OPEN ACCESS
IOP Publishing
Classical and Quantum Gravity
Class. Quantum Grav. **39** (2022) 175004 (23pp)
https://doi.org/10.1088/1361-6382/ac8094

Parameterised population models of transient non-Gaussian noise in the LIGO gravitational-wave detectors

Gregory Ashton[1,2,*], **Sarah Thiele**[3], **Yannick Lecoeuche**[4], **Jess McIver**[3] and **Laura K Nuttall**[1]

[1] Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth PO1 3FX, United Kingdom
[2] Department of Physics, Royal Holloway, University of London, TW20 0EX, United Kingdom
[3] University of British Columbia, Vancouver, BC V6T 1Z4, Canada
[4] LIGO Hanford Observatory, Richland, WA 99352, United States of America

E-mail: gregory.ashton@ligo.org

Received 1 March 2022, revised 28 June 2022
Accepted for publication 12 July 2022
Published 9 August 2022

Abstract

The two interferometric LIGO gravitational-wave observatories provide the most sensitive data to date to study the gravitational-wave universe. As part of a global network, they have completed their third observing run in which they observed many tens of signals from merging compact binary systems. It has long been known that a limiting factor in identifying transient gravitational-wave signals is the presence of transient non-Gaussian noise, which reduce the ability of astrophysical searches to detect signals confidently. Significant efforts are taken to identify and mitigate this noise at the source, but its presence persists, leading to the need for software solutions. Taking a set of transient noise artefacts categorised by the GravitySpy software during the O3a observing era, we produce parameterised population models of the noise projected into the space of astrophysical model parameters of merging binary systems. We compare the inferred population properties of transient noise artefacts with observed astrophysical systems from the GWTC2.1 catalogue. We find that while the population of astrophysical systems tend to have near equal masses and moderate spins, transient noise artefacts are typically characterised by extreme mass

*Author to whom any correspondence should be addressed.

ratios and large spins. This work provides a new method to calculate the consistency of an observed candidate with a given class of noise artefacts. This approach could be used in assessing the consistency of candidates found by astrophysical searches (i.e. determining if they are consistent with a known glitch class). Furthermore, the approach could be incorporated into astrophysical searches directly, potentially improving the reach of the detectors, though only a detailed study would verify this.

Keywords: gravitational-waves, black holes, data analysis

(Some figures may appear in colour only in the online journal)

1. Introduction

The LIGO, Virgo, and KAGRA gravitational-wave detectors [1–3] have opened the door to the gravitational-wave universe and provided our first insights into the mergers of black holes and neutron stars. These kilometre-scale interferometers, sensitive to signals in the tens of Hz to kHz band, survey the sky nearly isotropically during *observing runs* [4]. Between each observing run, the detectors are improved, reducing the noise level and yielding improvements in the sensitivity. Therefore, each new observing run probes deeper than before, yielding an increase in the rate of detections.

The LIGO detectors contain noise which is often assumed to be stationary and Gaussian for analysis, but often exhibits non-stationarity and frequent non-Gaussian transient noise artefacts termed *glitches* [5–9]. All glitches are believed to be instrumental or environmental in origin. In some cases, the cause of a glitch class can be identified and improvements made to the detector to reduce or eradicate the class. In other cases, so-called witness channels can identify when such glitches occurred and veto coincident candidate events (see, e.g. Davis *et al* [9]). Finally, glitches in classes without a known origin or witness channel remain a feature of the science-mode data from interferometric gravitational-wave observatories.

Glitches are detrimental to achieving the scientific goals of gravitational-wave observatories. Suppose glitches occur in coincidence with an actual astrophysical signal. In that case, this may result in the signal being missed by searches, reducing the number of detectable signals, or biasing our astrophysical inferences about the signal itself. However, glitches that do not occur in coincidence with real signals also reduce the sensitivity of the astrophysical searches. To identify signals, *search pipelines* use information about waveform morphology and coincidences between detectors (see, e.g. [10–14]) to calculate a detection statistic. The presence of glitches in the detector data produces values of the detection statistic larger than the expectation from Gaussian noise alone. To estimate the non-astrophysical background of the detection statistic due to glitches, most search pipelines use a bootstrap approach, for example, by time-shifting the data from independent detectors. These approaches enable a significance to be attached to a candidate (for example, by the false alarm rate or FAR). Therefore, a candidate's significance is affected both by how alike it is to predictions for astrophysical signals and the level of background noise created by glitches. Vetoing glitches or utilising detection statistics that better identify transient noise outliers (see, e.g. [15]) reduce the level of background noise from glitches and hence improve the number of sources that can be identified at a given confidence level. But, in either case, glitches make it more challenging to identify astrophysical signals.

Significant effort has been made to classify glitches into separate *classes* based on their morphology [7, 16–22]. Glitch classification can potentially identify the cause of glitches, enabling detector improvements to reduce their impact. Classification can also help mitigate the impact of glitches on a search, for example, by building a glitch-robust detection statistic [23] or help qualitatively interpret putative astrophysical signals by placing them in the context of known glitch classes.

Davis *et al* [24] investigated the impact of four glitch classes, blips, koi fish, scattered light, and scratchy glitches on the PyCBC search pipeline [12]. Calculating a rate at which the glitches pass a given detection threshold over the physical parameter space of sources, they demonstrate a method to calculate the significance of a possible signal that coincides with a glitch. An alternative approach explored in Ashton *et al* [25] builds information about the distribution of glitches into a Bayesian odds, eschewing a bootstrapped estimation of the background. However it is done, detection approaches which reduce the impact of glitches are critical to improving the search sensitivity and hence identifying more astrophysical candidates.

In this work, we take a modelled approach to glitch classification. We aim to deliver simple probabilistic models for the glitch classes most harmful to astrophysical analyses. To do this, we analyse single-detector glitch triggers pre-classified by the GravitySpy citizen science project [16] from the LIGO Hanford and Livingston detectors [1] during the O3a observing run. We note that the data from the O3b observing run is also publicly available, but we choose only to analyse O3a data. We expect that the population properties will be broadly consistent. However, an avenue for future research is to study whether any differences exist between observing runs and what these can tell us about the underlying mechanisms that generate glitches. We analyse each glitch using a model of a precessing binary black hole (BBH) merger (see Merritt *et al* [26] for an alternative approach developing parametric glitch models). Using an astrophysical black hole merger model enables us to project the properties of the glitch into the physical parameter space of astrophysical signals. Taking these projections, we then build simple parameterised population models for each glitch class and study this population with respect to known astrophysical signals.

The rest of this paper is structured as follows. In section 2, we introduce the methodology based on the principles of population inference. In section 3, we describe our results fitting population models to four different glitch classes. In section 4, we compare the glitch population with the population of binary black-hole mergers described in the GWTC2.1 transient gravitational-wave catalogue. Finally, we conclude in section 5 with a discussion about using these results to validate triggers and improve astrophysical searches.

2. Methodology

The methodology used herein follows the principles of *hierarchical population inference* for compact binary coalescence (CBC) signals (see references [27–29] and Thrane and Talbot [30] for a review). We aim to identify the population properties of four glitch classes *blip*, *tomte*, *fast scattering*, and *scattered light* which are known to most adversely impact search sensitivity [9, 24]. In figure 1, we provides time–frequency spectrograms demonstrating typical examples of each glitch class. These classes were the most numerous during the O1 to O3 observing runs and will likely persist for future observing runs. We perform the population inference separately on glitches in the LIGO Hanford (H1) and LIGO Livingston (L1) detectors. This produces two sets of population properties that we can use to understand the consistency of glitch classes between the detectors. In practice, when analysing triggers for which we

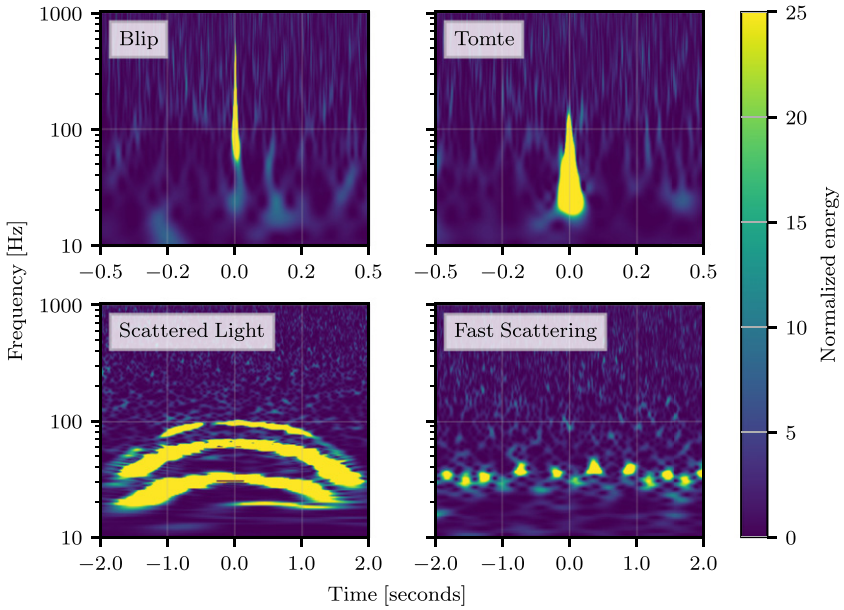


Figure 1. Time–frequency spectrograms of the four glitch classes which we analyse in this work.

do not know if they are a glitch in a single detector or an astrophysical signal, the properties will be inferred, including data from other detectors with coincident observations. In the case where the other detectors contain only Gaussian noise, inferences about the source's intrinsic parameters (e.g. the mass and spins) will remain unchanged (compared to analysing only the glitch-containing detector). This is because the intrinsic parameters determine the waveform morphology [31]. In this work, we analyse only three of the intrinsic parameters (described later) and therefore expect our single-detector results to hold for multi-detector analyses where the other detectors contain only Gaussian noise. However, we do not expect this to hold for the analysis of extrinsic parameters (where the lack of a signal in the other detectors will produce informative posteriors). Finally, the behaviour when multiple detectors are analysed, each with an independent glitch, is not yet understood. Hence, our results may not hold (except when the glitch in one detector is significantly louder and dominates). Nevertheless, the results herein will still help characterise whether the data in any one detector is consistent with a glitch, e.g., by performing single-detector analyses and looking at consistency between detectors.

