

This is the accepted manuscript made available via CHORUS. The article has been published as:

Distribution of effective spins and masses of binary black holes from the LIGO and Virgo O1–O3a observing runs

Javier Roulet, Horng Sheng Chia, Seth Olsen, Liang Dai, Tejaswi Venumadhav, Barak Zackay, and Matias Zaldarriaga

Phys. Rev. D **104**, 083010 — Published 4 October 2021

DOI: [10.1103/PhysRevD.104.083010](https://doi.org/10.1103/PhysRevD.104.083010)

On the Distribution of Effective Spins and Masses of Binary Black Holes from the LIGO and Virgo O1–O3a Observing Runs

Javier Roulet,¹ Horng Sheng Chia,² Seth Olsen,¹ Liang Dai,³
Tejaswi Venumadhav,^{4,5} Barak Zackay,⁶ and Matias Zaldarriaga²

¹*Department of Physics, Princeton University, Princeton, NJ 08540, USA*

²*School of Natural Sciences, Institute for Advanced Study, Princeton, NJ 08540, USA*

³*Department of Physics, University of California, Berkeley, 366 LeConte Hall, Berkeley, CA 94720, USA*

⁴*Department of Physics, University of California at Santa Barbara, Santa Barbara, CA 93106, USA*

⁵*International Centre for Theoretical Sciences, Tata Institute of Fundamental Research, Bangalore 560089, India*

⁶*Department of Particle Physics & Astrophysics, Weizmann Institute of Science, Rehovot 76100, Israel*

The distribution of effective spin χ_{eff} , a parameter that encodes the degree of spin–orbit alignment in a binary system, has been widely regarded as a robust discriminator between the isolated and dynamical formation pathways for merging binary black holes. Until the recent release of the GWTC-2 catalog, such tests have yielded inconclusive results due to the small number of events with measurable nonzero spins. In this work, we study the χ_{eff} distribution of the binary black holes detected in the LIGO–Virgo O1–O3a observing runs. Our focus is on the degree to which the χ_{eff} distribution is symmetric about $\chi_{\text{eff}} = 0$ and whether the data provides support for a population of negative- χ_{eff} systems. We find that the χ_{eff} distribution is asymmetric at 95% credibility, with an excess of aligned-spin binary systems ($\chi_{\text{eff}} > 0$) over anti-aligned ones. Moreover, we find that there is no evidence for negative- χ_{eff} systems in the current population of binary black holes. Thus, based solely on the χ_{eff} distribution, dynamical formation is disfavored as being responsible for the entirety of the observed merging binary black holes, while isolated formation remains viable. We also study the mass distribution of the current binary black hole population, confirming that a single truncated power law distribution in the primary source-frame mass, m_{1s} , fails to describe the observations. Instead, we find that the preferred models have a steep feature at $m_{1s} \sim 40 M_{\odot}$ consistent with a step and an extended, shallow tail to high masses.

I. INTRODUCTION

The growing number of gravitational wave sources observed by the LIGO and Virgo detectors is leading to an improved picture of the astrophysical population of binary mergers. The recent release of the second Gravitational-Wave Transient Catalog, GWTC-2 [1], by the LIGO–Virgo Collaboration (LVC) has roughly tripled the sample size of observed binary black hole mergers [2–8] and is starting to offer hints about the astrophysical origin of these binary systems [9–12].

Indeed, the distribution of binary black hole parameters (e.g. masses, spins, redshift) is an observable that allows us to test models of formation pathways for these systems. Proposed scenarios include dynamical assembly and hardening of binary black holes in dense stellar environments, such as globular clusters [13–21], nuclear star clusters [22, 23], and young stellar clusters [24–27]; isolated evolution of a binary star in the galactic field, which undergoes either a common envelope phase [28–43] or a chemically homogeneous evolution [44–46]; and binary mergers prompted by interactions with a supermassive black hole [47], gas and stars in the accretion disk of an active galactic nucleus [48–50], or additional companions in higher-multiplicity systems [51–55].

Since the individual components of the dimensionless spin vectors χ_1 and χ_2 are hard to measure [56–58] and their directions generally evolve with time due to precession [59, 60], a well-known effective aligned-spin param-

eter was introduced [61–63]

$$\chi_{\text{eff}} := \frac{\chi_1 + q\chi_2}{1+q} \cdot \hat{\mathbf{L}}, \quad (1)$$

where $\hat{\mathbf{L}}$ is the unit vector along the Newtonian orbital angular momentum of the binary, $q = m_2/m_1 \leq 1$ is the mass ratio. The effective spin is motivated by the fact that it can be measured relatively precisely in the data, and is approximately conserved throughout the binary coalescence after orbit averaging [61]. No less important, two of the main broad classes of binary black hole formation channels make predictions about qualitative features of the effective spin distribution that are robust to model uncertainties. Dynamical formation channels in general predict that the spins and orbit should be isotropically distributed and uncorrelated with each other. In particular, this implies that for these systems the χ_{eff} distribution is symmetric about 0. Isolated formation channels instead predict correlations in the spins and orbit directions due to mass transfer episodes or tidal interactions between the component stars. As a result, the isolated scenario predicts a distribution of χ_{eff} with little support at negative values. Within this channel, a small fraction of mergers with negative χ_{eff} could still possibly be explained by anisotropic supernova explosions at the black holes formation, which impart a natal kick that can change the plane of the orbit and thus the value of χ_{eff} [64, 65]. However, if these kicks were strong enough to invert the direction of the orbit in a sizeable fraction of the cases, they would also unbind the binaries so fre-

quently that the observed rates would be hard to explain [66, 67].

In this work we will study in detail the degree to which these two qualitative features of the χ_{eff} distribution, namely its symmetry about 0 and support at negative values, hold for the observed sample. Both features become hard to test if black hole spins are small, which is predicted from stellar evolution models [68, 69] and is also the case of most observations. Indeed, until the recent release of the GWTC-2 catalog these simple but general tests were mostly inconclusive due to the small number of events with measurable nonzero χ_{eff} [70–73]. Including events from the O3a observing run, Abbott *et al.* [9] first reported evidence for both features in the population: they found the χ_{eff} distribution to have a positive mean and support at negative values. Together, these observations suggest that neither dynamical nor isolated formation channels alone can explain the entirety of the detections. Combining this information with the observed mass distribution, Zevin *et al.* [11], Bouffanais *et al.* [12] reached a similar conclusion, and applied a further layer of interpretation to constrain uncertain parameters of physical models of binary black hole formation. Here, we instead constrain a phenomenological description of the binary black hole population, more akin to the analysis of Abbott *et al.* [9].

The mass distribution is another observable that can inform binary black hole formation channels, as well as physical processes of stellar evolution. Of special interest is the high-mass end of the mass distribution observable by LIGO–Virgo, $m \gtrsim 40 M_{\odot}$. Due to the (pulsational) pair instability supernova process, black holes with mass between $\sim 45 M_{\odot}$ and $135 M_{\odot}$ are not expected to form from stellar collapse (“upper mass gap”) [74–78]. A natural way to produce black holes in this mass range is through mergers of lighter black holes. In dense environments these so-called “second-generation” black holes can become paired and merge again, emitting an observable gravitational wave signal. This process is contingent on retention of the remnant black hole, so its efficiency depends on the interplay between the merger kick (a recoil of the remnant black hole due to asymmetric gravitational wave emission at merger) and the local escape velocity. The magnitude of the kicks is sensitive to the spins of the merging black holes, smaller spins usually yielding smaller kicks. In turn, different types of dense environments have different escape velocities, typical numbers being $10\text{--}10^2 \text{ km s}^{-1}$ for globular clusters and up to $\sim 10^3 \text{ km s}^{-1}$ for nuclear clusters. Second generation mergers do not happen for binaries formed in isolation. Some alternative pathways to produce black holes in this mass range may involve accretion of gas [79] or extreme values of the $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$ nuclear cross section, which can shift the location of the mass gap [78], see [80] and references therein for a recent review. On the observational side, current interferometers are particularly sensitive to mergers in this high-mass region of parameter space, which makes it a promising discriminator [81]. Indeed,

some events were observed to have significant support for one or both component black holes in this mass range (e.g. GW190521, GW190602_175927, GW190706_222641, GW190519_153544, GW190929_012149 [1], GW170817A [7]). While analyses prior to O3a found evidence for a cut-off in the mass distribution at $\sim 40 M_{\odot}$ [72, 73, 81, 82], this picture changed with the inclusion of O3a and models with more structure, including a tail at high mass, became favored [9]. Here, we will also explore parametric models of the primary mass distribution in order to validate these results.

Our main findings are:

1. The χ_{eff} distribution is inconsistent with being symmetric about zero at the 95% credible level, with aligned-spin binary systems ($\chi_{\text{eff}} > 0$) predominating over those with anti-aligned spins ($\chi_{\text{eff}} < 0$). This result provides some evidence against the formation scenario in which the entire population of binary black holes has isotropically-distributed spins, as predicted if all merging binary black holes are formed dynamically in dense stellar environments;
2. We find no evidence for negative χ_{eff} in the population, in contrast to Abbott *et al.* [9]. We are able to reproduce their results, but find that the parametrized model they used in order to reach this conclusion is disfavored by the data and that the inferred presence of negative spins is contingent on this parametrization;
3. We find that the primary-mass distribution steepens at $\sim 40 M_{\odot}$ and then flattens, with an extended tail to high masses whose detailed shape is hard to constrain with current data.

This paper is organized as follows: in Section II, we describe the data investigated in this work, our sample selection criteria, and the parameter estimation method used to infer the source parameters of the binary black holes. In Section III, we conduct a model-free exploration of the data, with a special focus on the empirical distribution of χ_{eff} . In Section IV, we describe our statistical methods for model selection and apply them to several parametrized models for the distributions of the effective spin and primary mass. We conclude in Section V. We provide details of the sample of events that we use in Appendix A.

II. DATA

The data explored in this work consists of the binary black hole events reported in the LVC GWTC-1 [2] and GWTC-2 [1] catalogs, and those identified in the independent IAS O1–O2 catalog [4–7]. Some of the events reported in the IAS O1–O2 catalog have been independently confirmed by Nitz *et al.* [3, 8]. Following the main analysis conducted by the LVC in their population

study [9], we exclude GW190814 [83] in this work as it is an outlier with respect to the rest of the observed population, and for ease of comparison between our results and the LVC’s results (see Section IV). We do not include the recent 3-OGC catalog [84], which was published as this work was being completed. A summary of the events used in this work is provided in Appendix A.

