

Characterization of merging black holes with two precessing spins

Renzis, Viola De; Gerosa, Davide; Pratten, Geraint; Schmidt, Patricia; Mould, Matthew

DOI:

[10.1103/PhysRevD.106.084040](https://doi.org/10.1103/PhysRevD.106.084040)

License:

Other (please specify with Rights Statement)

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Renzis, VD, Gerosa, D, Pratten, G, Schmidt, P & Mould, M 2022, 'Characterization of merging black holes with two precessing spins', *Physical Review D*, vol. 106, 084040. <https://doi.org/10.1103/PhysRevD.106.084040>

[Link to publication on Research at Birmingham portal](#)

Publisher Rights Statement:

This is the final published version of record of the following article: Viola De Renzis, Davide Gerosa, Geraint Pratten, Patricia Schmidt, and Matthew Mould, 'Characterization of merging black holes with two precessing spins', *Phys. Rev. D* 106, 084040, Published 21 October 2022 by American Physical Society. The published version is available at <https://doi.org/10.1103/PhysRevD.106.084040>

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.
- User 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.

Characterization of merging black holes with two precessing spins

Viola De Renzi^{1,2,*} Davide Gerosa^{1,2,3} Geraint Pratten,³ Patricia Schmidt³ and Matthew Mould³

¹*Dipartimento di Fisica “G. Occhialini,” Università degli Studi di Milano-Bicocca,
Piazza della Scienza 3, 20126 Milano, Italy*

²*INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, 20126 Milano, Italy*

³*School of Physics and Astronomy & Institute for Gravitational Wave Astronomy,
University of Birmingham, Birmingham B15 2TT, United Kingdom*



(Received 30 June 2022; accepted 28 September 2022; published 21 October 2022)

Spin precession in merging black-hole binaries is a treasure trove for both astrophysics and fundamental physics. There are now well-established strategies to infer from gravitational-wave data whether at least one of the two black holes is precessing. In this paper we tackle the next-in-line target, namely the statistical assessment that the observed system has two precessing spins. We find that the recently developed generalization of the effective precession spin parameter χ_p is a well-suited estimator to this task. With this estimator, the occurrence of two precessing spins is a necessary (though not sufficient) condition to obtain values $1 < \chi_p \leq 2$. Confident measurements of gravitational-wave sources with χ_p values in this range can be taken as a conservative assessment that the binary presents two precessing spins. We investigate this argument using a large set of > 100 software injections assuming anticipated LIGO/Virgo sensitivities for the upcoming fourth observing run, O4. Our results are very encouraging, suggesting that, if such binaries exist in nature and merge at a sufficient rate, current interferometers are likely to deliver the first confident detection of merging black holes with two precessing spins. We investigate prior effects and waveform systematics and, though these need to be better investigated, did not find any confident false-positive case among all the configurations we tested. Our assessment should thus be taken as conservative.

DOI: [10.1103/PhysRevD.106.084040](https://doi.org/10.1103/PhysRevD.106.084040)

I. INTRODUCTION

Black-hole (BH) binary spin precession is a key feature of the relativistic two-body problem [1,2]. Spin-spin and spin-orbit couplings in general relativity cause the orbital angular momentum \mathbf{L} and the BH spins $\mathbf{S}_{1,2}$ to jointly precess about the direction of the total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}_1 + \mathbf{S}_2$. This motion induces modulations to both the amplitude and the phase of the emitted gravitational waves (GWs).

Measurements of spin precession have important repercussions in both astrophysics and fundamental physics. For the stellar-mass BH binaries observed by LIGO and Virgo [3,4], spin precession provides unique leverage to discriminate between BH binaries formed in isolation and those assembled in dynamically in stellar clusters [5–8]. For the supermassive BH binaries targeted by LISA [9], spin measurements will provide information on, e.g., the occurrence of prolonged phases of disk accretion [10,11]. GW observations of precessing binary BHs also allow us to constrain modified theories of gravity, especially those with parity-violating interactions caused by additional fields [12].

While the masses of LIGO/Virgo events are usually well measured, spin effects provide a subdominant contribution to the emitted radiation and are thus considerably more challenging to characterize. At present, an unambiguous measurement of BH-binary spin precession is one of the holy grails of observational GW astronomy.

Data from the first three observing runs of the LIGO/Virgo network have provided some evidence for individual BH mergers with highly precessing spins [13–16]. The most suggestive indication is that from GW200129_065458, where Refs. [17,18] found strong evidence for BH-binary spin precession, while Ref. [19] raised potential issues in the glitch mitigation analysis. For the case of GW190521, a strong precession signature was also reported [14], though potential degeneracies with the eccentricity still need to be fully understood [20]. Collective evidence for spin precession was reported in the context of BH binary populations, with all current fits requiring some misaligned spins at high confidence [21–23].

Upcoming instrumental upgrades to the LIGO/Virgo (and hopefully KAGRA) network [24] are posed to provide increasingly accurate spin measurements. It is therefore not unreasonable to predict that the next observing run will deliver a confident, unambiguous identification of BH-binary spin precession. Crucially, measuring orbital-plane

*v.derenzis@campus.unimib.it

precession corresponds to inferring that at least one of the two BHs has a misaligned spin [1]. Inferring the presence of two misaligned spins requires extracting even feebler signatures from the signal, which are related to spin-spin (as opposed to spin-orbit) terms in the BH binary equations of motion.

This paper tackles such a next-in-line target. We perform > 100 software injections with realistic LIGO/Virgo sensitivity and demonstrate that signals with large-but-not-extreme signal-to-noise ratio (SNR) $\gtrsim 20$ allow us to detect two-spin effects *already in the next LIGO/Virgo observing run* (O4). Of course, this statement relies on the assumption that merging binaries with two large precessing spins exist and can merge efficiently. But if such GW sources are out there in the Universe, the next LIGO/Virgo run might provide the first observational constraints of their properties.

Compared to previous analyses which include two precessing spins (e.g., [25–27]) our investigation relies on a state-of-the-art reformulation of the precession estimator χ_p [28]. This generalizes the commonly used expression [29] by employing a rigorous post-Newtonian (PN) average over the joint evolution of *both* spins. Measurements of such an augmented χ_p for current GW events have been presented in Refs. [28,30]. Crucially for this paper, the precession-averaged estimator presents an exclusion region $1 < \chi_p \leq 2$ that can *only* be populated by binaries with two precessing spins. Measuring a binary with $\chi_p > 1$ at some large confidence (GW astronomers often use the 90% Bayesian credible interval), would allow us to claim the first detection of binary BH physics sourced by two precessing spins.

Our paper is organized as follows. In Sec. II we present our methodology, including details on the adopted precession estimator as well as the implemented parameter-estimation pipeline. In Sec. III we present the results of our analysis. In particular, we characterize (i) the SNR dependence on the resulting χ_p posterior distributions, (ii) the statistical behavior of large ensembles of sources, (iii) the impact of the prior, and (iv) the relevance of waveform systematics. Our conclusions are reported in Sec. IV. In the following we employ geometric units $G = c = 1$.

II. METHODS

A. Spin precession estimators

While the full BH-binary spin properties are in principle described by six degrees of freedom (three components for two spin vectors), a considerable amount of effort has been devoted to identifying a reduced number of parameters that encapsulate most of the information. These are often derived in a PN framework, with the most widely used quantities being the effective aligned spin χ_{eff} [31–33] and the effective precessing spin χ_p [28,29]. Alternative approaches include extending the precession estimator

to a two-dimensional vector [34], exploiting the precession/nutation amplitudes and frequencies [35,36], and computing the fraction of the SNR contained in the spin modulations [37].

Let us consider a BH binary where $q = m_2/m_1 \leq 1$ is the mass ratio, $\chi_i \in [0, 1]$ are the dimensionless spin magnitudes, θ_i are the angles between the spins and the orbital angular momentum, and $\Delta\Phi$ is the angle between the projections of the two spins onto the orbital plane.

The effective aligned spin is defined as [31]

$$\chi_{\text{eff}} = \frac{\chi_1 \cos \theta_1 + q\chi_2 \cos \theta_2}{1 + q}. \quad (1)$$

This is the spin quantity that affects the GW phase at lowest order and is a constant of motion at 2PN [32,33]. The effective spin χ_{eff} was recognized as the best measured spin parameter since the very first GW detections, the key reason being that it directly impacts the length of the signal.

