Impact of Bayesian Priors on the Characterization of Binary Black Hole Coalescences
Vitale, Salvatore; Gerosa, Davide; Haster, Carl-Johan; Chatziioannou, Katerina; Zimmerman, Aaron

DOI: 10.1103/PhysRevLett.119.251103
License: None: All rights reserved

Document Version
Peer reviewed version

Citation for published version (Harvard):
Vitale, S, Gerosa, D, Haster, C-J, Chatziioannou, K & Zimmerman, A 2017, 'Impact of Bayesian Priors on the Characterization of Binary Black Hole Coalescences', Physical Review Letters, vol. 119, no. 25, 251103. https://doi.org/10.1103/PhysRevLett.119.251103

Link to publication on Research at Birmingham portal

Publisher Rights Statement:
Checked for eligibility: 29/05/2019
Vitale, Salvatore, et al. "Impact of bayesian priors on the characterization of binary black hole coalescences." Physical review letters 119.25 (2017): 251103.
DOI: https://doi.org/10.1103/PhysRevLett.119.251103

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

• Users may freely distribute the URL that is used to identify this publication.
• Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
• Users may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
• Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 02. Oct. 2023
Impact of Bayesian priors on the characterization of binary black hole coalescences

Salvatore Vitale,1,* Davide Gerosa,2,1 Carl-Johan Haster,3,† Katerina Chatziioannou,3,§ and Aaron Zimmerman3,‡

1LIGO, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
2TAPIR 350-17, California Institute of Technology, 1200 E California Boulevard, Pasadena, CA 91125, USA
3Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, ON M5S 3H8, Canada

(Dated: July 19, 2018)

In a regime where data are only mildly informative, prior choices can play a significant role in Bayesian statistical inference, potentially affecting the inferred physics. We show this is indeed the case for some of the parameters inferred from current gravitational-wave measurements of binary black hole coalescences. We reanalyze the first detections performed by the twin LIGO interferometers using alternative (and astrophysically motivated) prior assumptions. We find different prior distributions can introduce deviations in the resulting posteriors that impact the physical interpretation of these systems. For instance, (i) limits on the 90% credible interval on the effective black hole spin $\chi_{\text{eff}}$ are subject to variations of $\sim 10\%$ if a prior with black hole spins mostly aligned to the binary’s angular momentum is considered instead of the standard choice of isotropic spin directions, and (ii) under priors motivated by the initial stellar mass function, we infer tighter constraints on the black hole masses, and in particular, we find no support for any of the inferred masses within the putative mass gap $M \lessgtr 5M_\odot$.

Introduction—Prior distributions are at the cornerstone of Bayesian statistics, where experimental data is used to update prior knowledge into posterior beliefs about observed phenomena. As Bayes himself put it, “there is reason to expect an event with more or less confidence according to the greater or less number of times in which, under given circumstances, it has happened without failing” [1]. Prior assumptions inevitably enter into all types of statistical inference, and the prominent role of priors is one of the main advantages of the Bayesian framework: the data analyst is forced to consider and explicitly specify how priors are incorporated into the analysis, thus avoiding the pitfalls of poorly specified assumptions [2, 3].

While choosing priors may in some cases be straightforward (for instance, a mass should always be positive), care must be taken when strong knowledge of the expected distribution is not available. In particular, in a regime where data are weakly informative and a feeble signal needs to be dug out of instrumental noise, different priors may indeed lead to different conclusions.

In this Letter we show that the current statistical inference on some black hole (BH) binary parameters through gravitational-wave (GW) observations does fall in this regime. With four likely observations of binary BH coalescences announced, and many more expected in the coming years, GW astronomy is becoming a reality [4–7]. Bayesian analysis has been used to measure the physical parameters of the binary coalescences in LIGO’s observing runs [8, 9], the BH merger rate [6, 10], as well as to perform tests of general relativity [6, 11]. Some of the physical parameters of the sources, e.g., the total mass (for heavier systems [12–14]) or the chirp mass (for lighter binaries [15–17]), strongly affect the phasing evolution of the GW signal. As a consequence, these parameters are usually well measured. Others, such as the mass ratio and spins, have a smaller impact on the waveform and are therefore harder to measure. Measurements of these parameters are more directly affected by the chosen prior (see Refs. [18, 19]).

