

Inferring the pair-instability mass gap from gravitational wave data

Fabio Antonini¹, Thomas Callister², Fani Dosopoulou¹, Isobel M. Romero-Shaw^{3,4,5} and Debatri Chattopadhyay⁶

¹*Gravity Exploration Institute, School of Physics and Astronomy, Cardiff University, Cardiff, CF24 3AA, United Kingdom*

²*Kavli Institute for Cosmological Physics, The University of Chicago, Chicago, Illinois 60637, USA*

³*DAMTP, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, United Kingdom*

⁴*Kavli Institute for Cosmology, Madingley Road, Cambridge, CB3 0HA, United Kingdom*

⁵*H. H. Wills Physics Laboratory, Tyndall Avenue, Bristol BS8 1TL, United Kingdom*

⁶*Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA) and Department of Physics and Astronomy, Northwestern University, 1800 Sherman Avenue, Evanston, Illinois 60201, USA*



(Received 10 June 2025; accepted 28 August 2025; published 22 September 2025)

We use hierarchical Bayesian inference with nonparametric Gaussian process models to investigate the effective inspiral spin parameter, χ_{eff} , as a function of primary black hole mass in the third gravitational-wave transient catalog (GWTC-3). Our analysis reveals a transition in the population at a primary mass of $46_{-5}^{+7} M_{\odot}$. Beyond this mass, the χ_{eff} distribution broadens, becomes consistent with being symmetric around zero, and has a median of $-0.03_{-0.59}^{+0.36}$ (90% credibility). These results are consistent with the presence of a pair-instability mass gap that is repopulated by black holes that are the remnant of a previous merger, formed in dense star clusters. However, asymmetric distributions skewed toward positive χ_{eff} are not excluded by current data. Below the inferred transition mass, we constrain the fraction of second-generation black holes to be $\lesssim 10\%$. These results provide model-independent support for a high-mass and high-spin population of black holes in the data, consistent with earlier work using parametric models. Imminent gravitational-wave data releases will be essential to sharpen constraints on spin symmetry and clarify the origin of the black holes.

DOI: 10.1103/nxnr-pdyx

I. INTRODUCTION

Observations of gravitational waves (GWs) from binary black hole (BH) mergers have revolutionized our ability to probe the lives and deaths of massive stars [1–5]. These mergers encode critical information about the formation, evolution, and final fates of massive binary systems across cosmic time. However, interpreting the growing sample of detections remains challenging, largely due to significant uncertainties in the physics of binary stellar evolution and BH formation. These include poorly understood processes such as mass transfer, common-envelope evolution, natal kicks, and spin alignment, as well as the sensitivity of population synthesis models to uncertain initial conditions and parameter choices [e.g., [6–9]].

A key open question is the existence and extent of a predicted mass gap in the BH birth mass distribution,

caused by (pulsational) pair-instability supernovae [(P) PISNe]. Although the location of the mass gap is subject to several uncertainties [e.g., [10,11]], stellar evolution theory predicts that stars with helium core masses in the range $\approx 40\text{--}65 M_{\odot}$ undergo partial mass ejection (pulsational PISN), while those above $\sim 65 M_{\odot}$ are entirely disrupted (full PISN), preventing the formation of BHs in the approximate range $\sim 40\text{--}130 M_{\odot}$ [12–17]. Yet, current GW observations from the LIGO-Virgo-KAGRA Collaboration (LVK) show no sharp cutoff or dearth of mergers in this mass range [2,18–23], raising questions about either the location of the gap or the mechanisms responsible for populating it [e.g., [10,14,24]].

One compelling explanation is the formation of high mass BHs from previous mergers [25]. Stellar-origin BHs undergo successive mergers in dense stellar environments, such as globular clusters, nuclear star clusters, or active galactic nucleus (AGN) disks [26–35]. In these scenarios, first-generation BHs merge to form more massive, second-generation BHs, which can then participate in further mergers. This process naturally populates the pair-instability mass gap and leads to distinctive spin signatures. The BH components involved in such mergers are expected to have higher spins and isotropic spin orientations [29,36,37].

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](#). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Detecting such a subpopulation would not only confirm the presence of the (P)PISN mass gap but also offer a direct probe of dynamical processes in star clusters and galactic nuclei.

The most informative spin parameter for detecting such effects is the effective inspiral spin, χ_{eff} , defined as the mass-weighted projection of the component BH spins onto the orbital angular momentum [38,39]. χ_{eff} is relatively well constrained in GW signals and serves as a key discriminator between isolated and dynamical formation channels. Isolated binaries formed through common-envelope evolution are expected to have low and moderately aligned spins, leading to positive-skewed χ_{eff} distributions [40,41], while dynamically assembled binaries are expected to produce broader, symmetric χ_{eff} distributions centred around zero [29,42,43]. Additional parameters—such as the precessing spin parameter χ_p [44], the mass ratio $q = m_2/m_1$, the orbital eccentricity, and the redshift z —have also been leveraged to identify subpopulations and trends indicative of dynamical formation [45–51]. In recent years, several studies have investigated correlations between mass and spin as a signature of dynamical mergers [23,52–57].

A key prediction of the merger scenario in dense clusters is the emergence of a distinct population of second-generation BHs with isotropic spin orientations and well-defined spin magnitude boundaries. In particular, if the pair-instability mass gap is populated by BHs formed from a previous merger, the resulting effective spin distribution is expected to be broad, symmetric about zero, and truncated at characteristic limits set by the spin of the merger remnant—typically $a \simeq 0.7$, where a is the dimensionless spin parameter of the BH, $a = c j / G m^2$ with j the spin angular momentum. This feature arises from robust physical considerations and is largely independent of the specific properties of the host cluster. In Ref. [58], we showed that the χ_{eff} distribution of BH binaries in which the primary was formed from a previous merger is a uniform distribution with boundaries at $|\chi_{\text{eff}}| \simeq 0.47$. This corresponds to the cumulative distribution function (CDF)

$$\text{CDF}(\leq \chi_{\text{eff}}) \simeq 0.5 + \chi_{\text{eff}}. \quad (1)$$

We note that Refs. [37,59] obtained a similar result, but with a significantly smaller boundary of the uniform distribution $|\chi_{\text{eff}}| \simeq 0.3$. In Ref. [58] we showed that a larger $|\chi_{\text{eff}}| \simeq 0.47$ is required to match the results of cluster simulation models and theoretical expectations.

Other formation channels, such as gas-assisted mergers in AGN discs, or field evolution of triples or higher multiplicity systems, can, in principle, also produce high-mass systems in the mass gap [e.g., [60,61]]. However, these channels do not yield similarly robust or universal predictions for the χ_{eff} distribution. For example, the degree of spin alignment in AGN disks is sensitive to uncertain details of disk structure, migration, and accretion efficiency, while spin outcomes in isolated multiple star systems depend heavily on natal spin

distributions, tides, and stellar evolution prescriptions. As a result, while our analysis—focused on identifying clear, data-driven signatures in the χ_{eff} distribution—can be used to test the cluster formation hypothesis, it provides limited discriminatory power for other formation channels with poorly constrained spin predictions.

In Ref. [58], we developed a parametric mixture model for the effective spin distribution. In that study, the χ_{eff} distribution was modeled as a mass-dependent combination of two populations: a low-spin Gaussian component, and a broad, near-uniform component representing the spin distribution expected from dynamically assembled, second-generation BHs. This analysis found strong evidence for a high-mass component with a broad, near-isotropic spin distribution becoming dominant above $\simeq 45 M_{\odot}$, consistent with the onset of the (P)PISN mass gap. However, the extent to which these conclusions depend on the specific parametric form of the model remained an open question.

In this work, we consider hierarchical inference of the effective spin distribution of BH binaries using data from the third gravitational-wave transient catalog [GWTC-3, [62]], employing flexible, nonparametric models that make minimal assumptions about the underlying population. Our primary aim is to determine the mass scale at which a subpopulation with spin properties consistent with 2nd generation BHs begins to emerge and to evaluate the statistical significance and astrophysical interpretation of this transition. By modeling both the spin and mass distributions nonparametrically, we allow the data to guide the inference without imposing strong theoretical priors, enabling a more robust and data-driven exploration of spin–mass correlations. We find consistent a transition mass of $\simeq 45 M_{\odot}$, consistent with our earlier study.

The rest of this paper is organized as follows. In Sec. II we describe our methods and model used. In Sec. III we present our results, describing observed trends in BH effective spins and mass ratios as a function of primary mass. In Sec. IV, we comment on possible astrophysical interpretations of these results, and conclude in Sec. V. In Appendixes A and B we report on the inferred primary mass distribution and present an additional model where we model the mass-ratio distribution as a function of primary mass. More details about the data and the methodology, as well all priors on the hyperparameters of our models are reported in Appendixes C and D.

II. DATA AND BASELINE MODEL

In the analysis that follows, we consider the subset of binary BH (BBH) mergers in GWTC-3 with false alarm rates below 1 yr^{-1} , consistent with Ref. [20]. We exclude all events which include at least one component with mass $< 3 M_{\odot}$ and is therefore likely to involve a neutron star [20]. The resulting dataset contains 69 BBH mergers. Selection effects are accounted for using the set of successfully recovered BBH injections made publicly available by the

LIGO-Virgo-KAGRA collaboration, covering their first three observing runs [20,63]; see Appendix C for further details.

We assume the merger rate density can be factorized as

$$R(m_1, m_2, \chi_{\text{eff}}; z) = R_{\text{ref}} \frac{f(m_1)}{f(20M_{\odot})} \left(\frac{1+z}{1.2} \right)^{\kappa} \times p(m_2|m_1)p(\chi_{\text{eff}}|m_1), \quad (2)$$

where R_{ref} denotes the merger rate per unit mass evaluated at redshift $z = 0.2$ and primary mass $m_1 = 20M_{\odot}$. Our primary object of study is $p(\chi_{\text{eff}}|m_1)$, the probability distribution of effective spins as a function of BH primary mass. In Sec. III below, we will adopt several different models for the probability distribution $p(\chi_{\text{eff}}|m_1)$, ranging from parametric to flexible nonparametric approaches.

We will simultaneously infer the distributions of BBH primary masses, mass ratios q , and redshifts z . Unless otherwise specified, we model the primary mass distribution nonparametrically via a Gaussian process (GP). In particular, we take

$$f(m_1) = \exp[\Phi(\ln m_1)] \quad (3)$$

where $\Phi(\ln m_1)$ is drawn from a zero-mean GP prior,

$$\Phi(x) \sim \mathcal{GP}(0, k(x, x'; a_m, \ell_m)) \quad (4)$$

with a squared-exponential kernel

$$k(x, x') = a_m^2 \exp\left(-\frac{(x - x')^2}{2\ell_m^2}\right). \quad (5)$$

Here, a_m is the amplitude of the GP (controlling vertical variation), and ℓ_m is the length scale (controlling smoothness). We treat both a_m and ℓ_m as free hyperparameters, assigning priors that allow for flexibility while ensuring the inference remains robust and the effective sample size remains sufficient (see the Appendix).

In practice, we evaluate the GP on a uniform grid in $\log m_1$ over the range $2M_{\odot}$ to $100M_{\odot}$, using grid points $\{x_i = \log m_1^{(i)}\}$, for $i = 1, \dots, N_{\text{bin}}$, with $N_{\text{bin}} = 100$. The latent function values $\mathbf{y} = [\Phi(x_1), \dots, \Phi(x_{N_{\text{bin}}})]^T$ are then drawn from the multivariate normal distribution

$$\mathbf{y} \sim \mathcal{N}(\mathbf{0}, \mathbf{K}),$$

where the covariance matrix is defined by $\mathbf{K}_{ij} = k(x_i, x_j; a_m, \ell_m)$. Samples are generated via Cholesky decomposition of \mathbf{K} to ensure numerical stability. The values of \mathbf{y} are then interpolated to evaluate the GP at the locations of events' posterior samples and recovered injections.

We model the conditional distribution of the secondary mass m_2 following Ref. [48]

$$p(m_2|m_1) \propto m_2^{\beta_q}, \quad 2M_{\odot} \leq m_2 \leq m_1. \quad (6)$$

Meanwhile, we assume that the volumetric merger rate evolves as a power law in $(1+z)$ [64,65], such that probability distribution of merger redshifts is

$$p(z) \propto \frac{1}{1+z} \frac{dV_c}{dz} (1+z)^{\kappa}. \quad (7)$$

III. RESULTS

We begin by highlighting a key result of this study: the primary mass scale \tilde{m} at which BH spins transition, previously inferred using strongly parametric models as $\tilde{m} = 44_{-4}^{+6} M_{\odot}$ [58], is recovered here using significantly more flexible, nonparametric models. In Ref. [58], the distribution of χ_{eff} was modeled as a mixture of a Gaussian component, representing the bulk population at $m_1 \lesssim \tilde{m}$, and a uniform component, associated with dynamically assembled (1st + 2nd generation BH) mergers at $m_1 \gtrsim \tilde{m}$. The parameter \tilde{m} denotes the transition mass where the dominant formation channel shifts from isolated to dynamical.

Figure 1 displays the inferred posterior distribution of \tilde{m} for all models considered in this work. Across the different model classes, to be discussed further below, we consistently recover a transition in spin properties at a characteristic primary mass of $\tilde{m} = 46_{-5}^{+7} M_{\odot}$ (90% credible interval). Above this mass, we find strong evidence for the emergence of a distinct subpopulation of BHs with systematically higher spins.

In the sections that follow, we describe the individual models in detail and examine their respective predictions for the mass-spin correlation.

A. Mixture fraction

If a (P)PISNe mass gap exists, and it is repopulated by BHs formed from a previous merger, then above a certain mass threshold (i.e., the lower edge of the gap) the distribution of χ_{eff} should change and become consistent with Eq. (1).

