Inferring the population properties of binary black holes from unresolved gravitational waves

Rory J. E. Smith,1,2 Colm Talbot,3,1,2 Francisco Hernandez Vivanco1,2 and Eric Thrane1,2

1School of Physics and Astronomy, Monash University, Vic 3800, Australia
2OzGrav: The ARC Centre of Excellence for Gravitational Wave Discovery, Clayton VIC 3800, Australia
3LIGO, California Institute of Technology, Pasadena, CA 91125, USA

1 INTRODUCTION

Every year, around $2 \times 10^6$ binary neutron stars and $1.5 \times 10^5$ binary black holes merge somewhere in the Universe, radiating gravitational waves Abbott et al. (2018a). Only a small fraction of these signals are detected by observatories such as Advanced LIGO (aLIGO), Advanced Virgo, and KAGRA Acernese et al. (2014); Aasi et al. (2015); Akutsu et al. (2019). The rest are too faint to be resolved. Nonetheless, the ensemble of unresolved gravitational-wave signals forms an astrophysical background, which can be detected by advanced gravitational-wave detectors Abbott et al. (2018a; 2016b); Smith & Thrane (2018); Hernandez-Vivanco et al. (2019). Here, we use the word “background” to denote gravitational-wave signals that are not clearly detected and published in catalogs, e.g., Abbott et al. (2019). Since there are many connotations associated with the notion of a gravitational-wave background, it is worth pausing to make our meaning absolutely clear.

First, we note that this definition of “background” is detector-dependent; as gravitational-wave detectors become more sensitive, a greater fraction of binary mergers will be clearly resolved, and so what we might refer to as background now will become foreground in the future. Second, we note that the gravitational-wave background from compact binaries is often thought of as a foreground when looking for primordial gravitational waves from the early Universe; see, e.g., Maggiore (2000). Indeed, one scientist’s foreground is another’s background; here we use the word “foreground” to refer to resolved binaries. Finally, there is a common notion that the gravitational-wave background consists of a plethora of unimaginably faint sources. In reality, it derives from a continuum of binaries, ranging from the nearly-detectable to the clearly-not-detectable. Since there is no universally accepted definition of “detection,” the boundary between the resolved catalog and the unresolved background is fuzzy.

However one may choose to delineate this boundary, the background encodes rich information about the mass and spin distributions of compact binaries. These distributions, in turn, provide insights into binary evolution Stevenson et al. (2015, 2017); Vitale et al. (2017); Talbot & Thrane (2017a); Gerosa & Berti (2017); Farr et al. (2017); Wysocki et al. (2018); Lower et al. (2018), star formation history, the fate of massive stars Fishbach & Holz (2017); Talbot & Thrane (2018); Abott et al. (2018a), the behavior of matter at supranuclear densities Abbott et al. (2018b), and the existence of primordial black holes Raidal et al. (2017), amongst other things. Crucially, the foreground probes only the closest binaries. By analyzing the foreground and background together, it is possible to probe the entire population of binary mergers.

Here, we use hierarchical inference1 to extend the method outlined in Smith & Thrane (2018) in order to determine the ensemble properties of compact binaries. By eliminating the artificial distinction between foreground and background, we probe greater distances than possible with resolved events alone, while eliminating bias from selection effects. We demonstrate that it is possible to make population inferences even when excluding statistically significant, “gold-plated” detections. The key results are posterior probability distributions describing the shape of the binary black hole mass and spin distributions, derived using entirely unresolved events. We show that these posteriors are consistent with the true values used for the generation of the simulated data. We argue that this method is statistically optimal in the sense that it is not possible to obtain more narrow posteriors given a fixed dataset.

This work builds on Gaebel et al. (2018), which describes how population studies can be extended to include sub-threshold candi-

1 For a review of hierarchical inference in gravitational-wave astronomy, see Section V of Thrane & Talbot (2018).
date events, some of which are bona fide gravitational-wave signals, even though any single candidate is probably a noise fluctuation. This is part of a broader trend in gravitational-wave astronomy. For example, the arguably marginal event GW170729 was included \(^2\) in the first gravitational-wave transient catalog GWTC-1 {Abott et al. (2018b)} and the companion paper {Abott et al. (2018a)}.