The spin-precession parameter $\chi_p \propto |d\hat{\mathbf{L}}/dt|$ tracks the change of the direction of the orbital angular momentum \mathbf{L} over time t [28,29]. It was originally introduced by Schmidt *et al.* [29] as a building block toward the construction of precessing waveforms. Their definition reads

$$\chi_p^{(\text{heu})} = \max \left(\chi_1 \sin \theta_1, q \frac{4q+3}{4+3q} \chi_2 \sin \theta_2 \right), \quad (2)$$

which in this paper we refer to as “heuristic χ_p .” This precessing spin parameter is defined in the domain $\chi_p^{(\text{heu})} \in [0, 1]$. Unlike χ_{eff} , the parameter χ_p depends on the projections of the spins onto the orbital plane, $\chi_i \sin \theta_i$, implying that a confident measurement of $\chi_p > 0$ requires that at least one of the two BH spins was misaligned before merger, and hence that the system was precessing.

Gerosa *et al.* [28] recently pointed out that Eq. (2) was derived by preferentially selecting some terms when averaging over the spin motion. Mathematically, this is reflected in the maximization operation reported in Eq. (2), which selects one of the two BHs as dominant to the precession dynamics, thus obfuscating two-spin effects. Relaxing this approximation yields a generalized parameter [28]

$$\chi_p^{(\text{gen})} = \left[(\chi_1 \sin \theta_1)^2 + \left(q \frac{4q+3}{4+3q} \chi_2 \sin \theta_2 \right)^2 + 2q \frac{4q+3}{4+3q} \chi_1 \chi_2 \sin \theta_1 \sin \theta_2 \cos \Delta\Phi \right]^{\frac{1}{2}}, \quad (3)$$

where the angles $\theta_1(t)$, $\theta_2(t)$, and $\Delta\Phi(t)$ all vary jointly with time. This can be averaged over a single precession cycle to obtain

$$\chi_p^{(\text{av})} = \frac{1}{\tau} \int_0^\tau \chi_p^{(\text{gen})}(t) dt, \quad (4)$$

where τ is the precession period. We argue this should be regarded as a more solid estimator because, although it is not a constant of motion like χ_{eff} , it at most varies only over the longer radiation-reaction timescale. In the following, we refer to Eq. (4) as the ‘‘averaged χ_p ’’ parameter. In practice, we perform the integral in Eq. (4) using a 2PN quasiadiabatic approach where the precession cycle is parametrized by $S(t) = |\mathbf{S}_1(t) + \mathbf{S}_2(t)|$ [33,38]. We refer the reader to Ref. [33] for details on the derivation of Eqs. (3) and (4), but stress that the starting point is simply the derivative $d\hat{\mathbf{L}}/dt$.

The reformulation of the precession parameter defines an extended range $\chi_p^{(\text{av})} \in [0, 2]$. As shown in Ref. [28], the heuristic and averaged definitions of χ_p have the same single-spin limit, which implies that the range of the latter cannot be freely absorbed with a normalization factor. From Eq. (3), it is immediate to see that $\chi_p > 1$ requires both $\chi_1 \sin \theta_1 \neq 0$ and $\chi_2 \sin \theta_2 \neq 0$, i.e., the binary must have two precessing spins. Such sources *can* lie in $0 \leq \chi_p \leq 1$, but both spins being misaligned is requisite in the two-spin domain. From Eq. (3), there is a larger volume of parameter space where $\chi_p > 1$ for comparable mass binaries $q \lesssim 1$ compared to asymmetric sources with $q \ll 1$. This is expected, as two-spin effects are highly suppressed in the low-mass ratio limit where $S_2/S_1 \propto q^2 \ll 1$ (cf. Ref. [39] for more work on spin precession in asymmetric binaries).

B. Parameter estimation pipeline

As is common practice in GW parameter estimation, we employ the following fifteen parameters to describe compact-binary coalescences: detector-frame total mass $M = m_1 + m_2$, mass ratio $q = m_2/m_1$, dimensionless spin magnitudes $\chi_{1,2}$, tilt angles $\theta_{1,2}$, azimuthal spin angle $\Delta\Phi$, azimuthal angle ϕ_{JL} between the total and orbital angular momenta, luminosity distance D_L , right ascension α , declination δ , polar angle θ_{JN} between total angular momentum and the line of sight, polarization ψ , time t_c and phase ϕ_c of coalescence [40].

We explore the joint Bayesian posterior distribution of these parameters under a Gaussian noise likelihood (e.g., [41]) using the parallelized BILBY pipeline [42,43] and its underlying DYNesty implementation of nested sampling [44]. Our runs make use of 2048 live points, a number of autocorrelation equal to 50 a random walk sampling method, and a likelihood marginalized over time and distance. Runs are halted when the log-evidence gain falls below 0.1.

We consider a three-detector network consisting of LIGO Livingston, LIGO Handford, and Virgo with their projected sensitivities for the upcoming fourth observing

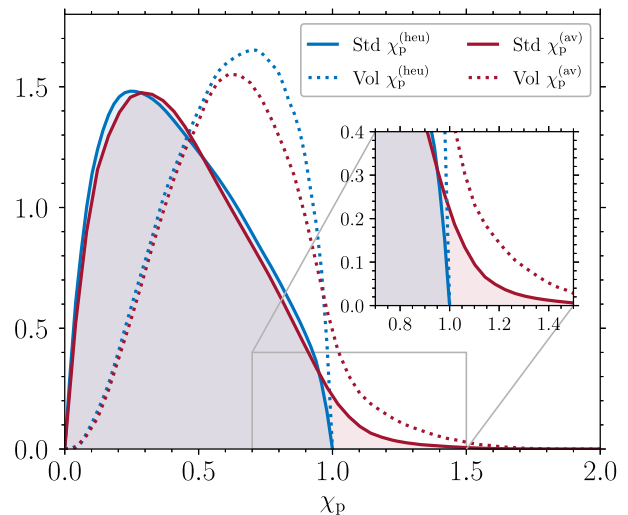


FIG. 1. Prior distributions for the heuristic (blue) and averaged (red) χ_p estimators. Dotted curves are computed assuming the standard spin prior $p(\chi) = \text{const}$; dashed curves instead assume a volumetric prior $p(\chi) \propto \chi^2$.

run O4 [24]. We consider data segments of 4 s, set a lower frequency cutoff of 20 Hz, assume a sampling frequency of 2048 Hz, and zero noise. Time-varying quantities are quoted when the detector-frame emission frequency of the dominant mode is 20 Hz. This choice has a negligible impact on the averaged χ_p estimator because it only varies on the long inspiral timescale of the binary evolution; see Ref. [28]. Unless stated otherwise, we quote our results using medians and 90% equal-tailed credible intervals.

We adopt uninformative priors as commonly used in current LIGO/Virgo analyses [13–16]. Specifically, priors on the masses are chosen to be uniform in $m_{1,2} \in [5, 100] M_\odot$, with further constraints imposed on the mass ratio $q \in [1/8, 1]$ and detector-frame chirp mass $M_c \in [10, 60] M_\odot$. For most of our runs, priors on the spins are taken to be uniform in magnitude $\chi_{1,2} \in [0, 0.99]$ and isotropic in directions. In the following, we will refer to this as our ‘‘standard’’ spin prior. To better explore prior effects, some of our runs are performed with a ‘‘volumetric’’ spin prior $p(\chi_i) \propto \chi_i^2$, corresponding to spin vectors that are uniformly drawn in volume (e.g., [13,45,46]). The luminosity distance prior is taken to be uniform in comoving volume with $D_L \in [100, 5000]$ Mpc.

