Constraining gravitational wave amplitude birefringence with GWTC-3

Thomas C. K. Ng,[1, *] Maximiliano Isi,[2, †] Kaze W. K. Wong,[2, ‡] and Will M. Farr[2, 3, §]

[1]*Department of Physics, The Chinese University of Hong Kong, Shatin, Hong Kong*
[2]*Center for Computational Astrophysics, Flatiron Institute,*
162 5th Ave, New York, NY 10010, United States
[3]*Department of Physics and Astronomy, Stony Brook University, Stony Brook NY 11794, United States*
(Dated: January 31, 2023)

The propagation of gravitational waves can reveal fundamental feature of the structure of space-time. For instance, differences in the propagation of gravitational-wave polarizations would be a smoking gun of parity violations in the gravitational sector, as expected from birefringent theories like Chern-Simons gravity. Here we look for evidence of amplitude birefringence in the latest LIGO-Virgo catalog (GWTC-3) through the use of birefringent templates inspired by dynamical Chern-Simons gravity. From 70 binary-black-hole signals, we obtain the most precise constraints on amplitude birefringence yet, an order of magnitude more stringent than previous results.

I. INTRODUCTION

Gravitational wave (GW) detections by the LIGO-Virgo-KAGRA Collaboration (LVK) [1–3] are now routinely used to test various aspects of Einstein's theory of general relativity (GR) [4–6]. Among those, measurements of the basic properties of GWs, like their speed and polarization, can directly probe the fundamental symmetries of the underlying theory of gravity [7]. For instance, unequal propagation of GW polarization states would reveal that spacetime is birefringent, a smoking gun for parity-odd theories like Chern-Simons gravity [8–10]. Here we constrain the magnitude of possible amplitude birefringence using binary black hole (BBH) signals from the latest LVK catalog, GWTC-3.

Previous studies have constrained amplitude birefringence by performing different statistical analyses. Yamada and Tanaka [11] and Wang *et al.* [12] both performed parameter estimation (PE) on the events in the first GW transient catalog [13], GWTC-1, using birefringent templates. Okounkova *et al.* [14] considered the distribution of observed inclinations of the GW events in the second GW transient catalog [15], GWTC-2, to look for signs of birefringence.

In this study, we used a frequency-dependent birefringence model to constrain the strength of GW amplitude birefringence by performing PE. This model is a better approximation of the birefringence effect than the frequency-independent model used in Okounkova *et al.* [14]. Compared to previous studies, we also performed PE on more events, including events in the third GW transient catalog [16], GWTC-3. We consider 70 binary black hole merger events with a false-alarm rate (FAR) $\leq 1\mathrm{yr}^{-1}$. These events are also listed in Table I in Ab-

bott *et al.* [17].[1] We then give the population constraint on the strength of GW amplitude birefringence by using the PE results.

In Sec. III, we describe the modification we made to the waveform model, mention the configuration we used in the PE and show the method we used to obtain the population constraint on GW amplitude birefringence. In Sec. IV, we present the population constraint on GW amplitude birefringence we obtained and show the results of individual PE. In Sec. V, we discuss the limitation of this study and provide suggestions for future studies.

II. BACKGROUND

A. Birefringence

In GR, GWs are comprised of two independent polarization modes, usually represented in the linear basis of plus (+) and cross (×) states. In the Fourier domain, they can be combined into left-handed (L) and right-handed (R) circular states [18],

$$h_{L/R} = \frac{1}{\sqrt{2}} \left(h_+ \pm i h_\times \right) , \qquad (1)$$

where h is the frequency domain strain, with the plus and minus signs for L and R respectively. These circular modes represent eigenstates of the helicity operator (helicity ± 2) and posses a definite parity. Einstein's theory, which is parity even, predicts no difference in the dynamics of these two states.

Yet, parity odd extensions of GR may make distinctions between the two circular polarizations, which can manifest in both the generation and propagation of GWs.

---

* thomas.ng@link.cuhk.edu.hk
† misi@flatironinstitute.org
‡ kwong@flatironinstitute.org
§ wfarr@flatironinstitute.org

---

[1] GW190720 is included in the list in Abbott *et al.* [17], but we did not include it in this study. Details in Sec. IV G.

The latter can manifest in changes to the relative amplitude and phase of the polarizations that accrue as the wave propagates, giving us hope to detect initially small effects that compound over long propagation distances.

