1 we see that the MACE prediction intervals becomes more and more uncertain as we move away from the training data. This reflects the fact

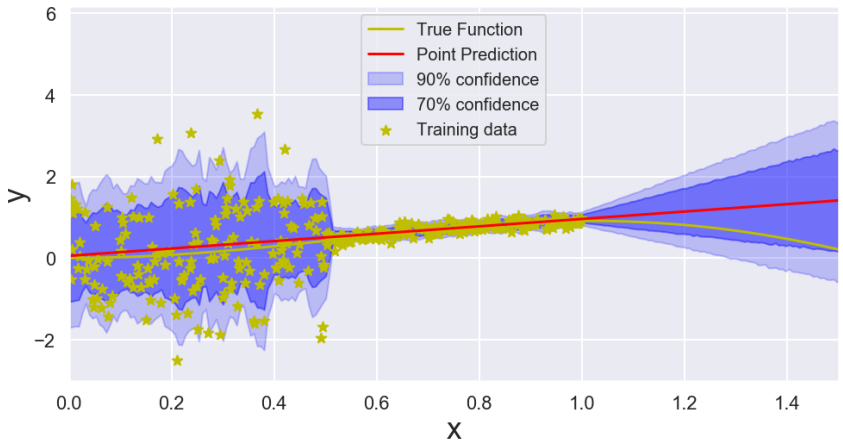

Figure 8.1: A simple example, where the model captures both local aleatoric and epistemic uncertainty

that we do not have any *relevant* data for this region and therefore any prediction contains a large amount of epistemic uncertainty.

In Fig.8.2 we look at the the California bike sharing data set (first presented in [310]) where the aim is to use this data predict the usage of shared bikes depending upon a variety of factors such as time, weather etc. To test how MACE responds to a changing model we remove variables one at a time from the data and re-train both the model and MACE-PI, we then compare the effect on the root mean squared error (RMSE), MPIW and PICP for a 90% confidence interval.

We see that as expected when we remove variables there is less explanatory power in the data and therefore the predictions become less accurate so the RMSE grows. Similarly the MPIW grows, this indicates that MACE is on average less confident in the predictions and therefore predicts a wider interval. Finally we then look at the calibration of the intervals for a 90% confidence interval, we see that this stays relatively constant until there are only a few remaining. At this point we highlight an interesting point about needing a well defined similarity metric. When we have removed so many variables, the distance between points with the remaining variables is not a good similarity metric and therefore the performance of MACE suffers. Up to this point we see the ideal behaviour as even with less information MACE is still aware of the limitations and is able to compensate for the lack of information by returning less confident predictions, i.e. wider prediction intervals. After this point the intervals become too wide suggesting that the missing information is not allowing MACE to be confident about any predictions.

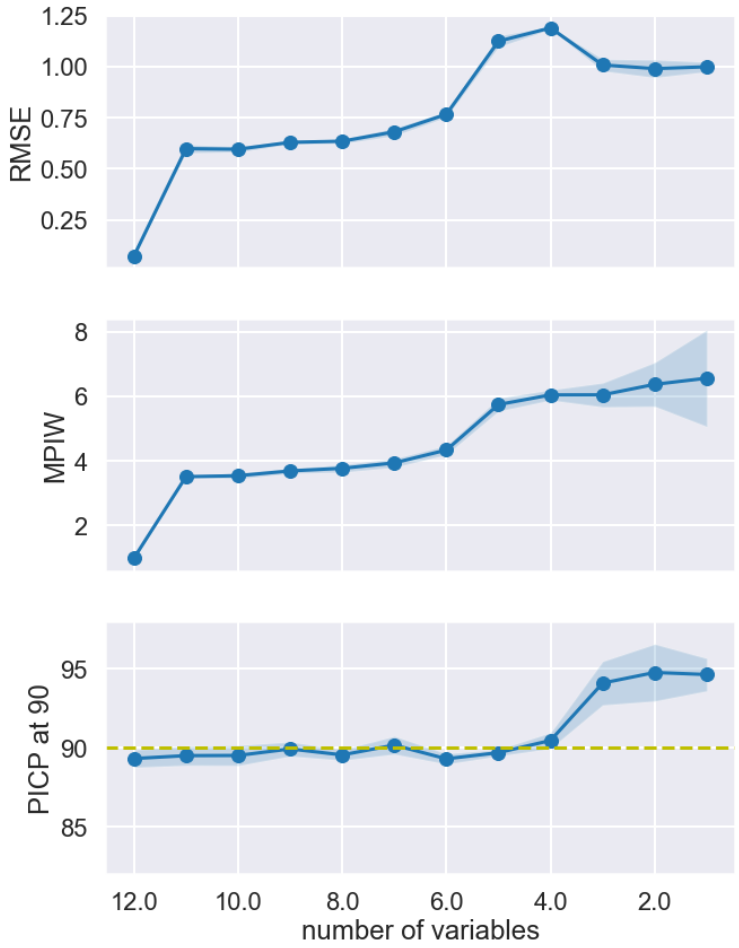

Figure 8.2: The effect of removing variables,error and width of the prediction intervals grow quickly but the calibration stays relatively constant until only a few variables remaining, at which point the noise in the data likely dominates. The uncertainty bars are obtained using k-fold cross validation for each iteration.

b Comparison with Gaussian Process Regression on benchmark datasets

Experimental setup

We will compare MACE with Gaussian Process (GP) regression. We will compare MACE and the GP using PICP, MPIW and root mean squared error (RMSE). For MACE, in all cases we will use a random forest regressor for the point prediction, implemented with Scikit learn ([274, 311]), the prediction intervals will then be learnt as described above. For MACE the RMSE is of course derived from the underlying point prediction model. By looking at this we can see examples where the random forest is more accurate that the GP and therefore highlight the utility in having the prediction interval derived independently of the method that generates the prediction. For all experiments we will use K-fold cross validation to generate a range of scores for each model. K-fold cross validation works by using k different samples of

train and test data. By evaluating our models on these different realisations this gives some estimate for the generalisation of the method. Here when stating figures we report the mean and standard error across the k folds.

For the GP will use the open source implementation from GPytorch ([312]). Due to the computational costs of using exact GP we will use Stochastic Variational Gaussian Process (SVGP) regression ([307] (though see [229] for methods to make exact GP scalabe to datasets of this size). In particular we will use Parametric Gaussian Process Regressors ([313]) as they can generally provide better prediction intervals. In all cases we will use the Radial-Basis (also known as Gaussian) kernel and optimise the GP hyper-parameters using the ADAM optimiser ([314]).

It should be noted that Gaussian Process models are generally sensitive to hyper-parameters and the choice of kernel so it may be possible to obtain results that are different to those shown here. For this reason and others such as the different training strategies, we are not necessarily doing an apples to apples comparison, we therefore will not claim that either method is superior in terms of performance. We are presenting MACE as an alternative for providing prediction intervals which can also generally produce good results as well as providing many other benefits which are discussed further in section 8.5

Results

We will evaluate our model on several datasets which can be found at the UCI database [300], these datasets vary in complexity, size ($n = O(10^2 - 10^5)$ and dimensionality (d = 2-384) (see table 8.1 for more details.). The aim for each of these datasets is to use the features provided in the data to fit a regression model and predict some target variable, e.g. the house price in the Boston dataset. For all datasets we scale the feature space such that all dimensions range between zero and one and standardise the target variable such that it has zero mean and unit standard deviation.

We will compare the models using the metrics described in 8.1; the prediction interval coverage probability (PICP) and the mean prediction interval width as well as the root mean squared error (RMSE). These three metrics evaluate the accuracy of the point prediction, the calibration of the intervals and the precision of the intervals, and therefore should provide a holistic assessment of the intervals.

We summarise these results in table 8.1 and plot the results in figures 8.3, 8.4, 8.5. In all of the figures we plot the distribution of scores, obtained by using the K folds, as a violin plot. I show each individual score as a dot in the centre of the violin. This allows us to visualise the range of scores that we might expect for each of the given datasets.

If we first look at figure 8.3, we have ordered the datasets in ascending order of the number of points. We can see that when the datasets have size $O(10^4)$ both

Table 8.1: Summary table of the comparison between a Gaussian Process and MACE-PI on our seven datasets, n is the number of data points and d is the dimensionality of the data. We report the Room Mean Squared Error, the Mean Prediciton Interval Width and the Prediction Interval Coverage Probability (see section 8.1) for each dataset.

	MACE				**GP**		
	RMSE	**PICP (at 90%)**	**MPIW (at 90%)**		**RMSE**	**PICP (at 90%)**	**MPIW (at 90%)**
Boston: n=506,d=13	0.17 ± 0.06	89.4 ± 3.3	1.27 ± 0.07		0.07 ± 0.01	81.0 ± 3.7	0.66 ± 0.02
Abalone:˜ n=4177,d=8	0.47 ± 0.05	91.8 ± 1.1	2.1 ± 0.12		0.47 ± 0.06	88.6± 2.9	2.07 ± 0.1
Wine: n=6497,d=12	0.57 ± 0.04	91.5 ± 1.3	2.6 ± 0.12		0.67 ± 0.05	88.5 ± 1.7	2.6 ± 0.1
Amsterdam: n=15148, n=16	0.60 ± 0.3	87.5 ± 1.2	1.70 ± 0.07		0.65 ± 0.4	86.9 ± 1.2	1.76 ± 0.1
Bike: n=17370, d=12	0.07 ± 0.001	89.3 ± 0.8	1.2 ± 0.3		0.16 ± 0.11	89.6± 0.7	1.03 ± 0.15
Slice: n= 53500, d=384	0.008 ± 0.002	89.6 ± 0.6	0.11 ± 0.01		0.086 ± 0.01	94.7 ± 0.4	0.26 ± 0.01
3d Road: n=434873, d=2	0.015 ± 0.0002	90.2 ± 0.5	0.29 ± 0.01		0.36± 0.007	91.9 ± 0.7	1.63 ± 0.005

models are performing similarly. The differences between models broadly fall within the errors obtained from the K-fold validation. This suggests that both models are learning a broadly similar model of the data, this changes however as the size of the dataset grows. We see that the random forest model has a considerably lower RMSE for the two largest datasets. This may be due to the sparse GP model not being complex enough to fit the data and if we used more inducing points we would see similar results.

