Spin it as you like: the (lack of a) measurement of the spin tilt distribution with LIGO-Virgo-KAGRA binary black holes

Salvatore Vitale,1,2* Sylvia Biscoveanu,1,2 Colm Talbot1,2

1LIGO Laboratory, Massachusetts Institute of Technology, 185 Albany St, Cambridge, MA 02139, USA
2Department of Physics and Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139, USA

November 9, 2022

ABSTRACT

Context. The growing set of gravitational-wave sources is being used to measure the properties of the underlying astrophysical populations of compact objects, black holes and neutron stars. Most of the detected systems are black hole binaries. While much has been learned about black holes by analyzing the latest LIGO-Virgo-KAGRA (LVK) catalog, GWTC-3, a measurement of the astrophysical distribution of the black hole spin orientations remains elusive. This is usually probed by measuring the cosine of the tilt angle (\(\cos \tau\)) between each black hole spin and the orbital angular momentum, \(\cos \tau = \pm 1\) being perfect alignment.

Aims. \cite{Abbott2021} has modeled the \(\cos \tau\) distribution as a mixture of an isotropic component and a Gaussian component with mean fixed at \(+1\) and width measured from the data. We want to verify if the data require the existence of such a peak at \(\cos \tau = +1\).

Methods. We use various alternative models for the astrophysical tilt distribution and measure their parameters using the LVK GWTC-3 catalog.

Results. We find that a) Augmenting the LVK model such that the mean \(\mu\) of the Gaussian is not fixed at \(+1\) returns results that strongly depend on priors. If we allow \(\mu > +1\) then the resulting astrophysical \(\cos \tau\) distribution peaks at \(+1\) and looks linear, rather than Gaussian. If we constrain \(-1 \leq \mu \leq +1\) the Gaussian component peaks at \(\mu = 0.48^{+0.06}_{-0.05}\) (median and 90% symmetric credible interval). Two other 2-component mixture models yield \(\cos \tau\) distributions that either have a broad peak centered at \(0.19^{+0.22}_{-0.12}\) or a plateau that spans the range \([-0.5, +1]\), without a clear peak at \(+1\). b) All of the models we considered agree on the fact that there is no excess of black hole tilts at around \(-1\). c) While yielding quite different posteriors, the models considered in this work have Bayesian evidences that are the same within error bars.

Conclusions. We conclude that the current dataset is not sufficiently informative to draw any model-independent conclusions on the astrophysical distribution of spin tilts, except that there is no excess of spins with negatively aligned tilts.

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

1. Introduction

More than 90 binary black holes (BBHs) have been detected in the data of the ground-based gravitational-wave detectors LIGO \cite{Aasi2015} and Virgo \cite{Acernese2015} by the LIGO-Virgo-Kagra (LVK) collaboration and other groups \cite{Abbott2021, Nitz2021, Olsen2022}. This dataset has been used to infer the properties of the underlying population—or populations—of BBHs. Among the parameters of interest, the masses and spins of the black holes play a prominent role since they can shed light on the binary formation channel\(^1\) (e.g., Vitale et al. 2017; Farr et al. 2017; Zevin et al. 2017; Farr et al. 2018; Wong et al. 2021; Zevin et al. 2021; Bouffanais et al. 2021).

There currently exist a few approaches toward measuring the population properties. The first is to use a functional form for a reasonable population distribution, parameterized by some phenomenological parameters. For example, the LVK parameterized the primary mass distribution of the black holes as a mixture of a power law distribution and a Gaussian component \cite{Abbott2021, Talbot2018, Fishbach2017}. This model includes several hyperparameters (since they pertain to the population as a whole, not to the individual events), which are measured from the data: the slope of the power law, the minimum and maximum black hole mass (including a smoothing parameter), the mean and standard deviation of the Gaussian component, and the branching ratio between the power law and Gaussian components. It is worth noting that not all parametric models make equally strong assumptions: an example of a more flexible parametrization is the Beta distribution that the LVK has used to describe the population distribution of black hole spin magnitudes \cite{Wysocki2019}. Those more elastic models might be more appropriate if one doesn’t have strong observational or theoretical expectations about what the astrophysical distribution of a parameter should look like, or simply if they want to be more conservative. Just as Bayesian priors can significantly affect the posterior when the likelihood doesn’t have a strong peak when analyzing individual compact binary coalescences, a hypermodel that is too strong could leave imprints on the inferred hyperparameters.

A second approach is to use non-parametric models, based on e.g. Gaussian processes \cite{Tiwari2021, Edelman2022}.
by using a mixture model composed of an isotropic distribution to characterize this distribution. We also find that using O3b only sources the cosine of the tilt distribution. We also find that those too yield similar evidences. Critically, all of these alternative—and equally supported by the data—models yield noticeably different posterior distributions for the tilt population relative to the default LVK model. In particular, different models give different support to the existence and position of a feature at positive cos τ. On the other hand, they all agree on the fact that there is no excess of systems with cos τ ≃ −1. We conclude that the current constraints on the distribution of tilt angles are significantly affected by the model used, and that more sources (or weaker models) are needed before any conclusions can be drawn about the astrophysical distribution of binary black hole tilts.

2. Methods

We use hierarchical Bayesian analysis—extensively described in Appendix A—to infer the astrophysical properties of the binary black holes reported in GWTC-3. To represent the astrophysical distribution of primary mass, mass ratio, redshift and spin magnitudes we use the flagship models used by the LVK in [Abbott et al. 2021e]. Namely, the primary mass distribution is their “Power law + Peak” (Tablot & Thrane 2018); the mass ratio is a power law (Fishbach & Holz 2020); the two spin magnitudes are independently and identically distributed according to a Beta distribution (Wysocki et al. 2019), and the redshift is evolving with a power law (Fishbach et al. 2018). For the (cos τ) tilt distributions, we consider several models of increasing complexity.

3. Results

In Appendix B we report our results when we use for the tilt distribution the “DEFAULT” spin model of Appendix B-2-a of Abbott et al. (2021e) (referred to as LVK default in the rest of the paper); this will serve as a useful comparison for the more complex models described in the reminder of this work. In this section we re-analyze the GWTC-3 BBH with different parameterized two-component mixture models for cos τ. We will report results for three-component models in Appendix C. It is assumed that the cos τ distribution is not correlated to any other of the astrophysical parameters. This assumption will be revisited in Appendix D.

3.1. Isotropic + Gaussian model

To check if the data requires that the Gaussian component of the LVK default mixture model peaks at cos τ = 1—Appendix B—we relax the assumption that the normal distribution must be centered at +1, That is, we treat the mean of the first parameter of the Gaussian component as a free parameter.

---

Footnotes:

2 For the BBHs detected in the second half of the third observing run (O3b), the LVK has also released tilt posteriors evaluated at minus infinity, i.e. at very large orbital separations (Mould & Gerosa 2022). We have run the Isotropic + Beta model of Sec. 3.1 on O3b sources only, and found that the analyses with tilts calculated at 20 Hz and minus infinity yield the same astrophysical cos τ distribution. We also find that using O3b only sources the cos τ distribution moves toward the left, compared to what is shown in Fig. 6 and peaks closer to 0.

