Inferring the astrophysical population of gravitational wave sources in the presence of noise transients

Jack Heinzel [1,2]★ Colm Talbot,[1,2] Gregory Ashton[3] and Salvatore Vitale [1,2]

[1]*LIGO, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA*
[2]*Kavli Institute for Astrophysics and Space Research and Department of Physics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA*
[3]*Department of Physics, Royal Holloway, University of London, Egham Hill, Surrey, TW20 0EX, UK*

ABSTRACT

The global network of interferometric gravitational wave (GW) observatories (LIGO, Virgo, KAGRA) has detected and characterized nearly 100 mergers of binary compact objects. However, many more real GWs are lurking sub-threshold, which need to be sifted from terrestrial-origin noise triggers (known as glitches). Because glitches are not due to astrophysical phenomena, inference on the glitch under the assumption it has an astrophysical source (e.g. binary black hole coalescence) results in source parameters that are inconsistent with what is known about the astrophysical population. In this work, we show how one can extract unbiased population constraints from a catalogue of both real GW events and glitch contaminants by performing Bayesian inference on their source populations simultaneously. In this paper, we assume glitches come from a specific class with a well-characterized effective population (blip glitches). We also calculate posteriors on the probability of each event in the catalogue belonging to the astrophysical or glitch class, and obtain posteriors on the number of astrophysical events in the catalogue, finding it to be consistent with the actual number of events included.

Key words: black hole mergers – gravitational waves – methods: data analysis – methods: statistical.

1 INTRODUCTION

Since the first direct detection of gravitational waves (GWs) from the merger of two stellar mass black holes (Abbott et al. 2016), the LIGO–Virgo–KAGRA (LVK) network has observed a large population of these stellar mass binary black holes (BBHs) (Abbott et al. 2019a, 2021; The LIGO Scientific Collaboration et al. 2021). With so many detections comes the ability to characterize the population of BBHs, and shed light on the dominant formation channels of stellar mass BBH mergers. While there is no theoretical consensus on the dominant formation channel, there are many proposals.

For instance, isolated binary evolution through a common envelope phase (Smarr & Blandford 1976; Van Den Heuvel 1976; Tutukov & Yungelson 1993; Ivanova et al. 2013), stable mass transfer (Van Den Heuvel, Portegies Zwart & de Mink 2017), dynamical many-body interactions in dense stellar environments (e.g. globular clusters, Kulkarni, Hut & McMillan 1993; Sigurdsson & Hernquist 1993; Portegies Zwart & McMillan 2000), chemically homogeneous stellar evolution (Mandel & de Mink 2016; Marchant et al. 2016), dynamical triples assisted by the Kozai–Lidov mechanism (Antonini, Toonen & Hamers 2017; Silsbee & Tremaine 2017), or primordial BBH systems (Bird et al. 2016; Ali-Haïmoud, Kovetz & Kamionkowski 2017) have been proposed. Traces of these different formation channels are imprinted in the population, distinguishing the relative rates and constraining the sub-population distributions (Mapelli 2021; Zevin et al. 2021; Mandel & Broekgaarden 2022). As more GWs are detected, the different astrophysical formation channels will begin to reveal themselves.

However, one is never sure of the origin of a potential GW detection. GWs are detected using search pipelines, which vary in their methodology, but in general scan the LVK data stream for matches to a GW template within some template bank dense over the expected source parameters (Allen 2005; Usman et al. 2016; Messick et al. 2017; Nitz et al. 2017; Hanna et al. 2020). This provides a point estimate on the source parameters with the best match template. If this best match passes some significance threshold, it is called a trigger.

GW interferometers are plagued by transient noise fluctuations (known as glitches), whose morphology occasionally mimics real events (Zevin et al. 2017; Cabero et al. 2019; Akutsu et al. 2021; Davis et al. 2021; Soni et al. 2021; Ashton et al. 2022; Acernese et al. 2022a, b). Most pipelines estimate the false alarm rate (FAR) of a trigger by time-sliding the data of different interferometers by more than the light-travel time between them. Any coincident triggers therefore cannot be caused by a GW propagating at the speed of light, and are deemed false alarms. By varying the time-slide and counting the total number of false alarms, pipelines can accurately estimate the FAR of a trigger. Comparing the FAR to the expected astrophysical rate of the trigger, search pipelines estimate the probability of astrophysical origin, or p_{astro}. In order to calculate the expected astrophysical rate of the trigger, pipelines must assume a model for the underlying astrophysical source population (The LIGO Scientific Collaboration et al. 2021).

To mitigate contamination from glitches, it is standard to use only the most significant events. Because p_{astro} estimates assume a population, it is unusual to use pipeline calculated p_{astro} as a threshold

★ E-mail: jack.heinzel@gmail.com

for population inference. Instead, a common threshold is FAR <1 yr^{-1}, yet even with this high threshold, one expects e.g. 4.6 false alarms in the catalogue used by Abbott et al. (2023b) under the assumption that the search pipelines produce events independently (Allen 2005; Usman et al. 2016; Messick et al. 2017; Nitz et al. 2017; Hanna et al. 2020; Abbott et al. 2023b). Therefore, one must tune the FAR threshold to minimize the systematic uncertainty of including more false alarms in the catalogue, and the statistical uncertainty of including fewer events.

There are also a plethora of sub-threshold (FAR $> 1yr^{-1}$) astro-physical events which contain information about the population of gravitational-wave sources in the Universe, especially in some of the more poorly measured regions of parameter space, where glitches are responsible for reduced search sensitivity. Sub-threshold mergers of binary neutron stars (BNS), neutron star black holes (NSBH), or stellar mass BBHs can improve known constraints on the population of these as GW progenitors. Indeed, there are many more events with lower significance; the rate of GW events scales with SNR^{-4}, assuming a constant merger rate in a Euclidean volume (Schutz 2011; Chen & Holz 2014). Though these lower significance events also encode less information about the progenitor, events as low as $SNR\sim6-7$ can have well-measured chirp masses (Huang et al. 2018).

Moreover, certain kinds of theoretical GW events may pass this FAR threshold only rarely, with the majority falling deep into the sub-threshold range. For instance, subsolar-mass compact objects are predicted by certain modifications to the standard model of particle physics or ΛCDM (Shandera, Jeong & Grasshorn Gebhardt 2018; Nitz & Wang 2021; Abbott et al. 2022a). Though no direct detections have been made of a sub-solar mass merger (Abbott et al. 2019b; Nitz & Wang 2021), it is possible there are some lurking within the large set of sub-threshold candidates; because of their low masses, the signal-to-noise ratio (SNR) and significance of the GW will be much lower.

Glitches in GW interferometers are commonly studied by mod-elling the data as some parametric and deterministic function plus a stationary and stochastic noise process (Cornish & Littenberg 2015; Merritt et al. 2021; Tolley et al. 2023; Udall & Davis 2023). This is preferable to modelling glitches as some general non-stationary noisy time series, where the statistical properties are unclear. A glitch model then requires a parametric function, called the glitch waveform, for the deterministic part of the sig-nal. Since significant false alarms will mimic real GWs, it is sensible to use a GW model for the glitch waveform. In this paper, we follow this prescription, modelling glitches with a GW waveform.

A more general glitch model distinguishes GWs from terrestrial glitches by signal coherence. Real GWs must be coherent between multiple detectors and the waveforms should be consistent with the same progenitor parameters, while the same is not true for coincident false alarms (Veitch & Vecchio 2010). Glitches may therefore be modelled as an independent GW waveform in each detector, relaxing this coherence requirement. This is justified as a worst-case scenario, where a background event is distinguished from an astrophysical one based purely on the signal coherence. This glitch model has been used to calculate the probability an event is astrophysical (Isi et al. 2018; Ashton, Thrane & Smith 2019; Pratten & Vecchio 2021), and to rule out marginal candidates (e.g. Ashton & Thrane 2020; Vajpeyi et al. 2022). The most general glitch models make no physical assumptions about the source and model glitches as a superposition of wavelets (Cornish & Littenberg 2015).

Whatever the waveform assumed for the glitches, a population would then be given by probability distributions on their parameters. Indeed, it is possible to study the population of glitches and astro-physical events simultaneously, allowing for each event to belong to either class. Previous work approached this problem from different perspectives. Farr et al. (2015) showed how to infer the rates of astrophysical and background populations when the shapes of the populations are known, but the identity of each event (i.e. which population it originates from) is unknown. Gaebel et al. (2019) show that it is indeed possible to do joint inference on an astrophysical and a glitch population, but leave a study with real GW data for a future analysis. Galaudage, Talbot & Thrane (2020) and Roulet et al. (2020) analyse real GW data, and fold in pipeline information—in particular, p_{astro} estimates, to build a glitch population model. However, this carries a fixed background event rate estimate by each search pipeline, rather than inferring the rate of events from the background population in a Bayesian manner.

In this paper, we present a general method to simultaneously model the population of background non-astrophysical triggers and the population of astrophysical objects, in a fully Bayesian manner. We use a population of short glitches ('blips') as identified by the GRAVITYSPY algorithm (Zevin et al. 2017) to contaminate the catalogue of astrophysical signals. While this is done for com-putational expedience, the method can be used for any type of non-astrophysical transients, as long as one can characterize their 'usual' properties. Similarly. while we focus on the population of BBHs, the method may be used to study any population of foreground events contaminated with undesirable background events. In Section 2, we briefly review Bayesian parameter estimation of GW sources and population inference. Then, we discuss how this picture is complicated when one allows for the possibility that the data set is contaminated by glitches. In Section 2.4, we discuss our glitch population parametrization and constrain the population hyperparameters using a large representative sample. In Section 3, we contaminate a catalogue of GWs with glitches, and show how our method consistently models and removes the bias due to the contaminants. Finally, in Section 4, we summarize and discuss future work.

2 METHODS

2.1 Parameter estimation

Consider a stretch of LVK frequency domain data d, which is a sum of noise n and waveform signal $h(\theta)$

$$d = h(\theta) + n, \tag{1}$$

where θ represents the unknown parameters of the GW source. Approximating the noise as stationary and Gaussian, the likelihood can be written

$$\log \mathcal{L}(d|\theta) = -\sum_j \left(2\Delta f \frac{|d_j - h_j(\theta)|^2}{\mathscr{P}_j} + \log(2\pi \mathscr{P}_j) \right), \tag{2}$$

where d_j and h_j represent the jth frequency component of the data and waveform, respectively, $\mathscr{P}_j$ is the power-spectral-density, and Δf is the frequency spacing (Whittle 1951). With this likelihood, a model for the waveform $h(\theta)$ given some GW parameters, and priors for the GW parameters, one can then sample from the posterior of the GW

parameters (Veitch et al. 2015; Thrane & Talbot 2019; Christensen & Meyer 2022).