In this section we will return to the model explored in Ref. [58], in which the distribution of BBH spins is a mass-dependent mixture of two components: a truncated Gaussian and the broad and uniform distribution represented by equation (1). Let $\zeta(m_1)$ be the fraction of events with uniformly distributed spins, such that

$$p(\chi_{\text{eff}}|m_1) = (1 - \zeta(m_1))\mathcal{N}(\chi_{\text{eff}}; \mu, \sigma) + \zeta(m_1)\mathcal{U}(\chi_{\text{eff}}; w = 0.47). \quad (8)$$

Here, $\mathcal{N}(\chi_{\text{eff}}; \mu, \sigma)$ denotes a normalized Gaussian distribution with mean μ and standard deviation σ truncated

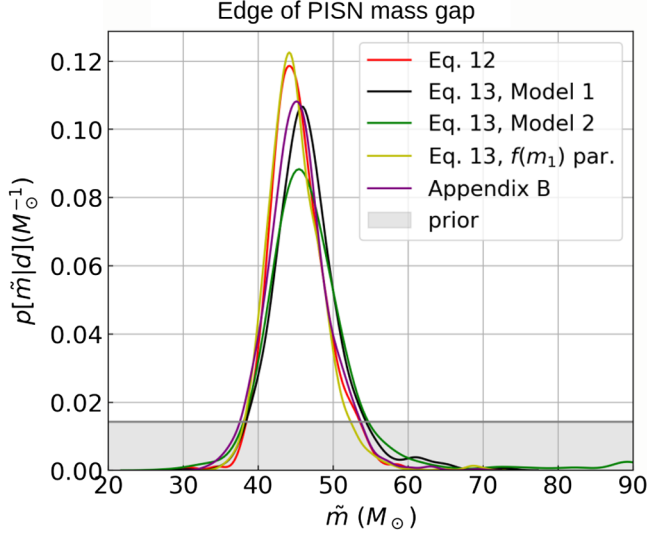


FIG. 1. Posteriors of the transitional mass \tilde{m} obtained under the different models considered in this work. All models yield a consistent estimate of $\tilde{m} \simeq 45M_\odot$. The black and purple lines correspond to models in which both the primary mass distribution and the χ_{eff} distribution are described nonparametrically; these differ only in their prior choices, referred to as model 1 and model 2 in the main text. The yellow line corresponds to a model in which the mass distribution is represented parametrically using the standard power-law plus peak prescription commonly used in previous studies. The red line corresponds to a model in which the χ_{eff} distribution is represented by the parametric form given in Eq. (12), i.e., a uniform distribution with independent bounds. Finally, the purple line denotes a model in which the mass ratio power-law slope parameter β_q is modeled as a GP over the primary BH mass (see Appendix B).

within $[-1, 1]$, and $\mathcal{U}(\chi_{\text{eff}}; w)$ is a uniform distribution defined over the range $|\chi_{\text{eff}}| < w$. We set $w = 0.47$, as predicted for mergers involving 2nd generation BHs.

In Ref. [58], we assumed that $\zeta(m_1)$ rose sharply from zero to one at the transition scale m_1 , restricting our model to a single transition from a purely Gaussian to a purely uniform effective spin distribution. Here, we instead model the mixture fraction $\zeta(m_1)$ via a GP over $\ln m_1$, allowing for considerably more complex behavior in the mass-dependent spin distribution. A GP model for $\zeta(m_1)$ allows for non-monotonic behavior, such as multiple transitions back and forth between spin morphologies. It moreover illustrates at which masses we have informative spin measurements, and thus well-measured $\zeta(m_1)$, and in which regions data are instead uninformative. Finally, $\zeta(m_1)$ will allow us to bound the fraction of systems possibly arising from dynamical formation across the full range of BH masses.

We specifically model $\zeta(m_1)$ via

$$\zeta(m_1) = S(\Psi[\ln(m_1)]) \quad (9)$$

where $\Psi(x)$ is drawn from a zero-mean GP prior, $\Psi(x) \sim \mathcal{GP}(0, k(x, x'; a_\zeta, \ell_\zeta))$, using a squared exponential

kernel with variance a_ζ^2 , and smoothing length ℓ_ζ . A sigmoid function

$$S(x) = \frac{1}{1 + e^{-x}} \quad (10)$$

is then applied to $\Psi(m_1)$ in order to ensure $0 \leq \zeta(m_1) \leq 1$.

The inferred $\zeta(m_1)$ is shown in the upper panel of Fig. 2. It exhibits a clear trend with primary BH mass. At $m_1 \lesssim 5M_\odot$, $\zeta(m_1)$ is unconstrained, spanning the full range from 0 to 1. This indicates that the data are not informative in this regime: this is expected, given the low number of detections and the difficulty in distinguishing spin contributions at small masses. Between $m_1 \approx 10M_\odot$ and $40M_\odot$, the 95% credible interval for $\zeta(m_1)$ remains consistent with 0 and constrained below $\simeq 0.15$. This implies that in this mass range the contribution from a dynamically assembled, isotropically spinning population is statistically small, i.e., less than 15% of the astrophysical population. The truncated Gaussian representing χ_{eff} is a narrow distribution with positive mean $\mu = 0.05^{+0.03}_{-0.03}$ (90% credibility).

Above $m_1 \approx 40M_\odot$, we observe a marked increase in $\zeta(m_1)$, with the mixture fraction bounded above $\zeta(m_1) > 0.1$ at 90% credibility, indicating a statistically significant nonzero contribution from a different population with a broader χ_{eff} distribution. This result indicates a transition in the BH population around $40M_\odot$. Above this mass, the lower 5% bound of the isotropic spin fraction approaches the 95% upper bound of the distribution below $40M_\odot$. This indicates a significant shift in spin properties, consistent with the emergence of a distinct subpopulation at higher masses.

In Fig. 2, we show the 90% prior confidence region. The broad prior on the kernel parameter a_ζ leads to intervals that, after applying the sigmoid transformation, lie close to the prior boundaries. The inferred 5% lower bound on the isotropic spin fraction in the interval ~ 40 to $80M_\odot$ remains significantly above zero. This indicates that the data are constraining this feature: the rise in $\zeta(m_1)$ at high masses is not simply reflecting the edge of a diffuse prior or of an informative region, but rather represents a data-driven trend supported by the observed population. At higher masses still, data again become uninformative due to a low number of observations, and $\zeta(m_1)$ asymptotes back to its prior.

B. Effective spin vs primary mass

To further investigate possible correlations between spin and primary mass, we consider a model in which the effective spin parameter χ_{eff} is instead described as a single Gaussian, but one whose mean and standard deviation vary as functions of the primary BH mass m_1

$$p(\chi_{\text{eff}}|m_1) = \mathcal{N}(\mu_\chi(m_1), \sigma_\chi(m_1)). \quad (11)$$

Here, $\mathcal{N}(\chi_{\text{eff}}; \mu, \sigma)$ denotes a truncated normal distribution where both the mean $\mu_\chi(m_1)$ and the standard deviation

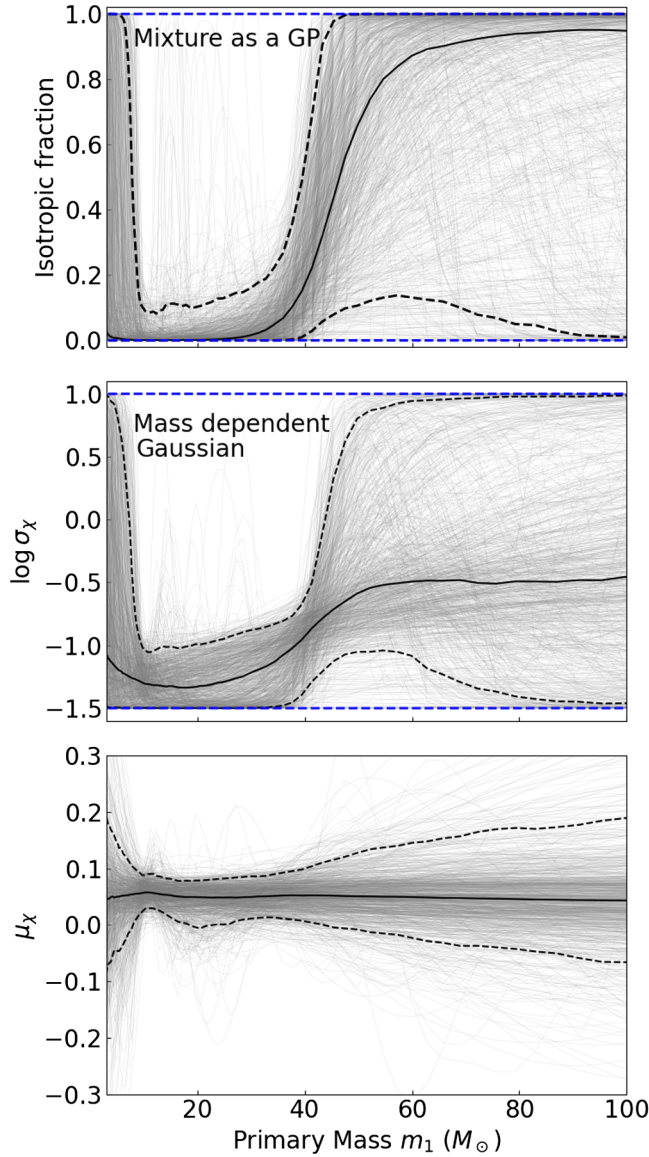


FIG. 2. Upper panel: median and 90% confidence intervals of the mass-dependent mixture fraction $\zeta(m_1)$ as a function of primary BH mass, representing the relative contribution of the isotropically spinning BH population to the total merger rate. Middle and lower panels: the mean and standard deviation of the χ_{eff} distribution, represented here as a truncated Normal distribution with mass dependent mean and variance. Blue-dashed lines show the prior 90% confidence region. For μ_χ these intervals are not shown as they are beyond the range of values plotted.

$\sigma_\chi(m_1)$ are modeled as nonparametric functions using independent GPs over $\log(m_1)$ and adopting an exponential kernel with parameters a_μ , ℓ_μ , a_σ , and ℓ_σ . The prior range of a_μ is then constrained to the interval $[-1, 1]$ using the sigmoid function $2S(x) - 1$, while $\log \sigma_\chi$ is constrained to the interval $[-1.5, 1]$ using $2.5S(x) - 1.5$. The mass distribution is again represented nonparametrically as described in Sec. II.

In contrast to mixture models that impose a population split (e.g., below and above a threshold mass), this framework enables us to search for continuous or gradual changes in the spin distribution that may signal a gradual shift in the population, rather than a sharp transition to a different population above a certain mass (as in Ref. [58]) or an evolving mixture between two starkly different spin distributions (as in Sec. III A). A sharp increase in the variance of χ_{eff} or a change in its mean above a certain value of m_1 would be consistent with the emergence of a different population.

The results of this analysis are presented in the middle and lower panels of Fig. 2. The mean of the χ_{eff} distribution is measured away from zero at $m_1 \simeq 10M_\odot$, where the merger rate has a strong peak (see Fig. 8). This suggests that the majority of merging binaries at these masses are unlikely to originate solely from dynamical interactions in dense star clusters. A positive mean for the χ_{eff} distribution indicates the presence of a subpopulation with preferentially aligned spins, consistent with formation channels such as isolated binary evolution in the field [e.g., [66,67]], or formation in AGN disks [e.g., [68]].

Across the entire mass range, the mean of the χ_{eff} distribution shows little variation with primary mass, although the uncertainty increases with mass such that the mean becomes consistent with zero above $35M_\odot$ (90% credibility). The standard deviation exhibits clear mass dependence, however, closely mirroring the behavior of $\zeta(m_1)$. Between 5 and $40M_\odot$, the distribution is well described by a narrow Gaussian with a positive mean of approximately 0.05 and a standard deviation corresponding to $\log \sigma \simeq -1.3$. At masses above $40M_\odot$, the dispersion increases and becomes $\log \sigma \gtrsim -1$ (90% credibility). These results are consistent with two distinct populations below and above $\simeq 45M_\odot$, that are characterized by different spin distributions. We note that a similar result was recently obtained by Ref. [69], who inferred the joint component mass and χ_{eff} distribution using an iterative kernel density estimation-based method.

In Ref. [58], we showed that the low-mass population is well characterized by a narrow Gaussian, while the high-mass population exhibits a broader distribution. We argued that this high-mass component is responsible for the trend observed in previous studies, where the spin dispersion was found to increase with mass under parametric models assuming a linear dependence of the mean and variance on primary mass [70]. Consistently, the present analysis shows that such linear models are unlikely to be a good representation of the real trend [see also [71]]. However, our analysis does not fully rule out a smoother transition or gradual trend in the spin properties of the overall population. For instance, as shown in Fig. 2, several posterior sample tracers exhibit a more continuous or approximately linear increase of $\log \sigma$ with mass.

C. Symmetry of the effective spin distribution

As identified in Ref. [58] and corroborated above with more flexible models, current data strongly support a correlation between BH spins and mass, with BHs above masses $\tilde{m} \approx 45M_\odot$ exhibiting a broader spin distribution than those below. In Ref. [58], we deliberately enforced the spin distribution of high-mass binary BHs to be isotropic (i.e. an effective spin distribution that is uniform and symmetric about zero). This choice was motivated by the possibility of second-generation BH mergers in dense environments, which robustly produce isotropic spins.

Here, we seek to more carefully test this hypothesis. In particular, do the effective spins of massive binary BHs indeed extend to both larger positive and larger negative values than their less massive counterparts? Or are conclusions about large and negative effective spins primarily driven by prior choices and modeling assumptions?

1. Parametric: Independent minimum and maximum bounds

We first proceed parametrically. As in Ref. [58], we adopt an effective spin model that transitions from a Gaussian to a uniform distribution below and above \tilde{m} , respectively. We extend this model, however, by regarding the upper ($\chi_{\text{eff,max}}$) and lower truncation bounds ($\chi_{\text{eff,min}}$) of the uniform distribution as independent parameters to be inferred from data

$$p_{\mathcal{N}+\mathcal{U}}(\chi_{\text{eff}}|m_1) = \begin{cases} \mathcal{N}(\chi_{\text{eff}}; \mu, \sigma), & m_1 < \tilde{m} \\ \mathcal{U}(\chi_{\text{eff}}; \chi_{\text{eff,min}}, \chi_{\text{eff,max}}) & m_1 \geq \tilde{m}. \end{cases} \quad (12)$$

Here, $\mathcal{U}(\chi_{\text{eff}}; \chi_{\text{eff,min}}, \chi_{\text{eff,max}})$ denotes a uniform distribution over the interval $[\chi_{\text{eff,min}}, \chi_{\text{eff,max}}]$. We place uniform priors on the lower and upper bounds of the uniform component: $\chi_{\text{eff,max}} \in \mathcal{U}(0.05, 1)$ and $\chi_{\text{eff,min}} \in \mathcal{U}(-1, \chi_{\text{eff,max}})$.