3 Even though we will only report results for the cosine of the tilt angle—cos τ—we might occasionally refer to tilts only, to lighten the text.
Gaussian component $\mu$ as another model parameter:

$$p(\cos \tau_1, \cos \tau_2 | \mu, \sigma, \vartheta) = \frac{1 - \vartheta}{4} + \vartheta \prod_j \mathcal{N}(\cos \tau_j, \mu, \sigma)$$  \hspace{1cm} (1)

The prior for $\mu$ is uniform in the range $[-1, 1]$ (Tab. G.2 reports the priors for the hyperparameters of all models used in the paper). The top panel of Fig. 1 shows the resulting posterior for the tilt angle. For the mean of the Gaussian component we measure $\mu = 0.48^{+0.48}_{-0.99}$ (unless otherwise stated, we quote median and 90% symmetric credible interval). Some of the uncertainty in this measurement is due to our choice to allow for large $\sigma$, since Gaussians with large $\sigma$ are rather flat, and can be centered anywhere without significantly affecting the likelihood. If we restrict the prior space to only allow for narrower Gaussians, then $\mu$ is much more constrained. E.g. if we limit to samples with $\sigma < 0.5$ ($\sigma < 1$) then $\mu = 0.28^{+0.33}_{-0.33}$ ($\mu = 0.41^{+0.47}_{-0.36}$). This can be also seen in a corner plot of the mean and standard deviation of the Gaussian component, Fig. 2 darker green. The data prefers positive means with standard deviations in the approximate range $\sigma \in [0.25, 1.5]$. Smaller values of $\sigma$ are excluded, as are Gaussians centered at negative values of $\mu$, i.e. such that $\chi_{\text{eff}}$—the mass-weighted projection of the total spin along the angular momentum (Damour 2001)—would be negative. The bottom panel of Fig. 1 shows the posterior on the differential merger rate per unit $\cos \tau$ for the same model.

It is worth noting that the astrophysical $\cos \tau$ posterior changes entirely if the prior for $\mu$ is extended to allow for values outside of the range $[-1, 1]$ (while the resulting population model is, of course, still truncated and properly normalized in that domain). For example, Fig. 3 shows the posterior for the tilt distribution obtained with a wider $\mu$ prior, uniform in the range $[-5, 5]$. Here again it is the case that the distribution is consistent with having a peak for aligned spins. A look at the joint distribution of $\mu$ and $\sigma$, Fig. 2 reveals that the peak at $\mu = 1$ is not obtained because the Gaussian component peaks there, but rather by truncating Gaussians that peak at $\mu > 1$ and have large standard deviations. This results in a much steeper shape for the posterior near the $\cos \tau = +1$ edge than what could possibly be obtained by forcing $-1 \leq \mu \leq 1$. The only feature that seems solid against model variations is the fact that there is no excess of systems at negative $\cos \tau$.

We can better visualize what happens at the edges of the $\cos \tau$ domain—i.e. for values of tilts close to aligned ($\cos \tau \approx 1$) or anti-aligned ($\cos \tau \approx -1$)—by plotting the ratio of the posterior support for aligned spin vs anti-aligned spins. This is equivalent to making a histogram of the ratio of the value that the thin black curves in Fig. 3 take on the far right and far left side. Specifically, for each of the posterior draws, we calculate an asymmetry coefficient $Y$ defined as

$$Y(\delta) \equiv \frac{p(\cos \tau \in [1 - \delta, 1])}{p(\cos \tau \in [-1, -1 + \delta])}$$  \hspace{1cm} (2)
Fig. 4. An (unnormalized) histogram of $Y(\delta)$, the ratio between the probability of $\cos \tau$ for $\cos \tau \in [1 - \delta, 1]$ and $\cos \tau \in [-1, -1 + \delta]$ for the models of Sec. 3.1, 3.2, and 3.3 for $\delta = 0.01$. We also show the result for the LVK default model (Appendix B) for comparison. The solid curves are obtained by sampling the hyperposteriors. The dashed line reports the same quantity, but drawing the model’s hyperparameters from their priors. Values of $Y > 1$ imply more support for aligned than anti-aligned spins.

Similarly, in histogram it. When $Y = 1$, the $\cos \tau$ distribution takes the same value at both of the edges; when $Y > 1$ ($Y < 1$) the $\cos \tau$ distribution has more support for aligned (anti-aligned) systems than for anti-aligned (aligned) ones. We notice that the LVK default model excludes a priori excess of anti-aligned tilts, i.e. $Y < 1$. This is instead not true for the other models we consider in this section. This is shown in the top panel of Fig. 4 for $\delta = 0.01$ (i.e. considering $\sim 8$ degree widths around $\pm L$, where $L$ is the angular momentum vector). We see that the prior (dashed green curve) of $Y$ for the Isotropic + Gaussian model can extend to values smaller than 1, i.e. can produce more anti-aligned spins than aligned ones (compare with the blue dashed curve and Appendix B). In fact, we see that the prior for this asymmetry probe is much less strong than in the default LVK default model as it does not exclude $Y < 1$. As for the LVK default, the posterior of $Y$ is not inconsistent with 1. Values of $Y$ smaller than 1—i.e. an excess of anti-aligned spins—are severely suppressed relative to the prior, and so is a large excess of positive tilts. The posterior for $Y$ has a rather broad peak in the range $\sim [1, 2]$ corresponding to distributions that are consistent with being either isotropic or having a mild excess of positive alignment. In Appendix B we show similar plots for different values of $\delta$.

In the top panel of Figure 5 we show the marginalized posteriors of the branching ratio for the isotropic component, i.e. of the models described in this section, together with the reference LVK default model. While small variations exist, they all have support across the whole prior range, with a preference for small values of $i$. By comparing Figs. 1, 3, and 5 we may be surprised that the branching ratio posteriors for the two Isotropic + Gaussian runs are basically the same, and yet Fig. 1 seems to have a higher density of horizontal (i.e. isotropic) curves. This happens because the Gaussian component becomes isotropic for small $\mu$s and large $\sigma$s. Comparing the two distributions for $\mu$ on the top-left panel of Fig. 2 we see that indeed the run where $\mu$ is constrained to $[-1, 1]$ has more support at small $\mu$s, which, together with large $\sigma$s, yield nearly horizontal posteriors in Fig. 1.

To a different extent the same is true for the models described below: there exist corners of the parameter space where the non-isotropic component can in fact generate a flat distribution, and unless otherwise said we won’t a priori exclude that possibility.

