GW190521 as a black-hole merger coincident with the ZTF19abanrhr flare

Juan Calderón Bustillo,1,2 Samson H.W. Leong,2 Koustav Chandra,3 Barry McKernan,4,5,6,7 and K. E. S. Ford4,5,6,7

1Instituto Galego de Física de Altas Enerxías, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Galicia, Spain
2Department of Physics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong
3Department of Physics, Indian Institute of Technology Bombay, Powai, Mumbai, Maharashtra 400076, India
4Department of Science, CUNY-BMCC, 199 Chambers St., New York, NY 10007, USA
5Department of Astrophysics, American Museum of Natural History, Central Park West, New York, NY 10028, USA
6Physics Program, The Graduate Center, CUNY, New York, NY 10016, USA
7Center for Computational Astrophysics, Flatiron Institute, New York, NY 10010, USA

We present an analysis that reconciles the gravitational wave signal GW190521 observed by the Advanced LIGO and Advanced Virgo detectors with the electromagnetic flare ZTF19abanrhr observed by the Zwicky Transient Facility. We analyze GW190521 under a mass-ratio prior uniform in \( Q \in [1, 4] \) using the state-of-the-art waveform model for black-hole mergers \( \text{NRSur7dq4} \). We find a 90% credible region for the black-hole masses extending far outside what originally reported by [1], where our maximum likelihood masses reside. We find a 15% probability that both black holes avoid the pair-instability supernova gap. We infer a three-dimensional sky-location highly consistent with ZTF19abanrhr, obtaining an odds-ratio \( O_{C/R} = 72 : 1 \) that strongly favors the hypothesis of a true coincidence over a random one. Combining this event with the neutron-star merger GW170817, we estimate a Hubble constant \( H_0 = 72.1^{+10.6}_{-6.4} \text{ km s}^{-1} \text{ Mpc}^{-1} \) at the 68% credible level.

I. INTRODUCTION

The gravitational-wave (GW) detectors Advanced LIGO [2], and Advanced Virgo [3] have made the observation of compact binary mergers almost routine. In only six years, these have reported \( \sim \mathcal{O}(90) \) such observations [4–6]. These have provided us with unprecedented knowledge on how BHs and neutron stars form and how they populate our Universe. Moreover, these observations have enabled the first tests of General Relativity in the strong-field regime, and qualitatively new studies of the Universe at a large scale [6, 7]. Unleashing such scientific potential from GW observations requires accurate inference of the source properties. This has been largely possible for most observations owing to both accurate, computationally efficient waveform models and to the fact that most detections displayed relatively long post-merger, inspiral stages that provided us with information about the individual merging bodies.

The detection of GW190521 by the LIGO, Virgo and KAGRA collaborations (LVK) represented the first departure from such “canonical” events [1, 8]. Owing to the large mass of its source, GW190521 barely displays any pre-merger dynamics, with the vast majority of the signal coming from the final distorted, merged object as it relaxes to its final BH form. In such a situation, there is little information about the parents of the final object, causing the inference of the source parameters to depend strongly on prior assumptions [9–14]. This has led to a large variety of interpretations of this event.

First, the LVK reported a quasi-circular binary black hole (BBH) merger with signatures of orbital precession involving at least one BH populating the pair-instability supernova (PISN) gap [15, 16]. Second, using a population-informed mass prior, Fishbach and Holz [17] hinted that GW190521 could involve one black-hole above the PISN gap and one below, known as a “straddling binary”. Next, [18] showed that for such short signals, orbital precession could be confused with high eccentricity. Consistently, Romero-Shaw et al. [11] and Gayathri et al. [12] showed that GW190521 is also consistent with an eccentric merger and Gamba et al. [13] even pointed to the possibility of a dynamical capture [19]. More important for this work, Nitz and Capano [10] showed that using a mass-ratio prior uniform in \( Q = m_1/m_2 \geq 1 \) could lead to an interpretation as an intermediate-high-mass ratio BBH; and similar results were found by Estellés et al. [20]. Finally, [14] showed that the event is consistent with the merger of horizonless exotic compact objects known as Proca stars [21, 22].

GW190521 could also be the first multi-messenger observation of a BBH event. Graham et al. [23] reported the observation of an electromagnetic signal, ZTF19abanrhr by the Zwicky Transient Facility (ZTF) [24, 25] in a region of the sky consistent with that initially reported by the LVK in an early warning [26], proposing it as a counterpart to GW190521. If true, this would have paramount implications in interpreting GWs from compact mergers, forecasts for future counterparts and measurements of the Hubble constant. However, due to larger inconsistencies with the sky-location finally estimated by the LVK for GW190521 [1], Ashton et al. [27] showed that a true association was unlikely. The same was concluded by [10, 20] under their analyses.
with alternative mass priors. Importantly, however, while GW190521 is a merger-ringdown dominated signal with signatures of precession, 20 (and in part 10) compared GW190521 to phenomenological waveform models 28 33 that are calibrated to numerical simulations of BBHs with no orbital precession, modelling this through analytical approximations that break during the merger-ringdown regime 34 35 (See Appendix A).