Accurate statistical inference about BH spins is of crucial astrophysical importance. While mass and rates distributions sensibly overlap in many different scenarios (see, e.g., Refs. [20, 21]), spins are arguably the best indicator of BH binary formation channels [22–25] (residual eccentricities may also provide a promising avenue; see, e.g., Ref. [26]). Several studies have already demonstrated the potential of spin measurements to discriminate between different pathways of BH binary formation and evolution [17, 27–33]. At the same time, precise and accurate estimation of the individual BH masses can be used to reconstruct their mass function [7] and to verify if BHs lighter than $5M_\odot$ exist [34]. Quantifying the effect of the prior choice is therefore a crucial step to make solid astrophysical statements using GW data.

Current priors—Parameter estimations for the first GW events GW150914 [4, 7, 9, 35], GW151226 [5, 7], GW170104 [6] and candidate LV151012 [7, 36] were performed using priors (i) uniform in component masses $m_1$ and $m_2$, (ii) uniform in dimensionless spin magnitude $\chi_i = |S_i|/m_i^2$ and (iii) isotropic in spin directions at the reference GW frequency of 20 Hz. While masses and spin magnitudes are constant up to high post-Newtonian (PN) order, spin directions are subject to change due to orbital plane and spin precession (see, e.g., [37, 38]). By the time BH binaries enter the LIGO band, their spin misalign-
ments may be different from those at formation, which are needed to discriminate formation channels [39]. Fortunately, isotropic spin distributions are kept isotropic when evolved under the 2PN spin precession equations [39–41], and the impact of this issue on the usual prior can be neglected (but see P$_3$ below).

Despite being at first sight reasonable, the isotropic spin prior distribution carries important astrophysical consequences. It is generally believed that binaries formed in isolation in the galactic field will have, on average, some tendency towards spins aligned with the orbital angular momentum. On the other hand, spin directions of binaries formed via dynamical interactions in stellar clusters are expected to be isotropically distributed. We are therefore in a risky situation: we may be biasing our astrophysical inference by assuming a priori one of the models we try to discriminate.

A similar note can be made regarding the choice of distributing spins uniformly in magnitude. If we were to base our previous knowledge on other observed BH systems, then moderately high spins should be favored as found in most x-ray binaries [42]. On the other hand, core-envelope interactions in massive stars may produce BHs with small spins [43, 44] (and it has been suggested that primordial BH spins should also be low [45]). One may therefore want to choose a prior which is peaked at either low or high spins, or perhaps even bimodal. An agnostic approach would be to consider the BH spins as vectors and draw them uniformly in volume, rather than uniformly in magnitude and isotropic in direction. Alternatively, one might naively assume black holes form in situations where a random amount of energy goes into the spin, and draw uniformly in specific rotational energy $E_{\text{rot}} \equiv 1 - \sqrt{1 + 1 - \chi^2} / \sqrt{2}$.

The spin parameter which is measured best (arguably the only spin parameter which is currently measured at all [6, 7]) is the so-called effective spin

$$\chi_{\text{eff}} = \frac{S_i}{m_1 + S_2/m_2} \cdot \hat{L} = \frac{\chi_1 \cos \theta_1 + q \chi_2 \cos \theta_2}{1 + q}, \quad (1)$$

where $q = m_2/m_1 \leq 1$ is the mass ratio, and $\theta_i = \arccos(\mathbf{S}_i \cdot \hat{L})$ are the angles between the spins $\mathbf{S}_i$ and the binary’s orbital angular momentum $\hat{L}$. It is clear from Eq. (1) that the mass ratio, spin magnitude and spin direction priors are all entangled in determining the prior distribution of $\chi_{\text{eff}}$.