3.2. Isotropic + Beta model

To allow for a more elastic model for the non-isotropic component, we now replace the Gaussian component of the previous section with a Beta distribution:

$$p(\cos \tau_1, \cos \tau_2|\alpha, \beta, b) = \frac{1 - b}{4} + b \prod_j B(\cos \tau_j, \alpha, \beta)$$

(3)

We offset the input of the Beta distribution and scale its maximum value such that it spans the domain $[-1, 1]$. We stress that we do not limit the range of $\alpha$ and $\beta$ to non-negative values, i.e. we do allow them to be smaller than 1—Tab [?]. In turn, this implies that we can get posteriors for $\cos \tau$ that peak at the edges of the range. The resulting posterior for $\cos \tau$ is shown in Fig. 6 which also shows for comparison the 90% CI when drawing hyperparameters from their priors, thin dashed lines. With this model, we recover a broad peak at small positive values of $\cos \tau$. One can convert the $\alpha$ and $\beta$ parameters of our rescaled Beta distribution to the corresponding mean as

$$\mu_0 = -1 + 2 \frac{\alpha}{\alpha + \beta}.$$ 

(4)

We find $\mu_0 = 0.19^{+0.22}_{-0.18}$. While some of the posterior draws do peak at $+1$, overall the upper edge of the 90% CI band does not show a peak in that region. Compared to what seen in the previous section, this model finds more support for for anti-aligned tilts, with a median value that is at the lower edge of the 90% CI for the Isotropic + Gaussian models. In the bottom panel of Fig. 4 we show in red the posterior (solid line) and prior (dashed line) of the asymmetry coefficient $Y$ defined in Eq. 2. For this model too we observe that the posterior disfavours configurations with an excess of anti-aligned black hole tilts, relative to the prior. We notice that for the Isotropic + Beta model values of $Y > 1$ do not necessarily represent a peak at positive tilts, see
3.3. Isotropic + Tukey model

We end our exploration of two-component models with a mixture of an isotropic distribution and a distribution based on the Tukey window function. Mathematically:

$$\begin{align*}
p(\cos \tau_1, \cos \tau_2|T_{\alpha}, T_r, T_k) &= \frac{1 - t}{4} \cdot \prod_{j=1}^{2} T(\cos \tau_j, T_{\alpha}, T_k, T_r) 
\end{align*}$$

where the exact expression for the functional form of the window function, and a few examples are given in Appendix E. Figure 7 reports the resulting posterior distribution for $\cos \tau$, together with the prior (dotted black lines). The 90% CI band feature a plateau that extends from $\cos \tau \sim 0.5$ to $+1$.

In Fig. 8, we show the posterior distribution for the parameters controlling the Tukey channel, together with the corresponding branching ratio. The Tukey component is centered at $T_{\alpha}$, whose 5th and 95th percentile are $-0.37$ and $0.94$, respectively. The marginal posterior for $T_k$ prefers values close to $1.8$, implying wider Tukey distributions. Smaller values of $T_k$ are possible for the near entirety of the support at $T_{\alpha} < 0$, since a Tukey that is basically a uniform distribution can be centered anywhere without affecting the likelihood. If we restrict the analysis to $T_k \leq 2$ ($T_k \leq 1$), we get that $T_{\alpha}$ is much better constrained to $T_{\alpha} = 0.46^{+0.04}_{-0.04}$ ($T_{\alpha} = 0.30^{+0.04}_{-0.07}$), which excludes negative values at $\sim 90\%$ credibility. The fact that our generous hyperparameter priors allow for Tukey distributions that resemble isotropic ones also explain the peak at $Y = 1$ in the bottom panel of Fig. 4, purple lines (solid for the posterior, dashed for the prior). We find, once again, that the data does not exclude negative values at $\sim 90\%$ credibility. The fact that our generous hyperparameter priors allow for Tukey distributions that resemble isotropic ones also explain the peak at $Y = 1$ in the bottom panel of Fig. 4, purple lines (solid for the posterior, dashed for the prior). We find, once again, that the data does not exclude negative values at $\sim 90\%$ credibility. The fact that our generous hyperparameter priors allow for Tukey distributions that resemble isotropic ones also explain the peak at $Y = 1$ in the bottom panel of Fig. 4, purple lines (solid for the posterior, dashed for the prior). We find, once again, that the data does not exclude negative values at $\sim 90\%$ credibility. The fact that our generous hyperparameter priors allow for Tukey distributions that resemble isotropic ones also explain the peak at $Y = 1$ in the bottom panel of Fig. 4, purple lines (solid for the posterior, dashed for the prior). We find, once again, that the data does not exclude negative values at $\sim 90\%$ credibility. The fact that our generous hyperparameter priors allow for Tukey distributions that resemble isotropic ones also explain the peak at $Y = 1$ in the bottom panel of Fig. 4, purple lines (solid for the posterior, dashed for the prior). We find, once again, that the data does not exclude negative values at $\sim 90\%$ credibility. The fact that our generous hyperparameter priors allow for Tukey distributions that resemble isotropic ones also explain the peak at $Y = 1$ in the bottom panel of Fig. 4, purple lines (solid for the posterior, dashed for the prior). We find, once again, that the data does not exclude negative values at $\sim 90\%$ credibility. The fact that our generous hyperparameter priors allow for Tukey distributions that resemble isotropic ones also explain the peak at $Y = 1$ in the bottom panel of Fig. 4, purple lines (solid for the posterior, dashed for the prior). We find, once again, that the data does not exclude negative values at $\sim 90\%$ credibility. The fact that our generous hyperparameter priors allow for Tukey distributions that resemble isotropic ones also explain the peak at $Y = 1$ in the bottom panel of Fig. 4, purple lines (solid for the posterior, dashed for the prior). We find, once again, that the data does not exclude negative values at $\sim 90\%$ credibility. The fact that our generous hyperparameter priors allow for Tukey distributions that resemble isotropic ones also explain the peak at $Y = 1$ in the bottom panel of Fig. 4, purple lines (solid for the posterior, dashed for the prior). We find, once again, that the data does not exclude negative values at $\sim 90\%$ credibility. The fact that our generous hyperparameter priors allow for Tukey distributions that resemble isotropic ones also explain the peak at $Y = 1$ in the bottom panel of Fig. 4, purple lines (solid for the posterior, dashed for the prior). We find, once again, that the data does not exclude negative values at $\sim 90\%$ credibility.
4. Conclusions

In this paper we have re-analyzed the LVK’s 69 BBHs of GWTC-3 using different models for the astrophysical distribution of the black hole spin tilt angle, i.e. the angle the spin vector forms with the orbital angular momentum at a reference frequency (20 Hz). Black hole spin tilts can yield precious information about their astrophysical formation channels. It is usually expected that dynamical formation of binaries results in an isotropic distribution of the spin vectors (e.g., Portegies Zwart & McMillan 2002; Rodriguez et al. 2015; Antonini & Rasio 2016; Rodriguez et al. 2019; Gerosa & Fishbach 2021). On the other hand, for black hole binaries formed in the field via isolated binary evolution, it is expected that the spins are nearly aligned with the angular momentum (e.g., Tutukov & Yungelson 1993; Kalogera 2000; Belczynski et al. 2020; Zaldarriaga et al. 2018; Stevenson et al. 2017; Gerosa et al. 2018), i.e. that tilts are small, as the angular momenta of the progenitor stars are aligned by star-star and star-disk interactions (Hut 1981; Paczynski 1981). Indeed, if the binary forms in the field, the only mechanism that could yield significant black hole tilts are asymmetries in the supernovae explosions that create the black holes. These asymmetries can impart a natal kick large enough to tilt the orbital plane (Katz 1975; Kalogera 2000; Hurley et al. 2002). However, the black hole natal kick distribution is poorly understood both theoretically (e.g., Dominik et al. 2012; Zevin et al. 2017; Mapelli & Giacobbo 2018; Repetto et al. 2012; Giacobbo & Mapelli 2020; Fragos et al. 2010) and observationally (Brandt et al. 1995; Nelemans et al. 1999; Mirabel et al. 2001; 2002; Wong et al. 2014).