We highlight a few innovations unique to this work. First, we eliminate selection effects entirely by making no distinction between detected events and sub-threshold events. Taking into account selection effects in population studies can be a somewhat subtle endeavour {Thrane (2018)}; {Mandel et al. (2018)}, involving challenging efficiency calculations {Ng et al. (2018)}; {Tiwari et al. (2018)}. These challenges are removed by eliminating the concept of a detection threshold. Second, by eliminating the minimum detection threshold entirely, we extend the range of the analysis to include events at large redshifts, well beyond what can be probed with unambiguous detections. This is an important first step toward studying the evolution of binary populations over cosmic time, though, more work is required to measure this redshift-dependence using hyper-parameters; see {Fishbach et al. (2018)}. Third, while {Gaebel et al. (2018)} generates pseudo posterior samples from a Fisher matrix approximation for the likelihood function, we calculate posterior samples using a full-fledged parameter estimation pipeline. By carrying out full parameter estimation (the main computational cost of the search), we show that our method is computationally feasible.

The remainder of this paper is organized as follows. In Section 2 we describe astrophysically motivated models of the binary black hole mass spectrum and spin distributions. In Section 3, we describe the method for population inference from a population of sub-threshold signals. In Section 4, we present the results of our Monte Carlo study. Concluding remarks are provided in Section 6.

### 2 POPULATION MODEL

We parameterize the mass and spin distributions using one of the prescriptions from {Abott et al. (2018a)}. In this section, we briefly summarize our population model. The reader is referred to the appendix for more details. Our models take the form of conditional priors \(p(\theta | \Lambda)\) where \(\theta\) are binary black hole parameters and \(\Lambda\) are hyper-parameters governing the shape of the \(\theta\) distribution. A list of hyper-parameters, their meaning, and injection values used in this study is provided in Tab. 1.

We model the black hole mass spectrum following {Talbot & Thrane (2018)}. The distribution is a mixture model of a truncated power-law and a Gaussian. An example of the source-frame primary mass distribution is shown in orange in Fig. 1 and the lab-frame distribution (distorted by cosmological redshift) is shown in blue. We model the distribution of black hole spin magnitudes following {Wysocki et al. (2018)}. The distribution is a beta distribution. We model the distribution of black hole spin orientations following {Talbot & Thrane (2017b)}. The distribution is a mixture model of an isotropic distribution and model with a preference for aligned spin. For this study, we choose a set of plausible population parameters based on {Abott et al. (2018a)}.

We assume a fixed, known redshift distribution of (or equivalently, luminosity distance). We assume that sources are uniformly distributed in co-moving volume to a maximum luminosity distance of \(d_L^{\max} \approx 5 \text{ Gpc} \) (redshift \(z = 0.8\)). Throughout, we assume the standard \(\Lambda\)CDM cosmology (\(\Omega_\Lambda = 0.69, \Omega_m = 0.31, H_0 = 7.7 \text{ km Mpc}^{-1}\text{s}^{-1}\)) {Ade et al. (2016)}. While this distance distribution ignores effects arising from the time-dependent star-formation rate, see {Fishbach et al. (2018)}; {You et al. (2020)}, it is satisfactory for our present purposes. By probing redshifts up to \(z = 0.8\) (lookback time = 7 Gyr), it is in-principle possible to glean information about a time when the Universe was younger and the star formation rate was higher {Madau & Dickinson (2014)}. In Fig. 2a and Fig. 2b we show the explicit redshift and luminosity distributions implied by our uniform-in-comoving volume source distribution with standard \(\Lambda\)CDM cosmology.

The final ingredient required to characterize our population of binary black holes is the duty cycle \(\xi\), the fraction of segments containing a binary black hole signal. In the next section, we describe how the data are divided into 16 s segments. Current observations of binary black hole mergers suggest that two black holes merge somewhere in the Universe on average once every \(223^{+352}_{-115}\) s. Most of these mergers probably take place at redshifts of \(z < 2\) (\(d_L \lesssim 15 \text{ Gpc}\)). Beyond \(z \approx 2\), it is believed that star-formation rate decreases {Madau & Dickinson (2014)}. With fewer stars, there are fewer black holes, and therefore fewer mergers. Assuming an average time between binary black hole of 100 s out to \(d_L = 15 \text{ Gpc}\), the duty cycle out to luminosity distances of 5 Gpc is approximately \(\xi = 6.67 \times 10^{-3}\), and so we use this value for our injection study.