**Prior choices**—In order to explore some of these issues and gauge the impact of priors on Bayesian inference, we have reanalyzed the BH coalescences detected by LIGO during its first observing run (O1) using a variety of alternative prior distributions. Results have been obtained using the nested sampling algorithm implemented in LALInference [8] and a reduced-order quadrature [47] implementation of the IMRPhenomPv2 waveform model, which partially accounts for spin precession effects through a single parameter $\chi_p$ [48]. We restrict our study to (detector-frame [49]) chirp masses $8M_\odot \leq M_1 \leq 45M_\odot$, mass ratios $q \geq 1/8$ and dimensionless spin magnitudes $\chi_i \leq 0.89$ [47]. These restrictions are not a problem for our study since for none of the runs we performed the posterior distributions had support near these boundaries. We have 142 analyzed the 32-sec data frames publicly released by the 143 LIGO/Virgo Collaboration (loc.s.ligo.org), using the BayesWave algorithm [50, 51] to estimate the on-source power spectral density needed for the likelihood evaluations [8], marginalizing over calibration uncertainties as in Refs. [6, 7].

Each event was analyzed multiple times, using one of the following priors: uniform in individual masses and spin magnitudes, isotropic in spin direction ($P_1$, the default choice used in LIGO analyses$^2$); uniform in individual masses and rotational energy of the BHs, isotropic in spin direction ($P_2$); uniform in individual masses, spin vectors uniform in volume ($P_3$, volumetric); uniform in individual masses, bimodal in spin magnitudes $P(\chi_{\text{eff}}) \propto \exp[-(\chi_1 - \mu_1)^2/(2\sigma^2)\, +\, \exp[-(\chi_2 - \mu_2)^2/(2\sigma^2)]$ with $\mu_1 = 0, \mu_2 = 0.89, \sigma = 0.1$, isotropic in spin directions ($P_4$) [52–54]; uniform in individual masses and spin magnitude, peaked around alignment for spin direction $P(\theta_i) \propto \exp[-(\cos \theta_i - \mu)^2/2\sigma^2]$ where $\mu = 1, \sigma = 1 - \cos(10^\circ)$ ($P_5$, c.f. Ref. [29]), which suggested these parameters could be inferred with $\mathcal{O}(50)$ observations; power law in primary’s mass $P(m_{\text{1}}) \propto m^\alpha$ with $\alpha = -2.3$ (as in Kroupa’s initial mass function for massive stars [55], c.f. also Ref. [56] and references therein), uniform in secondary’s mass, uniform in spin magnitude, isotropic in spin direction ($P_6$); power law in $m_1$ with $\alpha = -2.3$, logistic prior in the mass ratio $P(q) \propto 1/[1 + \exp(-k(q - q_0))]$ with $k = 20, q_0 = 0.8$ (this is meant to mimic numerical results of BH mergers in globular clusters [57]), uniform in spin magnitude, isotropic in spin direction ($P_7$); uniform in individual masses, Gaussian around zero for dimensionless spin magnitude $P(\chi_{\text{eff}}) \propto \exp[-(\chi_{\text{eff}} - \mu)^2/(2\sigma^2)]$ with $\mu = 0, \sigma = 0.1$ ($P_8$). In what follows we occasionally refer to the evidence of the priors. What is meant is the evidence of a Bayesian model where the spins and masses are distributed as in that prior.

**Spins**—Marginalized prior and posterior distributions in $\chi_{\text{eff}}$ are shown in Fig. 1. Table I shows the corresponding values of the Kullback-Leibler (KL) divergence [58] (a measurement of the information gain between prior and $^2$ This is essentially the same prior used in [4–7, 9, 35] apart from the $M_\odot$, $q$ and $\chi_i$ limitations required by the ROQ implementation.
GW150914 & GW151226 & LVT151012 & $D_{K1}^{\chi_{\text{eff}}}$ & $D_{K1}^{\chi_p}$ & log$_{10}$O & $D_{K1}^{\chi_{\text{eff}}}$ & $D_{K1}^{\chi_p}$ & log$_{10}$O & $D_{K1}^{\chi_{\text{eff}}}$ & $D_{K1}^{\chi_p}$ & log$_{10}$O \\