These expectations explain why in their most recent catalog, the LVK has modeled the astrophysical tilt distribution as a mixture of two components: an isotropic part and a Gaussian distribution centered at cosτ = 1 and with a width that is measured from the data. This is a rather strong model, as it forces onto the data a Gaussian that must be centered at +1. This might not be advisable, because individual tilt measurements are usually broad, which implies that the functional form of the assumed astrophysical model can leave a discernible imprint on the posterior. Given the rather large uncertainties about how much spin misalignment can be produced in each channel, and the fact that the BBH population currently on hand might contain back holes formed in different channels (Zevin et al. 2021; Wong et al. 2021; Bouffanais et al. 2021; Francolin & Paoli 2021), it is legitimate to question whether the data requires that (as opposed to: is consistent with) the cosτ distribution peaks at +1. We find that it does not.

We consider three 2-component mixture models: Isotropic + Gaussian, made of an isotropic component and a Gaussian component whose mean is not fixed at +1, but rather measured from the data; Isotropic + Beta, made of an isotropic component and a (singular) Beta component; Isotropic + Tukey, made of an isotropic component and a distribution based on the Tukey window (Katz 1975; Kalogera 2000; Hurley et al. 2002). We find that the only model with a posterior that peaks at cosτ = 1 is the Isotropic + Gaussian, and only if we allow the mean of the Gaussian to take values outside of the cosτ domain [−1, 1]: Fig. 2 shows that in this case the Gaussian component peaks at μ > 1 and is very broad, yielding a peak at +1, Fig. 3. The other two models yield either a posterior that peaks at cosτ = 0.2 and no peak at +1, or a plateau from cosτ = −0.5 to +1. For all of the above models the data is not decisively ruling out a fully isotropic tilt distribution, but is inconsistent with an excess of systems with large and negative tilts. The models we considered found features—on top of an isotropic distribution—whose exact shape depends to a large extent on the model’s flexibility. Therefore, the fact that the LVK default model finds support for a peak at cosτ = 1 can be explained because it can only add support at cosτ = 1 when trying to match any population features on top of the isotropic distribution. Our results agree with previous results for population models fitting the distribution of χ_eff, which find only a small fraction of sources with negative χ_eff, implying negative tilts (Roulet & Zaldarriaga 2019; Miller et al. 2020; Roulet et al. 2021; Callister et al. 2022). Indeed, if we recast our inferred distributions for cosτ and spin magnitude to the resulting χ_eff distribution, we obtain results consistent with Abbott et al. 2021e). In the appendices below we report on other models. In Appendix C we consider three 3-component mixture models: Isotropic + 2 Gaussians, made of an isotropic component and 2 Gaussian components; Isotropic + Gaussian + Beta, made of an isotropic component and a Beta component and a Gaussian component; Isotropic + Gaussian + Tukey, made of an isotropic component, a distribution based on the Tukey window and a Gaussian component. These more elastic models are consistent with a broad plateau in the cosτ posterior that extends from ~ −0.5 to +1. Whether there also are peaks or features at small positive cosτ and/or at +1 depends on the exact model.

In Appendix D we augment some of the 2-component mixture models to allow for correlations between the tilt angles and another of the binary parameters: component masses, component spins, mass ratio and total mass in turn. One might expect some correlation as the mechanisms that can misalign the binary orbital plane or the black hole spins are affected by the binary parameters (e.g. Janka & Mueller 1994; Burrows & Hayes 1996; Fryer & Kusenko 2006; Gerosa & Fishbach 2021). To keep the number of model parameters limited—consistently with the relatively limited number of sources—we only consider linear correlations, and only correlate the tilt distribution with one other parameter at a time. However, even with these limitations we find that the current dataset cannot significantly constrain eventual correlations. For all of the models considered in this work, we report Bayesian evidences, Tab. C.1 which might be used to calculate odds ratios. We find that, within sampling and numerical uncertainties—all of the models are equally supported by the data.

We conclude that the current dataset is not yet large and informative enough to prove that the astrophysical tilt distribution has features, nor that if features exist they manifest as an excess of systems with nearly aligned spin vectors. On the contrary, most of the models we considered yield a broad peak in the astrophysical cosτ distribution at small and positive values. The only conclusion that is consistently found across all models is that there is no excess of systems with negative tilts, relative to what is expected in an isotropic distribution. Our results agree with the literature (e.g. Callister et al. 2022; Tong et al. 2022; Abbott et al. 2021c; Mould et al. 2022) on the lack of an excess of cosτ < −1 but disagree on other details (e.g. whether there is an hard cutoff in the cosτ distribution at cosτ = 0 (cfr Callister et al. 2022)). The point of this work is to show that those disagreements are to be expected, given the information in the current dataset. The next observing run of LIGO, Virgo and KAGRA is scheduled to start in early 2023 (Abbott et al. 2018) and should yield hundreds of BBH sources. Those may yield a first firm measurement of the astrophysical distribution of the tilt angle, and possibly allow us to begin probing correlations with other astrophysical parameters.
Appendix A: Hierarchical inference

We aim to measure the hyper parameters \( \lambda \) that control the distribution of single-event parameters \( \theta \) (the black hole masses, spins, redshifts, etc.) given the dataset \( D \) consisting of the 69 GWTC-3 BBHs with false alarm ratio smaller than 1 per year—\( D \equiv \{d_i, i = 1, \ldots, 69\} \)—reported by the LVK collaboration [Abbott et al. 2021e].

The posterior for \( \lambda \) can be written as (Mandel et al. 2019; Fishbach et al. 2018; Vitale et al. 2020):

\[
p(\lambda | D) \propto \pi(\lambda) \prod_{i=1}^{69} p(d_i | \lambda) / \alpha(\lambda),
\]

where we have analytically marginalized over the overall merger rate, which is not relevant for our inference. The function \( \alpha(\lambda) \) represents the detection efficiency, i.e. the fraction of BBHs that are detectable, given the population parameters \( \lambda : \pi(\lambda) \) is the prior for the population hyper parameters, and \( p(d_i | \lambda) \) is the likelihood of the stretch of data containing the i-th BBH. This allows us to account for selection effects and infer the properties of the underlying, rather than the observed, population.

Using Bayes’ theorem and marginalizing over the single-event parameters, the single-event likelihood can be written as

\[
p(d_i | \lambda) = \int \, d\theta p(d_i | \theta) \pi(\theta | \lambda) \propto \int \, d\theta \frac{p(\theta | \lambda, \mathcal{H}_{\text{PE}}) \pi(\theta | \mathcal{H}_{\text{PE}})}{\pi(\theta | \mathcal{H}_{\text{PE}})}, \quad (A.1)
\]

where \( p(\theta | \lambda, \mathcal{H}_{\text{PE}}) \) is the posterior distribution for the binary parameters \( \theta \) of the i-th source. The population hyperparameters \( \lambda \) are typically inferred using a hierarchical process that first involves obtaining posteriors for \( \theta \) for each individual event under a non-informative prior, \( \pi(\theta | \mathcal{H}_{\text{PE}}) \). The hypothesis \( \mathcal{H}_{\text{PE}} \) represents the settings that were used during this individual-event parameter estimation step. The last term, \( \pi(\theta | \lambda) \), is the population prior, i.e., our model for how the parameters \( \theta \) are distributed in the population, given the hyper parameters.