Looking now at figure 8.4 where we are evaluating the PICP we see that generally MACE has a lower average PICP, which suggests the intervals are calibrated better to the ground truth probability, e.g. 90% confidence interval contains the truth closer to 90% of the time. For datasets $O(10^4)$ we see that the differences between models are roughly within the k fold errors. We note however that the standard deviation of the errors is much smaller for MACE suggesting that the model is more stable to small changes in the data. Again for the two largest datasets we find large differences, this suggests that MACE is able to better represent the uncertainty, again this could be due to the number of inducing points not being a complex enough model to capture the uncertainty accurately. ignoring the comparison we see that generally the PICP is very close to the target confidence level for MACE meaning that it is able to learn calibrated and reliable prediction intervals.

Finally when looking at the MPIW in 8.5, we again see that MACE performs very well relative to the GP given sufficient data, this suggests that the prediction intervals produced by MACE are *precise* in that when the model is well calibrated it is able to produce intervals that adjust well to the data where we can be very confident, or equivalently where the uncertainty about a prediction is low MACE is also able to produce restrictive prediction intervals.

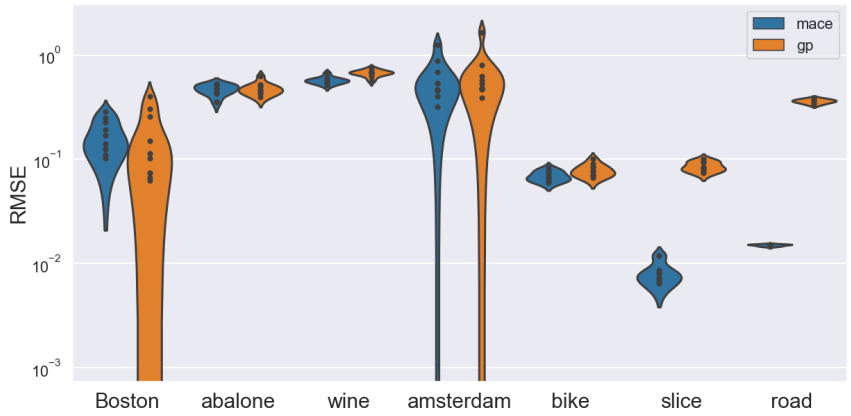

Figure 8.3: Violin plot showing the comparison of the root mean squared error for both models across the six example datasets, each dot represents the RMSE for one of the k folds, the violin plot then shows the distribution of these scores.

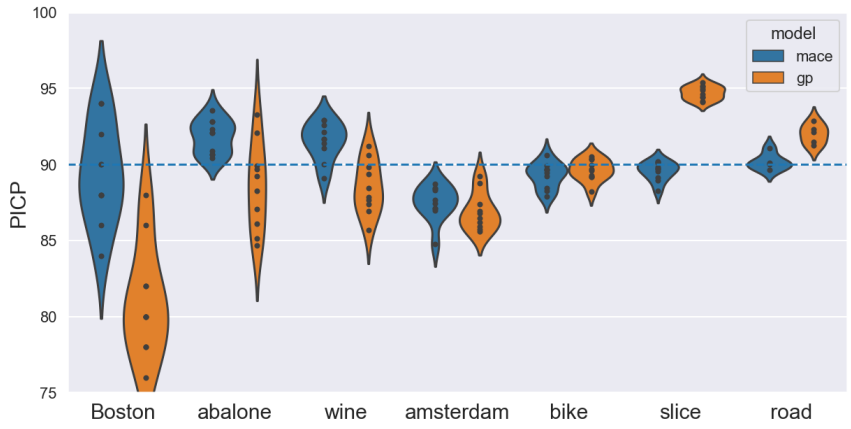

Figure 8.4: Violin plot showing the comparison of the prediction interval coverage probability (PICP) for both models across the six example datasets, each dot represents the PICP for one of the k folds, the violin plot then shows the distribution of these scores. The blue dotted line represents the coverage the models are aiming at.

8.5 Discussion

Our method differs from each of the state of the art methods in that it makes no assumptions about the model used to estimate the point prediction. It only requires a hold out dataset and a set of model predictions for that dataset. This makes it incredibly flexible and therefore simple to add onto any regression pipeline ad-hoc. We show that it is competitive on bench-marking datasets, despite this.

As well as this flexibility and ease of use, as in the case of the classification algorithm, there are several other benefits that will be added by using MACE. By directly modelling both types of uncertainty, the prediction interval becomes trivial to interpret. If the interval is large then either the model is not predicting well for data points similar to the desired prediction or the model does not have data similar

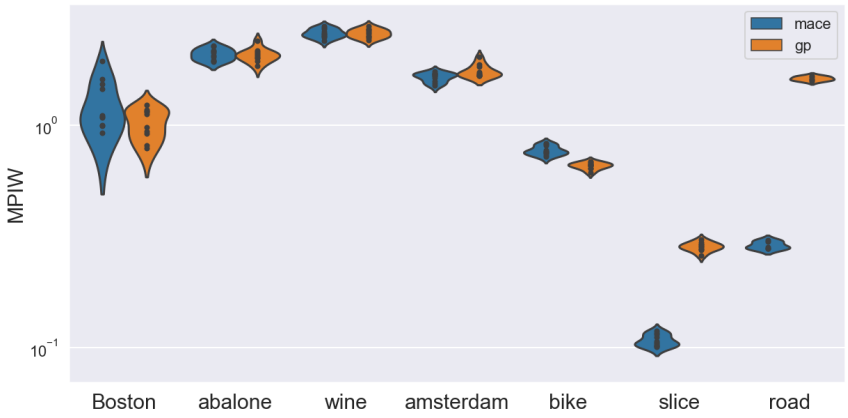

Figure 8.5: Violin plot showing the comparison of the mean prediction interval width for both models across the six example datasets, each dot represents the MPIW for one of the k folds, the violin plot then shows the distribution of these scores.

to the desired point. This can easily be translated into useful human in the loop pipelines, i.e. the user could easily seek to understand why the model struggles to predict well and/or collect more data in certain regions.

Anomaly detection is also trivial when using MACE, any anomalous data point will by definition be very different to the training data. This will mean that MACE predicts a large uncertainty, and in particular a large epistemic uncertainty. This could easily be converted into an automated test, for which points with epistemic uncertainty larger than some threshold are automatically flagged.

Another application where using MACE is that it can be useful is when considering the problem of data drift [303]). This is the notion that over time data will change and models trained at a particular point in time will become less effective. Models must then be re-trained on new data, however re-training models can be costly and when to re-train is a somewhat arbitrary choice. MACE can provide a useful metric here by noting that as data drift occurs the epistemic uncertainty will increase. One could then consider the previous n predictions and if their mean prediction width is larger than a modeller chosen threshold, this suggests that the data is sufficiently different and therefore re-training a model will be necessary. This can become an automated process which can allow models to be used for an extended period in the wild.

Here we suggest potential guidelines for the application of MACE to one's problems and highlight a couple of limitations that may determine whether it is a suitable method for the reader.

The algorithm requires splitting the dataset into four sets most machine learning pipelines pipelines involve three, training, validation and testing. These three can naturally be used for the respective chunks in MACE however there also needs to be a fairly large dataset not used in point prediction training in order to have a representative dataset to test for *similarity* . We have found that for most non-

trivial problems we need at least O(few $\times 10^3 - 10^4$) data points.

The other potential limitation is that MACE is not a black box, it requires a well defined notion of similarity between data points. This is usually a reasonable assumption however in some deep learning applications it is not trivial to define a notion of similarity that maps to reality. Strong correlations and/or many parameters that are not useful can also affect this may create an artificial similarity between points. Often the standard data pre-processing techniques will be sufficient for this to not be a problem, however there are applications where embeddings (e.g. word2vec , auto-encoders, etc [293, 294]) may improve performance substantially. We will extend this work with a specific emphasis on deep learning problems in future work.

8.6 Conclusion

This work has presented a novel method which can provide prediction intervals about any point prediction. Our method directly models the two cause of uncertainty, Aleatoric and epistemic, to produce prediction intervals. We have shown that this method is competitive with several state of the art methods whilst also being agnostic to the pipeline which produces the point prediction. As well as the utility of producing good prediction intervals for a given point we have highlighted other other situations where this method may be useful such as interpret-ability, anomaly detection, data drift, and human in the loop machine learning.

This method is a general method which is not gravitational-wave specific however many methods in gravitational-wave data analysis involve an interpolation or regression step and therefore there are several applications where applying this method would be useful. Firstly using the MACE-PI algorithm one could repeat the study shown in chapter 5 with any point prediction model (e.g. random forests) and then obtain uncertainty estimates using MACE. This would allow the method to be more flexible as it would not depend upon on the ability of the Gaussian process to fit the density surface.

Another application where this could be trivially added ad-hoc is in waveform surrogate modeling work such as [67]. This method involves doing an interpolation using neural networks, currently the neural networks only produce a point prediction and hence a single predicted waveform. One could add MACE into this pipeline and produce a range of plausible waveforms for a given set of parameters, this uncertainty could then be marginalised over during parameter estimation analysis.

Chapter 9

Concluding Remarks

This thesis was completed almost 18 months after the third observing run ended, and gravitational-wave astronomy is progressing rapidly. Since the first detection of gravitational waves [116], groundbreaking science has continued to be produced on an almost regular basis e.g. the first multi-messenger observation binary neutron stars [315] and either the heaviest neutron star or lightest black hole ever observed [155]. Looking to O4 and beyond, this progress shows no sign of slowing down [203]. In the years and decades to come we expect to observe events more frequently and using the information in these events, we expect to discover new and interesting physics.

This progress, however, as we have mentioned several times throughout this thesis, comes with many challenges. This thesis has mostly focused on the data analysis challenges that the future will bring and has looked at methods to address some of these problems. As well as addressing specific problems, this thesis has attempted to use both traditional mathematical/physics-based techniques as well as modern techniques such as machine learning to solve data analysis problems. Machine learning techniques have revolutionized many fields, such as computer vision, and will likely revolutionize many areas of the natural sciences including gravitational-wave science in the years to come however there is, of course, no free lunch. When using these methods we must ensure that they are used appropriately and where possible are guided by mathematics and physics that are well understood.

