Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **509,** 3957–3965 (2022) https://doi.org/10.1093/mnras/stab2977
Advance Access publication 2021 October 23

Optimized localization for gravitational waves from merging binaries

Zhi-Qiang You [1,2,3,4] Gregory Ashton [2,3,5★] Xing-Jiang Zhu [2,3,4] Eric Thrane[2,3]
and Zong-Hong Zhu [1,4,6]

[1] *School of Physics and Technology, Wuhan University, Wuhan 430072, China*
[2] *School of Physics and Astronomy, Monash University, Clayton, VIC 3800, Australia*
[3] *OzGrav: The ARC Centre of Excellence for Gravitational Wave Discovery, Clayton, VIC 3800, Australia*
[4] *Advanced Institute of Natural Sciences, Beijing Normal University at Zhuhai, Zhuhai 519087, China*
[5] *Department of Physics, Royal Holloway, University of London, Egham TW20 0EX, UK*
[6] *Department of Astronomy, Beijing Normal University, Beijing 100875, China*

Accepted 2021 October 13. Received 2021 October 13; in original form 2021 May 10

ABSTRACT

The Advanced LIGO and Virgo gravitational-wave observatories have opened a new window with which to study the inspiral and mergers of binary compact objects. These observations are most powerful when coordinated with multimessenger observations. This was underlined by the first observation of a binary neutron star merger GW170817, coincident with a short gamma-ray burst, GRB170817A, and the identification of the host galaxy NGC 4993 from the optical counterpart AT2017gfo. Finding the fast-fading optical counterpart critically depends on the rapid production of a sky map based on LIGO/Virgo data. Currently, a rapid initial sky map is produced, followed by a more accurate, high-latency, $\gtrsim$12 h sky map. We study optimization choices of the Bayesian prior and signal model, which can be used alongside other approaches such as reduced order quadrature. We find these yield up to a 60 per cent reduction in the time required to produce the high-latency localization for binary neutron star mergers.

Key words: gravitational waves – neutron star mergers.

1 INTRODUCTION

Transient multimessenger astronomy is delivering new insights into the nature of compact objects, the cosmological properties of our Universe, and general relativity. During their second observing run (O2), the Advanced LIGO (Aasi et al. 2015) and Virgo (Acernese et al. 2015) detectors observed gravitational-wave emission from a binary neutron star (BNS) merger, GW170817 (Abbott et al. 2017a). This observation coincided with a short gamma-ray burst, GRB170817A (Goldstein et al. 2017), and follow-up campaigns across the electromagnetic spectrum (Abbott et al. 2017b) led to the observation of an optical counterpart AT2017gfo (Perego, Radice & Bernuzzi 2017; Valenti et al. 2017; Yang et al. 2017) and identification of the host galaxy NGC 4993. This multimessenger view of the event provided critical insights into multimessenger astronomy, opening a new path by which to study and understand the mergers of neutron star binaries, short gamma-ray bursts, and their optical counterparts.

The LIGO and Virgo observatories, joined by KAGRA (Aso et al. 2013), have now completed their third observing run (Abbott et al. 2021a). During this observing run, open public alerts were issued (see emfollow.docs.ligo.org), enabling numerous follow-up campaigns (see e.g. Coughlin et al. 2019; Ackley et al. 2020; Antier et al. 2020; Gompertz et al. 2020). So far, there has yet to be a bona fide multimessenger observation from the O3 observing run [however, see Graham et al. (2020) and Ashton et al. (2020) for a discussion of a speculative connection between the binary black hole (BBH) merger GW190521 (Abbott et al. 2020b) and an active galactic nucleus flare]. In preparation for the fourth observing run (estimated to be 2022 June), the detectors are currently being upgraded with a projected BNS inspiral range of 190 Mpc (The LSC-Virgo-KAGRA Observational Science Working Groups 2020). During this observing run, the Advanced LIGO, Virgo, and KAGRA (HLVK) network is likely to observe tens of transient systems containing a neutron star and hundreds of BBH systems. BNS mergers are known to produce a multimessenger counterparts (Abbott et al. 2017b) known as a ‘kilonova’ (Li & Paczyński 1998). As yet, no unambiguous detection of a neutron star–black hole (NSBH) binary has been made; however, such systems may also produce an electromagnetic counterpart (Lattimer & Schramm 1974; Li & Paczyński 1998, see also Fernández & Metzger 2016 for a review). As such, mergers containing a neutron star tend to be prioritized by follow-up campaigns; however, as demonstrated by Graham et al. (2020), the next big surprise may instead come from a BBH merger.

Online gravitational-wave searches are able to identify events in the data and analyse their significance in a time-scale of $\mathcal{O}(\mathrm{s})$. Following the identification, the event is automatically vetted and published in a GCN notification in a time-scale of $\mathcal{O}(10\,\mathrm{s})$. Optical telescopes then perform (often automated) searches for the rapidly fading, time-scales of $\mathcal{O}(\mathrm{h})$, electromagnetic transient. The ability to identify the transient is highly dependent on the three-dimensional (3D) source localization. [Here, we parametrize in terms of the right ascension (RA), declination (Dec.), and luminosity distance

---

★ E-mail: gregory.ashton@ligo.org

© 2021 The Author(s)
Published by Oxford University Press on behalf of Royal Astronomical Society

d_L.] In the first three observing runs, results from two methods for producing localizations were published. First, the *low-latency* BAYESTAR (Singer & Price 2016; Singer et al. 2016) algorithm produces a 3D localization in $\mathcal{O}$(min) utilizing the maximum-likelihood template from the search pipelines. Secondly, the *high-latency* LALINFERENCE (Veitch et al. 2015) algorithm uses stochastic sampling to construct the full posterior probability distribution and produces results in a time-scale of $\mathcal{O}(>12\ \mathrm{h})$. While for the majority of systems, differences between the low-latency and high-latency localization are anticipated to be small (Abbott et al. 2018a), the high-latency localization is preferable since it can include an improved physical description of the signal and noise. Moreover, a systematic study (Morisaki & Raymond 2020), comparing the BAYESTAR algorithm with a stochastic-sampling-based approach (discussed below), demonstrates that BAYESTAR can overestimate the localization uncertainty when the best-fitting parameters of the signal are outside the online-detection-pipeline template bank. This can happen, for example, if the online pipeline does not include the spin precession of the source.