The above process also can apply to glitches, thinking of them as a deterministic signal buried in stochastic noise. Modelling glitches under some parameterization (e.g. a sine-Gaussian), one can perform parameter estimation exactly as above for the glitch parameters, which we denote ψ. Indeed, while glitches are usually ruled out by search pipelines by e.g. χ^2 discriminators (Allen 2005), there can be cases where glitches are mistaken for astrophysical GWs. Because population inferences generally assume that all events in the catalogue are truly astrophysical, a contaminant glitch in the catalogue will bias the inference. We want to relax this assumption, and jointly infer the population of astrophysical events and glitches.

2.2 Population inference without glitches

Before we discuss simultaneous inference of the astrophysical and glitch populations, we review the general GW population inference problem. Given posterior samples from a set of data time-series $\{d_i\}_{1 \le i \le N_{\text{events}}}$, one can write the likelihood for a population model. In general, a population model describes the rate of mergers within a small interval of GW parameter space $[\theta, \theta + d\theta]$. However, the rate is typically assumed to be a Poisson process, and we can instead write down a probability density $p_A(\theta|\Lambda)$, irrespective of the overall rate. Here, Λ are called the hyper-parameters; a finite list of parameters, which vary the shape of the population distribution (e.g. the mean and variance of a Gaussian, the power index to a power-law, and so on). We give the subscript A to refer to 'astrophysical'. This is in contrast to G for 'glitch', which we will use later in this paper.

Assuming a Poisson process for the events and marginalizing over the overall rate R with an uninformative (uniform in $\log R$) prior, one obtains the hierarchical likelihood

$$\mathcal{L}(\{d_i\}|\Lambda) \propto \prod_{i=1}^{N_{\text{events}}} \frac{\int d\theta \, \mathcal{L}(d_i|\theta) p_A(\theta|\Lambda)}{\alpha(\Lambda)} \tag{3}$$

and the selection function

$$\alpha(\Lambda) = \int d\theta \, p_{\text{det},A}(\theta) p_A(\theta|\Lambda) \tag{4}$$

is the fraction of events, which are detectable in the population with hyper-parameters Λ (for a derivation of the likelihood see Mandel, Farr & Gair 2019; Vitale et al. 2020). The quantity $p_{\text{det},A}(\theta)$ is the probability of detecting an astrophysical event with parameters θ, given by

$$p_{\text{det},A}(\theta) = \int_{\{d \in \mathcal{D} | \rho(d) > \rho_{\text{thr}}\}} \mathcal{L}(d|\theta) \mathrm{d}d \tag{5}$$

the integral over all possible data realizations, which exceed the detection threshold $\rho(d) > \rho_{\text{thr}}$ (i.e. FAR $< 1 \, \text{yr}^{-1}$; as in Abbott et al. 2023b).

In practice, the integrals in equations (3) and (4) are estimated with Monte Carlo estimators. In particular,

$$\int d\theta \, \mathcal{L}(d_i|\theta) p_A(\theta|\Lambda) \sim \frac{Z(d_i)}{N_{\text{samp}}} \sum_{j=1}^{N_{\text{samp}}} \left. \frac{p_A(\theta_j|\Lambda)}{\pi(\theta_j|\mathcal{H}_{\text{PE}})} \right|_{\theta_j \sim p(\theta|d_i)}, \tag{6}$$

where θ_j are samples from the ith event posterior,

$$Z(d_i) = \int d\theta \, \mathcal{L}(d_i|\theta) \pi(\theta|\mathcal{H}_{\text{PE}}) \tag{7}$$

is the evidence and $\pi(\theta|\mathcal{H}_{\text{PE}})$ is the sampling prior used for the parameter estimation. As for the selection function,

$$\alpha(\Lambda) \sim \frac{1}{N_{\text{draw}}} \sum_{j=1}^{N_{\text{det}}} \left. \frac{p_A(\theta_j|\Lambda)}{p_{\text{draw}}(\theta_j)} \right|_{\theta_j \sim p_{\text{draw}}(\theta)}, \tag{8}$$

where N_{draw} events are drawn from some fiducial distribution $p_{\text{draw}}(\theta)$, data drawn from the conditioned likelihood $\mathcal{L}(d|\theta)$ with a suitable power-spectral-density choice, and then search pipelines run to recover N_{det} of the total events (for details see e.g. Tiwari 2018; Farr 2019).

2.3 Population inference with glitches

The above procedure assumes every event, which passes the threshold is a real GW. This assumption can be relaxed by simultaneously fitting the glitch population. Suppose the glitch waveform is given by parameters ψ, and we obtain posteriors on $p(\psi|d_i)$ for each event in the catalogue, as well as posteriors on $p(\theta|d_i)$ for the GW parameters. With equation (79) in Vitale et al. (2020) and a relative rate η of GWs versus a GW-like glitches, one can marginalize over the total rate with a uniform in $\log R$ prior to generalize equation (3).

$$\mathcal{L}(\{d_i\}|\Lambda_A, \Lambda_G, \eta) \propto$$
$$\prod_{i=1}^{N_{\text{events}}} \frac{\eta \int d\theta \, \mathcal{L}(d_i|\theta) p_A(\theta|\Lambda_A) + (1-\eta) \int d\psi \, \mathcal{L}(d_i|\psi) p_G(\psi|\Lambda_G)}{\eta \alpha_A(\Lambda_A) + (1-\eta)\alpha_G(\Lambda_G)}, \tag{9}$$

where Λ_A and Λ_G refer to the astrophysical and glitch hyper-parameters, $p_G(\psi|\Lambda_G)$ is the population model for the glitch waveform parameters, and $\alpha_X(\Lambda_X)$ is the selection function for the X subpopulation:

$$\alpha_X(\Lambda_X) = \int d\theta \, p_{\text{det},X}(\theta) p_X(\theta|\Lambda_X), \tag{10}$$

$p_{\text{det},G}$ is analogous to the $p_{\text{det},A}$ we defined above, but we want to allow for the possibility that the detection criterion $\rho(d) > \rho_{\text{thr}}$ is different for glitches. In reality, the same detection criterion must be used for all events for a catalogue, but for reasons we will describe below, we must use a different detection criterion for glitches in this study.

The mixing fraction η represents the relative rate of all GWs from all GW-like sources (astrophysical and glitches), whether they are detected or not. It is useful to define a detectable mixing fraction:

$$\bar{\eta} = \frac{\eta \alpha_A(\Lambda_A)}{\eta \alpha_A(\Lambda_A) + (1-\eta)\alpha_G(\Lambda_G)}, \tag{11}$$

which is the fraction of detectable events which are GWs. In this case, and a bit of algebra, the likelihood of equation (9) can be recast as

$$\mathcal{L}(\{d_i\}|\Lambda_A, \Lambda_G, \eta) \propto$$
$$\prod_{i=1}^{N_{\text{events}}} \frac{\bar{\eta} \int d\theta \, \mathcal{L}(d_i|\theta) p_A(\theta|\Lambda_A)}{\alpha_A(\Lambda_A)}$$
$$+ \frac{(1-\bar{\eta}) \int d\psi \, \mathcal{L}(d_i|\psi) p_G(\psi|\Lambda_G)}{\alpha_G(\Lambda_G)}, \tag{12}$$

which is the form of the likelihood we will use in the sampling.

So far we have assumed glitches and GWs will be characterized with different parameters, θ and ψ. However, glitches which can contaminate a GW catalogue will necessarily be well-modelled by a GW waveform. For this proof-of-principle analysis, we thus model the waveform of a glitch as a GW (we set $\psi \to \theta$). Furthermore,
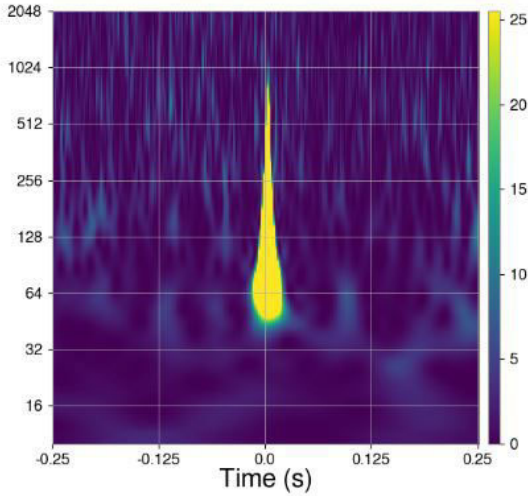
Figure 1. A time–frequency spectrogram of a GRAVITYSPY-identified blip glitch in LIGO-Hanford. This blip occurred on 2019 August 28 at UTC 16:56:49. Due to their short duration, blip glitches can be mistaken for high mass BBHs.

we only model the population in the intrinsic GW parameters; this will be explained further in Section 2.4. This simplifies the analysis: we do not need pieces of evidence and posterior samples for every event under both the glitch and GW hypotheses—both analyses are the same. Indeed, under these assumptions the analysis reduces to a GW population inference with a mixture population; equation (9) becomes equation (3) with

$$p(\theta|\Lambda) \rightarrow \eta p_A(\theta|\Lambda_A) + (1-\eta)p_G(\theta|\Lambda_G), \tag{13}$$

and a selection function

$$\alpha(\Lambda) \rightarrow \eta \alpha_A(\Lambda_A) + (1-\eta)\alpha_G(\Lambda_G). \tag{14}$$

Equation (13) treats the glitch population as an additional 'astrophysical' population, albeit occupying a different region of parameter space from the population of true astrophysical BBHs.

There is one additional caveat. In the LVK population analysis of Abbott et al. (2023b), events included in the catalogue are selected by their FAR ($<1 \, \mathrm{yr}^{-1}$), and so we would like to also select glitches by their FAR to match Abbott et al. (2023b). However, this requires us to calculate FARs for many injections from a fiducial glitch population. Running search pipelines to calculate FARs of injected glitches may be necessary for a future study, however for this proof-of-principle paper it is simply too expensive. Instead, we select glitches for inclusion with a cheaper threshold, the SNR. We can then estimate $\alpha_G(\Lambda_G)$ with a reweighted Monte Carlo estimator using a custom set of injections, and estimate $\alpha_A(\Lambda_A)$ with the injection set already provided in LVK (2021).

2.4 Characterizing the glitch population

In the citizen-science project GRAVITYSPY, glitches are classified according to their time frequency spectrograms (Zevin et al. 2017; Glanzer et al. 2023). For instance, blip glitches are short bursts of excess power, with a time frequency spectrogram morphology shown in Fig. 1.

In fact, blip glitches are more likely to contaminate a GW catalogue, since they can mimic high mass BBHs (Cabero et al. 2019). For this reason, we restrict this first study to blip glitches,

though the formalism can be extended to any glitch class, or even combination of classes. This would require a new population model and Λ_G for each additional class, plus a mixing fraction.

In order to understand various populations of glitches, Ashton et al. (2022) analysed a set of 1000 GRAVITYSPY identified blip glitches with the `IMRPhenomPv2` GW waveform (Hannam et al. 2014; Bohé et al. 2016; Husa et al. 2016; Khan et al. 2016). Since blip glitches are not due to any astrophysical process, they are usually present in a single detector, with multiple detector coincidences occurring randomly. As single detector triggers, only information about the intrinsic parameters (masses and spins) may be extracted. Therefore Ashton et al. (2022) provides posterior samples only over the intrinsic parameters and the redshift.