Figure 1 shows the resulting prior probability density for the heuristic and the averaged χ_p definitions. At low values of χ_p , the prior distributions of the two estimators are qualitatively very similar. This behavior was explicitly imposed in Ref. [28] when generalizing the χ_p definition. By construction, the region of $1 < \chi_p \leq 2$ is not allowed for the heuristic formulation and, consequently, the prior distribution of $\chi_p^{(\text{heu})}$ is steeply truncated at $\chi_p = 1$. On the other hand, the prior distribution of $\chi_p^{(\text{av})}$ extends into the

two-spin region $\chi_p > 1$. However, under these commonly used assumptions, the tail at large χ_p values is very sparsely populated. From Eq. (3), reaching $\chi_p \approx 2$ requires systems with $q \approx 1$, $\chi_{1,2} \approx 1$, $\theta_{1,2} \approx \pi/2$, and $\Delta\Phi \approx 0$. Such a strong prior suppression is a key element of our analysis and suggests that current GW data are being analyzed with a prior that strongly disfavors the region of parameter space that is exclusive to two-spin physics. Although still present, this effects is less prominent for the volumetric spin prior. More quantitatively, we find $p(\chi_p > 1) = 0.02$ and 0.07 for the averaged estimator under the standard and volumetric prior, respectively.

For the majority of our runs, we employ the IMRPhenomXPHM [47–49] waveform model for both injection and recovery. This is a state-of-the-art frequency-domain approximant that captures spin precession without relying on a single-spin approximation. Selecting the same model for both injection and recovery allows us to first isolate statistical effects without systematics. Waveform systematics are then explored with a dedicated analysis where we select different models for injection and recovery. In particular, we use the time-domain model IMRPhenomTPHM [50] as well as the numerical-relativity surrogate NRSur7dq4 [51]. When recovering with NRSur7dq4, we restrict our prior to $q > 1/6$, which corresponds to the

extended range of validity of the model. Because this approximant only covers ~ 20 orbits before merger, we also restrict our priors to $m_{1,2} \in [35, 150] M_\odot$ and $M_c \in [40, 60] M_\odot$ to ensure the signal is fully above the low-frequency cutoff of 20 Hz.

III. RESULTS

A. Single-system series

As a first step, we highlight the main implications of the χ_p reformulation on GW parameter estimation. To this end, we present a series of six software injections where the same binary is observed at different SNRs. We select a source with $\chi_p^{(\text{heu})} = 0.67$ and $\chi_p^{(\text{avu})} = 1.22 > 1$, which thus contains two prominently precessing spins. In particular, the injected system has $M = 54.1 M_\odot$, $q = 0.96$, $\chi_1 = 0.56$, $\chi_2 = 0.7$, $\theta_{1,2} = \pi/2$, $\Delta\Phi = 0.1$, $\theta_{JN} = 1.0$, $\phi_{JL} = 1.0$, $\alpha = 0.75$, $\delta = 0.5$, $\psi = 1.0$, $\phi_c = \pi/4$, $t_c = 0.0$. We select increasing values of the luminosity distance $D_L = 200, 500, 700, 900, 1300, 1700$ Mpc while keeping the detector-frame mass M fixed. The corresponding three-detector network SNRs are $\rho = 124, 44, 35, 27, 19$, and 14 . We use the IMRPhenomXPHM waveform model for both injection and recovery and employ standard uninformative priors.

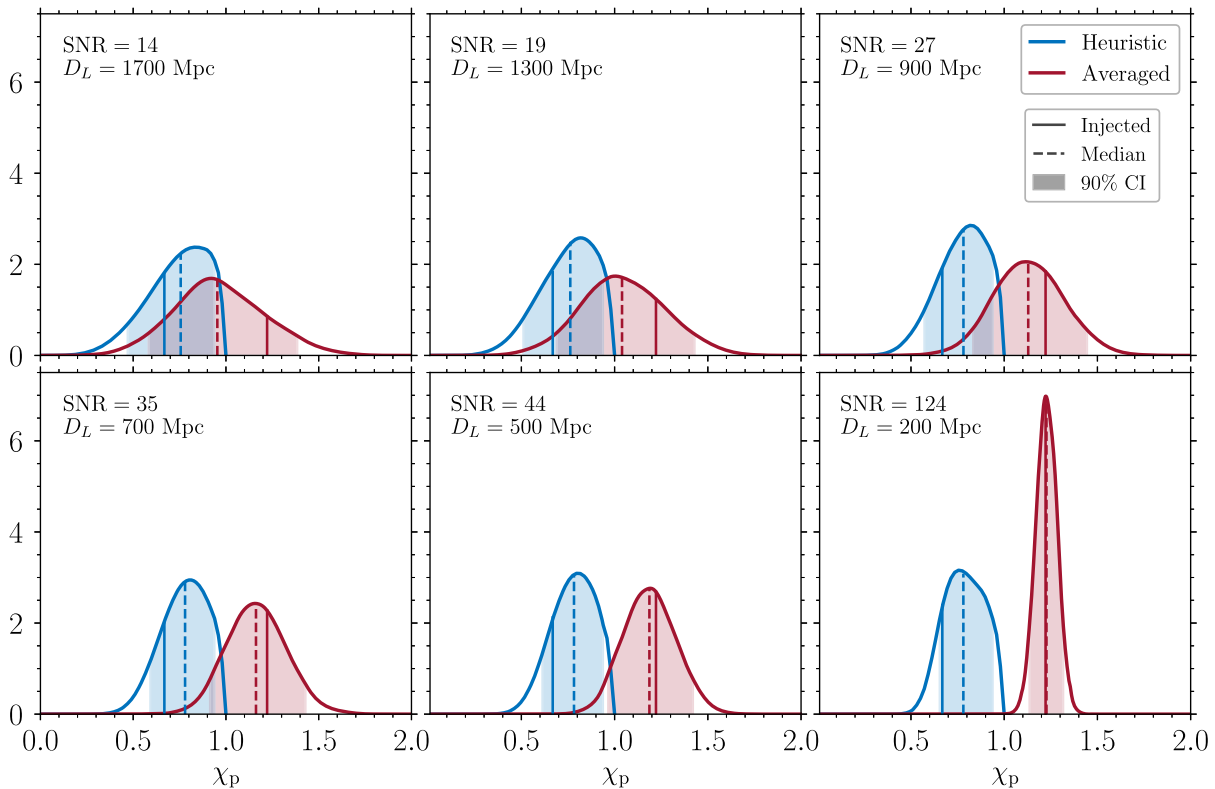


FIG. 2. Posterior distribution of the heuristic (blue) and averaged (red) χ_p precession estimator for the single-system series described in Sec. III A. Panels from left to right and top to bottom shows results for the same source injected at increasing SNRs and decreasing luminosity distance D_L . Solid and dashed lines indicate the true value and the median of the recovered posterior. The shaded areas indicate the 90% CI (credible intervals).

Our results are illustrated in Fig. 2. As one moves from the lowest to the highest value of the SNR, the recovered posteriors of both the averaged and the heuristic χ_p converge to the injected values. This is expected because we have used the same signal model for injection and recovery and we are not considering a specific noise realization.

For the system with the lowest SNR $\rho = 14$, the posteriors of our two χ_p definitions largely overlap. Quoting median and 90% credible interval, we find $\chi_p^{(av)} = 0.95^{+0.43}_{-0.37}$, which implies that we cannot confidently tell that the source has two misaligned spins. As the SNR increases, so does our ability to infer that the binary has two precessing spins. For the system with the largest SNR $\rho = 124$, the two marginalized χ_p distributions are almost completely detached. The posterior of the heuristic χ_p is by definition truncated at $\chi_p = 1$ because Eq. (2) allows only for the contribution from a single, dominant spin. On the contrary, considering our averaged definition yields $\chi_p = 1.22^{+0.09}_{-0.09}$ for $\rho = 124$, implying one infers the presence of two precessing spins with a credibility of $p(\chi_p > 1) = 99.9\%$.

Figure 2 also shows that the posterior of the averaged χ_p is closer to a Gaussian compared to that of the heuristic estimator. This indicates that, if a significant non-Gaussianity in the heuristic χ_p posterior were to appear in GW data, it could be taken as a potential indication that some additional two-spin physics is present but is being missed because of the suboptimality of the employed estimator.