This thesis begins by looking at the gravitational waves emitted by precessing compact binary systems, precession leads to complicated orbital dynamics which imprint non-trivial changes in the phase and amplitude of the waveform we observe. We showed that by decomposing a precessing waveform into a power series of five non-precessing harmonics, the non-trivial characteristics of precessing waveforms could be easily understood as the combination and superposition of simple non-precessing harmonics. With this new formulation of the problem, it is clear that for the vast majority of signals only the first two harmonics contribute significantly to the overall waveform. This then simplifies the problem further, if we have significant

power in the sub-dominant precession then we have observed precession, otherwise, we have not. This motivated to the *precession SNR*, ρ_p which we quantifies the amount of observable precession in a waveform. This metric was subsequently used in several LVK analsyes such as [155, 1, 57]. The thesis then looks at work which used the precession SNR to conduct studies that would not have been feasible prior to its conception.

In chapter 3 a population study was done which looked at the rate at which we might expect to observe precession under different spin population models. This then allowed us to evaluate the likelihood of these models being consistent with no clear observations of precession at the time of writing, providing evidence in favor of certain models over others. We also predicted that given the most likely model of the ones we tested, there is around an 80% chance that we will observe a precessing event. This prediction turned out to potentially be correct with the marginal evidence for precession [1].

Chapter 4 conducts a large-scale parameter estimation study, looking at where in parameter space we would expect to observe precessing binaries. By moving along several dimensions independently we were able to highlight clear trends regarding the types of systems that are more likely to have observable precession in their waveforms. Using these parameter estimation runs we were also able to highlight interesting degeneracies across the parameter space for example the degeneracy between the inclination angle θ_{JN} and the precession parameter χ_p can clearly be explained as a line of constant ρ_p. Finally, we show that there is a mapping between the precession SNR and Bayes factors. Bayes Factors have been the standard measure of evidence for whether there is precession in a system however they are computationally expensive to compute relative to the precession SNR. This means that in the future we may be able to save considerable computational resources by using the precession SNR instead.

The thesis then moves on to look at machine learning applications in gravitational-wave data analysis, in chapter 5 we use Gaussian Processes to interpolate posterior samples. This gives us a continuous representation of the parameter space from a discrete set of samples, having a representation like this is considerably more useful for population analysis and other downstream applications. We also highlight a further benefit of using Gaussian Processes by incorporating the uncertainty estimates produced by them into sky-maps, this gives us an uncertainty estimate on our 90% contours which could be useful for electromagnetic follow-up.

In chapter 6, we look at how recent developments in waveform modeling using machine learning can be used to make parameter estimation faster. We implement for the first time a vectorized parameter estimation analysis, that can gain large efficiency gains by doing parameter estimation in large *batches*. We also present a method of doing gradient-based sampling where we are able to use *Autodiff* to calculate the gradient of the likelihood function. This would only previously have been

possible using approximate or numerical methods, we then highlight the weaknesses and benefits of this method compared to simple brute force random walk sampling.

In chapters 7 and 8 we move on to look at very important questions which often limit the practical use of machine learning techniques; can we trust them? And do they produce reliable uncertainty estimates? Having reliable and trustworthy uncertainty estimates is essential if these techniques are to be used in gravitational-wave astronomy. In chapter 7 section, I show that often machine learning methods do not incorporate uncertainty properly and therefore are not able to produce reliable and trustworthy predictions. I then present a novel algorithm that can be applied to classification pipelines to address this problem. This algorithm explicitly accounts for both epistemic (whether your model has the relevant data to make a good prediction) and aleatoric (the intrinsic randomness inherent in a system). I highlight use-cases where doing this is essential for the model to be useful in real-world applications. In chapter 8 I then show how to adapt this algorithm for regression models and compare the prediction intervals produced by this method with Gaussian processes.

Much of the work within it could prove to be complimentary, for example having a better understanding of precession will allow us to better understand the parameter space that we are would like to fit and sample as in chapters 5 and 6 respectively. If we could incorporate this knowledge using physics based bijectors when sampling e.g. []. These bijectors would effectively map the complicated gravitational-wave parameter space to a simple space, allowing the chains to move around a simple space and collect effective samples quickly. In other fields this has been done previously with neural networks however combining and guiding these using our physical understanding would likely prove to be even more efficient. The insights generated using the precessing SNR we would be a prime candidate for this and would be able to make precessing parameter analysis considerably more efficient. Techniques such as this which exploit our physical insights combined with the modern techniques such as HMC, vectorisation, neural networks and exploitation of GPUs could lead to huge efficiency savings relative to the current parameter estimation routines. In the conclusion of chapter 8 I point to some future applications for MACE in gravitational-wave data analysis but as machine learning models become more popular, the need for reliable uncertainty estimates will become more and more relevant so there may be many more potential applications in the years to come.

Finally I hope that this thesis presents a body of work that shows how to use both traditional and data-driven techniques to solve gravitational wave data analysis problems. I believe the right balance between these two approaches will prove to be the best way forward in the years to come.

Bibliography

[1] R. Abbott *et al.* GW190412: Observation of a Binary-Black-Hole Coalescence with Asymmetric Masses. 4 2020.

[2] B. P. Abbott *et al.* GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs. *Phys. Rev.*, X9(3):031040, 2019.

[3] Stephen Fairhurst, Rhys Green, Charlie Hoy, Mark Hannam, and Alistair Muir. Two-harmonic approximation for gravitational waveforms from precessing binaries. *Phys. Rev. D*, 102:024055, 2020.

[4] Albert Einstein. Fundamental ideas of the general theory of relativity and the application of this theory in astronomy. *Preussische Akademie der Wissenschaften, Sitzungsberichte*, 315, 1915.

[5] Henri Poincaré. Note de sur la dynamique de l'électron; note de sur la dynamique de l'électron; note on" on the dynamics of the electron". *Academie des Sciences Paris Comptes Rendus*, 150:1504–1508, 1906.

[6] Albert Einstein, Max Born, Hedwig Born, *et al.* Born-einstein letters. 1971.

[7] Albert Einstein and Nathan Rosen. On gravitational waves. *Journal of the Franklin Institute*, 223(1):43–54, 1937.

[8] Luc Blanchet. Gravitational radiation from post-newtonian sources and inspiralling compact binaries. *Living reviews in relativity*, 17(1):1–187, 2014.

[9] Luis Lehner. Numerical relativity: a review. *Classical and Quantum Gravity*, 18(17):R25, 2001.

[10] Joseph Weber. Detection and generation of gravitational waves. *Physical Review*, 117(1):306, 1960.

[11] RWP Drever, J Hough, R Bland, and GW Lessnoff. Search for short bursts of gravitational radiation. *Nature*, 246(5432):340–344, 1973.

[12] Russell A Hulse and Joseph H Taylor. Discovery of a pulsar in a binary system. *The Astrophysical Journal*, 195:L51–L53, 1975.

[13] Joseph H Taylor and Joel M Weisberg. A new test of general relativity-gravitational radiation and the binary pulsar psr 1913+ 16. *The Astrophysical Journal*, 253:908–920, 1982.

[14] ME Gerstenshtein and VI Pustovoit. On the detection of low frequency gravitational waves. *Soviet Physics-JETP*, 16(2):433–435, 1963.

[15] Rainer Weiss and Dirk Muehlner. Electronically coupled broadband gravitational antenna. *Research Laboratory of Electronics (MIT),(105)*, 54, 1972.

[16] Ronald WP Drever. Interferometric detectors for gravitational radiation. *Lecture Notes in Physics, Berlin Springer Verlag*, 124:321–338, 1983.

[17] RWP Drever, FJ Raab, KS Thorne, R Vogt, and R Weiss. A laser interferometer gravitational-wave observatory (ligo), 1989.

[18] Junaid Aasi, *et al.* Advanced ligo. *Classical and quantum gravity*, 32(7):074001, 2015.

[19] Benjamin P Abbott, *et al.* A guide to ligo–virgo detector noise and extraction of transient gravitational-wave signals. *Classical and Quantum Gravity*, 37(5):055002, 2020.

[20] Benno Willke, *et al.* The geo 600 gravitational wave detector. *Classical and Quantum Gravity*, 19(7):1377, 2002.

[21] F Acernese, *et al.* Advanced virgo: a second-generation interferometric gravitational wave detector. *Classical and Quantum Gravity*, 32(2):024001, 2014.

[22] T Akutsu, *et al.* Kagra: 2.5 generation interferometric gravitational wave detector. *arXiv preprint arXiv:1811.08079*, 2018.

[23] Karsten Danzmann, LISA Study Team, *et al.* Lisa: laser interferometer space antenna for gravitational wave measurements. *Classical and Quantum Gravity*, 13(11A):A247, 1996.

[24] M Punturo, *et al.* The einstein telescope: a third-generation gravitational wave observatory. *Classical and Quantum Gravity*, 27(19):194002, 2010.

[25] JDE Creighton and WG Anderson. Gravitational-waves physics and astronomy: An introduction to theory, experiment and data analysis, 2011. *URL http://www. wiley-vch. de/publish/dt/books/ISBN3-527-40886-X. II A*, 2018.

[26] Michele Maggiore. *Gravitational waves: Volume 1: Theory and experiments*, volume 1. Oxford university press, 2008.

[27] Lee S Finn. Detection, measurement, and gravitational radiation. *Physical Review D*, 46(12):5236, 1992.

[28] Piotr Jaranowski, Andrzej Krolak, and Bernard F Schutz. Data analysis of gravitational-wave signals from spinning neutron stars: The signal and its detection. *Physical Review D*, 58(6):063001, 1998.

[29] Alessandra Buonanno and Thibault Damour. Effective one-body approach to general relativistic two-body dynamics. *Physical Review D*, 59(8):084006, 1999.

[30] Yi Pan, *et al.* Inspiral-merger-ringdown waveforms of spinning, precessing black-hole binaries in the effective-one-body formalism. *Physical Review D*, 89(8):084006, 2014.

[31] Parameswaran Ajith, *et al.* A phenomenological template family for black-hole coalescence waveforms. *Classical and Quantum Gravity*, 24(19):S689, 2007.

[32] Sebastian Khan, Katerina Chatziioannou, Mark Hannam, and Frank Ohme. Phenomenological model for the gravitational-wave signal from precessing binary black holes with two-spin effects. *Physical Review D*, 100(2):024059, 2019.

[33] Benjamin J Owen and Bangalore Suryanarayana Sathyaprakash. Matched filtering of gravitational waves from inspiraling compact binaries: Computational cost and template placement. *Physical Review D*, 60(2):022002, 1999.

[34] Parameswaran Ajith, *et al.* Template bank for gravitational waveforms from coalescing binary black holes: Nonspinning binaries. *Physical Review D*, 77(10):104017, 2008.