$P_1$ & 1.02 & 0.03 & — & 1.93 & 0.21 & — & 0.53 & 0.03 & — \\
$P_2$ & 1.36 & 0.06 & -0.3 & 1.78 & 0.04 & 0.0 & 0.89 & 0.05 & -0.1 \\
$P_3$ & 1.52 & 0.09 & -0.4 & 1.76 & 0.02 & 0.0 & 0.95 & 0.04 & 0.0 \\
$P_4$ & 0.88 & 0.12 & 0.0 & 2.56 & 0.70 & -0.1 & 0.61 & 0.12 & -0.1 \\
$P_5$ & 4.21 & 1.75 & -1.7 & 0.82 & 0.21 & 0.0 & 0.22 & 0.07 & 0.5 \\
$P_6$ & 0.96 & 0.01 & 0.1 & 2.12 & 0.08 & 0.4 & 0.24 & 0.00 & 0.4 \\
$P_7$ & 0.93 & 0.06 & 0.4 & 2.63 & 0.02 & 0.4 & 0.26 & 0.01 & 0.5 \\
$P_8$ & 0.14 & 0.07 & 0.3 & 4.82 & 0.70 & -1.7 & 0.03 & 0.02 & -0.1 \\

TABLE I. For each of the three O1 events, we show the KL divergence $D_{KL}$ on (i) $\chi_{\text{eff}}$ and (ii) $\chi_p$, measuring the information gain between prior to posterior in bits, and (iii) the Bayesian odds ratio log$_{10}$O of each single analysis $P_n$ compared to the standard one $P_1$. $P_1$ ($P_2$) is preferred if log$_{10}$O is positive (negative). For comparison, the KL divergence between the $\chi_{\text{eff}}$ $P_1$ prior and a uniform distribution over the same range is 0.82 bits. The log odds have an uncertainty of ±0.04.

The case of the $\chi_{\text{eff}}$ 0 P$_3$ prior distribution is typically returned as a posterior almost unchanged by the data. As already mentioned, an exception is $P_5$ for GW150914, which however is disfavored at $O \sim 0.02$. For GW151226, the $P_4$ prior yields a posterior on $\chi_p$ which is quite different from the prior, $D_{K1}^{\chi_p} = 0.7$ bits, and models that data comparably well. This happens because the prior of the spins magnitude is bimodal, while the posterior prefers the high spin mode for the primary.

Component masses—Different choices for the component mass priors also carry important astrophysical implications. Figure 2 shows posterior distributions of the two BH masses for the three O1 events using the default $P_1$ prior, as well as $P_6$ and $P_7$. Both $P_6$ and $P_7$ have larger prior support for binary mass ratio close to unity compared to $P_1$. This additional weight at comparable masses has a visible effect on the posteriors. While presenting odds similar to $P_1$, the resulting posterior distributions for $P_6$ and $P_7$ now prefer the region closer to the $m_1 = m_2$ line in Fig. 2. In particular, when GW151226 is analyzed with any of these two priors, the 99% credible interval for the source-frame mass of the secondary object is above 5$M_{\odot}$.

This point has serious consequences in the astrophysical interpretation of the results. Electromagnetic ob-

---

$^3$ We notice that the $P_1$ posteriors on LVT151012 are wider than those presented in [7]. This is due to differences in the way the power spectral density was estimated compared to this analysis.
Observations of neutron stars and stellar-mass BHs (c.f. Ref. [42] and references therein) hint at a putative mass gap between highest neutron star masses ($m \lesssim 3M_\odot$) and the lowest BH masses ($m \gtrsim 5M_\odot$). Current measurements, however, are not conclusive since the lack of BHs at lower masses could be entirely due to selection effects [60]. The confirmation or exclusion of the mass gap by GW observations is expected to provide unique insights on stellar collapse and compact-object formation [61]. Figure 2 shows that, when analyzed with priors motivated by stellar physics like Kroupa’s initial mass function, GW data for GW151226 are fully consistent with the existence of a mass gap. Both $P_6$ and $P_7$ are slightly favored over the default prior, with $O \approx 2.5$. A careful consideration of priors may thus be important to securely discriminate between BHs and neutron stars [62–67] and thus establish the presence of the mass gap between the two classes of sources.