With the posterior samples in hand, Ashton et al. (2022) fit a population model in the detector frame chirp mass, mass ratio, and primary spin (see their Figs 2–4). Qualitatively, the population of GRAVITYSPY blip glitches shows different features from the population of BBHs (extreme mass ratios, spins, and low redshifts, inconsistent with e.g. Abbott et al. 2023b). We will use this to our advantage to separate the populations.

We slightly modify the population model of Ashton et al. (2022). Instead of modelling the primary spin magnitude, we model in the effective spin parameter:

$$\chi_{\mathrm{eff}} = \frac{a_1 \cos\theta_1 + q a_2 \cos\theta_2}{1+q}, \tag{15}$$

where a_1 and a_2 are the spin magnitudes of the primary and secondary BHs in Kerr units, $q = m_2/m_1$ is the mass ratio (where $0 < q < 1$ by convention), and θ_1 and θ_2 are the spin tilts measured from the orbital angular momentum. χ_{eff} is the spin parameter, which occurs at lowest order in the waveform, and is measured better than individual spins (Racine 2008; Pürrer, Hannam & Ohme 2016; Vitale et al. 2017; Ng et al. 2018).

We model the glitch population in the detector-frame chirp mass, mass ratio, effective spin parameter, and redshift: $\theta = (\mathcal{M}_{\mathrm{c,det}}, q, \chi_{\mathrm{eff}}, z)$. In particular, we use a skewed Gaussian (equation (16)) for both the detector-frame chirp mass $\mathcal{M}_{\mathrm{c,det}}$ and the redshift z with hyper-parameters μ_m, σ_m, κ_m and μ_z, σ_z, and κ_z, respectively

$$p(x|\mu, \sigma, \kappa) = \frac{2}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right) \Phi\left(\kappa \frac{x-\mu}{\sigma}\right), \tag{16}$$

where ϕ and Φ are the standard Gaussian and Gaussian integral, respectively. We model χ_{eff} and q with a correlated mixture model of two two-dimensional Gaussians in the $\chi_{\mathrm{eff}}-q$ plane with hyper-parameters denoted $\vec{\lambda}_{q\chi}$ for brevity

$$\begin{aligned} p_{q\chi}(q, \chi_{\mathrm{eff}}|\vec{\lambda}_{q\chi}) = {} & N_1 \eta_{q,\chi} \phi(f_1[q, \chi_{\mathrm{eff}}]) \phi(g_1[q, \chi_{\mathrm{eff}}]) \\ & + N_2(1-\eta_{q,\chi}) \phi(f_2[q, \chi_{\mathrm{eff}}]) \phi(g_2[q, \chi_{\mathrm{eff}}]) \\ \vec{\lambda}_{q\chi} = {} & (\mu_{q,1}.\mu_{q,2}, \mu_{\chi,1}, \mu_{\chi,2}, \sigma_{q,1}, \sigma_{q,2}, \sigma_{\chi,1}, \\ & \sigma_{\chi,2}, \theta_{q,\chi}, \eta_{q,\chi}) \end{aligned} \tag{17}$$

where

$$f_i[q, \chi_{\mathrm{eff}}] = \frac{(q-\mu_{q,i})\cos(\theta_{q,\chi}) + (\chi_{\mathrm{eff}}-\mu_{\chi,i})\sin(\theta_{q,\chi})}{\sigma_{q,i}} \tag{18}$$

$$g_i[q, \chi_{\mathrm{eff}}] = \frac{(q-\mu_{q,i})\sin(\theta_{q,\chi}) + (\chi_{\mathrm{eff}}-\mu_{\chi,i})\cos(\theta_{q,\chi})}{\sigma_{q,i}}, \tag{19}$$

and N_1 and N_2 are normalization coefficients, numerically calculated because χ_{eff} and q are required to be positive. Equation (18) describes a pair of two dimensional Gaussians with branching fraction $\eta_{q,\chi}$, parametrized by the variances along the eigenvectors of the covariance matrix ($\sigma_{q,i}^2$ and $\sigma_{\chi,i}^2$), and the angle they are
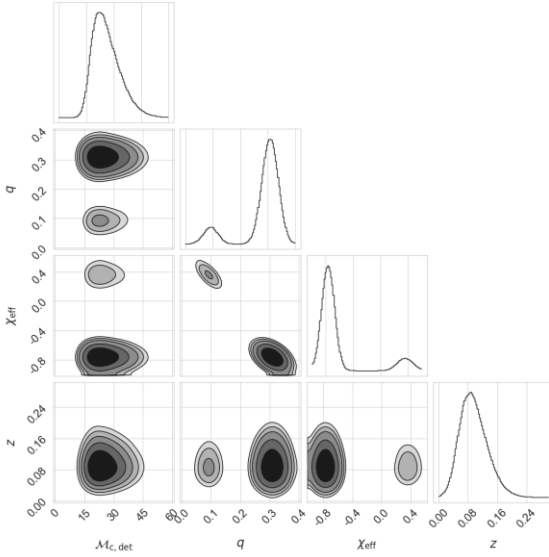
Figure 2. The posterior population distribution of the population inference on the glitch population alone, using all 1000 blip posteriors from Ashton et al. (2022). Contours show the 1–5σ regions.

'tilted' by ($\theta_{q,\chi}$, assumed to be the same for both Gaussians). The glitch population model for is the product of the $\mathcal{M}_{c,det}$, z, and $q-\chi_{eff}$ models, and Λ_G is the union of their hyper-parameters. We chose to leave precession unmodelled in the population by projecting the six-dimensional spin population on to the effective aligned spin parameter. However, a future study could examine how the populations further separate including the spin precession parameter and correlations therein, or in the full six-dimensional spin space.

We are now ready to measure the population of blip glitches with our model, as is done in Ashton et al. (2022). By using all 1000 posteriors from Ashton et al. (2022) we obtain tight constraints on the glitch population alone. This is a critical step of our analysis. We must measure the population of glitches well to optimally separate it from the population of GWs. Fortunately, we have access to the unbiased population of blip glitches before any selection criteria are enforced.[1] We may assume the 1000 glitches from Ashton et al. (2022) are a representative sample. The constraints we measure in this step inform the boundaries of the priors we use during simultaneous inference. We show the posterior population distribution (PPD) in Fig. 2.

As for the astrophysical population parametrization, we use the power-law plus peak model of Talbot & Thrane (2018), Abbott et al. (2023b), and the redshift model of Fishbach, Holz & Farr (2018). We modify the spin distribution model by modelling χ_{eff} with a Gaussian, following Roulet & Zaldarriaga (2019), Miller, Callister & Farr (2020), and Callister et al. (2021). This gives us the set of Λ_A and Λ_G, which will be inferred together with the detectable mixing fraction $\overline{\eta}$ in our joint analysis.

---

[1]Note there is a cut on these GRAVITYSPY glitches with SNR > 8, and another given they are GRAVITYSPY identified. We can still treat this as the unbiased population with no changes to our analysis. In a real analysis, one would still have access to an unbiased sample of the population of glitches.

2.5 Simultenous inference and selection effects

We model the selection effects of glitches in entirely the same way we model the selection effects of GWs. We emphasize that selection effects depend on the data alone. If we believe glitches have data well-modelled by a GW plus Gaussian noise, then the probability of detecting a glitch is well-approximated by the probability of detecting a GW with the corresponding parameters.

We also define $p_{astro,i}(\Lambda)$ for each event in the catalogue, a population dependent quantity,

$$p_{astro,i}(\Lambda) = p(astro|d_i, \Lambda) = \frac{\eta\mathcal{L}(d_i|\Lambda_A)}{\eta\mathcal{L}(d_i|\Lambda_A) + (1-\eta)\mathcal{L}(d_i|\Lambda_G)}. \tag{20}$$

This comes directly from Bayes' Theorem. It is perhaps more intuitive to use $\overline{\eta}$ instead of η, however, in that case the likelihood terms must each acquire a $1/\int d\theta p_{det,X} p_X(\theta|\Lambda_X)$ term, and it reduces again to equation (20). This folds in the dependence on source parameters and uncertainty in the population hyper-parameters, and so in general p_{astro} is a posterior, based on the posterior on Λ. Search pipelines output a point-estimate of this quantity for each event, using the point estimate on the progenitor parameters with the matched template, a fixed underlying astrophysical population, and a direct calculation of the glitch-rate term with the FAR. Farr et al. (2015) and Kapadia et al. (2020) define similar quantities.

3 RESULTS

We contaminate the catalogue of 69 BBH events (with FAR < 1 yr^{-1}) analysed in Abbott et al. (2023b) with blip posteriors obtained from Ashton et al. (2022). Note the posteriors in Abbott et al. (2023b) are sampled used state-of-the-art waveforms including higher order modes, while Ashton et al. (2022) uses the rapid IMRPhenomPv2, a waveform approximant including only the dominant $(l, m) = (2, 2)$ mode (Hannam et al. 2014; Bohé et al. 2016; Husa et al. 2016; Khan et al. 2016). Indeed, blip posteriors converge on unequal mass ratios ($q \sim 0.1$), where higher order modes become significant. While this will bias the glitch population model, this paper is intended to be a proof of concept and so we use the posterior samples as provided.

We inject $N_{blip} = [0, 1, 2,..., 19, 20]$ contaminant posteriors from Ashton et al. (2022) into the set of 69 BBH posteriors analysed in Abbott et al. (2023b). We then sample the hyper-posterior of Λ using the nested sampler DYNESTY (Speagle 2020; Koposov et al. 2022) and the code GWPOPULATION (Talbot et al. 2019). We must cut regions of parameter space above the total variance of the hierarchical likelihood-estimator. Without handling variance the sampler can converge on regions of parameter space with poor Monte Carlo estimates, and thereby bias the posterior sampling from the true posterior (Golomb & Talbot 2022). We do this as well as cut out regions with poorly behaved selection function estimates, as described in Farr (2019) and Essick & Farr (2022).

We describe several methods of quantifying the bias (or lack thereof) of performing the simultaneous inference.

3.1 The detectable mixing fraction

For each catalogue and its inference, we obtain a posterior on the detectable mixing fraction $\overline{\eta}$. We plot these posteriors as violin plots in Fig. 3. The dashed-black line is the true mixing fraction in our catalogue, given by the number of BBHs divided by the total number of events. Note the posteriors peak at the dashed line, i.e. it is recovering the correct number of contaminants.
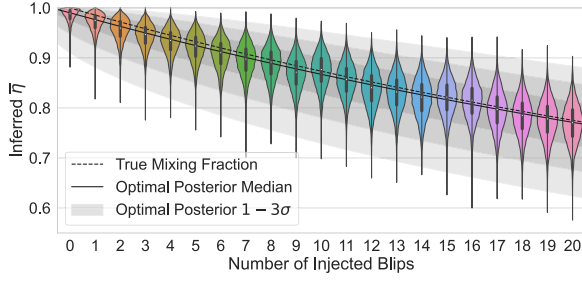
Figure 3. The violins show the inferred detectable mixing fraction $\overline{\eta}$ for each run. The *x*-axis indexes the number of injected blips and each violin refers to a different inference. The black-dashed line is the injected mixing fraction, given by $1-N_{blips}/N_{events}$. Notice the inference recovers the injected mixing fraction well. We compare against the optimal posterior, which would be inferred with perfect knowledge on which events are BBHs and are glitches. We show the $1-3\sigma$ and median of this optimal posterior in black (see the Appendix and equation (A1)).