The integral in Eq. (A.1) can be approximated as a discrete sum

\[
\int \, d\theta p(d_i | \theta) \pi(\theta | \lambda) \approx N_{\text{samples}}^{-1} \sum_{k=1}^{N_{\text{samples}}} \frac{\pi(\theta_k | \lambda)}{\pi(\theta_k | \mathcal{H}_{\text{PE}})},
\]

where the \( N_{\text{samples}} \) samples are drawn from the posterior distribution of the i-th event. We use the posterior samples of the 69 BBHs reported in GWTC-3, as released in Abbott et al. (2020b). For the sources reported in GWTC-1, we use the samples labelled 1MRPhenomPv2_posterior in the data release; for GWTC-2 we use PublicationSamples; for GWTC-2.1 we use PrecessingSpin1MRHM, and for GWTC-3 we use C01:Mixed. To sample the hyper posterior we use the dynesty [Speagle 2020] sampler available with the GWPopulation package [Talbot et al. 2019].

The detection efficiency \( \alpha(\lambda) \) can also be calculated through an approximated sum starting from a large collection of simulated BBHs for which the SNR (or another detection statistic) is recorded, as described in Farr (2019), Abbott et al. (2021c). We use the endo3_bbhpopt_LIGO-12100113-v12-1238166018-158436000 hd5 sensitivity file released by the LVK (Abbott et al. 2021c) to calculate \( \alpha(\lambda) \), using a false alarm threshold of 1 per year to identify detectable sources, consistently with Abbott et al. (2021c).

Appendix B: Reference tilt model

We will be comparing our results against the LVK’s model (LVK default) of Abbott et al. (2021c) a mixture between an isotropic component and Gaussian distribution with \( \mu = 1 \) and an unknown standard deviation:

\[
p(\cos \tau_1, \cos \tau_2 | \sigma, g) = 1 - g \frac{2}{4} \int \sum_j \mathcal{N}(\cos \tau_j, \mu = 1, \sigma)^2
\]

The Gaussian component is truncated and normalized in the range \([-1,1]\). The two hyperparameters of LVK default are thus the branching ratio \( g \) of the Gaussian component and its standard deviation \( \sigma \), the same for both black holes. We notice that in Talbot & Thrane (2017) the two normal distributions can assume different values of \( \sigma \). However, since in general the spins of the least massive objects are measured with extremely large uncertainty, there are no reasons to expect that imposing the same distribution to both tilts will introduce biases.

In Fig. B.1 we show the resulting inference on the cos \( \tau \), which—modulo differences in sampling settings—is directly comparable to what is presented by the LVK in Abbott et al. (2021) (their Fig. 15). The colored area shows the 90% credible interval (CI), the thick dashed line is the median, and the thin lines represent individual draws from the posterior. The two dashed lines represent the edges of the 90% credible interval obtained by sampling the hyperparameters from their priors. It is worth noticing that the LVK default model excludes a priori the possibility of an excess of tilts relative to isotropy (i.e. a posterior larger than 0.5) at negative values, as well as a dearth of tilts relative to isotropy for cos \( \tau \geq 0.45 \). Just as Abbott et al. (2021c), we find that the posterior is not inconsistent with a fully isotropic tilt distribution, while preferring an excess of positive alignment.

This is shown in the top panel of Fig. B.4, where the solid blue line is obtained using samples from the hyperparameters’ posterior whereas the dashed blue line is obtained by sampling their priors. The fact that there is a hard cutoff at \( \tau = 1 \) (the finite bin size causes the curves to extend to values slightly smaller than 1) is just a symptom of the fact that the LVK default model ex-
includes a priori an excess of negative tilts and a dearth at positive tilts, as mentioned above.

The curve is consistent with $Y = 1$, i.e. isotropic posteriors are perfectly consistent with the data, even though it should be appreciated that the model prefers that region a priori. The level of consistency can also be assessed with Figure 5 which reports with dashed blue lines the marginalized posterior on the branching ratio of the isotropic component (as opposed to the Gaussian component, to allow direct comparisons with other models). While broad, it favors small values for the fraction of sources in the isotropic component, though fully isotropic distributions ($i = 1$) are not excluded. The other curves in the figure are discussed in the main body. For all of our models, Tab. C.1 reports the Bayesian evidence, maximum log-likelihood and the number of parameters for the cos $\tau$ model, as a differential relative to the default LVK model. That table also includes a fully isotropic model (Isotropic, with $p(\cos \tau_1, \cos \tau_2) = 1/4$), which we include as a useful reference. The Isotropic model performs the worst, though not at the point that it can be ruled out with high confidence.

**Appendix C: Three-component models**

The results presented in Sec. 3 show that, depending on the exact model being used, the tilt distribution seems to show either a peak at $+1$, a peak at a smaller positive values of cos $\tau$, or a broad plateau for positive cos $\tau$. As this might suggest that two peaks, or features, are present in the data, in this Appendix we consider models that comprise of an isotropic components, plus two other components. To explore the effect of the model on the resulting posterior, we consider different functional forms.

**C.1. Isotropic + Gaussian + Beta model**

We use a mixture model with an isotropic component, a Beta distribution component and a Gaussian component:

\[
p(\cos \tau_1, \cos \tau_2|0, b, \mu, \sigma, \alpha, \beta) = \frac{1 - \alpha - \beta}{4} + b \sum_{j=1}^{2} \mathcal{B}(\cos \tau_j, \alpha, \beta) + \mathcal{N}(\cos \tau_j, \mu, \sigma)
\]  

(C.1)

In order to reduce degeneracy between the Gaussian and the Beta components, we set the uniform prior for the mean of the Gaussian component to $\mu \sim \mathcal{U}(0, 5)$; meanwhile we restrict the prior of the Beta parameters to non-singular values, $\alpha, \beta \sim \mathcal{U}(1, 20)$. In practice, this reduces the possibility that the two components can both create peaks in the same region of the cos $\tau$ domain, which would increase degeneracy and hence make sampling more inefficient.

The resulting cos $\tau$ posterior in shown in Fig. C.1. The individual posterior draws are colored according to their value of $\alpha$ (we stress that small (large) $\alpha$ does not necessarily imply large (small) b since the isotropic fraction $i \equiv 1 - \alpha - b$ needs not be zero). The 90% CI shows traces of the two features we encountered previously, namely a peak at $+1$ and one at smaller positive values of cos $\tau$. As with all of the other models explored in this paper (and, to our knowledge, in the literature) we find that the data excludes an excess of black holes with cos $\tau \approx -1$. The corner plot in Fig. C.2 shows the three branching ratios for this model. The prior for $g$ and $b$ was uniform in the plane, with the constrain that $g + b \leq 1$; we show the resulting marginal priors as dotted black lines in the diagonal panels. This model prefers small values of $b$ coupled with large values of $g$, as shown in the top-left off-diagonal panel. There is little posterior support for even moderate values of $b$: the 95th percentile for the marginal posterior $p(b)$ is 0.50. As already visible in Fig. C.1, the Beta component peaks at small positive values of cos $\tau$: we find $\mu_\beta = 0.15^{+0.39}_{-0.45}$.