In particular, *amplitude* birefringence would enhance one polarization mode over the other. To first order in theories like Chern-Simons gravity, the Fourier-domain waveform observed a comoving distance d_C away from the source takes the form of

$$h_{L/R}^{\mathrm{br}}(f) = h_{L/R}^{\mathrm{GR}}(f) \times \exp\left(\pm\kappa \frac{d_C}{1\,\mathrm{Gpc}} \frac{f}{100\,\mathrm{Hz}}\right), \quad (2)$$

where the emitted waveform $h_{L/R}^{\mathrm{GR}}$ is modified by an exponential birefringent factor to yield the observed waveform $h_{L/R}^{\mathrm{br}}$. The overall magnitude of this effect for a given distance and frequency f is set by a dimensionless opacity parameter, κ, which encodes the intrinsic strength of the birefringence. The emitted waveform for a given source (i.e., the waveform observed very close to the source) will generally differ from the analogous waveform predicted by GR; however, since we expect most viable modifications to GR to be intrinsically small, it is standard to approximate the emitted waveform by the prediction from GR (hence the notation "h^{GR}" above).

Although the intrinsic modification is small, the effect targeted by Eq. (2) accumulates as the GW propagates. During propagation, the effect of birefringence will be built up with the number of cycles which depends on the distance traveled and the frequency of the GWs. According to Eq. (2), a positive κ means the left-handed polarization is enhanced over the right-handed polarization, while a negative κ means the opposite; when $\kappa = 0$, the observed waveform is the same as GR predicts, meaning there is no birefringence.

Equation (2) can be derived as the first order effect in an expansion for multiple extensions of GR. In general, κ will be a function of the theory parameters and the cosmological history, e.g., the value of the pseudo-scalar field and its derivative in Chern-Simons gravity. Since it originates from an expansion, Eq. (2) is a good approximation only for small exponents,

$$|\kappa|\,(d_C/1\,\mathrm{Gpc})\,(f/100\,\mathrm{Hz}) < 1. \quad (3)$$

Otherwise, more frequency-dependent terms would enter the exponent of Eq. (2) in a theory-dependent way. [**MAX**: specify what the parameter for the expansion is, add citations, and check with a theorist]

B. Relation to inclination

Under certain conditions, the effect of birefringence can be degenerate with a change in the orientation of the source with respect to the line of sight. Concretely, for a nonprecessing compact binary inspiral in GR, the observed amplitude ratio of the left-handed and right-handed polarizations is only a function of the inclination ι, the angle between the orbital angular momentum
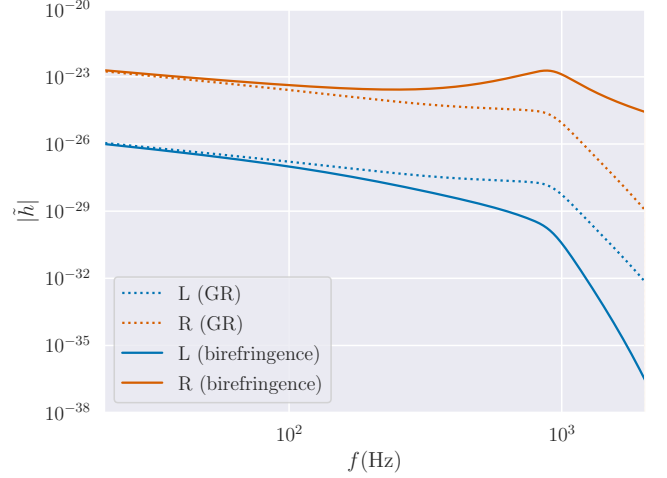

FIG. 1. The observed Fourier amplitude of two polarizations of a simulated GW signal with and without birefringence. The amplitudes, as GR predicts, are shown in dotted lines with blue and orange colors representing the left-handed and right-handed polarizations. The amplitudes with birefringence are shown in solid lines with blue and orange colors representing the left-handed and right-handed polarizations. This plot shows the effect of amplitude birefringence with frequency dependence. [**MAX**: @Thomas can you add a panel with the time-domain waveforms (maybe just for L or R)?]

of the source and the line of sight. For the dominant $\ell = |m| = 2$ angular mode of the radiation, the relation between the amplitude ratio and the inclination is

$$\frac{h_L^{\mathrm{GR}}}{h_R^{\mathrm{GR}}} = \left(\frac{1 - \cos\iota}{1 + \cos\iota}\right)^2 \quad (4)$$

for all frequencies (see, e.g., Sec. IIIC in [18]).