Depending on how easily our models for astrophysical signals and detector noise can account for the properties of a given trigger, some detections are more statistically significant than the others. Roulet *et al.* [73] provided a framework to take this into account when using triggers of arbitrary significance. However, in order to simplify the interpretation of the results shown in Section III, we find it convenient to define a “gold sample” of events that are confidently astrophysical in origin. For a similar reason, we also exclude from the gold sample those events that happened when a detector exhibited non-Gaussian noise transients, which makes estimation of their parameters and significance more challenging. We include an event in the gold sample if (i) it was identified by at least two search pipelines with a false-alarm rate $\text{FAR} < 0.1 \text{ yr}^{-1}$; and (ii) none of the detectors exhibited non-Gaussian transient noise in its vicinity (see Appendix A for details). These criteria are neither explicitly dependent on nor expected to correlate significantly with the binary black hole intrinsic parameters; as such, our gold sample constitutes an unbiased representation of the distribution for the intrinsic parameters of detectable mergers. Indeed, as we shall see in Section IV, our conclusions are not strongly affected by this choice of sample. Out of the total 55 events considered in this work, 30 are in the gold sample (see Appendix A).

We infer the source parameters of each binary system with the `IMRPhenomXPHM` model, which describes the gravitational waves emitted by a quasi-circular binary black hole [85]. This model accounts for spin-orbit precession and the $(\ell, |m|) = \{(2, 2), (2, 1), (3, 3), (3, 2), (4, 4)\}$ harmonics of the gravitational radiation. We use the relative binning algorithm to evaluate the likelihood [86], and `PyMultiNest` [87] to sample the posterior distribution. For the events identified near non-Gaussian transient noise (summarized in Ref. [1, table V]), we do not make special mitigation efforts, though we verify that we obtain parameter estimation results that are similar to those reported by Abbott *et al.* [1], who applied glitch subtraction algorithms before performing parameter estimations [88–90].

For each event, we sample the posterior distribution using a prior that is uniform in detector-frame component masses, χ_{eff} and luminosity volume. For the remaining spin components, we adopt a uniform prior for the poorly-measured variable $\chi_{\text{diff}} := (q\chi_1 - \chi_2) \cdot \hat{\mathbf{L}}/(1+q)$, conditioned on χ_{eff} and enforcing the Kerr limit on the individual spin magnitudes, $|\chi_1| \leq 1$ and $|\chi_2| \leq 1$. χ_{eff} and χ_{diff} together determine the two spin components that are aligned with the orbital angular momentum, χ_{1z} and χ_{2z} . We then take the prior of the in-plane spin com-

ponents of the black holes, χ_{ix} and χ_{iy} with $i = 1, 2$, to be uniformly distributed in the disk $\chi_{ix}^2 + \chi_{iy}^2 \leq 1 - \chi_{iz}^2$.

Our parameter estimation results are broadly consistent with LVC’s after accounting for the differences in spin priors, with two notable exceptions. Firstly, we find that the posterior distribution for GW151226 [91] significantly changes towards more unequal mass ratio, larger positive χ_{eff} and more misaligned primary spin when higher harmonics and precession are included in the parameter estimation [92]. Another remarkable event is GW190521, which was reported to have component source-frame masses $m_{1s} = 85_{-14}^{+21} M_{\odot}$, $m_{2s} = 66_{-18}^{+17} M_{\odot}$ by the LVC [93]. Using a different prior for the masses and distance, and allowing for a broader parameter range, Nitz and Capano [94] found a qualitatively different trimodal solution, with roughly similar total mass and peaks at $q \sim 1/2, 1/5$, and $1/12$. Instead, we find a bimodal solution which is approximately consistent with the first two of these modes [95], similar to that reported in [96].

III. MODEL-FREE EXPLORATION

In this section, we carry out a model-free exploration of the data. Our emphasis is on the symmetry, or lack thereof, between positive and negative values of χ_{eff} in the observed χ_{eff} distribution. We also investigate if the data requires a distribution with support at negative values of χ_{eff} . To ease the interpretation of the plots shown in this section, we shall restrict ourselves to the events identified in the gold sample (see Section II). We defer a model-dependent analysis of the data to Section IV.

A. Support for nonzero χ_{eff}

We first test the simplest hypothesis that all binary black holes have $\chi_{\text{eff}} = 0$, with any apparent deviation away from zero arising due to measurement uncertainty. This test is motivated by the fact that, while the χ_{eff} measurements of some of the events have most of their support at $\chi_{\text{eff}} < 0$ (GW170121, GW150914, GW170818, GW190421.213856, GW170104, GW151012, GW190915.235702, GW170727, GW190521, GW190408.181802), none of them confidently excludes $\chi_{\text{eff}} = 0$. In the left panel of Fig. 1, we explore whether the observed scatter in the χ_{eff} distribution is consistent with noisy measurements of a $\chi_{\text{eff}} = 0$ population. We plot the empirical distribution of the quantity $\langle \chi_{\text{eff}} \rangle / \sigma$, i.e. the mean χ_{eff} of each of the event’s posterior samples divided by their standard deviation, and compare it with the cumulative of a standard Gaussian distribution with zero mean and amplitude N_0 , where N_0 is the number of events in this distribution. Provided that the likelihood is approximately Gaussian as a function of χ_{eff} , these distributions should match if the true χ_{eff} were 0. In particular, with the current number of observed events,

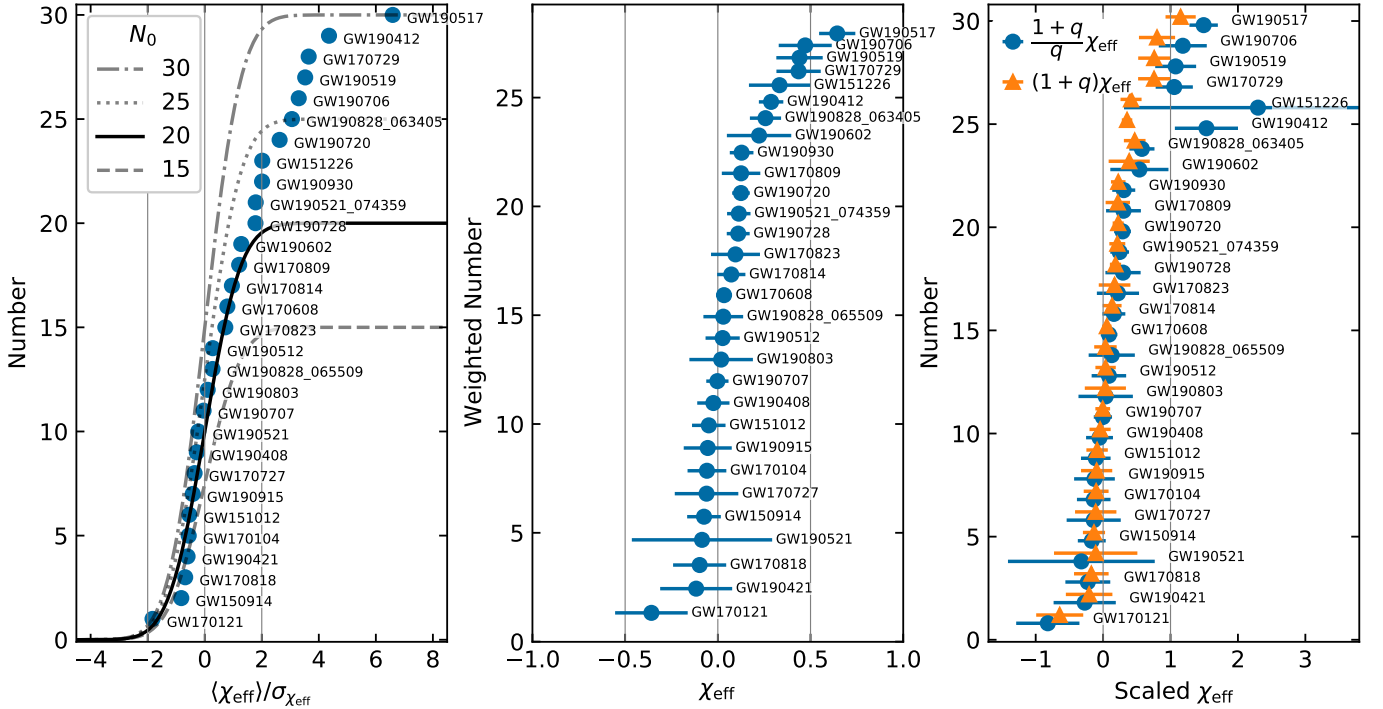


FIG. 1. Empirical spin distributions of the events in the gold sample (see Appendix A). For each event, parameter estimation samples were obtained using a uniform prior in χ_{eff} . To avoid clutter, event names were abbreviated when this did not cause ambiguity. *Left panel:* Mean effective spin scaled by the standard deviation for each event’s posterior. We see that about $N_0 \approx 20$ events in the gold sample are consistent with noisy measurements of a $\chi_{\text{eff}} = 0$ subpopulation, but a tail in the positive χ_{eff} end of the distribution is clearly needed in order to accommodate the remaining ≈ 10 events. Conversely, no such tail seems to be needed at the negative end. *Middle panel:* χ_{eff} distribution, where markers and error bars indicate mean and standard deviation. In the cumulative, each event is weighted by the ratio of the event’s sensitive volume to a similar event with zero spins in order to cancel spin selection effects. We see that there are several events with small but well-measured $\chi_{\text{eff}} > 0$ for which spin selection effects are not important. *Right panel:* ratio of observed χ_{eff} to its characteristic value if strong tides were acted on the secondary (blue circles) or primary (orange triangles) black hole progenitor. A number of events are inconsistent with any of these variables being 0 or 1, thereby excluding the strong-tide model as the only mechanism generating black hole spins.

we would not expect to find events that are more than 2σ away from $\chi_{\text{eff}} = 0$. In the left panel of Fig. 1, we observe that although $N_0 \approx 20$ out of the 30 events in the gold sample are consistent with noisy measurements of a $\chi_{\text{eff}} = 0$ distribution, there is an excess of about 10 events with $\chi_{\text{eff}} > 0$ that cannot be explained by measurement uncertainty. On the other hand, no such tail seems to be needed in the $\chi_{\text{eff}} < 0$ interval.

A phenomenon that will be important in the interpretation of this tail is the so-called “orbital hangup” [97], which entails that other parameters being equal, mergers with large and positive values of χ_{eff} have a louder emission of gravitational waves. This effect induces a selection bias, as events with $\chi_{\text{eff}} > 0$ are detectable to larger distances. Under the hypothesis tested in this Section, however, all events have $\chi_{\text{eff}} = 0$ and thus the bias is unimportant.¹

B. Symmetry of the χ_{eff} distribution

Due to the orbital-hangup effect, the observed excess of $\chi_{\text{eff}} > 0$ events relative to those with $\chi_{\text{eff}} < 0$ in the left panel of Fig. 1 does not immediately imply that the astrophysical χ_{eff} distribution is asymmetric about $\chi_{\text{eff}} = 0$. This effect leads to a selection bias favoring more observations of $\chi_{\text{eff}} > 0$ events, even if the astrophysical χ_{eff} distribution is symmetric **about zero** [72, 98]. The observed excess of $\chi_{\text{eff}} > 0$ events thus

ter through the parameter estimation prior: in order to match the observed amplitude of a signal, higher- χ_{eff} solutions are located farther out in distance and thus have more phase space volume available. In other words, a flat prior for the astrophysical χ_{eff} distribution is implicitly skewed towards positive χ_{eff} values when conditioned on the strain amplitude measured at the detector. As a result, noisy measurements of a $\chi_{\text{eff}} = 0$ distribution would be slightly biased towards $\chi_{\text{eff}} > 0$. However, under the hypothesis that the true $\chi_{\text{eff}} = 0$ this effect is small, as the measured χ_{eff} would be small.