Here we analyse GW190521 under several mass priors using the state-of-the-art waveform model \texttt{NRSur7dq4} 36. This model, labelled as “preferred” by the LVK in their original analysis of GW190521 [11,8], is directly calibrated to 1528 numerical simulations of precessing BBHs 37 with mass ratios $Q = m_1/m_2 \leq 4$, therefore including all the physics present in these systems. Imposing a mass-ratio prior uniform in $Q = m_1/m_2 \in [1,4]$, we find that GW190521 is consistent with a BBH with masses $m_1 = 106^{+30}_{-25} \, M_{\odot}$ and $m_2 = 61^{+21}_{-19} \, M_{\odot}$ at a distance of $2.8^{+3.4}_{-1.7} \, \text{Gpc}$. Considering a PISN gap spanning the mass range $[65, 130] \, M_{\odot}$, we find a 15% probability that GW190521 is a “straddling binary”. Moreover, we recover a sky-location highly consistent with ZTF19abanrhr. Reproducing the analysis in Ashton et al. 27, we obtain an odds of $\mathcal{O}_{C/R} = 72$ against a random spatial coincidence, indicative of a strong evidence for a true association 38. We note that while 10 did also use the model \texttt{NRSur7dq4} they extended it past its calibration region, up to $Q \leq 6$. Doing so we find a lower but still strong odds-ratio of $\mathcal{O}_{C/R} = 47$ (see also Appendix A).

II. ANALYSIS SETUP

a. Data and waveform model We perform full Bayesian parameter inference on 4 seconds of publicly available data 39 from the Advanced LIGO 2 and Advanced Virgo 3 detectors around the time of GW190521 sampled at 1024 Hz and with a subtraction of the 60 Hz US-power line 40 41. We employ the same power-spectral-density estimate as in 1. We compare GW190521 to theoretical waveform templates for BBH predicted by General Relativity generated with the waveform model \texttt{NRSur7dq4} 36. This model is calibrated to numerical simulations of precessing quasi-circular BBHs with mass ratios $Q \leq 4$ and dimensionless spin-magnitudes $a_i \leq 0.8$ but can be extrapolated up to $Q \leq 6$ and $a_i \leq 0.99$.

b. Priors Following standard practice, in their original analysis of GW190521, the LVK placed uniform priors on the red-shifted component masses 11 8. We will refer to this as the LVK prior. This, however, translates into a strong prior for rather equal-mass systems that severely punishes high mass ratio regions where the best-fitting parameters may reside 10. Here, we test two different mass-ratio priors: one uniform in $q = m_2/m_1 \in [0.25,1]$, that also favours equal-mass ratios, and another one uniform in $Q = m_1/m_2 \in [1,4]$ that removes such preference. Throughout the text, we will respectively refer to these as $Q$ and $q$-runs. As a consistency test, since the \texttt{NRSur7dq4} model can be extrapolated to $Q \leq 6$, we perform additional runs extending our priors to respectively $q = 1/6$ and $Q = 6$. We place a uniform prior on the total red-shifted mass and a prior on luminosity distance uniform in co-moving volume assuming Planck 2015 cosmology. In contrast, the LVK employed a prior $p(d_L) \propto d_L^2$, uniform in Euclidean Volume. Finally, we place standard priors in all remaining parameters. Consistently with the LVK analysis 11, we compute spin-values at a fiducial reference frequency of $f_{\text{ref}} = 11 \, \text{Hz}$ and use a minimum frequency cutoff of $f_{\text{min}} = 11 \, \text{Hz}$.

c. Sampler Settings We employ the parameter estimation software \texttt{Parallel Bilby} 42 43 equipped with the nested sampler \texttt{Dynesty} 44, using 4096 live points. In contrast, we note that the LVK analysis was performed with the LALInference software 45 using its nested sampling scheme \texttt{LALInferenceNest} 46 with 2048 live points.

d. LVK cross-check As a consistency check, we also reproduce the LVK analysis by placing the same mass and distance priors.

III. RESULTS

Table I summarises our results, contrasting them with those of the LVK. We report median values and symmetric 90% credible intervals (CIs) except for Hubble constant $H_0$, for which following standard practice, we report 68% CIs. The first two columns report results corresponding to our $Q$-analyses. Throughout the text, we will use the first column ($Q \leq 4$) as our “preferred” results, with those in the second column ($Q \leq 6$) serving as a robustness check. Results in the central columns correspond to our $q$-analyses, and the last two columns show the result of our reproduction of the LVK analysis and the original LVK result, respectively.

First, we note that our reproduction of the LVK analysis matches well with their result 1. Second, as expected, $q$-analyses also report results consistent with the LVK ones. Finally, results from both $Q$-analyses are self-consistent but deviate from the other four.