Since birefringence impacts the observed amplitude ratio of left- and right-handed modes, it could also affect inferences about the source inclination [9]. However, the two effects are degenerate only if the frequency dependence of Eq. (2) is neglected. This is easy to see from Eq. (2), since the implied polarization ratio for the $\ell = |m| = 2$ mode of a nonprecessing source is

$$\frac{h_L^{\mathrm{br}}}{h_R^{\mathrm{br}}} = \left(\frac{1 - \cos\iota}{1 + \cos\iota}\right)^2 \exp\left(2\kappa \frac{d_C}{1\,\mathrm{Gpc}} \frac{f}{100\,\mathrm{Hz}}\right). \quad (5)$$

For an isolated Fourier mode of definite frequency f, the effect of birefringence will be degenerate with a change in inclination; however, if multiple modes come into play, then no change in inclination can mask the effect of birefringence, which will generate dephasing in the time domain waveform (Fig 1).

Okounkova *et al.* [14] took the effect of birefringence to be independent of the frequency, which is a zeroth-order approximation of the birefringence model in Chern-Simons gravity. This assumption results in a full degeneracy between κ and ι: to reconstruct the amplitude ratio

from the interferometer data, an ι representing a more face-off inspiral can pair with a positive κ, or an ι representing a more face-on inspiral with a negative κ. That fact can be used to constrain frequency-independent birefringence by searching for features in the distribution of inferred inclinations [14].

By implementing Eq. (2), which is a first-order approximation of the birefringence model, we generally break the degeneracy between birefringence and source orientation; this was also the case in the frequency-dependent relations studied in [11, 12]. Nevertheless, there exist systems for which the degeneracy cannot be perfectly broken because not enough frequencies are available in the data. As we will see in Sec. IV, this is the case for heavy BBHs, which are only in the sensitive band of the detectors for a few cycles.

III. METHOD

A. Single-event parameter estimation

To constrain birefringence, we reanalyze events from GWTC-3 [16, 19] implementing Eq. (2) to directly obtain a posterior on κ from the strain of each event. We analyze the 70 BBHs that were detected with FAR $< 1/\mathrm{yr}$; to avoid extended computations on longer signals and considering these are generally at closer distances, we do not analyze systems involving neutron stars in this work. We procure strain data from the Gravitational Wave Open Science Center (GWOSC) [20].

We carry out PE using a custom version of BILBY [21], that applies Eq. (2) for any GR baseline waveform. We depart from the PE configuration as in Abbott *et al.* [16, 19] **[MAX: @Thomas cite actual zenodo link]**, with IMRPHENOMXPHM as the reference waveform. We apply a distance prior uniform in comoving volume **[MAX: @Thomas is this right?]**, and set the prior on κ to a uniform distribution between -1 and 1.

[MAX: @Thomas list here any special considerations for individual events]

In order to validate our PE implementation, we reproduce the LVK PE results obtained assuming GR by enforcing $\kappa = 0$; this also has the advantage of producing GR runs that are directly comparable to our birefringent runs. All our PE results, including the GR valudation runs, are published in Ng *et al.* [22].

B. Collective analysis

1. Shared birefringence parameter

Since birefringence is a property of spacetime and we are targeting a propagation effect not intrinsic to any source, we should expect κ to take the same value for all signals, whether it vanishes or not. **[MAX: caveats?]**

Under this assumption, the posterior on κ inferred collectively from all events with a uniform prior is simply obtained from the product of the individual likelihoods, $p(d_i \mid \kappa) \propto p(\kappa \mid d_i)/p(\kappa)$, such that

$$p(\kappa \mid \{d_i\}) \propto p(\kappa) \prod_i \frac{p(\kappa \mid d_i)}{p(\kappa)}, \qquad (6)$$

where d_i is the strain data for the ith event, and $p(\kappa)$ is the prior on κ; since the prior is uniform, in our case Eq. (6) reduces to the product of the posteriors, namely $p(\kappa \mid \{d_i\}) \propto \prod_i p(\kappa \mid d_i)$. We use Eq. (6) to obtain the primary constraint presented in this work.