¹ Strictly speaking, this observational bias may in principle en-

requires careful interpretation. In the middle panel of Fig. 1, we plot the empirical distribution of χ_{eff} correcting for this observational bias. The bias is computed as follows: for each event, we compute the weight factor

$$w = \langle V_{\text{no spin}}/V \rangle, \quad (2)$$

which is inversely proportional to the sensitive volume V to the corresponding event. Here, we approximate V of a source that has (detector-frame) intrinsic parameters $\theta_{\text{int}} = \{m_1, m_2, \chi_{1z}, \chi_{2z}\}$ as

$$V(\theta_{\text{int}}) \propto \rho_0^3(\theta_{\text{int}}), \quad (3)$$

with ρ_0 the single-detector signal-to-noise ratio (SNR) of an overhead, face-on source at a fiducial luminosity distance with a fiducial sensitivity. $V_{\text{no spin}}$ is defined similarly but with $\chi_{1z} = \chi_{2z} = 0$. For simplicity, we set the in-plane spin components to zero and neglect cosmological evolution throughout this computation. We then average the ratio of these two volumes over the posterior distribution of each event in order to obtain the weight w . In the middle panel of Fig. 1, we see that many of the events that deviate most significantly away from $\chi_{\text{eff}} = 0$ have small values of χ_{eff} and hence a small impact in the sensitive volume. In particular, GW190728.064510, GW190521.074359, GW190720.000836, GW190930.133541, GW190828.063405 and GW190412 are $\gtrsim 2\sigma$ away from $\chi_{\text{eff}} = 0$ and have relatively small values of $\chi_{\text{eff}} \sim 0.1$ – 0.25 . The vertical spacing between events in this plot is given by the volume weight w of the event: for the first four of these events w is approximately 0.95, and for the last two approximately 0.75. Since these are small volume corrections, there is no compelling reason as to why the same number of corresponding events on the negative side of χ_{eff} should not be seen, for an astrophysical χ_{eff} distribution that is symmetric about zero. There are also events that are more than 2σ away from zero χ_{eff} but with relatively large values of $\chi_{\text{eff}} \sim 0.5$ (GW170729, GW190519.153544, GW190706.222641 and GW190517.055101). The selection effect for these events is more appreciable, $w \sim 0.5$, so it would be easier to miss similar events with the opposite sign of χ_{eff} . Altogether, the left and middle panels of Fig. 1 hint that the empirical effective spin distribution is consistent with a distribution with no support for negative spins, but not so much with one symmetric about $\chi_{\text{eff}} = 0$.

C. Testing tidal models

Finally, we explore whether the observed events with positive χ_{eff} can be explained by a simple model of tides acting on the progenitor of one of the component black holes. The simplest and most extreme model for tides assumes that tides sourced by the companion are either very efficient at spinning up the progenitor star or negligible depending on the orbital separation after a common envelope phase, because tidal torques are very sensitive

to the orbital separation. Then, a fraction of the component black holes would come from tidally-torqued progenitors and would have a large, aligned spin $\chi_z \approx 1$ [99, 100]. If, barring tides, natal black hole spins were small [68, 69], the χ_{eff} distribution would have peaks at $\chi_{\text{eff}} = 0, q/(1+q), 1/(1+q), 1$ when the tides were inefficient, torqued the progenitor of the secondary black hole, torqued the progenitor of the primary, or torqued both, respectively. In the right panel of Fig. 1 we show the empirical distribution of χ_{eff} rescaled by the value under the hypotheses that either the secondary or the primary black hole is maximally spinning and aligned with the orbit. We find that several of the events with well-measured nonzero spin do not seem to be well explained by this model (GW190728.064510, GW190521.074359, GW190720.000836, GW170809, GW190930.133541 and GW190828.063405). This is in agreement with earlier findings that either a less extreme model of tidal torques (as argued in [69, 101]) or a distribution of natal spins with some dispersion is needed in order to explain the observed spins with tides [73].

IV. MODEL SELECTION

In order to validate and quantify our findings in Section III, in this section we perform selection of parametric models for the observed binary black hole population. We first provide a brief review of our statistical framework, and then constrain the parameters of several models for the astrophysical effective spin and primary mass distributions.

A. Statistical framework

Following Roulet *et al.* [73], we evaluate the likelihood $P(\{d_i\} | \lambda)$ of an observed set of triggers $\{d_i\}$, given a phenomenological population model λ for the distributions of binary black hole source parameters, as:

$$P(\{d_i\} | \lambda) \propto e^{-N_a(\lambda)} \prod_{i=1}^{N_{\text{trig}}} \left(\left. \frac{dN_a(\lambda)}{dN_a(\lambda_0)} \right|_{d_i} p_{\text{astro},i}(\lambda_0) + 1 - p_{\text{astro},i}(\lambda_0) \right). \quad (4)$$

Here, $N_a(\lambda)$ is the expected number of triggers of astrophysical origin under the population model λ (as opposed to detector noise), over a fixed and arbitrarily chosen detection threshold; $dN_a(\lambda)/dN_a(\lambda_0)|_{d_i}$ is the ratio of expected densities, in data space, of astrophysical triggers similar to that of the i th event between the population model λ and a fixed, arbitrary reference model λ_0 ; and $p_{\text{astro},i}(\lambda_0)$ is the probability of astrophysical origin of the i th trigger under the reference population model. The data space contains observable quantities that carry information about the astrophysical population, like measured detector strains and derived detection statistics.

All the quantities described above depend on the search pipeline used; in addition, $N_a(\lambda)$ and the set of triggers itself depend on the detection threshold chosen. Three ingredients are required in order to estimate these quantities: a set of software injections labeled by whether they exceed the detection threshold, to quantify the sensitivity of the search; posterior samples characterizing the parameters of each individual event; and the set of $\{p_{\text{astro},i}(\lambda_0)\}$ encoding the events' significance [73].

Equation (4) naturally factors into the product of likelihoods from searches on disjoint datasets, such as different observing runs. Since the full strain data from observing runs O1 and O2 are publicly accessible [102], for these data sets we base our analysis on our searches for binary black holes [7, 103, 104]. The strain data for O3a has also recently been released, and analyzed by Nitz *et al.* [84] when this work was close to completion; we do not include these results here. The LVC provides a set of software injections with FAR estimates from the search pipelines they used (cWB, GstLAL, PyCBC and PyCBC BBH) [105], and the GWTC-2 catalog itself which reports $\{p_{\text{astro},i}\}$ for the latter 3 pipelines [1]. For O3a we use these data products, which are adequate for computing the quantities in Eq. (4) with the following caveat. Our method requires knowing $\{p_{\text{astro},i}(\lambda_0)\}$ under some specific astrophysical model, which was not explicated in the GWTC-2 release. We take two alternative approaches: (i) we conservatively consider only O3a events that are in the gold sample, so that all $p_{\text{astro}} = 1$ under any model allowed by observations; or (ii) we consider the same O3a binary black hole mergers as in Ref. [9]—i.e. with an inverse false-alarm rate IFAR > 1 yr in any pipeline and excluding GW190814—taking the reported p_{astro} at face value and assigning it to an arbitrary model λ_0 featuring a broad distribution in black hole parameters, described below. We will refer to these two samples as GWTC-1 + IAS + Gold O3a and GWTC-1 + IAS + GWTC-2, respectively. We will find that our conclusions are not strongly affected by the sample used. We implement the sample choices by setting appropriate thresholds on the IFAR, which are reported for both events and injections in GWTC-2. The O3a injections do not report whether they fall near a glitch (one of the criteria of the gold sample), but these should be present in only a few percent of the events given the reported rate of ~ 1 glitch/min [1].

Following [73], we adopt a fiducial population model λ_0 that is described by the following distribution function:

$$f(m_{1s}, q, \chi_{\text{eff}}, D_L \mid \lambda'_0) \propto m_{1s}^{-\alpha_0} D_L^2, \quad (5)$$

where D_L is the luminosity distance and $\alpha_0 = 2.35$. We adopted the λ' notation for the parameters that control the shape of the distribution, while the rate R controls its normalization, i.e. $\lambda = (R, \lambda')$. The ranges of the parameters in Eq. (5) are taken to be $m_{1\min} < m_{1s} < m_{1\max}$ and $q_{\min} < q < 1$, where $m_{1\min} = 3 M_\odot$, $m_{1\max} = 120 M_\odot$ and $q_{\min} = 1/20$.

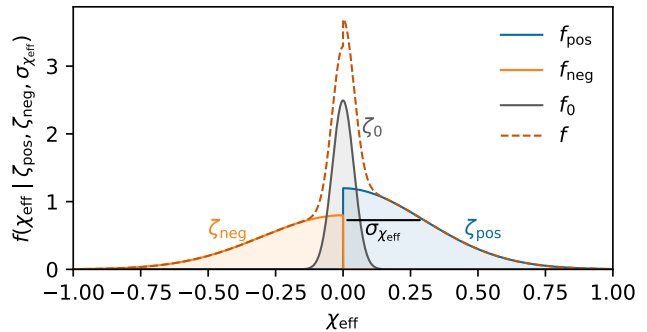


FIG. 2. Sketch of the functional form Eq. (6), which we use to parametrize the χ_{eff} distribution as the sum of three subpopulations. These subpopulations have positive, negative or zero effective spins, with each described by truncated Gaussians that peak at $\chi_{\text{eff}} = 0$. We use three independent shape parameters: the width of the positive and negative distributions, $\sigma_{\chi_{\text{eff}}}$, which are constrained to be equal, and the three branching ratios ζ_j which sum to unity. For technical reasons, we fix the width of the $\chi_{\text{eff}} \approx 0$ subpopulation to have a small but non-vanishing dispersion of $\sigma_0 = 0.04$.

B. Spin distribution

Motivated by Fig. 1 and the discussion in Sections I and III, as well as Refs. [9, 71], we will consider a phenomenological model for the effective spin distribution that allows us to explore the degree of symmetry of the distribution about $\chi_{\text{eff}} = 0$. This model will also allow us to quantify the support for positive and negative values of χ_{eff} in the population.