### 3 INFERENCES FROM THE GRAVITATIONAL-WAVE BACKGROUND

#### 3.1 Overview

This section describes the statistical formalism that allows us to calculate the hyper-posterior distribution \(p(\Lambda | d)\) for population parameters \(\Lambda\) described in Section 2 given some dataset \(d\). We follow the method described in {Smith & Thrane (2018)}. The calculation is divided into the following steps.

1. We divide the data into 16 s segments. These segments are a convenient size so that any given segment is unlikely to contain more than one binary black hole signal. However, they are long enough that it is relatively unlikely for a binary black hole signal to fall on the boundary of two segments; see {Smith & Thrane (2018)}.
Following Smith & Thrane (2018), we employ a likelihood function to describe the probability of some large dataset $\tilde{d}$ given a population of binary black hole described by hyper-parameters $\xi$ (the fraction of data segments containing a signal) and $\Lambda$, which describes the shape of the binary black hole mass and spin distributions

$$
\Psi^{\text{tot}}(\tilde{d}|\Lambda, \xi) = \prod_{i} \int_{0}^{1} \xi \ L(d_i|\Lambda, \mathcal{H}_S) + (1 - \xi) Z(d_i|\mathcal{H}_N). 
$$

There is a lot to explain in this equation and the rest of this subsection is devoted to this task. The tot superscript denotes that this is the likelihood for the entire dataset $\tilde{d}$. The expression includes a product over $i$ data segments running from $i = 1$ to $n$. The term $L(d_i|\Lambda, \mathcal{H}_S)$ is the single-segment Bayesian evidence for the data $d_i$ in segment $i$ given the signal hypothesis $\mathcal{H}_S$ and hyper-parameters $\Lambda$. The term $Z(d_i|\mathcal{H}_N)$ is the single-segment noise evidence for the data $d_i$ in segment $i$. The hyper-parameter $\xi$ is often referred to as “duty cycle,” and may be converted into a rate Smith & Thrane (2018).

The single-segment noise evidence $Z(d_i|\mathcal{H}_N)$ is straightforwardly calculated for each segment using a Gaussian-noise likelihood

$$
Z(d_i|\mathcal{H}_N) = \exp \left( -\frac{1}{2} \langle d_i, d_i \rangle \right)
$$

3 We note that this is missing a normalisation factor, however, as this only depends on the PSD and not on the template, we can freely factor this out of the both the signal and noise evidences.

---

### Table 1. Hyper parameters $\Lambda_i$ of the binary black hole mass and spin population distributions.

| Hyper parameter $\Lambda_i$ | Description | Injection value |
|-----------------------------|-------------|----------------|
| $\xi$                       | Astrophysical duty cycle | $6.67 \times 10^{-3}$ |
| $m_{\text{min}}(M_\odot)$  | Minimum black hole mass | $8.68M_\odot$ |
| $m_{\text{max}}(M_\odot)$  | Maximum mass of black holes in the power law component | $39.5M_\odot$ |
| $\mu_{\text{min}}(M_\odot)$| Mean of the Gaussian component of the primary mass distribution | $33.4M_\odot$ |
| $\sigma_m$                  | Standard deviation of the Gaussian component of the primary mass distribution | $1.08M_\odot$ |
| $\lambda_m$                 | Fraction of black holes in the Gaussian component of the primary mass distribution | 0.340 |
| $\alpha_m$                  | Slope of the power law component of the primary mass distribution | 2.00 |
| $\beta_m$                   | Slope of the mass ratio distribution | -0.198 |
| $\sigma_{\text{max}}$       | Maximum spin magnitude | 1.00 |
| $\alpha_\mu$                | Spin-magnitude beta distribution slope parameter (rise) | 1.50 |
| $\beta_\mu$                 | Spin-magnitude beta distribution slope parameter (fall) | 3.50 |
| $\sigma_{\text{tilt}}$      | Standard deviation of the spin-tilt angle distribution | 1.00 |
| $\xi_{\text{tilt}}$         | Fraction of BBHs with Guassian distributed spin tilts | 0.50 |