While the results here suggest the sampling is correctly recovering the blips, there is a caveat. The quantity $\overline{\eta}$ represents a statement on the underlying relative rates, it is not the fraction of BBHs in the catalogue. In other words, this is not a like to like comparison. We want to understand what our inference predicts are the number of BBHs and blips in our catalogue.

For instance, suppose we may unambiguously identify which events are BBHs and which are blips solely using the event parameters. That is, the populations are disjoint to the point that every event posterior overlaps with only one of the astrophysical or glitch populations. It turns out that $\overline{\eta}$ does not converge on a delta function: it will have some width due to Poisson rate uncertainty. Rather, it converges on an analytic optimal posterior, which we calculate by assuming the populations are so disjoint that every event posterior uniquely determines which population the event originates from. Details on this calculation are in the Appendix.

From this theoretical optimal posterior, we can calculate the median and $1-3\sigma$ levels, which we show as a function of the number of added contaminants (*x*-axis) in Fig. 3. Note how similar the measured posteriors on $\overline{\eta}$ are to the optimal posterior given perfect knowledge on which events are BBHs and glitches. The populations of blip glitches and BBHs are nearly disjoint; this suggests the inference can uncover, which events are in which population much more precisely than the $\overline{\eta}$ posteriors naively indicate.

3.2 Inferred number of contaminants and BBHs

Calculating p_{astro}[2] (i.e. equation (20)) for each event in each run, we notice that the posteriors on each event tends to be sharply peaked, e.g. GW150914 peaks at $p_{astro}(\Lambda) \to 1$, the blips peak at $p_{astro}(\Lambda) \to 0$. We show posteriors on $1-p_{astro} = p_{blip}$ for two example events in Fig. 4, GW151226 and GW200302. GW200302 is the event with the highest probability of being a blip, see the Appendix for details. Note as the number of injected blips increases, the p_{blip} increases for GW200302. This is because the $\overline{\eta}$ posterior converges on lower mixing fractions; lowering the odds that any given event is

---

[2]We emphasize that statements made in this paper about p_{astro} should be understood as the probability of the event not being a blip, rather than the probability of the event being astrophysical in origin. This is rather cumbersome to write, so we continue with the abuse of notation in p_{astro}.
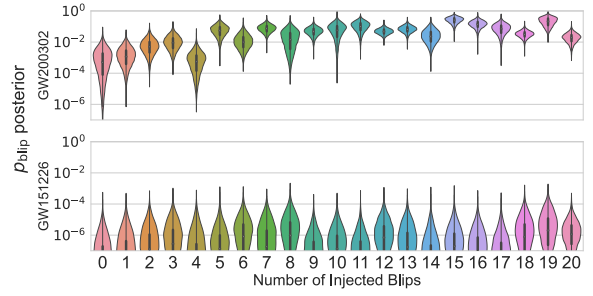
Figure 4. Calculated $1-p_{astro} = p_{blip}$ for two events, GW200302 and GW151226, in each inference. GW200302 consistently had the lowest p_{astro} of all the BBH events, while we selected GW151226 to be a representative event for the standard BBH in the catalogue. Note a subtle trend for decreasing p_{astro} as the number of injected blips increased.

astrophysical. This is much more apparent in GW200302, where p_{blip} is mostly dominated by these odds. GW151226 is a representative event for what most BBH p_{astro} posteriors look like. In fact, many posteriors are even more extreme than GW151226; $\log_{10}(p_{blip}) \to -\infty$ for many events, see Table A1 in the Appendix for the full event list.
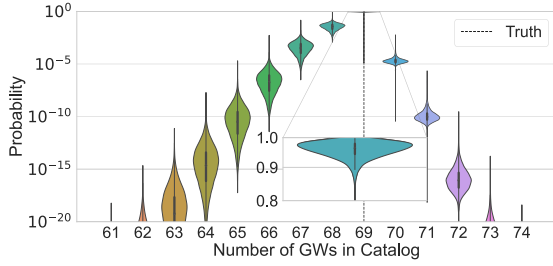
Most p_{astro} posteriors are sharply peaked, nearly delta functions. Translating this into a calculation on the number of BBHs and blips in the catalogue, this suggests that the inferred number of BBHs and blips in the catalogue is also sharply peaked. Indeed, using the $p_{astro,\,i}$ defined in equation (20) we may calculate the probability that exactly k of N_{events} are astrophysical. Since each data realization is independent, the $p_{astro,\,i}$ of each event will be statistically independent. The probability that exactly k of N_{events} total events in the catalogue are BBHs is then

$$p_k(\Lambda) = \sum_{\gamma \in \Gamma(k, N_{events})} \left[\prod_{j=1}^{k} p_{astro, \gamma(j)}(\Lambda) \prod_{j=k+1}^{N_{events}} 1 - p_{astro, \gamma(j)}(\Lambda) \right], \tag{21}$$

where $\Gamma(k, N_{events})$ is the set of *k*-combinations of N_{events} (it contains N_{events} choose k elements), a subset of the set of permutations of N_{events}. Thinking of permutations as one-to-one and on to functions from the set $\{1,..., N_{events}\}$ to itself, *k*-combinations are permutations where two permutations γ_1 and γ_2 are equivalent if there is the set equality $\gamma_1(\{1,..,k\}) = \gamma_2(\{1,..,k\})$. Informally, the probability that exactly k of N_{events} are BBHs is the probability a specific set of *k* events are BBHs and the others are glitches, summed over all the possible sets of *k* events. Note that if all $p_{astro,\,i}$ are the same, equation (21) reduces to the binomial distribution as expected. However, equation (21) is much too computationally expensive to evaluate directly. We use a trick with symmetric polynomials to vastly simplify the calculation, see the appendix for details. We also note that Galaudage et al. (2020) consider the sum of the $p_{astro,\,i}$. This is the expectation value over *k* of equation (21), which is also discussed in further detail in the Appendix.

After contaminating the catalogue of 69 BBHs passing the LVK selection criteria (Abbott et al. 2023b) with 0–20 independently drawn random blips, and running 21 inferences on the hyper-parameters Λ on the 21 variably contaminated catalogues, we calculate equation (21) for each Λ sample. We show an example in Fig. 5, the run with 20 contaminant blips. In this run and in most runs, the probability for exactly 69 BBHs in the catalogue rails against 1, while for some other runs it can be more uncertain. Variability between runs is due

Figure 5. The probability of having *k* events, which are astrophysical in the catalogue. The horizontal axis is the number of events in the catalogue, and the vertical axis represents the $p_k(\Lambda)$ probability of there being exactly *k* astrophysical events in the catalogue. Since the probability for exactly 69 BBHs rails against 1, we show an inset zoom on the $p_{69}(\Lambda)$ violin. The uncertainty in the value of the probability $p_k(\Lambda)$ comes from the uncertainty in the population parameters Λ. This particular run was with 69 BBHs injected and 20 contaminant blips injected.
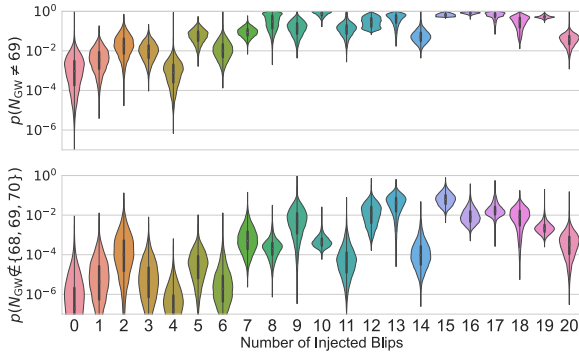

Figure 6. In the top panel, we show violins for the inferred posterior probabilities of the catalogue not having 69 BBHs in it; $1-p_{69}(\Lambda)$. The vertical axis shows the logarithm of the probability, and the horizontal axis is the number of injected blips in the catalogue. In the bottom panel, we show the posterior probabilities of the catalogue having some number of BBHs which is not 68, 69, or 70; $1-p_{68}(\Lambda) - p_{69}(\Lambda) - p_{70}(\Lambda)$. Note the increase in the probabilities as the number of injected blips increases; this is due to higher odds that any given event is a blip (lower $\overline{\eta}$). The dip at exactly 20 injected blips is because those 20 contaminants happen to be easily resolvable from the GW population, and so $p_{69}(\Lambda)$ peaks strongly at 1.

to the differences in how 'BBH-like' the blip contaminants are, and how well they fit into the blip population model.

We also show posteriors on the probabilities of having exactly 69 BBHs in the catalogue. Specifically, since many of the probabilities rail against 1, we show the logarithm of the negation: the $\log_{10}$ probability of not having 69 BBHs in the catalogue, shown in the top panel of Fig. 6. As the number of contaminants increases, the resolving power drops, meaning the probability becomes more spread out between ~68 and 70. Furthermore, the odds any given event is a BBH drops, as the mixing fraction between BBHs and blips becomes more blip-favoured. That said, up to 20 injected blips we observe significant probabilities of exactly 69 BBHs in the catalogue, and near unity probabilities of 68 or 69 or 70 BBHs in the catalogue (Fig. 6).

While there is some variation in the probabilities, this method consistently recovers the correct number of injected contaminants, so long as the populations are sufficiently dissimilar. It is not clear that
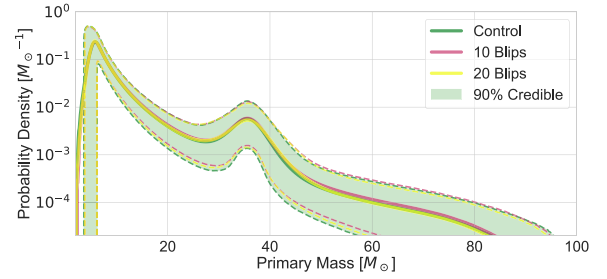
Figure 7. The inferred astrophysical mass distribution. In green we show the control run, with no contaminants injected and no glitch model included. We also show the runs with with the glitch model included and injected contaminants; we show runs with 10 and 20 blips included. The solid line is the PPD and the dashed lines show the upper and lower limits on the 90 per cent credible region. The inferred distributions appear consistent.

the correctly recovering the number of contaminants prevents slight biases from arising in the population inference, especially given there is some small variability in the inferred number of contaminants in the catalogue.