This argument is further explored in Fig. 3, where we show the skewness and kurtosis for the same six injections presented in Fig. 2. These quantities are related to the third and fourth moments of the distribution and describe the departure from Gaussianity; both are zero for normally distributed data, with the skewness quantifying the left-right asymmetry and the kurtosis quantifying the weight of the tails [52]. Figure 3 shows that both skewness and kurtosis of the averaged χ_p are approximately distributed around 0. On the other hand, the skewness (kurtosis) of the heuristic χ_p strongly increases (decreases) with the SNR. This indicates that (i) the heuristic χ_p posteriors have a thinner tails compared a normal distribution and that (ii) their left tail is more pronounced compared to the right tail. These features can be taken as a quantification of the artificial cutoff at $\chi_p = 1$, an assumption that is naturally relaxed when considering the averaged χ_p estimator.

B. Parameter-space exploration

Using the same settings, we now target the statistical properties emerging from a large number of injected signals. Ideally, one would want to inject signals drawn from the prior (this is necessary, for instance, to present a probability-probability plot [53]). In our case, such a

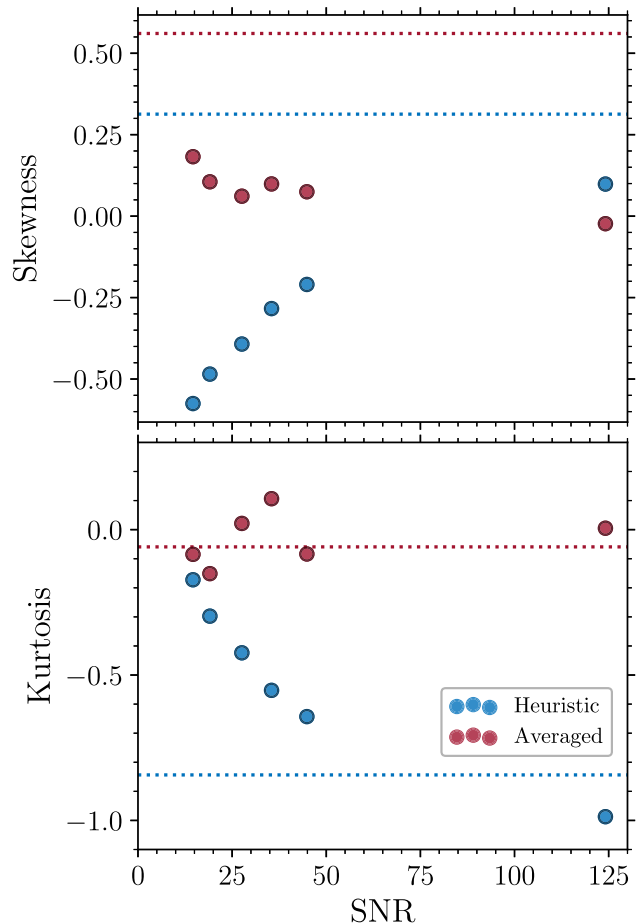


FIG. 3. Skewness (upper panel) and kurtosis (lower panel) of the χ_p posterior distribution as a function of the SNR for the six injections described in Sec. III A. Blue and red scatter points refer to the heuristic and averaged χ_p definition, respectively. The dotted lines represent the values of the skewness and kurtosis calculated from the priors.

procedure would be highly suboptimal and ultimately computationally intractable because, as shown in Fig. 1, the two-spin region with $\chi_p > 1$ corresponds to a very low prior volume (where from now on in the paper we only refer to the average formulation of χ_p). Most of the injections would thus be placed in the region where only one of the two spins dominates. We thus opt for an injection distribution with χ_p uniform in $[0, 2]$ which, although of dubious astrophysical relevance, is well suited to assess the statistical property of the proposed estimator. More precisely, we draw values of χ_p and then reweight samples of the intrinsic binary properties drawn from the uninformative prior (Sec. II B) to the injection distribution using an acceptance/rejection scheme with an absolute numerical tolerance of 0.04 between the original and resampled values of χ_p . We have verified that this choice does not significantly impact our results. Since precession effects are subdominant in the waveform, when selecting the extrinsic

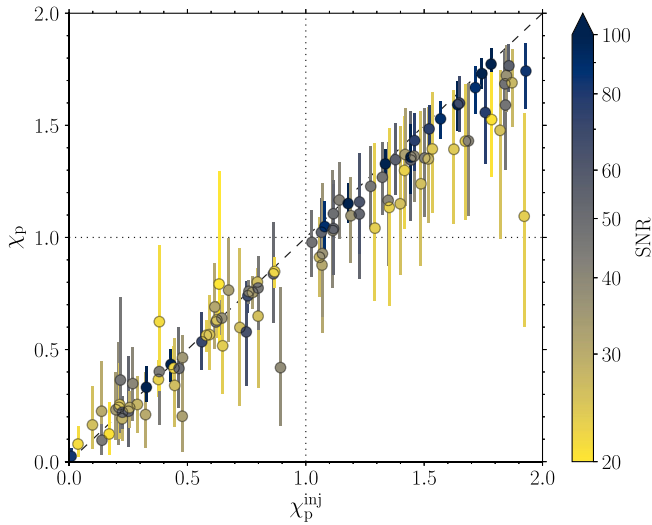


FIG. 4. Set of 100 injections obtained by reweighting the averaged χ_p prior toward a uniform distribution in $[0, 2]$. The medians (scatter points) and symmetric 90% credible intervals (error bars) of the recovered posteriors are plotted against the true values χ_p^{inj} . Vertical and horizontal dotted lines indicate $\chi_p = 1$ while the dashed diagonal line corresponds to $\chi_p = \chi_p^{\text{inj}}$, i.e., successful recovery. The three-detector SNRs of the injected sources are reported on the color scale.

properties for injections we only consider sources with $\rho > 20$, i.e., ≈ 2 times larger than the current detection threshold [16].

Figure 4 shows the recovered posteriors of the averaged χ_p parameter as a function of the true values for 100 such injections. One can divide the parameter space into four distinct regions, acting much like a confusion matrix in statistics.

- (i) *True negatives* (bottom-left quadrant in Fig. 4): injected $\chi_p < 1$ and recovered $\chi_p < 1$. The injected configurations are not unique to sources with two precessing spins and are recovered as such.
- (ii) *False positives* (top-left quadrant in Fig. 4): injected $\chi_p < 1$ and recovered $\chi_p > 1$. For these sources, one infers the presence two precessing spins even if they might not be present.
- (iii) *False negatives* (bottom-right quadrant in Fig. 4): injected $\chi_p > 1$ and recovered $\chi_p < 1$. In this region sources have two precessing spins but one is not able infer their occurrence from the signal.
- (iv) *True positives* (top-right quadrant in Fig. 4): injected $\chi_p > 1$ and recovered $\chi_p > 1$. These sources are characterized by two precessing spins and one can successfully infer that this is the case.

For each posterior distribution, we compute the fraction of the samples in each of these four regions and then compute the arithmetic mean over the injected sample (this is equivalent to assuming a flat population prior on χ_p because our injections are distributed uniformly). We report

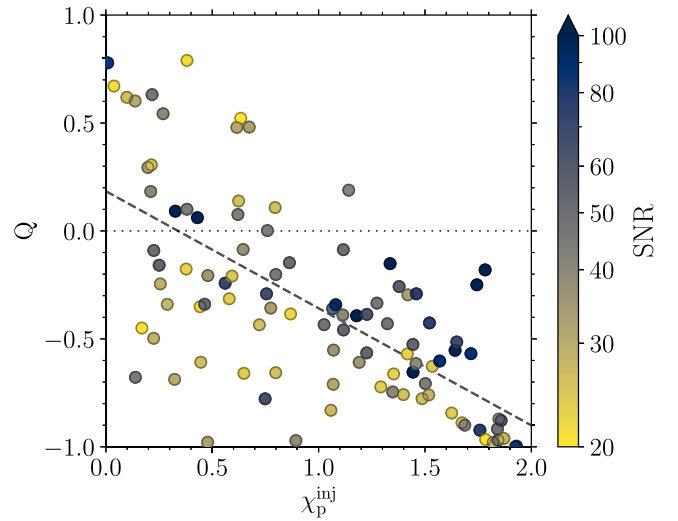


FIG. 5. Adjusted posterior quantile Q for the posterior distribution of the averaged χ_p parameter. Sources above (below) the horizontal dashed line indicate cases where χ_p is overestimated (underestimated). To guide the eye, the diagonal dashed line shows a simple linear fit $Q = -0.54\chi_p^{\text{inj}} + 0.18$. The color scale indicates the SNRs of the sources.