Finally, we have verified that the marginalized chirp mass posteriors are stable over the change of priors. For GW150914, all cases except $P_5$ (which, however, presents lower odds) yield posterior medians within a $\sim 0.5M_\odot$ interval. The median for $P_3$ is $1M_\odot$ larger than that of $P_1$ (for comparison the 90% credible interval for the $P_1$ run is $\sim 3M_\odot$). For GW151226, all runs yield posterior medians within a $\sim 0.03M_\odot$ interval, compared with a 90% credible interval of $\sim 0.6M_\odot$ for the $P_1$ run.

Conclusions and future prospects—In this Letter we have shown how different prior choices can influence the statistical inference on the parameters of binary BHs. We have estimated the parameters of the first GW events detected under different prior assumptions and verified that both the component masses and the effective spin are impacted by the choice of the prior. For example, for GW150914 and the software replica of GW170104 the effective spin can be made close to zero using a prior that prefers small spin magnitudes, without significant loss of evidence. Conversely, for GW151226 all priors we used give a posterior density for $\chi_{\text{eff}}$ which is positive at the 99% credible interval. Using the default priors for GW151226 results in a source-frame mass posterior for the lighter BH which has support in the suggested mass gap between BHs and neutron stars. We have shown that this conclusion does not hold if a different, yet reasonable, prior on the masses (i.e. a power law modeled on the initial stellar mass function with and without a logistic function on the mass ratio) is used.

As exemplified by the case of LVT151012, the effect of different prior choices is more severe for weak GW signals. Data will be more informative for future loud events and, eventually, more and more physical conclusions will become robust with respect to the details of the prior choice. This point will be specifically addressed in
future work, together with a wider variety of prior distributions more carefully modeled on BH formation pathways, along with the effect of the PN spin evolution [68]. Other investigations of prior effects in GW data analysis are also underway [69].

We hope our work may spark new efforts at incorporating a range of prior choices into inferences about the progenitors of binary BHs. Instead of attempting to explain constraints obtained from a single prior choice, we argue different astrophysical models should be used as priors in the data analysis process. Model selection should then be applied to assess which model better explains the observations. For instance, posteriors calculated under different priors naturally enter in hierarchical analyses [29, 30, 33, 70–72], or when one wants to calculate the branching ratio between populations [27]. More simply, one can just use the evidences calculated under different priors naturally enter in hierarchical analyses [29, 30, 33, 70–72], or when one wants to calculate the odds ratios between models. In this scenario, a cumulative odds given multiple detections can be easily produced (e.g. Ref. [73]).

Ultimately, statistical inference consists of the continual update of one’s current (prior) knowledge in the face of new observational evidence. The dependence of one’s beliefs on prior knowledge should be viewed as a strength rather than a weakness as we approach new observations, since priors allow us to account for all past evidence on the subject.

Acknowledgments—We thank Thomas Dent, Jim Fuller, Chris Pankow, Harald Pfeiffer, Vivien Raymond and Leo Stein for various stimulating discussions. We thank Christopher Berry, James Clark, Will Farr, Heather Fong, Prayush Kumar, Richard O’Shaughnessy, Harald Pfeiffer, and John Veitch for discussions on the use of the volumetric prior and making their implementation public. We thank Lisa Barsotti for providing a simulated noise spectral density for the analysis of the GW170104 software replica. S.V. acknowledges support of MIT through the Solomon Buchsbaum Research Fund, the National Science Foundation, and the LIGO Laboratory. LIGO was constructed by the California Institute of Technology and Massachusetts Institute of Technology with funding from the National Science Foundation and operates under cooperative agreement PHY-0757058. D.G. is supported by NASA through Einstein Postdoctoral Fellowship Grant No. PF6-170152 awarded by the Chandra X-ray Center, which is operated by the Smithsonian Astrophysical Observatory for NASA under Contract NAS8-03060. The authors acknowledge the LIGO Data Grid clusters. This project started during the workshop “Strong Gravity and Binary Dynamics with Gravitational Wave Observations”, funded through NSF CAREER Grant No. PHY-1055103. This paper carries LIGO Document Number P1700176.