3.3 Biases in the BBH population

While the correct number of blips is recovered in each run, we want to be sure that no biases are introduced in the inferred astrophysical distributions. For example, we show inferred distributions of the primary masses for a control run and with 10 and 20 injected blips in Fig. 7. Qualitatively speaking, they appear to be essentially identical. The control run is a population inference on the catalogue of 69 BBHs in Abbott et al. (2023b), using the same astrophysical population model parametrization described in Section 2.4.

We quantify any differences by calculating the Jensen–Shannon (JS) divergence between the inferred distributions of a control population inference and the inferred astrophysical sub-populations from contaminated catalogues. The JS divergences show no trends, with a median consistently at ~0.09−0.1 bits. We show the JS divergences in the middle column of Table A3 in the Appendix, and in the first row we show the JS divergences between two draws from the control hyperparameters.

3.4 Biases from unmodelled blip contaminants

Some glitches appear significantly more astrophysical than others. For the run with 20 blips injected and the 69 BBH mergers, we plot the posteriors on the effective 'BBH' parameters of the glitches, and population-averaged p_{astro} values overlaid on the blip PPD, see Fig. 8.

There are some general patterns, most notably that extreme χ_{eff} seems to be the strongest predictor of low p_{astro}, and if the primary mass m_1 falls above the maximum mass cutoff m_{max} in the astrophysical model, the p_{astro} is zero. We show a table of the median and 90 per cent credible region parameters of each blip, along with the SNR and p_{astro} in Table A2 in the Appendix.

We want to understand the kind of biases, which are induced by including blips into the population, without controlling for those contaminants with a glitch model. Of the run with 20 injected blips and 69 GWs, we select the blips, which could most plausibly be astrophysical, i.e. they have the highest p_{astro}. We selected the blip with the highest p_{astro} (the top row in Table A2), and the 10 blips with the highest p_{astro} (the top 10 rows in Table A2), and contaminated
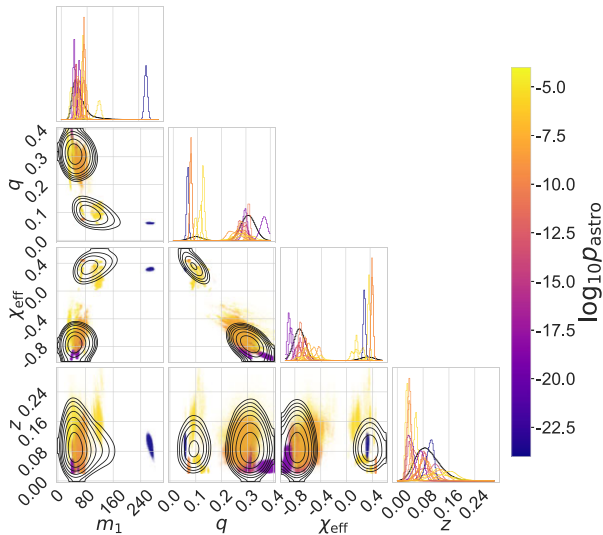
Figure 8. We show the 20 posteriors on the blip effective BBH parameters injected, and their corresponding mean p_{astro}, labelled in the figure by the colour of the posterior points. See the colourbar on the right. Note some general patterns: very low χ_{eff} values and very high masses correspond to low p_{astro} values. Note also that all the blip p_{astro} values are still very low, less than 10^{-4}.
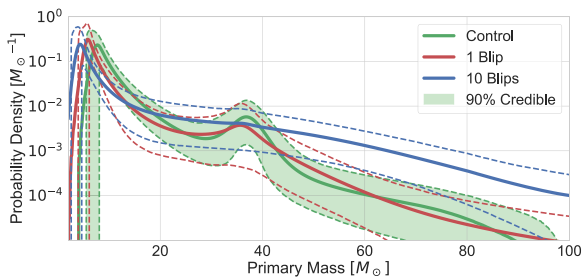
Figure 9. The inferred mass distribution for a control run compared to the inferred mass distribution when 1 and 10 astrophysically plausible blips are included into the catalogue, without controlling for their bias with a glitch model. Note the increased support at high mass, and the broadening of the Gaussian peak. The low mass end of the distribution is much less affected.

the catalogue of 69 BBH mergers passing the LVK selection criteria (Abbott et al. 2023b) with these 1 and 10 blips. We then sample from the population hyper-posterior without any glitch model.

In order to prevent population hyper-parameters from railing against prior ranges, we extended the prior range of m_{max} significantly (the maximum cutoff mass parameter in the model of Talbot & Thrane 2018) to allow values up to 500 $M_\odot$.

All the inferred distributions are biased. For instance, we show the inferred primary mass distribution for the control run, and for 1 and 10 contaminants, see Fig. 9.

We compute the JS divergences for these inferred distributions, compared to the control distribution. We show them in the right-hand column of Table A3 in the Appendix.

4 CONCLUSION AND FUTURE WORK

In this article, we presented a method for inference of a population of GW sources, which is contaminated by non-astrophysical events.

We contaminated the catalogue of 69 BBHs of Abbott et al. (2023b) with an increasing number of single-interferometer blip glitches from Ashton et al. (2022). We showed how to generalize a population inference to not only infer the shape parameters of a GW population, but to simultaneously infer the population of the glitch background events. We tested this method, and showed that it in practice identifies and removes systematic biases from population inference. As GW astronomy matures, interesting results may reveal themselves only on the level of populations, and satisfactory statistical significance may require delving into sub-threshold events.

As a proof of principle analysis, we chose only to consider the blip glitch class from GRAVITYSPY, since Ashton et al. (2022) had already produced parameter estimation samples for these. We caution that the method we presented here will only be robust to blip glitch contamination; we leave it to a future study to do a full simultaneous analysis with a model for an extended population of glitches.

There is another caveat, in the appropriate estimation of the selection effects. In an end-to-end analysis, the detection criterion is the same for glitches and GWs, and so must be estimated consistently. The current most common method requires a massive set of simulated GWs from a population similar to the population of astrophysical GWs into detector noise, and re-weighting for different population hyper-parameters (LIGO Scientific Collaboration, Virgo Collaboration & KAGRA Collaboration 2021; Abbott et al. 2023b). The set of glitches comes from regions of parameter space poorly sampled by the injection set, and so to properly estimate the selection effects, one needs an auxiliary suite of injections over the appropriate regions of parameter space. This is a significant computational expense, although it is regularly done by the LVK collaboration to estimate the selection effects of astrophysical GWs.

Though it is a challenge, there are many applications for a method to simultaneously infer the population of astrophysical GWs and non-astrophysical glitches. The most immediate application would be to lower the threshold for including a trigger into the catalogue, e.g. select on FAR $<2yr^{-1}$, or FAR $<5yr^{-1}$. There are real GW events lurking below the FAR $< 1yr^{-1}$ threshold, and these can aid in constraining the population. This would require an accurate model for the glitches that actually pass the threshold, rather than using our fiducial blip glitch model, and while conceptually similar to this work, the full treatment would also require running end-to-end search pipelines on injections from the glitch population. We leave this to a future study. There are other useful applications as well. Some GWs occur while only a single detector is online (Callister et al. 2017; Nitz et al. 2020; Cabourn Davies & Harry 2022). These single detector events often cannot enter a catalogue for population inference, and so they cannot be used for constraining the population. Our approach of modelling the intrinsic population of glitches is a step towards the use of single detector triggers in population analyses.

This method can also help characterize triggers found in searches for exotic objects. As an example, BBHs beyond the upper mass gap remain elusive (Ezquiaga & Holz 2021). The search sensitivity for these objects is reduced by the presence of short duration glitches much like blip glitches (Cabero et al. 2019), and so a joint analysis of a population of these background glitches and the astrophysical 'beyond-the-gap' BBHs would measure tighter constraints on their rates. As another example, an analogous procedure is conceivable for continuous wave (CW) sources. One may be able to characterize the population of CWs and the 'glitches' associated, which are due to monochromatic coherent power between detectors (Abbott et al. 2020, 2022b, c; Cieślar et al. 2021). This may benefit a search for CWs or population level characterization of CW sources.

For analyses like the one presented, it is critical to have both an accurate waveform model for glitches and an accurate glitch population model. In this paper, we model glitches with a GW waveform, however, it may be useful to use alternative glitch waveforms. One option is to use non-coherent GW waveforms to model the glitches, where the signal in each interferometer is fit with independent GW waveforms (Veitch & Vecchio 2010). One can also use non-GW waveform models, such as GLITSCHEN (Merritt et al. 2021) or BAYESWAVE (Cornish & Littenberg 2015). In cases where the glitch waveform model is different from the GW waveform, equation (9) must be used in its more general form. Second, we must have an appropriate model for the glitch population, and using as accurate as possible a model will be crucial. For example, if one continues to use a coherent GW waveform, one could fold in the analysis information about extrinsic parameters, e.g. the fact that the population of glitches is not expected to be isotropic (Payne et al. 2020; Vitale, Biscoveanu & Talbot 2022; Essick et al. 2023). We plan to explore both these avenues in a future work.

ACKNOWLEDGEMENTS

5 DATA AVAILABILITY

The data underlying this work can be found on a zenodo data release at https://doi.org/10.5281/zenodo.7860652. The public gravitational wave data can be found in the Gravitational Wave Open Science Center (Abbott et al. 2023a).