47.45% of true negatives, 0.55% of false positives, 7.01% of false negatives, and 44.98% of true positives.

From Fig. 4, the signals with higher SNR lie closer to the injected values and present thinner posterior distributions, as expected. In the true negative region, the recovered posteriors are distributed around the true value without evident systematic trends. On the other hand, when $\chi_p > 1$, the recovered posteriors systematically underestimate the true value. While this is, in general, true for most systems, in a few cases this is sufficient to cause false negatives.

One can further quantify this behavior using the adjusted posterior quantile

$$Q = 2 \int_0^{\chi_p^{\text{inj}}} p(\chi_p) d\chi_p - 1 \in [-1, 1], \quad (5)$$

where $p(\chi_p)$ is the posterior distribution and χ_p^{inj} is the true value. The ideal case where the median of $p(\chi_p)$ coincides with the true value corresponds to $Q = 0$. Obtaining $Q > 0$ ($Q < 0$) instead implies that the amount of precession in the system is being overestimated (underestimated), and $100|Q| < X$ implies that the injected value is inside the $X\%$ symmetric confidence interval of the recovered posterior. The values of Q for our 100 injections are shown in Fig. 5. We find a strong decreasing trend of Q for increasing values of χ_p , which becomes particularly evident in the $\chi_p > 1$ region. Our analysis indicates that, in general, statistical errors cause an underestimate of χ_p whenever $\chi_p > 1$. In other words, given a waveform model, sources with two-spin precession require larger SNR for accurate measurement (cf. [54,55]).

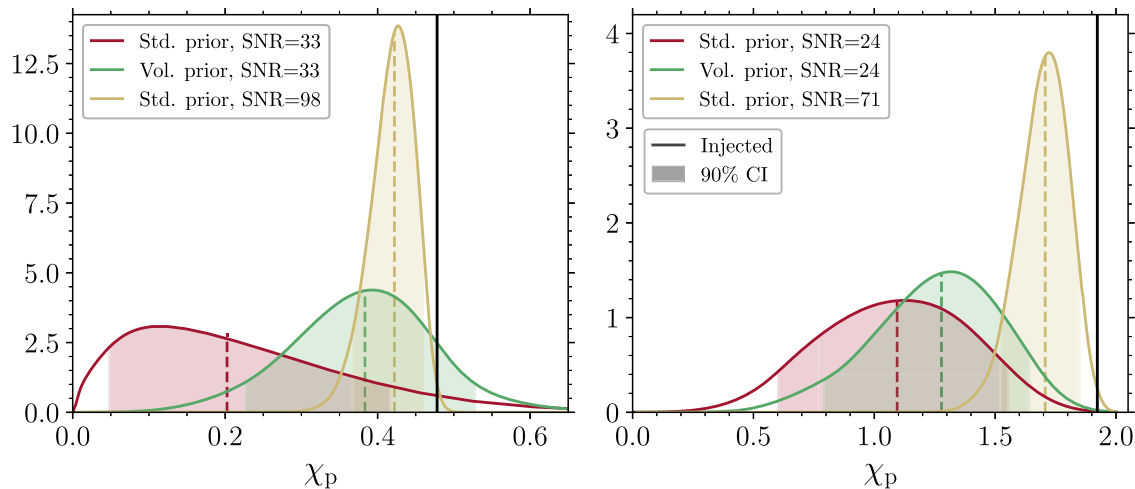


FIG. 6. Recovery of the averaged χ_p estimator with different priors and SNRs. The left and right panel shows results for sources injected with $\chi_p = 0.48$ and $\chi_p = 1.92$, respectively (black vertical lines). The red and green posterior distributions are obtained under the standard and volumetric priors, respectively, and the same SNRs used in Sec. III B. For the yellow distributions, the SNR was boosted by a factor ≈ 3 . Dashed lines indicate the medians of the posteriors while the shaded area indicates the 90% credible interval.

With a completeness of 86.5% and a contamination of 1.2%, our results are, overall, are extremely encouraging.¹ The broader conclusion is that sources with $\chi_p > 1$ and sufficiently high SNR $\rho \gtrsim 20$ in O4 can, in principle, be correctly identified as affected by two precessing spins.

C. Impact of the prior

Our parameter-space exploration highlights a generic tendency to underestimate precession effects whenever $\chi_p > 1$. The steep feature at $\chi_p \approx 1$ shown in Fig. 1 strongly suggests that this statistical bias is driven by the employed prior. To verify this, we select two injections among the 100 we have just presented with posterior quantile $Q \approx -1$, i.e., where the displacement between the injected and recovered values of χ_p is maximized. More specifically, the two systems we consider have $\chi_p \in \{0.48, 1.92\}$, $Q \in \{-0.98, -1\}$, and $\rho \in \{32.8, 23.6\}$, respectively.

In Fig. 6 we compare the posterior distributions obtained under the standard uniform prior as in Sec. III B against additional inference runs where we instead take a volumetric prior on the spins. The latter choice enhances the prior weight assigned to configurations with large spins (cf. Fig. 1). For the injection with $\chi_p = 0.48$ (left panel), we recover $\chi_p = 0.20_{-0.16}^{+0.22}$ with the standard prior and $\chi_p = 0.38_{-0.16}^{+0.14}$ with the volumetric prior. This example shows that simply changing the prior to an alternative that is equally well motivated—why should vectors like the spins be distributed uniformly in magnitude instead of

volume?—can significantly mitigate the inferred bias. In this case, the posterior quantile increases from $Q = -0.98$ to $Q = -0.75$. The improvement is less evident, but still present, for the injection with $\chi_p = 1.92$, where a volumetric prior yields $\chi_p = 1.28_{-0.49}^{+0.37}$ compared to $\chi_p = 1.10_{-0.49}^{+0.46}$ for the standard prior.

Figure 6 also shows additional runs where the same sources are considered at higher SNR, larger by a factor ≈ 3 , using the standard priors. As already shown in Sec. III A, the posterior tends toward the true values for louder sources. Prior effects are still evident, with the true value remaining outside the 90% credible interval. More specifically, for these high-SNR runs we find $\chi_p = 0.42_{-0.05}^{+0.04}$ and $\chi_p = 1.71_{-0.19}^{+0.14}$ for the $\chi_p = 0.48$ and $\chi_p = 1.92$ case, respectively.

D. Waveform systematics

All the analyses illustrated so far were performed using the same waveform model for both injection and recovery and, therefore, do not capture systematic errors due to any mismodeling of the signal. Binaries with prominent spin effects are harder to model, implying that the $\chi_p > 1$ region we are interested in is also where discrepancies between different approximants are more likely to appear.

Figure 7 and Table I illustrate the posterior distribution of the averaged χ_p parameter when different models are used in injection and recovery. We test various combinations of the IMRPhenomXPHM [49], IMRPhenomTPHM [50], and NRSur7dq4 [51] waveform models. We concentrate on two systems selected from the 100 injections presented in Sec. III B. In particular, we consider one source with $\chi_p = 1.57$ characterized by two prominently precessing spins as well as a control case with $\chi_p = 0.43$. The SNRs

¹As common in binary classification [52], we define completeness = true positives/(true positives + false negatives) and contamination = false positives/(true positives + false positives).