---

1 We note that in this analysis, we find a probability $p(Q \geq 4) < 0.03$.

2 We attribute the small difference in maximum SNR to the different number of sampling live points (4096 vs 2048).
a. Model selection. Fig. 1 shows the posterior distributions for the component masses of GW190521 according to three different analyses. In blue is our preferred analysis, while green and orange respectively denote the original LVK analysis and our reproduction of it. The contours in the main panel represent the corresponding two-dimensional 90% credible regions (max $\mathcal{L}$), best-fitting templates. We show these in Appendix A overlaid onto the detector data. The colour darkness denotes the probability density for our preferred analysis. On the sides, we show the corresponding distributions for the two masses. We delimit the 90% credible intervals with vertical bars. Finally, red lines denote the ends of the PISN gap. The stars denote the maximum likelihood masses.

First, as expected from Table I, we note that green and orange contours overlap well and that the corresponding max $\mathcal{L}$ masses lay close to each other. This, together with Table I indicates that we correctly reproduce the LVK results. Second, while our blue contour encompasses the LVK one, it also contains a secondary region of high probability density that extends to larger mass-ratio regions, with our max $\mathcal{L}$ masses residing deep in such region. Third, while the Bayes Factors in the bottom row of Table I show no strong preference for any of these analyses, we note that our best-fitting waveform yields a much larger max $\mathcal{L}$ and signal-to-noise ratio (SNR) than both the LVK analysis and our reproduction of it. In particular, the latter yields a $\ln(\text{max } \mathcal{L}) = 114.1$ that is 5.1 units smaller than the $\ln(\text{max } \mathcal{L}) = 119.2$ we obtain $\Box$ Similarly, $q$-analyses also report lower max $\mathcal{L}$. This showcases that, while not discarded by their Bayes’ Factors, these analyses can miss even the highest-likelihood regions of the parameter space due to their prior settings, potentially missing important information about the source. We understand that this, together with the lack of observational constraints in this region of the parameter space, justifies the investigation of alternative priors that can explore such regions and the following detailed analysis of the corresponding results.

![Diagram showing posterior distributions for the component masses of GW190521.](image)

FIG. 1. Posterior distributions of the component masses of GW190521. The blue, green, and orange contours denote the 90% credible regions respectively obtained by us using a mass-ratio prior uniform in $Q \in [1, 4]$, by the LVK analysis and by us when using the LVK prior settings. The respective lines in the side plots denote symmetric 90% credible intervals around the median, while the red ones denote the limits of the PISN gap. The stars denote the maximum likelihood masses.

![Diagram showing posterior distributions for the magnitude and tilt of the GW190521 spins.](image)

FIG. 2. Posterior distributions for the magnitude and tilt of the GW190521 spins. The radial coordinate represents the dimensionless spin magnitudes $a_{1,2}$. The angle denotes the spin tilts $\theta_{1,2}$, defined as the angle between the spin and the orbital angular momentum of the binary at a reference frequency $f_{\text{ref}} = 11\text{ Hz}$. A zero tilt denotes alignment with the orbital angular momentum. Non-zero tilts induce orbital precession. All pixels have equal prior probability. When replicating the LVK setup, we obtain results consistent with those in [1].

---

$^3$ This is, our best-fit template is $e^{5.1} \simeq 164$ times more likely given the detectors data.
b. Source parameters

We infer individual masses \( m_1 = 106^{+36}_{-28} \ M_\odot \) and \( m_2 = 61^{+21}_{-19} \ M_\odot \). This yields a mass ratio \( Q = 1.71^{+1.50}_{-0.65} \) and total mass \( M_{\text{tot}} = 168^{+29}_{-27} \ M_\odot \), both larger than those reported in [1]. If we assume that the PISN gap spans the range [65, 130] \( M_\odot \), the primary black hole (BH) has a 15% probability of being above the gap while the secondary has a 65% probability of residing below it. The probability that both BHs are out of the gap is 13%, in high contrast with the 0.001% reported in [1]. We also infer a larger orbit inclination \( \iota = 70^{+4}_{-4} \) deg than reported in [1]. Altogether this translates into a much weaker source (See Appendix C) that yields a significantly closer distance of \( 2.8^{+3.4}_{-1.7} \) Gpc and a smaller redshift \( z = 0.49^{+0.44}_{-0.26} \) that peaks at the value of ZTF19abanrhr (see later).

Fig. 2 shows the posterior distribution for the magnitude and the tilt of spins w.r.t. orbital angular momentum. Because our mass-ratio posterior extends to larger values than the LVK, we find the data to be more informative about the primary spin but we can barely retrieve information about the secondary [17]. Similar to [1], we find that, while largely unconstrained, the spin magnitudes \( a_{1,2} \) peak near the Kerr limit \( a = 1 \). We constrain the tilt of the primary spin to \( \theta_1 = 107^{+36}_{-53} \) deg. Tilted spins – i.e. with non-zero components within the orbital plane – induce orbital precession via spin-orbit coupling [18, 49]. The impact of precession in the waveform is captured by effective precessing spin \( \chi_p \) [33, 34], for which we obtain a posterior distribution consistent with the LVK. Finally, we obtain a 71% probability for the off-plane component to be anti-aligned with the orbital angular momentum. Combining this with the barely informative posteriors on the secondary spin yields a lower effective-spin parameter \( \chi_{\text{eff}} = -0.13^{+0.34}_{-0.33} \).