Firstly, we model the effective spin distribution as a mixture of three subpopulations with negative, zero, and positive χ_{eff} :

$$f_{\chi_{\text{eff}}}(\chi_{\text{eff}} \mid \zeta_{\text{pos}}, \zeta_{\text{neg}}, \sigma_{\chi_{\text{eff}}}) = \zeta_0 \mathcal{N}(\chi_{\text{eff}}; \sigma_0 = 0.04) + \zeta_{\text{neg}} \mathcal{N}_{<0}(\chi_{\text{eff}}; \sigma_{\chi_{\text{eff}}}) + \zeta_{\text{pos}} \mathcal{N}_{>0}(\chi_{\text{eff}}; \sigma_{\chi_{\text{eff}}}). \quad (6)$$

Here, the various parameters $\zeta_j \in [0, 1]$ are the branching ratios for each subpopulation, constrained to have unit sum; $\mathcal{N}(x; \sigma)$ is the normal distribution with zero mean, dispersion σ , truncated at $x = \pm 1$; $\mathcal{N}_{<0}(x; \sigma)$ is a similar normal distribution but truncated at $x = -1$ and $x = 0$, while $\mathcal{N}_{>0}(x; \sigma)$ is truncated at $x = 0$ and $x = 1$. The functional form Eq. (6) is sketched in Fig. 2 for a particular choice of parameters. Note that we have enforced the dispersion parameters of the positive and negative subpopulations to be equal, such that setting $\zeta_{\text{pos}} = \zeta_{\text{neg}}$ yields a χ_{eff} distribution **symmetric about zero**. For the $\chi_{\text{eff}} \approx 0$ subpopulation, we adopt a small (relative to typical measurement uncertainties) but nonvanishing dispersion $\sigma_0 = 0.04$ in order to ensure that the reweighting procedure used in our algorithm is well behaved [73].

In this Section we will only vary the effective spin distribution, while the remaining spin components are assumed to follow the parameter estimation prior described

in Section II. For the other binary black hole parameters, we will assume the following fixed distribution :

$$f(\chi_{\text{eff}}, m_{1s}, q, D_L) = f_{\chi_{\text{eff}}}(\chi_{\text{eff}}) f_{m_{1s}}(m_{1s}) f_q(q) f_{D_L}(D_L). \quad (7)$$

Following Abbott *et al.* [9] we adopt a broken power-law distribution for the primary mass:

$$f_{m_{1s}}(m_{1s}) \propto \begin{cases} 0, & m_{1s} < 5 M_{\odot} \\ \left(\frac{m_{1s}}{m_{\text{break}}}\right)^{-\alpha_1}, & 5 M_{\odot} < m_{1s} < m_{\text{break}} \\ \left(\frac{m_{1s}}{m_{\text{break}}}\right)^{-\alpha_2}, & m_{\text{break}} < m_{1s}, \end{cases} \quad (8)$$

with $\alpha_1 = 1.6$, $\alpha_2 = 5.6$, $m_{\text{break}} = 40 M_{\odot}$. For simplicity, we adopt a mass-ratio distribution that is uniform in $1/20 < q < 1$ and take the distance distribution to be uniform in comoving volume-time.

We use the likelihood in Eq. (4) to obtain a posterior distribution for the population parameters, by adopting a Jeffreys prior for the overall merger rate $\pi(R | \lambda') \propto \sqrt{N_a(R, \lambda')}/R$; recall that λ' are the shape parameters $(\zeta_{\text{pos}}, \zeta_{\text{neg}}, \sigma_{\chi_{\text{eff}}})$. For these we adopt a uniform prior $\pi(\lambda') = \text{const.}$ This prior is invariant to the choice of which two out of the three branching ratios are used to

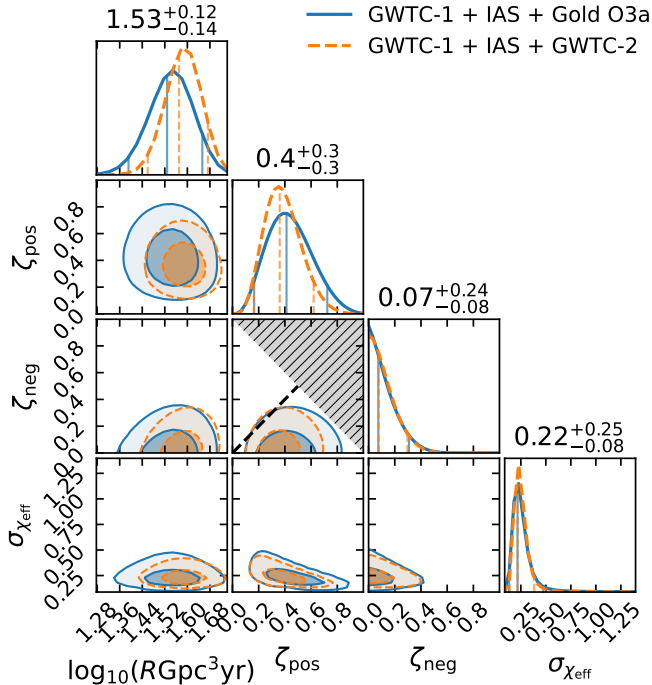


FIG. 3. Constraints on the model parameters of the population model Eq. (6). We see that a χ_{eff} distribution **symmetric about 0** (black dashed line), with $\zeta_{\text{pos}} = \zeta_{\text{neg}}$, is disfavored by the data. In addition, the population is consistent with having no negative-spin subpopulation. The two-dimensional contours enclose the 50% and 90% credible regions. Parameter values (median and 90% confidence level) are reported for the GWTC-1 + IAS + Gold O3a sample.

	$\Delta \max \ln L$	$\Delta \ln Z$
Symmetric χ_{eff}	0	0
Positive χ_{eff}	$2.1^{+0.5}_{-0.4}$	$1.6^{+0.5}_{-0.3}$
Positive/Negative mixture χ_{eff}	$2.1^{+0.5}_{-0.4}$	$1.4^{+0.4}_{-0.2}$
Gaussian χ_{eff}	$0.2^{+0.7}_{-0.6}$	$-0.2^{+0.6}_{-0.8}$

TABLE I. Scores for models of the χ_{eff} distribution. Difference in the maximum log likelihood and log evidence relative to the χ_{eff} model **symmetric about zero**, $\zeta_{\text{pos}} = \zeta_{\text{neg}}$. Error bars indicate the 90% confidence level and account for stochastic errors due to the finite number of injections and parameter samples used, and are estimated with 100 bootstrap realizations of the analysis similarly to [73].

parametrize the distribution.

We show our constraints on the parameters of this model in Fig. 3, for the two samples used. We find two remarkable results: first, 95% of the posterior lies at $\zeta_{\text{pos}} > \zeta_{\text{neg}}$ and a χ_{eff} **distribution symmetric about zero** ($\zeta_{\text{pos}} = \zeta_{\text{neg}}$, dashed line) is disfavored; second, the population is consistent with $\zeta_{\text{neg}} = 0$, i.e. no spins anti-aligned with the binary orbit. These conclusions do not depend on which of the two event samples are considered.

We quantify these statements using the Bayesian evidence and maximum likelihood as model scores: we report in Table I the scores achieved by the following models: a χ_{eff} distribution **symmetric about zero** given by Eq. (6) with $\zeta_{\text{pos}} = \zeta_{\text{neg}}$, a positive χ_{eff} distribution setting $\zeta_{\text{neg}} = 0$, and the full mixture. The symmetric χ_{eff} model is representative of a scenario completely dominated by dynamical formation in clusters, while the positive χ_{eff} model represents a case dominated by isolated binaries—with the caveat that in this channel there exist mechanisms to achieve some spin-orbit misalignment, e.g. supernova kicks.

The first result that positive χ_{eff} predominates over negative is in general agreement with the analysis of Abbott *et al.* [9]. Indeed, parametrizing the χ_{eff} distribution with a Gaussian, they find that a positive mean is preferred; likewise, they favor spin orientation distributions with at least some degree of anisotropy.

On the other hand, our second finding that there is yet no evidence for negative χ_{eff} in the population is in contrast with the results of Abbott *et al.* [9], who found that all Gaussian fits to the observed χ_{eff} distribution had a sizable support at negative χ_{eff} . We suggest that their result is contingent on the assumed parametrization of the population as a Gaussian distribution, while our parametrization has more freedom to accommodate features near $\chi_{\text{eff}} = 0$. In particular, the maximum likelihood solution has parameters $(\zeta_{\text{pos}}, \zeta_{\text{neg}}, \zeta_0, \sigma_{\chi_{\text{eff}}}) = (0.45, 0.00, 0.55, 0.23)$, featuring a sharp peak at $\chi_{\text{eff}} \approx 0$, a rapid decline at negative χ_{eff} and an extended tail at positive χ_{eff} which are hard to capture with a single Gaussian. To test this hypothesis we try a similar Gaussian model for the χ_{eff} distribution,

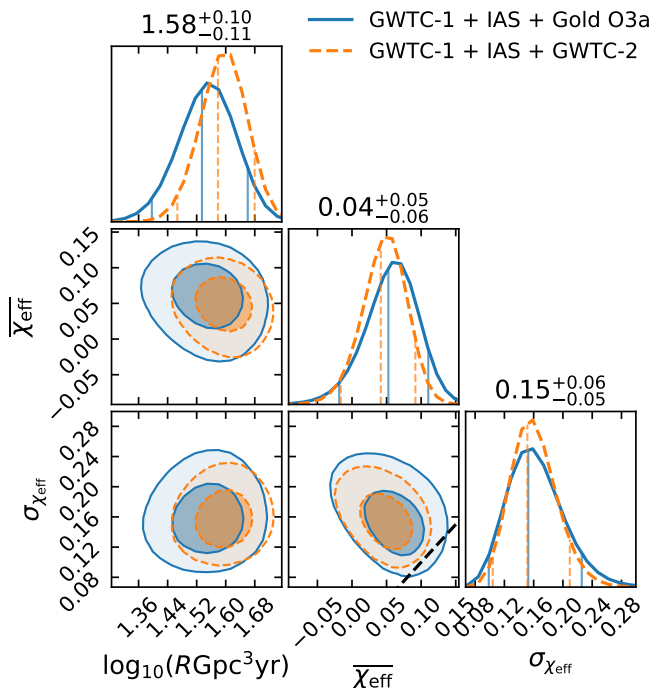


FIG. 4. Constraints assuming a Gaussian model for the χ_{eff} distribution. The black dashed line corresponds to $\sigma_{\chi_{\text{eff}}} = \overline{\chi_{\text{eff}}}$, models above the line have sizable support at negative χ_{eff} . Thus, contrary to Fig. 3, under this model one would conclude that negative χ_{eff} are present in the population.

shown in Fig. 4. With this model, we indeed find good quantitative agreement with Abbott *et al.* [9, figure 11] and would recover their same conclusions: we find that models without support at negative χ_{eff} ($\sigma_{\chi_{\text{eff}}} \ll \overline{\chi_{\text{eff}}}$) are excluded. In Table I we see that the Gaussian model performs worse than other models we tried, in particular the model restricted to positive χ_{eff} . Abbott *et al.* [9] did consider the possibility that their finding of negative spins could be driven by the Gaussian parametrization. Indeed, in Abbott *et al.* [9, figure 27] they show that adding a free parameter $\chi_{\text{eff}}^{\text{min}}$ below which the Gaussian is truncated, they exclude $\chi_{\text{eff}}^{\text{min}} \geq 0$ at 99% credibility and find that small negative values $-0.2 \lesssim \chi_{\text{eff}}^{\text{min}} \lesssim 0$ are preferred. We interpret that the large number of events at $\chi_{\text{eff}} \approx 0$ drives the exclusion of positive $\chi_{\text{eff}}^{\text{min}}$, furthermore, the fact that small negative values of $\chi_{\text{eff}}^{\text{min}}$ are preferred over large negative values indicates that the Gaussian model $\chi_{\text{eff}}^{\text{min}} = -1$, which motivated the claim of existence of negative χ_{eff} systems, does not fit well the observed population. We conclude that, while it is certainly possible that there are negative χ_{eff} systems in the population, there is not enough evidence for them yet.