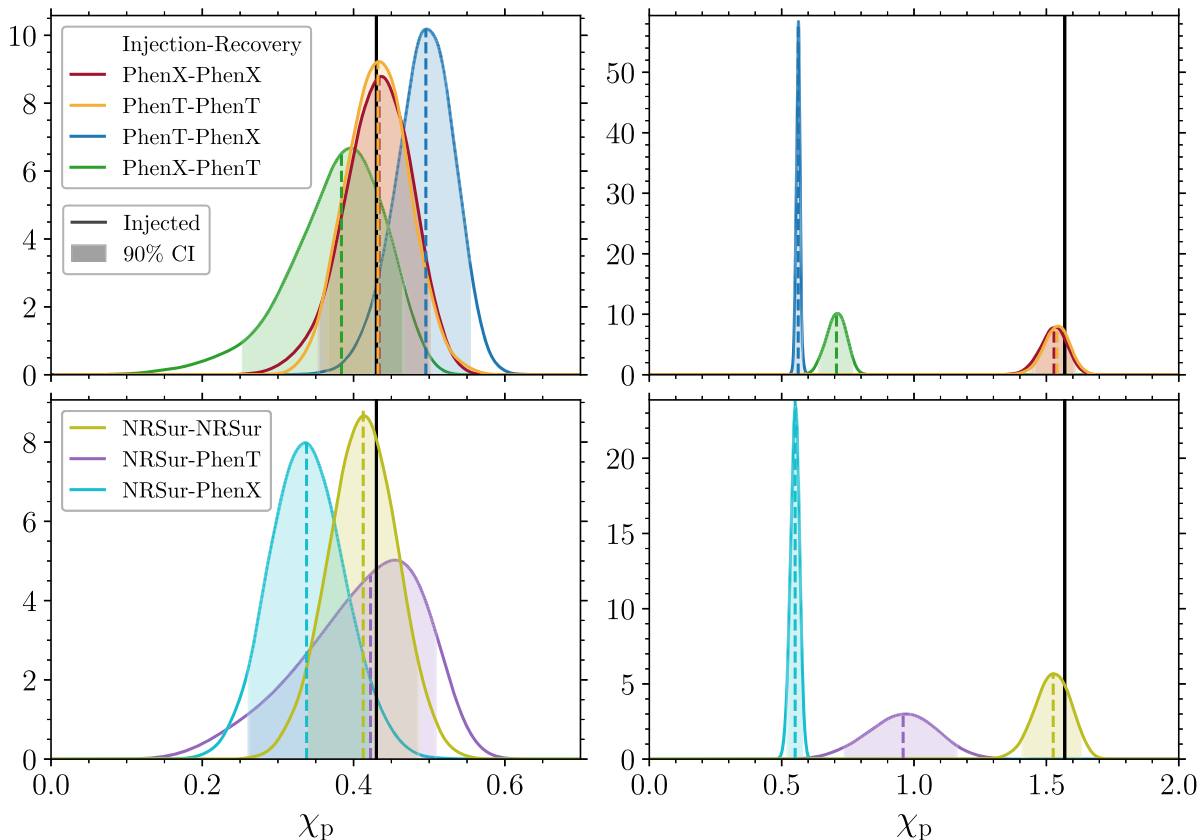


FIG. 7. Posterior distributions of the averaged χ_p parameter obtained with seven different combinations of waveform models. The left (right) panels show results for a source injected with $\chi_p = 0.43$ and $M = 131.1 M_\odot$ ($\chi_p = 1.57$ and $M = 130.8 M_\odot$). The top panels show results obtained with the two phenomenological models IMRPhenomXPHM (“PhenX”) for $\chi_p = 0.43$ ($\rho = 107.3$) and for $\chi_p = 1.57$ ($\rho = 90.3$) and IMRPhenomTPHM (“PhenT”) for $\chi_p = 0.43$ ($\rho = 93.6$) and $\chi_p = 1.57$ ($\rho = 81.7$). The bottom panels show results obtained in combination with the numerical-relativity surrogate model NRSur7dq4 (“NRSur”) for $\chi_p = 0.43$ ($\rho = 100.2$) and for $\chi_p = 1.57$ ($\rho = 75.6$). For each case, the label reported before (after) the hyphen in the legend refers to the waveform model used at the injection (recovery) stage. The injected values are indicated with black vertical lines. Medians and 90% credible interval of the posterior distribution are indicated with dashed lines and shaded areas, respectively.

computed using IMRPhenomXPHM are $\rho = 90.3$ and $\rho = 107.3$, respectively. Both sources have sufficiently large detector-frame total masses $M \gtrsim 125 M_\odot$ such that the signal is short enough to be simulated with NRSur7dq4. The cases where the signal is injected with NRSur7dq4 are arguably more realistic as this model is proven to be more accurate, i.e., closer to numerical-relativity simulations [51], though the model does require extrapolation at the low- q and high- $\chi_{1,2}$ edges of the parameter space we consider. For both the analyzed cases, the posteriors are relatively well centered on the true values whenever the injection and recovery are performed with the same waveform model.

The most evident feature from Fig. 7 is that systematic biases increase dramatically for higher values of χ_p . This statement holds even though our analyzed low- χ_p (high- χ_p) case has a higher (smaller) SNR and should thus be more (less) susceptible to waveform systematics.

For the $\chi_p = 0.43$ source (left panels in Fig. 7), the injected value lies inside the 90% credible interval of the posterior for most of the waveform combinations we tested. The only exception is the case where we inject with the NRSur7dq4 and recover with IMRPhenomXPHM. This run shows the largest quantile $Q = -0.93$ which tentatively suggests a lower accuracy of that model to spin precession, at least for this specific set of parameters. This conclusion is reinforced by our results obtained when the source is generated with IMRPhenomTPHM but recovered with IMRPhenomXPHM: the injected value is barely inside the 90% credible interval with posterior quantile is $Q = -0.88$.

For the second case studied here with $\chi_p = 1.57$ (right panels in Fig. 7), waveform systematics are severe. All waveform combinations where we inject and recover with different models return posterior distributions that are inconsistent with the true value at extremely high confidence (so high that we cannot meaningfully quantify it

TABLE I. Posterior quantiles Q , medians, and 90% credible intervals of the averaged χ_p estimator from analyses performed with three different waveform models: IMRPhenomXPHM (“PhenX”), IMRPhenomTPHM (“PhenT”), and NRSur7dq4 (“NRSur”). The top (bottom) table shows results for injections with $\chi_p = 0.43$ ($\chi_p = 1.57$). In each table, the rows (columns) indicate the waveform used for signal injection (recovery).

$\chi_p^{\text{inj}} = 0.43$	PhenX		PhenT		NRSur	
	Q	χ_p	Q	χ_p	Q	χ_p
PhenX	0.06	$0.43^{+0.07}_{-0.08}$	-0.6	$0.38^{+0.08}_{-0.13}$
PhenT	0.88	$0.5^{+0.06}_{-0.07}$	0.04	$0.43^{+0.07}_{-0.07}$
NRSur	-0.93	$0.34^{+0.08}_{-0.08}$	-0.08	$0.42^{+0.09}_{-0.16}$	-0.3	$0.41^{+0.07}_{-0.08}$

$\chi_p^{\text{inj}} = 1.57$	PhenX		PhenT		NRSur	
	Q	χ_p	Q	χ_p	Q	χ_p
PhenX	-0.6	$1.53^{+0.08}_{-0.09}$	-1	$0.71^{+0.06}_{-0.07}$
PhenT	-1	$0.56^{+0.01}_{-0.01}$	-0.47	$1.54^{+0.08}_{-0.08}$
NRSur	-1	$0.55^{+0.03}_{-0.03}$	-1	$0.96^{+0.2}_{-0.22}$	-0.48	$1.53^{+0.1}_{-0.12}$

with the samples at our disposal). The worst cases appears to be those when we recover with IMRPhenomXPHM, which produce a χ_p posterior that is entirely below unity.

While a more complete investigation on waveform systematics is beyond the scope of this work, the selected cases studied here tentatively indicate that current state-of-the-art approximants struggle at providing a consistent modeling of the signal in the $\chi_p > 1$ region, to a level that will be significant for the heavy, loud sources expected in O4. The discrepancies between the waveform models in the high-precession limit are potentially expected due to the differences in the prescriptions for the spin dynamics between the models. In IMRPhenomXPHM, the precession angles are calculated by applying the stationary phase approximation to the multitime-scale analysis of the precession equations [33,56]. The prescription is then artificially extended through the merger and ringdown beyond its regime of validity. In IMRPhenomTPHM, the precession angles are calculated by direct integration of the equations of motion, coupled to a semianalytical approximation for the merger-ringdown that relies on an angular velocity determined by the quasinormal mode frequencies of the remnant BH [57].