2. Nonshared birefringence parameters

In order to further characterize our set of measurements, and identify potential outliers, we carry out an additional collective analysis that allows κ to take different values for each event [23, 24]. To do this, we apply hierarchical Bayesian inference [25] to model the distribution of κ's consistent with the observed data: we posit that there is a specific value of the parameter, κ_i, associated to each event, drawn from some unknown distribution of underlying values; from the imperfect measurements of κ_i for each event, we may reconstruct the underlying distribution. If we are interested in constraining the first two moments of the distribution, it is convenient to parametrize the κ_i's as drawn from a Gaussian with unknown mean μ and variance σ^2, i.e., $\kappa_i \sim \mathcal{N}(\mu, \sigma^2)$ [26], and measure those hyperparameters from the collection of observed data.

If GR is correct and there is no birefringence, we should find the observed κ distribution to be consistent with a delta function at the origin ($\kappa_i = 0$ for all i, or $\mu = \sigma = 0$); on the other hand, if spacetime is birefringent, we expect to find a delta function at some nonzero value ($\kappa_i = \kappa \neq 0$, or $\mu = \kappa$ and $\sigma = 0$). But this analysis also has the power to reveal unexpected physics or systematics in our measurements: if σ is confidently found to be nonzero, this would imply that our set of measurements are statistically unlikely to originate from a unique κ value. This could signal further physics than is implied by Eq. (2) **[MAX: which physics? @Will, any ideas?]** or, more prosaically, that there are outliers in our measurements due to mismodeling, e.g., in the waveform approximant or the noise of the detector.

Starting from the posterior on κ from each event, the posterior on the hyperparameters μ and σ can be calculated by

$$p(\mu, \sigma \mid \{d\}) \propto p(\mu, \sigma) \prod_i \int \frac{p(\kappa_i \mid d_i)}{p(\kappa_i)} p(\kappa_i \mid \mu, \sigma) \, d\kappa_i, \tag{7}$$

where $p(\kappa)$ is the prior initially applied to κ during PE, which in our case is a uniform distribution $\mathcal{U}[-1, 1]$.

Furhter choosing the hyperpriors on μ and σ to be uniform, Eq. (7) simplifies to

$$p(\mu, \sigma \mid \{d\}) \propto \prod_i \int p(\kappa_i \mid d_i) \mathcal{N}(\mu, \sigma) \, d\kappa_i \,. \quad (8)$$

We also calculate the expected population distribution of κ marginalized over μ and σ. This is given by

$$p(\kappa \mid \{d\}) = \int \mathcal{N}(\mu, \sigma) p(\mu, \sigma \mid \{d\}) \, d\mu \, d\sigma \,, \quad (9)$$

and represents our overall expectation for κ, given the observed set of individual measurements. To sample the posterior distribution of μ and σ, we use the sampling package FLOWMC [27].

IV. RESULTS

In this section, we present the results of our study. We first show the PE results of all events and present the results of Bayesian hierarchical modeling. Then, we will discuss some special events individually.

A. Result on GWTC-3

In figure 2, we show the posterior of κ for each event in the form of a violin plot. The median and confidence interval of κ are calculated with the restricted posterior distribution of κ, which is given by equation 6. We excluded GW200129 from the calculation of the median and confidence interval of κ. Details of the exclusion are discussed in Sec. IV F.

B. Bayesian hierarchical modeling

We then show the result of Bayesian hierarchical modeling in figure 3. The median of μ within 90% confidence interval is $-0.002^{+0.034}_{-0.034}$. The median of σ within 90% confidence interval is $0.033^{+0.024}_{-0.027}$. The posterior of μ is concentrated around 0, which means the population constraint on κ is consistent with GR.

C. Constaint on κ

In figure 4, we show both the restricted posterior of κ and the generic posterior of κ. The restricted posterior is calculated by equation 6 with the assumption that κ is a single value constant. The generic posterior is calculated by equation 9 with the assumption that κ is a true value drawn from a Gaussian distribution. The median of the restricted posterior within 90% confidence interval is $-0.010^{+0.044}_{-0.035}$. The median of the generic posterior within 90% confidence interval is $-0.005^{+0.066}_{-0.058}$. Note that the result is consistent with GR when $\kappa = 0$.