[35] Derek Davis, *et al.* Ligo detector characterization in the second and third observing runs. *Classical and Quantum Gravity*, 38(13):135014, 2021.

[36] Bruce Allen, Warren G Anderson, Patrick R Brady, Duncan A Brown, and Jolien DE Creighton. Findchirp: An algorithm for detection of gravitational waves from inspiraling compact binaries. *Physical Review D*, 85(12):122006, 2012.

[37] Samantha A Usman, *et al.* The pycbc search for gravitational waves from compact binary coalescence. *Classical and Quantum Gravity*, 33(21):215004, 2016.

[38] Surabhi Sachdev, *et al.* The gstlal search analysis methods for compact binary mergers in advanced ligo's second and advanced virgo's first observing runs. *arXiv preprint arXiv:1901.08580*, 2019.

[39] Curt Cutler and Eanna E Flanagan. 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, 1994.

[40] Eric Thrane and Colm Talbot. An introduction to bayesian inference in gravitational-wave astronomy: parameter estimation, model selection, and hierarchical models. *Publications of the Astronomical Society of Australia*, 36, 2019.

[41] Ilya Mandel and Tassos Fragos. An alternative interpretation of gw190412 as a binary black hole merger with a rapidly spinning secondary. *The Astrophysical Journal Letters*, 895(2):L28, 2020.

[42] Michael Zevin, Christopher PL Berry, Scott Coughlin, Katerina Chatziioannou, and Salvatore Vitale. You can't always get what you want: The impact of prior assumptions on interpreting gw190412. *The Astrophysical Journal Letters*, 899(1):L17, 2020.

[43] Charles J Geyer. Markov chain monte carlo maximum likelihood. 1991.

[44] Steve Brooks, Andrew Gelman, Galin Jones, and Xiao-Li Meng. *Handbook of markov chain monte carlo*. CRC press, 2011.

[45] John Skilling *et al.* Nested sampling for general Bayesian computation. *Bayesian analysis*, 1(4):833–859, 2006.

[46] James R Norris and John Robert Norris. *Markov chains*. Number 2. Cambridge university press, 1998.

[47] Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.

[48] W Keith Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

[49] Michael Betancourt. A conceptual introduction to hamiltonian monte carlo. *arXiv preprint arXiv:1701.02434*, 2017.

[50] Marc van der Sluys, *et al.* Parameter estimation of spinning binary inspirals using Markov-chain Monte Carlo. *Class. Quant. Grav.*, 25:184011, 2008.

[51] J. Veitch and A. Vecchio. Bayesian coherent analysis of in-spiral gravitational wave signals with a detector network. *Phys. Rev.*, D81:062003, 2010.

[52] J. Veitch *et al.* Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library. *Phys. Rev.*, D91(4):042003, 2015.

[53] Gregory Ashton, *et al.* Bilby: A user-friendly bayesian inference library for gravitational-wave astronomy. *The Astrophysical Journal Supplement Series*, 241(2):27, 2019.

[54] C. M. Biwer, *et al.* PyCBC Inference: A Python-based parameter estimation toolkit for compact binary coalescence signals. *Publ. Astron. Soc. Pac.*, 131(996):024503, 2019.

[55] David J Earl and Michael W Deem. Parallel tempering: Theory, applications, and new perspectives. *Physical Chemistry Chemical Physics*, 7(23):3910–3916, 2005.

[56] Walter R Gilks, Sylvia Richardson, and David Spiegelhalter. *Markov chain Monte Carlo in practice*. CRC press, 1995.

[57] R. Abbott *et al.* GWTC-2: Compact Binary Coalescences Observed by LIGO and Virgo During the First Half of the Third Observing Run. 10 2020.

[58] R Abbott, *et al.* Tests of general relativity with binary black holes from the second ligo-virgo gravitational-wave transient catalog. *Physical Review D*, 103(12):122002, 2021.

[59] R. Abbott *et al.* Population Properties of Compact Objects from the Second LIGO-Virgo Gravitational-Wave Transient Catalog. 10 2020.

[60] BP Abbott, *et al.* A gravitational-wave measurement of the hubble constant following the second observing run of advanced ligo and virgo. *The Astrophysical Journal*, 909(2):218, 2021.

[61] Bangalore Suryanarayana Sathyaprakash and Bernard F Schutz. Physics, astrophysics and cosmology with gravitational waves. *Living reviews in relativity*, 12(1):1–141, 2009.

[62] Theocharis A. Apostolatos, Curt Cutler, Gerald J. Sussman, and Kip S. Thorne. Spin induced orbital precession and its modulation of the gravitational wave forms from merging binaries. *Phys. Rev.*, D49:6274–6297, 1994.

[63] Elena Cuoco, *et al.* Enhancing gravitational-wave science with machine learning. *arXiv preprint arXiv:2005.03745*, 2020.

[64] Christopher J Fluke and Colin Jacobs. Surveying the reach and maturity of machine learning and artificial intelligence in astronomy. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(2):e1349, 2020.

[65] John Jumper, *et al.* Highly accurate protein structure prediction with alphafold. *Nature*, pages 1–11, 2021.

[66] Alvin JK Chua, Chad R Galley, and Michele Vallisneri. Reduced-order modeling with artificial neurons for gravitational-wave inference. *Physical review letters*, 122(21):211101, 2019.

[67] Sebastian Khan and Rhys Green. Gravitational-wave surrogate models powered by artificial neural networks: The ann-sur for waveform generation. *arXiv preprint arXiv:2008.12932*, 2020.

[68] Radford M. Neal. MCMC using Hamiltonian dynamics. *arXiv:1206.1901 [physics, stat]*, June 2012.

[69] Andreas Griewank *et al.* On automatic differentiation. *Mathematical Programming: recent developments and applications*, 6(6):83–107, 1989.

[70] Atilim Gunes Baydin, Barak A Pearlmutter, Alexey Andreyevich Radul, and Jeffrey Mark Siskind. Automatic differentiation in machine learning: a survey. *Journal of machine learning research*, 18, 2018.

[71] Lawrence E. Kidder. Coalescing binary systems of compact objects to post-Newtonian 5/2 order. 5. Spin effects. *Phys. Rev.*, D52:821–847, 1995.

[72] T. A. Apostolatos. Search templates for gravitational waves from precessing, inspiraling binaries. *Phys. Rev.*, D52:605–620, 1995.

[73] Alessandra Buonanno, Yan-bei Chen, and Michele Vallisneri. Detecting gravitational waves from precessing binaries of spinning compact objects: Adiabatic limit. *Phys. Rev.*, D67:104025, 2003. [Erratum: Phys. Rev.D74,029904(2006)].

[74] B. P. Abbott *et al.* Binary Black Hole Population Properties Inferred from the First and Second Observing Runs of Advanced LIGO and Advanced Virgo. *Astrophys. J.*, 882(2):L24, 2019.

[75] Rory Smith, *et al.* Fast and accurate inference on gravitational waves from precessing compact binaries. *Phys. Rev.*, D94(4):044031, 2016.

[76] Eric Poisson and Clifford M. Will. Gravitational waves from inspiraling compact binaries: Parameter estimation using second postNewtonian wave forms. *Phys. Rev.*, D52:848–855, 1995.

[77] Ryan N. Lang and Scott A. Hughes. Measuring coalescing massive binary black holes with gravitational waves: The Impact of spin-induced precession. *Phys. Rev.*, D74:122001, 2006. [Erratum: Phys. Rev.D77,109901(2008)].

[78] Stephen Fairhurst. Triangulation of gravitational wave sources with a network of detectors. *New J. Phys.*, 11:123006, 2009. [Erratum: New J. Phys.13,069602(2011)].

[79] Emily Baird, Stephen Fairhurst, Mark Hannam, and Patricia Murphy. Degeneracy between mass and spin in black-hole-binary waveforms. *Phys. Rev.*, D87(2):024035, 2013.

[80] Mark Hannam, Duncan A. Brown, Stephen Fairhurst, Chris L. Fryer, and Ian W. Harry. When can gravitational-wave observations distinguish between black holes and neutron stars? *Astrophys. J.*, 766:L14, 2013.

[81] Leo P. Singer and Larry R. Price. Rapid Bayesian position reconstruction for gravitational-wave transients. *Phys. Rev.*, D93(2):024013, 2016.

[82] Michele Vallisneri. Testing general relativity with gravitational waves: a reality check. *Phys. Rev.*, D86:082001, 2012.

[83] Samantha A. Usman, Joseph C. Mills, and Stephen Fairhurst. Constraining the Inclinations of Binary Mergers from Gravitational-wave Observations. *Astrophys. J.*, 877(2):82, 2019.

[84] Mark Hannam, *et al.* Simple Model of Complete Precessing Black-Hole-Binary Gravitational Waveforms. *Phys. Rev. Lett.*, 113(15):151101, 2014.

[85] Andrea Taracchini *et al.* Effective-one-body model for black-hole binaries with generic mass ratios and spins. *Phys. Rev.*, D89(6):061502, 2014.

[86] Sebastian Khan, Frank Ohme, Katerina Chatziioannou, and Mark Hannam. Including higher order multipoles in gravitational-wave models for precessing binary black holes. *Phys. Rev.*, D101(2):024056, 2020.

[87] Vijay Varma, *et al.* Surrogate models for precessing binary black hole simulations with unequal masses. *Phys. Rev. Research.*, 1:033015, 2019.

[88] Geraint Pratten *et al.* Let's twist again: computationally efficient models for the dominant and sub-dominant harmonic modes of precessing binary black holes. *arXiv*, 4 2020.

[89] Salvatore Vitale, Ryan Lynch, John Veitch, Vivien Raymond, and Riccardo Sturani. Measuring the spin of black holes in binary systems using gravitational waves. *Phys. Rev. Lett.*, 112(25):251101, 2014.

[90] Daniele Trifirò, *et al.* Distinguishing black-hole spin-orbit resonances by their gravitational wave signatures. II: Full parameter estimation. *Phys. Rev.*, D93(4):044071, 2016.

[91] Salvatore Vitale, *et al.* Parameter estimation for heavy binary-black holes with networks of second-generation gravitational-wave detectors. *Phys. Rev. D*, 95(6):064053, 2017.

[92] R. O'Shaughnessy, *et al.* Parameter estimation of gravitational waves from precessing black hole-neutron star inspirals with higher harmonics. *Phys. Rev.*, D89(10):102005, 2014.