To infer the properties of glitches in a given class, we first identify a list of *glitch times* pre-categorised by GravitySpy with a confidence threshold of 95%. We check that these glitches would typically contribute to the noise backgrounds estimated by search pipelines, i.e. they would not be vetoed by the single-detector signal-to-noise ratio (SNR) cuts. Each glitch is individually vetted by eye using a time–frequency visualisation to ensure it is consistent with the corresponding glitch class. In our analyses, we then analyse each single glitch from a single detector: from these analyses, we build our population models per-detector. Specifically, we analyse the data $\mathbf{d}_i$ surrounding the ith glitch in our list using a stochastic sampling algorithm (specifically, the nested sampling DYNESTY [32] algorithm through the BILBY [33] package) and a model M. In this work, we use the **`IMRPhenomPv2`** [34, 35] model, which

includes the inspiral, merger and ringdown of merging BBHs and has been a standard benchmark waveform for several observing runs [36, 37]. **IMRPhenomPv2**, however, only models the dominant $(\ell, m) = (2, 2)$ modes in the gravitational-wave radiation. Recent advances in waveform modelling have produced improved models, including the effect of higher-order modes [38–43]. Nevertheless, we use **IMRPhenomPv2** as it is well standardised and fast to evaluate, enabling us to analyse hundreds of glitches without a high computational cost. However, this implies that we infer the properties of glitches given the **IMRPhenomPv2** waveform model. It is well known that systematic uncertainty exists between waveform models and that models including higher-order modes produce quite different posterior distributions, especially when the system has asymmetric masses [44–46]. We, therefore, caution when comparing the glitch population obtained herein with candidate triggers that are analysed using waveforms other than **IMRPhenomPv2**, especially if those waveforms include the effect of higher-order modes.

The **IMRPhenomPv2** model has an associated set of model parameters θ, which describe the mass, spins, location, and orientation of the source. For each glitch from each class, we produce an estimate of the posterior distribution $p(\theta|\mathbf{d}_i, M)$ (in the form of a set of posterior samples) and the Bayesian evidence $\mathcal{Z}(\mathbf{d}_i|M)$. During the analysis of individual glitches, we use a Bayesian prior $\pi(\theta)$ appropriate for analysing astrophysical BBHs. Specifically, the prior is uniform in the detector-frame[5] chirp mass [47]

$$\mathcal{M}^{\rm det} = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}, \tag{1}$$

between 8 and $200M_\odot$ (where m_1 and m_2 are the detector-frame masses of the primary and secondary black holes) and uniform in mass ratio $q = m_2/m_1$ between $1/20$ and 1 and assumes an isotropic distribution on the black holes spins. For all other parameters, we use the standard non-informative astrophysical prior distributions defined in Veitch *et al* [31].

To infer the population properties of glitches in a given class, we take the set of posterior samples and Bayesian evidence from each individual analysis and perform a hierarchical population inference. Specifically, for a parameter $\theta_i \in \theta$, we define a population hyper-models are designed to capture the broad features in a simple parameterised form. The three hypermodels are:

M1 1 A *power law* distribution, it has a single hyperparameter $\Lambda = \{\alpha\}$ and density

$$\pi(\theta_i|\alpha, \mathcal{H}_1) \propto \theta_i^\alpha, \tag{2}$$

with bounds given by the physical bounds on θ_i. This distribution is useful for modelling populations which tend to *rail* against the physical bounds (for example, the primary spin). For α, we use a normally distributed prior with zero mean and a standard deviation of 5. This removes the requirement for an arbitrary upper bound on the magnitude while exponentially suppressing very large values.

M1 2 A truncated *skew-normal* distribution [48, 49], it has three hyperparameters $\Lambda = \{\mu, \sigma, \kappa\}$ and density

$$\pi(\theta_i|\mu, \sigma, \kappa, \mathcal{H}_2) = \frac{2}{\sigma}\phi\left(\frac{\theta_i - \mu}{\sigma}\right)\Phi\left(\kappa\frac{\theta_i - \mu}{\sigma}\right) \tag{3}$$

[5] We infer the glitch population properties of the chirp mass as measured in the detector frame, $\mathcal{M}^{\rm det}$, rather than the source-frame $\mathcal{M} = \mathcal{M}^{\rm det}/(1 + z)$ where z is the inferred redshift.

where $\phi(x)$ is the standard normal probability density function and $\Phi(x)$ is the error function. We truncate the distribution again to the physical bounds on θ_i. This distribution is useful for modelling unimodal populations with a distinct peak. We use a uniform prior for μ over the prior support on θ_i, a normal prior for κ with zero mean and a standard deviation of 10 and a truncated normal prior for σ with standard deviation given by the maximum support on θ_i.

M1 3 A mixture-model of two normal distributions with five hyperparameters $\Lambda = \{\xi,\ \mu_0, \sigma_0,\ \mu_1, \sigma_1\}$ and density

$$\pi(\theta_i|\Lambda, \mathcal{H}_3) = \frac{\xi}{\sigma_0}\phi\left(\frac{\theta_i - \mu_0}{\sigma_0}\right) + \frac{1-\xi}{\sigma_1}\phi\left(\frac{\theta_i - \mu_1}{\sigma_1}\right). \tag{4}$$

This is effective in modelling multi-modal distributions; we do not include a skew as this was found to be poorly constrained in the cases where the M3 model was required. As for the M2 model, we use a uniform prior for μ_0 and μ_1 and a truncated normal distribution for σ_0 and σ_1. For ξ, we apply a uniform prior on [0, 0.5].

For each model, the likelihood is constructed from equation (32) of Thrane and Talbot [30] using 'recycling'. This enables a two-step approach (first analysing individual events, then the population properties), which is critical to reducing the wall-time of the analysis. For each hypermodel, we report the median values of $p(\Lambda|\{\mathbf{d}_i\}, \mathcal{H})$. These, together with the model descriptions above, constitute the core product of this work: simple parameterised models of glitches in the LIGO detectors.

To verify that the hypermodel captures the broad-scale features of the population, we use predictive posterior checking. First, we take the 'measured' population (i.e. the posteriors for each glitch) and bin each individual glitch-posterior in a histogram. Then, for each bin, calculate the mean and the 90% credible interval (between 5% and 95%) and plot this as a solid line and shaded band. This provides a summary of the population properties, which includes the variation in the individual posteriors[6]. An example of this population visualisation can be seen in the right-hand column of table 2: the 'measured' curve shows the mean (solid grey line) and 90% credible interval (filled grey region) for the population of blip glitches across separate parameters. Second, to plot the 'predicted' population, we repeat the process above but simulate the individual posteriors from the population model. Specifically, we draw θ_i from $p(\theta_i|\Lambda')$ (where Λ' is the median)[7] then simulate a posterior and repeat the steps above. The simulated posterior is derived by randomly sampling a posterior from the 'measured' population and aligning its mean with a random draw from the posterior population distribution. This is a non-optimal method, but the alternative (simulating the signal predicted by the posterior population distribution in real detector noise unaffected by other transient noise and performing Bayesian inference) is too computationally demanding to be feasible. Examples of the resulting posterior predictive checks can be found in the right-hand column of table 2: the 'predicted' curve shows the mean (solid orange line) and 90% credible (filled orange region) predicted by the hyper model across separate parameters. Here we see that the models capture the broad features (e.g. the number and location of modes and their width)

[6] Some alternatives, e.g. a histogram of the means, exclude information about the posterior uncertainty captured by our method.

[7] An improvement to the predictive check would be to draw Λ' from the inferred posterior. However, we wish to understand the ability of our simple models, parameterised by the median of the hyperparameter distribution, to explain the data.

but do not always capture the details (e.g. the measured and predicted curves do not overlap perfectly).