D. Result on GW150914

In this study, we included frequency dependence of birefringence, which would affect the posterior of κ obtained from PE. Consider GW150914, the first GW detected by LIGO, as an example. In figure 5, we show the posteriors of the parameters of GW150914. With the frequency-independent birefringence model, the posteriors for $\cos \iota$ look different from the posteriors assuming GR. This is because, for a nonprecessing system, there is a degeneracy between κ and ι if the frequency dependence is not included.

On the other hand, with the frequency-dependent birefringence model, the posterior looks similar to the GR posterior. This is because the frequency dependence broke the degeneracy, as the effect of birefringence will differ from the effect of changing ι. In this case, the most probable value of κ is close to 0, which means the birefringence is weak or absent. Therefore, the PE result with frequency dependence is consistent with GR.

E. Result on GW190521

Even with the frequency-dependent birefringence model, some events still have a bimodal κ distribution. Consider GW190521, the most massive binary black hole merger in the events we included. The degeneracy between κ and ι cannot be broken by the frequency-dependent birefringence model, as shown in figure 6. The reason could be this GW event's frequency range is too narrow to break the degeneracy. The effect of birefringence at different frequencies within the range is similar, so the degeneracy cannot be broken. Therefore, the narrow frequency range can be a possible reason why the κ distribution of GW190521 is bimodal.

F. Exclusion of GW200129

The reason for the exclusion is that past research suggests a potential glitch in GW200129 Virgo data. [28] We perform three sets of PE to check if the glitch would affect the population posterior of κ. We pick only two out of three detector data in each set to perform PE. The two sets of PE results with Virgo data return a κ posterior that strongly suggests a negative κ. Further investigation is needed to understand the glitch. Thus, we did not include this event in our analysis.

G. Exclusion of GW190720

For GW190720, we could not perform PE with Bilby successfully. In LVK PE papers [16, 19], the PE of some events was performed with LALInference [29] instead of Bilby. GW190720 was one of them. We successfully migrated the LALInference configuration to Bilby for most
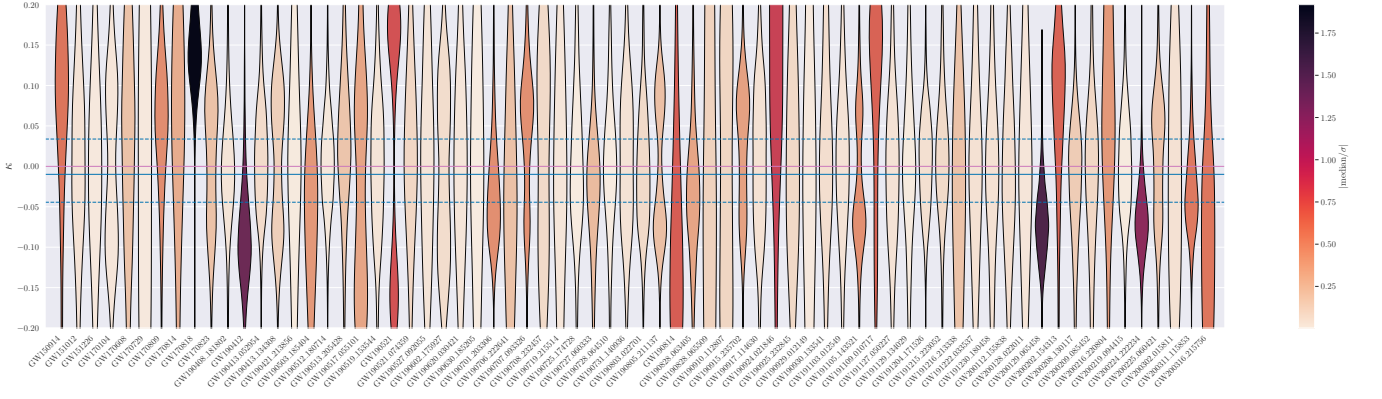
FIG. 2.   The violin plot shows the posterior of κ for 69 events in this study. Each violin represents a different event. The color of the violins represents the quotient of the median and standard deviation of the posterior. The blue horizontal solid line represents the median of the restricted posterior of κ, and the blue dashed lines enclose the 90% confidence interval of the restricted posterior of κ. The pink horizontal solid line represents $\kappa = 0$.
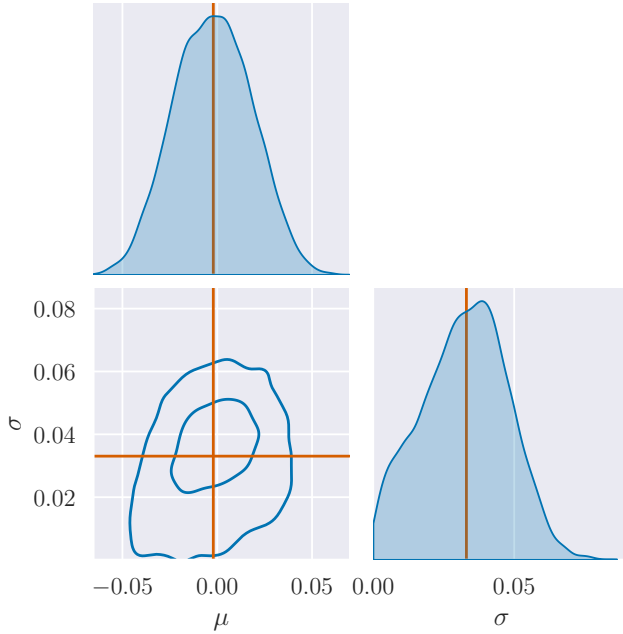