**C.2. Isotropic + 2 Gaussians model**

Next, we use a mixture model with an isotropic component, and two Gaussian distributions. Here too, to avoid perfect degeneracy, we restrict somewhat the allowed range of the Gaussian means. The Gaussian on the right (index “R”) has a mean that can only vary in the range $[0, 5]$. The prior for the Gaussian one on the left (index “L”) spans the range $[-1, 1]$, however, we set to 0 the likelihood for samples for which $\mu \notin [A, B]$, where $A$
and $B$ are hyperparameters of the model. In practice, this implies that the Gaussian on the left is truncated (and hence normalized) in the range $[A, B]$ and has a mean in the same range, for each sample. Mathematically:

$$p(\cos \tau_1, \cos \tau_2 | \theta_L, \theta_R, \mu_L, \mu_R, \sigma_L, \sigma_R, A, B) = \frac{1 - \theta_L - \theta_R}{4} + \theta_L \prod_j N(\cos \tau_j, \mu_L, \sigma_L) + \theta_R \prod_j N(\cos \tau_j, \mu_R, \sigma_R). \quad (C.2)$$

We explicitly add hyperparameters for the domain of the left Gaussian in order to verify if the data prefers solutions that do not add posterior support to the anti-aligned ($\cos \tau \geq -1$) region (cf. Callister et al. 2022). The resulting posterior for $\cos \tau$ is shown in Fig. C.3, where the individual posterior draws are colored according to the branching ratio of the right Gaussian – $\theta_R$. The 90% CI shows again two features: a rather broad peak for small positive value of $\cos \tau$ and a second peak at $+1$.

We note that the data is informative for the parameters $A$ and $B$. Fig. C.4. While their priors are uniform, the posteriors for both $A$ and $B$ show clear peaks. For $A$ we measure $A = -0.49^{+0.14}_{-0.40}$ which notably excludes $-1$ at 90% credibility. Meanwhile, the posterior for $B$ rails against $+1$. The standard deviation for the left Gaussian, $\sigma_L$ is large (the 5th percentile of $p(\sigma_L | d)$ is 0.46) which implies that even though functionally speaking this component of our model is a Gaussian, the data seems to prefer very wide Gaussians, resembling pieces of segments. The mean of the left Gaussian prefers small positive values, $\mu_L = 0.13^{+0.52}_{-0.30}$ and shows no obvious correlation with $\sigma_L$.

Figure C.5 shows the branching ratios for the two Gaussian and the isotropic component $i \equiv 1 - \theta_L - \theta_R$, together with the corresponding priors (dashed lines). The measurement is not precise, and only small departures from the priors are apparent. In particular, for both $\theta_L$ and $\theta_R$, the posteriors yield a wide peak at $\sim 0.5$, whereas $i$ peaks at 0 more than the prior.

We stress that the posterior on $\cos \tau$ we obtained for this model is heavily impacted by the fact that the left Gaussian is truncated in a range, whose position is measured from the data. If instead we set $-A = B = 1$, i.e. we extend (and normalize) the left Gaussian to the full $\cos \tau$ range, we obtain a radically different posterior, Fig. C.6. The branching ratio for left Gaussian component in this case doesn’t show significant differences relative to the prior. While some of the posterior draws show prominent peaks for small positive values of $\cos \tau$, those are not frequent enough to create a visible peak in the 90% CI band, as was instead the case in Fig. C.3.

Given that the only difference between the models behind Fig. C.6 and Fig. C.3 is the truncation of the left Gaussian’s domain, it is tempting to think that the tails of the left Gaussian—if free to extend all the way to $\cos \tau = -1$—would give too much...
posterior weight in that region, which is not supported by the data. This explanation is also consistent with the fact that our model of Sec. C.1 does find the peak, since a Beta distribution can produce tails which are less wide than a Gaussian.

### C.3. Isotropic + Gaussian + Tukey model

To end our exploration of 3-component models, we modify the model of the previous section and replace the left Gaussian with a distribution based on the Tukey window function. Mathematically:

$$p(\cos \tau_1, \cos \tau_2 | \theta, T_{t0}, T_k, T_r, \mu, \sigma) = \frac{1 - \theta - \gamma}{4} + \prod_j T(\cos \tau_j, T_{t0}, T_k, T_r) + \theta \prod_j N(\cos \tau_j, \mu, \sigma)$$  \hspace{1cm} (C.3)

The priors for all of the hyper-parameters are uniform, with the exception of the branching ratios $t$ and $\gamma$, which are jointly uniform in the triangle $t + \gamma \leq 1$.

Figure C.7 shows the posteriors for the branching ratios, including that of the isotropic component $i = 1 - t - \gamma$. As for the Isotropic + 2 Gaussians model, the branching ratios are not measured with precision. The data prefers smaller values of $t$ and $i$ and moderate values of $\gamma$. Comparing Fig. C.8 with the corresponding plot for the Isotropic + Tukey run—Fig. 8—we find qualitatively consistent results. In particular, $T_{t0}$ has mostly support at positive values, except when $T_k$ can take large values or $t$ is small. Using the full posterior, we find $T_{t0} = 0.34_{-0.58}^{+0.58}$, while restricting to samples with $T_k \leq 2$ ($T_k \leq 1$) yields $T_{t0} = 0.37_{-0.81}^{+0.53}$ ($T_{t0} = 0.26_{-0.52}^{+0.54}$) consistent with the simpler 2-component model.

Similarly, we find that the posterior for $\cos \tau$ mainly differs from that of Fig. 7 because of some additional – but not large – support at $\cos \tau = 1$, due to the contribution of the Gaussian component.

### Appendix D: Correlated mixture models

In this Appendix we revisit some of the models presented in the main body of the paper, and we extend them to allow for the possibility that the hyperparameters governing the population-level $\cos \tau$ distribution are correlated with some of the astrophysical binary parameters. Previous works have considered correlations between the effective aligned spin, $\chi_{\text{eff}}$, and the BBH masses and redshifts (Safarzadeh et al. 2020; Callister et al. 2021; Franciolini & Pani 2022; Biscoveanu et al. 2022; Adamiec & Thrane 2022), but not a direct correlation between the tilts and these other intrinsic parameters. Given that the number of BBH sources is still relatively small, we only consider a subset of 2-component models, in order to keep the number of
hyperparameters limited. For some of the correlated models, the distributions for the two tilt angles are not assumed to be identical (this happens when each tilt is allowed to be correlated with the corresponding component mass or spin magnitude): we will only report the distribution for the tilt of primary (i.e. most massive) black hole, as it is usually best measured.

D.1. Isotropic + correlated Gaussian model

We first allow for the possibility that the mean and standard deviation of the Gaussian component might be correlated with other astrophysical parameters – $\kappa$, described below – since those should be related to the details of the supernovae explosions that would have tilted the orbit (e.g., Gerosa et al. 2018). We minimally modify the Isotropic + Gaussian model to allow the mean and standard deviation to linearly vary with the parameter that is correlated to the spin tilt. This introduces another set of hyperparameters, which control the linear part of the mean and standard deviation:

$$p(\cos \tau_1, \cos \tau_2 | \mu_a, \sigma_a, \sigma_b, \kappa_1, \kappa_2) = \frac{1 - \kappa}{4} + \kappa \sum_j \mathcal{N}(\cos \tau_j, \mu(k_j), \sigma(k_j)), \quad (D.1)$$

where $\mu(k) = \mu_a + \frac{\kappa}{2} \mu_b$ and $\sigma(k) = \sigma_a + \frac{\kappa}{2} \sigma_b$.