The posterior of the parameters governing the χ_{eff} distribution for the high-mass BH population are shown in Fig. 3. The posterior distribution of χ_{eff} is shown in Fig. 4. The inferred transition mass is found to be $\tilde{m} = 45^{+7}_{-5}M_\odot$, and the median of the χ_{eff} distribution is inferred to be $-0.01^{+0.21}_{-0.24}$ (90% credibility). The upper bound of the uniform distribution is relatively well constrained by the data, with $\chi_{\text{eff,max}} = 0.57^{+0.21}_{-0.19}$. The lower bound $\chi_{\text{eff,min}}$, however, is more weakly constrained. We find that $\chi_{\text{eff,min}} < 0$ at 98% credibility, indicating evidence for negative effective spin among high-mass binary BHs. At the same time, we cannot confidently exclude a minimum effective spin that is zero or small and positive. We therefore conclude that the broadening effective spin distribution at high masses is primarily driven by massive BHs exhibiting larger, more positive spins. Whether massive BHs also exhibit larger and more negative spins remains unclear with current data.

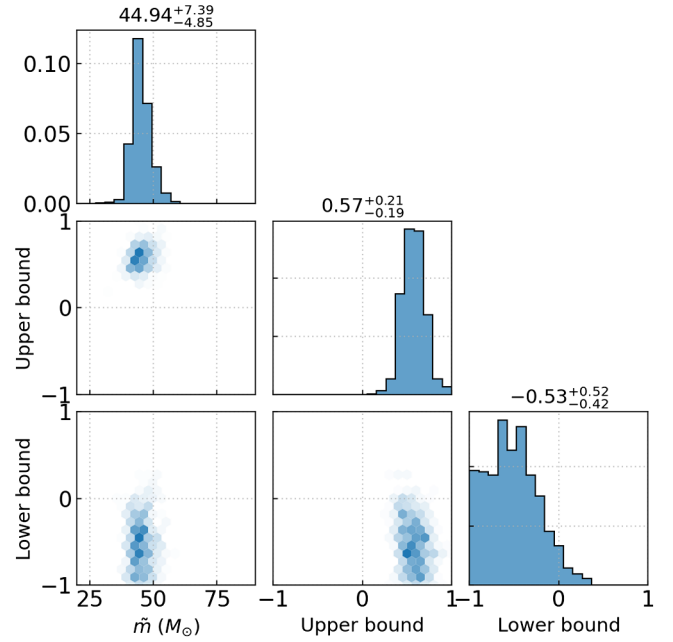


FIG. 3. The lower and upper bounds of the χ_{eff} distribution inferred using the parametric model of Eq. (12), as well as the posterior of the transition mass \tilde{m} between the low mass and high-mass/high-spin BH populations.

Despite the weak constraints on $\chi_{\text{eff,min}}$, we note that current data remain consistent with expectations from repeated mergers in dense clusters. In a dynamical formation scenario where the primary BH formed from a previous merger we should expect $\chi_{\text{eff,max}} = -\chi_{\text{eff,min}} \simeq 0.5$ [58]. The results of Figs. 3 and 4 illustrate that the data are

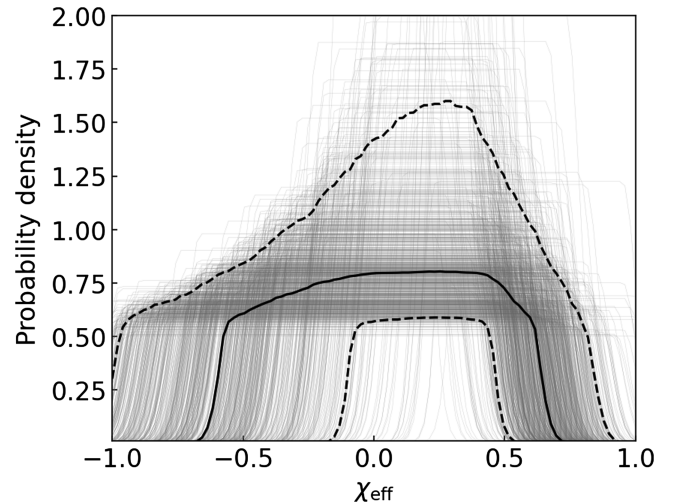


FIG. 4. The χ_{eff} distribution inferred using the parametric model of equation (12), where the high mass population is represented by a uniform distribution with independent lower and upper bounds. The solid line is the posterior median, while the dashed lines show the 90% confidence intervals. Individual posterior draws are shown via light gray traces.

compatible with this basic prediction; additional data will be required to further confirm or rule out this hypothesis.

2. Nonparametric: Gaussian process effective spin distributions

As a further check on these conclusions, we consider a more flexible version of Eq. (12) in which the spin distribution at high masses is described via a GP prior

$$p(\chi_{\text{eff}}|m_1) = \begin{cases} \mathcal{N}(\chi_{\text{eff}}; \mu, \sigma) & m_1 < \tilde{m} \\ e^{\Theta(\chi_{\text{eff}})} / \int_{-1}^1 e^{\Theta(\chi_{\text{eff}})} d\chi_{\text{eff}} & m_1 \geq \tilde{m}. \end{cases} \quad (13)$$

Here, the function $\Theta(\chi_{\text{eff}})$ is generated from a GP,

$$\Theta(\chi_{\text{eff}}) \sim \mathcal{GP}(0, k(\chi_{\text{eff}}, \chi'_{\text{eff}}; a_\chi, \ell_\chi)),$$

with zero mean and a squared-exponential covariance kernel. We evaluate these GPs on a regular grid of $N_{\text{bin}} = 100$ points in χ_{eff} within the range -1 to $+1$. We also consider an additional fit where the nonparametric model for the primary mass distribution was replaced by the power law + Gaussian peak model often used in the literature [e.g., [72]]. The results, however, remain nearly unchanged and are not reported here. This shows, however, that the

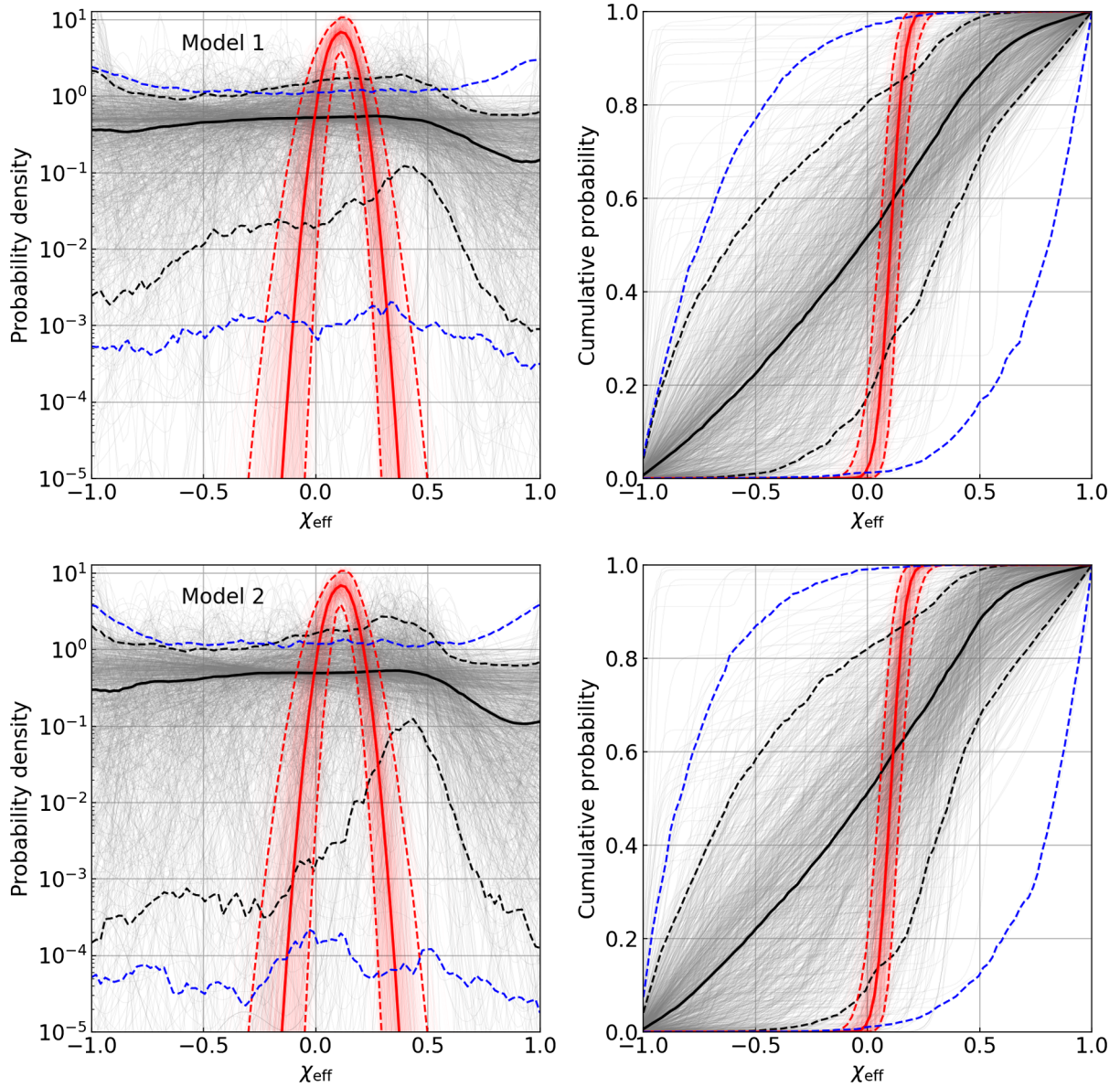


FIG. 5. The distribution of χ_{eff} obtained under model (13) for both the Gaussian component at $m_1 < \tilde{m}$ (red) and the nonparametric component at $m_1 \geq \tilde{m}$ (black and green). We show the results for two different choice of priors on the kernel parameters used in hierarchical analysis, showing how these affect the resulting distributions. Individual posterior draws are shown via light gray and red traces. Blue-dashed lines show the prior 90% confidence intervals.

inference about spin properties is robust across different assumptions about the BH mass spectrum.

We show the inferred distribution of χ_{eff} in Fig. 5. The low and high mass populations can be clearly separated based on their χ_{eff} distributions. We find that the lower bound of the χ_{eff} distribution for the high mass population depends somewhat on the prior choice for the kernel parameters. Thus, in Fig. 5 we show the results for two different choices of prior that are given in Table I of Appendix D. As shown in the figure, model 2 allows for larger amplitude variations of $e^{\Theta(\chi_{\text{eff}})}$ over the χ_{eff} grid than model 1, resulting in a broader prior.

The two prior models produce overall similar distributions for χ_{eff} . Specifically, the inferred median of the χ_{eff} distribution under the two models are $-0.03^{+0.36}_{-0.59}$ (model 1) and $-0.02^{+0.34}_{-0.59}$ (model 2). However, we observe noticeable differences in the lower 5% credible bound outside the range $0 \lesssim \chi_{\text{eff}} \lesssim 0.7$, where model 2 leads to significantly lower probability values. This indicates that the current gravitational-wave data provide only weak constraints on the shape of the χ_{eff} distribution in these regions. Consequently, the inferred distribution in this range is more sensitive to the choice of prior, reflecting the limited information content of the data.

An interesting feature of the χ_{eff} posterior is the asymmetry observed in the lower tail of the distribution, which tends to extend more prominently toward positive values. This asymmetry has previously been interpreted as a physical signature in the χ_{eff} distribution, possibly pointing to formation channels involving AGN disks, where accretion disks can efficiently align BH spins with the orbital angular momentum [73]. Such alignment mechanisms would naturally lead to a positive skew in χ_{eff} due to preferentially aligned spins. Alternatively, it could still be possible that the (P)PISN gap stars at much larger mass values and that the positive alignment is produced by binary stellar evolution where tides can align the stellar progenitor spins with the orbital angular momentum [74]. However, our interpretation is that the observed asymmetry is not necessarily physical, but rather a consequence of the data-limited ability to constrain the negative end of the χ_{eff} distribution. This is particularly evident when examining the cumulative distributions, which reveal that the lower edge of the posterior spans a much broader range than the upper edge. This broad lower tail reflects the fact that large negative values of χ_{eff} remain poorly constrained, leading to an apparent-but not statistically significant-asymmetry.

In Fig. 6 we show the posterior distributions of the CDF evaluated at specific values of χ_{eff} . In particular, we display the posteriors for $\text{CDF}(\chi_{\text{eff}} = 0)$, as well as for $\text{CDF}(-\chi_{\text{eff}})$ and $1 - \text{CDF}(\chi_{\text{eff}})$ evaluated at $\chi_{\text{eff}} = 0.3, 0.5$, and 0.7 . For a population with isotropic spin orientations and randomly aligned angular momenta, the χ_{eff} distribution is expected to be symmetric about zero. This

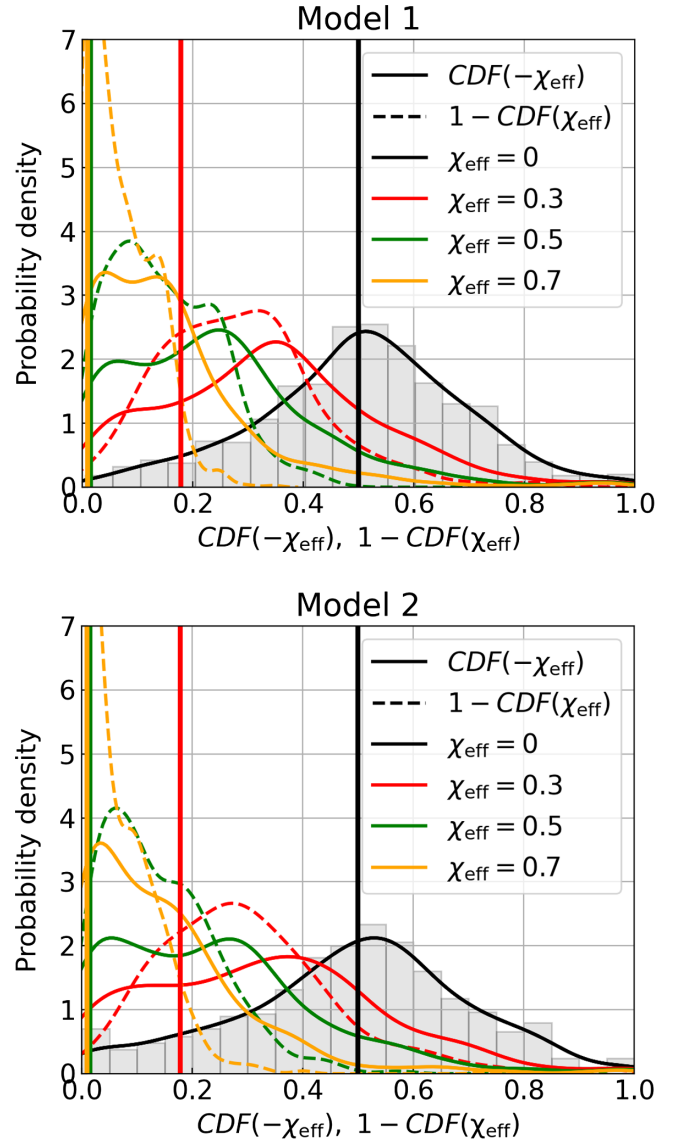


FIG. 6. Posterior distributions of the CDF of χ_{eff} evaluated at selected values: $\chi_{\text{eff}} = 0, 0.3, 0.5$, and 0.7 . We show the posteriors for $\text{CDF}(-\chi_{\text{eff}})$ and $1 - \text{CDF}(\chi_{\text{eff}})$ to assess the symmetry of the χ_{eff} distribution about zero. For an isotropic distribution, we expect $\text{CDF}(\chi_{\text{eff}} = 0) = 0.5$ and $\text{CDF}(-\chi_{\text{eff}}) = 1 - \text{CDF}(\chi_{\text{eff}})$ at each value. The vertical lines give the values expected under the dynamical formation hypothesis, derived from Eq. (1). In gray we show the unsmoothed histogram giving the distribution of the cumulative function evaluated at $\chi_{\text{eff}} = 0$, i.e., $\text{CDF}(\chi_{\text{eff}} = 0)$. The two panels correspond to different prior choices for the kernel parameters used in the GP as explained in the main text.

symmetry implies the two conditions: (1) a median $\text{CDF}(\chi_{\text{eff}} = 0) = 0.5$, and (2) $\text{CDF}(-\chi_{\text{eff}}) = 1 - \text{CDF}(\chi_{\text{eff}})$ at a given χ_{eff} .