Within isolated formation channels, the fraction of negative χ_{eff} systems ζ_{neg} is an indicator of typical natal (supernova) kick velocities, larger kicks generally giving larger ζ_{neg} . Gerosa *et al.* [65, figure 6] find that measurements of ζ_{neg} to a precision better than 0.1 would start putting meaningful constraints on kick velocities.

Our current bound $\zeta_{\text{neg}} \lesssim 0.3$ is compatible with even extreme kicks, but with a factor of few more detections this would be a promising source of information.

We point out that the GWTC-1 + IAS + GWTC-2 sample differs from that of the analysis in Abbott *et al.* [9] in that it includes events in the IAS catalog. However, having included these events only weakens our conclusions due to the presence of GW170121, the confident detection with the most support for negative χ_{eff} in the sample.

C. Mass distribution

We now turn to the distribution of merging binary black hole masses. Using data from the first two observing runs, several past studies have identified that the primary mass distribution was well described by a power-law truncated at $m_{\text{max}} \approx 40 M_{\odot}$ [72, 73, 81, 82, 106]. The third observing run revealed that the mass distribution has a tail that extends to higher masses, and that models with more features, e.g. a broken power-law, were favored. One diagnostic that a single truncated power-law did not fit the O3a data was that its inferred parameter values experienced a large shift when including the new events, in particular, m_{max} was found to increase from $40.8^{+11.8}_{-4.4} M_{\odot}$ to $78.5^{+14.1}_{-9.4} M_{\odot}$ [9].

As this development evidenced, one has to bear in mind that with a finite number of events one cannot probe the tail of the distribution arbitrarily far out. Thus, constraints obtained on the population are to be interpreted as a characterization of the bulk of the distribution, up to a quantile that depends on the number of events: with N_{trig} triggers, a fraction $\sim \mathcal{O}(1/N_{\text{trig}})$ of the distribution cannot be probed; with the present sample this is at the few-percent level. At this point we introduce a feature in our analysis that makes this notion explicit: we add to the model a second subpopulation of astrophysical triggers that come from a broad parameter distribution λ'_0 accounting for a small fraction ϵ of the total rate:

$$\frac{dN_a}{d\theta}(\theta | \lambda, \lambda'_0, \epsilon) = R[(1 - \epsilon)f(\theta | \lambda') + \epsilon f(\theta | \lambda'_0)]; \quad (9)$$

for $\epsilon = 0$ we recover the previous analysis. Recall that we call the distribution shape parameters λ' , so that $\lambda = (R, \lambda')$. For simplicity, we will fix the parameter $\epsilon = 0.05$. This change makes little difference for events that are well described by the population model λ , but since the broad subpopulation can accommodate any of its outliers, the model λ is no longer forced to explain all the observations. A practical advantage of this is that we get a sensitive diagnostic that some specific events may be poorly accommodated by the (ultimately arbitrary) parametrizations we chose, if they get classified with high confidence as belonging to the other subpopulation λ'_0 —evidencing that a model with more freedom is needed to explain all events. We also construct a simple goodness-of-fit test for the λ model based on the Bayes

factor between a model with $\epsilon = 0$ or a small fixed value $\epsilon = 0.05$. If the $\epsilon = 0$ model is already a good description of all observed events, adding a broad subpopulation should not increase the evidence significantly.

The likelihood for this augmented model can be evaluated in post-processing from the same auxiliary quantities $w_i(\lambda', \lambda'_0, \mathcal{VT}(\lambda'), p_{\text{astro},i}(\lambda_0))$ we use in the evaluation of Eq. (4) (see [73]):

$$\begin{aligned}
 &P(\{d_i\} \mid \lambda, \lambda'_0, \epsilon) \\
 &\propto \exp\{-R[(1-\epsilon)\overline{\mathcal{VT}}(\lambda') + \epsilon\overline{\mathcal{VT}}(\lambda'_0)]\} \\
 &\times \prod_{i=1}^{N_{\text{trig}}} \left\{ \frac{R}{R_0} [w_i(\lambda', \lambda'_0)(1-\epsilon) + \epsilon] p_{\text{astro},i}(\lambda_0) \right. \\
 &\quad \left. + 1 - p_{\text{astro},i}(\lambda_0) \right\}.
 \end{aligned} \quad (10)$$

Likewise, we can also extract the classification of each event as coming from the main component λ or the broader component λ'_0 : the probability that the i th event came from the λ'_0 population is

$$\begin{aligned}
 &p_{\text{outlier},i}(\lambda, \lambda'_0, \epsilon) \\
 &= \frac{R\epsilon}{R[(1-\epsilon)w_i(\lambda', \lambda'_0) + \epsilon] + R_0(1/p_{\text{astro},i}(\lambda_0) - 1)}.
 \end{aligned} \quad (11)$$

We apply this procedure to three models of the mass distribution that are simplified versions of the TRUNCATED, BROKEN POWER LAW and POWER LAW + PEAK models studied in Abbott *et al.* [9]. Our broken power law model is given by Eq. (8), with $\alpha_1, \alpha_2, m_{\text{break}}$ promoted to free parameters. Our truncated model corresponds to $\alpha_2 \rightarrow \infty$. Our power law + peak model corresponds to $\alpha_2 = \alpha_1$, plus the addition of a Gaussian component with mean m_{peak} and dispersion $\sigma = 5M_\odot$ that accounts for a fraction ζ_{peak} of the total rate. In all three cases we assume a uniform distribution for χ_{eff} , and identical distributions as in Section IV B for the remaining parameters. With these choices, the models λ' and λ'_0 only differ in the primary source-frame mass distribution, which will ease the interpretation of our results.

Fig. 5 shows the constraints we obtain using the GWTC-1 + IAS + Gold O3a sample; these plots are largely unchanged if we use the GWTC-1 + IAS + GWTC-2 sample (not shown). For the case $\epsilon = 0$ we find large quantitative agreement with Abbott *et al.* [9] in the constraints for the corresponding model parameters; in particular, that the data favor a break with $\alpha_1 > \alpha_2$ (region above the dashed line in Fig. 5b). When we set $\epsilon = 0.05$, allowing these models to not fit all events, we find that the model parameter constraints are affected: for the truncated power-law model the effect is catastrophic, in the sense that the posteriors for $\epsilon = 0$ and 0.05 are inconsistent with each other; while for the broken power-law model there remains a region of overlap and for the power law + peak model the inferred parameters remain largely unaffected. This is in

	ϵ	$\Delta \max \ln L$	$\Delta \ln Z$
Truncated power law	0	$-7.4^{+0.3}_{-0.3}$	$-6.21^{+0.29}_{-0.19}$
	0.05	0	0
Broken power law	0	$-2.51^{+0.14}_{-0.12}$	$-3.12^{+0.13}_{-0.14}$
	0.05	$-0.23^{+0.07}_{-0.11}$	$0.02^{+0.03}_{-0.04}$
Power law + peak	0	$-1.03^{+0.18}_{-0.17}$	$-3.18^{+0.22}_{-0.13}$
	0.05	$0.04^{+0.21}_{-0.17}$	$-1.56^{+0.15}_{-0.17}$

TABLE II. Scores for models of the primary mass distribution. Maximum log likelihood and log evidence for truncated power law and broken power law models, plus a fraction $\epsilon = 0$ or 0.05 of the population coming from a broad distribution λ'_0 per Eq. (9). The scores are referred to the preferred truncated power law model with $\epsilon = 0.05$. In all three cases $\epsilon = 0$ is disfavored, implying that the models struggle to accommodate all observations.

line with the discussion of [9, figure 2] and suggests that the truncated power-law model with $\epsilon = 0$ fails to describe the astrophysical distribution. It is interesting to note that, with $\epsilon = 0.05$, the truncated and broken power law parametrizations are consistent with the same physical solution $\alpha \approx \alpha_1$, $m_{\text{max}} \approx m_{\text{break}}$, $\alpha_2 \gg 1$, which exhibits a sharp step at m_{break} and a tail that extends to high masses. The power law + peak parametrization cannot produce a step. We show these inferred distributions in Fig. 6 to further illustrate the point that both parametrizations give consistent answers, especially for the bulk of the distribution. Note that the differential merger rate is best constrained around $m_{1s} \sim 20M_\odot$, where most observations lie [73].