FIG. 3.   The posterior of μ and σ of the population distribution of κ. 2D contour plot shows 39.35% and 90% confidence level. The orange line represents the median of μ and σ. The plot shows that the population constraint on κ is consistent with GR.

FIG. 4.   The generic and restricted posterior of κ. The blue solid line shows the restricted posterior of κ, and the blue dashed lines enclose the 90% confidence interval. The orange solid line shows the generic posterior of κ, and the orange dashed lines enclose the 90% confidence interval. The pink vertical solid line represents $\kappa = 0$.

V.   DISCUSSION

A.   Comparison with previous studies

Okounkova *et al.* [14] gave a constraint on GW amplitude birefringence by performing statistical analysis on GWTC-2. We convert the constraint they gave to the same units as ours to make a comparison. They were able to constrain κ to be $|\kappa| \lesssim 0.74$ at 1σ. We obtained a tighter constraint on κ with $|\kappa| \lesssim 0.04$ at 1σ. Our result is an order of magnitude improvement in constraining κ. The main reason is that we have more events from GWTC-3 than GWTC-2.
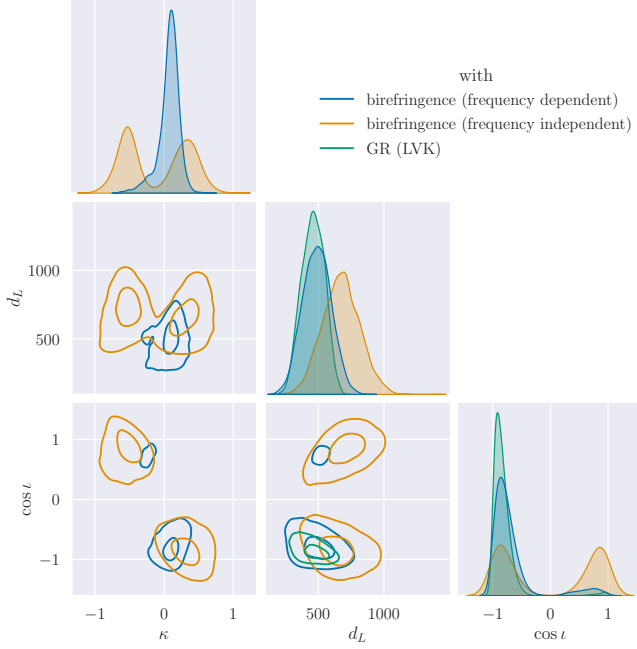
of these events and recovered similar PE results, except for GW190720. Bilby could not evaluate the likelihood function on Virgo data. As a result, we excluded this event from our analysis.

FIG. 5. The posterior of κ, luminosity distance d_L and $\cos\iota$ for GW150914. Colors in the plot represent the PE result with GR done by LVK without cosmological reweighing [16, 19], the PE results done by us with both frequency independent and dependent birefringence respectively. 2D contour plot shows 39.35% and 90% confidence level. Note that there is no posterior of κ for the PE result from LVK, as the LVK PE is based on GR, which does not suggest GW amplitude birefringence. This plot shows that frequency-independent birefringence creates a degeneracy between κ and ι, while frequency-dependent birefringence can break it.