Constructing the full posterior probability distribution is a computationally challenging task. Stochastic sampling methods such as Markov chain Monte Carlo (Metropolis et al. 1953; Hastings 1970) and nested sampling (Skilling 2006) require $\sim 10^6$–10^8 evaluations of the likelihood to analyse gravitational-wave signals (Christensen & Meyer 1998; Veitch & Vecchio 2008). Typically, each evaluation of the likelihood is dominated by the time required to model the source. This time varies between a few milliseconds and several tens of seconds, depending on the signal duration and sophistication of the waveform model. As such, the *wall time* required to draw a sufficient number of samples to approximate the posterior can be between several hours and many tens of days. A number of advances have been made in reducing the wall time:

(i) *Reduced-order-quadrature (ROQ) methods* (Antil et al. 2012; Canizares et al. 2013, 2015; Smith et al. 2016; Qi & Raymond 2021) interpolate the likelihood to high accuracy and can speed up evaluation times by factors of several hundred. While ROQ-based methods have enjoyed considerable success (see e.g. Abbott et al. 2020c), they do require that the ROQ basis be pre-constructed, often at a significant computational cost. As such, their utility can be limited for online production of the 3D localization if the pre-computed basis set does not cover the required parameter space. Morisaki & Raymond (2020) recently demonstrated that so-called focused-ROQ (FROQ), in which many bases covering narrow ranges of the parameter space offer greater speed-ups still: with gains of up to 10^4 seen for low-mass systems.

(ii) *Heterodyned likelihoods* (Cornish 2010, 2021), also known as the *relative-binning* method (Zackay, Dai & Venumadhav 2018; Finstad & Brown 2020), exploit the computation for likelihoods of similar waveforms, whose phases and amplitudes differ smoothly with frequency, by pre-computing frequency-binned overlaps of the best-fitting waveform with the data. These methods do not require a pre-computation step and offer speed-ups of up to $\sim 10^4$ in likelihood evaluation times. The accuracy of these approaches depends on the expansion order: Just a few terms are required to sufficiently approximate the likelihood. Demonstrations of this method are very promising, but work is needed to verify the accuracy and limitations of the method against the full likelihood.

(iii) Then, there is *brute-force parallelization*. Parallelization can be done at the level of the likelihood itself (Talbot et al. 2019), the stochastic sampler, or using multiple independent stochastic samplers. The latter two aspects have been generously employed in standard inference packages (Veitch et al. 2015; Ashton et al. 2019; Biwer et al. 2019) using the few tens of cores available on typical central processing units, while Smith et al. (2020) demonstrated the capacity to scale to the many hundreds of cores available in high-performance computers using the `dynesty` (Speagle 2020) nested sampling algorithm. Such approaches are useful as they do not require pre-computation and make no requirements about the waveform itself. Unlike the other techniques to reduce wall time, this technique is not 'free' (achieved through clever design) – it requires additional computing resources.

(iv) The RIFT family of stochastic samplers (Pankow et al. 2015; Lange, O'Shaughnessy & Rizzo 2018) employs aspects of brute-force parallelization (with extensions to graphical processing units; Wysocki et al. 2019) alongside pre-computing aspects of the waveform in order to carry out inference with iterative fitting.

(v) Significant speed-ups may be realized by the use of machine-learning-based approaches. Such approaches do not apply the standard principles of stochastic sampling; instead, the algorithm is pre-trained on example of signals in noise and can then produce posterior samples within a few seconds (Gabbard et al. 2019; Green, Simpson & Gair 2020; Green & Gair 2021). Such algorithms present a significant opportunity as they could handle non-Gaussian noise and arbitrarily complex waveforms by developing realistic training sets. For these approaches, the optimizations discussed herein are not directly applicable, in the sense that they do not consider a specific prior or waveform model during the analysis. However, the lessons learned can be applied in selecting the complexity of training data.

(vi) Finally, there are many other approaches to speeding up stochastic sampling such as the use of machine-learning-coupled nested sampling (Williams, Veitch & Messenger 2021), adaptive frequency banding (Morisaki 2021), and representing the signals in the time–frequency domain (Cornish 2020).

In the fourth observing run, one (or many) of these approaches may be used to reduce the latency of full parameter-estimation results. However, the wall time they require (or the training time in the case of machine-learning-based approaches) still depends on the choice of waveform model and the astrophysical prior. In this work, we study how to optimize the choice of model and prior to reduce the wall time. We will develop these ideas in the context of ROQ methods and parallelization, but the ideas apply equally to many of the methods listed above.

This paper is organized as follows. In Section 2, we describe the optimization of the prior, and then in Section 3, we validate which of these optimizations produces acceptably small bias and calculate the improvement in wall time. In Section 4, we discuss optimization of the waveform model. In Section 5, we conclude with a discussion on how these choices can be used during the next observing run to minimize the time to produce localization.

2 OPTIMIZED CHOICES FOR PRIORS

A compact binary coalescence signal is described by up to 10 intrinsic parameters and 7 extrinsic parameters. The intrinsic parameters are chirp mass $\mathcal{M}$, mass ratio q, component spin magnitude (a_1 for the primary heavier-mass object and a_2 for the secondary lighter-mass object), and four angles describing the spin orientation (θ_1, θ_2, ϕ_{12}, and ϕ_{JL}). For systems containing one or more neutron stars, the intrinsic parameters also include the neutron star tidal deformability Λ_1 and Λ_2. The extrinsic parameters are the 3D localization coordinates (RA, Dec., d_L), the polarization angle of the source ψ, the GPS reference time t_{geo} of the merger, the angle between the total angular momentum and the line of sight $\theta_{\rm JN}$, and the binary phase ϕ (defined at a fixed reference frequency). [See table E1 of Romero-Shaw et al. (2020) for a detailed description of these parameter and alternative parametrizations.][1] For stochastic samplers, the number of likelihood evaluations (and hence the wall time) depends on the complexity of the posterior. As a general rule of thumb, the number of parameters is a good leading-order description of the complexity: More parameters require more likelihood evaluations and hence a longer wall time. Optimizing the choice of prior can help minimize the wall time while ensuring the result remains unbiased.

First, we discuss different options for the spin prior. In an *aligned-spin* prior, the source model is restricted to solutions excluding precession (Schmidt, Ohme & Hannam 2015). In a *low-spin* model, the magnitudes of the spin are restricted (typically to a dimensionless spin of 0.05; Abbott et al. 2018b). The upper end of the low-spin prior corresponds to a conservative limit on the effective spins of pulsars in known Galactic double neutron stars that are capable of merging within a Hubble time (Zhu & Ashton 2020). Further, one can neglect the effects of spin entirely. It was shown by Farr et al. (2016) that the 3D localization for low-spin BNS sources is unbiased given a zero-spin prior. However, black hole (BH) binaries can sustain significant dimensionless spin (see e.g. Abbott et al. 2021c).