[93] Benjamin Farr, Evan Ochsner, Will M. Farr, and Richard O'Shaughnessy. A more effective coordinate system for parameter estimation of precessing compact binaries from gravitational waves. *Phys. Rev.*, D90(2):024018, 2014.

[94] Tyson B. Littenberg, Ben Farr, Scott Coughlin, and Vicky Kalogera. Systematic errors in low latency gravitational wave parameter estimation impact electromagnetic follow-up observations. *Astrophys. J.*, 820(1):7, 2016.

[95] Ben Farr *et al.* Parameter estimation on gravitational waves from neutron-star binaries with spinning components. *Astrophys. J.*, 825(2):116, 2016.

[96] Salvatore Vitale and Matthew Evans. Parameter estimation for binary black holes with networks of third generation gravitational-wave detectors. *Phys. Rev.*, D95(6):064052, 2017.

[97] Duncan A. Brown, Andrew Lundgren, and R. O'Shaughnessy. Nonspinning searches for spinning binaries in ground-based detector data: Amplitude and mismatch predictions in the constant precession cone approximation. *Phys. Rev.*, D86:064020, 2012.

[98] Richard O'Shaughnessy, Prakash Nepal, and Andrew Lundgren. A semianalytic fisher matrix for precessing binaries with a single significant spin. *Classical and Quantum Gravity*, 37(11):115006, 2020.

[99] A. Lundgren and R. O'Shaughnessy. Single-spin precessing gravitational waveform in closed form. *Phys. Rev.*, D89(4):044021, 2014.

[100] Ian W. Harry, *et al.* Investigating the effect of precession on searches for neutron-star-black-hole binaries with Advanced LIGO. *Phys. Rev.*, D89(2):024010, 2014.

[101] Diego Fazi. *Development of a Physical-Template Search for Gravitational Waves from Spinning Compact-Object Binaries with LIGO.* PhD thesis, Bologna U., 2009.

[102] I. W. Harry and S. Fairhurst. A coherent triggered search for single spin compact binary coalescences in gravitational wave data. *Class. Quant. Grav.*, 28:134008, 2011.

[103] Ian Harry, Stephen Privitera, Alejandro Bohé, and Alessandra Buonanno. Searching for Gravitational Waves from Compact Binaries with Precessing Spins. *Phys. Rev.*, D94(2):024012, 2016.

[104] Kostas D. Kokkotas and Bernd G. Schmidt. Quasi-normal modes of stars and black holes. *Living Reviews in Relativity*, 2(1):2, Sep 1999.

[105] Curt Cutler and Eanna E. Flanagan. Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral wave form? *Phys. Rev.*, D49:2658–2697, 1994.

[106] Patricia Schmidt, Mark Hannam, and Sascha Husa. Towards models of gravitational waveforms from generic binaries: A simple approximate mapping between precessing and non-precessing inspiral signals. *Phys. Rev.*, D86:104063, 2012.

[107] Sascha Husa, *et al.* Frequency-domain gravitational waves from nonprecessing black-hole binaries. I. New numerical waveforms and anatomy of the signal. *Phys. Rev.*, D93(4):044006, 2016.

[108] Sebastian Khan, *et al.* Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era. *Phys. Rev.*, D93(4):044007, 2016.

[109] Prayush Kumar, *et al.* Accuracy of binary black hole waveform models for aligned-spin binaries. *Phys. Rev.*, D93(10):104050, 2016.

[110] Patricia Schmidt, Frank Ohme, and Mark Hannam. Towards models of gravitational waveforms from generic binaries II: Modelling precession effects with a single effective precession parameter. *Phys. Rev.*, D91(2):024043, 2015.

[111] Stanislav Babak, Andrea Taracchini, and Alessandra Buonanno. Validating the effective-one-body model of spinning, precessing binary black holes against numerical relativity. *Phys. Rev.*, D95(2):024010, 2017.

[112] Sebastian Khan, Katerina Chatziioannou, Mark Hannam, and Frank Ohme. Phenomenological model for the gravitational-wave signal from precessing binary black holes with two-spin effects. *Phys. Rev.*, D100(2):024059, 2019.

[113] R. O'Shaughnessy, L. London, J. Healy, and D. Shoemaker. Precession during merger: Strong polarization changes are observationally accessible features of strong-field gravity during binary black hole merger. *Phys. Rev.*, D87(4):044038, 2013.

[114] Davide Gerosa, Richard O'Shaughnessy, Michael Kesden, Emanuele Berti, and Ulrich Sperhake. Distinguishing black-hole spin-orbit resonances by their gravitational-wave signatures. *Phys. Rev.*, D89(12):124025, 2014.

[115] Davide Gerosa, Michael Kesden, Ulrich Sperhake, Emanuele Berti, and Richard O'Shaughnessy. Multi-timescale analysis of phase transitions in precessing black-hole binaries. *Phys. Rev.*, D92:064016, 2015.

[116] B. P. Abbott *et al.* Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys. Rev. Lett.*, 116(6):061102, 2016.

[117] B. P. Abbott *et al.* Binary Black Hole Mergers in the first Advanced LIGO Observing Run. *Phys. Rev.*, X6(4):041015, 2016. [erratum: Phys. Rev.X8,no.3,039903(2018)].

[118] Michael Boyle, Robert Owen, and Harald P. Pfeiffer. A geometric approach to the precession of compact binaries. *Phys. Rev.*, D84:124011, 2011.

[119] Ian W. Harry and Stephen Fairhurst. A targeted coherent search for gravitational waves from compact binary coalescences. *Phys. Rev.*, D83:084002, 2011.

[120] Serge Droz, Daniel J. Knapp, Eric Poisson, and Benjamin J. Owen. Gravitational waves from inspiraling compact binaries: Validity of the stationary phase approximation to the Fourier transform. *Phys. Rev.*, D59:124016, 1999.

[121] Will M. Farr, *et al.* Distinguishing Spin-Aligned and Isotropic Black Hole Populations With Gravitational Waves. *Nature*, 548:426, 2017.

[122] Vaibhav Tiwari, Stephen Fairhurst, and Mark Hannam. Constraining black-hole spins with gravitational wave observations. *Astrophys. J.*, 868(2):140, 2018.

[123] Stephen Fairhurst, Rhys Green, Mark Hannam, and Charlie Hoy. When will we observe binary black holes precessing? *Phys. Rev. D*, 102(4):041302, 2020.

[124] Fabio Antonini, Carl L. Rodriguez, Cristobal Petrovich, and Caitlin L. Fischer. Precessional dynamics of black hole triples: binary mergers with near-zero effective spin. *Mon. Not. Roy. Astron. Soc.*, 480(1):L58–L62, 2018.

[125] Carl L. Rodriguez and Fabio Antonini. A Triple Origin for the Heavy and Low-Spin Binary Black Holes Detected by LIGO/Virgo. *Astrophys. J.*, 863(1):7, 2018.

[126] Eanna E. Flanagan and Scott A. Hughes. Measuring gravitational waves from binary black hole coalescences: 2. The Waves' information and its extraction, with and without templates. *Phys. Rev.*, D57:4566–4587, 1998.

[127] Mark A. Miller. Accuracy requirements for the calculation of gravitational waveforms from coalescing compact binaries in numerical relativity. *Phys. Rev.*, D71:104016, 2005.

[128] P. Ajith, N. Fotopoulos, S. Privitera, A. Neunzert, and A.J. Weinstein. Effectual template bank for the detection of gravitational waves from inspiralling compact binaries with generic spins. *Phys. Rev. D*, 89(8):084041, 2014.

[129] Cody Messick *et al.* Analysis Framework for the Prompt Discovery of Compact Binary Mergers in Gravitational-wave Data. *Phys. Rev.*, D95(4):042001, 2017.

[130] Samantha A. Usman *et al.* The PyCBC search for gravitational waves from compact binary coalescence. *Class. Quant. Grav.*, 33(21):215004, 2016.

[131] Benjamin J. Owen. Search templates for gravitational waves from inspiraling binaries: Choice of template spacing. *Phys. Rev.*, D53:6749–6761, 1996.

[132] Benjamin J. Owen and B. S. Sathyaprakash. Matched filtering of gravitational waves from inspiraling compact binaries: Computational cost and template placement. *Phys. Rev.*, D60:022002, 1999.

[133] S. Babak, R. Balasubramanian, D. Churches, T. Cokelaer, and B. S. Sathyaprakash. A Template bank to search for gravitational waves from inspiralling compact binaries. I. Physical models. *Class. Quant. Grav.*, 23:5477–5504, 2006.

[134] Bruce Allen, Warren G. Anderson, Patrick R. Brady, Duncan A. Brown, and Jolien D. E. Creighton. FINDCHIRP: An Algorithm for detection of gravitational waves from inspiraling compact binaries. *Phys. Rev.*, D85:122006, 2012.

[135] Bruce Allen. χ^2 time-frequency discriminator for gravitational wave detection. *Phys. Rev.*, D71:062001, 2005.

[136] S. Babak *et al.* Searching for gravitational waves from binary coalescence. *Phys. Rev.*, D87(2):024033, 2013.

[137] Tito Dal Canton and Ian W. Harry. Designing a template bank to observe compact binary coalescences in Advanced LIGO's second observing run. *arXiv e-prints*, page arXiv:1705.01845, May 2017.

[138] Debnandini Mukherjee, *et al.* The GstLAL template bank for spinning compact binary mergers in the second observation run of Advanced LIGO and Virgo. *arXiv e-prints*, page arXiv:1812.05121, December 2018.

[139] Yi Pan, Alessandra Buonanno, Yan-bei Chen, and Michele Vallisneri. A Physical template family for gravitational waves from precessing binaries of spinning compact objects: Application to single spin binaries. *Phys. Rev.*, D69:104017, 2004. [Erratum: Phys. Rev.D74,029905(2006)].

[140] Alexander H. Nitz, Thomas Dent, Tito Dal Canton, Stephen Fairhurst, and Duncan A. Brown. Detecting binary compact-object mergers with gravitational waves: Understanding and Improving the sensitivity of the PyCBC search. *Astrophys. J.*, 849(2):118, 2017.

[141] Rhys Green, *et al.* Identifying when precession can be measured in gravitational waveforms. *Physical Review D*, 103(12):124023, 2021.

[142] Colm Talbot and Eric Thrane. Determining the population properties of spinning black holes. *Phys. Rev.*, D96(2):023012, 2017.