[1] T. Bayes, Philosophical Transactions of the Royal Society of London Series 1 53, 370 (1763).
[2] E. T. Jaynes, Probability Theory: The Logic of Science (CUP, 2003).
[3] G. D’Agostini, (1995), hep-ph/9512295.
[4] B. P. Abbott et al. (LIGO and Virgo Scientific Collaboration), PRL 116, 061102 (2016), arXiv:1602.03837 [gr-qc].
[5] B. P. Abbott et al. (LIGO and Virgo Scientific Collaboration), PRL 116, 241103 (2016), arXiv:1606.04855 [gr-qc].
[6] B. P. Abbott et al. (LIGO and Virgo Scientific Collaboration), PRL 118, 221101 (2017), arXiv:1706.01812 [gr-qc].
[7] B. P. Abbott et al. (LIGO and Virgo Scientific Collaboration), PRX 6, 041015 (2016), arXiv:1606.04856 [gr-qc].
[8] J. Veitch, V. Raymond, B. Farr, W. Farr, P. Graff, S. Vitale, et al., PRD 91, 042003 (2015), arXiv:1409.7215 [gr-qc].
[9] B. P. Abbott et al. (LIGO and Virgo Scientific Collaboration), PRL 116, 241102 (2016), arXiv:1602.03840 [gr-qc].
[10] B. P. Abbott et al. (LIGO and Virgo Scientific Collaboration), ApJ 833, L1 (2016), arXiv:1602.03842 [astro-ph.HE].
[11] B. P. Abbott et al. (LIGO and Virgo Scientific Collaboration), PRL 116, 221101 (2016), arXiv:1603.03841 [gr-qc].
[12] P. B. Graff, A. Buonanno, and B. S. Sathyaprakash, PRD 92, 022002 (2015), arXiv:1504.04766 [gr-qc].
[13] J. Veitch, M. Purrer, and I. Mandel, PRL 115, 141101 (2015), arXiv:1503.05953 [astro-ph.HE].
[14] C.-J. Haster, Z. Wang, C. P. L. Berry, S. Stevenson, J. Veitch, and I. Mandel, MNRAS 457, 4499 (2016), arXiv:1511.01431 [astro-ph.HE].
[15] C. Cutler and É. E. Flanagan, PRD 49, 2565 (1994), gr-qc/9402014.
[16] S. Vitale, R. Lynch, V. Raymond, R. Sturani, J. Veitch, and P. Graff, PRD 95, 064053 (2017), arXiv:1611.01122 [gr-qc].
[17] S. Vitale, R. Lynch, J. Veitch, V. Raymond, and R. Sturani, PRL 112, 251101 (2014), arXiv:1403.0129 [gr-qc].
[18] H. Fong, H. Pfeiffer, P. Kumar, and W. Farr, private communication.
[19] J. Veitch and R. O’Shaughnessy, private communication.
[20] K. Belczynski, D. E. Holz, T. Bulik, and R. O’Shaughnessy, Nature 534, 512 (2016), arXiv:1602.04531 [astro-ph.HE].
[21] C. L. Rodriguez, C.-J. Haster, S. Chatterjee, V. Kalogera, and F. A. Rasio, ApJ 824, L8 (2016), arXiv:1604.04254 [astro-ph.HE].
[22] D. Gerosa, M. Kesden, E. Berti, R. O’Shaughnessy, and U. Sperhake, PRD 87, 104028 (2013), arXiv:1302.4442 [gr-qc].
[23] R. O’Shaughnessy, D. Gerosa, and D. Wysocki, PRL 119, 011101 (2017), arXiv:1704.03879 [astro-ph.HE].
[24] C. L. Rodriguez, M. Zevin, C. Pankow, V. Kalogera, and F. A. Rasio, ApJ 832, L2 (2016), arXiv:1609.05916 [astro-ph.HE].
[25] D. Kushnir, M. Zaldarriaga, J. A. Kollmeier, and R. Waldman, MNRAS 462, 844 (2016), arXiv:1605.05683 [astro-ph.HE].
[26] A. Nishizawa, A. Sesana, E. Berti, and A. Klein, MN-