Secondly, we examine the priors for tidal parameters. For BNS and NSBH binaries, the tidal parameters are relatively poorly measured (Abbott et al. 2018b, 2020c). A straightforward prior choice is to neglect tidal parameters (in effect, we set $\Lambda_1 = 0$ and $\Lambda_2 = 0$, i.e. that both components are BHs). The logic here is that, if we cannot accurately measure the tides, the tidal parameters probably do not have a strong effect on the sky localization.

Thirdly, we consider the prior for mass ratio. For the two confident BNS events in Gravitational-Wave Transient Catalog 2 (Abbott et al. 2021a), the masses of the two components are nearly equal. Under the astrophysical low-spin prior, the mass ratio is constrained to be between 0.7–1 and 0.8–1 for GW170817 and GW190425, respectively (Abbott et al. 2018b, 2020c). This suggests another potential prior optimization: restricting to equal-mass systems $q = 1$.

We investigate these optimization strategies. In Table 1, we list four prior distributions. A precessing, but low-spin *astrophysical prior* captures our broad expectation for the typical population parameters of neutron star binaries; it is from this prior that we draw simulation parameters. The remaining three priors combine increasing restrictions on the spin and mass ratio. For all three choices, we assume zero tidal deformability: This prior optimization was already used during the third observing run for high-latency 3D localization.

3 VALIDATION OF PRIOR OPTIMIZATION

The $a_i < 0.05$, $a_i = 0$, and $a_i = 0$, $q = 1$ priors defined in Table 1 increasingly constrain the astrophysical properties of the source. Almost certainly, these overconstrain the source properties. For example, GW190425 shows some support for spin effects and is unlikely to be equal mass while the data from GW170817 do constrain the tidal deformability. The question is: Do these unphysical prior constraints bias the 3D sky localization? To answer this question, we perform tests on simulated data and look at the localization of detections during previous observing runs that have the potential for electromagnetic counterparts.

Table 1. Definition of the astrophysical, low-spin ($a_i < 0.05$), zero-spin ($a_i = 0$), and zero-spin and equal-mass ($a_i = 0$, $q = 1$) priors investigated in this work.

	Astrophysical	$a_i < 0.05$	$a_i = 0$	$a_i = 0, q = 1$
a_1	[0, 0.05]	[0, 0.05]	0	0
a_2	[0, 0.05]	[0, 0.05]	0	0
q	[0.125, 1.0]	[0.125, 1]	[0.125, 1]	1
Λ_1	[0, 5000]	0	0	0
Λ_2	[0, 5000]	0	0	0

We define a_i and Λ_i to be the absolute spin magnitude and tidal deformability of the ith component in the binary and $q \leq 1$ to be the mass ratio. For the astrophysical prior and low-spin prior, the spins are fully precessing but restricted in magnitude. For all other parameters, the priors follow the standard distributions described in table E1 of Romero-Shaw et al. (2020).

3.1 Parameter–parameter test

For our first test, we simulate 100 BNSs using the IMRPHENOMPV2_NRTIDAL waveform (Dietrich et al. 2019). The simulated signals have spin magnitudes, mass ratios, and tidal deformability parameters drawn from the astrophysical prior (see Table 1). For the remaining parameters, we draw them from the standard astrophysical distributions.

The signals are simulated in 128 s of data using the projected O4 sensitivity coloured Gaussian noise (The LSC-Virgo-KAGRA Observational Science Working Groups 2020) for the HLV detector network. Signals with a network signal-to-noise ratio (S/N) less than 12 are discarded, with replacement; this ensures that the injection set reflects the population of events from which we are likely to observe electromagnetic counterparts: high-S/N systems with good sky localization. We analyse the simulated data sets using the IMRPHENOMPV2 (Schmidt, Hannam & Husa 2012; Hannam et al. 2014; Khan et al. 2016) waveform model (which excludes tidal deformability), an ROQ basis (Smith et al. 2016), and the DYNESTY nested sampling algorithm as implemented in BILBY (Ashton et al. 2020). The ROQ basis is limited to chirp mass values of 1.42–2.60 $M_\odot$; as such, we limit the prior distribution on chirp mass (and similarly the distribution from which we draw simulation parameters) to this range.

We find that for all choices of prior specified in Table 1, the 3D localization is unbiased. We test this using a parameter–parameter (PP) test (Cook, Gelman & Rubin 2006; Talts et al. 2018) over the 100 simulated signals. In Fig. 1, we show the results of the PP test for the most restrictive zero-spin and equal-mass ($a_i = 0$, $q = 1$) prior. Qualitatively, bias manifests in a PP plot as a deviation in the parameter curve from the diagonal: For all parameters in Fig. 1, the curves remain inside the 3σ uncertainty region. One way to quantify the bias from a PP plot is to calculate the p-value expressing the probability that the fraction of events in a particular confidence interval is drawn from a uniform distribution. We calculate the p-value using the Kolmogorov–Smirnov test as implemented in SCIPY (Jones et al. 2001). For the most restrictive $a_i = 0$, $q = 1$ prior, the p-values are 0.122, 0.551, and 0.710 for Dec., $d_{\rm L}$, and RA, respectively.

---

[1] In addition, there may be up to 20 parameters per detector to describe the detector calibration [see Cahillane et al. (2017) for a description, Romero-Shaw et al. (2020) for a discussion of how these are marginalized in the BILBY software used herein, and Payne et al. (2020) and Vitale et al. (2021) for new more physical approaches to marginalizing calibration uncertainty]. However, Payne et al. (2020) suggest that the effect of calibration uncertainty is likely to be negligible during the advanced-detector era, and that calibration uncertainty can be added to inference calculations in post-processing.

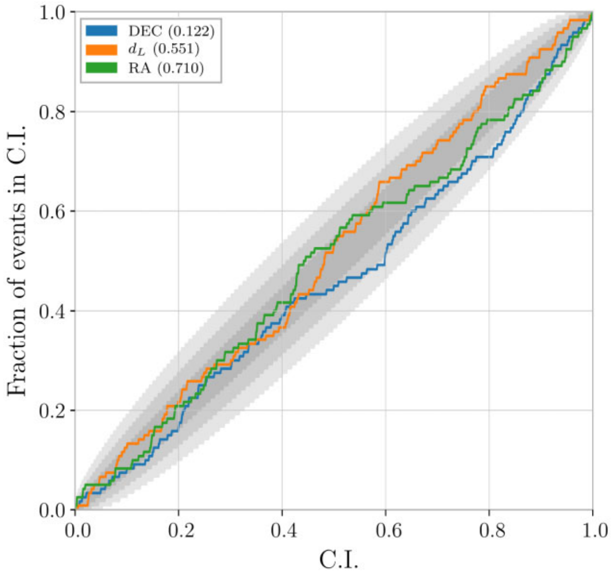

Figure 1. A PP test of the 3D sky localization parameters of 100 simulated BNS signals using the $a_i = 0$, $q = 1$ prior. We simulate signals drawn from the astrophysical prior (see Table 1) using the `IMRPhenomPv2_NRTidal` waveform model and analyse the data using the `IMRPhenomPv2` waveform model. The grey region indicates the 1σ, 2σ, and 3σ confidence intervals. Individual p-values for each of the 3D localization parameters are shown in the legend.