[143] Colm Talbot, Rory Smith, Eric Thrane, and Gregory B. Poole. Parallelized Inference for Gravitational-Wave Astronomy. *Phys. Rev. D*, 100(4):043030, 2019.

[144] Daniel Wysocki, Jacob Lange, and Richard O'Shaughnessy. Reconstructing phenomenological distributions of compact binaries via gravitational wave observations. *Phys. Rev. D*, 100(4):043012, 2019.

[145] Gregory Ashton *et al.* BILBY: A user-friendly Bayesian inference library for gravitational-wave astronomy. *Astrophys. J. Suppl.*, 241(2):27, 2019.

[146] Ben Farr, Daniel E. Holz, and Will M. Farr. Using Spin to Understand the Formation of LIGO and Virgo's Black Holes. *Astrophys. J.*, 854(1):L9, 2018.

[147] Maya Fishbach and Daniel E. Holz. Where Are LIGO's Big Black Holes? *Astrophys. J. Lett.*, 851(2):L25, 2017.

[148] Maya Fishbach, Daniel E. Holz, and Will M. Farr. Does the Black Hole Merger Rate Evolve with Redshift? *Astrophys. J. Lett.*, 863(2):L41, 2018.

[149] K. S. Thorne. Multipole Expansions of Gravitational Radiation. *Rev. Mod. Phys.*, 52:299–339, 1980.

[150] P. Ajith, *et al.* Data formats for numerical relativity waves. *arXiv e-prints*, page arXiv:0709.0093, September 2007.

[151] J. Aasi *et al.* Advanced LIGO. *Class. Quant. Grav.*, 32:074001, 2015.

[152] F. Acernese *et al.* Advanced Virgo: a second-generation interferometric gravitational wave detector. *Class. Quant. Grav.*, 32(2):024001, 2015.

[153] Charlie Hoy and Vivien Raymond. PESummary: the code agnostic Parameter Estimation Summary page builder. *arXiv preprint arXiv:2006.06639*, 6 2020.

[154] B.P. Abbott *et al.* GW190425: Observation of a Compact Binary Coalescence with Total Mass $\sim 3.4 M_{\odot}$. *Astrophys. J. Lett.*, 892(1):L3, 2020.

[155] R. Abbott *et al.* GW190814: Gravitational Waves from the Coalescence of a 23 Solar Mass Black Hole with a 2.6 Solar Mass Compact Object. *Astrophys. J.*, 896(2):L44, 2020.

[156] R. Abbott *et al.* Properties and astrophysical implications of the 150 Msun binary black hole merger GW190521. *Astrophys. J. Lett.*, 900:L13, 2020.

[157] R. Abbott *et al.* GW190521: A Binary Black Hole Merger with a Total Mass of 150 $M_{\odot}$. *Phys. Rev. Lett.*, 125(10):101102, 2020.

[158] Alexander H Nitz, *et al.* 2-ogc: Open gravitational-wave catalog of binary mergers from analysis of public advanced ligo and virgo data. *The Astrophysical Journal*, 891(2):123, 2020.

[159] Tejaswi Venumadhav, Barak Zackay, Javier Roulet, Liang Dai, and Matias Zaldarriaga. New binary black hole mergers in the second observing run of advanced ligo and advanced virgo. *Physical Review D*, 101(8):083030, 2020.

[160] Barak Zackay, Tejaswi Venumadhav, Liang Dai, Javier Roulet, and Matias Zaldarriaga. Highly spinning and aligned binary black hole merger in the advanced ligo first observing run. *Physical Review D*, 100(2):023007, 2019.

[161] Barak Zackay, Liang Dai, Tejaswi Venumadhav, Javier Roulet, and Matias Zaldarriaga. Detecting gravitational waves with disparate detector responses: two new binary black hole mergers. *arXiv preprint arXiv:1910.09528*, 2019.

[162] Bernard F. Schutz. Determining the Hubble constant from gravitational wave observations. *Nature*, 323(6086):310–311, September 1986.

[163] M. Soares-Santos *et al.* First measurement of the Hubble constant from a dark standard siren using the Dark Energy Survey galaxies and the LIGO/Virgo binary-black-hole merger GW170814. *Submitted to: Astrophys. J.*, 2019.

[164] B. P. Abbott, *et al.* Gravitational Waves and Gamma-Rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A. *The Astrophysical Journal*, 848(2):L13, October 2017.

[165] Benjamin P Abbott, *et al.* Gw170817: Measurements of neutron star radii and equation of state. *Physical review letters*, 121(16):161101, 2018.

[166] Benjamin P Abbott, *et al.* Gw170817: implications for the stochastic gravitational-wave background from compact binary coalescences. *Physical review letters*, 120(9):091101, 2018.

[167] LIGO Scientific Collaboration, *et al.* A gravitational-wave standard siren measurement of the hubble constant. *Nature*, 551(7678):85–88, 2017.

[168] BP Abbott, *et al.* Tests of general relativity with the binary black hole signals from the ligo-virgo catalog gwtc-1. *Physical Review D*, 100(10):104036, 2019.

[169] B.P. Abbott *et al.* Prospects for Observing and Localizing Gravitational-Wave Transients with Advanced LIGO, Advanced Virgo and KAGRA. *Living Rev. Rel.*, 21(1):3, 2018.

[170] Yoichi Aso, *et al.* Interferometer design of the kagra gravitational wave detector. *Physical Review D*, 88(4):043007, 2013.

[171] B. S. Sathyaprakash *et al.* Cosmology and the Early Universe. *arXiv*, 2019.

[172] Eugenio Bianchi, Anuradha Gupta, Hal M. Haggard, and B. S. Sathyaprakash. Quantum gravity and black hole spin in gravitational wave observations: a test of the Bekenstein-Hawking entropy. *arXiv*, 2018.

[173] K. E. Saavik Ford, *et al.* Multi-Messenger Astrophysics Opportunities with Stellar-Mass Binary Black Hole Mergers. *arXiv*, 2019.

[174] LIGO Scientific Collaboration and Virgo Collaboration, *et al.* Properties of the Binary Black Hole Merger GW150914. *Phys. Rev. Lett.*, 116(24):241102, June 2016.

[175] Ben Farr, *et al.* Parameter estimation on gravitational waves from neutron-star binaries with spinning components. *The Astrophysical Journal*, 825(2):116, 2016.

[176] Ken KY Ng, *et al.* Gravitational-wave astrophysics with effective-spin measurements: Asymmetries and selection biases. *Physical Review D*, 98(8):083007, 2018.

[177] Philip B. Graff, Alessandra Buonanno, and B. S. Sathyaprakash. Missing Link: Bayesian detection and measurement of intermediate-mass black-hole binaries. *Phys. Rev. D*, 92(2):022002, 2015.

[178] Carl-Johan Haster, *et al.* Inference on gravitational waves from coalescences of stellar-mass compact objects and intermediate-mass black holes. *Mon. Not. Roy. Astron. Soc.*, 457(4):4499–4506, 2016.

[179] Hang Yu *et al.* Prospects for detecting gravitational waves at 5 Hz with ground-based detectors. *Phys. Rev. Lett.*, 120(14):141102, 2018.

[180] Yi Pan, *et al.* Inspiral-merger-ringdown waveforms of spinning, precessing black-hole binaries in the effective-one-body formalism. *Phys. Rev. D*, 89(8):084006, 2014.

[181] Alberto Vecchio. Lisa observations of rapidly spinning massive black hole binary systems. *Physical Review D*, 70(4):042001, 2004.

[182] Katerina Chatziioannou, *et al.* Measuring the properties of nearly extremal black holes with gravitational waves. *Phys. Rev. D*, 98(4):044028, 2018.

[183] Salvatore Vitale, Ryan Lynch, John Veitch, Vivien Raymond, and Riccardo Sturani. Measuring the spin of black holes in binary systems using gravitational waves. *Physical Review Letters*, 112(25):251101, 2014.

[184] Benjamin P Abbott, *et al.* Effects of waveform model systematics on the interpretation of gw150914. *Classical and Quantum Gravity*, 34(10):104002, 2017.

[185] Emanuele Berti, Alessandra Buonanno, and Clifford M Will. Estimating spinning binary parameters and testing alternative theories of gravity with lisa. *Physical Review D*, 71(8):084025, 2005.

[186] Geraint Pratten, Patricia Schmidt, Riccardo Buscicchio, and Lucy M. Thomas. On measuring precession in GW190814-like asymmetric compact binaries. *arXiv*, 6 2020.

[187] LIGO Scientific Collaboration and Virgo Collaboration, *et al.* GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence. *Phys. Rev. Lett.*, 116(24):241103, June 2016.

[188] Alessandra Buonanno, Yan-bei Chen, Yi Pan, and Michele Vallisneri. A Quasi-physical family of gravity-wave templates for precessing binaries of spinning compact objects. 2. Application to double-spin precessing binaries. *Phys. Rev. D*, 70:104003, 2004. [Erratum: Phys.Rev.D 74, 029902 (2006)].

[189] Bohé, Alejandro and Hannam, Mark and Husa, Sascha and Ohme, Frank and Puerrer, Michael and Schmidt, Patricia. Phenompv2 - technical notes for lal implementation. Technical Report LIGO-T1500602, LIGO Project, 2016.

[190] Benjamin P Abbott, *et al.* Gw170814: a three-detector observation of gravitational waves from a binary black hole coalescence. *Physical review letters*, 119(14):141101, 2017.

[191] A Buikema, *et al.* Sensitivity and performance of the advanced ligo detectors in the third observing run. *Physical Review D*, 102(6):062003, 2020.

[192] Benjamin P Abbott, *et al.* Prospects for observing and localizing gravitational-wave transients with advanced ligo, advanced virgo and kagra. *Living Reviews in Relativity*, 21(1):3, 2018.

[193] LIGO Scientific Collaboration. LIGO Algorithm Library, 2018.

[194] Chinmay Kalaghatgi, Mark Hannam, and Vivien Raymond. Parameter estimation with a spinning multi-mode waveform model: Imrphenomhm. *arXiv preprint arXiv:1909.10010*, 2019.

[195] Salvatore Vitale and Hsin-Yu Chen. Measuring the hubble constant with neutron star black hole mergers. *Physical review letters*, 121(2):021303, 2018.