3. Results

We now describe the results for each glitch class in the sections below. For each glitch class, we provide a table with a summary (the glitch model and median inferred population parameters) and the results of the posterior predictive checks. We also provide this information in a machine-readable table [50]. For each model, we report the results of a single hypermodel. However, we fitted various models during development and used the posterior predictive checks as a qualitative guide to select the best fitting model.

3.1. Tomte glitches

In table 1, we report the hypermodel, median, and posterior predictive checks for the population properties in $\mathcal{M}^{\rm det}$, q, and a_1 for tomte glitches in the H1 and L1 detectors. For each detector, we analyse a set of 1000 tomte glitches identified with GravitySpy[8]. The SNR of these glitches ranges from 7.7 to 74; the distribution is peaked with a median SNR of 21, but 90% of the tomte glitches have an SNR less than 36[9]. We also verify that we find similar conclusions about the properties when we downsample to 100 glitches; we use this later to justify analysing smaller sets of the fast scattering and scattered light glitches.

For the chirp mass, the measured glitch population is unimodal with a peak at $\sim 75 M_\odot$ and some positive skew. We fit the population using the M2 skew-normal distribution and find a qualitatively acceptable fit for both detector populations. However, the model under-predicts the rate of glitches with $\mathcal{M}^{\rm det} \sim [125, 175]$ in the H1 detector. While the broad features are consistent between H1 and L1, the amount of skew is noticeably different, and tomte glitches in the H1 detectors have chirp mass values up to ~ 150, while tomte glitches in the L1 detector, $\mathcal{M}^{\rm det} \lesssim 125$.

For the mass ratio, the population is unimodal with some skew and the bulk of support on $q \lesssim 1/2$. We fit the mass ratio distribution using the M2 skew normal model and find a good qualitative fit in H1 and L1. As with the chirp mass, we note some qualitative differences between the H1 and L1 detectors, with greater support in the H1 detectors for mass ratios up to 0.5, which is not seen in the L1 glitch population.

The primary spin, a_1 piles up at the extreme spin case $a_1 = 1$; we fit this using the M1 power-law model and find a reasonable agreement. However, we note a lack of events at the $a_1 = 1$ bound. Our fitted model does not capture this feature; however, it is sufficiently narrow that the model remains a reasonable fit. The power-law index $\alpha \sim 5$–6 provides a rough guide to the extent to which the spin 'piles up'; this value is less extreme than in the blip glitch population discussed in the next section. We do not observe significant differences between the H1 and L1 detector populations for the spin.

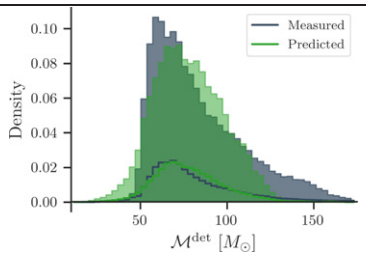
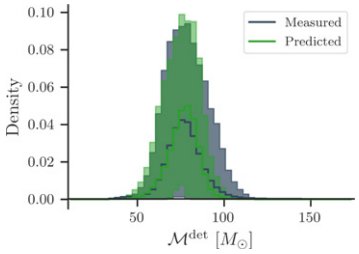
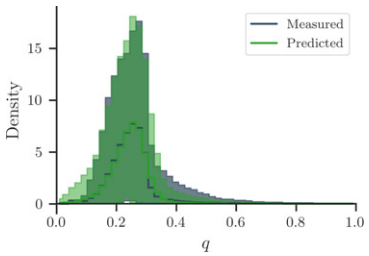
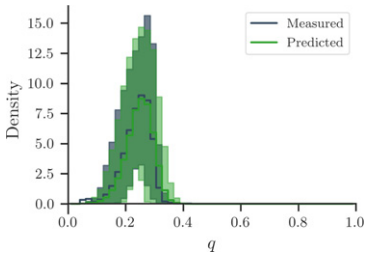
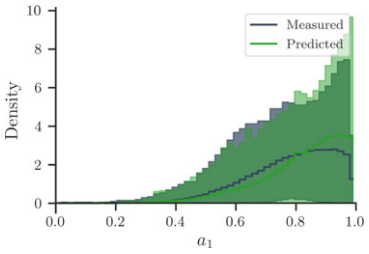
3.2. Blip glitches

In table 2, we report the hypermodel, median, and posterior predictive checks for the population properties in $\mathcal{M}^{\rm det}$, q, and a_1 for blip glitches in the H1 and L1 detectors. We analyse a set of

[8] Of these 1000, computational issues resulted in the failure to analyse 1 tomte glitch from the L1 set; as such, our results pertain to 1000 H1 tomte glitches and 999 L1 tomte glitches.

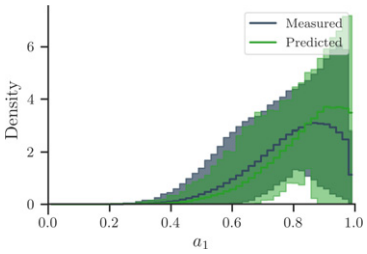
[9] Throughout, we refer to the glitch SNR as the value measured by the `Omicron` [51] software. The lower bound of 7.5 is the result of a threshold applied when selecting the glitches for analysis.

Table 1. Table of parameterised models for tomte glitches in the H1 and L1 interferometers during the O3a era. Uncertainties on the median denote the symmetric 90% credible interval. See section 2 for a description of the posterior predictive plots.

Detector	Type	Parameter	Model	Median	Posterior predictive
H1	Tomte	$\mathcal{M}^{\rm det}$	M2: skew-normal	$\mu = 56^{+1}_{-1}$ $\sigma = 29^{+2}_{-2}$ $\kappa = 3.6^{+0.7}_{-0.6}$	
L1	Tomte	$\mathcal{M}^{\rm det}$	M2: skew-normal	$\mu = 83.5^{+0.6}_{-0.6}$ $\sigma = 8.8^{+0.7}_{-0.6}$ $\kappa = -10^{+6}_{-10}$	
H1	Tomte	q	M2: skew-normal	$\mu = 0.289^{+0.003}_{-0.003}$ $\sigma = 0.057^{+0.004}_{-0.003}$ $\kappa = -14^{+6}_{-10}$	
L1	Tomte	q	M2: skew-normal	$\mu = 0.283^{+0.003}_{-0.003}$ $\sigma = 0.043^{+0.003}_{-0.003}$ $\kappa = -16^{+7}_{-10}$	
H1	Tomte	a_1	M1: powerlaw	$\alpha = 5.3^{+0.4}_{-0.4}$	

(*continued on next page*)

Table 1. Continued.

Detector	Type	Parameter	Model	Median	Posterior predictive
L1	Tomte	a_1	M1: powerlaw	$\alpha = 6.9^{+0.6}_{-0.6}$	

1000 blip glitches in both H1 and L1 identified with GravitySpy. [10]. The SNR of these glitches ranges from 7.5 to 47. The distribution peaks at the minimum bound, and 90% of the glitches have an SNR less than 20. We also verify that our results are robust when using only a subset of 100 randomly selected glitches.

The measured glitch population is unimodal for the chirp mass, with a peak at $\sim 25 M_\odot$ and some positive skew. We find that the M2 skew-normal model provides a good qualitative fit. Comparing the measured and predicted posterior checks, the distribution means are well aligned, though the skew-normal under-predicts the density in the 40–$60 M_\odot$ region.

For the mass ratio, the population is bimodal, with identical features found in both detectors. The bimodality originates from bimodal features observed in the posterior distributions of individual events. Though we see significant scatter in the shape of individual posteriors, averaging over multiple glitches, we observe similar distributions in both the H1 and L1 detectors, indicating a common origin. The bulk of the support for both modes is found in the region $q \lesssim 1/2$, with some support up to the equal-mass bound. We apply the M3 normal mixture model, which provides a reasonable fit to the data, though it tends to underpredict the number of nearly equal-mass glitches in the L1 detector.

Finally, for a_1, the magnitude of the primary spin, we find that the measured glitch values pile up onto the extreme spin case $a_1 \sim 1$. We also found this feature in the tomte population, but here it is more extreme. We find the simplistic M1 power-law model provides a suitable fit to this data; the extreme power-law index $\alpha \sim 37$ demonstrates the strength of the pile-up.