In Table II we report the maximum likelihood and evidence for each of the models studied. We find that, although the broken power law model outperforms the truncated model when $\epsilon = 0$, both perform poorly relative to their $\epsilon = 0.05$ counterparts. This suggests that neither is a good description of the mass distribution. The power law + peak model achieves similar scores as the broken power law model with $\epsilon = 0$, but it gets only a slight improvement from $\epsilon = 0.05$, thus getting similarly disfavored. Among all the variations, thus, the preferred models in terms of evidence are either the truncated or broken power law with $\epsilon = 0.05$, i.e. with a small fraction of events in a broad tail that extends to high masses. The fact that these two models achieve similar likelihood and evidence, together with the above observation that they are consistent with the same physical solution, suggests that both are comparably good descriptions of the bulk of the distribution and their different scores for $\epsilon = 0$ are driven by the few outlier events. This is confirmed in Fig. 6. Comparing the $\epsilon = 0$ entries in Table I to Abbott *et al.* [9, table 2], we find agreement in that the truncated power law model is rejected, however, Abbott *et al.* [9] find a preference for the power law + peak model over the broken power law, which we instead find comparable. Some differences are expected because, for simplicity, in our implementation of these models we fixed or omitted some parameters, so the models and associated phase

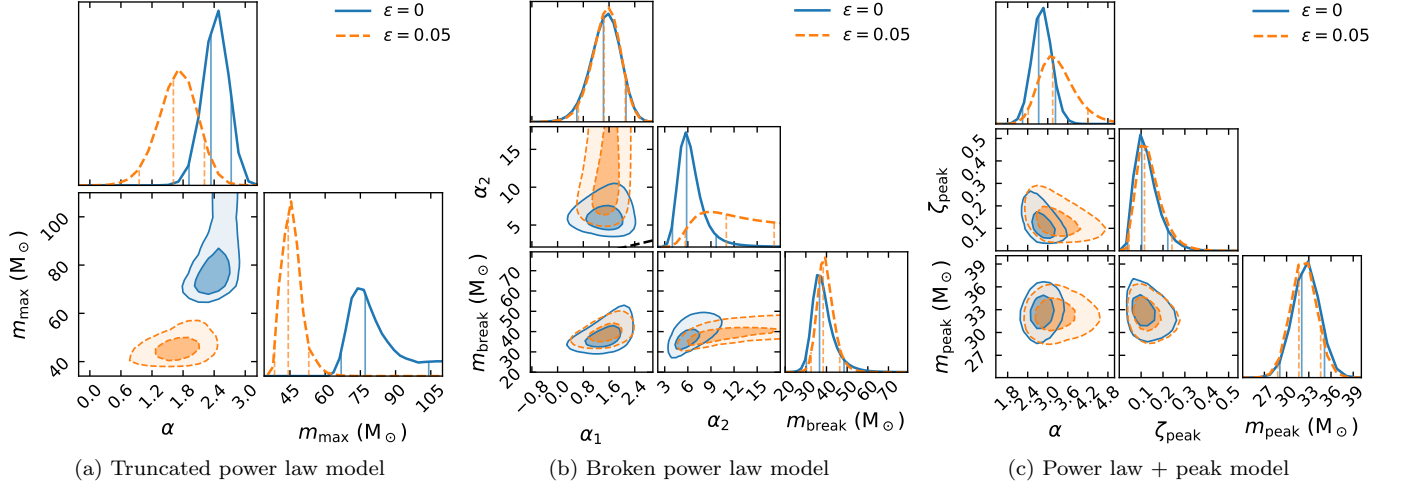


FIG. 5. Adding a broad subpopulation λ'_0 with a fraction $\epsilon = 0.05$ of the astrophysical rate affects the inferred parameters of the mass distribution. This is a major effect for the truncated power law model (a), moderate for the broken power law model (b) and minor for the power law + peak model (c). These constraints are derived using the GWTC-1 + IAS + Gold O3a sample of events; we find similar results with GWTC-1 + IAS + GWTC-2. **Repeating the analysis with $\epsilon = 0.1$ yields similar results as with $\epsilon = 0.05$, corroborating that it is the freedom to accommodate outliers that drives these changes rather than the particular choice of ϵ .**

spaces are not equivalent.

We can get some insight by inspecting the probabilities p_{outlier} of coming from the broad subpopulation λ'_0 assigned to each event, which we report in Table III. Events with a high value of p_{outlier} are better explained by the broad subpopulation and drive a preference for $\epsilon \neq 0$. However, note that even if the true astrophysical population was well described by the parametrization λ , in a catalog of many events some are bound to be in the

	Truncated power law	Broken power law	Power law + peak
GW190521	1.00	0.94	0.68
GW190602_175927	0.95	0.72	0.66
GW190706_222641	0.88	0.72	0.75
GW190519_153544	0.76	0.54	0.59
GW190929_012149	0.57	0.46	0.51
GW190620_030421	0.43	0.34	0.47
GW190701_203306	0.33	0.19	0.29
GW190413_134308	0.27	0.25	0.31

TABLE III. Probability that each event is a model outlier, as defined in Eq. (11) and marginalized over model parameters λ , with $\epsilon = 0.05$, for the mass models studied. Only events in the gold sample with the highest values of p_{outlier} are shown, for brevity. Note that this naturally selects the events with highest primary mass.

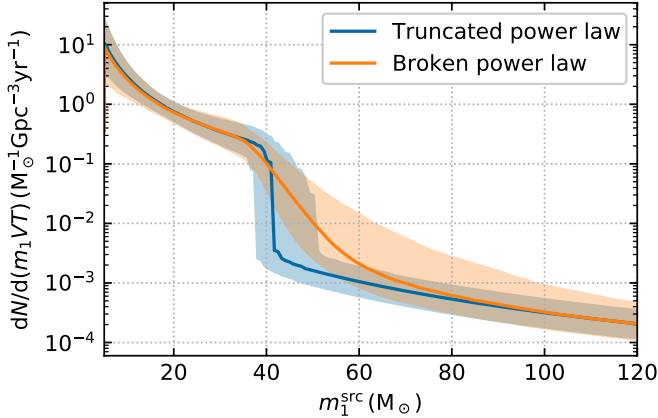


FIG. 6. Mass distribution predicted by the models favored by the data: truncated power law or broken power law, in both cases with an additional broad subpopulation with a fraction $\epsilon = 0.05$ of the total rate responsible for the shallow tail to high masses. These distributions feature a steepening around $40 M_\odot$, consistent with a step, and a flattening at higher mass. Note that we do not fit for the power-law index nor the normalization of the high-mass tail.

tail of the distribution and might individually be better described by a broad distribution. The expected distribution of p_{outlier} under a model λ is hard to compute, which is why we do not use the values of p_{outlier} as a quantitative model test. This said, it is apparent that GW190521 is an extreme outlier of the truncated power law model, and there are two other events that are in some tension. For the broken power law model, GW190521 is in some tension but the other values of p_{outlier} are milder. For the power law + peak model, no single event is a strong outlier.

Another interesting effect is that GW170817A, a candidate event with $m_{1s} = 56^{+16}_{-10} M_\odot$ and a rather low false-alarm rate of $1/(36 \text{ O2})$ observing runs [7], had an estimated probability of astrophysical origin marginalized

over population parameters of $\overline{p_{\text{astro}}} = 0.07$, under the truncated power law model favored after O1 and O2 [73]. This low value was driven by the lack of observations of other events with similar properties, mainly mass. Under the newly favored models, it has a moderately different $\overline{p_{\text{astro}}} = 0.22$ for the truncated power law and 0.26 for the broken power law, both with $\epsilon = 0.05$. This showcases that p_{astro} values for marginal events in the tails of the distribution are bound to get updated as our knowledge of the population improves.

To summarize, Table II, Figs. 5 and 6 suggest that the mass distribution exhibits a steepening around $40 M_{\odot}$ and an extended, shallow high-mass tail. From Table III we conclude that the need for this tail is dominated by GW190521, so at this point we do not attempt to model its shape based on a single event. Future data releases will allow to probe these features in the mass distribution.

V. CONCLUSIONS

We have investigated the properties of the effective spin and primary mass distributions of binary black holes identified in the GWTC-1 [2], GWTC-2 [1], and IAS O1–O2 [4–7] event catalogs. Our study involved re-analyzing all binary black hole signals with the recently developed IMRPhenomXPHM waveform model [85], which includes orbital precession and higher-order modes.

We designed a parametric model of the χ_{eff} distribution which has three components — with negative, approximately zero and positive χ_{eff} — to test some general predictions of the dynamic and isolated formation channels for merging binary black holes. Namely, dynamical formation channels predict a χ_{eff} distribution that is symmetric about 0, while negative χ_{eff} (i.e., large spin–orbit misalignment) should be very rare for isolated field binaries. Interestingly, we found that a symmetric distribution is disfavored: the data suggests that the number of positive χ_{eff} events is larger than that with negative χ_{eff} at 95% credibility. Although the evidence at this point is not conclusive, this simple test is already becoming powerful enough to hint that not all binary black holes are dynamically assembled, in agreement with other analyses of these data [9, 11, 12]. The number of detections is expected to roughly double with the forthcoming release of the O3b catalog, which should settle this question if the same trend continues.

Moreover, we find no evidence for negative χ_{eff} in the population. This result is in tension with Ref. [9]; we attribute the discrepancy to the different parametrizations of the spin distribution chosen. We were able to reproduce the results of Ref. [9] with a Gaussian model for χ_{eff} , but found that this model fares worse at describing the features in the spin distribution, in particular, a large concentration of events near $\chi_{\text{eff}} = 0$. Our conclusion is in agreement with a model-free inspection of the empirical χ_{eff} distribution, which suggests that all events with significant support at $\chi_{\text{eff}} < 0$ are consistent with coming

from a population with $\chi_{\text{eff}} = 0$. Therefore, we conclude that the observed effective spin distribution does not rule out that all observations are explained by isolated binary formation.

Regarding the distribution of primary masses, we confirmed the result of [9] that a truncated power law fails to describe the observations. Moreover, we found evidence that a broken power law model or a power law plus a Gaussian peak, which assume a continuous distribution, compare poorly to a model in which a small fraction of the events comes from a broad subpopulation, with an extended tail at high masses. This suggests that the tail of the mass distribution has interesting features that will be probed with the coming data releases.

ACKNOWLEDGMENTS

We are grateful to Ilya Mandel, Will Farr, Davide Gerosa and the attendees of the gravitational wave meeting at the Center for Computational Astrophysics for helpful comments and discussion.

HSC gratefully acknowledges support from the Rubicon Fellowship awarded by the Netherlands Organisation for Scientific Research (NWO). SO acknowledges support from the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-2039656. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. LD acknowledges support from the Michael M. Garland startup research grant at the University of California, Berkeley. TV acknowledges support by the National Science Foundation under Grant No. 2012086. BZ is supported by a research grant from the Ruth and Herman Albert Scholarship Program for New Scientists. MZ is supported by NSF grants PHY-1820775 the Canadian Institute for Advanced Research (CIFAR) Program on Gravity and the Extreme Universe and the Simons Foundation Modern Inflationary Cosmology initiative.

This research has made use of data, software and/or web tools obtained from the Gravitational Wave Open Science Center (<https://www.gw-openscience.org/>), a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration. LIGO Laboratory and Advanced LIGO are funded by the United States National Science Foundation (NSF) 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.

Appendix A: Sample Selection

In this appendix, we inventorize the binary black hole mergers used in this work, which are listed in Table IV. We define the gold sample (third column of Table IV) as the set of events that (i) were detected by at least two search pipelines with a FAR $< 0.1 \text{ yr}^{-1}$ (fourth column); and (ii) on strain data that are free of non-Gaussian transient noise (fifth column). We consider the following pipelines: cWB [1], GstLAL [1], PyCBC [1, 3, 8], PyCBC BBH [1, 8], and IAS [6]. Events with non-Gaussian artifacts are reported in [9, table V]. We do not include GW190814 in the GWTC-1 + IAS + Gold O3a sample because it was detected near non-Gaussian transient noise [1]. Nor do we include GW190814 in the GWTC-1 + IAS + GWTC-2 sample (Section IV B) as it was not included in the main GWTC-2 population analysis due to being an outlier in the mass ratio distribution [9]. For events in the O1 and O2 observing runs, $p_{\text{astro}}(\lambda_0)$ is computed in [73]. For events in O3a, it is taken at face value from [1] as the maximum p_{astro} over pipelines, and may not accurately correspond to the model λ_0 .