B. Future work

Future work is to perform PE on binary neutron star mergers, such as GW170817. The frequency range of the signal is much wider compared to the binary black hole mergers. Thus, the difference in the effect of birefringence at different frequencies within the range can be more significant. The PE on binary neutron star mergers can allow us to further constrain the birefringence effect and the beyond-GR theories that predict it. However, performing PE on binary neutron star mergers requires much more computational resources. Therefore, we may need to wait for future PE methods and tools to further our work.

Another future work is to apply the same method to more GW events and events with a higher signal-to-noise ratio (SNR). LVK will release more events with higher SNR in the future. Using data with higher SNR allows us to obtain more precise PE results and constrain the birefringence effect more precisely. And using data from more events will allow us to calculate a more constrained population posterior of κ.
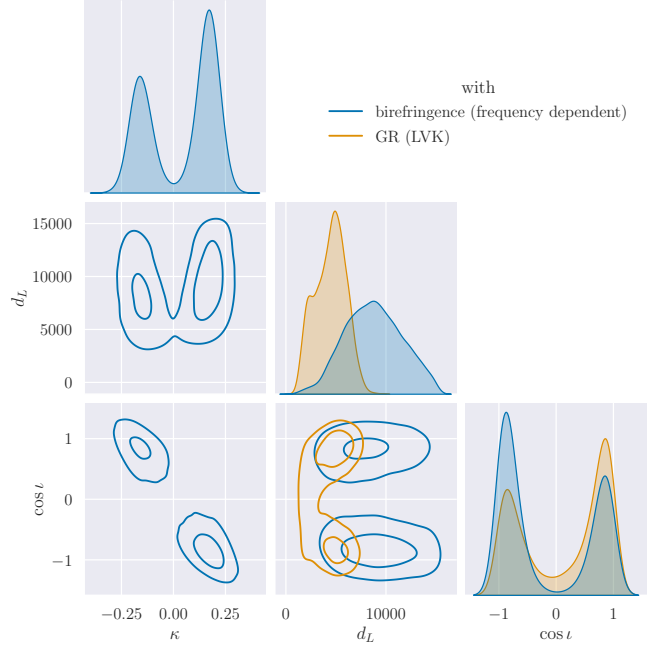

FIG. 6. The posterior of κ, luminosity distance d_L and $\cos\iota$ for GW190521. Colors in the plot are the PE result with GR done by LVK without cosmological reweighing [16, 19] and the PE result done by us with the frequency-dependent birefringence. 2D contour plot shows 39.35% and 90% confidence level. This plot shows that frequency-dependent birefringence cannot break the degeneracy between κ and ι.

[1] J. Aasi *et al.* (LIGO Scientific Collaboration), Advanced LIGO, Classical and Quantum Gravity **32**, 074001 (2015).

[2] F. Acernese *et al.*, Advanced virgo: a second-generation interferometric gravitational wave detector, Classical and Quantum Gravity **32**, 024001 (2014).

[3] T. Akutsu *et al.*, Overview of kagra: Detector design and construction history (2020).

[4] B. P. Abbott *et al.* (LIGO Scientific, Virgo), Tests of general relativity with GW150914, Phys. Rev. Lett. **116**, 221101 (2016), [Erratum: Phys.Rev.Lett. 121, 129902 (2018)], arXiv:1602.03841 [gr-qc].

[5] B. P. Abbott *et al.* (LIGO Scientific, Virgo), Tests of General Relativity with GW170817, Phys. Rev. Lett. **123**, 011102 (2019), arXiv:1811.00364 [gr-qc].

[6] R. Abbott *et al.* (LIGO Scientific, VIRGO, KAGRA), Tests of General Relativity with GWTC-3, (2021), arXiv:2112.06861 [gr-qc].

[7] C. M. Will, *Theory and Experiment in Gravitational Physics* (Cambridge University Press, 2018).

[8] R. Jackiw and S. Y. Pi, Chern-Simons modification of general relativity, Phys. Rev. D **68**, 104012 (2003), arXiv:gr-qc/0308071.

[9] S. Alexander and N. Yunes, Chern-Simons Modified General Relativity, Phys. Rept. **480**, 1 (2009), arXiv:0907.2562 [hep-th].