---

![Figure 2. Prior distributions on redshift (left) and luminosity distance (right)](image-url)
Here, we employ a noise-weighted inner product

$$\langle a, b \rangle = \frac{1}{4\pi} \int \frac{\alpha(f_k) b(f_k)}{S_n(f_k)} \, df,$$

where the sum is over frequency bins $k$ with bin widths of $\Delta f$ and $S_n(f)$ is the strain noise power spectral density.

The single-segment signal likelihood $\mathcal{L}(d_i|\Lambda, H_S)$ is given by (Eq. 5) yielding:

$$\mathcal{L}(d_i|\Lambda, H_S) \approx Z(d_i|H_S) \sum_{k=1}^{n_s} \pi(\theta_{k,i}|\Lambda) \pi(\theta_{k,i}).$$

Here, $Z(d_i|H_S)$ is the Bayesian evidence for a binary black hole signal in segment $i$ calculated using some default prior for the binary black hole parameters $\theta$ denoted $\pi(\theta)$. Assuming Gaussian noise, it is given by

$$Z(d_i|H_S) = \int d\theta_i \mathcal{L}(d_i|\theta_i, H_S) \pi(\theta_i) = \int d\theta_i \exp \left( -\frac{1}{2} \left( d_i - h(\theta_i), d_i - h(\theta_i) \right) \right) \pi(\theta_i),$$

(5)

where $h(\theta)$ is the gravitational waveform, in this case, calculated \texttt{IMRPhenomPv2} approximant Hanam et al. (2014); Smith et al. (2016). The integral in Eq. 5 is calculated numerically using the Bayesian inference library, \texttt{bilby} Ashton et al. (2018) implementation of \texttt{dynesty} Speagle (2020). In addition to calculating $Z(d_i|H_S)$, \texttt{bilby} outputs a list of $n_s$ posterior samples $\{\theta_{k,i}\}$, which describe the posterior $p(\theta|d_i)$ given the default prior. It is sometimes said that the ratio of priors $\pi(\theta_{k,i}|\Lambda)/\pi(\theta_{k,i})$ in Eq. 4 serves to "reweight" the posterior samples calculated using the default prior $\pi(\theta)$ Thrane & Talbot (2018).

\subsection{3.3 The hyper-posterior}

Using the hyper likelihood defined in Eq. 1, it is straightforward to obtain the hyper-posterior for duty cycle and the other hyper-parameters $\Lambda$

$$p(\Lambda, \xi|d) \propto \frac{\psi_{\text{tot}}(\Lambda, \xi)}{Z_{\text{pop}}^\Lambda} \pi(\Lambda) \pi(\xi).$$

(6)

Here $\pi(\Lambda)$ is the hyper-parameter prior, which we take to be uniform for each hyper-parameter. The distribution $\pi(\xi)$ is the duty cycle prior. In a real analysis, one should choose a distribution, which uses a Poisson distribution to relate duty cycle to astrophysical rate; see Smith & Thrane (2018). However, for our present purposes, it is convenient to simply employ a uniform prior. The variable $Z_{\text{pop}}^\Lambda$ is the hyper-evidence. The hyper-evidence can be used to carry out model selection between different population models; see Talbot & Thrane (2017b, 2018); Stevenson et al. (2015, 2017); Abott et al. (2018a); Stevenson et al. (2017); Vitale et al. (2017); Talbot & Thrane (2017a); Gerosa & Berti (2017); Farr et al. (2017); Wysocki et al. (2018); Lower et al. (2018).