Waveform developers are actively working toward calibrating PN-based waveform models using numerical-relativity simulations with precessing spins (e.g., [58]), which will hopefully alleviate the systematic deviations reported here. It is also important to stress that, for this exercise, we had to select sources with high-enough mass such that the signal is fully covered by NRSur7dq4, but these are also the systems where the precession signature is expected to be weaker. This is because precession cycles are contained in the low-frequency part

of the signal that gradually falls out of band as the total mass increases.

On a more positive note, the conclusion that emerges here is that waveform systematic do not produce false positive: if a future observation will deliver $\chi_p > 1$ at high confidence, it appears safe to claim that the BH binary had two precessing spin.

IV. CONCLUSIONS

If an incoming LIGO/Virgo source is composed of merging BHs with two precessing spins, will we able to tell? In this paper we have provided a statistical assessment of this question using a large set of software injections.

For dimensionality reduction and interpretation purposes, it is useful to have a single parameter that can capture the effect of precession in GW data. We employ a recent generalization [28] of the effective precessing spin χ_p [29]. Unlike its predecessor, the augmented formulation does not assume that one of the two spin dominates the dynamics. In particular, the region of the parameter space $1 < \chi_p \leq 2$ is exclusive to binaries with two precessing spins. Because spin-spin couplings in GW data provide a weak contribution to the waveform, measuring a source with $\chi_p > 1$ also requires sensitive detectors. While such a detection has not occurred in current data [28,30], our software injections at O4 sensitivity demonstrate that this goal is well within our reach.

In this paper, we have concentrated solely on inference of precession with the χ_p estimator. With the BH-binary parameter space spanning 15 dimensions, our large set of injections naturally contains much more information that could potentially be extracted (including, but not limited to, different spin precession estimators, correlation between the effective spins and other binary parameters, and the vector spin components themselves). In order to facilitate further exploitation, our posterior chains are made publicly available in their entirety at github.com/ViolaDeRenzi/twoprecessingspins [59]. The total computational budget to collect these data amounts to about half a million CPU hours.

It is important to stress that the occurrence of two precessing spins is a necessary, but not sufficient, condition to obtain values $\chi_p > 1$, (i.e., being in this region implies the source has two precessing spins, but not vice versa). Such large values of χ_p require a considerable fine-tuning of the binary’s intrinsic parameters (large spin magnitudes, mass ratios close to unity, spins coplanar with the orbit and aligned with each other). This makes our assessment very conservative and lets us identify sources with smoking-gun evidence of two-spin precession.

As the detectors’ sensitivities increase and one moves beyond effective-spin parametrizations, inference on the higher-dimensional spin parameter space will hopefully allow us to relax such conservative assumptions.

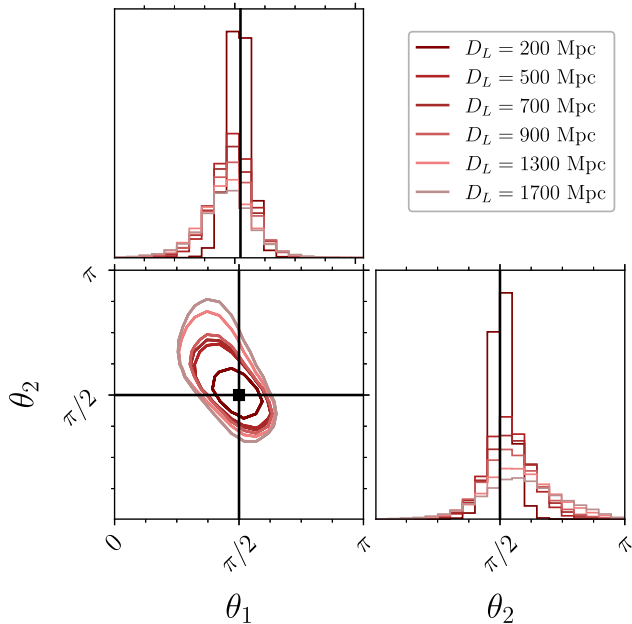


FIG. 8. Joint posterior distribution of the tilt angles θ_1 and θ_2 for the single-system series described in Sec. III A. Darker (lighter) curves and histograms refer to sources at smaller (larger) distance. Contour level correspond to 90% credible intervals.

Looking ahead in this direction, constraining a source away from the edges of the θ_1 – θ_2 plane can also be taken as a telltale sign of two-spin precession (recall that $0 \leq \theta_i \leq \pi$). Figure 8 shows the joint posterior distribution of the spin tilts for the series of injections described in Sec. III A where, indeed, nonprecessing configurations can be largely excluded. One point worth stressing is the short-timescale dependence of the tilt angles, which results in deeper issues when performing population studies [60]. In contrast, the formulation of χ_p explored here only varies on the longer

radiation-reaction timescale—which is the best one can hope for in the absence of additional constant of motions besides χ_{eff} [32].

Our large-scale injection study shows that BH binaries with two prominently precessing spins at sufficiently high SNR can be generically identified as such. We also pointed out how the departure from Gaussianity of the χ_p distribution can be a precious indicator of the suboptimality of the adopted indicator. At the same time, prior effects and waveform systematics introduce some interpretation issues that need to be further explored.

Of all the software injections we performed with different source parameters and waveform models, we did not detect a single confident false positive (i.e., a source with $\chi_p < 1$ which is erroneously recovered as having $\chi_p > 1$). Our study strongly indicates that, should a confident detection with $\chi_p > 1$ be made in O4, this would provide a conservative and safe claim of the first observation of a merging BH binary with two precessing spins.

ACKNOWLEDGMENTS

We thank Daria Gangardt, Nathan Steinle, Floor Broekgaarden, and Roberto Cotesta for discussions. V. D. R., D. G., and M. M. are supported by European Union’s H2020 ERC Starting Grant No. 945155-GWmining, Cariplo Foundation Grant No. 2021-0555, and Leverhulme Trust Grant No. RPG-2019-350. V. D. R. acknowledges support from H2020 project HPC-EUROPA3 (INFRAIA-2016-1-730897). G. P. and P. S. acknowledge support from STFC Grant No. ST/V005677/1. Computational work was performed at CINECA with allocations through INFN, Bicocca, and IS CRA project HP10BEQ9JB. This manuscript has LIGO document number P2200196.

-
- [1] T. A. Apostolatos, C. Cutler, G. J. Sussman, and K. S. Thorne, *Phys. Rev. D* **49**, 6274 (1994).
 - [2] L. E. Kidder, *Phys. Rev. D* **52**, 821 (1995).
 - [3] J. Aasi, B. P. Abbott, R. Abbott, T. Abbott, M. R. Abernathy, K. Ackley, C. Adams, T. Adams, P. Addesso *et al.* (LIGO Scientific Collaboration), *Classical Quantum Gravity* **32**, 074001 (2015).
 - [4] F. Acernese, M. Agathos, K. Agatsuma, D. Aisa, N. Allemandou, A. Allocca, J. Amarni, P. Astone, G. Balestri, G. Ballardin *et al.*, *Classical Quantum Gravity* **32**, 024001 (2015).
 - [5] D. Gerosa, M. Kesden, E. Berti, R. O’Shaughnessy, and U. Sperhake, *Phys. Rev. D* **87**, 104028 (2013).
 - [6] C. L. Rodriguez, M. Zevin, C. Pankow, V. Kalogera, and F. A. Rasio, *Astrophys. J. Lett.* **832**, L2 (2016).
 - [7] S. Vitale, R. Lynch, R. Sturani, and P. Graff, *Classical Quantum Gravity* **34**, 03LT01 (2017).
 - [8] S. Stevenson, C. P. L. Berry, and I. Mandel, *Mon. Not. R. Astron. Soc.* **471**, 2801 (2017).
 - [9] P. Amaro-Seoane *et al.*, [arXiv:1702.00786](https://arxiv.org/abs/1702.00786).
 - [10] A. Sesana, E. Barausse, M. Dotti, and E. M. Rossi, *Astrophys. J.* **794**, 104 (2014).
 - [11] M. Sayeb, L. Blecha, L. Z. Kelley, D. Gerosa, M. Kesden, and J. Thomas, *Mon. Not. R. Astron. Soc.* **501**, 2531 (2021).
 - [12] N. Loutrel, T. Tanaka, and N. Yunes, *Phys. Rev. D* **98**, 064020 (2018).
 - [13] B. P. Abbott, R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt *et al.*, *Phys. Rev. X* **9**, 031040 (2019).