Table 2. The fraction of the truth value out of 90 per cent and 68 per cent credible intervals for $a_i < 0.05$ and $a_i = 0$, $q = 1$ priors for the 100 simulated BNS signals discussed in Section 3.1.

Model	Fraction	Dec.	RA	d_{L} (Mpc)
$a_i < 0.05$	Out of 90 per cent	9 per cent	12 per cent	10 per cent
	Out of 68 per cent	32 per cent	33 per cent	31 per cent
$a_i = 0, q = 1$	Out of 90 per cent	13 per cent	12 per cent	9 per cent
	Out of 68 per cent	37 per cent	35 per cent	21 per cent

That none of these p-values are small $\ll$0.05 indicates that there is no measurable sign of bias.

As an additional way to understand the bias, in Table 2, we calculate the fraction of events outside of the 90 per cent and 68 per cent credible intervals for each prior. For an unbiased inference, we would expect the fraction out of the 90 per cent and 68 per cent credible intervals to be 10 per cent and 32 per cent, respectively. The difference then provides a quantitative estimate of the bias. For the $a_i < 0.05$ prior, this is up to 2 per cent, while for the $a_i = 0$, $q = 1$ prior, it is up to 11 per cent.

In Appendix A, we show additional two PP plots for the $a_i = 0$, $q = 1$ prior for a two-detector network (Fig. A1) and for 500 injections (Fig. A2). Furthermore, in Table A1, we list the p-values and fractions of truth parameters that are out of the 90 per cent credible intervals, for 50, 100, 300, and 500 injections. These illustrate that our results are robust against the choice of detector network configuration and the number of injections.

To compare the wall time between the $a_i < 0.05$ and $a_i = 0$, $q = 1$ model, in Fig. 2, we plot the optimal S/N of the simulated signals and the sampling time. As expected for a nested sampling algorithm, the sampling time is correlated with the signal S/N.

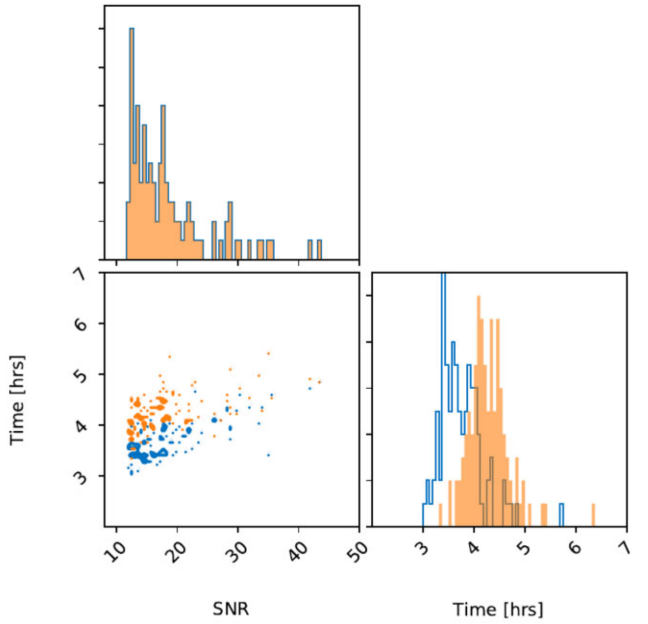

Figure 2. The S/N of simulated signals and the measured wall time in hours for the 100 simulated BNS events obtained using the $a_i = 0$, $q = 1$ (blue) and $a_i < 0.05$ (orange) priors studied in Section 3.1. All wall times are evaluated on an Intel Gold 6140 CPU processor using 32 threads.

[The sampling time is generally proportional to the ratio of prior to posterior volume (Speagle 2020); at a fixed prior volume, the posterior volume decreases as the S/N increases, leading to longer wall times.] The mean sampling times for the $a_i < 0.05$ and $a_i = 0$, $q = 1$ priors are $\sim$4.5 and 3.5 h, respectively. This demonstrates that, in addition to the low-spin savings already identified by Farr et al. (2016), utilizing an equal-mass prior can provide a further $\sim$20 per cent performance improvement.

3.2 Implications for BNS and NSBH localization

To date, the LIGO and Virgo detectors have observed two BNS events, GW170817 (Abbott et al. 2017a) and GW190425 (Abbott et al. 2020c), and two NSBH events, GW200105 and GW200115 (Abbott et al. 2021d). These systems are likely to be accompanied by a rapidly fading electromagnetic counterpart and are hence prioritized for a follow-up by optical telescopes (Fernández & Metzger 2016). As such, we have the most to gain in optimizing the production of the high-latency 3D localization.

To study our optimized priors for BNS signals, we re-analyse the public data (Abbott et al. 2021b) for the BNS events GW170817 and GW190425 using the $a_i = 0$, $q = 1$ prior and the IMRPhenomPv2 waveform model. In Figs 3 and 4, we plot the sky localization uncertainty from our re-analysis, the O3 low-latency (BAYESTAR), and the O3 high-latency (LALINFERENCE) results for GW170817 and GW190425, respectively.

High-latency results better constrain the sky localization: This can be seen in Table 3 where we report the area coverage and was demonstrated systematically by Morisaki & Raymond (2020). Table 3 demonstrates that the $a_i = 0$, $q = 1$ prior provides an equivalent constraint on the sky area to the O3 high-latency results for BNS systems. We conclude then that for BNS, any method attempting to reduce the wall time of high-latency localization results (e.g. the

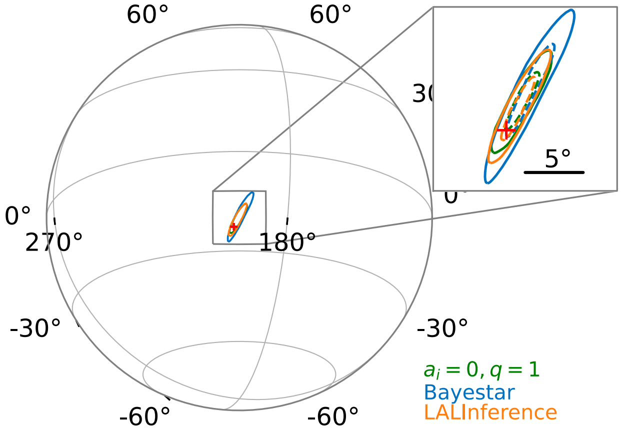

Figure 3. Sky location uncertainty for GW170817. We show posteriors from the $a_i = 0$, $q = 1$ prior in green, the O3 low-latency BAYESTAR analysis in blue, and the O3 high-latency LALINFERENCE analysis in orange. The red ‘plus’ is the location of the identified host galaxy (Abbott et al. 2017b). The solid (dashed) line represents the 90 per cent (50 per cent) credible region.