From the posterior distributions shown in Fig. 6, we find that both symmetry conditions (1) and (2) are satisfied within the uncertainties of the data. Specifically, the median

values of the CDFs are consistent with the expectations for an isotropic spin distribution across the selected values of χ_{eff} . However, the posterior distributions remain relatively broad, highlighting the limited constraining power of current observations. In particular, the posterior for $\text{CDF}(\chi_{\text{eff}} = 0)$ retains nonzero support at zero, indicating that highly asymmetric spin distributions—such as those strongly peaked at positive χ_{eff} values—are disfavored but not definitively excluded by the data. Conversely, the data do rule out a complete absence of systems with positive χ_{eff} , implying that a subpopulation of spin-aligned binaries is present in the high mass population.

IV. ASTROPHYSICAL IMPLICATIONS

The mass transition we identified in this work has important implications for the formation and evolution of BBHs, and for the astrophysical environments in which they form. The emergence of a broader χ_{eff} distribution above $\approx 45M_{\odot}$, observed consistently across multiple models, points to the presence of a distinct high-mass subpopulation. This transition mass lies near the expected onset of the (P)PISN mass gap [14,17,75] and suggests that BHs above this threshold are not formed through ordinary stellar collapse.

One possible origin of the high mass population we identified is dynamical formation in a dense stellar environment where the primary BH was formed from a previous merger. Such environments include globular clusters [e.g., [76,77]], young massive clusters [e.g., [78–80]], and nuclear star clusters [e.g., [27,31]], where frequent dynamical interactions can assemble and harden BH binaries, leading to repeated mergers over time. To assess whether this hypothesis is consistent with current data, we evolve a grid of star cluster models using the code `cBHBd` [81]. Since the χ_{eff} distribution is largely insensitive to the choice of initial cluster conditions [58], we simulate a single population of 10,000 clusters with fixed metallicity $Z = 1\%Z_{\odot}$ and with masses uniformly distributed in logarithmic space between 10^3 and $10^6 M_{\odot}$. The clusters are evolved for 10 Gyr. In Fig. 7, we show the resulting χ_{eff} distribution for all mergers occurring in the pair-instability mass gap, which in our models lies at $\approx 50M_{\odot}$. We find that the predicted distributions lie almost entirely within the 90% credible intervals of the distribution inferred from the data. Small differences in the tails of the distributions are not considered significant, as those regions are partially shaped by prior assumptions and are not strongly constrained by the data.

If a significant fraction of the high-mass BBH mergers are indeed dynamically assembled—as our analysis tentatively indicates—this would reinforce the idea that dense stellar clusters play a central role in populating the upper end of the BH mass distribution. Dynamical formation models in which clusters dominate the high-mass BH merger rate often also predict a significant, if not dominant,

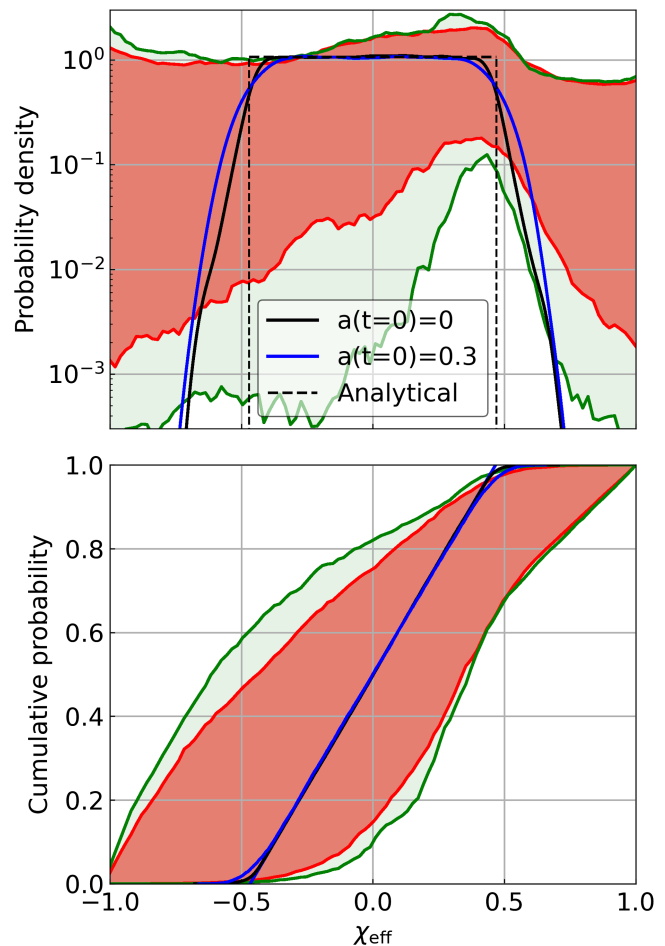


FIG. 7. Comparison between the inferred χ_{eff} distribution and the distribution predicted by cluster models within the pair-instability mass gap. We show the 90% confidence intervals inferred using Eq. (13) and for our two prior choices: model 1 (red) and model 2 (green). The blue dotted line is the simple analytical model of equation (1). The continuous black and blue lines are obtained from cluster models that were evolved with `cBHBd` [81]. We show the results of two different choices for the natal spins of the BHs, in one all BHs are initially nonspinning, and in the other model the all have an initial dimensionless spin parameter $a = 0.3$.

contribution from cluster-formed binaries at lower masses, below the expected pair-instability gap threshold [82]. Thus, identifying a dynamically assembled subpopulation not only sheds light on the high-mass regime but will also put constraints on the contribution of clusters across the full BH mass spectrum. More broadly, evidence for a cluster origin provides a valuable probe of internal cluster dynamics—offering constraints on BH retention efficiency following gravitational recoil, the role of mass segregation in binary formation, and the initial structural and dynamical conditions required for clusters to retain and recycle merger remnants.

On the other hand, if future data were to rule out a pure cluster origin for the high-mass, broad-spin population, this

would also carry significant astrophysical consequences. This would imply that the apparent lack of a sharp pair-instability mass gap in the observed population is not the result of mass growth through BH mergers. In this case, the gap may occur at significantly higher masses than predicted, or might be absent altogether, calling into question existing models of stellar evolution, particularly the treatment of (P)PISN. Alternatively, other formation channels could be contributing to the high-mass population. These include mergers in AGN disks, where gas torques and accretion can aid binary formation and potentially align spins [61,83], or consecutive mergers occurring in the field through the evolution of triples and higher-order systems [32,60]. Another scenario that can produce high-spin, aligned binaries is chemically homogeneous evolution [e.g., [9,84–87]]. In systems with orbital periods $\lesssim 1\text{--}2$ days, strong tidal forces spin the stars up to near their critical velocities, driving efficient Eddington-Sweet circulations and shear instabilities that mix hydrogen from the envelope into the core and helium outward. This rotational mixing prevents the development of a deep convective envelope and suppresses radial expansion on the main sequence, so the components remain compact and tidally locked throughout their lives. Consequently, at core collapse, each star retains most of its angular momentum, producing BHs with near-maximal spins that are almost perfectly aligned with the orbital angular momentum. The resulting χ_{eff} distribution is then sharply peaked at high positive values, in stark contrast to the broad, symmetric distribution expected from dynamical channels. Such scenarios typically make different predictions for spin and eccentricity distributions, redshift evolution, and merger rates, and distinguishing between them will be a major goal for future observational campaigns.

Identifying the location of the inferred transition mass will provide indirect constraints on stellar evolution and nuclear astrophysics. The pair-instability mass gap depends sensitively on the properties of massive stars, particularly on the helium core mass at collapse, which is set by the rate of the $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$ nuclear reaction during helium burning [12]. A transition to second-generation mergers at $\simeq 45M_{\odot}$ would be consistent with the presence of a (P)PISN mass gap at the lower end of the final BH masses than is sometimes assumed. If confirmed, this could favor models with relatively high $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$ rates, leading to more massive CO cores and earlier onset of instability. While current data do not yet allow strong constraints on this reaction rate, future improvements in mass and spin measurements could strengthen the connection between gravitational-wave astronomy and nuclear physics.

Our results also have potential implications for standard siren cosmology [88]. If the sharp transition from low to high effective spin values is indeed associated with the onset of 1st + 2nd generation mergers, it would constitute a relatively redshift-invariant mass scale in the source frame.

As such, it provides a physically motivated and stable reference point for spectral siren cosmology, complementary but more robust than, for example, the Gaussian excess at $\sim 35M_{\odot}$, whose astrophysical origin remains uncertain. However, we note that if metallicity evolution causes the mass gap location to drift with redshift, employing the gap as a standard mass scale for spectral sirens cosmology could introduce systematic offsets in distance estimates. As a result, improving nuclear reaction rate measurements and mass loss models is vital not only to map BH formation channels, but also to establish the (P)PISN gap as a robust tool in both astrophysical population studies and precision cosmology [89,90].

It is also important to recognize that theoretical uncertainties in the precise location and width of the (P)PISN mass gap—driven by factors such as the $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$ reaction rate [91], wind mass loss prescriptions, progenitor rotation, and internal mixing [92]—can shift both the lower and upper edges of the gap by tens of solar masses [93]. Such shifts will naturally smooth out the predicted black-hole mass function, turning what might appear as a sharp threshold in χ_{eff} into a more gradual transition. In practice, this means that the high spin population above $\simeq 45M_{\odot}$ could partially arise from the smearing of the gap itself, rather than solely from a distinct second-generation subpopulation; disentangling these effects will require tighter constraints on stellar evolution physics and more extensive gravitational wave catalogs [12,94].

Taken together, our results illustrate the potential of gravitational-wave observations not only to inform BH formation scenarios but also to probe stellar interiors and the dynamical environments in which massive binaries form. Continued accumulation of high-mass merger detections—particularly those with nonzero spin measurements—will be crucial for refining constraints on the properties of the high-mass BH population and clarifying its astrophysical origin.

V. SUMMARY AND CONCLUSIONS

We have investigated the effective spin distribution of BBH mergers using nonparametric models that allow for flexibility in both the spin and mass distributions. We find that above a primary BH mass of $46_{-6}^{+7}M_{\odot}$ the spin distribution becomes broader and more consistent with isotropic orientation of the spins and orbit. This value of the transitional mass is in good agreement with our previous work based on strongly parametric models where the high-mass and isotropically spinning population was modeled using a uniform distribution, and where we inferred $\tilde{m} = 44_{-4}^{+6}M_{\odot}$ [58].

We explore the phenomenology of this mass-spin correlation using a variety of increasingly flexible models. When adopting a Gaussian process model for the mass-dependent fraction $\zeta(m_1)$ of binary BH mergers with isotropic spins,

we find that this fraction rises sharply above $m_1 \approx 40M_\odot$, reaching $\zeta(m_1) > 0.1$ at 90% confidence and continuing to increase at higher masses. Similar behavior is seen when instead modeling effective spins as a Gaussian distribution with a mass-dependent mean $\mu(m_1)$ and standard deviation $\sigma(m_1)$. In the low-mass regime ($m_1 < 40M_\odot$), we find the effective spin distribution to be tightly clustered around $\mu_\chi = 0.05^{+0.03}_{-0.03}$ with a standard deviation $\log \sigma \simeq -1.3$, consistent with low natal spins and spin alignment; at higher masses, the width of the χ_{eff} distribution increases, with $\log \sigma > -1$ at 90% confidence, indicative of a broader spin distribution.

We furthermore explored the evidence for spin isotropy (i.e. symmetry in the effective spin distribution) among high-mass BHs. When independently measuring the maximum and minimum effective spin among high-mass, we constrain the upper bound to $\chi_{\text{eff,max}} = 0.57^{+0.21}_{-0.19}$, while the lower bound remains weakly constrained at $\chi_{\text{eff,min}} = -0.53^{+0.52}_{-0.42}$. However, 98% of the posterior mass for $\chi_{\text{eff,min}}$ lies below zero, providing weak evidence for negative effective spins in this subpopulation. Under our most flexible nonparametric models, the effective spin distribution for the high-mass BBH population is consistent with being symmetric about zero, although current constraints remain limited by statistical uncertainties.

The data do not exclude alternative possibilities, such as a spin distribution with an excess of aligned systems. In particular, weak constraints on $\chi_{\text{eff,min}}$ mean that we cannot rule a minimum effective spin among high-mass systems that is zero or small and positive. In contrast, the upper bound is better constrained, with a posterior consistent with $\chi_{\text{eff,max}} \simeq 0.5$, as predicted for 1st + 2nd generation mergers in a dense cluster. This asymmetry in constraint quality—tighter on the upper end and looser on the lower—can give rise to an apparent skew in the posterior distribution of χ_{eff} . However, this should not be taken as evidence for an intrinsically asymmetric spin population. Rather, it likely reflects the limited number of current gravitational-wave observations. Given these limitations, we cannot make a definitive statement about the symmetry of the underlying distribution a part that the data remain consistent with a symmetric, isotropic spin distribution. Putting more stringent constraints on a dynamical formation scenario will require a larger sample of high-mass mergers from current and future observing runs.