[196] Michele Vallisneri. Use and abuse of the Fisher information matrix in the assessment of gravitational-wave parameter-estimation prospects. *Phys. Rev. D*, 77:042001, 2008.

[197] Manuela Campanelli, C.O. Lousto, and Y. Zlochower. Spinning-black-hole binaries: The orbital hang up. *Phys. Rev. D*, 74:041501, 2006.

[198] Edwin T Jaynes. *Probability theory: The logic of science*. Cambridge university press, 2003.

[199] Neil Cornish, Laura Sampson, Nicolas Yunes, and Frans Pretorius. Gravitational wave tests of general relativity with the parameterized post-einsteinian framework. *Physical Review D*, 84(6):062003, 2011.

[200] Cameron Mills and Stephen Fairhurst. Measuring gravitational-wave higher-order modes. *arXiv preprint arXiv:2007.04313*, 2020.

[201] Chinmay Kalaghatgi and Mark Hannam. Investigating the effect of in-plane spin directions for Precessing BBH systems. *arXiv*, 8 2020.

[202] Benjamin P Abbott, *et al.* Observation of gravitational waves from a binary black hole merger. *Physical review letters*, 116(6):061102, 2016.

[203] Benjamin P Abbott, *et al.* Prospects for observing and localizing gravitational-wave transients with advanced ligo, advanced virgo and kagra. *Living reviews in relativity*, 23(1):1–69, 2020.

[204] Daniel George and E. A. Huerta. Deep Learning for real-time gravitational wave detection and parameter estimation: Results with Advanced LIGO data. *Physics Letters B*, 778:64–70, Mar 2018.

[205] Hunter Gabbard, Chris Messenger, Ik Siong Heng, Francesco Tonolini, and Roderick Murray-Smith. Bayesian parameter estimation using conditional variational autoencoders for gravitational-wave astronomy. *arXiv e-prints*, page arXiv:1909.06296, Sep 2019.

[206] Stephen R Green, Christine Simpson, and Jonathan Gair. Gravitational-wave parameter estimation with autoregressive neural network flows. *Physical Review D*, 102(10):104057, 2020.

[207] XiLong Fan, Jin Li, Xin Li, YuanHong Zhong, and JunWei Cao. Applying deep neural networks to the detection and space parameter estimation of compact binary coalescence with a network of gravitational wave detectors. *SCIENCE CHINA Physics, Mechanics & Astronomy*, 62(6):969512, 2019.

[208] Christopher J. Moore, Christopher P.L. Berry, Alvin J.K. Chua, and Jonathan R. Gair. Improving gravitational-wave parameter estimation using Gaussian process regression. *Physical Review D*, 93(6):1–25, 2016.

[209] Yoshinta Eka Setyawati, Michael Pürrer, and Frank Ohme. Regression methods in waveform modeling: a comparative study. *Classical and Quantum Gravity*, 2020.

[210] Jim W. Barrett, Ilya Mandel, Coenraad J. Neijssel, Simon Stevenson, and Alejandro Vigna-Gómez. Exploring the Parameter Space of Compact Binary Population Synthesis. *Proceedings of the International Astronomical Union*, 12(S325):46–50, 2016.

[211] Stephen R. Taylor and Davide Gerosa. Mining gravitational-wave catalogs to understand binary stellar evolution: A new hierarchical Bayesian framework. *Physical Review D*, 98(8):1–19, 2018.

[212] Philippe Landry and Reed Essick. Nonparametric inference of the neutron star equation of state from gravitational wave observations. *Physical Review D*, 99(8):084049, 2019.

[213] Jacob Lange, Richard O'Shaughnessy, and Monica Rizzo. Rapid and accurate parameter inference for coalescing, precessing compact binaries. *arXiv preprint arXiv:1805.10457*, 2018.

[214] Iain Murray, David MacKay, and Ryan P Adams. The gaussian process density sampler. *Advances in Neural Information Processing Systems*, 21:9–16, 2008.

[215] George Papamakarios, Theo Pavlakou, and Iain Murray. Masked autoregressive flow for density estimation. In *Advances in Neural Information Processing Systems*, pages 2338–2347, 2017.

[216] Zhipeng Wang and David W Scott. Nonparametric density estimation for high-dimensional data—algorithms and applications. *Wiley Interdisciplinary Reviews: Computational Statistics*, 11(4):e1461, 2019.

[217] R Abbott, *et al.* Population properties of compact objects from the second ligo-virgo gravitational-wave transient catalog. *arXiv preprint arXiv:2010.14533*, 2020.

[218] Walter Del Pozzo, *et al.* Dirichlet process gaussian-mixture model: An application to localizing coalescing binary neutron stars with gravitational-wave observations. *Monthly Notices of the Royal Astronomical Society*, 479(1):601–614, 2018.

[219] Colm Talbot and Eric Thrane. Fast, flexible, and accurate evaluation of malmquist bias with machine learning: Preparing for the pending flood of gravitational-wave detections. *arXiv preprint arXiv:2012.01317*, 2020.

[220] Motonobu Kanagawa, Philipp Hennig, Dino Sejdinovic, and Bharath K Sriperumbudur. Gaussian processes and kernel methods: A review on connections and equivalences. *arXiv preprint arXiv:1807.02582*, 2018.

[221] M Pitkin, C Messenger, and X Fan. Hierarchical bayesian method for detecting continuous gravitational waves from an ensemble of pulsars. *Physical Review D*, 98(6):063001, 2018.

[222] Peter TH Pang, Tim Dietrich, Ingo Tews, and Chris Van Den Broeck. Parameter estimation for strong phase transitions in supranuclear matter using gravitational-wave astronomy. *Physical Review Research*, 2(3):033514, 2020.

[223] Ryan Lynch, Salvatore Vitale, Reed Essick, Erik Katsavounidis, and Florent Robinet. Information-theoretic approach to the gravitational-wave burst detection problem. *Physical Review D*, 95(10):104046, 2017.

[224] Matt P Wand and M Chris Jones. *Kernel smoothing*. CRC press, 1994.

[225] John Veitch, *et al.* Parameter estimation for compact binaries with ground-based gravitational-wave observations using the lalinference software library. *Physical Review D*, 91(4):042003, 2015.

[226] Christopher KI Williams and Carl Edward Rasmussen. *Gaussian processes for machine learning*, volume 2 of *1*. MIT press Cambridge, MA, 2006.

[227] Joaquin Quiñonero-Candela and Carl Edward Rasmussen. A unifying view of sparse approximate gaussian process regression. *Journal of Machine Learning Research*, 6(Dec):1939–1959, 2005.

[228] James Hensman, Nicolo Fusi, and Neil D Lawrence. Gaussian processes for big data. *arXiv preprint arXiv:1309.6835*, 2013.

[229] Ke Alexander Wang, *et al.* Exact gaussian processes on a million data points. *arXiv preprint arXiv:1903.08114*, 2019.

[230] Vinod Nair and Geoffrey E Hinton. Rectified linear units improve restricted boltzmann machines. In *Icml*, 2010.

[231] Andrew McHutchon and Carl Rasmussen. Gaussian process training with input noise. *Advances in Neural Information Processing Systems*, 24:1341–1349, 2011.

[232] Haitao Liu, Yew-Soon Ong, and Jianfei Cai. Large-scale heteroscedastic regression via gaussian process. *IEEE transactions on neural networks and learning systems*, 2020.

[233] Pauli Virtanen, *et al.* Scipy 1.0: fundamental algorithms for scientific computing in python. *Nature methods*, 17(3):261–272, 2020.

[234] Charlie Hoy and Vivien Raymond. Pesummary: the code agnostic parameter estimation summary page builder. *arXiv preprint arXiv:2006.06639*, 2020.

[235] IM Romero-Shaw, *et al.* Bayesian inference for compact binary coalescences with bilby: Validation and application to the first ligo–virgo gravitational-wave transient catalogue. *arXiv preprint arXiv:2006.00714*, 2020.

[236] Benjamin P Abbott, *et al.* Properties of the binary black hole merger gw150914. *Physical review letters*, 116(24):241102, 2016.

[237] Patricia Schmidt, Frank Ohme, and Mark Hannam. Towards models of gravitational waveforms from generic binaries: Ii. modelling precession effects with a single effective precession parameter. *Physical Review D*, 91(2):024043, 2015.

[238] Benjamin P Abbott, *et al.* Gw170817: observation of gravitational waves from a binary neutron star inspiral. *Physical Review Letters*, 119(16):161101, 2017.

[239] K Grover, *et al.* Comparison of gravitational wave detector network sky localization approximations. *Physical Review D*, 89(4):042004, 2014.

[240] B. P. Abbott, R. Abbott, T. D. Abbott, *et al.* Multi-messenger Observations of a Binary Neutron Star Merger. *apj*, 848:L12, October 2017.

[241] Marcelle Soares-Santos, *et al.* First measurement of the hubble constant from a dark standard siren using the dark energy survey galaxies and the ligo/virgo binary–black-hole merger gw170814. *The Astrophysical Journal Letters*, 876(1):L7, 2019.

[242] Andrew Gordon Wilson, Zhiting Hu, Ruslan Salakhutdinov, and Eric P Xing. Deep kernel learning. In *Artificial intelligence and statistics*, pages 370–378, 2016.

[243] Francisco Hernandez Vivanco, *et al.* Measuring the neutron star equation of state with gravitational waves: The first forty binary neutron star merger observations. *Physical Review D*, 100(10):103009, 2019.

[244] Thomas McClintock and Eduardo Rozo. Reconstructing probability distributions with gaussian processes. *Monthly Notices of the Royal Astronomical Society*, 489(3):4155–4160, 2019.

[245] Virginia d'Emilio, Rhys Green, and Vivien Raymond. Reconstructing the gravitational waves posterior probability distribution with gaussian processes. *In Preparation*, 2021.

[246] Anthony O'Hagan. Bayes–hermite quadrature. *Journal of statistical planning and inference*, 29(3):245–260, 1991.

[247] Philip Graff, Farhan Feroz, Michael P Hobson, and Anthony Lasenby. Bambi: blind accelerated multimodal bayesian inference. *Monthly Notices of the Royal Astronomical Society*, 421(1):169–180, 2012.

[248] Ben Farr, Will Farr, Douglas Rudd, Adrian Price-Whelan, and Duncan Macleod. kombine: Kernel-density-based parallel ensemble sampler. *Astrophysics Source Code Library*, pages ascl–2004, 2020.