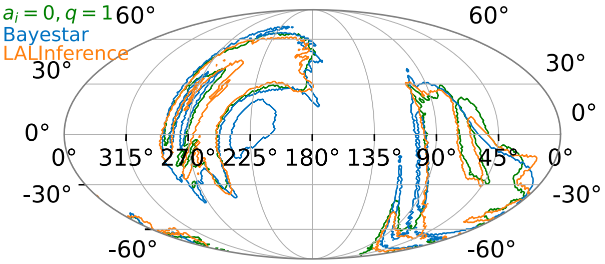

Figure 4. Sky location uncertainty for GW190425. We show posteriors from the $a_i = 0$, $q = 1$ prior in green, the O3 low-latency BAYESTAR analysis in blue, and the O3 high-latency LALINFERENCE analysis in orange. The solid line represents the 90 per cent credible region.

FROQ method; Morisaki & Raymond 2020) could further gain from an optimized choice of prior.

The obvious concern in using an optimized prior with zero spin and equal mass is the induced bias if the detectors observe an event that is highly spinning or of unequal mass ratio. We have controlled this, to some extent, in our PP tests (Section 3.1) by allowing our simulated signals to be drawn from the ‘astrophysical prior’ (cf. Table 1). In particular, we consider mass ratios down to 1/8. In comparison, the mass ratios for two recent NSBH events are around 1/5 (Abbott et al. 2021d). The spin measurements are dominated by the contribution from the more massive BHs: Whereas the dimensionless BH spin magnitude a_1 is constrained to less than 0.23 in GW200105 at the 90 per cent credible level, a_1 could be as high as 0.5 (though with a large uncertainty and still being consistent with zero) for GW200115 (Abbott et al. 2021d).

This suggests that the $a_i = 0$, $q = 1$ prior may be appropriate for NSBH systems, provided the mass ratio is not too extreme and that the BH does not have significant spin. Nevertheless, a real signal that falls outside of our astrophysical prior may be one of the most exciting opportunities for an electromagnetic follow-up. As such, while proposing the use of the $a_i = 0$, $q = 1$ prior, we also recommend parallel computations be made with a broad non-informative prior. In most instances, this will lead to a small amount of wasted computation as the localization will be essentially identical. But, this provides a safeguard against ‘overoptimization’.

3.3 Implications for BBH localization

Having studied prior optimization for the localization of BNS and NSBH events, we now turn to BBH events. BBH systems have significant spin (Abbott et al. 2021c) and have been observed with highly asymmetric masses (Abbott et al. 2020b,d). This suggests that the optimized priors may perform poorly. To test this, we apply the $a_i = 0$ prior to the public data (Abbott et al. 2021b) for the GW190814 (Abbott et al. 2020d) and GW190412 (Abbott et al. 2020a) events. We report the credible intervals in Table 3. For both events, we find that the $a_i = 0$ prior constrains the posterior to a region almost as large as the O3 low-latency results.

In Figs 5 and 6, we show the sky localization for the $a_i = 0$ model for GW190412 and GW190814 and compare these to the high- and low-latency results. Here we see in detail that the $a_i = 0$ prior performs about as well as the low-latency localization (finding an extra mode not present in the high-latency result, and overall a much broader area). From this, we conclude that optimizing the prior for BH systems (which can exhibit significant spin and mass asymmetry) does not perform as well as a high-latency analysis with a complete prior specification.

4 OPTIMIZING THE CHOICE OF WAVEFORM MODEL

In Section 3, we demonstrated that for BNS and NSBH signals with small spins and moderate mass ratios, the sky localization is unbiased by an $a_i = 0$, $q = 1$ prior. We analysed the data using the `IMRPhenomPv2` *waveform approximant* model, which models a fully precessing BBH merger. (Waveform approximants allow the generation of a predicted signal to within a few to tens of milliseconds. Their computation time is typically determined by their level of sophistication: Approximants that better model the underlying physics typically are slower to generate.) In this section, we aim to investigate if simpler waveform models can be used in place of more physically plausible models for localization.

In Table 4, we provide a breakdown of the timing of the IMRPhenomPv2 waveform approximant, the IMRPhenomPv2_NRTidal (used to simulate signals in Section 3.1), and the non-spinning `TaylorF2` (Buonanno et al. 2009) waveform, which models only the inspiral of non-spinning point particles. The TaylorF2 is not as physically accurate as the IMRP waveforms, which include the inspiral, merger, and ring-down (and tidal effects in the case of `IMRPhenomPv2_NRTidal`) of precessing binary systems. This lack of physics translates into a likelihood, which can be evaluated ~30 per cent faster than the IMRPhenomPv2 waveform. If we are using a non-spinning and equal-mass prior, then the only remaining difference between TaylorF2 and IMRPhenomPv2 is the physical modelling of the merger and ring-down. However, the localization is predominantly determined by triangulation from the inspiral signal. This suggests a further optimization: use the TaylorF2 waveform.

To demonstrate that such a waveform optimization does not bias the result, we first look at a fiducial simulated signal. We simulate a spinning BNS (simulation parameters: $q = 0.7$, $a_1 = 0.04$, $a_2 = 0.01$, $\Lambda_1 = 1500$, $\Lambda_2 = 750$, and $d_L = 150\,\mathrm{Mpc}$) in 128 s of a two-detector network (HL) assuming O4 design-sensitivity noise. The signal has a simulated network optimal S/N of ~18. We analyse the signal using IMRPhenomPv2 and TaylorF2 waveforms under an $a_i = 0$, $q = 1$ prior and plot the 2D sky localization in Fig. 7.

Table 3. The 68 per cent credible intervals for the luminosity distance ($d_{\rm L}$) and sky position from our optimized priors, the publicly available O3 low-latency software (BAYESTAR), and O3 high-latency software (LALINFERENCE).