While the present work was being completed, Nitz *et al.* [84] reported their analysis of the O3a data, providing independent confirmation of all the sources reported in GWTC-2 except for GW190426.152155 and GW190909.114149, and further finding four previously unreported events. We defer the inclusion of these results to future work. Including this catalog, the two-pipeline condition would be fulfilled by most of the O3a events in Table IV, thereby enlarging the gold sample. Still, note that the sample restriction did not change the qualitative conclusions of our analysis.

Run	Name	Gold ≥ 2 pip. Clean $p_{\text{astro}}(\lambda_0)$			
O1	GW150914	✓	✓	✓	1.00
	GW151012	✓	✓	✓	1.00
	GW151226	✓	✓	✓	1.00
	GW151216			✓	0.50
O2	GW170823	✓	✓	✓	1.00
	GW170809	✓	✓	✓	1.00
	GW170729	✓	✓	✓	1.00
	GW170814	✓	✓	✓	1.00
	GW170104	✓	✓	✓	1.00
	GW170727	✓	✓	✓	0.99
	GW170121	✓	✓	✓	0.97
	GW170304			✓	1.00
	GW170818	✓	✓	✓	0.92
	170412B			✓	0.02
	GW170403			✓	0.61
	GW170425			✓	0.60
	GW170202			✓	0.61
	GW170817A			✓	0.74
	GW170608	✓	✓	✓	1.00
O3a	GW190408.181802	✓	✓	✓	1.00
	GW190412	✓	✓	✓	1.00
	GW190413.052954			✓	0.98
	GW190413.134308				0.98
	GW190421.213856	✓	✓	✓	1.00
	GW190424.180648				0.91
	GW190503.185404		✓		1.00
	GW190512.180714	✓	✓	✓	1.00
	GW190513.205428		✓		1.00
	GW190514.065416				0.96
	GW190517.055101	✓	✓	✓	1.00
	GW190519.153544	✓	✓	✓	1.00
	GW190521	✓	✓	✓	1.00
	GW190521.074359	✓	✓	✓	1.00
	GW190527.092055			✓	0.99
	GW190602.175927	✓	✓	✓	1.00
	GW190620.030421			✓	1.00
	GW190630.185205			✓	1.00
	GW190701.203306				1.00
	GW190706.222641	✓	✓	✓	1.00
	GW190707.093326	✓	✓	✓	1.00
	GW190708.232457			✓	1.00
	GW190719.215514			✓	0.82
	GW190720.000836	✓	✓	✓	1.00
	GW190727.060333		✓		1.00
	GW190728.064510	✓	✓	✓	1.00
	GW190731.140936			✓	0.97
	GW190803.022701	✓	✓	✓	0.99
	GW190828.063405	✓	✓	✓	1.00
	GW190828.065509	✓	✓	✓	1.00
	GW190909.114149			✓	0.89
	GW190910.112807			✓	1.00
	GW190915.235702	✓	✓	✓	1.00
	GW190924.021846		✓		1.00
	GW190929.012149			✓	1.00
	GW190930.133541	✓	✓	✓	1.00

TABLE IV. Binary black hole events used in this work. Checkmarks from the third to fifth columns indicate events that are in the gold sample, were identified by at least two pipelines with IFAR $> 10 \text{ yr}$, and were observed in the absence of glitches, respectively. The p_{astro} values shown here are evaluated with the reference model λ_0 described in Eq. (5).

-
- [1] R. Abbott *et al.*, “GWTC-2: Compact binary coalescences observed by LIGO and Virgo during the first half of the third observing run,” (2020), arXiv:2010.14527.
- [2] B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), *Phys. Rev. X* **9**, 031040 (2019).
- [3] A. H. Nitz, C. Capano, A. B. Nielsen, S. Reyes, R. White, D. A. Brown, and B. Krishnan, *Astrophys. J.* **872**, 195 (2019), arXiv:1811.01921 [gr-qc].
- [4] T. Venumadhav, B. Zackay, J. Roulet, L. Dai, and M. Zaldarriaga, *Phys. Rev. D* **100**, 023011 (2019), arXiv:1902.10341 [astro-ph.IM].
- [5] B. Zackay, T. Venumadhav, L. Dai, J. Roulet, and M. Zaldarriaga, *Phys. Rev. D* **100**, 023007 (2019), arXiv:1902.10331 [astro-ph.HE].
- [6] T. Venumadhav, B. Zackay, J. Roulet, L. Dai, and M. Zaldarriaga, *Phys. Rev. D* **101**, 083030 (2020), arXiv:1904.07214 [astro-ph.HE].
- [7] B. Zackay, L. Dai, T. Venumadhav, J. Roulet, and M. Zaldarriaga, “Detecting Gravitational Waves With Disparate Detector Responses: Two New Binary Black Hole Mergers,” (2019), arXiv:1910.09528 [astro-ph.HE].
- [8] A. H. Nitz, T. Dent, G. S. Davies, S. Kumar, C. D. Capano, I. Harry, S. Mozzon, L. Nuttall, A. Lundgren, and M. Tápai, *Astrophys. J.* **891**, 123 (2020), arXiv:1910.05331 [astro-ph.HE].
- [9] R. Abbott *et al.*, “Population properties of compact objects from the second LIGO–Virgo Gravitational-Wave Transient Catalog,” (2020), arXiv:2010.14533.
- [10] K. W. K. Wong, K. Breivik, K. Kremer, and T. Callister, *Phys. Rev. D* **103**, 083021 (2021).
- [11] 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, *The Astrophysical Journal* **910**, 152 (2021).
- [12] Y. Bouffanais, M. Mapelli, F. Santoliquido, N. Giacobbo, U. N. D. Carlo, S. Rastello, M. C. Artale, and G. Iorio, “New insights on binary black hole formation channels after GWTC-2: young star clusters versus isolated binaries,” (2021), arXiv:2102.12495.
- [13] S. F. P. Zwart and S. L. McMillan, *The Astrophysical Journal Letters* **528**, L17 (1999).
- [14] R. M. O’Leary, F. A. Rasio, J. M. Fregeau, N. Ivanova, and R. O’Shaughnessy, *The Astrophysical Journal* **637**, 937 (2006).
- [15] A. Sadowski, K. Belczynski, T. Bulik, N. Ivanova, F. A. Rasio, and R. O’Shaughnessy, *The Astrophysical Journal* **676**, 1162 (2008).
- [16] J. Downing, M. Benacquista, M. Giersz, and R. Spurzem, *Monthly Notices of the Royal Astronomical Society* **407**, 1946 (2010).
- [17] J. Downing, M. Benacquista, M. Giersz, and R. Spurzem, *Monthly Notices of the Royal Astronomical Society* **416**, 133 (2011).
- [18] J. Samsing, M. MacLeod, and E. Ramirez-Ruiz, *Astrophys. J.* **784**, 71 (2014).
- [19] C. L. Rodriguez, M. Morscher, B. Pattabiraman, S. Chatterjee, C.-J. Haster, and F. A. Rasio, *Phys. Rev. Lett.* **115**, 051101 (2015).
- [20] C. L. Rodriguez, S. Chatterjee, and F. A. Rasio, *Physical Review D* **93**, 084029 (2016).
- [21] A. Askar, M. Szkudlarek, D. Gondek-Rosińska, M. Giersz, and T. Bulik, *Monthly Notices of the Royal Astronomical Society: Letters* **464**, L36 (2016).
- [22] F. Antonini and F. A. Rasio, *The Astrophysical Journal* **831**, 187 (2016).
- [23] C. Petrovich and F. Antonini, *The Astrophysical Journal* **846**, 146 (2017).
- [24] B. M. Ziosi, M. Mapelli, M. Branchesi, and G. Tormen, *Monthly Notices of the Royal Astronomical Society* **441**, 3703 (2014).
- [25] M. Mapelli, *Monthly Notices of the Royal Astronomical Society* **459**, 3432 (2016).
- [26] S. Banerjee, *Monthly Notices of the Royal Astronomical Society* **467**, 524 (2017).
- [27] S. Chatterjee, C. L. Rodriguez, V. Kalogera, and F. A. Rasio, *The Astrophysical Journal Letters* **836**, L26 (2017).
- [28] G. Nelemans, L. Yungelson, and S. P. Zwart, *Astronomy & Astrophysics* **375**, 890 (2001).
- [29] K. Belczynski, V. Kalogera, and T. Bulik, *The Astrophysical Journal* **572**, 407 (2002).
- [30] R. Voss and T. M. Tauris, *Monthly Notices of the Royal Astronomical Society* **342**, 1169 (2003).
- [31] K. Belczynski, R. E. Taam, V. Kalogera, F. A. Rasio, and T. Bulik, *The Astrophysical Journal* **662**, 504 (2007).
- [32] K. Belczynski, V. Kalogera, F. A. Rasio, R. E. Taam, A. Zezas, T. Bulik, T. J. Maccarone, and N. Ivanova, *The Astrophysical Journal Supplement Series* **174**, 223 (2008).
- [33] M. Dominik, K. Belczynski, C. Fryer, D. E. Holz, E. Berti, T. Bulik, I. Mandel, and R. O’Shaughnessy, *The Astrophysical Journal* **779**, 72 (2013).
- [34] K. Belczynski, A. Buonanno, M. Cantiello, C. L. Fryer, D. E. Holz, I. Mandel, M. C. Miller, and M. Walczak, *The Astrophysical Journal* **789**, 120 (2014).
- [35] N. Mennekens and D. Vanbeveren, *Astronomy & Astrophysics* **564**, A134 (2014).
- [36] M. Spera, M. Mapelli, and A. Bressan, *Monthly Notices of the Royal Astronomical Society* **451**, 4086 (2015).
- [37] J. Eldridge and E. Stanway, *Monthly Notices of the Royal Astronomical Society* **462**, 3302 (2016).
- [38] S. Stevenson, A. Vigna-Gómez, I. Mandel, J. W. Barrett, C. J. Neijssel, D. Perkins, and S. E. de Mink, *Nature communications* **8**, 14906 (2017).
- [39] M. Mapelli, N. Giacobbo, E. Ripamonti, and M. Spera, *Monthly Notices of the Royal Astronomical Society* **472**, 2422 (2017).
- [40] N. Giacobbo, M. Mapelli, and M. Spera, *Monthly Notices of the Royal Astronomical Society* **474**, 2959 (2017).
- [41] M. Mapelli and N. Giacobbo, *Monthly Notices of the Royal Astronomical Society* (2018).
- [42] M. U. Kruckow, T. M. Tauris, N. Langer, M. Kramer, and R. G. Izzard, *Monthly Notices of the Royal Astronomical Society* **481**, 1908 (2018).
- [43] N. Giacobbo and M. Mapelli, *Monthly Notices of the Royal Astronomical Society* **480**, 2011 (2018).
- [44] P. Marchant, N. Langer, P. Podsiadlowski, T. M. Tauris, and T. J. Moriya, *Astronomy & Astrophysics* **588**, A50 (2016).
- [45] S. De Mink and I. Mandel, *Monthly Notices of the Royal*