We use the SurfinBH package [52] and the fit NRSur7dq4Remnant [36] to estimate the properties of the remnant BH from the BBH ones. We obtain a final spin \( \lambda_f = 0.65^{+0.12}_{-0.23} \) and a final mass \( M_f = 162^{+29}_{-28} \ M_\odot \) with zero support below \( 100 \ M_\odot \) thus maintaining its original interpretation as an intermediate-mass black hole [1].

c. Association with ZTF19abanrhr

Fig. 3 shows the one and two-dimensional posterior distributions for the location of GW190521 in Cartesian coordinates. The black lines denote the location of ZTF19abanrhr. The coordinates are defined in the principal axes of the sky-map.

We assess the probability of a true coincidence. Using the formalism in [27, 53], we compute the odds of a common-source hypothesis, \( C \), against a random coincidence, \( R \), as \( O_{C/R} = \pi_{C/R} \times I_\theta \). Here, \( I_\theta \) (or three-dimensional overlap integral) denotes the ratio between the posterior probability density (derive from GW190521 data \( d_{GW190521} \)) of the sky-location and distance, and the common-source hypothesis prior; both evaluated at the parameter values \( \theta_{\text{ZTF}} \) of ZTF19abanrhr. This is

\[
I_\theta = \frac{p(\theta_{\text{ZTF}} \mid d_{GW190521}, C)}{\pi(\theta_{\text{ZTF}} \mid C)},
\]

where \( \theta = \{ \text{RA, DEC, } d_L \} \). The factor \( \pi_{C/R} \) denotes the prior odds against a true coincidence which, as in [27], we set this to \( \pi_{C/R} = 1/13 \), which is determined by the number of flares similar to ZTF19abanrhr in the ZTF alert stream. Finally, as in [53], we re-weight our posterior samples to consider a distance prior uniform in co-moving volume and source-frame time [54]

\[
p(z) \propto dV(z)/[dz(1+z)].
\]

We assume that the PISN gap spans the range [65, 130] \( M_\odot \), the primary black hole (BH) has a 15% probability of being above the gap while the secondary has a 65% probability of residing below it. The probability that both BHs are out of the gap is 13%, in high contrast with the 0.001% reported in [1].
For our reproduction of the LVK analysis, we obtain a weak evidence $O_{C/R}^{\text{LVK}, \text{ours}} = 2.9$, qualitatively equivalent to the value $O_{C/R}^{\text{LVK}} = 2.4$ obtained by [27]. In contrast, due to the much better sky-location and distance agreement from our $Q$-analysis, we obtain a strong odds of $O_{C/R} = 72$. To check the robustness of this result, we repeated this analysis by extending the waveform model up to $Q = 6$ – out of its calibration region – obtaining similarly strong evidence of $O_{C/R} = 47$. As a final check, we obtain values close to those in [27] when using mass-ratio priors uniform in $q \leq 1$. Finally, in Table I in Appendix C we report the values of individual overlap integrals for the distance and the angular sky-location (RA, DEC).


\textbf{d. Physical interpretation} BBHs can efficiently merge in AGN disks as a result of gas torques and dynamical encounters [53, 55]. EM radiation must be produced in such mergers and may be detectable even against bright AGN disks, only if such disks are very thin and have relatively low luminosity or if the merger product is kicked out of the optically thick mid-plane and accretes at highly super-Eddington rates while also generating a jetted outflow [28, 69]. Keplerian orbital velocities for BBH in AGN disks span $O(10^{13.4}\text{km/s})$, so modest kicks at an angle to BBH orbital angular momentum can lead to significant orbital perturbation and emergence from the disk, but not escape from the nucleus. From [23], constraints on a BBH merger come from flare start and end times, flare luminosity, and color changes (if any).

BBH mergers at time $t_{\text{GW}}$ will tend to occur in the optically thick mid-plane. For typical disk models, a flare can only be observed if the kicked merger product is in the optically thin disk atmosphere. So, the flare begins after a flight through the optically thick parts of the disk. Flare start time, therefore, depends on kick velocity, disk scale height, and the resultant of the orbital velocity ($v_{\text{orb}}$) and the kick velocity ($v_{\text{kick}}$), which in turn depends on the merger location in the disk. The flare ends at $t_{\text{end}}$ after the mass of gas within the sphere of influence of the remnant drops below the amount required to supply the large mass accretion rate required to power the flare. Total distance travelled is $v_{\text{kick}}(t_{\text{GW}} - t_{\text{end}})$, and assuming a mid-plane merger, the merger product leaves the merger site at angle $\theta = \sin^{-1}(v_{\text{kick}}/v_{\text{orb}})$.