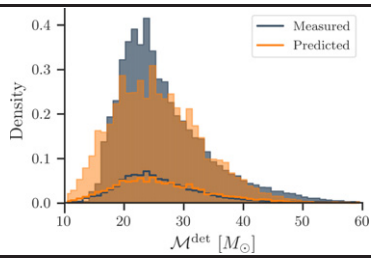
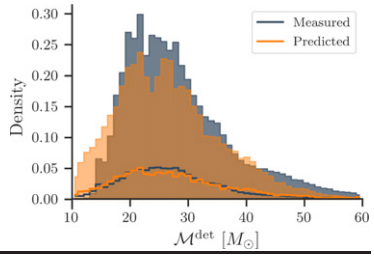
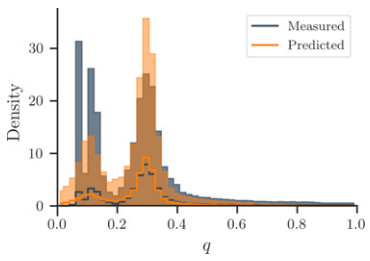
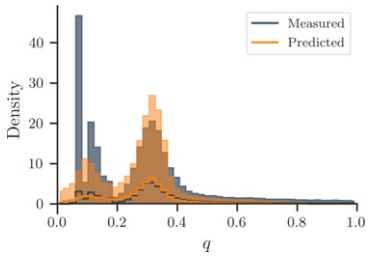
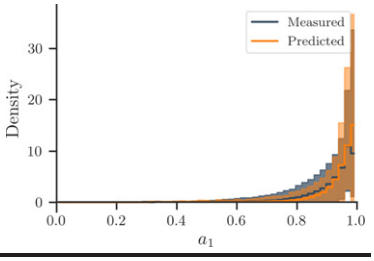
Across all three parameters, we find remarkably consistent glitch behaviour between the H1 and L1 detectors. This, of course, may be simply due to the robust classification of events by GravitySpy. However, we note that the tomte glitches did not show such consistency.

3.3. Fast scattering glitches

In table 3, we report the hypermodel, median, and posterior predictive checks for the population properties in $\mathcal{M}^{\rm det}$, q, and a_1 for fast scattering glitches in the H1 and L1 detectors. We use a set of 100 fast scattering glitches identified with GravitySpy [22]. We limit ourselves to 100 glitches based on our analysis of the blip and tomte glitches, which indicated 100 glitches is sufficient to represent the population properties. This ten-fold reduction in events correspondingly reduces the computational burden of the analysis. The SNR distribution of these glitches

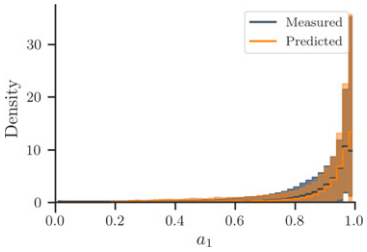
[10] Of these 1000, computational issues resulted in the failure to analyse 61 blip glitches from the L1 set; as such, our results pertain to 1000 H1 blip glitches and 939 L1 blip glitches.

Table 2. Table of parameterised models for blip glitches in the H1 and L1 interferometers during the O3a era. Uncertainties on the median denote the symmetric 90% credible interval. See section 2 for a description of the posterior predictive plots. See section 2 for a description of the posterior predictive plots.

Detector	Type	Parameter	Model	Median	Posterior predictive
H1	Blip	$\mathcal{M}^{\rm det}$	M2: skew-normal	$\mu = 17.2^{+0.5}_{-0.4}$ $\sigma = 11.8^{+0.6}_{-0.6}$ $\kappa = 4.1^{+0.7}_{-0.6}$	
L1	Blip	$\mathcal{M}^{\rm det}$	M2: skew-normal	$\mu = 16.1^{+0.6}_{-0.6}$ $\sigma = 14.7^{+0.9}_{-0.8}$ $\kappa = 4.3^{+0.8}_{-0.7}$	
H1	Blip	q	M3: normal + normal	$\mu_0 = 0.104^{+0.005}_{-0.005}$ $\sigma_0 = 0.030^{+0.005}_{-0.003}$ $\mu_1 = 0.299^{+0.002}_{-0.002}$ $\sigma_1 = 0.018^{+0.002}_{-0.001}$ $\xi = 0.20^{+0.03}_{-0.02}$	
L1	Blip	q	M3: normal + normal	$\mu_0 = 0.101^{+0.006}_{-0.01}$ $\sigma_0 = 0.038^{+0.008}_{-0.005}$ $\mu_1 = 0.317^{+0.003}_{-0.002}$ $\sigma_1 = 0.025^{+0.002}_{-0.002}$ $\xi = 0.21^{+0.03}_{-0.03}$	
H1	Blip	a_1	M1: powerlaw	$\alpha = 58^{+4}_{-5}$	

(*continued on next page*)

Table 2. Continued.

Detector	Type	Parameter	Model	Median	Posterior predictive
L1	Blip	a_1	M1: powerlaw	$\alpha = 60^{+5}_{-4}$	

is strongly peaked at the minimum bound of 7.5, with 90% of the glitches have an SNR less than 13. However, the distribution has a long tail with a maximum value of 121.

For the chirp mass, we find that the population has support from $75M_\odot$ up to the artificial prior bound of $200M_\odot$. This indicates that the actual population likely have support at larger masses still. Such extreme heavy systems are beyond the space of likely astrophysical signals that ground-based detectors will observe in the advanced era. As such, extending our prior beyond this artificial bound is unlikely to help distinguish glitches from signals in the near term, so we choose not to do so. The observed distribution is 'lumpy' with multiple components. We choose to fit the distribution with a skew normal distribution. This does not capture the 'lumpy' behaviour but does broadly capture the typical support.

For the mass ratio, the fast scattering population have extreme mass ratios with a median of $\sim 1/10$ with some support up to $q \sim 0.8$. We fit this distribution using the M2 skew normal model and find a reasonable agreement between the predicted and measured distribution.

The primary spin, as with previously studied glitches, peaks at the extremal $a_1 = 1$ bound. However, the measured population has significant support for systems with a_1 right down to zero. We fit this distribution using the M1 power-law model and find a good fit.

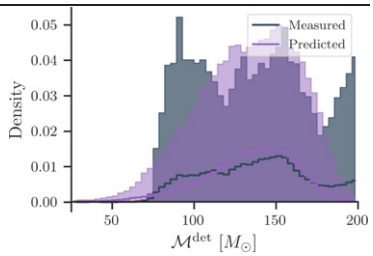
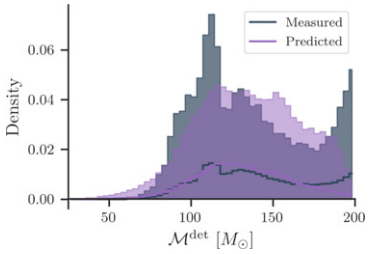
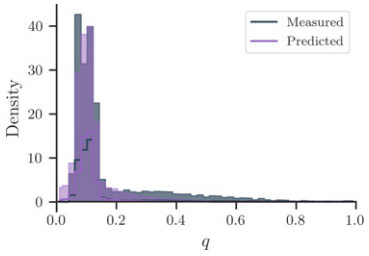
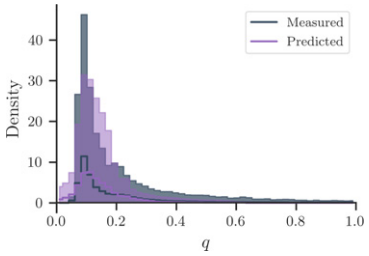
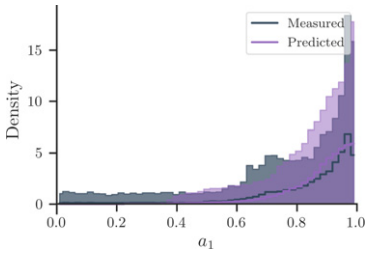
3.4. Scattered light glitches

In table 4, we report the hypermodel, median, and posterior predictive checks for the population properties in $\mathcal{M}^{\rm det}$, q, and a_1 for scattered light glitches in the H1 and L1 detectors. We use a set of 100 scattered light glitches identified with GravitySpy. The choice to use just 100 glitches mirrors the motivation in section 3.2. The SNR of these glitches ranges from 7.5 to 370; as with the fast scattering glitch type, the distribution peaks at the lower bound with a long tail. 90% of the glitches have an SNR less than 47.

For the chirp mass, we find that the population has support from $75M_\odot$ up to the artificial prior bound of $200M_\odot$ with significant 'lumpiness'. The behaviour in the chirp mass distribution is similar in form between the fast scattering and scattered light glitches. For the L1 detector, there is a significant overabundance at the equal-mass bound. We choose to fit the distribution with a skew normal distribution. This does not capture the 'lumpy' behaviour but does broadly capture the typical support.

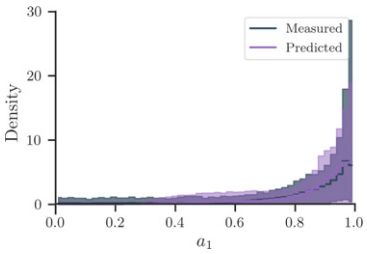
For the mass ratio, the scatter light population have extreme mass ratios with a median of $\sim 1/10$. But, unlike the scattered light glitches, $q \lesssim 0.2$. We fit this distribution using the M2 skew normal model and find a reasonable agreement between the predicted and measured distribution.

Table 3. Table of parameterised models for fast scattering glitches in the H1 and L1 interferometers during the O3a era. Uncertainties on the median denote the symmetric 90% credible interval. See section 2 for a description of the posterior predictive plots.