	Parameters	$a_i = 0, q = 1$	$a_i = 0$	O3 low latency	O3 high latency
GW170817 (BNS)	$d_{\rm L}$ (Mpc)	39^{+6}_{-9}	–	40^{+8}_{-8}	40^{+8}_{-14}
	Area (deg^2)	16	–	31	16
GW190425 (BNS)	$d_{\rm L}$ (Mpc)	158^{+43}_{-46}	–	155^{+45}_{-45}	156^{+41}_{-41}
	Area (deg^2)	8012	–	10183	7461
GW190814 (BBH/NSBH)	$d_{\rm L}$ (Mpc)	–	276^{+54}_{-69}	236^{+53}_{-53}	267^{+52}_{-52}
	Area (deg^2)	–	37	38	23
GW190412 (BBH)	$d_{\rm L}$ (Mpc)	–	738^{+237}_{-211}	812^{+194}_{-194}	734^{+138}_{-173}
	Area (deg^2)	–	142	156	110

Below the event name, we give the credible source classification. For BNS events, we use the optimized $a_i = 0$, $q = 1$ prior, while we use the $a_i = 0$ for events containing a BH. We note that some of the reduction in the localization of between the low-latency and high-latency results for GW170817 arises from recalibration of the data from the Virgo interferometer (Abbott et al. 2019).

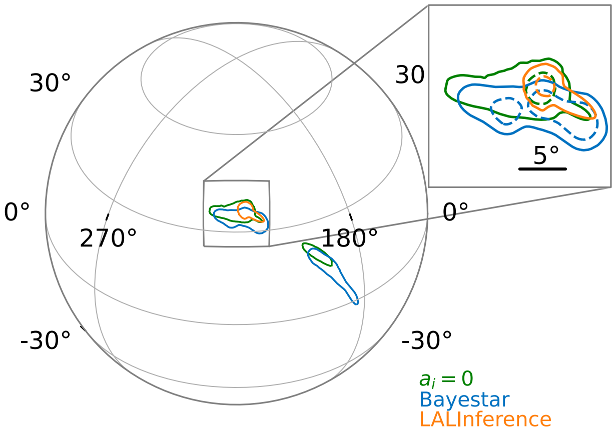

Figure 5. Sky location uncertainty for GW190412 (Abbott et al. 2020a). We show posteriors from the $a_i = 0$ prior in green, the O3 low-latency BAYESTAR analysis in blue, and the O3 high-latency LALINFERENCE analysis in orange. The solid (dashed) line represents the 90 per cent (50 per cent) credible region.

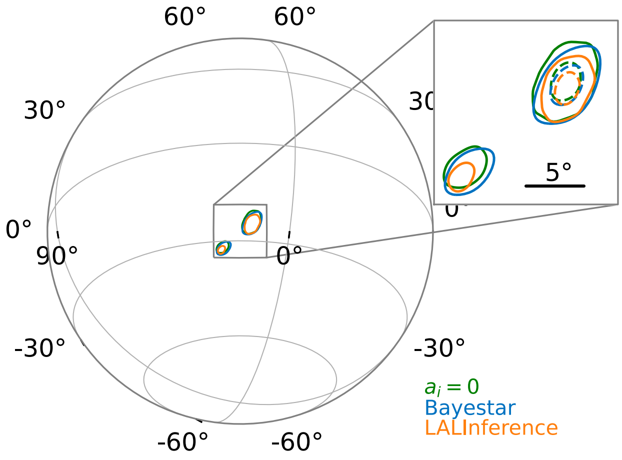

Figure 6. Sky location uncertainty for GW190814 (Abbott et al. 2020d). We show posteriors from the $a_i = 0$ prior in green, the O3 low-latency BAYESTAR analysis in blue, and the O3 high-latency LALINFERENCE analysis in orange. The solid (dashed) line represents the 90 per cent (50 per cent) credible region.

Table 4. Per-likelihood and per-waveform evaluation times.

Waveform approximant	Per-likelihood evaluation (ms)	Per-waveform evaluation (ms)
IMRPHENOMPV2_NRTIDAL	93 ± 5	53 ± 4
IMRPHENOMPV2	87 ± 6	47 ± 4
TAYLORF2	60 ± 8	13.3 ± 0.7

The per-likelihood captures the full cost of computing the likelihood in BILBY for simulated data lasting 128 s. The difference between the per-likelihood and per-waveform values (which amounts to the BILBY data processing overhead) is ∼40 ms. For the TAYLORF2 waveform, the data processing overhead is the dominant contribution to the per-likelihood evaluation. These timings apply only to the standard configuration (without any analytical marginalization). They can be drastically reduced by ROQ and heterodyne methods as discussed in Section 1. All timings are benchmarked on an Intel Core i7-7820HQ CPU @ 2.90 GHz.

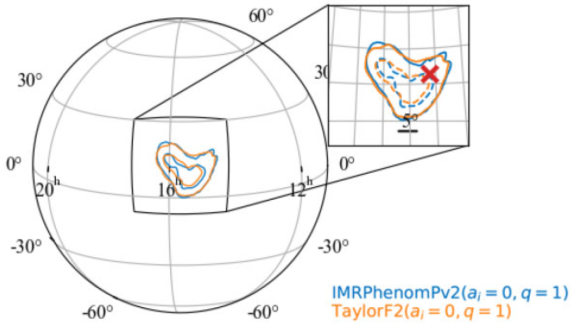

Figure 7. 2D sky localization of a simulated BNS system comparing the IMRPHENOMPV2 and TAYLORF2 waveform models with identical $a_i = 0$, $q = 1$ priors. The red 'cross' indicates the simulated signal value.

This figure demonstrates that we do not see a systematic difference between the two waveforms, despite their differing physical assumptions. (The inferred luminosity distances show a similar level of agreement.)

Both the IMRPHENOMPV2 and TAYLORF2 analyses require a similar number of likelihood evaluations ($\sim 30 \times 10^6$), but the TAYLORF2 run had an overall wall time 30 per cent less than that of the IMRPHENOMPV2 analysis. This confirms that the reduction in per-likelihood evaluations demonstrated in Table 4 translates directly