REFERENCES

Abbott B. P. et al., 2016, Phys. Rev. Lett., 116, 061102
Abbott B. P. et al., 2019a, Phys. Rev. X, 9, 031040
Abbott B. P. et al., 2019b, Phys. Rev. Lett., 123, 161102
Abbott R. et al., 2020, ApJ, 902, L21
Abbott R. et al., 2021, preprint (arXiv:2108.01045)
Abbott R. et al., 2022a, preprint (arXiv:2212.01477)
Abbott R. et al., 2022b, Phys. Rev. D, 106, 102008
Abbott R. et al., 2022c, ApJ, 935, 1
Abbott R. et al., 2023a, preprint (arXiv:2302.03676)
Abbott R. et al., 2023b, Phys. Rev. X, 13, 011048
Acernese F. et al., 2022a, preprint (arXiv:2210.15633)
Acernese F. et al., 2022b, preprint (arXiv:2210.15634)
Akutsu T. et al., 2021, Progr. Theor. Exp. Phys., 2021, 05A102
Ali-Haïmoud Y., Kovetz E. D., Kamionkowski M., 2017, Phys. Rev. D, 96, 123523
Allen B., 2005, Phys. Rev. D, 71, 062001
Antonini F., Toonen S., Hamers A. S., 2017, ApJ, 841, 77
Ashton G., Thrane E., 2020, MNRAS, 498, 1905
Ashton G., Thrane E., Smith R. J. E., 2019, Phys. Rev. D, 100, 123018
Ashton G., Thiele S., Lecoeuche Y., McIver J., Nuttall L. K., 2022, Class. Quant. Grav., 39, 175004
Bird S., Cholis I., Muñoz J. B., Ali-Haïmoud Y., Kamionkowski M., Kovetz E. D., Raccanelli A., Riess A. G., 2016, Phys. Rev. Lett., 116, 201301
Bohé A., Hannam M., Husa S., Ohme F., Pürrer M., Schmidt P., 2016, PhenomPv2-Technical notes for the LAL implementation. LIGO Scientific Collaboration, Washington
Cabero M. et al., 2019, Class. Quant. Grav., 36, 155010
Cabourn Davies G. S., Harry I. W., 2022, Class. Quant. Grav., 39, 215012
Callister T. A., Kanner J. B., Massinger T. J., Dhurandhar S., Weinstein A. J., 2017, Class. Quant. Grav., 34, 155007
Callister T. A., Haster C.-J., Ng K. K. Y., Vitale S., Farr W. M., 2021, ApJL, 922, L5
Chen H.-Y., Holz D. E., 2014, The loudest gravitational wave events. American Physical Society, College Park
Christensen N., Meyer R., 2022, Rev. Mod. Phys., 94, 025001
Cieślar M., Bulik T., Curyło M., Sieniawska M., Singh N., Bejger M., 2021, A&A, 649, A92
Cornish N. J., Littenberg T. B., 2015, Class. Quant. Grav., 32, 135012
Davis D. et al., 2021, Class. Quant. Grav., 38, 135014
Essick R., Farr W., 2022, preprint (arXiv:2204.00461)
Essick R., Farr W. M., Fishbach M., Holz D. E., Katsavounidis E., 2023, Phys. Rev. D, 107, 043016
Ezquiaga J. M., Holz D. E., 2021, ApJ, 909, L23
Farr W. M., 2019, Res. Notes AAS, 3, 66
Farr W. M., Gair J. R., Mandel I., Cutler C., 2015, Phys. Rev. D, 91, 023005
Fishbach M., Holz D. E., Farr W. M., 2018, ApJL, 863, L41
Gaebel S. M., Veitch J., Dent T., Farr W. M., 2019, MNRAS, 484, 4008
Galaudage S., Talbot C., Thrane E., 2020, Phys. Rev. D, 102, 083026
Glanzer J. et al., 2023, Class. Quant. Grav., 40, 065004
Golomb J., Talbot C., 2022, preprint (arXiv:2210.12287)
Hanna C. et al., 2020, Phys. Rev. D, 101, 022003
Hannam M., Schmidt P., Bohé A., Haegel L., Husa S., Ohme F., Pratten G., Pürrer M., 2014, Phys. Rev. Lett., 113, 151101
Huang Y., Middleton H., Ng K. K. Y., Vitale S., Veitch J., 2018, Phys. Rev. D, 98, 123021
Husa S., Khan S., Hannam M., Pürrer M., Ohme F., Forteza X. J., Bohé A., 2016, Phys. Rev. D, 93, 044006
Isi M., Smith R., Vitale S., Massinger T. J., Kanner J., Vajpeyi A., 2018, Phys. Rev. D, 98, 042007
Ivanova N. et al., 2013, A&A Rev., 21, 59
Kapadia S. J. et al., 2020, Class. Quant. Grav., 37, 045007
Khan S., Husa S., Hannam M., Ohme F., Pürrer M., Forteza X. J., Bohé A., 2016, Phys. Rev. D, 93, 044007
Koposov S. et al., 2022, joshspeagle/dynesty: v2.0.3. Zenodo.
Kulkarni S. R., Hut P., McMillan S., 1993, Nature, 364, 421

LIGO Scientific Collaboration, Virgo Collaboration, KAGRA Collaboration, 2021, GWTC-3: Compact binary coalescences observed by LIGO and Virgo during the second part of the third observing run – O1+O2 + O3 search sensitivity estimates [Data set]. Zenodo

Mandel I., Broekgaarden F. S., 2022, Liv. Rev. Relat., 25, 1

Mandel I., de Mink S. E., 2016, MNRAS, 458, 2634

Mandel I., Farr W. M., Gair J. R., 2019, MNRAS, 486, 1086

Mapelli M., 2021, in Bambi C., Katsanevas S., Kokkotas K., eds, Handbook of Gravitational Wave Astronomy. Springer Nature, Singapore, p. 705

Marchant P., Langer N., Podsiadlowski P., Tauris T. M., Moriya T. J., 2016, A&A, 588, A50

Merritt J. D., Farr B., Hur R., Edelman B., Doctor Z., 2021, Phys. Rev. D, 104, 102004

Messick C. et al., 2017, Phys. Rev. D, 95, 042001

Miller S., Callister T. A., Farr W. M., 2020, ApJ, 895, 128

Ng K. K. Y., Vitale S., Zimmerman A., Chatziioannou K., Gerosa D., Haster C.-J., 2018, Phys. Rev. D, 98, 083007

Nitz A. H., Wang Y.-F., 2021, ApJ, 915, 54

Nitz A. H., Dent T., Dal Canton T., Fairhurst S., Brown D. A., 2017, ApJ, 849, 118

Nitz A. H., Dent T., Davies G. S., Harry I., 2020, ApJ, 897, 169

Payne E., Banagiri S., Lasky P. D., Thrane E., 2020, Phys. Rev. D, 102, 102004

Portegies Zwart S. F., McMillan S. L. W., 2000, ApJ, 528, L17

Pratten G., Vecchio A., 2021, Phys. Rev. D, 104, 124039

Pürrer M., Hannam M., Ohme F., 2016, Phys. Rev. D, 93, 084042

Racine E., 2008, Phys. Rev. D, 78, 044021

Roulet J., Zaldarriaga M., 2019, MNRAS, 484, 4216

Roulet J., Venumadhav T., Zackay B., Dai L., Zaldarriaga M., 2020, Phys. Rev. D, 102, 123022

Schutz B. F., 2011, Class. Quant. Grav., 28, 125023

Shandera S., Jeong D., Grasshorn Gebhardt H. S., 2018, Phys. Rev. Lett., 120, 241102

Sigurdsson S., Hernquist L., 1993, Nature, 364, 423

Silsbee K., Tremaine S., 2017, ApJ, 836, 39

Smarr L. L., Blandford R., 1976, ApJ, 207, 574

Soni S. et al., 2021, Class. Quant. Grav., 38, 195016

Speagle J. S., 2020, MNRAS, 493, 3132

Talbot C., Thrane E., 2018, ApJ, 856, 173

Talbot C., Smith R., Thrane E., Poole G. B., 2019, Phys. Rev. D, 100, 043030

Abbott R. et al. 2021, preprint (arXiv:2111.03606)

Thrane E., Talbot C., 2019, PASA, 36, e010

Tiwari V., 2018, Class. Quant. Grav., 35, 145009

Tolley A. E., Cabourn Davies G. S., Harry I. W., Lundgren A. P., 2023, preprint (arXiv:2301.10491)

Tutukov A. V., Yungelson L. R., 1993, MNRAS, 260, 675

Udall R. P., Davis D., 2023, Appl. Phys. Lett., 122, 094103

Usman S. A. et al., 2016, Class. Quant. Grav., 33, 215004

Vajpeyi A. et al., 2022, MNRAS, 516, 5309

Van Den Heuvel E. P. J., 1976, Symp. Int. Astron. Union, 73, 35

Van Den Heuvel E. P. J., Portegies Zwart S. F., de Mink S. E., 2017, MNRAS, 471, 4256

Veitch J., Vecchio A., 2010, Phys. Rev. D, 81, 062003

Veitch J. et al., 2015, Phys. Rev. D, 91, 042003

Vitale S., Lynch R., Raymond V., Sturani R., Veitch J., Graff P., 2017, Phys. Rev. D, 95, 064053

Vitale S., Gerosa D., Farr W. M., Taylor S. R., 2020, Inferring the Properties of a Population of Compact Binaries in Presence of Selection Effects. Springer, Singapore, p. 1

Vitale S., Biscoveanu S., Talbot C., 2022, preprint (arXiv:2204.00968)

Whittle P., 1951, Hypothesis Testing in Time Series Analysis. Statistics/Uppsala Universitet. Almqvist and Wiksells Boktryckeri AB, Oxford

Zevin M. et al., 2017, Class. Quant. Grav., 34, 064003

Zevin M. et al., 2021, ApJ, 910, 152

APPENDIX A

Event information

We show upper bounds on the calculated p_{astro} for each GW event when we included 20 contaminant blips in Table A1. We also show upper bounds on each p_{astro}, given by the 90 per cent and 99 per cent upper bounds. In Table A2, we show the parameters of the 20 blips that contaminate the catalogue in the run with 20 blips. Note we use a random set of blips for each catalogue, e.g. the 19 contaminants for the run with 19 blips are not a subset of the 20 contaminants for the run with 20 blips.

Optimal detectable mixing fraction posterior

Consider the scenario where the populations are disjoint such that every event posterior uniquely determines which population the event originates from. The event parameters tell us with no ambiguity whether an event is a glitch or a GW. Therefore, we want to infer the relative rate of detectable events given we detected N_{events}, with k unambiguous astrophysical events, and the rest unambiguous glitches. This is a common problem in Bayesian inference and it admits an analytical posterior, given by equation (A1). This is the best the inference could possibly constrain $\overline{\eta}$, and so it is a useful benchmark to compare to.

$$p(\overline{\eta}) = \frac{\overline{\eta}^k (1 - \overline{\eta})^{N_{events} - k}}{B(k + 1, N_{events} - k + 1)}, \tag{A1}$$

which assumes a uniform prior in $\overline{\eta}$ from 0 to 1, and the denominator is a normalization.
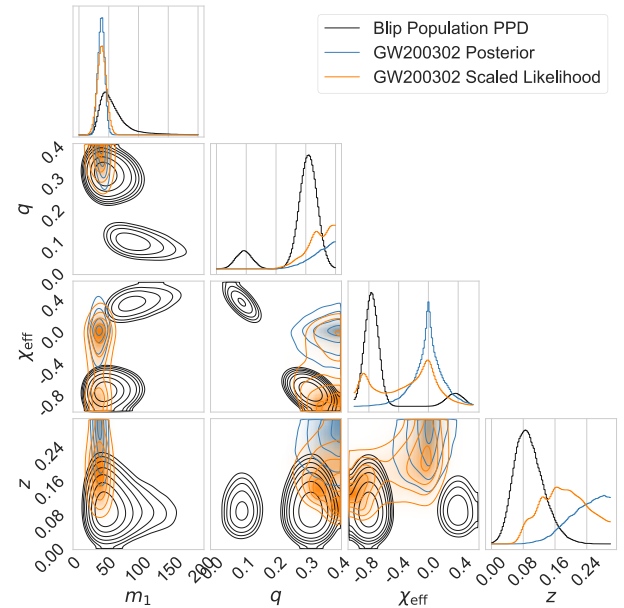
Figure A1. A corner plot with the primary mass, mass ratio, effective aligned spin, and redshift posterior of GW200302 overlaid on the population predictive distribution of the blip population. The posterior on GW200302 is shown in blue, with the first 4σ contours and sample points. The blip population predictive distribution is in black contours, showing the first 7σ. Since it is not the posterior overlap but the likelihood overlap which contributes to the p_{blip}, we include the posterior reweighted by the inverse of the prior. This highlights the regions of high overlap for the glitch likelihood term.