\section{Results: Demonstration with Simulated Data}

We analyze 5.5 days of simulated Advanced LIGO (aLIGO) design-sensitivity data Aasi et al. (2015) containing an ensemble of 200 simulated binary black hole signals. We divide the data into $3 \times 10^4$ sixteen-second segments. This yields a duty cycle $\xi = 200/30000 = 6.67 \times 10^{-3}$. We derive the duty cycle by first assuming an average merger range of binary black holes of 1 per 100s. We then assume that the merger rate drops significantly beyond a redshift of $z \sim 2$ so that their contribution can be effectively ignored. The fraction of all binaries contained in the volume with maximum redshift considered here, $z = 0.8$, is approximately $4\%$. The average merger rate out to $z = 0.8$ is then approximately one merger per 45min. In 5.5 days this yields 176 binary mergers, however we choose to round up to 200.

The masses and spins of the binary black hole’s are drawn from the mass and spin distributions described in Sec. 2. The remaining “extrinsic” parameters are drawn using standard distributions. All of the signals in our injection set are below the usual threshold for matched-filter network SNR: $h'_{\text{injection}} = 12$. Based on results from Smith & Thrane (2018), we expect the binary black hole background to be detectable with approximately one day of aLIGO design sensitivity data.

We estimate the signal and noise evidence $Z_S, Z_N$, and obtain posterior samples for binary black hole source parameters for every data segment. The priors, summarized in Table 2, and are chosen to be relatively uninformative so we can recycle the posterior samples later. We then use the sets of evidence and posterior samples as input to Eq. 3.3 to compute the posterior for $\Lambda$ — the population mass and spin distribution parameters — and $\xi$, the astrophysical duty cycle.

The computational cost of running full parameter estimation on $3 \times 10^4$ 16-second data segments is kept manageable by explicitly marginalizing over three parameters, which are difficult to sample: comoving distance, coalescence time, and coalescence phase; see e.g., Thrane & Talbot (2018) for the details of these marginalization schemes. By marginalizing over these parameters, we significantly decrease the convergence time, and hence run time, of computing evidences and drawing posterior samples in step 1.

We find that the background is detectable within one week out to comoving distances of 5 Gpc, assuming masses and spins drawn

| Parameter $\theta_i$ | Prior $\pi(\theta_i)$ |
|---------------------|---------------------|
| $m_1$               | $\text{Uniform}(6M_\odot, 50M_\odot)$ |
| $q$                 | $\text{Uniform}(0, 2)$ |
| $D_C^c$             | $\text{Uniform}(1Gpc^2, 5Gpc^2)$ |
| $t_c$               | $\text{Uniform}(0, 16)$ |
| $\cos i$            | $\text{Uniform}(-1, 1)$ |
| $\phi_e$            | $\text{Uniform}(0, 2\pi)$ |
| $\phi$              | $\text{Uniform}(0, 2\pi)$ |
| $\phi_{12}$         | $\text{Uniform}(0, 2\pi)$ |
| $a_1$               | $\text{Uniform}(0, 0.1)$ |
| $a_2$               | $\text{Uniform}(0, 0.1)$ |
| $a$                 | $\text{Uniform}(0, 2\pi) |
| $\cos \delta$       | $\text{Uniform}(-1, 1)$ |
from the distribution described in Sec. 2. The posterior distribution on $\xi$ is consistent with the true value of $\xi = 0.67\%$, and the log Bayes factor (Eq. 15 of Smith & Thrane (2018)) overwhelmingly supports a detection of a population of compact binaries: In BF $\approx 700$, confirming the previous result from Smith & Thrane (2018) with a different, more realistic population of BBH.

We find that we can begin to constrain some of the mass and spin population parameters are using the the 200 unresolved mergers in our simulated data. In Fig. 3a we show posterior predictive distributions for different mass and spin parameters. The posterior predictive distributions reflect our updated prior based on information from our hyper-priorers; see Thrane & Talbot (2018). The contours represent the $1 - \sigma$ and $2 - \sigma$ credible intervals.

In Fig. A1, we show posterior distributions for hyper-parameters associated with the duty cycle and mass parameters. In A2, we show posterior distributions for the parameters associated with the Gaussian component of the mass population model. In Fig. A3, we show posterior distributions for hyper-parameters describing black hole spins.