Main software: ASTROPY [95]; BILBY [96]; cBHBD [81]; JAX [97]; NUMPY [98]; NUMPYRO [99,100]; SCIPY [101].

ACKNOWLEDGMENTS

The authors would like to thank Jack Heinzl, and Matthew Mould for reviewing an early version of this manuscript and for their useful comments that helped to improve the paper. F. A. and F. D. are supported by the UK's Science and Technology Facilities Council Grant

No. ST/V005618/1. I. M. R. S. acknowledges the support of the Herchel Smith Postdoctoral Fellowship Fund and the Science and Technology Facilities Council Grant No. ST/Y001990/1. This material is based upon work supported by NSF's LIGO Laboratory which is a major facility fully funded by the National Science Foundation, as well as the Science and Technology Facilities Council (STFC) of the United Kingdom, the Max-Planck-Society (MPS), and the State of Niedersachsen/Germany for support of the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. Virgo is funded, through the European Gravitational Observatory (EGO), by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale di Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by institutions from Belgium, Germany, Greece, Hungary, Ireland, Japan, Monaco, Poland, Portugal, Spain. K. A. G. R. A. is supported by Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan Society for the Promotion of Science (JSPS) in Japan; National Research Foundation (NRF) and Ministry of Science and ICT (MSIT) in Korea; Academia Sinica (AS) and National Science and Technology Council (NSTC) in Taiwan. The authors are grateful for computational resources provided by Cardiff University and supported by STFC Grant No. ST/V005618/1. D. C. thanks the Gordon and Betty Moore Foundation for funding this research through Grant No. GBMF12341.

DATA AVAILABILITY

The data that support the findings of this article are openly available [102].

APPENDIX A: MASS DISTRIBUTION

In Fig. 8, we present the primary mass distribution inferred under our suite of models. The recovered features are consistent with previous findings [20], revealing prominent peaks in the differential merger rate around $10M_\odot$ and $35M_\odot$. These structures align with earlier studies that attributed such features to underlying astrophysical imprints in the BH mass function [103,104].

For systems with $m_1 \lesssim 40M_\odot$, the population is well described by a single truncated Gaussian effective spin distribution. In this lower-mass regime, the contribution from the isotropic spin component remains minimal.

When the model includes a mass-dependent transition between two spin populations, the inference robustly recovers a sharp change in the χ_{eff} distribution at a characteristic mass of $\tilde{m} \simeq 45M_\odot$. This is reflected in the narrow posterior for \tilde{m} , as shown in Fig. 1. Importantly, this transition mass appears to be largely insensitive to the

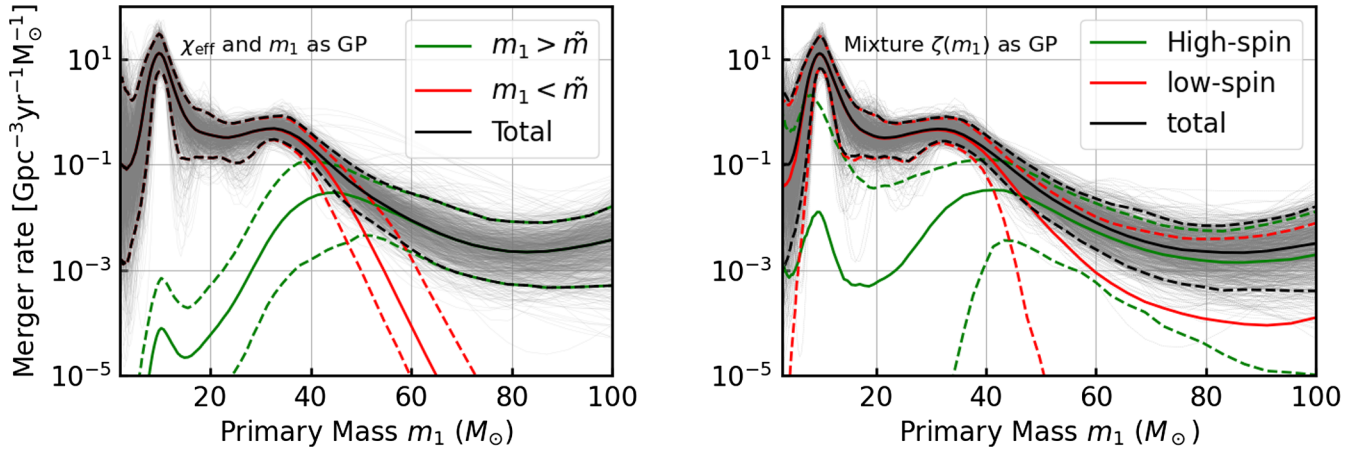


FIG. 8. The differential merger rate as a function of primary mass. In the left panel we use the flexible representation of the χ_{eff} distribution given in Eq. (13). In the right panel we show the merger rate inferred under the model of Eq. (8) where the χ_{eff} is a mixture between a low-spin and a high spin populations, with a mass dependent mixture fraction $\zeta(m_1)$. In both cases the data favor a transition to a different population above $\simeq 45M_\odot$, although the mass value at which this occurs becomes more uncertain under the latter model.

specific functional form used to model the primary mass distribution.

Furthermore, even when using the nonparametric model described in Sec. III A, which allows the mixture fraction $\zeta(m_1)$ between spin subpopulations to vary with mass, the data continue to favor a distinct transition to a highly spinning subpopulation near the same mass scale. However, in this case, the exact fraction of the isotropic component at high masses is less well constrained. This is likely due to the increased flexibility of the model and the limited number of high-mass mergers in the current dataset. Nonetheless, the posterior distributions still require a nonzero contribution from the isotropic population at $m_1 \gtrsim 45M_\odot$.

Finally, we note that the increased variance in the inferred mass distribution at very low ($\lesssim 5M_\odot$) and very high ($\gtrsim 100M_\odot$) masses reflects a reversion to the prior in regions where observational constraints are weak.

APPENDIX B: MASS RATIO vs PRIMARY MASS

Mergers involving a BH that is itself the product of a previous merger are expected to form a distinct population compared to mergers involving first generation BHs only. Not only should this population exhibit larger primary masses and spins, it should additionally be characterized by a distinct mass-ratio distribution preferentially favoring unequal-mass binaries.

If the distinct BH spin distributions described above are indeed indicative of a high-mass population of second-generation BHs, then we additionally expect to see signs of a mass ratio distribution that similarly varies with primary mass. To search for evidence of such behavior, we extend the power-law model [Eq. (6)] of secondary masses, promoting the power-law index β_q from a constant to itself be a function of primary mass: $\beta_q(m_1)$. We model $\beta_q(m_1)$ as

$$\beta_q(m_1) = \Xi[\ln(m_1)], \quad (\text{B1})$$

where, as before, Ξ is drawn from a GP prior evaluated on a regular grid in $\ln m_1$ and we use an exponential kernel with variance and smoothing length a_{β_q} and ℓ_{β_q} , respectively. This flexible approach allows us to explore whether the distribution of mass ratio has a detectable shift above a characteristic mass scale.

We present the results of this analysis in Fig. 9. There are indications of possible mass-dependence of $\beta_q(m_1)$. In particular, we see that data constrain $\beta_q \geq 0$ at both $10M_\odot$ and $35M_\odot$ at 95% credibility, whereas β_q is permitted to be smaller or negative between approximately $15\text{--}30M_\odot$ and above $40M_\odot$. However, the data are not sufficiently informative to determine robustly whether the mass-ratio distribution exhibits systematic trends or transitions; the

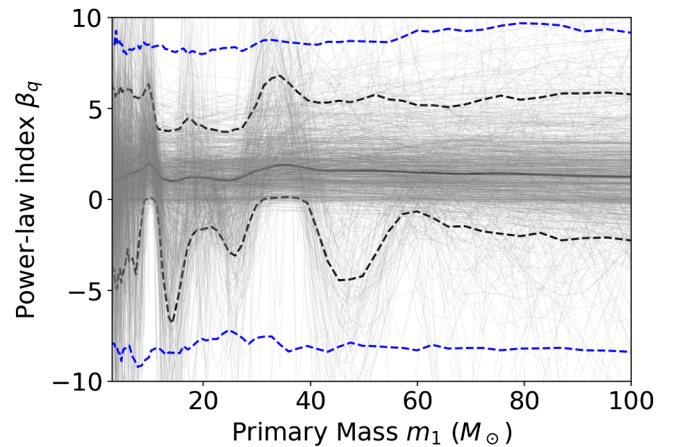


FIG. 9. Power-law index of the mass-ratio distribution evaluated using a GP over a uniform grid in $\ln m_1$. For reference, the blue-dashed lines show the 90% confidence intervals of the prior.

results above may be due simply to the elevated uncertainties in the $15\text{--}30M_\odot$ and $40\text{--}100M_\odot$ ranges due to the small number of events observed with these masses.

APPENDIX C: HIERARCHICAL BAYESIAN INFERENCE

We carry out hierarchical inference on the BBH population using Hamiltonian Monte Carlo (HMC), as implemented in NUMPYRO, a probabilistic programming framework built on JAX. HMC requires the likelihood to be a differentiable function of the population hyperparameters. However, the piecewise-defined spin models we use [e.g., Eq. (13)] introduce discontinuities at $m_1 = \tilde{m}$, which result in non-differentiable behavior in the likelihood with respect to \tilde{m} . To address this issue, we replace these discontinuous transitions with smooth but sharp interpolations between the functional forms of the effective spin distribution above and below \tilde{m} , ensuring differentiability while preserving the essential structure of the model. We redirect the reader to [58] for the details of how this is implemented.

We conduct our analysis within the framework of standard hierarchical Bayesian inference. For each gravitational-wave event, let $p(\theta_i|d_i)$ denote the posterior distribution over the source parameters θ_i (such as component masses, redshift, and spins), conditioned on the observed data d_i . The posterior on the population hyperparameters Λ , given data from all detected events, is then given by [e.g., [48,64,105]]

$$p(\Lambda|\{d_i\}) \propto p(\Lambda) \xi^{-N_{\text{obs}}}(\Lambda) \prod_{i=1}^{N_{\text{obs}}} \int d\theta_i p(\theta_i|d_i) \frac{p(\theta_i|\Lambda)}{p_{\text{pe}}(\theta_i)}, \quad (\text{C1})$$

where $p_{\text{pe}}(\theta_i)$ is the prior used in the original parameter estimation analysis, and $p(\Lambda)$ is our prior on the hyperparameters governing the population distribution. Rather than performing the full integral over each posterior, we approximate the expectation using discrete posterior samples for each event. This yields

$$p(\Lambda|\{d_i\}) \propto p(\Lambda) \xi^{-N_{\text{obs}}}(\Lambda) \prod_{i=1}^{N_{\text{obs}}} \left\langle \frac{p(\theta_i|\Lambda)}{p_{\text{pe}}(\theta_i)} \right\rangle, \quad (\text{C2})$$

where the angle brackets denote an average over the posterior samples of each event.

The detection efficiency $\xi(\Lambda)$ accounts for the fraction of astrophysical signals that we expect to detect under the population model Λ . It is estimated using injection campaigns consisting of N_{inj} simulated signals drawn from a reference distribution $p_{\text{inj}}(\theta_i)$. Considering only the N_{found} injections that are successfully recovered by at least one search pipeline with a false alarm rate below 1 yr^{-1} , we compute

$$\xi(\Lambda) = \frac{1}{N_{\text{inj}}} \sum_{i=1}^{N_{\text{found}}} \frac{p(\theta_i|\Lambda)}{p_{\text{inj}}(\theta_i)}, \quad (\text{C3})$$

where the ratio reweights the found injections from the injection prior to the proposed population model. This approach follows the methodology described in [48], and uses the injection campaign presented in [63,106].

Finally, to guard against potential biases from finite sampling effects in the evaluation of Eq. (C2), we compute the number of effective samples N_{eff} that contribute to the Monte Carlo estimate of the likelihood for each gravitational-wave event. Given a set of N_i posterior samples $\{\theta_{i,j}\}_{j=1}^{N_i}$ for event i , the effective sample count under a proposed population model Λ is defined as

$$N_{\text{eff},i}(\Lambda) \equiv \frac{[\sum_{j=1}^{N_i} w_{i,j}(\Lambda)]^2}{\sum_{j=1}^{N_i} [w_{i,j}(\Lambda)]^2}, \quad (\text{C4})$$

where the weights $w_{i,j}(\Lambda) = p(\theta_{i,j}|\Lambda)/p_{\text{pe}}(\theta_{i,j})$ reweight each sample according to the proposed population model. A small value $N_{\text{eff},i}(\Lambda) \lesssim 10$ indicates that only a handful of posterior samples are informative under the model Λ , which may lead to increased sensitivity to Monte Carlo noise [107].

Similarly, the effective number of injections contributing to the calculation of Eq. (C3) is defined as

$$N_{\text{eff}}^{\text{inj}}(\Lambda) = \frac{(\sum_i w_i(\Lambda))^2}{\sum_j w_j^2(\Lambda)}. \quad (\text{C5})$$

To ensure that the systematic uncertainty in $N_{\text{exp}}(\Lambda)$ remains a subdominant effect in our hierarchical analysis, it is necessary that $N_{\text{eff}}^{\text{inj}}(\Lambda) \gtrsim 4N_{\text{obs}}$ [107].

We track the statistic $\mathcal{N} \equiv \min \log [N_{\text{eff},i}(\Lambda)]$ and compute $N_{\text{eff}}^{\text{inj}}$ for each population model. We then safeguard the inference by strongly penalizing models for which $N_{\text{eff}}^{\text{inj}} < 4N_{\text{obs}}$ and/or $\min \log N_{\text{eff}} < 0.6$. Specifically, we define the function

$$S(x) = \frac{1}{1 + x^{-30}}, \quad (\text{C6})$$

which approaches unity for large x and falls rapidly to zero as x approaches zero. We then include the terms

$$\ln S\left(\frac{N_{\text{eff}}^{\text{inj}}}{4N_{\text{obs}}}\right) + \ln S\left(\frac{\mathcal{N}}{0.6}\right) \quad (\text{C7})$$

in the log-likelihood calculation implemented in NUMPYRO. This ensures the log-likelihood diverges to $-\infty$ if either

condition is violated, thereby excluding models with pathologically low effective sample sizes.

APPENDIX D: PRIORS AND HYPERPARAMETER POSTERiors

The specific priors adopted for our analysis are summarized in Table I. Figures 10–14 show the posteriors for the key parameters of our models.