Table A1. Inferred $1 - p_{astro} = p_{blip}$ for each event in the catalogue of Abbott et al. ([2023b](#)), calculated from the run with 20 injected blips. We show the upper bounds on the inferred p_{blip} at both 90 per cent and 99 per cent credence. Note GW200302 has $p_{blip} \lesssim 7.2$ per cent, the highest non-astrophysical probability event, and the event GW190503_185404 has the second highest $p_{blip} \lesssim 7.0$ per cent.

Event	90 per cent upper bound	99 per cent upper bound	Event	90 per cent upper bound	99 per cent upper bound
GW150914	0	0	GW190731_140936	6.7×10^{-16}	2.3×10^{-13}
GW151012	3.1×10^{-4}	6.1×10^{-4}	GW190803_022701	2.3×10^{-10}	6.8×10^{-9}
GW151226	1.2×10^{-5}	5.9×10^{-5}	GW190805_211137	0	0
GW170104	1.3×10^{-6}	4.8×10^{-6}	GW190828_063405	0	0
GW170608	7.3×10^{-8}	4.4×10^{-7}	GW190828_065509	3.7×10^{-9}	3.4×10^{-8}
GW151226	9.7×10^{-4}	2.9×10^{-3}	GW190910_112807	0	0
GW170809	2.6×10^{-10}	4.4×10^{-9}	GW190915_235702	2.3×10^{-11}	3.6×10^{-10}
GW170814	0	0	GW190924_021846	8.7×10^{-8}	1.1×10^{-6}
GW170818	0	5.1×10^{-15}	GW190925_232845	1.7×10^{-7}	8.6×10^{-7}
GW170823	1.6×10^{-7}	9.0×10^{-7}	GW190929_012149	1.3×10^{-7}	1.6×10^{-6}
GW190408_181802	0	0	GW190930_133541	8.2×10^{-11}	8.4×10^{-10}
GW190412	7.2×10^{-5}	2.8×10^{-4}	GW191103_012549	1.8×10^{-8}	1.6×10^{-7}
GW190413_052954	1.8×10^{-9}	3.3×10^{-8}	GW191105_143521	4.2×10^{-9}	3.2×10^{-8}
GW190413_134308	3.2×10^{-14}	1.7×10^{-12}	GW191109_010717	6.0×10^{-5}	4.6×10^{-4}
GW190421_213856	1.9×10^{-10}	4.0×10^{-9}	GW191127_050227	1.8×10^{-4}	7.7×10^{-4}
GW190503_185404	4.0×10^{-2}	7.0×10^{-2}	GW191129_134029	1.6×10^{-11}	2.3×10^{-10}
GW190512_180714	2.0×10^{-15}	1.4×10^{-13}	GW191204_171526	0	0
GW190513_205428	4.0×10^{-13}	2.8×10^{-11}	GW191215_223052	0	2.2×10^{-16}
GW190517_055101	1.1×10^{-8}	3.6×10^{-7}	GW191216_213338	3.2×10^{-6}	1.8×10^{-5}
GW190519_153544	0	0	GW191222_033537	3.6×10^{-11}	1.1×10^{-9}
GW190521	5.6×10^{-16}	4.9×10^{-14}	GW191230_180458	3.4×10^{-7}	4.3×10^{-6}
GW190521_074359	0	0	GW200112_155838	0	0
GW190527_092055	1.8×10^{-9}	2.1×10^{-8}	GW200128_022011	2.2×10^{-16}	2.0×10^{-14}
GW190602_175927	1.8×10^{-10}	1.2×10^{-8}	GW200129_065458	4.8×10^{-9}	4.7×10^{-8}
GW190620_030421	2.4×10^{-12}	1.8×10^{-10}	GW200202_154313	1.9×10^{-7}	1.1×10^{-6}
GW190630_185205	0	0	GW200208_130117	2.0×10^{-7}	1.7×10^{-6}
GW190701_203306	4.8×10^{-12}	1.6×10^{-10}	GW200209_085452	7.7×10^{-9}	1.1×10^{-7}
GW190706_222641	2.3×10^{-15}	5.4×10^{-13}	GW200216_220804	4.1×10^{-7}	2.3×10^{-6}
GW190707_093326	4.6×10^{-11}	4.8×10^{-10}	GW200219_094415	1.3×10^{-12}	1.1×10^{-10}
GW190708_232457	0	6.7×10^{-16}	GW200224_222234	1.2×10^{-6}	5.4×10^{-6}
GW190719_215514	7.5×10^{-12}	2.3×10^{-10}	GW200225_060421	4.7×10^{-8}	3.7×10^{-7}
GW190720_000836	1.4×10^{-10}	1.4×10^{-9}	GW200302_015811	4.1×10^{-2}	7.2×10^{-2}
GW190725_174728	2.3×10^{-6}	8.5×10^{-6}	GW200311_115853	4.9×10^{-4}	1.4×10^{-3}
GW190727_060333	0	0	GW200316_215756	1.6×10^{-8}	1.2×10^{-7}
GW190728_064510	4.8×10^{-12}	7.4×10^{-11}			

In fact, we can see how this arises directly from equation (12). If the population models for the glitches and the BBHs are completely disjoint for all event posteriors, then in each term in the product of equation (12), either the glitch term $\int d\psi \mathcal{L}(d_i|\psi) p_G(\psi|\Lambda_G)$ or the astrophysical term $\int d\theta \mathcal{L}(d_i|\theta) p_A(\theta|\Lambda_A)$ will vanish. The likelihood then factorizes:

$$\mathcal{L}(\{d_i\}|\Lambda_A, \Lambda_G, \eta) \propto \overline{\eta}^k(1 - \overline{\eta})^{N_{events}-k}$$
$$\prod_{i=1}^{k} \frac{\int d\theta \mathcal{L}(d_i|\theta) p_A(\theta|\Lambda_A)}{\alpha_A(\Lambda_A)} \prod_{i=k+1}^{N_{events}} \frac{\int d\psi \mathcal{L}(d_i|\psi) p_G(\psi|\Lambda_G)}{\alpha_G(\Lambda_G)},$$
$$\text{(A2)}$$

and so the inference may proceed independently for the astrophysical hyper-parameters Λ_A, the glitch hyper-parameters Λ_G, and the detectable mixing fraction $\overline{\eta}$. This matches the intuitive result that independent populations may be characterized independently. Pulling out the $\overline{\eta}$ term in the likelihood and normalizing with a uniform prior between 0 and 1, we recover equation (A1).

In general, the glitch and BBH populations are not completely disjoint and the glitch/astrophysical terms in the product in equation (12) do not vanish. With the additional uncertainty in the 'identity' of each event in the catalogue, the posterior on $\overline{\eta}$ will broaden. The degree of broadening tells us how close the inferences are coming towards knowing there are exactly k BBHs of N_{events} total events.

GW200302

GW200302 has the largest support for $1 - - - p_{astro} = p_{blip}$. To understand this, we show the corner plot overlay of the GW200302 posterior and the blip population predictive distribution in Fig. A1. Note the tails of the GW200302 posterior overlaps with the blip population distribution; this is why the p_{astro} for GW200302 is relatively low.

Further, it is not the posterior 'overlap' which is taken into the population likelihood, but the event likelihood 'overlap'. The tails of the posterior in χ_{eff} and in mass ratio q—the samples, which happen to fall neatly into the blip population—are therefore weighted much more highly, since the sampling prior there is much lower. Indeed, the χ_{eff} posterior is essentially the recovered χ_{eff} sampling prior. This drives up the glitch population term in equation (20), and therefore lowers the p_{astro}. This is expected: if there are poor constraints on the source parameters, we must be more agnostic about the event's origin based on the parameters alone.

Table A2. Median and 90 per cent credible intervals for the GW effective parameters of the 20 blip contaminants for the 20 injection run, organized by the median $p_{\rm astro}$. The SNR is the optimal SNR.