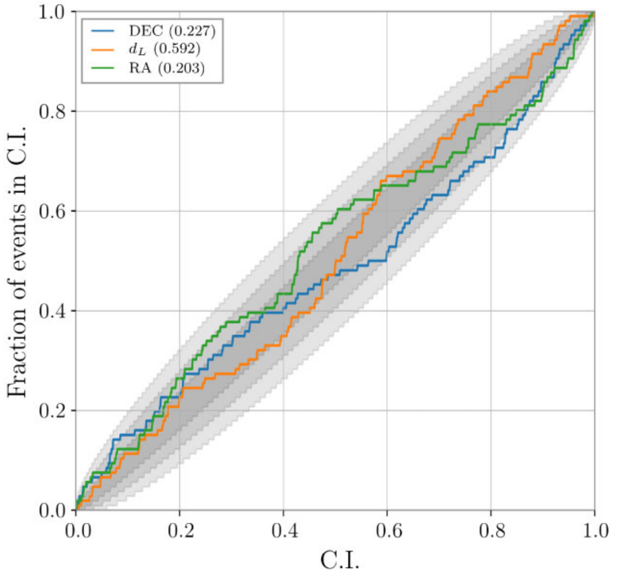

Figure 8. A PP test of the 3D sky localization parameters of 100 simulated BNS signals using the $a_i = 0$, $q = 1$ prior. We simulate signals drawn from the astrophysical prior (see Table 1) using the IMRPhenomPv2_NRTidal waveform model and analyse the data using the TaylorF2 waveform model. The grey region indicates the 1σ, 2σ, and 3σ confidence intervals. Individual p-values for each of the 3D localization parameters are shown in the legend.

into wall time reductions. For reference, we also analyse the data with the IMRPhenomPv2_NRTidal waveform and a full prior (i.e. using the full range of spins and tidal parameters). Comparing the wall times, the $a_i = 0$, $q = 1$ TaylorF2 analysis is 60 per cent faster.

Finally, we verify that using TaylorF2 does not introduce a bias. We repeat the PP test introduced in Fig. 1, but use the TaylorF2 waveform approximant to analyse the data. We report the results in Fig. 8. Again, the 3D localization remains inside the 3σ uncertainty region. The p-values for Dec., $d_{\rm L}$, and RA are 0.227, 0.592, and 0.203, respectively.

5 SUMMARY

We investigate optimization choices for the Bayesian prior and signal model used to produce sky localizations from ground-based gravitational-wave observatories. The rapid production of these sky localizations is critical to aiding in the search for possible counterparts. For BNS systems (where we expect small spins and near-equal mass), we demonstrate that a restrictive prior with zero spin and equal mass can reduce the number of required likelihood evaluations and hence the wall time by ∼40 per cent. At the same time, we demonstrate that this optimized prior is unbiased, provided the signal does not have extreme spins or mass ratios. We also demonstrate that using physically simpler waveform models provides equivalent sky maps with up to a ∼30 per cent reduction in wall time. Taken together, for BNS and moderate-mass-ratio NSBH systems, optimized choices of the prior and waveform can reduce the wall time by up to 60 per cent. This efficiency saving can be directly applied to the current stochastic sampling methods used in high latency (e.g. by implementing the optimized prior in the LALInference or Bilby samplers). For BBH systems, we demonstrate that the same prior optimizations do not apply: Zero-spin and equal-mass assumptions produce poorer sky localization.

We considered these optimization choices in the context of standard stochastic sampling. As discussed in Section 1, a number of new ideas are being developed, which offer wall time speed-ups up to factors of a few hundred. Prior and waveform optimization can be applied to both heterodyning and brute-force parallelization. For ROQ-based methods, a speed-up is achieved by reducing the size of the basis (since it no longer needs to model unequal-mass or spinning BNS), resulting in a speed-up of the basis itself. For machine-learning-based approaches, the optimizations described herein can be used to simplify the training set used at the learning stage. Before the next observing run, we advocate for comparative head-to-head mock data challenges to ascertain which method is the fastest and most robust. This would include extending the studies of the optimized prior choices herein.

Whichever sampling method is used during the O4 observing run, we advocate that for BNS and NSBH systems, two localization analyses be performed: (1) an optimized localization that uses a zero-spin and equal-mass prior (and an inspiral-only waveform if applicable); and (2) a complete localization that uses a moderate-low-spin and unequal-mass configuration. The optimized localization can be produced in about half the time of the complete localization. We anticipate that the two will differ only at the level of stochastic sampling. Nevertheless, by running both analyses, it can be assured that for systems with significant spins and highly asymmetric masses, an updated localization can be produced.

ACKNOWLEDGEMENTS

We thank Soichiro Morisaki for useful feedback during the development of this work. In analysing GW190425, we made use of the ROQ basis provided by Baylor, Smith & Chase (2019). This work was supported by the Australian Research Council (ARC) Future Fellowship FT150100281 and Centre of Excellence CE170100004 and also the National Natural Science Foundation of China under Grant Nos 11633001 and 11920101003. This work was performed on the OzSTAR national facility at the Swinburne University of Technology. The OzSTAR programme receives funding in part from the Astronomy National Collaborative Research Infrastructure Strategy (NCRIS) allocation provided by the Australian Government. This material is based upon work supported by NSF's LIGO Laboratory, which is a major facility fully funded by the National Science Foundation.

DATA AVAILABILITY

This research has made use of data, software, and/or web tools obtained from the Gravitational Wave Open Science Center (https://www.gw-openscience.org/), a service of LIGO Laboratory, the LIGO Scientific Collaboration, and the Virgo Collaboration. LIGO Laboratory and Advanced LIGO are funded by the United States National Science Foundation (NSF) as well as the Science and Technology Facilities Council (STFC) of the United Kingdom, the Max-Planck-Society (MPS), and the State of Niedersachsen/Germany for the support in the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. Virgo is funded, through the European Gravitational Observatory (EGO), by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale di Fisica Nucleare (INFN), and the Dutch Nikhef, with contributions by institutions from Belgium, Germany, Greece, Hungary, Ireland, Japan, Monaco, Poland, Portugal, and Spain.