Notice that even though we could also have allowed for correlations in the branching ratio $\kappa$, we decided not to, as that parameter is already very poorly measured, cfr. Fig 5. Similarly, and following Callister et al. (2021); Biscoveanu et al. (2021) we only consider linear correlations. As more sources are detected, both of these assumptions might be trivially relaxed. The constant $N$ is chosen to guarantee that the coefficient of $\mu_b$, $\sigma_b$ is always smaller than 1. We explore the following possible correlations:

- Component masses, $\kappa_1 = m_1, \kappa_2 = m_2, N = 100 M_\odot$
- Component spins, $\kappa_1 = \chi_1, \kappa_2 = \chi_2, N = 1$
- Mass ratio, $\kappa_1 = \kappa_2 = q, N = 1$
- Total mass, $\kappa_1 = \kappa_2 = m_{\text{tot}}, N = 200 M_\odot$

We find that we cannot constrain in any significant way the parameters that enact the correlations (i.e. $\mu_b$ and $\sigma_b$), for which we recover posteriors which highly resemble the corresponding priors. This is shown in Fig. D.1 where we report the parameters of the Gaussian component for the analysis where they are correlated with the component masses. The dashed horizontal lines in the diagonal panels represent the corresponding priors. Especially for the standard deviation term $\sigma_b$, no information is gained relative to the prior. Because we restrict the prior on $\sigma_b$ to non-negative values (see Tab. G.2) to ensure that the width of the Gaussian does not become negative for any values of $\kappa$, this implies that the overall standard deviation for the Gaussian component can only increase with the mass, Fig. D.2. However, as made clear by comparing the 90% CI band with the extent of the 90% CI obtained with prior draws, the increase of $\sigma$ is entirely prior-driven.

D.2. Isotropic + correlated Beta model

Finally, we augment the Isotropic + Beta model of Sec. 3.2 to allow for correlations in the parameters that control the Beta component:

$$p(\cos \tau_1, \cos \tau_2 | \alpha_a, \beta_a, \beta_b, \kappa_1, \kappa_2, b) = \frac{1 - b}{4} + b \sum_j \mathcal{B}(\cos \tau_j, \alpha(k_j), \beta(k_j)), \quad (D.2)$$

with $\alpha(k) = \alpha_a + \frac{\kappa}{2} \alpha_b$ and $\beta(k) = \beta_a + \frac{\kappa}{2} \beta_b$.

We consider the same possible correlations (i.e. values of $\kappa$ and N) described in the previous section. As for the previous correlated model, we restrict the priors for $\alpha_b$ and $\beta_b$ to the non-
Vitale et al.: Spin it as you like

Fig. D.2. Posterior of the standard deviation of the Gaussian component for the Isotropic + correlated Gaussian model, when we allow for correlation with the component masses, as a function of the primary mass $m_1$. The thin black lines are individual posterior draws, the colored band is the 90% CI and the thick dashed line is the median. The two horizontal thin dotted lines enclose the 90% CI for the Isotropic + Gaussian model, which does not allow for correlations. Finally, the two thin blue dashed line enclose the 90% CI obtained sampling the prior.

Fig. D.3. Joint and marginal posteriors for the Beta parameters obtained in the analysis where they are correlated with the mass ratio. Dashed lines in the diagonal panels represent the priors.

Fig. D.4. Same as D.2 but for the $\beta$ parameter of the Isotropic + correlated Beta model, when correlated with the mass ratio $q$.

Fig. D.5. Conditional posteriors for $\cos \tau_1$ for the Isotropic + correlated Beta model, when $\cos \tau$ is correlated with the mass ratio. Colored bands show the 90% CI posterior for $\cos \tau_1$ conditional on a fixed value of the mass ratio; yellow dashed lines enclose the 90% CI obtained sampling the correlated parameter from its inferred astrophysical distribution; black dashed lines enclose the 90% CI of the Isotropic + Beta model, which does not allow for correlations. The increased support at $+1$ as $q$ decreases is mostly prior-driven.

negative domain to ensure that the overall $\alpha$ and $\beta$ parameters of the Beta distribution do not become negative, enforcing that only positive correlations can exist between the $\cos \tau$ distribution and $\kappa$. For this model we find that the current dataset cannot significantly constrain the correlation parameters, even though we don’t recover exactly the priors. For example, in Fig. D.3 we show the posterior and priors (thin dashed lines) for the parameters of the Beta distribution when we allow correlations with the mass ratio $q$. The parameters that enact the correlation, $\alpha_b$ and $\beta_b$ have wide posteriors, which however are not as similar as their prior as $\sigma_b$ was for the Isotropic + correlated Gaussian models, Fig. D.1. Figure D.4 shows that the main impact of the measurement, relative to the prior, is to exclude large values of $\beta$. However, it is still the case that the overall trend in the 90% CI of the Beta parameters are prior dominated. Just as for the Isotropic + correlated Gaussian models, this results in more support at $\cos \tau \approx +1$ for small masses, mass ratios or spins. Functionally, this happens because the parameters controlling the Beta component take smaller values at small values of the correlated parameter, and that moves the peak toward the edge of the domain—e.g. Fig. D.5 for correlations with the mass ratio. However, just as for the Isotropic + correlated Gaussian model, these trends are mostly a result of the prior and of the model. 
Appendix E: Tukey window

We implement the Tukey window used in Eq. 5 as

\[
T(x, T_{x0}, T_k, T_r) \propto \begin{cases} 
0, & x < \max(-1, T_{x0} - T_k) \\
\frac{1}{2} \left[ 1 + \cos \left( \frac{\pi}{T_k T_r} (x - T_{x0} + T_k T_r) \right) \right], & \max(-1, T_{x0} - T_k) \leq x < T_{x0} - T_k (1 - T_r) \\
1, & T_{x0} - T_k (1 - T_r) \leq x < T_{x0} + T_k (1 - T_r) \\
\frac{1}{2} \left[ 1 + \cos \left( \frac{\pi}{T_k T_r} (x - T_{x0} - T_k T_r) \right) \right], & T_{x0} + T_k (1 - T_r) \leq x \leq \min(+1, T_{x0} + T_k) \\
0, & x > \min(+1, T_{x0} + T_k) 
\end{cases}
\]

(E.1)

Fig. E.1. Four examples of the distribution in Eq. E.1

This represents a Tukey window that is symmetric around \(T_{x0}\) and whose domain is \(2T_k\) wide. The parameter \(T_r\) controls the shape of the window (\(T_r = 0\) gives a rectangular window while \(T_r = 1\) gives a cosine). The distribution is then truncated and normalized in the range \([-1, 1]\). Figure E.1 shows four examples. Since the width, the shape, and the position can all be varied, this model is quite elastic and can latch onto both broad and narrow features. We highlight that in the default setting, we allow the uniform prior of \(T_k\) to go up to 4, Tab G.2. This implies, that just as for the Isotropic + Gaussian model, there are parts of the parameter space where the non-isotropic component can be made very similar to, or indistinguishable from, the isotropic component. In this case, that happens when \(T_k\) is large and \(T_r\) is small. This distribution can also produce curves that ramp up from zero to a plateau, with various degrees of smoothness: the green thick line in Fig E.1 is an example and—if \(t\) were zero—would produce a \(\cos \tau\) distribution similar to the second row in Fig. 5 of Callister et al. (2022).
Appendix F: Asymmetry $Y(\delta)$ for various values of $\delta$

![Graphs showing $Y(\delta)$ for various values of $\delta$](image)