Number	m_1	q	$\chi_{\rm eff}$	z	SNR	Median $p_{\rm astro}$	$p_{\rm astro}$ 5 per cent	$p_{\rm astro}$ 95 per cent
1	$49.9^{+12.0}_{-5.5}$	$0.28^{+0.05}_{-0.05}$	$-0.64^{+0.28}_{-0.09}$	$0.12^{+0.07}_{-0.06}$	$10.9^{+1.7}_{-1.8}$	1.0×10^{-5}	5.2×10^{-6}	2.0×10^{-5}
2	$31.2^{+12.1}_{-4.1}$	$0.3^{+0.1}_{-0.1}$	$-0.58^{+0.23}_{-0.22}$	$0.13^{+0.06}_{-0.04}$	$8.62^{+1.83}_{-1.89}$	5.8×10^{-6}	2.7×10^{-6}	1.4×10^{-5}
3	$65.1^{+5.6}_{-19.3}$	$0.11^{+0.2}_{-0.02}$	$0.19^{+0.12}_{-0.95}$	$0.058^{+0.01}_{-0.008}$	$15.4^{+1.7}_{-1.7}$	8.3×10^{-7}	3.9×10^{-8}	1.1×10^{-5}
4	$66.6^{+0.8}_{-2.0}$	$0.12^{+0.01}_{-0.0}$	$0.38^{+0.0}_{-0.02}$	$0.031^{+0.007}_{-0.006}$	$27.5^{+1.6}_{-1.7}$	4.2×10^{-8}	3.1×10^{-11}	5.9×10^{-6}
5	$56.4^{+4.8}_{-3.5}$	$0.26^{+0.03}_{-0.04}$	$-0.62^{+0.13}_{-0.07}$	$0.093^{+0.033}_{-0.029}$	$13.0^{+1.7}_{-1.7}$	1.3×10^{-8}	1.0×10^{-10}	2.7×10^{-7}
6	$116.0^{+9.8}_{-40.5}$	$0.094^{+0.175}_{-0.013}$	$0.15^{+0.21}_{-0.73}$	$0.16^{+0.07}_{-0.05}$	$10.5^{+1.7}_{-1.8}$	5.6×10^{-9}	3.5×10^{-13}	1.1×10^{-5}
7	$51.4^{+21.0}_{-2.9}$	$0.28^{+0.02}_{-0.03}$	$-0.67^{+0.17}_{-0.06}$	$0.082^{+0.052}_{-0.031}$	$14.5^{+1.7}_{-1.8}$	2.9×10^{-9}	4.8×10^{-15}	2.6×10^{-8}
8	$76.8^{+5.2}_{-4.7}$	$0.28^{+0.04}_{-0.04}$	$-0.54^{+0.11}_{-0.12}$	$0.13^{+0.04}_{-0.05}$	$15.5^{+1.7}_{-1.7}$	1.1×10^{-10}	1.8×10^{-18}	5.2×10^{-8}
9	$45.9^{+4.4}_{-1.7}$	$0.31^{+0.01}_{-0.01}$	$-0.75^{+0.03}_{-0.04}$	$0.051^{+0.016}_{-0.017}$	$21.2^{+1.6}_{-1.7}$	3.0×10^{-12}	2.7×10^{-25}	1.2×10^{-10}
10	$61.1^{+4.8}_{-7.4}$	$0.25^{+0.05}_{-0.04}$	$-0.51^{+0.11}_{-0.28}$	$0.083^{+0.025}_{-0.029}$	$14.8^{+1.7}_{-1.7}$	3.4×10^{-13}	1.1×10^{-20}	6.9×10^{-9}
11	$69.4^{+5.3}_{-7.0}$	$0.27^{+0.03}_{-0.03}$	$-0.72^{+0.16}_{-0.08}$	$0.079^{+0.016}_{-0.03}$	$20.6^{+1.7}_{-1.6}$	9.7×10^{-16}	2.5×10^{-23}	2.5×10^{-12}
12	$71.9^{+2.9}_{-1.4}$	$0.074^{+0.003}_{-0.003}$	$0.44^{+0.01}_{-0.01}$	$0.037^{+0.005}_{-0.008}$	$36.0^{+1.7}_{-1.7}$	2.0×10^{-16}	0	1.0×10^{-10}
13	$69.8^{+1.2}_{-1.3}$	$0.073^{+0.001}_{-0.002}$	$0.44^{+0.0}_{-0.0}$	$0.038^{+0.017}_{-0.011}$	$26.7^{+1.6}_{-1.7}$	2.9×10^{-17}	0	2.9×10^{-11}
14	$43.7^{+11.8}_{-2.1}$	$0.3^{+0.0}_{-0.0}$	$-0.67^{+0.06}_{-0.11}$	$0.084^{+0.019}_{-0.019}$	$15.1^{+1.7}_{-1.7}$	8.3×10^{-21}	8.0×10^{-32}	3.7×10^{-15}
15	$57.3^{+6.1}_{-4.4}$	$0.28^{+0.04}_{-0.04}$	$-0.77^{+0.15}_{-0.08}$	$0.11^{+0.03}_{-0.04}$	$14.3^{+1.7}_{-1.7}$	5.2×10^{-21}	6.3×10^{-32}	2.2×10^{-15}
16	$52.9^{+11.2}_{-6.8}$	$0.29^{+0.03}_{-0.03}$	$-0.77^{+0.13}_{-0.09}$	$0.088^{+0.045}_{-0.041}$	$14.2^{+1.7}_{-1.7}$	7.2×10^{-22}	4.0×10^{-33}	4.7×10^{-16}
17	$45.1^{+2.4}_{-1.9}$	$0.29^{+0.02}_{-0.02}$	$-0.71^{+0.07}_{-0.06}$	$0.041^{+0.023}_{-0.013}$	$17.9^{+1.7}_{-1.7}$	3.0×10^{-26}	2.2×10^{-42}	5.3×10^{-18}
18	$56.3^{+2.6}_{-2.3}$	$0.3^{+0.0}_{-0.0}$	$-0.9^{+0.1}_{-0.0}$	$0.07^{+0.02}_{-0.03}$	$19.3^{+1.7}_{-1.7}$	2.2×10^{-36}	6.8×10^{-60}	9.0×10^{-25}
19	$40.3^{+1.7}_{-1.5}$	$0.37^{+0.02}_{-0.03}$	$-0.94^{+0.05}_{-0.03}$	$0.039^{+0.014}_{-0.012}$	$22.9^{+1.6}_{-1.7}$	4.1×10^{-45}	2.0×10^{-77}	7.3×10^{-29}
20	$262.0^{+5.7}_{-5.5}$	$0.06^{+0.0}_{-0.0}$	$0.31^{+0.01}_{-0.01}$	$0.1^{+0.0}_{-0.0}$	$19.9^{+1.7}_{-1.7}$	0	0	0

JS divergences

We show the JS divergences measured between the inferred astrophysical sub-populations of each run, where we included the glitch model to account for the injected contaminants. We also show the JS divergence for 1 and 10 injected blips where we did not include the glitch model.

Calculating the inferred number of events in the catalogue

In equation (21), we show an expression for the probability on the number of events in the catalogue, depending on the population hyper-parameters Λ. However, this expression is a sum of $\mathcal{O}(10^{20})$ terms, and as such is not computationally feasible to evaluate. Fortunately, there is a much more efficient method to complete the sum.

The sum has a largest term, which we can easily find by first ordering the list of $p_{{\rm astro},\,i}$ from largest to smallest. This term then corresponds to the identity k-combination, denoted $\gamma_0 = I$. We notate this term by $p^{(0)} = \prod_{i=1}^{k} p_{{\rm astro},\,i} \prod_{i=k+1}^{N_{\rm events}} (1 - p_{{\rm astro},\,i})$. Because many of the $p_{\rm astro}$ posteriors have non-negligible posterior width, this term does not completely dominate the entire sum, however, we can express all the other k-combinations in the sum in terms of this $p^{(0)}$. In particular, we can think of each $\gamma \in \Gamma(k, N_{\rm events})$ in terms of the number r of events, which must be exchanged from the astrophysical bin to the glitch bin in order to match γ_0. Because we are summing unique k-combinations, the probability associated with the family of k-combinations which are r events different from γ_0 is given by

$$p^{(r)} = p^{(0)} e_r \left(\frac{1 - p_{{\rm astro},1}}{p_{{\rm astro},1}}, ..., \frac{1 - p_{{\rm astro},k}}{p_{{\rm astro},k}} \right)$$
$$\times e_r \left(\frac{p_{{\rm astro},k+1}}{1 - p_{{\rm astro},k+1}}, ..., \frac{p_{{\rm astro},N_{\rm events}}}{1 - p_{{\rm astro},N_{\rm events}}} \right), \qquad \text{(A3)}$$

where the e_r is the rth symmetric polynomial. Symmetric polynomials are defined such that every term has degree r and every r combination of the variables appears once in the sum, e.g. $e_2(x, y, z) = xy + xz + yz$. Note how symmetric polynomials naturally capture the idea of summing over unique sets. The first polynomial term is the sum over all unique sets of size r sending events from the astrophysical bin to the glitch bin. The second polynomial term is similar, sending all unique sets of size r from the glitch bin to the astrophysical bin. Their product, then, is the sum over all combinations of unique set exchanges of size r between the astrophysical and glitch bin.

This is nice, but it is not helpful unless one can rapidly evaluate the symmetric polynomials. It turns out that one can easily find the rth symmetric polynomial recursively from the previous $r-1$ symmetric polynomials, using Newton and Girard's Theorem:

$$r e_r(x_1, ..., x_n) = \sum_{j=1}^{r} (-1)^{j-1} e_{r-j}(x_1, ..., x_n) f_j(x_1, ..., x_n), \qquad \text{(A4)}$$

where the $f_j(x_1, ..., x_n) = x_1^j + ... + x_n^j$ are computationally trivial to evaluate. With this in hand, we can rapidly evaluate equation (21) as

$$p_k(\Lambda) = \sum_{r=0}^{\min(k, N_{\rm events}-k)} p^{(r)}. \qquad \text{(A5)}$$

If one wishes to calculate the k-expectation over the $p_k(\Lambda)$

$$\langle p_k(\Lambda) \rangle_k = \sum_{k=0}^{N_{\rm events}} k p_k(\Lambda), \qquad \text{(A6)}$$

it is simple enough to evaluate given all the $p_k(\Lambda)$, however, it is clear that this should also equal the sum of the $p_{{\rm astro},\,i}$, thinking of the $p_{{\rm astro},\,i}$ as independent Bernoulli trials. We can show they are equivalent by

Table A3. JS divergences in units of bits (base-2 logarithm) between the inferred distributions of a control run and the astrophysical sub-population of the simultaneous fitting runs. In the middle column are the runs, which include a glitch model, but in the right-hand column we show runs, which do not have a glitch model and so contaminants must be fitted with the astrophysical population. In the control row, we show the JS divergence posterior from two random samples from the hyper-parameter posterior in the control run. Note the consistency between each run in the middle column, in particular, the lack of any sort of (increasing) trend. In the right-hand column, notice that the JS divergences increase as expected.

N_{blip}	JS w/o glitch model (bits)	JS w/o glitch model (bits)
Control	$0.097^{+0.123}_{-0.063}$	$0.097^{+0.123}_{-0.063}$
0	$0.088^{+0.119}_{-0.057}$	–
1	$0.091^{+0.117}_{-0.058}$	$0.104^{+0.133}_{-0.065}$
2	$0.090^{+0.119}_{-0.059}$	–
3	$0.090^{+0.117}_{-0.059}$	–
4	$0.088^{+0.118}_{-0.057}$	–
5	$0.093^{+0.119}_{-0.059}$	–
6	$0.092^{+0.121}_{-0.060}$	–
7	$0.091^{+0.126}_{-0.059}$	–
8	$0.094^{+0.126}_{-0.059}$	–
9	$0.090^{+0.116}_{-0.057}$	–
10	$0.093^{+0.118}_{-0.059}$	$0.406^{+0.188}_{-0.132}$
11	$0.088^{+0.113}_{-0.056}$	–
12	$0.090^{+0.119}_{-0.057}$	–
13	$0.088^{+0.120}_{-0.056}$	–
14	$0.090^{+0.120}_{-0.058}$	–
15	$0.091^{+0.117}_{-0.059}$	–
16	$0.095^{+0.121}_{-0.060}$	–
17	$0.097^{+0.120}_{-0.062}$	–
18	$0.091^{+0.126}_{-0.057}$	–
19	$0.089^{+0.115}_{-0.056}$	–
20	$0.090^{+0.119}_{-0.057}$	–

writing down a generating polynomial for $p_k(\Lambda)$

$$\sum_{k=0}^{N_{events}} x^k p_k(\Lambda) = \prod_{i=1}^{N_{events}} \left[x p_{astro,i} + (1 - p_{astro,i}) \right]. \qquad (A7)$$

Evaluating the polynomial for $x = 1$ shows the $p_k(\Lambda)$ are indeed normalized, and evaluating the first derivative at $x = 1$ shows the k-expectation is equal to the sum of the $p_{astro,i}$.

One may be tempted to use the k-expectation as it has continuous support, however, we caution that using only the k-expectation can be somewhat misleading. For some population inferences, there was very little support for 69 GW events in the k-expectation posterior, while there was a reasonable probability for having exactly 69 GW events in the catalogue. These are different statistical statements and should not be mistaken for one another.

This kind of calculation can in principle be done for any population inference with a mixing fraction. That said, our populations are nearly disjoint and as such the posterior width on $\bar{\eta}$ is dominated by Poisson uncertainty, not uncertainty on which events in the catalogue should belong to which sub-populations. For other population inferences with mixing fractions, the events may not be as easy to differentiate into sub-populations, and the uncertainty on the mixing fraction will have a larger contribution from this uncertainty. The $p_k(\Lambda)$ will have broader support and will more closely mimic the (appropriately rescaled) detectable mixing fraction posterior.

This paper has been typeset from a T$_{E}$X/L^{A}T$_{E}$X file prepared by the author.