[249] Ilya Mandel and Tassos Fragos. An alternative interpretation of gw190412 as a binary black hole merger with a rapidly spinning secondary. *arXiv preprint arXiv:2004.09288*, 2020.

[250] F. Pedregosa, *et al.* Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[251] Radford M Neal. *Bayesian learning for neural networks*, volume 118 of *1*. Springer Science & Business Media, 2012.

[252] Alexander H Nitz, *et al.* 1-ogc: The first open gravitational-wave catalog of binary mergers from analysis of public advanced ligo data. *The Astrophysical Journal*, 872(2):195, 2019.

[253] Gregory Ashton and Colm Talbot. Bilby-mcmc: An mcmc sampler for gravitational-wave inference. *arXiv preprint arXiv:2106.08730*, 2021.

[254] Edward K Porter and Jérôme Carré. A hamiltonian monte–carlo method for bayesian inference of supermassive black hole binaries. *Classical and Quantum Gravity*, 31(14):145004, 2014.

[255] Yann Bouffanais and Edward K Porter. Bayesian inference for binary neutron star inspirals using a hamiltonian monte carlo algorithm. *Physical Review D*, 100(10):104023, 2019.

[256] Hunter Gabbard, Chris Messenger, Ik Siong Heng, Francesco Tonolini, and Roderick Murray-Smith. Bayesian parameter estimation using conditional variational autoencoders for gravitational-wave astronomy. *arXiv preprint arXiv:1909.06296*, 2019.

[257] Junpeng Lao, *et al.* tfp.mcmc: Modern markov chain monte carlo tools built for modern hardware, 2020.

[258] Joshua V. Dillon, *et al.* Tensorflow distributions, 2017.

[259] Bob Carpenter, *et al. Stan* : A Probabilistic Programming Language. *Journal of Statistical Software*, 76(1), 2017.

[260] Alejandro Bohé, *et al.* Improved effective-one-body model of spinning, nonprecessing binary black holes for the era of gravitational-wave astrophysics with advanced detectors. *Physical Review D*, 95(4):044028, 2017.

[261] Martín Abadi, *et al.* Tensorflow: A system for large-scale machine learning. In *12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16)*, pages 265–283, 2016.

[262] MJ Betancourt, Simon Byrne, and Mark Girolami. Optimizing the integrator step size for hamiltonian monte carlo. *arXiv preprint arXiv:1411.6669*, 2014.

[263] Matthew D. Hoffman and Andrew Gelman. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *arXiv:1111.4246 [cs, stat]*, November 2011.

[264] Matthew Hoffman, Alexey Radul, and Pavel Sountsov. An adaptive-mcmc scheme for setting trajectory lengths in hamiltonian monte carlo. In *International Conference on Artificial Intelligence and Statistics*, pages 3907–3915. PMLR, 2021.

[265] Ravin Kumar, Colin Carroll, Ari Hartikainen, and Osvaldo Martin. Arviz a unified library for exploratory analysis of bayesian models in python. *Journal of Open Source Software*, 4(33):1143, 2019.

[266] Andrew Gelman, John B. Carlin, Hal S. Stern, and Donald B. Rubin. *Bayesian Data Analysis*. Chapman and Hall/CRC, 3rd ed. edition, 2013.

[267] Amir Hajian. Efficient cosmological parameter estimation with hamiltonian monte carlo technique. *Physical Review D*, 75(8):083525, 2007.

[268] Matthew Hoffman, *et al.* Neutra-lizing bad geometry in hamiltonian monte carlo using neural transport, 2019.

[269] Michael J Williams, John Veitch, and Chris Messenger. Nested sampling with normalizing flows for gravitational-wave inference. *Physical Review D*, 103(10):103006, 2021.

[270] Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell. Simple and scalable predictive uncertainty estimation using deep ensembles. In *Advances in neural information processing systems*, pages 6402–6413, 2017.

[271] Alex Kendall and Yarin Gal. What uncertainties do we need in bayesian deep learning for computer vision? In *Advances in neural information processing systems*, pages 5574–5584, 2017.

[272] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Information Science and Statistics. Springer, New York, 2006.

[273] Wei-Yin Loh. Fifty years of classification and regression trees. *International Statistical Review*, 82(3):329–348, 2014.

[274] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[275] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

[276] Heinrich Jiang, Been Kim, Melody Y. Guan, and Maya Gupta. To Trust Or Not To Trust A Classifier. *arXiv:1805.11783 [cs, stat]*, May 2018.

[277] Chuan Guo, Geoff Pleiss, Yu Sun, and Kilian Q. Weinberger. On Calibration of Modern Neural Networks. *arXiv:1706.04599 [cs]*, August 2017.

[278] Carl Edward Rasmussen. Gaussian processes in machine learning. In *Summer School on Machine Learning*, pages 63–71. Springer, 2003.

[279] Mahdi Pakdaman Naeini, Gregory Cooper, and Milos Hauskrecht. Obtaining well calibrated probabilities using bayesian binning. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015.

[280] Jeremy Nixon, Michael W Dusenberry, Linchuan Zhang, Ghassen Jerfel, and Dustin Tran. Measuring calibration in deep learning. In *CVPR Workshops*, pages 38–41, 2019.

[281] Tilmann Gneiting and Adrian E Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American statistical Association*, 102(477):359–378, 2007.

[282] Allan H Murphy and Robert L Winkler. Reliability of subjective probability forecasts of precipitation and temperature. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 26(1):41–47, 1977.

[283] Jerome Friedman, Trevor Hastie, and Robert Tibshirani. *The elements of statistical learning*, volume 1. Springer series in statistics New York, 2001.

[284] John Platt *et al.* Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Advances in large margin classifiers*, 10(3):61–74, 1999.

[285] Bianca Zadrozny and Charles Elkan. *Transforming Classifier Scores into Accurate Multiclass Probability Estimates*. 2002.

[286] Alexandru Niculescu-Mizil and Rich Caruana. Predicting good probabilities with supervised learning. In *Proceedings of the 22nd International Conference on Machine Learning - ICML '05*, pages 625–632, Bonn, Germany, 2005. ACM Press.

[287] Meelis Kull, *et al.* In . H. Wallach, *et al.*, editors, *Advances in Neural Information Processing Systems 32*, pages 12295–12305. Curran Associates, Inc., 2019.

[288] Michael E Tipping. Bayesian inference: An introduction to principles and practice in machine learning. In *Summer School on Machine Learning*, pages 41–62. Springer, 2003.

[289] Peter Congdon. *Applied bayesian modelling*, volume 595. John Wiley & Sons, 2014.

[290] Yarin Gal and Zoubin Ghahramani. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. *arXiv:1506.02142 [cs, stat]*, June 2015.

[291] Yann LeCun, Corinna Cortes, and CJ Burges. Mnist handwritten digit database. *ATT Labs [Online]. Available: http://yann.lecun.com/exdb/mnist*, 2, 2010.

[292] IT Jolliffe. Principal component analysis. *Technometrics*, 45(3):276, 2003.

[293] Geoffrey E Hinton and Ruslan R Salakhutdinov. Reducing the dimensionality of data with neural networks. *science*, 313(5786):504–507, 2006.

[294] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pages 3111–3119, 2013.

[295] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

[296] Yury A Malkov and Dmitry A Yashunin. Efficient and robust approximate nearest neighbor search using hierarchical navigable small world graphs. *IEEE transactions on pattern analysis and machine intelligence*, 2018.

[297] Jeff Johnson, Matthijs Douze, and Hervé Jégou. Billion-scale similarity search with gpus. *arXiv preprint arXiv:1702.08734*, 2017.

[298] Ruiqi Guo, *et al.* Accelerating large-scale inference with anisotropic vector quantization. In *International Conference on Machine Learning*, 2020.

[299] Martin Aumüller, Erik Bernhardsson, and Alexander Faithfull. Ann-benchmarks: A benchmarking tool for approximate nearest neighbor algorithms. In *International Conference on Similarity Search and Applications*, pages 34–49. Springer, 2017.

[300] Dheeru Dua and Casey Graff. Uci machine learning repository, 2017.

[301] Jacob Goldberger, Geoffrey E Hinton, Sam Roweis, and Russ R Salakhutdinov. Neighbourhood components analysis. *Advances in neural information processing systems*, 17:513–520, 2004.

[302] Mahmut Kaya and Hasan Şakir Bilge. Deep metric learning: A survey. *Symmetry*, 11(9):1066, 2019.

[303] Nithya Sambasivan, *et al.* " everyone wants to do the model work, not the data work": Data cascades in high-stakes ai. 2021.

[304] Aaron Fisher, Cynthia Rudin, and Francesca Dominici. All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously. *Journal of Machine Learning Research*, 20(177):1–81, 2019.

[305] Tongwen Huang, Zhiqi Zhang, and Junlin Zhang. Fibinet: combining feature importance and bilinear feature interaction for click-through rate prediction. In *Proceedings of the 13th ACM Conference on Recommender Systems*, pages 169–177, 2019.

[306] Christian Szegedy, *et al.* Intriguing properties of neural networks. *arXiv preprint arXiv:1312.6199*, 2013.

[307] James Hensman, Alexander Matthews, and Zoubin Ghahramani. Scalable variational gaussian process classification. *JMLR*, 2015.

[308] Abbas Khosravi, Saeid Nahavandi, Doug Creighton, and Amir F Atiya. Comprehensive review of neural network-based prediction intervals and new advances. *IEEE Transactions on neural networks*, 22(9):1341–1356, 2011.

[309] Glenn Shafer and Vladimir Vovk. A tutorial on conformal prediction. *Journal of Machine Learning Research*, 9(Mar):371–421, 2008.

[310] Hadi Fanaee-T and Joao Gama. Event labeling combining ensemble detectors and background knowledge. *Progress in Artificial Intelligence*, pages 1–15, 2013.

[311] Fabian Pedregosa, *et al.* Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830, 2011.

[312] Jacob Gardner, Geoff Pleiss, Kilian Q Weinberger, David Bindel, and Andrew G Wilson. Gpytorch: Blackbox matrix-matrix gaussian process inference with gpu acceleration. In *Advances in Neural Information Processing Systems*, pages 7576–7586, 2018.

[313] Martin Jankowiak, Geoff Pleiss, and Jacob R Gardner. Parametric gaussian process regressors. *arXiv*, pages arXiv–1910, 2019.

[314] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[315] B. P. Abbott, *et al.* GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. *Physical Review Letters*, 119(16), October 2017.