5 HOW SENSITIVE ARE WE TO SUBTHRESHOLD EVENTS?

In this section, we investigate where the information for our analysis comes from. Is our resolving power coming primarily from binaries just below the detection threshold, or do we gain information from weaker events as well? To address this question, we carry out a follow-up study where we introduce a new hyper-parameter, $d_{\text{max}}$, the maximum comoving distance for binary mergers. In our new population model, the rate of binary mergers drops to zero for the maximum comoving distance for binary mergers. In our updated prior based on information from our hyper-priorers; see Thrane & Talbot (2018). The contours represent the $1 - \sigma$ and $2 - \sigma$ credible intervals.

The distribution of mass ratios follows a power-law distribution with unknown spectral index $\beta$. Additionally, there is a smoothing parameter $\delta m$ which enables the distribution to have a smooth turn-on at low masses.

The first equation describes the prior probability of the primary mass $m_1$ (corresponding to the heavier of the two black holes in a binary black hole) given the hyper-parameters $\Lambda$. The second equation describes the prior probability of the mass ratio $q = m_2/m_1$ given $m_1$ and $\Lambda$.

The fraction of black holes in the Gaussian component is $\lambda_m$.

7 ACKNOWLEDGEMENTS

RS, CT, FHV, and ET are supported by the Australian Research Council (ARC) CE170100004. ET is supported by ARC FT150100281. We thank Stuart Anderson and the LIGO Data Grid for assistance with computing infrastructure, and Maya Fishbach, Thomas Callister and Thomas Dent for helpful comments and suggestions. We acknowledge the OzStar cluster for providing graphical processor units to carry out some of our calculations.

APPENDIX A: POPULATION HYPER PARAMETER ESTIMATION

The one and two dimensional PDFs for the population hyper parameters used in this study are shown below.

APPENDIX B: POPULATION MODEL DETAILS

B1 Source-frame mass

The conditional prior for binary black hole mass is:

$$
\pi_m(m_1|\Lambda) = \left[1 - \lambda_m \right] A(\Lambda) m_1^{-\sigma} \Theta(m_{\text{max}} - m_1) + \lambda_m B(\Lambda) \exp \left(-\frac{(m_1 - \mu_m)^2}{2\sigma_m^2} \right) S(m_1|m_{\text{min}}, \delta m),
$$

$$
\pi_q(q|m_1, \Lambda) = C(m_1, \Lambda) q^\beta S(m_2|m_{\text{min}}, \delta m).
$$

The distribution of mass ratios follows a power-law distribution with unknown spectral index $\beta$. Additionally, there is a smoothing parameter $\delta m$ which enables the distribution to have a smooth turn-on at low masses.

The prior for primary mass $\pi(m_1|\Lambda)$ is constructed from two
Figure 3. Posterior predictive distributions of binary black hole parameters. These results are obtained using five and a half days of simulated aLIGO data containing 395 binary black holes signals. The dashed line is the true distribution, while the red contours represent the 50\% (light) and 90\% credible intervals on the inferred distributions. The parameters are: (top) primary black hole mass $m_1$, (center) mass ratio $q$, (lower left) spin magnitude $a$, (lower right) cosine spin tilt $\cos \theta$.

The first term
\begin{equation}
(1 - \lambda_m)A(\Lambda) m_1^{-\alpha} \Theta(m_{\text{max}} - m_1), \tag{B2}
\end{equation}
describes a power-law distribution with index $\alpha \in \Lambda$. The Heaviside step-function cuts off the distribution at $m_{\text{max}} \in \Lambda$. One minus the term $\lambda_m \in \Lambda$ is the fraction of events that are part of this power-law distribution. The term $A(\Lambda)$ is a normalization constant. This term is motivated by the fact that the stellar mass function is power-law distributed as well as evidence of a cut-off in the black hole mass spectrum Fishbach et al. (2017); Talbot & Thrane (2017b); Abott et al. (2018a).