As shown in Sec. III C 2, certain features of the inferred posteriors are sensitive to the choice of priors for the kernel parameters, which govern the behavior of the latent function in the Gaussian process. This kind of sensitivity is typical in nonparametric models, where one must decide how smooth the function should be across the grid, and where the posterior can revert to the prior in regions of the GP grid where the data are not informative.

To address this, we ran our models multiple times with different prior choices and selected those that allowed for

flexibility and broad variation in the latent function, while still yielding accurate and stable likelihood estimates. In model 1, we adopt normal distributions for the priors on the kernel amplitude a_χ and the logarithm of the length scale $\ln \ell_\chi$, following [22]. Specifically, we take $a_\chi \sim \mathcal{N}(0, 3)$ and $\ln \ell_\chi \sim \mathcal{N}(-1, 0.5)$, applied over the χ_{eff} grid.

These prior choices are quantitatively motivated as follows: the maximum fractional variation we physically expect in the merger rate across the grid is of order $\sim 10^3$. The grid size is $\Delta x = 2$. For a stationary GP with a squared exponential kernel and amplitude a_χ , the prior typical fluctuation (in log rate) between two points separated by a distance Δx is $\text{Std}[\Delta f] = \sqrt{2}a_\chi(1 - \exp[-\frac{(\Delta x)^2}{2\ell_\chi^2}])^{1/2}$. For large $\Delta x/\ell_\chi$, this saturates to $\sqrt{2}a_\chi$. Setting $a_\chi = 3$, we expect a prior standard deviation in log rate of about 4.2 (since $\sqrt{2} \times 3 \approx 4.2$), comfortably covering the expected physical variation of up to three orders of magnitude, but without being overwhelmingly broad. The prior on ℓ_χ centers the length scale at $\exp(-1) \approx 0.37$ (with most prior weight between about 0.2 and 0.6), so that the GP can model features ranging from fairly localized to nearly flat across the grid, but discourages extremely rapid oscillations or excessively global trends unless strongly favored by the data. We find that these choices provide a balance between flexibility and regularization, and ensure that the GP adapts to physically plausible variations while penalizing models that are too rigid.

Similar prior choices were adopted for the remaining models, and can be similarly justified. In model 2, we adopted log-uniform priors on the kernel parameters a^2 and ℓ to model both the χ_{eff} and m_1 distributions. We set a lower bound on ℓ to avoid short-scale variations that would result in small, unphysical features and poor sampling efficiency. The upper bound is fixed at $\log \ell = 2$, which is significantly larger than the grid full width $\Delta x = 2$ used for both $\log m_1$ and χ_{eff} . To allow for large but physically plausible variations in the log rate, we impose an upper bound which is again derived from the maximum expected variation across the parameter space, $\delta_{\text{max}} R \sim 1000$. This requires $a^2 \gtrsim \ln(1000)^2 \sim 50$ to ensure that the GP prior is flexible enough to capture significant features in the data, while maintaining numerical stability and well-behaved posteriors.

TABLE I. Priors adopted for the hyperparameters with which we describe the primary mass, mass ratio, and redshift distributions of the BBH population.

Prior	Defined in	Parameter
$\mathcal{HN}(3)$	Eq. (13) model 1	a_χ
$\mathcal{N}(-0.5, 1)$	Eq. (13) model 1	$\ln \ell_\chi$
$\mathcal{LU}(0, 2)$	Eq. (13) model 2	a_χ^2
$\mathcal{LU}(-1, 1)$	Eq. (13) model 2	ℓ_χ
$\mathcal{HN}(3)$	Eq. (3) model 1	a_m
$\mathcal{N}(0, 1)$	Eq. (3) model 1	$\ln \ell_m$
$\mathcal{LU}(0, 2)$	Eq. (3) model 2	a_m^2
$\mathcal{LU}(-1, 1)$	Eq. (3) model 2	ℓ_m
$\mathcal{HN}(4)$	Eq. (9), (11)	a_ζ, a_μ, a_σ
$\mathcal{N}(-0.5, 1)$	Eqs. (9), (11)	$\ln \ell_\zeta, \ln \ell_\mu, \ln \ell_\sigma$
$\mathcal{HN}(5)$	Eq. (B1)	a_{β_q}
$\mathcal{N}(-0.5, 1)$	Eq. (B1)	$\ln \ell_{\beta_q}$
$\mathcal{U}(20, 100)$	Eq. (12)	\tilde{m}
$\mathcal{U}(-1, 1)$	Eq. (8)	μ
$\mathcal{LU}(-1.5, 0)$	Eq. (8)	σ
$\mathcal{U}(0.05, 1)$	Eq. (12)	$\chi_{\text{eff}, \text{max}}$
$\mathcal{U}(-1, \chi_{\text{eff}, \text{max}})$	Eq. (12)	$\chi_{\text{eff}, \text{min}}$
$\mathcal{N}(0, 3)$	Eq. (6)	β_q
$\mathcal{N}(0, 6)$	Eq. (7)	κ

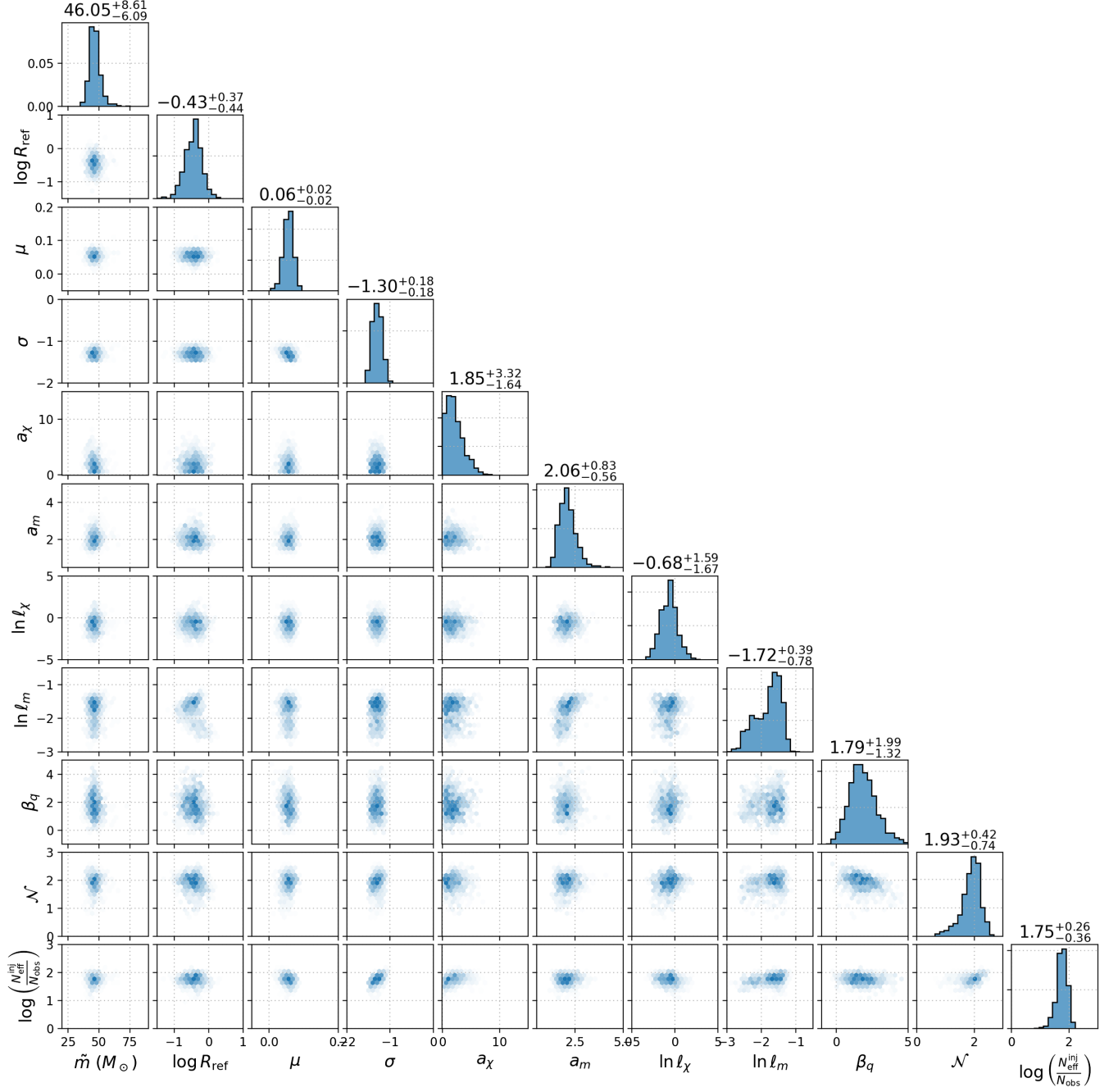


FIG. 10. Posteriors on the parameters that govern the hierarchical model where the χ_{eff} distribution is represented by a fixed Gaussian below \tilde{m} and a nonparametric distribution above \tilde{m} . In this case the priors on the kernel parameters of the GP prior corresponding to model 1 in Fig. 5 and Table I. Here $N_{\text{eff}}^{\text{inj}}/N_{\text{obs}}$ gives the total number of injection divided by the number of detections. R_{ref} represents the differential merger rate value in units of $\text{Gpc}^{-3}\text{yr}^{-1}M_{\odot}^{-1}$ evaluated at $20M_{\odot}$ and at redshift $z = 0.2$.

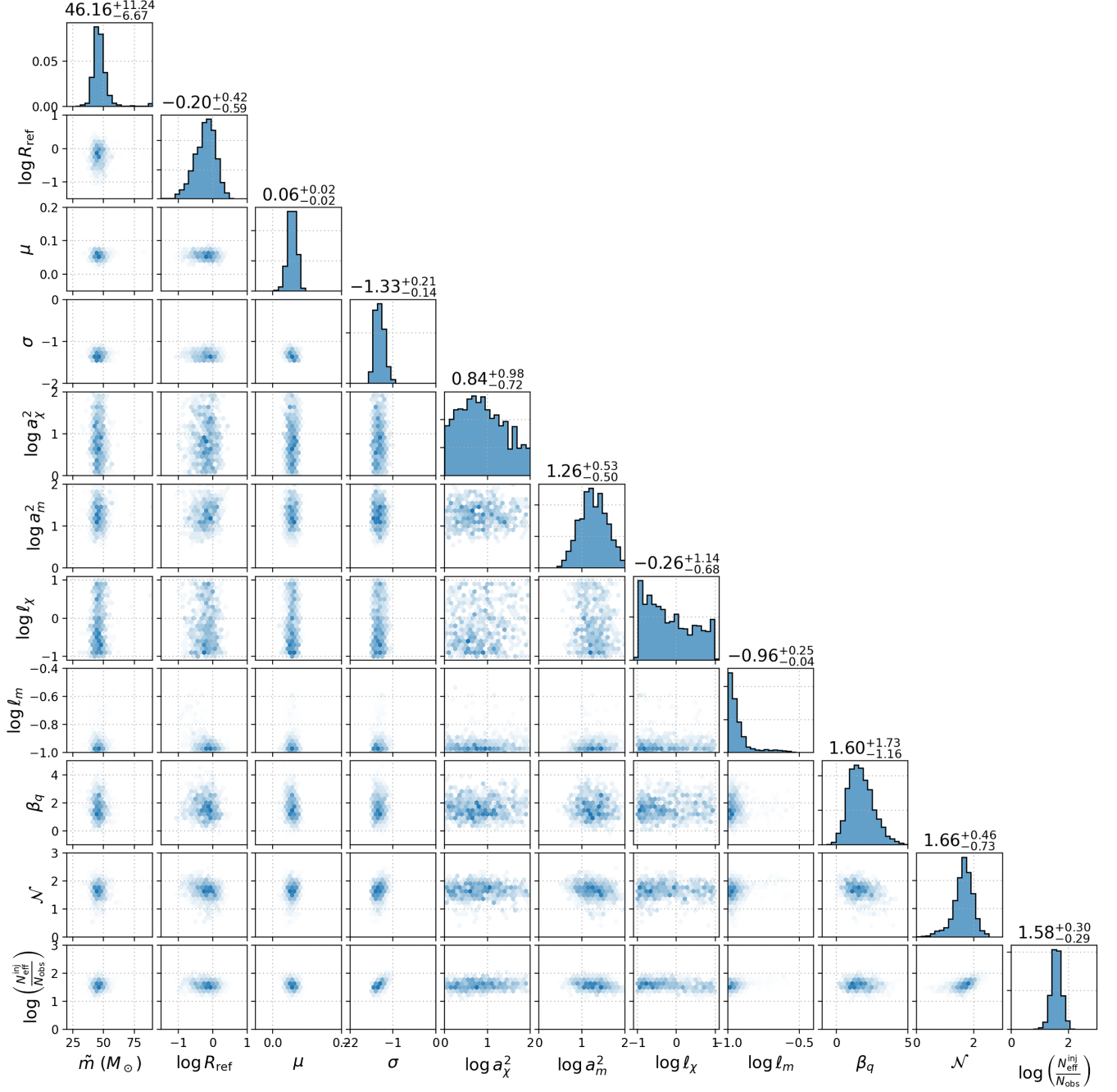


FIG. 11. Posteriors on the parameters that govern the hierarchical model where the χ_{eff} distribution is represented by a fixed Gaussian below \tilde{m} and a nonparametric distribution above \tilde{m} . In this case the priors on the kernel parameters of the GP prior corresponding to model 2 in Fig. 5 and Table I. Here $N_{\text{eff}}^{\text{inj}}/N_{\text{obs}}$ gives the total number of injection divided by the number of detections. R_{ref} represents the differential merger rate value in units of $\text{Gpc}^{-3}\text{yr}^{-1}M_{\odot}^{-1}$ evaluated at $20M_{\odot}$ and at redshift $z = 0.2$.