**Fig. F.1.** Same as previous but for various values of $\delta$.

Appendix G: Tables

Table G.1 reports for all models (including those discussed later in other appendices) the number of spin parameters, the natural log of the Bayesian evidence and the natural log of the maximum likelihood point. All quantities are expressed as deltas relative to the reference LVK default model. Table G.2 lists the priors used for the hyperparameters of all models.
the data.

one may draw is that a purely isotropic model yields the worst match to these uncertainties, the only reliable – yet unsurprising – conclusions

\[ \log \Delta \text{gral} \] in Eq. A.1. Golomb & Talbot (2022) estimates that to be roughly \( \pm \) uncertainty of \( \Delta \log L \). With our settings, the evidences carry a statistical number of parameters for the \( \cos \tau \) models (e.g. LVK default and Isotropic + Gaussian) the broader model finds a slightly lower maximum likelihood point. Given these uncertainties, the only reliable – yet unsurprising – conclusions one may draw is that a purely isotropic model yields the worst match to the data.

| Run | lnEvidence | lnMaxL | # spin pars |
|-----|------------|--------|-------------|
| Isotropic | -0.8 | -3.2 | -2 |
| LVK default | ref | ref | ref |
| Isotropic + Gaussian \( \mu \in [1,1] \) | -0.3 | -0.4 | +1 |
| Isotropic + Gaussian \( \mu \in [-5,5] \) | -0.4 | -0.2 | +1 |
| Isotropic + Beta | +0.4 | +0.4 | +1 |
| Isotropic + Tukey | -0.2 | +0.7 | +2 |
| Isotropic + correlated Gaussian \( m \) | -0.4 | -0.6 | +3 |
| Isotropic + correlated Gaussian \( q \) | -0.5 | -0.1 | +3 |
| Isotropic + correlated Gaussian \( \chi \) | -0.7 | -0.1 | +3 |
| Isotropic + correlated Beta \( M_{\text{tot}} \) | -0.4 | +0.0 | +3 |
| Isotropic + correlated Beta \( g \) | -0.1 | -0.3 | +3 |
| Isotropic + correlated Beta \( \chi \) | +0.3 | -0.6 | +3 |
| Isotropic + correlated Beta \( M_{\text{tot}} \) | -0.3 | -0.7 | +3 |
| Isotropic + Gaussian + Beta | +0.3 | -0.6 | +4 |
| Isotropic + Gaussian + Beta \( \chi \) | -0.1 | -0.3 | +4 |
| Isotropic + 2 Gaussians \( -\chi = A = 1 \) | +0.1 | +0.3 | +5 |
| Isotropic + Gaussian + Tukey | +0.6 | +0.2 | +6 |

**Table G.1.** Bayesian evidence, maximum log likelihood value an number of parameters for the \( \cos \tau \) models (relative to the reference LVK model of Eq. B.1). With our settings, the evidences carry a statistical uncertainty of \( \pm 0.15 \). Additional uncertainties in the log likelihood – and hence evidence – arise from the numerical evaluation of the integral in Eq. A.1. Golomb & Talbot (2022) estimates that to be roughly \( \Delta \log L \leq 1 \) when using the publicly released LVK injection sets to estimate selection effects. We also note that nested sampling algorithms do not aim at finding the highest likelihood point, so it is possible that for nested models (e.g. LVK default and Isotropic + Gaussian) the broader model finds a slightly lower maximum likelihood point. Given these uncertainties, the only reliable – yet unsurprising – conclusions one may draw is that a purely isotropic model yields the worst match to the data.

| Run | lnEvidence | lnMaxL | # spin pars |
|-----|------------|--------|-------------|
| LVK default-Eq.B.1 | | | |
| \( \sigma \) | \( \mathcal{U}(0.1,4) \) | | |
| \( \mu \) | \( \mathcal{U}(0.2,0.9) \) | | |
| Isotropic + Gaussian- Eq.1 | | | |
| \( \sigma \) | \( \mathcal{U}(0.1,4) \) | | |
| \( \mu \) | \( \mathcal{U}(0.9,5) \) | | |
| Isotropic + Beta- Eq.3 | | | |
| \( \sigma \) | \( \mathcal{U}(0.1,5) \) | | |
| \( \beta \) | \( \mathcal{U}(0.05,5) \) | | |
| Isotropic + Tukey- Eq.5 | | | |
| \( T_{\chi} \) | \( \mathcal{U}(0.01,1) \) | | |
| \( T_{r} \) | \( \mathcal{U}(0.9,5) \) | | |
| \( T_{k} \) | \( \mathcal{U}(0.1,4) \) | | |
| \( t \) | \( \mathcal{U}(0.1) \) | | |
| Isotropic + Gaussian + Beta- Eq.C.1 | | | |
| \( \mu \) | \( \mathcal{U}(0.1,20) \) | | |
| \( \beta \) | \( \mathcal{U}(0.1,20) \) | | |
| \( \sigma \) | \( \mathcal{U}(0.1,5) \) | | |
| \( \gamma \) | \( \mathcal{U}(0.1,1) \) | | |
| Isotropic + 2 Gaussians- Eq.C.2 | | | |
| \( \mu \) | \( \mathcal{U}(-1.1,0.5 \mu < B \) | | |
| \( \sigma \) | \( \mathcal{U}(0.1,4) \) | | |
| \( \mu \) | \( \mathcal{U}(0.9,5) \) | | |
| \( \sigma \) | \( \mathcal{U}(0.1,5) \) | | |
| \( \mu \) | \( \mathcal{U}(0.1,4) \) | | |
| \( \mu \) | \( \mathcal{U}(0.01,1) \) | | |
| \( \mu \) | \( \mathcal{U}(0.01,1) \) | | |
| \( \mu \) | \( \mathcal{U}(0.9,5) \) | | |
| Isotropic + Gaussian + Tukey- Eq.C.3 | | | |
| \( \mu \) | \( \mathcal{U}(0.1,5) \) | | |
| \( \mu \) | \( \mathcal{U}(0.1,4) \) | | |
| \( \mu \) | \( \mathcal{U}(0.1,5) \) | | |
| \( \mu \) | \( \mathcal{U}(0.1,4) \) | | |
| \( \mu \) | \( \mathcal{U}(0.1,5) \) | | |
| Isotropic + correlated Gaussian- Eq.D.1 | | | |
| \( \mu \) | \( \mathcal{U}(0.5,5) \) | | |
| \( \mu \) | \( \mathcal{U}(0.5,5) \) | | |
| \( \sigma \) | \( \mathcal{U}(0.1,4) \) | | |
| \( \sigma \) | \( \mathcal{U}(0.1,5) \) | | |
| \( \mu \) | \( \mathcal{U}(0.1,4) \) | | |
| \( \mu \) | \( \mathcal{U}(0.1,5) \) | | |
| \( \mu \) | \( \mathcal{U}(0.1,4) \) | | |
| Isotropic + correlated Beta- Eq.D.2 | | | |
| \( \mu \) | \( \mathcal{U}(0.5,5) \) | | |
| \( \mu \) | \( \mathcal{U}(0.1,4) \) | | |
| \( \mu \) | \( \mathcal{U}(0.1,5) \) | | |
| \( \mu \) | \( \mathcal{U}(0.1,4) \) | | |
| \( \mu \) | \( \mathcal{U}(0.1,5) \) | | |
| \( \mu \) | \( \mathcal{U}(0.1,4) \) | | |

**Table G.2.** The priors used for the hyper parameters of the tilt models. All variables are dimensionless.