[10] C. F. Sopuerta and N. Yunes, Extreme and Intermediate-Mass Ratio Inspirals in Dynamical Chern-Simons Modified Gravity, Phys. Rev. D **80**, 064006 (2009), arXiv:0904.4501 [gr-qc].

[11] K. Yamada and T. Tanaka, Parametrized test of parity-violating gravity using GWTC-1 events, Progress of Theoretical and Experimental Physics **2020**, 10.1093/ptep/ptaa103 (2020), 093E01, https://academic.oup.com/ptep/article-pdf/2020/9/093E01/33689419/ptaa103.pdf.

[12] Y.-F. Wang, R. Niu, T. Zhu, and W. Zhao, Gravitational wave implications for the parity symmetry of gravity in the high energy region, The Astrophysical Journal **908**, 58 (2021).

[13] B. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), GWTC-1: A gravitational-wave transient catalog of compact binary mergers observed by LIGO and virgo during the first and second observing runs, Physical Review X **9**, 031040 (2019).

[14] M. Okounkova, W. M. Farr, M. Isi, and L. C. Stein, Constraining gravitational wave amplitude birefringence and chern-simons gravity with GWTC-2, Physical Review D **106**, 10.1103/physrevd.106.044067 (2022).

[15] R. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), GWTC-2: Compact binary coalescences observed by LIGO and virgo during the first half of the third observing run, Physical Review X **11**, 021053 (2021).

[16] R. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration and KAGRA Collaboration), Gwtc-3: Compact binary coalescences observed by ligo and virgo during the second part of the third observing run (2021).

[17] R. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration and KAGRA Collaboration), The population of merging compact binaries inferred using gravitational waves through gwtc-3 (2021).

[18] M. Isi, Parametrizing gravitational-wave polarizations, (2022), arXiv:2208.03372 [gr-qc].

[19] R. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Gwtc-2.1: Deep extended catalog of compact binary coalescences observed by ligo and virgo during the first half of the third observing run (2021).

[20] R. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Open data from the first and second observing runs of advanced LIGO and advanced virgo, SoftwareX **13**, 100658 (2021).

[21] G. Ashton, M. Hübner, P. D. Lasky, C. Talbot, K. Ackley, S. Biscoveanu, Q. Chu, A. Divakarla, P. J. Easter, B. Goncharov, F. H. Vivanco, J. Harms, M. E. Lower, G. D. Meadors, D. Melchor, E. Payne, M. D. Pitkin, J. Powell, N. Sarin, R. J. E. Smith, and E. Thrane, Bilby: A user-friendly bayesian inference library for gravitational-wave astronomy, The Astrophysical Journal Supplement Series **241**, 27 (2019).

[22] T. C. K. Ng, M. Isi, K. W. K. Wong, and W. M. Farr, Constraining gravitational wave amplitude birefringence with GWTC-3: Data Release, 10.5281/zenodo.7338924 (2023).

[23] A. Zimmerman, C.-J. Haster, and K. Chatziioannou, On combining information from multiple gravitational wave sources, Phys. Rev. D **99**, 124044 (2019), arXiv:1903.11008 [astro-ph.IM].

[24] M. Isi, W. M. Farr, and K. Chatziioannou, Comparing Bayes factors and hierarchical inference for testing general relativity with gravitational waves, Phys. Rev. D **106**, 024048 (2022), arXiv:2204.10742 [gr-qc].

[25] T. J. Loredo, Accounting for source uncertainties in analyses of astronomical survey data, AIP Conf. Proc. **735**, 195 (2004), arXiv:astro-ph/0409387.

[26] M. Isi, K. Chatziioannou, and W. M. Farr, Hierarchical test of general relativity with gravitational waves, Phys. Rev. Lett. **123**, 121101 (2019), arXiv:1904.08011 [gr-qc].

[27] K. W. K. Wong, M. Gabrié, and D. Foreman-Mackey, flowmc: Normalizing-flow enhanced sampling package for probabilistic inference in jax (2022).

[28] E. Payne, S. Hourihane, J. Golomb, R. Udall, D. Davis, and K. Chatziioannou, Curious case of GW200129: Interplay between spin-precession inference and data-quality issues, Physical Review D **106**, 104017 (2022).

[29] LIGO Scientific Collaboration, LALSuite: LIGO Scientific Collaboration Algorithm Library Suite, Astrophysics Source Code Library, record ascl:2012.021 (2020), ascl:2012.021.