We can also reasonably assume that a BBH formed due to migration within the disk plane will form with its orbital angular momentum parallel or anti-parallel to the local plane of the AGN disk. As the BBH semi-major axis shrinks, there is a competition between gas torques which tend to maintain the binary orbit in a co-planar orientation and dynamical interactions with the spherical nucleus component, which can drive the BBH out of the plane [60]. Once the orbit is small enough that gas hardening torques become negligible, other torques (notably GR-induced) can cause precession of the binary orbit plane, so the orbit may not be precisely co-planar.

The re-interpretation of GW190521 presented here suggests several interesting possibilities for the EM flare association. First, the host AGN is unobscured and optically bright and therefore likely viewed at an angle $\lesssim 60^\circ$ to the observer (i.e., a Type I AGN). In contrast, when restricting to the sky-location of the AGN, we find that the orbit, which should be nearly aligned with the AGN, is viewed at $i = 80_{-30}^{+20}\text{deg}$. Such viewing angle suggests either that the inner AGN disk is warped [61] and that the outer disk is misaligned at $\sim 80^\circ$ to line-of-sight, or that strong dynamical encounters with the spherical nuclear component are common [62, 63], most likely in a short-lived (< 5 Myr) dense AGN disk. Second, the larger mass inferred here for $m_1$ in GW190521 lends greater support to an AGN origin for this merger. A primary mass $> 10^2 M_\odot$ is itself plausibly a fourth-generation BH from hierarchical mergers, and such mergers can only occur in galactic nuclei, most likely AGN [64].

\textbf{e. Hubble constant} The expansion rate of the Universe is parametrised by the Hubble constant $H_0$, which relates the distance of a source with its recession velocity from Earth. Because GW and EM waves respectively provide distance and redshift estimates, the joint observation of a source enables an estimate of $H_0$ independent of those purely based on EM information [67, 69]. This was first performed with the joint observation of the neutron-star merger GW170817-GRB170817A [70, 72], yielding a value $H_0 = 70.0_{-8.0}^{+12.0}\text{km s}^{-1}\text{Mpc}^{-1}$ [23]. Also e.g. [74, 75] performed estimates of $H_0$ using GW190521-ZTF19abanrhr based on the LVK parameter estimates for GW190521 [8]. Following the methods in [73, 74], we perform this measurement using our results for GW190521.
We consider GW selection effects according to two different BBH merger rates: one constant up to a redshift $z = 10$ and one following the Madau-Dickinson star-formation rate $\psi(z) \propto (1 + z)^{2.7}$ [65]. For both cases, we obtain $H_0 = 73^{+24}_{-15}\text{km s}^{-1}\text{Mpc}^{-1}$ at the 68% credible level. This is in sharp contrast to the value of $H_0 = 61^{+17}_{-20}\text{km s}^{-1}\text{Mpc}^{-1}$ obtained from the LVK results. Combining this with the kilonova GW170817-GRB170817A yields a joint estimate of $H_0 = 72.1^{+10.6}_{-6.4}\text{km s}^{-1}\text{Mpc}^{-1}$, shown in Fig. 6. This implies a 15% uncertainty improvement over the GW170817-GRB170817A estimate.

**Discussion.** GW190521 has challenged the fields of gravitational-wave astronomy and astrophysics. Its “simplest” interpretation of being a quasi-circular BBH required waveform models not available at the time of its detection [28, 36, 77–79] and poses puzzles such as a merging BH populating the PISN gap. Also, its short signal length makes its interpretation very sensitive to prior assumptions. Along the lines of [10, 20] we have shown that usage of standard prior assumptions in Bayesian inference can highly suppress parameter regions of high likelihood.

Removing such suppression, we have shown that GW190521 is consistent with a quasi-circular BBH with a larger mass ratio, larger inclination, and much closer distance than originally reported by the LVK. More importantly, the sky location of GW190521 is consistent with that of its proposed electromagnetic counterpart ZTF19abanrhr. We obtain an odds of $O_{C/R} = 71$ favouring a true coincidence in high contrast the $O_{C/R} = 2.8$ reported for the original LVK results [1, 27]. This result points, for the first time, to GW190521 as a strong candidate for the first multi-messenger observation of a compact merger with masses in the black-hole range.

An association between GW190521 and an EM counterpart leads to several astrophysical constraints. First, in order to explain the flare, the kicked, merged BH must be accreting at super-Eddington rates, with radiation escaping in a jet. Second, the orientation of the binary

---

5 To compute the likelihood for the distance conditioned to the sky-location of ZTF19abanrhr $p(d_L | \theta_{ZTF})$ we fit the posterior probability density $p(\text{RA}, \text{DEC}, d_L)$ with a Gaussian Mixture Model (GMM) and compute $p(\text{RA}_{ZTF}, \text{DEC}_{ZTF}, d_L)$. We re-weight by the appropriate distance prior when needed.