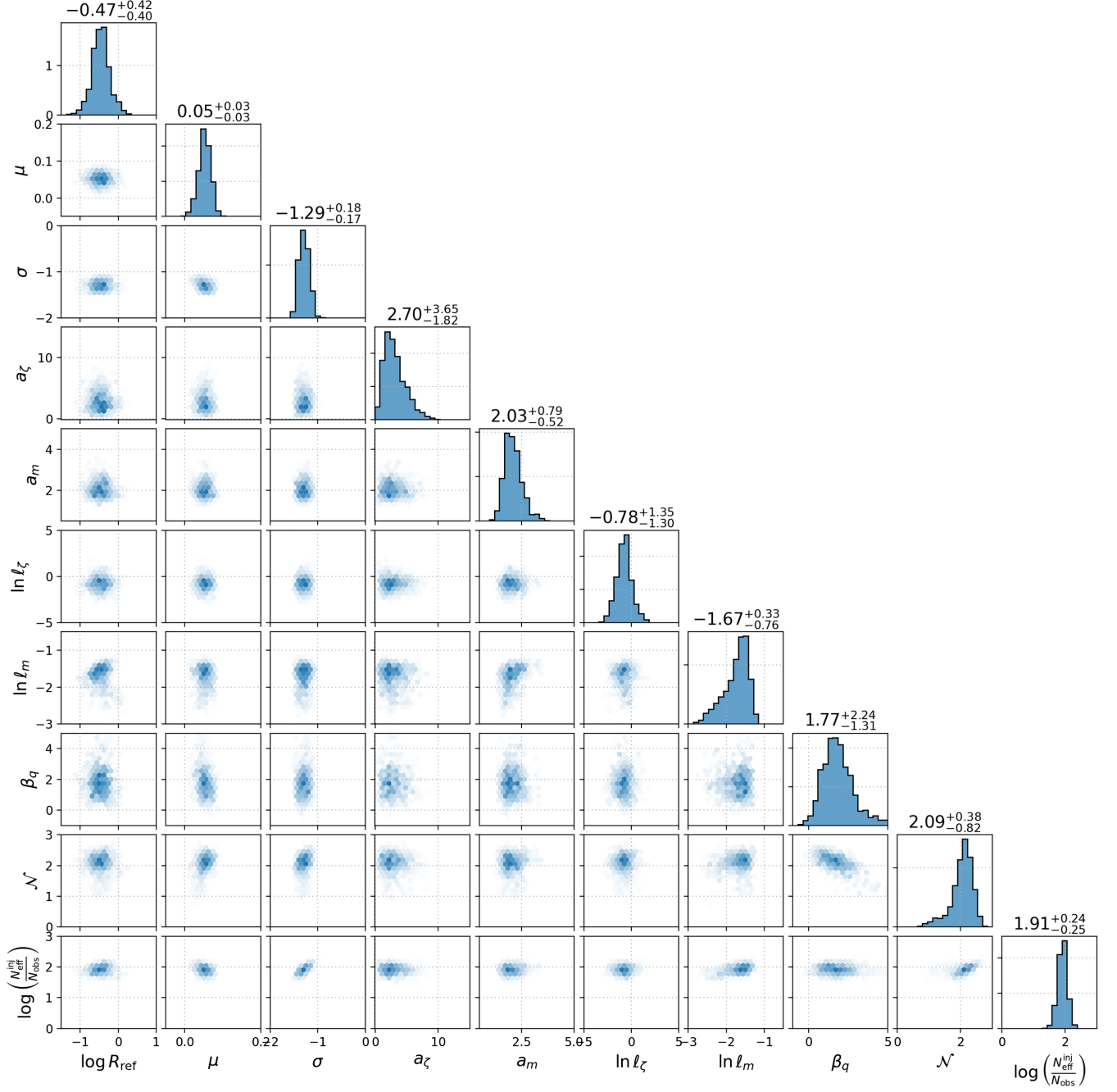


FIG. 12. Posteriors on the key parameters that govern the hierarchical model of Eq. (8) where the χ_{eff} distribution is represented as mixture between a Gaussian and a uniform distribution with the mixing fraction modeled as a GP over $\log m_1$. Here $N_{\text{eff}}^{\text{inj}}/N_{\text{obs}}$ gives the total number of injection divided by the number of detections. R_{ref} represents the differential merger rate value in units of $\text{Gpc}^{-3}\text{yr}^{-1}M_{\odot}^{-1}$ evaluated at $20M_{\odot}$ and at redshift $z = 0.2$.

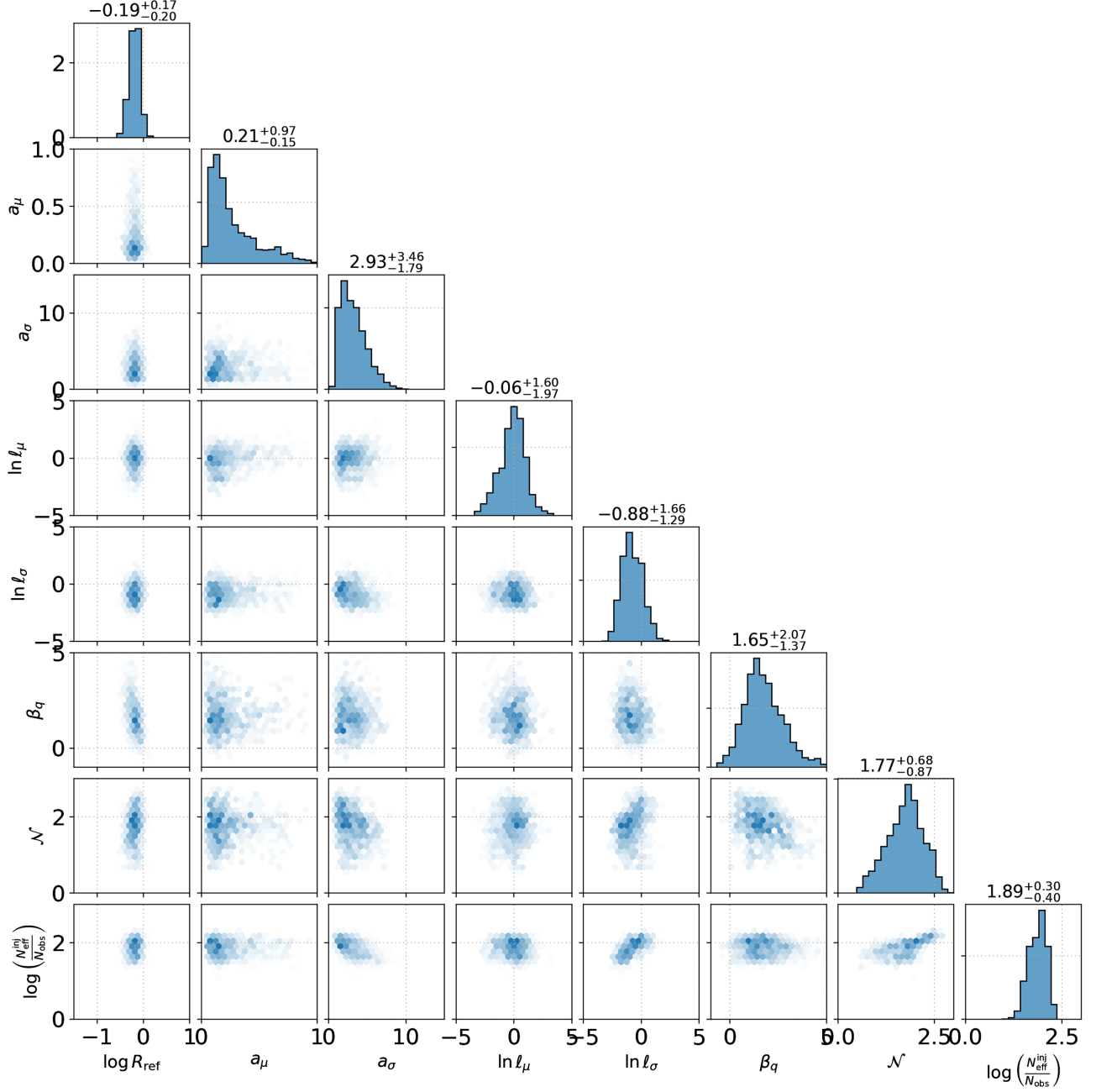


FIG. 13. Posteriors on the key parameters that govern the hierarchical model of Sec. III B where the χ_{eff} distribution is represented as truncated Gaussian with mass dependent mean and variance, both modeled nonparametrically as a GP. Here $N_{\text{eff}}^{\text{inj}}/N_{\text{obs}}$ gives the total number of injection divided by the number of detections. R_{ref} represents the differential merger rate value in units of $\text{Gpc}^{-3}\text{yr}^{-1}M_{\odot}^{-1}$ evaluated at $20M_{\odot}$ and at redshift $z = 0.2$.

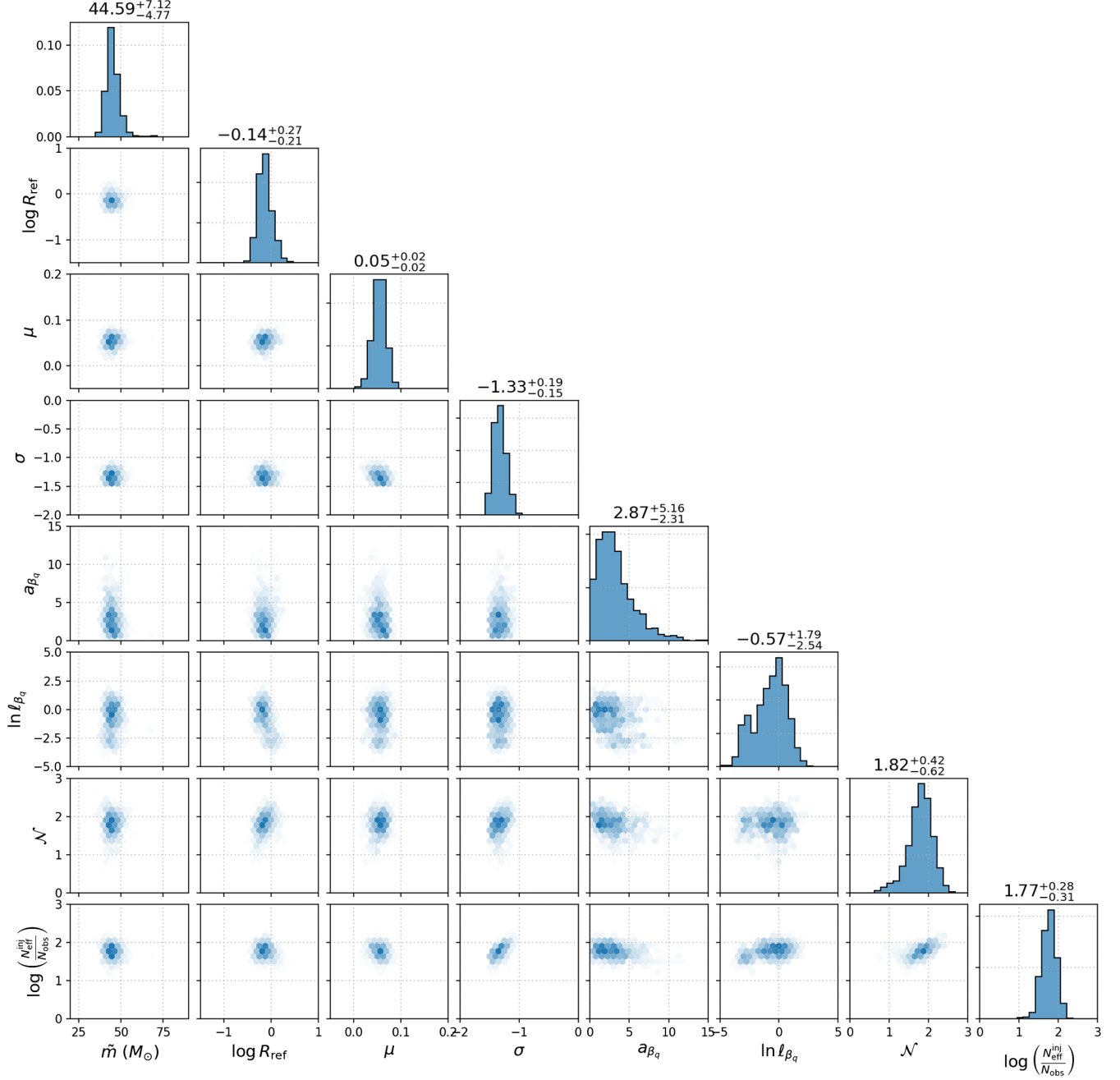


FIG. 14. Posteriors on the key parameters that govern the hierarchical model of Appendix B with a mass dependent power-law index of the mass ratio distribution, which is represented nonparametrically using a GP over a mass grid. Here $N_{\text{eff}}^{\text{inj}}/N_{\text{obs}}$ gives the total number of injection divided by the number of detections. R_{ref} represents the differential merger rate value in units of $\text{Gpc}^{-3}\text{yr}^{-1}M_{\odot}^{-1}$ evaluated at $20M_{\odot}$ and at redshift $z = 0.2$.

- [1] B. P. Abbott *et al.*, *Phys. Rev. X* **9**, 031040 (2019).
 [2] LIGO Scientific and Virgo Collaborations, *Astrophys. J. Lett.* **882**, L24 (2019).

- [3] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *arXiv:2010.14533*.
 [4] B. P. Abbott *et al.*, *Phys. Rev. X* **13**, 041039 (2023).