Detector	Type	Parameter	Model	Median	Posterior predictive
H1	Fast scattering	$\mathcal{M}^{\rm det}$	M2: skew-normal	$\mu = 164^{+7}_{-8}$ $\sigma = 39^{+8}_{-6}$ $\kappa = -7^{+5}_{-10}$	
L1	Fast scattering	$\mathcal{M}^{\rm det}$	M2: skew-normal	$\mu = 102^{+30}_{-8}$ $\sigma = 43^{+10}_{-20}$ $\kappa = 4^{+6}_{-4}$	
H1	Fast scattering	q	M2: skew-normal	$\mu = 0.120^{+0.005}_{-0.009}$ $\sigma = 0.027^{+0.006}_{-0.006}$ $\kappa = -7^{+6}_{-10}$	
L1	Fast scattering	q	M2: skew-normal	$\mu = 0.072^{+0.006}_{-0.006}$ $\sigma = 0.077^{+0.02}_{-0.02}$ $\kappa = 11^{+10}_{-5}$	
H1	Fast scattering	a_1	M1: powerlaw	$\alpha = 11^{+3}_{-2}$	

(*continued on next page*)

Table 3. Continued.

Detector	Type	Parameter	Model	Median	Posterior predictive
L1	Fast scattering	a_1	M1: powerlaw	$\alpha = 21^{+6}_{-4}$	

As with previously studied glitches, the primary spin peaks at the extremal $a_1 = 1$ bound but with little support for smaller spins. We fit this distribution using the M1 power-law model and find a good fit. The power-law index $\alpha \sim 28$ is nearly as large as that of the blip glitches.

4. Comparison with the GWTC2.1 catalogue

So far, 90 gravitational-wave signals have been reported and analysed by the LIGO, Virgo and KAGRA collaborations [36, 37, 52, 53]. Predominantly, these have been BBH systems except for two binary neutron star and two neutron star black hole binaries. Here we compare our glitch population with the 47 events with $p_{\rm astro} > 0.5$ reported in either the GWTC2 [37] or GWTC2.1 [52] catalogues during the O3a observing run[11].

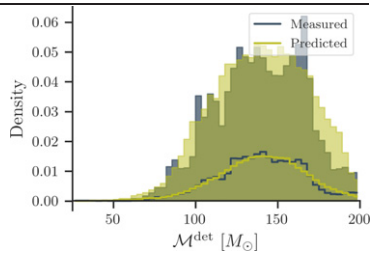
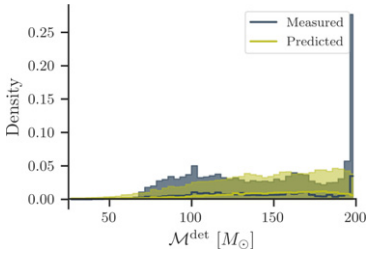
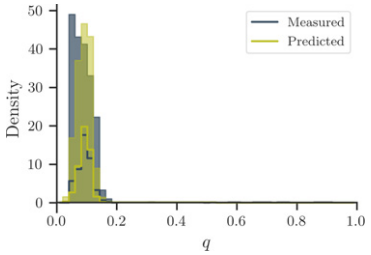
We provide two-dimensional plots comparing our glitch population with the observed signal in $\mathcal{M}^{\rm det}$-q (figure 2), $\mathcal{M}^{\rm det}$-a_1 (figure 3), and q-a_1 (figure 4). In each plot, we give the 90% credible interval of the four glitch classes (for both H1 and L1) along with the combined equal-weighted glitch population in black. Projections of the overall glitch population are given for each one-dimensional plane. Overlaid on this, we then plot the symmetric 90% credible interval for all events with $p_{\rm astro} > 0.5$ reported in either the GWTC2 or GWTC2.1 catalogues.

Comparing the glitch population with the astrophysical BBH distribution, three features stand out: (i) glitches have large spin magnitudes while astrophysical signals have median spin magnitudes $\lesssim 0.9$; (ii) glitches have extreme mass ratios while astrophysical signals tend to have more equal-mass; and finally, (iii) the combined glitch population spans the entire chirp mass range considered in this work, though individual glitch classes tend to cluster.

Glitches have extreme mass ratios and spin because the CBC signals models are a poor fit to the glitch morphology (see, e.g. Merritt *et al* [26]). In the equal-mass zero-spin limit, the CBC waveform is sinusoidal with a characteristic monotonically increasing frequency evolution. Non-equal mass ratios and spins modify this behaviour introducing 'twisting-up' effects which produce irregular waveform morphology's. Glitches are caused by terrestrial disturbances, which are not expected to characteristically chirp up in frequency. So, when fitting the CBC signal model, it is unsurprising that the best fits happen in the non-equal mass and large spin limits.

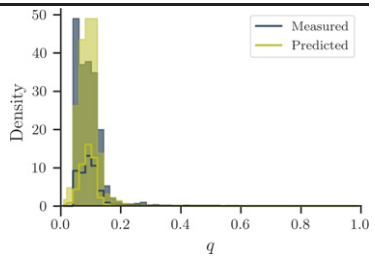
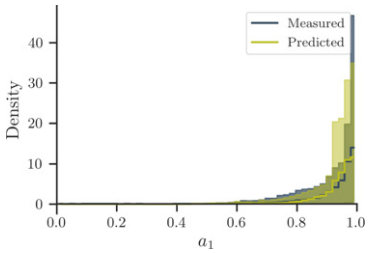
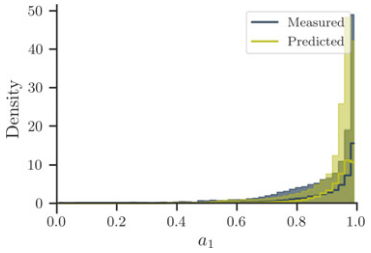
[11] The GWTC2.1 catalogue reported on 44 events with $p_{\rm astro} > 0.5$ observed during O3a, we then added the three events reported in GWTC2, but which had a re-calculated $p_{\rm astro} < 0.5$. This superset, therefore, includes all events during O3a reported by the collaboration with $p_{\rm astro} > 0.5$.

Table 4. Table of parameterised models for scattered light glitches in the H1 and L1 interferometers during the O3a era. Uncertainties on the median denote the symmetric 90% credible interval.

Detector	Type	Parameter	Model	Median	Posterior predictive
H1	Scattered light	$\mathcal{M}^{\rm det}$	M2: skew-normal	$\mu = 151^{+10}_{-30}$ $\sigma = 26^{+9}_{-7}$ $\kappa = -1^{+2}_{-2}$	
L1	Scattered light	$\mathcal{M}^{\rm det}$	M2: skew-normal	$\mu = 199^{+1}_{-10}$ $\sigma = 67^{+9}_{-9}$ $\kappa = -11^{+8}_{-10}$	
H1	Scattered light	q	M2: skew-normal	$\mu = 0.113^{+0.005}_{-0.006}$ $\sigma = 0.029^{+0.005}_{-0.005}$ $\kappa = -8^{+5}_{-10}$	

(*continued on next page*)

Table 4. Continued.

Detector	Type	Parameter	Model	Median	Posterior predictive
L1	Scattered light	q	M2: skew-normal	$\mu = 0.116^{+0.007}_{-0.05}$ $\sigma = 0.035^{+0.008}_{-0.01}$ $\kappa = -4^{+6}_{-10}$	
H1	Scattered light	a_1	M1: powerlaw	$\alpha = 29^{+5}_{-5}$	
H1	Scattered light	a_1	M1: powerlaw	$\alpha = 27^{+6}_{-5}$	

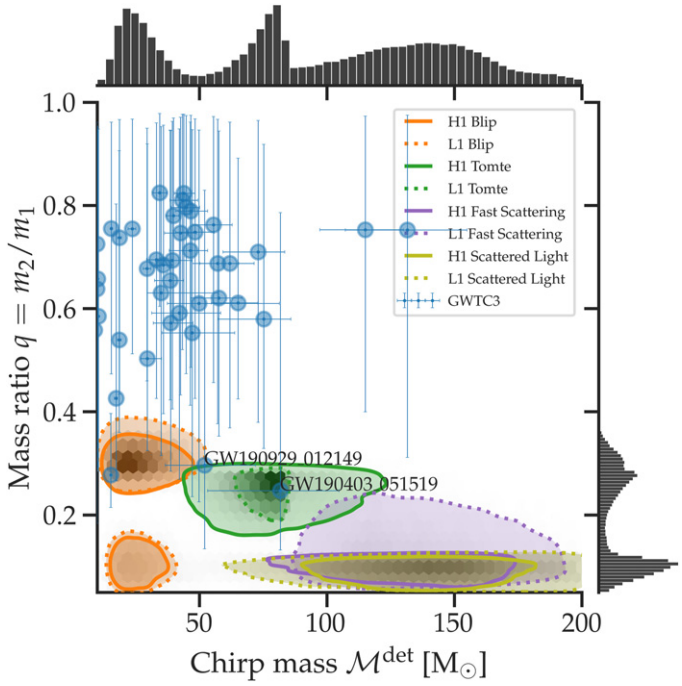


Figure 2. Visualisations of the parameterised glitch models developed in this work with the set of all events observed in the GWTC2 and GWTC2.1 catalogues [37, 52] for the chirp mass and mass ratio. A solid (dashed) line marks the 90% credible interval for the predicted distribution in the H1 (L1) detector, coloured by the glitch classification. A black density plot (and 1D marginal histograms) indicate the distribution of all glitches. Blue markers indicate all confident detections from the GWTC2 and GWTC2.1 catalogues. We remind the reader that the chirp mass reported here is measured in the red-shifted detector frame, not the source frame.