6 For the LVK runs, the prior $\pi(d_L)$ or $d_L^2$ is removed by the projection along the line-of-sight and no re-weighting is needed. For our runs, we re-weight by $d_L^2 / \pi(d_L)$ where $\pi(d_L)$ is our prior uniform in co-moving volume.
implies that either the outer disk of the AGN (where the merger occurred) is strongly warped compared to the inner disk, or dynamical interactions with the spheroidal component are common. The latter is more likely in a short-lived, dense AGN disk. Third, the higher primary mass combined with a break in the LVK observed mass function around 35 M⊙ is suggestive of a high generation primary that could only merge in a galactic nucleus.

Finally, we note that this study is limited to the “standard” scenario of a quasi-circular BBH. However, GW190521 has also been shown to be consistent an eccentric merger [11,13] of BBHs and even an exotic headon merger of Proca-stars [14]. The future development of waveform models accounting for both precession and eccentricity shall yield further clues on the true nature of GW190521 and the robustness of our results.
ACKNOWLEDGEMENTS

We thank Will Farr, and Greg Ashton for the respective public release of their codes to compute the value of the Hubble constant [80] and the coincidence odds [27] as well as for valuable discussions. We also thank Héctor Estelles, Sergei Ossokine, and Vijay Varma for their help to understand the impact of possible waveform systematics and Leo Singer for his advice on the usage of his ClusteredKDE code [81] to compute the values in Table II. The analyzed data is publicly available at the online Gravitational Wave Open Science Center [39]. LVK results quoted throughout the paper and the corresponding histograms and contours have made use of the publicly available sample release in https://dcc.ligo.org/P2000158-v4. JCB is supported by a fellowship from “la Caixa” Foundation (ID 100010434) and by the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 847648. The fellowship code is LCF/BQ/P120/11760016. JCB is also supported by the research grant PID2020-118635GB-I00 from the Spain-Ministerio de Ciencia e Innovación. This work has received financial support from Xunta de Galicia (Centro singular de investigación de Galicia accreditation 2019-2022), by European Union ERDF, and by the “María de Maeztu” Units of Excellence program MDM-2016-0692 and the Spanish Research State Agency. KC acknowledges the MHRD, Government of India, for the fellowship support. SHWL is partially funded by the Department of Physics at The Chinese University of Hong Kong. BM and KESF are supported by NSF AST-1831415 and Simons Foundation Grant 533845. We acknowledge the use of IUCAA LDG cluster Sarathi for the computational/numerical work. The authors acknowledge computational resources provided by the CIT cluster of the LIGO Laboratory and supported by National Science Foundation Grants PHY-0757058 and PHY0823459, and the support of the NSF CIT cluster for the provision of computational resources for our parameter inference runs. This material is based upon work supported by NSF’s LIGO Laboratory which is a major facility fully funded by the National Science Foundation. This manuscript has LIGO DCC number P2100467.

Appendix A: Maximum Likelihood waveforms

Fig. 7 shows the whitened maximum likelihood waveforms corresponding to the three analyses shown in Fig. 1 in the main text together with the corresponding detector strain data. We show our maximum likelihood waveform in red while those in black and orange correspond to those of the LVK analysis and our reproduction of it.

We note two main aspects. First, while the black and yellow waveforms overlay well during the whole signal duration, ours differ visibly in the Hanford and Livingston cases in two main aspects. These are a steeper signal rise before the merger and a different height at the signal peaks. Second, for the case of Virgo, our waveform shows some complex structure (see between $t = 0.425$ and $t = 0.450$) that is not present in the other two. These differences, visible by eye, are behind the highly differing maximum likelihood values reported in Table I.

1. Comparison with Nitz et. al and Estelles et. al.

a. Phenomenological models

As noted in the main text, Nitz and Capano [10] and Estellés et al. [20] did also analyse GW190521 using a mass prior uniform in mass ratio $Q \geq 1$ but reached a different conclusion on the potential coincidence with ZTF19abanrhr. The analysis of [20] and part of that of [10] where conducted with phenomenological waveform models that are not calibrated to numerical simulations of precessing black holes and therefore may not accurately account for all the physics in such systems. Instead, these waveform models are calibrated to simulations of non-precessing binaries and precession is then accounted for by “twisting” the resulting waveforms using analytical expressions whose validity breaks during the merger and ringdown regimes. In addition, for instance, these models assume a conjugation symmetry $h_{\ell,m} = (-1)^m h^{\ast}_{\ell,-m}$ between right-handed and left-handed GW emission modes which is not present in precessing systems. For these reasons, we argue that these models should not be taken as a reference when analysing a system like GW190521 for which precession is a relevant effect and for which higher-harmonics can play an important role due to its high mass [82, 84].