The second term in $\pi(m_1|\Lambda)$
\begin{equation}
\lambda_m B(\Lambda) \exp \left(-\frac{(m_1 - \mu_m)^2}{2\sigma_m^2}\right), \tag{B3}
\end{equation}
corresponds to a Gaussian distribution with mean $\mu_m \in \Lambda$ and width $\sigma_m \in \Lambda$. The fraction of events that are part of the Gaussian distribution is given by $\lambda_m$. The $B(\Lambda)$ term is a normalization constant.
This term is motivated by the possibility of a bump in the black hole mass spectrum from pulsational pair instability supernovae Talbot & Thrane (2018); Abbott et al. (2018a); Marchant, Renzo, Farmer, Pappas, Taam, de Mink & Kalogera (Marchant et al.).

To the far right of the expression for \( \pi(m_1|\Lambda) \) is a third term

\[
S(m,m_{\min},\delta m) = (\exp f(m - m_{\min},\delta m) + 1)^{-1}
\]

\[
f(m,\delta m) = \frac{\delta m}{m - \delta m}.
\]

The \( m_{\min} \) parameter enforces a minimum black hole mass and \( \delta m \) is the mass range over which the black hole mass spectrum falls to zero. This term is motivated by the fact that there is likely a minimum black hole mass, at least for black holes made through stellar collapse Talbot & Thrane (2018).

The conditional prior for mass ratio is described by a power law with index \( \beta \in \Lambda \). The smoothing function \( S \) applies a low-mass cut-off in the secondary mass \( m_2 \), again using minimum mass \( m_{\min} \) and \( \delta m \) for the mass range over which the mass spectrum falls to zero. The variable \( C(m_1,\Lambda) \) is a normalization constant.

**B2 Lab-frame mass**

The binary black hole lab-frame mass is a function of redshift because

\[
m_I = (1 + z)m_s,
\]

where \( m_s \) is the source-frame mass and \( m_I \) is the lab-frame mass. When considering events at cosmological distances, the prior distributions for lab-frame masses become covariant with luminosity distance \( D_L \) due to cosmological redshift. In the source frame, the distributions of black hole mass and redshift are separable so that

\[
\pi(z,m_s) = \pi_m(m_s)\pi_z(z)
\]

Whatever form the distributions we choose for \( \pi_z(z) \) and \( \pi_m(m_s) \), they imply some prior for the lab-frame mass:

\[
\pi(z,m_I) = \pi(z,m_s(m_I)) \left| \frac{dm_s}{dm_I} \right| = (1 + z)^{-1}\pi(z,m_I/(1 + z)).
\]

**B3 Spin**

The distribution of spin magnitudes \((a_1, a_2)\) are assumed to follow a beta distribution described by three parameters \((\alpha_a, \beta_a, a_{\max}) \in \Lambda \). By treating \( a_{\max} \) as a free parameter, our model is a generalization of the prescription from Wysocki et al. (2018). The conditional prior for spin magnitude is

\[
\pi_a(a|\alpha_a, \beta_a, a_{\max}) = \frac{a^{\alpha_a - 1}(a_{\max} - a)^{\beta_a - 1}}{a_{\max}^{\alpha_a + \beta_a - 1}B(\alpha_a, \beta_a)}.
\]

Here \( B(\alpha_a, \beta_a) \) is the Beta function.

We characterize the black hole spin orientation in terms of the cosine of the polar angle between the orbital angular momentum and the black hole spin \( z_{1,2} = \cos(t_{1,2}) \) where \( t_{1,2} \) is the polar angle. We ignore the azimuthal angle, which has a comparatively small effect on the gravitational waveform. We assume that the distribution of spin orientations is a mixture of an isotropic component and a preferentially aligned component modeled as a truncated half-Gaussian with unknown width \( \sigma_{\tilt} \) and which peaks at \( t_1 = t_2 = 1 \).

\[
\pi(z_1, z_2|\sigma_{\tilt}, a_{\tilt}) = \frac{1 - \lambda_{\tilt}}{4} e^{-\left(1-z_1^2\right)/2\sigma_{\tilt}^2} + \frac{\lambda_{\tilt}}{2\pi} \int_0^{\pi} \int_0^{\pi} e^{-\left(1-z_1^2\right)/2\sigma_{\tilt}^2} \sigma_{\tilt} \sin t \sin t dt dt
\]

The isotropic distribution is a model for mergers in dense stellar environments such as globular clusters, where spin orientations are expected to be isotropically oriented. The aligned distribution models binaries formed in the field. The fraction of binaries in the preferentially aligned component is \( \lambda_{\tilt} \). We assume that both component spins are independently drawn from the same distribution.