Of the BBH systems reported so far in O3a, just one, GW190403_051519, falls firmly inside the 90% credible interval for the tomte population in chirp mass, mass ratio, and primary spin magnitude. Additionally, GW190929_012149 sits at the edge of the tomte distribution, but no signals are consistent with the other three glitch classes. Several other signals fall within the interval for the chirp mass and primary spin magnitude but are clearly distinguished when looking at the mass ratio and primary spin magnitude[12]. However, it should be noted that both GW190403_051519 and GW190929_012149 have lower SNRs than any of the glitches included in our population analysis since we select only high-confidence glitches. However, Merritt *et al* [26] argue that both quiet and loud glitches (of the same class) exhibit similar morphology.

[12] We note that the GWTC2.1 analysis of these events used the higher-order mode waveforms `IMRPhenomXPHM` [42] and `SEOBNRv4PHM` [39]. As previously discussed, adding higher-order mode content not modelled by the **`IMRPhenomPv2`** waveform used in this work can result in significant shifts in the posterior distribution. However, we re-analysed GW190403_051519 using the **`IMRPhenomPv2`** waveform and found the event remained consistent with the tomte population. Between the GWTC2.1 analysis and our re-analysis, the median detector-frame chirp mass shifted from 86 to 56$M_\odot$, the mass ratio from 0.31 to 0.19, and the primary spin magnitude from 0.93 to 0.85.

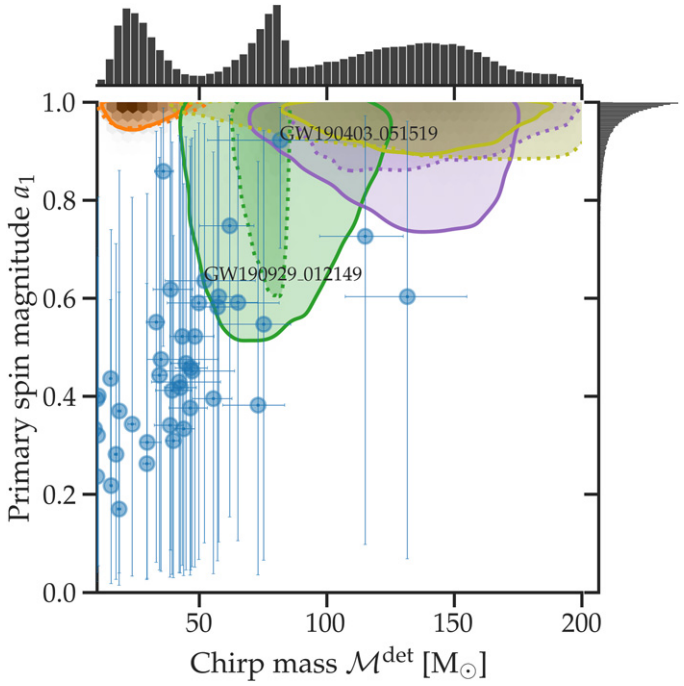


Figure 3. Visualisations of the parameterised glitch models developed in this work for the chirp mass and primary spin magnitude. See figure 2 for a complete description.

In figure 5, we plot the time–frequency spectrogram of GW190403_051519 and a characteristic signal track. This illustrates that the event is primarily identified in the L1 detector, with no clear signal found in the H1 detector.

GW190403_051519 was first identified in the GWTC2.1 catalogue by the GWTC2.1 PyCBC-BBH search alone [52] with a probability of astrophysical origin ($p_{\rm astro}$) of 0.61. No other pipelines identified GW190403_051519 as a candidate. Meanwhile, GW190929_012149 was identified in the GWTC2 catalogue by the GstLAL search alone [37] with $p_{\rm astro} = 1$. However, in the GWTC2.1 catalogue, GW190929_012149 was subsequently identified by all four pipelines used with $p_{\rm astro}$ values ranging from 0.14 to 0.87. GW190403_051519 is notable because the mass of the primary (heavier) object has more than 50% support for a mass greater than $100M_\odot$ and is either inside or above the 'mass gap' predicted by pair-instability supernova theory [55].

That GW190403_051519 and GW190929_012149 are consistent with the tomte glitch population does not, of course, mean that the events are tomte glitches. However, they have $p_{\rm astro}$ values, which allow for a terrestrial cause (i.e. $1 - p_{\rm astro} \gtrsim 0.1$). Tomte glitches were numerous during O3a and are known to impact search sensitivity adversely. Therefore, it seems reasonable to conclude that if GW190403_051519 or GW190929_012149 are not of astrophysical origin, of the four glitch classes considered here, they are most consistent with tomte glitches. However, we add that this assumes the signal morphology of the glitch classes studied herein extends to the low SNR regime of GW190403_051519 and GW190929_012149.

This work does not change the estimated $p_{\rm astro}$ values for GW190403_051519 or GW190929_012149. The search pipelines that identify signals are robust to all relevant glitch

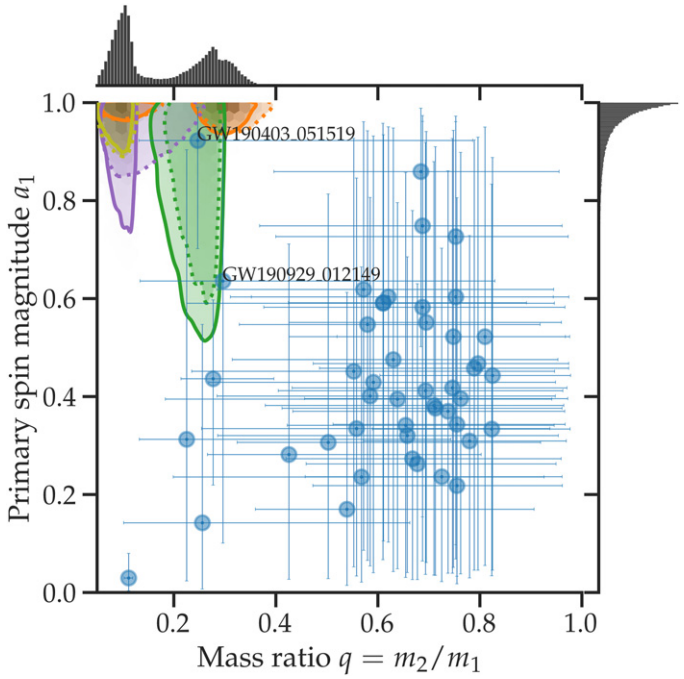


Figure 4. Visualisations of the parameterised glitch models developed in this work for the mass ratio and primary spin magnitude. See figure 2 for a complete description.

classes, use information about the consistency between separate detectors, and apply glitch discriminators [15]. Using bootstrap approaches (e.g. time-shifting the data from independent detectors), the search pipelines estimate the background noise from all glitch classes, and this is intrinsically part of the $p_{\rm astro}$ calculation. Nevertheless, figure 4 demonstrates that to identify signals with large spins and unequal mass ratios, astrophysical searches must contend with an elevated background of glitches compared to other areas of the parameter space. Search pipelines account for the distribution of noise triggers (including glitches) over the searched parameter space in their detection statistics and the resulting significance estimates (both the FAR, and $p_{\rm astro}$), though this is done in a variety of ways [13, 14, 56, 57]. Therefore, the variation in glitch rate across the parameter space indicates that search pipelines are likely not uniformly sensitive across the parameter space [24, 58].