As a cross-check, we performed our reproduction of the above two studies by repeating our runs using the models IMRPhenomXPHM used by Nitz and Capano [10] and IMRPhenomTPHM used by Estellés et al. [20]. We did this for mass ratio priors uniform in $Q \in [1, 4]$ and $Q \in [1, 6]$ (to allow for an apples-to-apples comparison with our results) and for $Q \in [1, 10]$. In all cases, we obtain weak evidence for a true coincidence with ZTF19abanrhr. Finally, we showcase the vastly different predictions that these models make at the points in the parameter space we are exploring. Fig. 8 shows the waveforms computed with NRSur7dq4, IMRPhenomTPHM and IMRPhenomXPHM at the maximum likelihood parameters values reported by NRSur7dq4 together with the corresponding Livingston strain data. It is clear that both phenomenological models predict extremely different waveforms. Consequently, it is not surprising that they lead to very different conclusions.
First, while we do not recover a peak in the mass-ratio distribution at \( Q > 4 \) (most likely due to our still lower number of live points) but we find that the maximum likelihood point lies beyond \( Q = 4 \). Second, when restricting the sky-location to ZTF19abanrhr, we find a bimodal distribution with a strong peak at \( d_L \simeq 1000 \) Mpc that is not present in 4k-live-points analysis presented in the main text. We show this feature in Fig. 8 in the blue-solid curve. The dashed and dotted curves show that this feature mostly comes from the beyond-calibration \( Q > 4 \) region. We note, nevertheless, that even in the presence of this secondary peak, we obtain an odds ratio of 50 : 1 in favour of the coincidence hypothesis, consistent with the 47 : 1 we obtain using 4k live points.

Appendix B: Detailed coincidence overlap integrals

In the main text we have reported the full three-dimensional odds-ratio \( \mathcal{O}_{C/R} \) for a true coincidence between GW190521 and ZTF19abanrhr, computed as \( \mathcal{O}_{C/R} = I_\theta \times \pi_C / \pi_R \) [27]. Here, \( \pi_C / \pi_R \) denotes the prior odds while \( I_\theta \) is known as the overlap integral.

This is equal to the ratio between the probability for the distance and sky-location \( \theta = (d_L^{ZTF}, \Omega^{ZTF}) \) of ZTF19abanrhr given the corresponding posterior distributions \( p(d_L, \Omega) \) obtained from the analysis of the GW190521 data \( d_{GW190521} \) and the corresponding prior probability \( \pi(d_L, \Omega) \). In particular one can express:

\[
I_\theta = \int \frac{p(\theta \mid d_{GW190521}, C) p(\theta \mid d_{ZTF}, C)}{\pi(\theta \mid C)} \, d\theta. \tag{B1}
\]

Since the distributions for the ZTF parameters can be considered as delta functions, we can simplify the above to:

\[
I_\theta = \int \frac{p(\theta \mid d_{GW190521}, C)}{\pi(\theta \mid C)} \, d\theta.
\]
FIG. 9. Impact of live points on distance estimates conditioned to the location of ZTF19abanhr. We show posterior distributions for the distance of GW190521 conditioned to the sky-location of ZTF19abanhr for analyses conducted with different number of live points and for different cuts in the mass-ratio Q.

\[
I_\theta \equiv I_{dL,\Omega} = \frac{p(d^\text{ZTF}_L, \Omega | d_{\text{GW190521}})}{\pi(d^\text{ZTF}_L, \Omega | C)}
\]

and assuming \(d_L\) and \(\Omega\) are independent,

\[
I_{dL,\Omega} \simeq \frac{p(d^\text{ZTF}_L | d_{\text{GW190521}})}{\pi(d^\text{ZTF}_L | C)} \frac{p(\Omega | d_{\text{GW190521}})}{\pi(\Omega | C)} \]

\[= I_{dL} I_\Omega.\]

Here \((I_{dL}, I_\Omega)\) denote the “partial” overlap integrals. For a detailed derivation, please see [27]. In Table II we show the values of these integrals, together with the full three-dimensional ones of the various analyses presented in this work. It is evident that the \(Q\)-analyses report values for the individual integrals are larger than those of the remaining analyses by factors of, respectively \(\simeq 2\) and \(\simeq 4\). Finally, for completeness, we re-calculated the LVK results are reported in [27] (see next).

1. Accuracy of overlap integral \(I_{dL,\Omega}\) and corresponding odd-ratios

The calculation of \(I_{dL,\Omega}\) involves an estimation of the joint three-dimensional probability density distribution for the sky-angles and luminosity distance \(p(d_L, \Omega)\). To this, both the authors of [20, 27] and us used the multi-dimensional Kernel Density Estimator (KDE) ClusteredKDE described in [33]. A crucial element of such KDE estimation is the maximum number of clusters \(k_{\text{max}}\) that can be defined within the sample set. The algorithm finds then an optimal number of clusters \(k < k_{\text{max}}\) that maximises the Bayesian Information Criterion, or BIC. To ensure the robustness of our results, we computed our \(I_{dL,\Omega}\) (or equivalently, \(O_{C/R}\)) for varying values of \(k_{\text{max}} \in [5, 160]\). Fig. 10 shows the percent difference between \(I_{dL,\Omega}\) computed for different \(k_{\text{max}}\) and \(k_{\text{max}} = 160\). The two pairs horizontal of lines denote respectively denote \(\pm 5\%\) and \(\pm 10\%\) differences. For reference, the vertical lines denote the values \(k_{\text{max}} = 15\) used in [27] and \(k_{\text{max}} = 15\) used by default in [81].