**REFERENCES**

Aasi J., et al., 2015, Class. Quant. Grav., 32, 074001
Abbott B. P., et al., 2016a, Phys. Rev. X, 6, 041015
Abbott B. P., et al., 2016b, Phys. Rev. Lett., 116, 131102
Abbott B. P., et al., 2018a, Phys. Rev. Lett., 120, 091101
Abbott B. P., et al., 2018b, Phys. Rev. Lett., 121, 161101
Abbott B. P., et al., 2019, Phys. Rev. X, 9, 031040
Abbott B. P., et al., 2018a
Abbott B. P., et al., 2018b
Acernese F., et al., 2014, Classical and Quantum Gravity, 32, 024001
Ade P. A. R., et al., 2016, Astronomy & Astrophysics, 594, A13
Akutsu T., et al., 2019, Nature Astronomy, 3, 354-340
Aston G., et al., 2018
Farr W. M., Stevenson S., Miller M. C., Mandel I., Farr B., Vecchio A., 2017, Nature, 548, 426
Fischbach M., Holz D. E., 2017, Astrophys. J. Lett., 851, L25
Fischbach M., Holz D., Farr B., 2017, Astrophys. J. Lett., 840, L24
Fischbach M., Holz D. E., Farr W. M., 2018, Astrophys. J. Lett., 863, L41
Gaebel S. M., Veitch J., Dent T., Farr W. M., 2018
Gerosa D., Berti E., 2017, Phys. Rev. D, 95, 124046
Hannam M., Schmidt P., Bohé A., HaegeL H., Husa S., Ohme F., Pratten G., Pürrer M., 2014, Phys. Rev. Lett., 113, 151101
Figure A1. One- and two-dimensional (hyper-) posterior distribution. This figure showcases duty cycle \( \xi \) and hyper-parameters related to the mass-spectrum peak. From left to right; the astrophysical duty cycle \( \xi \); the slope of the power law component of the primary mass distribution \( \alpha_m \); the slope of the mass ratio distribution \( \beta_m \); the minimum black hole mass \( m_{\text{min}} \); and the maximum black hole mass in the power-law component \( m_{\text{max}} \). The dashed lines are the 90\% credible intervals.
Figure A2. One- and two-dimensional (hyper-) posterior distributions. This figure showcases duty cycle hyper-parameters related to shape of the binary black hole mass spectrum. From left to right: The fraction of black holes in the Gaussian component of the primary mass distribution $\lambda_m$; the mean of the Gaussian component of the primary mass distribution $\mu_m$; and the standard deviation of the Gaussian component of the primary mass distribution $\sigma_m$. The dashed lines are the 90% credible intervals.

Talbot C., Thrane E., 2017b, Phys. Rev. D, 96, 023012
Talbot C., Thrane E., 2018, The Astrophysical Journal, 856, 173
Thrane E., Talbot C., 2018
Tiwari V., Fairhurst S., Hannam M., 2018, Astrophys. J., 868, 140
Vitale S., Lynch R., Sturani R., Graff P., 2017, Class. Quant. Grav., 34, 03LT01
Wysocki D., Lange J., O. ’shaughnessy R., 2018
You Z.-Q., Zhu X.-J., Ashton G., Thrane E., Zhu Z.-H., 2020
Figure A3. One- and two-dimensional (hyper-) posterior distributions. This figure shows hyper-parameters related to the distribution of black hole spins. From left to right: the fraction of BBHs with Gaussian distributed spin tilts $\xi_{\text{tilt}}$; the standard deviation of the spin-tilt angle distribution $\sigma_{\text{tilt}}$; the maximum spin magnitude $a_{\text{max}}$; the spin-magnitude beta distribution slope parameter (rise) $\alpha_a$; and the spin-magnitude beta distribution slope parameter (fall) $\beta_a$. The dashed lines are the 90% credible intervals.