5. Discussion and conclusion

We present population models for non-Gaussian transient noise (glitches) behaviour in the Hanford and Livingston LIGO detectors during the O3a observing run. We first identify lists of blip, tomte, fast scattering, and scattered light glitches pre-categorised by the GravitySpy citizen-science project and then vetted for class consistency using time–frequency visualisations. These glitch classes are known to most adversely affect search sensitivity. We then analyse each glitch class using a precessing BBH model using Bayesian inference. In effect, this allows us to project the properties of the glitch onto the parameter space of astrophysical signals. Finally, we use hierarchical population inference methods to fit simple parameterised

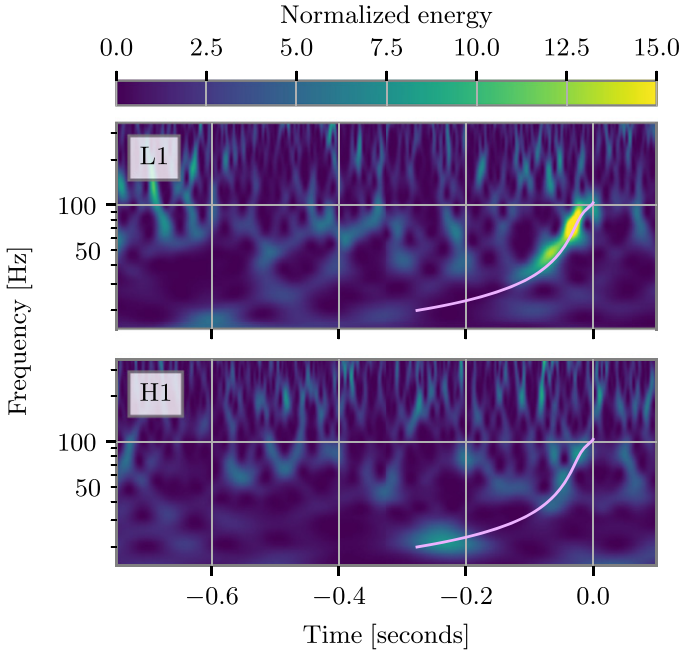


Figure 5. Time–frequency spectrograms of the GW190403_051519 event. We overlay the predicted signal using the `SEOBNRv4_ROM` waveform approximant [54] and the median-estimated signal properties from our **`IMRPhenomPv2`** re-analysis of the event. (We note a mismatch between the waveform used for analysis and the waveform used to plot the signal track. However, the signal track here is intended only to guide the eye and not a quantitative test.)

models to the glitch populations. We do this separately for the chirp mass, mass ratio, and primary spin magnitude parameters; all other parameters we find to be only weakly informative at the population level. We verify the performance of the models using posterior predictive checks and confirm that they reproduce the bulk features of the population.

Finally, we compare the population of glitches with events observed with $p_{\rm astro} > 0.5$ during the O3a observing run. We find two events, GW190403_051519 and GW190929_012149, which have a chirp mass, mass ratio, and primary spin magnitude consistent with the tomte glitch class. GW190403_051519 is a fascinating case as it has a mass that may be inconsistent with the predictions of pair-instability supernova theory.

The parameterised models developed in this work can be used to validate astrophysical signals and potentially improve search algorithms. Already, in section 4, we have demonstrated how the consistency of a trigger with the glitch populations can provide insights into the likely cause if the trigger is not astrophysical. For example, GW190929_051519 has a 39% probability of being terrestrial; we argue that if it is terrestrial, it is most consistent with the tomte glitch class. This type of approach could be used to better-inform retractions in open public alerts (where the final astrophysical probability is not yet available) by identifying when a trigger bears a significant consistency with a known glitch class. In addition, the glitch populations developed here could be used by search pipelines to distinguish astrophysical signals from terrestrial glitches better. Specifically, the Bayesian *astrophysical odds* approach [25] requires a parameterised glitch model (see Ashton and Thrane [59] for the use of a simplistic glitch model in relation to events during O1). The improved models presented here could be directly

incorporated, better separating glitches from signals and resulting in an enhanced capacity of the approach to identity signals. Meanwhile, traditional match-filter pipelines could, for example, add the log of the inverse glitch probability (calculated from the probability of the template parameters under the inferred population models in this work, for example) to their detection statistic. Triggers which are highly consistent with the glitch population would therefore be down-weighted, resulting in a reduced non-astrophysical background and corresponding improvement in the capacity to identify signals. (We note that for current pipelines, this would require a mapping between the precessing spin parameters used in this work to the aligned spin components used by the search pipelines). Already, many search pipelines already apply a version of the proposal above in which the detection statistics and background distributions are calculated over the template banks. This enables a lowering of the background rate in regions of parameter space less contaminated by glitches. Including the probabilistic models here into the detection statistic itself could refine this approach, providing more fine-grained modelled information. Ultimately, this may lead to a greater number of astrophysical signals being detected. However, the impact of this will depend on how much additional information is provided by these glitch models relative to the approaches already taken. To determine if such an approach can improve the detection efficiency, a large-scale simulation study is needed, which is beyond the scope of the current work.

We also identify multimodality in the mass ratio of blip glitches which may be used to identify this glitch class. We do not yet understand why this multimodality is exhibited, but it seems likely the waveform fits two separate parts of the glitch morphology in different regions of parameter space. This information may help formulate an improved glitch model (e.g., breaking the astrophysical waveform in a non-physical way that fits the glitch). Such a model could then be used to identify glitches and improve astrophysical searches.

At the same time, in the future, we hope to improve these models by including correlations between parameters, using higher-order mode waveform models and extending the study to sub-dominant parameters. These improvements will provide a more refined glitch consistency test and may reveal new glitch population features.

Acknowledgments

We a grateful to Derek Davis for help with the script used to produce figures 1 and 5, to Andrew Williamson for insightful comments about the nature of the bimodality seen in the blip glitch mass ratio, to Tom Dent for helpful feedback on the manuscript, and to the CBC and detector characterisation groups of the LIGO and Virgo Collaborations for helpful feedback on this work. We also think Richard O'Shaughnessy, Jason Hathaway, Ben Henderson, Brian Cowburn for useful discussions related to this work. GA and LKN thank the UKRI Future Leaders Fellowship for support through the Grant MR/T01881X/1. This research has made use of data, software and web tools obtained from the Gravitational Wave Open Science Center (https://gw-openscience.org), a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration. LIGO is funded by the U.S. National Science Foundation. Virgo is funded by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale della Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by Polish and Hungarian institutes. We are also grateful to computing resources provided by the LIGO Laboratory computing clusters at California Institute of Technology and LIGO Hanford Observatory supported by National Science Foundation Grants PHY-0757058 and PHY-0823459. The majority of analysis performed for this research was done using resources provided by the Open Science Grid [60, 61], which is supported by the National Science Foundation award

#2030508. This work makes use of the `SciPy` [62], `Numpy` [63–65], `GWpy` [66], and `PyCBC` [12] software for data analysis and visualisation.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: 10.5281/zenodo.5550510.

ORCID iDs

Gregory Ashton https://orcid.org/0000-0001-7288-2231
Sarah Thiele https://orcid.org/0000-0001-7442-6926
Jess McIver https://orcid.org/0000-0003-0316-1355
Laura K Nuttall https://orcid.org/0000-0001-7472-0201

References

[1] Aasi J *et al* (Advanced LIGO) 2015 *Class. Quantum Grav.* **32** 074001
[2] Acernese F *et al* 2015 Advanced Virgo: a second-generation interferometric gravitational wave detector *Class. Quantum Grav.* **32** 024001
[3] Aso Y, Michimura Y, Somiya K, Ando M, Miyakawa O, Sekiguchi T, Tatsumi D and Yamamoto H 2013 Interferometer design of the KAGRA gravitational wave detector *Phys. Rev.* D **88** 043007
[4] Abbott B P *et al* 2020 Prospects for observing and localizing gravitational-wave transients with Advanced LIGO, Advanced Virgo and KAGRA *Living Rev. Relativ.* **23** 3
[5] Blackburn L *et al* 2008 The LSC glitch group: monitoring noise transients during the fifth LIGO science run *Class. Quantum Grav.* **25** 184004
[6] Abbott B P *et al* 2016 Characterization of transient noise in Advanced LIGO relevant to gravitational wave signal GW150914 *Class. Quantum Grav.* **33** 134001
[7] Cabero M *et al* 2019 Blip glitches in Advanced LIGO data *Class. Quantum Grav.* **36** 155010
[8] Abbott B P *et al* 2020 A guide to LIGO–Virgo detector noise and extraction of transient gravitational-wave signals *Class. Quantum Grav.* **37** 055002
[9] Davis D *et al* 2021 LIGO detector characterization in the second and third observing runs *Class. Quantum Grav.* **38** 135014
[10] Allen B, Anderson W G, Brady P R, Brown D A and Creighton J D E 2012 FINDCHIRP: an algorithm for detection of gravitational waves from inspiraling compact binaries *Phys. Rev.* D **85** 122006
[11] Cannon K, Hanna C and Keppel D 2013 Method to estimate the significance of coincident gravitational-wave observations from compact binary coalescence *Phys. Rev.* D **88** 024025
[12] Nitz A H, Dent T, Dal Canton T, Fairhurst S and Brown D A 2017 Detecting binary compact-object mergers with gravitational waves: understanding and improving the sensitivity of the PyCBC search *Astrophys. J.* **849** 118
[13] Aubin F *et al* 2021 The MBTA pipeline for detecting compact binary coalescences in the third LIGO–Virgo observing run *Class. Quantum Grav.* **38** 095004
[14] Sachdev S *et al* 2019 The GstLAL search analysis methods for compact binary mergers in Advanced LIGO's second and Advanced Virgo's first observing runs (arXiv:1901.08580)
[15] Allen B 2005 χ^2 time–frequency discriminator for gravitational wave detection *Phys. Rev.* D **71** 062001
[16] Zevin M *et al* 2017 Gravity Spy: integrating advanced LIGO detector characterization, machine learning, and citizen science *Class. Quantum Grav.* **34** 064003
[17] Mukherjee S, Obaid R and Matkarimov B 2010 Classification of glitch waveforms in gravitational wave detector characterization *J. Phys.: Conf. Ser.* **243** 012006