- [5] B. P. Abbott *et al.*, *Phys. Rev. X* **11**, 021053 (2021).
- [6] M. Spera, M. Mapelli, and A. Bressan, *Mon. Not. R. Astron. Soc.* **451**, 4086 (2015).
- [7] K. Belczynski, D. E. Holz, T. Bulik, and R. O’Shaughnessy, *Nature (London)* **534**, 512 (2016).
- [8] S. Stevenson, F. Ohme, and S. Fairhurst, *Astrophys. J.* **810**, 58 (2015).
- [9] P. Marchant, N. Langer, P. Podsiadlowski, T. M. Tauris, and T. J. Moriya, *Astron. Astrophys.* **588** (2016).
- [10] E. Farrell, J. H. Groh, R. Hirschi, L. Murphy, E. Kaiser, S. Ekström, C. Georgy, and G. Meynet, *Mon. Not. R. Astron. Soc.* **502**, L40 (2021).
- [11] K. Belczynski, R. Hirschi, E. A. Kaiser, J. Liu, J. Casares, Y. Lu, R. O’Shaughnessy, A. Heger, S. Justham, and R. Soria, *Astrophys. J.* **890**, 113 (2020).
- [12] S. E. Woosley, *Astrophys. J.* **836**, 244 (2017).
- [13] M. Spera and M. Mapelli, *Mon. Not. R. Astron. Soc.* **470**, 4739 (2017).
- [14] R. Farmer, M. Renzo, S. E. de Mink, M. Fishbach, and S. Justham, *Astrophys. J. Lett.* **902**, L36 (2020).
- [15] D. D. Hendriks, L. A. C. van Son, M. Renzo, R. G. Izzard, and R. Farmer, *Mon. Not. R. Astron. Soc.* **526**, 4130 (2023).
- [16] R. Farmer, M. Renzo, S. E. de Mink, P. Marchant, and S. Justham, *Astrophys. J.* **887**, 53 (2019).
- [17] S.-C. Leung, K. Nomoto, and S. Blinnikov, *Astrophys. J.* **887**, 72 (2019).
- [18] B. P. Abbott *et al.*, *Astrophys. J. Lett.* **913**, L7 (2021).
- [19] B. Edelman, Z. Doctor, and B. Farr, *Astrophys. J. Lett.* **913**, L23 (2021).
- [20] B. P. Abbott *et al.*, *Phys. Rev. X* **13**, 011048 (2023).
- [21] A. Ray, I. M. Hernandez, S. Mohite, J. Creighton, and S. Kapadia, *Astrophys. J.* **957**, 37 (2023).
- [22] T. A. Callister and W. M. Farr, *Phys. Rev. X* **14**, 021005 (2024).
- [23] Y.-J. Li, Y.-Z. Wang, S.-P. Tang, and Y.-Z. Fan, *Phys. Rev. Lett.* **133**, 051401 (2024).
- [24] E. R. J. Winch, J. S. Vink, E. R. Higgins, and G. N. Sabhahit, *Mon. Not. R. Astron. Soc.* **529**, 2980 (2024).
- [25] In the literature, such events are often termed “hierarchical mergers,” implying growth through multiple successive mergers. However, we avoid this terminology, as theoretical models predict that the vast majority of BHs in the mass gap originate from a single prior merger—i.e., they are predominantly second-generation BHs.
- [26] R. M. O’Leary, F. A. Rasio, J. M. Fregeau, N. Ivanova, and R. O’Shaughnessy, *Astrophys. J.* **637**, 937 (2006).
- [27] F. Antonini and F. A. Rasio, *Astrophys. J.* **831**, 187 (2016).
- [28] C. L. Rodriguez, M. Morscher, B. Pattabiraman, S. Chatterjee, C.-J. Haster, and F. A. Rasio, *Phys. Rev. Lett.* **115**, 051101 (2015).
- [29] D. Gerosa and E. Berti, *Phys. Rev. D* **95**, 124046 (2017).
- [30] Y. Yang, I. Bartos, V. Gayathri, K. E. S. Ford, Z. Haiman, S. Klimentko, B. Kocsis, S. Márka, Z. Márka, B. McKernan *et al.*, *Phys. Rev. Lett.* **123**, 181101 (2019).
- [31] F. Antonini, M. Gieles, and A. Gualandris, *Mon. Not. R. Astron. Soc.* **486**, 5008 (2019).
- [32] G. Fragione and B. Kocsis, *Mon. Not. R. Astron. Soc.* **486**, 4781 (2019).
- [33] M. Dall’Amico, M. Mapelli, U. N. Di Carlo, Y. Bouffanais, S. Rastello, F. Santoliquido, A. Ballone, and M. Arca Sedda, *Mon. Not. R. Astron. Soc.* **508**, 3045 (2021).
- [34] M. Mapelli, M. Dall’Amico, Y. Bouffanais, N. Giacobbo, M. Arca Sedda, M. C. Artale, A. Ballone, U. N. Di Carlo, G. Iorio, F. Santoliquido *et al.*, *Mon. Not. R. Astron. Soc.* **505**, 339 (2021).
- [35] H. Tagawa, Z. Haiman, I. Bartos, B. Kocsis, and K. Omukai, *Mon. Not. R. Astron. Soc.* **507**, 3362 (2021).
- [36] M. Fishbach, D. E. Holz, and B. Farr, *Astrophys. J. Lett.* **840**, L24 (2017).
- [37] V. Baibhav, D. Gerosa, E. Berti, K. W. K. Wong, T. Helfer, and M. Mould, *Phys. Rev. D* **102**, 043002 (2020).
- [38] P. Ajith, M. Hannam, S. Husa, Y. Chen, B. Brügmann, N. Dorband, D. Müller, F. Ohme, D. Pollney, C. Reisswig *et al.*, *Phys. Rev. Lett.* **106**, 241101 (2011).
- [39] S. Vitale, R. Lynch, J. Veitch, V. Raymond, and R. Sturani, *Phys. Rev. Lett.* **112** (2014).
- [40] W. M. Farr, S. Stevenson, M. C. Miller, I. Mandel, B. Farr, and A. Vecchio, *Nature (London)* **548**, 426 (2017).
- [41] J. Golomb and C. Talbot, *Phys. Rev. D* **108**, 103009 (2023).
- [42] S. Vitale, R. Lynch, R. Sturani, and P. Graff, *Classical Quantum Gravity* **34**, 03LT01 (2017).
- [43] C. L. Rodriguez, M. Zevin, C. Pankow, V. Kalogera, and F. A. Rasio, *Astrophys. J. Lett.* **832**, L2 (2016).
- [44] P. Schmidt, F. Ohme, and M. Hannam, *Phys. Rev. D* **91**, 024043 (2015).
- [45] M. Zevin, S. S. Bavera, C. P. L. Berry, V. Kalogera, T. Fragos, P. Marchant, C. L. Rodriguez, F. Antonini, D. E. Holz, and C. Pankow, *Astrophys. J.* **910**, 152 (2021).
- [46] D. Gerosa and M. Fishbach, *Nat. Astron.* **5**, 749 (2021).
- [47] T. A. Callister, C.-J. Haster, K. K. Y. Ng, S. Vitale, and W. M. Farr, *Astrophys. J. Lett.* **922**, L5 (2021).
- [48] T. A. Callister, S. J. Miller, K. Chatziioannou, and W. M. Farr, *Astrophys. J. Lett.* **937**, L13 (2022).
- [49] I. M. Romero-Shaw, P. D. Lasky, and E. Thrane, *Astrophys. J.* **940**, 171 (2022).
- [50] C. Hoy, S. Fairhurst, M. Hannam, and V. Tiwari, *Astrophys. J.* **928**, 75 (2022).
- [51] C. Hoy, S. Fairhurst, and I. Mandel, *Phys. Rev. D* **111**, 023037 (2025).
- [52] C. Kimball, C. Talbot, C. P. L. Berry, M. Carney, M. Zevin, E. Thrane, and V. Kalogera, *Astrophys. J.* **900**, 177 (2020).
- [53] B. Liu and D. Lai, *Mon. Not. R. Astron. Soc.* **502**, 2049 (2021).
- [54] G. Franciolini and P. Pani, *Phys. Rev. D* **105**, 123024 (2022).
- [55] G. Pierra, S. Mastrogiovanni, and S. Perriès, *Astron. Astrophys.* **692**, A80 (2024).
- [56] J. Heinzl, M. Mould, S. Álvarez-López, and S. Vitale, *Phys. Rev. D* **111**, 063043 (2025).
- [57] J. Heinzl, M. Mould, and S. Vitale, *Phys. Rev. D* **111**, L061305 (2025).
- [58] F. Antonini, I. M. Romero-Shaw, and T. Callister, *Phys. Rev. Lett.* **134**, 011401 (2025).
- [59] V. Baibhav, E. Berti, D. Gerosa, M. Mould, and K. W. K. Wong, *Phys. Rev. D* **104**, 084002 (2021).
- [60] G. Fragione, A. Loeb, and F. A. Rasio, *Astrophys. J. Lett.* **895**, L15 (2020).

- [61] H. Tagawa, B. Kocsis, Z. Haiman, I. Bartos, K. Omukai, and J. Samsing, *Astrophys. J.* **908**, 194 (2021).
- [62] Abbott *et al.*, [arXiv:2111.03634](https://arxiv.org/abs/2111.03634).
- [63] The LIGO Scientific, Virgo, and KAGRA Collaborations, 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 (2021), [10.5281/zenodo.5636816](https://zenodo.org/record/5636816).
- [64] M. Fishbach, D. E. Holz, and W. M. Farr, *Astrophys. J. Lett.* **863**, L41 (2018).
- [65] T. Callister, M. Fishbach, D. E. Holz, and W. M. Farr, *Astrophys. J. Lett.* **896**, L32 (2020).
- [66] J. Godfrey, B. Edelman, and B. Farr, [arXiv:2304.01288](https://arxiv.org/abs/2304.01288).
- [67] I. Mandel and S. E. De Mink, *Mon. Not. R. Astron. Soc.* **458**, 2634 (2016).
- [68] I. Bartos, B. Kocsis, Z. Haiman, and S. Márka, *Astrophys. J.* **835**, 165 (2017).
- [69] J. Sadiq, T. Dent, and A. Lorenzo-Medina, [arXiv:2506.02250](https://arxiv.org/abs/2506.02250).
- [70] S. Biscoveanu, T. A. Callister, C.-J. Haster, K. K. Y. Ng, S. Vitale, and W. M. Farr, *Astrophys. J. Lett.* **932**, L19 (2022).
- [71] J. Heinkel, S. Biscoveanu, and S. Vitale, *Phys. Rev. D* **109**, 103006 (2024).
- [72] R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, N. Adhikari, R. X. Adhikari, V. B. Adya, C. Affeldt, D. Agarwal *et al.*, *Phys. Rev. X* **13**, 011048 (2023).
- [73] Y.-J. Li, Y.-Z. Wang, S.-P. Tang, T. Chen, and Y.-Z. Fan, *Astrophys. J.* **987**, 65 (2025).
- [74] P. Hut, *Astron. Astrophys.* **99**, 126 (1981).
- [75] F. R. N. Schneider, P. Podsiadlowski, and E. Laplace, *Astrophys. J. Lett.* **950**, L9 (2023).
- [76] S. Chatterjee, C. L. Rodriguez, and F. A. Rasio, *Astrophys. J.* **834**, 68 (2017).
- [77] C. L. Rodriguez, P. Amaro-Seoane, S. Chatterjee, and F. A. Rasio, *Phys. Rev. Lett.* **120**, 151101 (2018).
- [78] S. Rastello, P. Amaro-Seoane, M. Arca-Sedda, R. Capuzzo-Dolcetta, G. Fragione, and I. Tosta e Melo, *Mon. Not. R. Astron. Soc.* **483**, 1233 (2019).
- [79] U. N. Di Carlo, M. Mapelli, M. Pasquato, S. Rastello, A. Ballone, M. Dall’Amico, N. Giacobbo, G. Iorio, M. Spera, S. Torniamenti *et al.*, *Mon. Not. R. Astron. Soc.* **507**, 5132 (2021).
- [80] S. Banerjee, *Mon. Not. R. Astron. Soc.* **503**, 3371 (2021).
- [81] F. Antonini and M. Gieles, *Mon. Not. R. Astron. Soc.* **492**, 2936 (2020).
- [82] F. Antonini, M. Gieles, F. Dosopoulou, and D. Chattopadhyay, *Mon. Not. R. Astron. Soc.* **522**, 466 (2023).
- [83] B. McKernan and K. E. S. Ford, *Mon. Not. R. Astron. Soc.* **531**, 3479 (2024).
- [84] I. Mandel, *Mon. Not. R. Astron. Soc.* **456**, 578 (2016).
- [85] J. Fuller and L. Ma, *Astrophys. J. Lett.* **881**, L1 (2019).
- [86] J. Riley, I. Mandel, P. Marchant, E. Butler, K. Nathaniel, C. Neijssel, S. Shortt, and A. Vigna-Gómez, *Mon. Not. R. Astron. Soc.* **505**, 663 (2021).
- [87] M. Dall’Amico, M. Mapelli, G. Iorio, G. Costa, S. Charlot, E. Korb, C. Sgalletta, and M. Lecroq, *Astron. Astrophys.* **695**, A221 (2025).
- [88] H. Tong, M. Fishbach, and E. Thrane, *Astrophys. J.* **985**, 220 (2025).
- [89] K. Belczynski, A. Heger, W. Gladysz, A. J. Ruiter, S. Woosley, G. Wiktorowicz, H. Y. Chen, T. Bulik, R. O’Shaughnessy, D. E. Holz *et al.*, *Astron. Astrophys.* **594**, A97 (2016).
- [90] R. Farmer, M. Renzo, S. E. de Mink, M. Fishbach, and S. Justham, *Astrophys. J. Lett.* **902**, L36 (2020).
- [91] Y. Shen, B. Guo, R. J. deBoer, E. Li, Z. Li, Y. Li, X. Tang, D. Pang, S. Adhikari, C. Basu *et al.*, *Astrophys. J.* **945**, 41 (2023).
- [92] P. Marchant and T. J. Moriya, *Astron. Astrophys.* **640**, L18 (2020).
- [93] S. E. Woosley and A. Heger, *Astrophys. J. Lett.* **912**, L31 (2021).
- [94] P. Marchant, M. Renzo, R. Farmer, K. M. W. Pappas, R. E. Taam, S. E. de Mink, and V. Kalogera, *Astrophys. J.* **882**, 36 (2019).
- [95] P. Greenfield, T. Robitaille, E. Tollerud, T. Aldcroft, K. Barbary, P. Barrett, E. Bray, N. Crighton, A. Conley, S. Conseil *et al.*, *Astropy: Community Python Library for Astronomy*, Astrophysics Source Code Library, record ascl:1304.002 (Michigan State University, Michigan, USA, 2013).
- [96] G. Ashton, M. Hübner, P. D. Lasky, C. Talbot, K. Ackley, S. Biscoveanu, Q. Chu, A. Divakarla, P. J. Easter, B. Goncharov *et al.*, *Astrophys. J. Suppl. Ser.* **241**, 27 (2019).
- [97] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Néculea, A. Paszke, J. VanderPlas, S. Wanderman-Milne *et al.*, *JAX: Composable transformations of Python+NumPy programs* (2018), [http://github.com/google/jax](https://github.com/google/jax).
- [98] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith *et al.*, *Nature (London)* **585**, 357 (2020).
- [99] D. Phan, N. Pradhan, and M. Jankowiak, [arXiv:1912.11554](https://arxiv.org/abs/1912.11554).
- [100] E. Bingham, J. P. Chen, M. Jankowiak, F. Obermeyer, N. Pradhan, T. Karaletsos, R. Singh, P. A. Szerlip, P. Horsfall, and N. D. Goodman, *J. Mach. Learn. Res.* **20**, 28 (2019).
- [101] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright *et al.*, *Nat. Methods* **17**, 261 (2020).
- [102] <https://zenodo.org/records/11254021>
- [103] V. Tiwari and S. Fairhurst, *Astrophys. J. Lett.* **913**, L19 (2021).
- [104] S. Galadage and A. Lamberts, *Astron. Astrophys.* **694**, A186 (2025).
- [105] I. Mandel, W. M. Farr, and J. R. Gair, *Mon. Not. R. Astron. Soc.* **486**, 1086 (2019).
- [106] Abbott *et al.*, *Phys. Rev. D* **109**, 022001 (2024).
- [107] R. Essick and W. Farr, [arXiv:2204.00461](https://arxiv.org/abs/2204.00461).