- [14] R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, A. Adams, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt *et al.*, *Phys. Rev. X* **11**, 021053 (2021).
- [15] R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, N. Adhikari, R. X. Adhikari, V. B. Adya *et al.* (The LIGO Scientific, the Virgo Collaborations), [arXiv:2108.01045](https://arxiv.org/abs/2108.01045).
- [16] R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, N. Adhikari, R. X. Adhikari *et al.* (The LIGO Scientific, the Virgo, the KAGRA Collaborations), [arXiv:2111.03606](https://arxiv.org/abs/2111.03606).
- [17] M. Hannam, C. Hoy, J. E. Thompson, S. Fairhurst, V. Raymond (Members of the LIGO and Virgo Collaborations), [arXiv:2112.11300](https://arxiv.org/abs/2112.11300).
- [18] V. Varma, S. Biscoveanu, T. Islam, F. H. Shaik, C.-J. Haster, M. Isi, W. M. Farr, S. E. Field, and S. Vitale, *Phys. Rev. Lett.* **128**, 191102 (2022).
- [19] E. Payne, S. Hourihane, J. Golomb, R. Udall, D. Davis, and K. Chatziioannou, [arXiv:2206.11932](https://arxiv.org/abs/2206.11932).
- [20] I. Romero-Shaw, P. D. Lasky, E. Thrane, and J. Calderón Bustillo, *Astrophys. J. Lett.* **903**, L5 (2020).
- [21] R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, N. Adhikari, R. X. Adhikari *et al.* (The LIGO Scientific, the Virgo, the KAGRA Collaborations), [arXiv:2111.03634](https://arxiv.org/abs/2111.03634).
- [22] T. A. Callister, S. J. Miller, K. Chatziioannou, and W. M. Farr, *Astrophys. J. Lett.* **937**, L13 (2022).
- [23] M. Mould, D. Gerosa, F. S. Broekgaarden, and N. Steinle, [arXiv:2205.12329](https://arxiv.org/abs/2205.12329).
- [24] B. P. Abbott, R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, V. B. Adya, C. Affeldt, M. Agathos *et al.*, *Living Rev. Relativity* **23**, 3 (2020).
- [25] S. Khan, K. Chatziioannou, M. Hannam, and F. Ohme, *Phys. Rev. D* **100**, 024059 (2019).
- [26] S. Biscoveanu, M. Isi, V. Varma, and S. Vitale, *Phys. Rev. D* **104**, 103018 (2021).
- [27] V. Varma, M. Isi, S. Biscoveanu, W. M. Farr, and S. Vitale, *Phys. Rev. D* **105**, 024045 (2022).
- [28] D. Gerosa, M. Mould, D. Gangardt, P. Schmidt, G. Pratten, and L. M. Thomas, *Phys. Rev. D* **103**, 064067 (2021).
- [29] P. Schmidt, F. Ohme, and M. Hannam, *Phys. Rev. D* **91**, 024043 (2015).
- [30] C. Henshaw, R. O’Shaughnessy, and L. Cadonati, *Classical Quantum Gravity* **39**, 125003 (2022).
- [31] T. Damour, *Phys. Rev. D* **64**, 124013 (2001).
- [32] É. Racine, *Phys. Rev. D* **78**, 044021 (2008).
- [33] D. Gerosa, M. Kesden, U. Sperhake, E. Berti, and R. O’Shaughnessy, *Phys. Rev. D* **92**, 064016 (2015).
- [34] L. M. Thomas, P. Schmidt, and G. Pratten, *Phys. Rev. D* **103**, 083022 (2021).
- [35] D. Gangardt, N. Steinle, M. Kesden, D. Gerosa, and E. Stoikos, *Phys. Rev. D* **103**, 124026 (2021).
- [36] D. Gangardt, D. Gerosa, M. Kesden, V. De Renzi, and N. Steinle, *Phys. Rev. D* **106**, 024019 (2022).
- [37] S. Fairhurst, R. Green, M. Hannam, and C. Hoy, *Phys. Rev. D* **102**, 041302 (2020).
- [38] M. Kesden, D. Gerosa, R. O’Shaughnessy, E. Berti, and U. Sperhake, *Phys. Rev. Lett.* **114**, 081103 (2015).
- [39] G. Pratten, P. Schmidt, R. Busicchio, and L. M. Thomas, *Phys. Rev. Res.* **2**, 043096 (2020).
- [40] J. Veitch *et al.*, *Phys. Rev. D* **91**, 042003 (2015).
- [41] J. Veitch and A. Vecchio, *Phys. Rev. D* **78**, 022001 (2008).
- [42] G. Ashton *et al.*, *Astrophys. J. Suppl. Ser.* **241**, 27 (2019).
- [43] R. J. E. Smith, G. Ashton, A. Vajpeyi, and C. Talbot, *Mon. Not. R. Astron. Soc.* **498**, 4492 (2020).
- [44] J. S. Speagle, *Mon. Not. R. Astron. Soc.* **493**, 3132 (2020).
- [45] S. Vitale, D. Gerosa, C.-J. Haster, K. Chatziioannou, and A. Zimmerman, *Phys. Rev. Lett.* **119**, 251103 (2017).
- [46] K. Chatziioannou, G. Lovelace, M. Boyle, M. Giesler, D. A. Hemberger, R. Katebi, L. E. Kidder, H. P. Pfeiffer, M. A. Scheel, and B. Szilágyi, *Phys. Rev. D* **98**, 044028 (2018).
- [47] G. Pratten, S. Husa, C. García-Quirós, M. Colleoni, A. Ramos-Buades, H. Estellés, and R. Jaume, *Phys. Rev. D* **102**, 064001 (2020).
- [48] C. García-Quirós, M. Colleoni, S. Husa, H. Estellés, G. Pratten, A. Ramos-Buades, M. Mateu-Lucena, and R. Jaume, *Phys. Rev. D* **102**, 064002 (2020).
- [49] 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, *Phys. Rev. D* **103**, 104056 (2021).
- [50] H. Estellés, M. Colleoni, C. García-Quirós, S. Husa, D. Keitel, M. Mateu-Lucena, M. d. L. Planas, and A. Ramos-Buades, *Phys. Rev. D* **105**, 084040 (2022).
- [51] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, D. Gerosa, L. C. Stein, L. E. Kidder, and H. P. Pfeiffer, *Phys. Rev. Res.* **1**, 033015 (2019).
- [52] Ž. Ivezić, A. J. Connolly, J. T. VanderPlas, and A. Gray, *Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data, Updated Edition* (Princeton University Press, 2019), <https://press.princeton.edu/books/hardcover/9780691198309/statistics-data-mining-and-machine-learning-in-astronomy>.
- [53] J. Chambers, W. Cleveland, B. Kleiner, and P. Tukey, *Graphical Methods for Data Analysis* (Wadsworth International Group, 1983).
- [54] S. Vitale, R. Lynch, J. Veitch, V. Raymond, and R. Sturani, *Phys. Rev. Lett.* **112**, 251101 (2014).
- [55] M. Pürrer, M. Hannam, and F. Ohme, *Phys. Rev. D* **93**, 084042 (2016).
- [56] K. Chatziioannou, A. Klein, N. Yunes, and N. Cornish, *Phys. Rev. D* **95**, 104004 (2017).
- [57] S. Marsat and J. G. Baker, [arXiv:1806.10734](https://arxiv.org/abs/1806.10734).
- [58] E. Hamilton, L. London, J. E. Thompson, E. Fauchon-Jones, M. Hannam, C. Kalaghatgi, S. Khan, F. Pannarale, and A. Vano-Vinuales, *Phys. Rev. D* **104**, 124027 (2021).
- [59] V. De Renzi and D. Gerosa, github.com/ViolaDeRenzi/twoprecessingspins, 10.5281/zenodo.6777952 (2022).
- [60] M. Mould and D. Gerosa, *Phys. Rev. D* **105**, 024076 (2022).