[18] Rampone S, Pierro V, Troiano L and Pinto I M 2013 Neural network aided glitch-burst discrimination and glitch classification *Int. J. Mod. Phys.* C **24** 1350084
[19] Powell J, Trifirò D, Cuoco E, Heng I S and Cavaglià M 2015 Classification methods for noise transients in advanced gravitational-wave detectors *Class. Quantum Grav.* **32** 215012
[20] Powell J, Torres-Forné A, Lynch R, Trifirò D, Cuoco E, Cavaglià M, Heng I S and Font J A 2017 Classification methods for noise transients in advanced gravitational-wave detectors:II. Performance tests on Advanced LIGO data *Class. Quantum Grav.* **34** 034002
[21] Mukund N, Abraham S, Kandhasamy S, Mitra S and Philip N S 2017 Transient classification in LIGO data using difference boosting neural network *Phys. Rev.* D **95** 104059
[22] Soni S *et al* 2021 Discovering features in gravitational-wave data through detector characterization, citizen science and machine learning *Class. Quantum Grav.* **38** 195016
[23] Jadhav S, Mukund N, Gadre B, Mitra S and Abraham S 2020 Improving significance of binary black hole mergers in Advanced LIGO data using deep learning : confirmation of GW151216 (arXiv:2010.08584)
[24] Davis D, White L V and Saulson P R 2020 Utilizing aLIGO glitch classifications to validate gravitational-wave candidates *Class. Quantum Grav.* **37** 145001
[25] Ashton G, Thrane E and Smith R J E 2019 Gravitational wave detection without boot straps: a Bayesian approach *Phys. Rev.* D **100** 123018
[26] Merritt J, Farr B, Hur R, Edelman B and Doctor Z 2021 Transient glitch mitigation in Advanced LIGO data with *glitschen* (arXiv:2108.12044)
[27] Mandel I 2010 Parameter estimation on gravitational waves from multiple coalescing binaries *Phys. Rev.* D **81** 084029
[28] Mandel I and O'Shaughnessy R 2010 Compact binary coalescences in the band of ground-based gravitational-wave detectors *Class. Quantum Grav.* **27** 114007
[29] Adams M R, Cornish N J and Littenberg T B 2012 Astrophysical model selection in gravitational wave astronomy *Phys. Rev.* D **86** 124032
[30] Thrane E and Talbot C 2019 An introduction to Bayesian inference in gravitational-wave astronomy: parameter estimation, model selection, and hierarchical models *Publ. Astron. Soc. Aust.* **36** e010
[31] Veitch J *et al* 2015 Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library *Phys. Rev.* D **91** 042003
[32] Speagle J S 2020 DYNESTY: a dynamic nested sampling package for estimating Bayesian posteriors and evidences *Mon. Not. R. Astron. Soc.* **493** 3132
[33] Ashton G *et al* 2019 BILBY: a user-friendly Bayesian inference library for gravitational-wave astronomy *Astrophys. J. Suppl. Ser.* **241** 27
[34] Hannam M, Schmidt P, Bohé A, Haegel L, Husa S, Ohme F, Pratten G and Pürrer M 2014 Simple model of complete precessing black-hole-binary gravitational waveforms *Phys. Rev. Lett.* **113** 151101
[35] Schmidt P, Hannam M and Husa S 2012 Towards models of gravitational waveforms from generic binaries: a simple approximate mapping between precessing and nonprecessing inspiral signals *Phys. Rev.* D **86** 104063
[36] Abbott B P *et al* 2019 GWTC-1: a gravitational-wave transient catalog of compact binary mergers observed by LIGO and Virgo during the first and second observing runs *Phys. Rev.* X **9** 031040
[37] Abbott R *et al* 2021 GWTC-2: compact binary coalescences observed by LIGO and Virgo during the first half of the third observing run *Phys. Rev.* X **11** 021053
[38] London L *et al* 2018 First higher-multipole model of gravitational waves from spinning and coalescing black-hole binaries *Phys. Rev. Lett.* **120** 161102
[39] Cotesta R, Buonanno A, Bohé A, Taracchini A, Hinder I and Ossokine S 2018 Enriching the symphony of gravitational waves from binary black holes by tuning higher harmonics *Phys. Rev.* D **98** 084028
[40] Nagar A *et al* 2018 Time-domain effective-one-body gravitational waveforms for coalescing compact binaries with nonprecessing spins, tides, and self-spin effects *Phys. Rev.* D **98** 104052
[41] Varma V, Field S E, Scheel M A, Blackman J, Kidder L E and Pfeiffer H P 2019 Surrogate model of hybridized numerical relativity binary black hole waveforms *Phys. Rev.* D **99** 064045
[42] Pratten G *et al* 2021 Computationally efficient models for the dominant and subdominant harmonic modes of precessing binary black holes *Phys. Rev.* D **103** 104056
[43] Estellés H, Husa S, Colleoni M, Keitel D, Mateu-Lucena M, García-Quirós C, Ramos-Buades A and Borchers A 2020 Time domain phenomenological model of gravitational wave subdominant harmonics for quasi-circular non-precessing binary black hole coalescences (arXiv:2012.11923)

[44] Abbott R *et al* 2020 GW190412: observation of a binary-black-hole coalescence with asymmetric masses *Phys. Rev.* D **102** 043015
[45] Abbott R *et al* 2020 GW190521: a binary black hole merger with a total mass of $150M_{\odot}$ *Phys. Rev. Lett.* **125** 101102
[46] Abbott R *et al* 2020 GW190814 gravitational waves from the coalescence of a 23 solar mass black hole with a 2.6 solar mass compact object *Astrophys. J. Lett.* **896** L44
[47] Cutler C and Flanagan É E 1994 Gravitational waves from merging compact binaries: how accurately can one extract the binary's parameters from the inspiral waveform? *Phys. Rev.* D **49** 2658
[48] O'hagan A and Leonard T 1976 Bayes estimation subject to uncertainty about parameter constraints *Biometrika* **63** 201
[49] Azzalini A and Capitanio A 1999 Statistical applications of the multivariate skew normal distribution *J. R. Stat. Soc.* B **61** 579
[50] Ashton G (2021) Data release: parameterised population models of transient non-Gaussian noise in the LIGO gravitational-wave detectors https://doi.org/10.5281/zenodo.5550510
[51] Robinet F, Arnaud N, Leroy N, Lundgren A, Macleod D and McIver J 2020 Omicron: a tool to characterize transient noise in gravitational-wave detectors *SoftwareX* **12** 100620
[52] The LIGO Scientific Collaboration, The Virgo Collaboration *et al* 2021 GWTC-2.1: deep extended catalog of compact binary coalescences observed by LIGO and Virgo during the first half of the third observing run (arXiv:2108.01045)
[53] Abbott R and Abbott T D *et al* (The LIGO Scientific Collaboration, The Virgo Collaboration, The KAGRA Collaboration) 2021 GWTC-3: compact binary coalescences observed by LIGO and Virgo during the second part of the third observing run (arXiv:2111.03606)
[54] Bohé A *et al* 2017 Improved effective-one-body model of spinning, nonprecessing binary black holes for the era of gravitational-wave astrophysics with advanced detectors *Phys. Rev.* D **95** 044028
[55] Woosley S E 2017 Pulsational pair-instability supernovae *Astrophys. J.* **836** 244
[56] Nitz A H, Capano C, Nielsen A B, Reyes S, White R, Brown D A and Krishnan B 2019 1-OGC the first open gravitational-wave catalog of binary mergers from analysis of public Advanced LIGO data *Astrophys. J.* **872** 195
[57] Nitz A H *et al* 2020 2-OGC: open gravitational-wave catalog of binary mergers from analysis of public Advanced LIGO and Virgo data *Astrophys. J.* **891** 123
[58] Abbott B P *et al* 2018 Effects of data quality vetoes on a search for compact binary coalescences in Advanced LIGO's first observing run *Class. Quantum Grav.* **35** 065010
[59] Ashton G and Thrane E 2020 The astrophysical odds of GW151216 *Mon. Not. R. Astron. Soc.* **498** 1905
[60] Pordes R *et al* 2007 The open science grid *J. Phys.: Conf. Ser.* **78** 012057
[61] Sfiligoi I, Bradley D C, Holzman B, Mhashilkar P, Padhi S and Wurthwein F 2009 The pilot way to grid resources using glideinWMS *2019 WRI World Congress on Computer Science and Information Engineering* vol 2 pp 428–32
[62] Virtanen P *et al* 2020 SciPy 1.0: fundamental algorithms for scientific computing in Python *Nat. Methods* **17** 261
[63] Harris CR, Millman K J and van der Walt S J 2020 Array programming with NumPy *Nature* **585** 357–62
[64] Van Der Walt S, Colbert S C and Varoquaux G 2011 The NumPy array: a structure for efficient numerical computation *Comput. Sci. Eng.* **13** 22
[65] Harris C R *et al* 2020 Array programming with NumPy *Nature* **585** 357
[66] Macleod D, Urban A L, Coughlin S, Massinger T, Pitkin M and rngeorge (2021) gwpy/gwpy: 2.0.