- Astronomical Society **460**, 3545 (2016).
- [46] I. Mandel and S. E. de Mink, Monthly Notices of the Royal Astronomical Society **458**, 2634 (2016).
 - [47] F. Antonini and H. B. Perets, The Astrophysical Journal **757**, 27 (2012).
 - [48] B. McKernan, K. Ford, W. Lyra, and H. Perets, Monthly Notices of the Royal Astronomical Society **425**, 460 (2012).
 - [49] N. C. Stone, B. D. Metzger, and Z. Haiman, Monthly Notices of the Royal Astronomical Society **464**, 946 (2016).
 - [50] I. Bartos, B. Kocsis, Z. Haiman, and S. Márka, The Astrophysical Journal **835**, 165 (2017).
 - [51] F. Antonini, N. Murray, and S. Mikkola, The Astrophysical Journal **781**, 45 (2014).
 - [52] T. O. Kimpson, M. Spera, M. Mapelli, and B. M. Ziosi, Monthly Notices of the Royal Astronomical Society **463**, 2443 (2016).
 - [53] F. Antonini, S. Toonen, and A. S. Hamers, The Astrophysical Journal **841**, 77 (2017).
 - [54] B. Liu and D. Lai, The Astrophysical Journal **863**, 68 (2018).
 - [55] A. S. Hamers, H. B. Perets, F. Antonini, and S. F. P. Zwart, Monthly Notices of the Royal Astronomical Society **449**, 4221 (2015).
 - [56] S. Vitale, R. Lynch, J. Veitch, V. Raymond, and R. Sturani, Phys. Rev. Lett. **112**, 251101 (2014), arXiv:1403.0129 [gr-qc].
 - [57] M. Pürrer, M. Hannam, and F. Ohme, Phys. Rev. D **93**, 084042 (2016), arXiv:1512.04955 [gr-qc].
 - [58] S. Vitale, R. Lynch, V. Raymond, R. Sturani, J. Veitch, and P. Graff, Phys. Rev. D **95**, 064053 (2017), arXiv:1611.01122 [gr-qc].
 - [59] T. A. Apostolatos, C. Cutler, G. J. Sussman, and K. S. Thorne, Phys. Rev. D **49**, 6274 (1994).
 - [60] L. E. Kidder, Phys. Rev. D **52**, 821 (1995), arXiv:gr-qc/9506022.
 - [61] E. Racine, Phys. Rev. D **78**, 044021 (2008), arXiv:0803.1820 [gr-qc].
 - [62] P. Ajith *et al.*, Phys. Rev. Lett. **106**, 241101 (2011), arXiv:0909.2867 [gr-qc].
 - [63] L. Santamaria *et al.*, Phys. Rev. D **82**, 064016 (2010), arXiv:1005.3306 [gr-qc].
 - [64] C. L. Rodriguez, M. Zevin, C. Pankow, V. Kalogera, and F. A. Rasio, The Astrophysical Journal **832**, L2 (2016).
 - [65] D. Gerosa, E. Berti, R. O’Shaughnessy, K. Belczynski, M. Kesden, D. Wysocki, and W. Gladysz, Phys. Rev. D **98**, 084036 (2018).
 - [66] K. Belczynski and T. Bulik, Astronomy and Astrophysics **346**, 91 (1999).
 - [67] T. A. Callister, W. M. Farr, and M. Renzo, “State of the field: Binary black hole natal kicks and prospects for isolated field formation after GWTC-2,” (2020), arXiv:2011.09570.
 - [68] J. Fuller and L. Ma, The Astrophysical Journal **881**, L1 (2019).
 - [69] S. S. Bavera, T. Fragos, Y. Qin, E. Zapartas, C. J. Neijssel, I. Mandel, A. Batta, S. M. Gaebel, C. Kimball, and S. Stevenson, Astronomy & Astrophysics **635**, A97 (2020).
 - [70] W. M. Farr, S. Stevenson, M. C. Miller, I. Mandel, B. Farr, and A. Vecchio, Nature **548**, 426 (2017), arXiv:1706.01385.
 - [71] B. Farr, D. E. Holz, and W. M. Farr, The Astrophysical Journal **854**, L9 (2018).
 - [72] J. Roulet and M. Zaldarriaga, Monthly Notices of the Royal Astronomical Society **484**, 4216 (2019).
 - [73] J. Roulet, T. Venumadhav, B. Zackay, L. Dai, and M. Zaldarriaga, Phys. Rev. D **102**, 123022 (2020).
 - [74] W. A. Fowler and F. Hoyle, The Astrophysical Journal Supplement Series **9**, 201 (1964).
 - [75] Z. Barkat, G. Rakavy, and N. Sack, Phys. Rev. Lett. **18**, 379 (1967).
 - [76] J. R. Bond, W. D. Arnett, and B. J. Carr, The Astrophysical Journal **280**, 825 (1984).
 - [77] A. Heger, C. L. Fryer, S. E. Woosley, N. Langer, and D. H. Hartmann, The Astrophysical Journal **591**, 288 (2003).
 - [78] R. Farmer, M. Renzo, S. E. de Mink, P. Marchant, and S. Justham, The Astrophysical Journal **887**, 53 (2019).
 - [79] M. Safarzadeh and Z. Haiman, The Astrophysical Journal **903**, L21 (2020).
 - [80] D. Gerosa and M. Fishbach, “Hierarchical mergers of stellar-mass black holes and their gravitational-wave signatures,” (2021), arXiv:2105.03439.
 - [81] M. Fishbach and D. E. Holz, The Astrophysical Journal **851**, L25 (2017).
 - [82] D. Wysocki, J. Lange, and R. O’Shaughnessy, Phys. Rev. D **100**, 043012 (2019).
 - [83] R. Abbott *et al.* (LIGO Scientific, Virgo), Astrophys. J. Lett. **896**, L44 (2020), arXiv:2006.12611 [astro-ph.HE].
 - [84] A. H. Nitz, C. D. Capano, S. Kumar, Y.-F. Wang, S. Kastha, M. Schäfer, R. Dhurkunde, and M. Cabero, “3-OGC: Catalog of gravitational waves from compact-binary mergers,” (2021), arXiv:2105.09151.
 - [85] G. Pratten, C. García-Quirós, M. Colleoni, A. Ramos-Buades, H. Estellés, M. Mateu-Lucena, R. Jaume, M. Haney, D. Keitel, J. E. Thompson, and S. Husa, “Computationally efficient models for the dominant and sub-dominant harmonic modes of precessing binary black holes,” (2020), arXiv:2004.06503.
 - [86] B. Zackay, L. Dai, and T. Venumadhav, “Relative binning and fast likelihood evaluation for gravitational wave parameter estimation,” (2018), arXiv:1806.08792 [astro-ph.IM].
 - [87] J. Buchner, A. Georgakakis, K. Nandra, L. Hsu, C. Rangel, M. Brightman, A. Merloni, M. Salvato, J. Donley, and D. Kocevski, Astronomy & Astrophysics **564**, A125 (2014).
 - [88] N. J. Cornish and T. B. Littenberg, Classical and Quantum Gravity **32**, 135012 (2015).
 - [89] T. B. Littenberg, J. B. Kanner, N. J. Cornish, and M. Millhouse, Phys. Rev. D **94**, 044050 (2016).
 - [90] N. J. Cornish, T. B. Littenberg, B. Bécsy, K. Chatziioannou, J. A. Clark, S. Ghonge, and M. Millhouse, Phys. Rev. D **103**, 044006 (2021).
 - [91] B. P. Abbott *et al.* (LIGO Scientific, Virgo), Phys. Rev. Lett. **116**, 241103 (2016), arXiv:1606.04855 [gr-qc].
 - [92] H. S. Chia, S. Olsen, J. Roulet, L. Dai, T. Venumadhav, B. Zackay, and M. Zaldarriaga, “Boxing day surprise: Higher multipoles and orbital precession in GW151226,” (2021), arXiv:2105.06486.
 - [93] R. Abbott *et al.*, Physical Review Letters **125** (2020), 10.1103/physrevlett.125.101102.
 - [94] A. H. Nitz and C. D. Capano, The Astrophysical Journal **907**, L9 (2021).
 - [95] S. R. Olsen *et al.*, In preparation.

- [96] H. Estellés, S. Husa, M. Colleoni, M. Mateu-Lucena, M. de Lluc Planas, C. García-Quirós, D. Keitel, A. Ramos-Buades, A. K. Mehta, A. Buonanno, and S. Ossokine, “A detailed analysis of GW190521 with phenomenological waveform models,” (2021), arXiv:2105.06360.
- [97] M. Campanelli, C. O. Lousto, and Y. Zlochower, *Physical Review D* **74** (2006), 10.1103/physrevd.74.041501.
- [98] K. K. Y. Ng, S. Vitale, A. Zimmerman, K. Chatziioannou, D. Gerosa, and C.-J. Haster, *Phys. Rev. D* **98**, 083007 (2018).
- [99] D. Kushnir, M. Zaldarriaga, J. A. Kollmeier, and R. Waldman, *Monthly Notices of the Royal Astronomical Society* **462**, 844 (2016).
- [100] M. Zaldarriaga, D. Kushnir, and J. A. Kollmeier, *Monthly Notices of the Royal Astronomical Society* **473**, 4174 (2017).
- [101] Y. Qin, T. Fragos, G. Meynet, J. Andrews, M. Sørensen, and H. F. Song, *Astronomy & Astrophysics* **616**, A28 (2018).
- [102] R. Abbott *et al.*, *SoftwareX* **13**, 100658 (2021).
- [103] T. Venumadhav, B. Zackay, J. Roulet, L. Dai, and M. Zaldarriaga, *Phys. Rev. D* **100**, 023011 (2019).
- [104] T. Venumadhav, B. Zackay, J. Roulet, L. Dai, and M. Zaldarriaga, *Phys. Rev. D* **101**, 083030 (2020).
- [105] “Gravitational Wave Open Science Center (GWOSC),” www.gw-openscience.org/02/.
- [106] B. P. Abbott *et al.*, *The Astrophysical Journal* **882**, L24 (2019).