We highlight two main aspects. First, we note that large variations happen for \(k_{\text{max}} \leq 30\). For instance, the value obtained from the LVK samples for \(k_{\text{max}} \leq 20\) is twice that reported in [27], which we label by LVK Ashton et. al. in Table II. This, nevertheless, does not change the conclusion that evidence for a true coincidence results.

| Analysis         | \(I_{dL}\) | \(I_\Omega\) | \(I_{dL,\Omega}\) | \(O_{C/R}^{\text{D}}\) |
|------------------|-----------|-------------|----------------|-----------------|
| LVK [27]         | 1.75      | 29.26       | 31.3           | 2.4             |
| LVK Priors       | 1.75      | 29.26       | 61.1           | 4.6             |
| Q \(\in [1, 4]\) | 4.07      | 118.83      | 937.1          | 72.0            |
| Q \(\in [1, 6]\) | 3.54      | 86.59       | 611.7          | 47.0            |
| Q \(\in [1/4, 1]\)| 1.63      | 12.96       | 44.2           | 3.4             |
| Q \(\in [1/6, 1]\)| 1.98      | 27.06       | 50.0           | 3.9             |

TABLE II. Detailed GW190521 vs ZTF coincidence results. Overlap integrals for the sky-location of GW190521 its potential counterpart and odds-ratio in favour of the coincidence hypothesis. The last five rows report values for our different analyses of GW190521, that of our reproduction of it. The first two show the values reported by [27] using the original LVK results, and the second shows our recalculation of these values.
between GW190521 and ZTF19abanhr under the LVK results is weak. Second, we show that our results vary by less than $\simeq 5\%$ for $k_{\text{max}} \geq 80$ up to our reference value of $k_{\text{max}} \geq 160$. The exception to this $Q \leq 6$ analysis for which there are excursions towards $+10\%$ levels that, in any case, would not change our qualitative conclusions.

Appendix C: Intrinsic source loudness

As mentioned in the main text, the decrease in luminosity distance reported by our analysis is due to a decrease in the estimated intrinsic loudness of the sources. This is due to the fact that the LVK analysis uses a prior that favors intrinsically louder $q \approx 1$ sources. Fig 11 shows the posterior distribution of the intrinsic loudness from each analyses, computed as the product of the optimal SNR of the corresponding waveform multiplied by the corresponding estimated luminosity distance, as in [88]. It is clear that, in fact, our analysis shows a strong preference for weaker sources.

FIG. 11. Posterior distribution for the intrinsic loudness of the source of GW190521. This is computed as the product of the optimal network signal-to-noise ratio and the luminosity distance of the posterior samples. The LVK analysis strongly prefers sources that are louder by a factor of $\simeq 3$, explaining the vast difference between the corresponding distance estimates.
Subtracting Narrow-band Noise

A. Viets and M. Wade, "Subtracting Narrow-band Noise from LIGO Strain Data in the Third Observing Run," Tech. Rep. DCC-T2100058 (LIGO, 2021).

R. J. E. Smith, G. Ashton, A. Vajpeyi, and C. Talbot, Monthly Notices of the Royal Astronomical Society 498, 4492 (2020).

G. Ashton et al., Astrophys. J. Suppl. 241, 27 (2019).

J. S. Speagle, Monthly Notices of the Royal Astronomical Society 493, 3132 (2020).

J. Veitch, M. Purrrer, and I. Mandel, Phys. Rev. Lett. 115, 141101 (2015) [arXiv:1503.05953 [astro-ph.HE]].

I. Mandel, V. Raymond, C. Roever, M. van der Sluis, and J. Veitch. https://lscsoft.docs.ligo.org/lalsuite/lalinferece/_j.a_l.inference.nest_8c_source.html (2009).

G. Ashton et al., Phys. Rev. D 102, 044055 (2020).

J. Veitch et al.,Mon. Not. Roy. Astron. Soc. 460, L65 (2017).

J. M. Bardeen and J. A. Petterson, Astrophysical Journal Letters 195, L5 (1975).

W. Zheng, "A comprehensive measurement of the local value of the Hubble constant with 1 km/s/Mpc uncertainty from the Hubble space telescope and the SH0ES value of the Hubble constant with 1 km/s/Mpc uncertainty from the Hubble space telescope and the SH0ES team." (2021), arXiv:2112.04510.

B. Abbott et al., Physical Review Letters 119 (2017), 10.1103/physrevlett.119.161101.

J. Veitch et al., Mon. Not. Roy. Astron. Soc. 425, 460 (2012) [arXiv:1206.2309 [astro-ph.GA]].