REFERENCES

Aasi J. et al., 2015, Class. Quantum Gravity, 32, 074001
Abbott B. P. et al., 2017a, Phys. Rev. Lett., 119, 161101
Abbott B. P. et al., 2017b, ApJ, 848, L12
Abbott B. P. et al., 2018a, Living Rev. Relativ., 21, 3
Abbott B. P. et al., 2018b, Phys. Rev. Lett., 121, 161101
Abbott B. P. et al., 2019, Phys. Rev. X, 9, 011001
Abbott R. et al., 2020a, Phys. Rev. D, 102, 043015
Abbott R. et al., 2020b, Phys. Rev. Lett., 125, 101102
Abbott B. P. et al., 2020c, ApJ, 892, L3
Abbott R. et al., 2020d, ApJ, 896, L44
Abbott R. et al., 2021a, Phys. Rev. X, 11, 021053
Abbott R., Abbott T. D., Abraham S. et. al., 2021b, SoftwareX, 13, 100658
Abbott R. et al., 2021c, ApJ, 913, L7
Abbott R. et al., 2021d, ApJ, 915, L5
Acernese F. et al., 2015, Class. Quantum Gravity, 32, 024001
Ackley K. et al., 2020, A&A, 643, A113
Antier S. et al., 2020, MNRAS, 497, 5518
Antil H., Field S. E., Herrmann F., Nochetto R. H., Tiglio M., 2012, preprint (arXiv:1210.0577)
Ashton G. et al., 2019, ApJS, 241, 27
Ashton G., Ackley K., Magaña Hernandez I., Piotrzkowski B., 2020, preprint (arXiv:2009.12346)
Aso Y., Michimura Y., Somiya K., Ando M., Miyakawa O., Sekiguchi T., Tatsumi D., Yamamoto H., 2013, Phys. Rev. D, 88, 043007
Baylor A., Smith R., Chase E., 2019, Available at: https://doi.org/10.5281/zenodo.3478659(Last accessed: 01/09/2020)
Biwer C. M., Capano C. D., De S., Cabero M., Brown D. A., Nitz A. H., Raymond V., 2019, Publ. Astron. Soc. Pac., 131, 024503
Buonanno A., Iyer B. R., Ochsner E., Pan Y., Sathyaprakash B. S., 2009, Phys. Rev. D, 80, 084043
Cahillane C. et al., 2017, Phys. Rev. D, 96, 102001
Canizares P., Field S. E., Gair J. R., Tiglio M., 2013, Phys. Rev. D, 87, 124005
Canizares P., Field S. E., Gair J., Raymond V., Smith R., Tiglio M., 2015, Phys. Rev. Lett., 114, 071104
Christensen N., Meyer R., 1998, Phys. Rev. D, 58, 082001
Cook S. R., Gelman A., Rubin D. B., 2006, J. Comput. Graph. Stat., 15, 675
Cornish N. J., 2010, preprint (arXiv:1007.4820)
Cornish N. J., 2020, Phys. Rev. D, 102, 124038
Cornish N. J., 2021, Phys. Rev. D, 103, 104057
Coughlin M. W. et al., 2019, ApJ, 885, L19
Dietrich T., Samajdar A., Khan S., Johnson-McDaniel N. K., Dudi R., Tichy W., 2019, Phys. Rev. D, 100, 044003
Farr B. et al., 2016, ApJ, 825, 116
Fernández R., Metzger B. D., 2016, Annu. Rev. Nucl. Part. Sci., 66, 23
Finstad D., Brown D. A., 2020, ApJ, 905, L9
Gabbard H., Messenger C., Heng I. S., Tonolini F., Murray-Smith R., 2019, preprint (arXiv:1909.06296)
Goldstein A. et al., 2017, ApJ, 848, L14
Gompertz B. P. et al., 2020, MNRAS, 497, 726
Graham M. J. et al., 2020, Phys. Rev. Lett., 124, 251102
Green S. R., Gair J., 2021, Mach. Learn.: Sci. Technol., 2, 03LT01
Green S. R., Simpson C., Gair J., 2020, Phys. Rev. D, 102, 104057
Hannam M., Schmidt P., Bohé A., Haegel L., Husa S., Ohme F., Pratten G., Pürrer M., 2014, Phys. Rev. Lett., 113, 151101
Hastings W. K., 1970, Biometrika, 57, 97
Jones E. et al., 2001, SciPy: Open source scientific tools for Python, Available at: http://www.scipy.org/(Last accessed: 01/09/2020)
Khan S., Husa S., Hannam M., Ohme F., Pürrer M., Forteza X. J., Bohé A., 2016, Phys. Rev. D, 93, 044007
Lange J., O'Shaughnessy R., Rizzo M., 2018, preprint (arXiv:1805.10457)
Lattimer J. M., Schramm D. N., 1974, ApJ, 192, L145
Li L.-X., Paczyński B., 1998, ApJ, 507, L59
Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E., 1953, J. Chem. Phys., 21, 1087
Morisaki S., 2021, Phys. Rev. D, 104, 044062
Morisaki S., Raymond V., 2020, Phys. Rev. D, 102, 104020
Pankow C., Brady P., Ochsner E., O'Shaughnessy R., 2015, Phys. Rev. D, 92, 023002
Payne E., Talbot C., Lasky P. D., Thrane E., Kissel J. S., 2020, Phys. Rev. D, 102, 122004
Perego A., Radice D., Bernuzzi S., 2017, ApJ, 850, L37
Qi H., Raymond V., 2021, Phys. Rev. D, 104, 063031
Romero-Shaw I. M. et al., 2020, MNRAS, 499, 3295
Schmidt P., Hannam M., Husa S., 2012, Phys. Rev. D, 86, 104063
Schmidt P., Ohme F., Hannam M., 2015, Phys. Rev. D, 91, 024043
Singer L. P., Price L. R., 2016, Phys. Rev. D, 93, 024013
Singer L. P. et al., 2016, ApJ, 829, L15
Skilling J., 2006, Bayesian Anal., 1, 833
Smith R., Field S. E., Blackburn K., Haster C.-J., Pürrer M., Raymond V., Schmidt P., 2016, Phys. Rev. D, 94, 044031
Smith R. J. E., Ashton G., Vajpeyi A., Talbot C., 2020, MNRAS, 498, 4492
Speagle J. S., 2020, MNRAS, 493, 3132
Talbot C., Smith R., Thrane E., Poole G. B., 2019, Phys. Rev. D, 100, 043030
Talts S., Betancourt M., Simpson D., Vehtari A., Gelman A., 2018, preprint (arXiv:1804.06788)
The LSC-Virgo-KAGRA Observational Science Working Groups, 2020, Available at: https://dcc.ligo.org/LIGO-T2000424/public (Last accessed: 01/09/2020)
Valenti S. et al., 2017, ApJ, 848, L24
Veitch J., Vecchio A., 2008, Phys. Rev. D, 78, 022001
Veitch J. et al., 2015, Phys. Rev. D, 91, 042003
Vitale S., Haster C.-J., Sun L., Farr B., Goetz E., Kissel J., Cahillane C., 2021, Phys. Rev. D, 103, 063016
Williams M. J., Veitch J., Messenger C., 2021, Phys. Rev. D, 103, 103006
Wysocki D., O'Shaughnessy R., Lange J., Fang Y.-L. L., 2019, Phys. Rev. D, 99, 084026
Yang S. et al., 2017, ApJ, 851, L48
Zackay B., Dai L., Venumadhav T., 2018, preprint (arXiv:1806.08792)
Zhu X.-J., Ashton G., 2020, ApJ, 902, L12

APPENDIX A: ADDITIONAL DETAILS ON THE PP TEST

Here, we present additional details of the PP test we perform for our optimized choice of priors. In Fig. A1, we show the PP plot for the

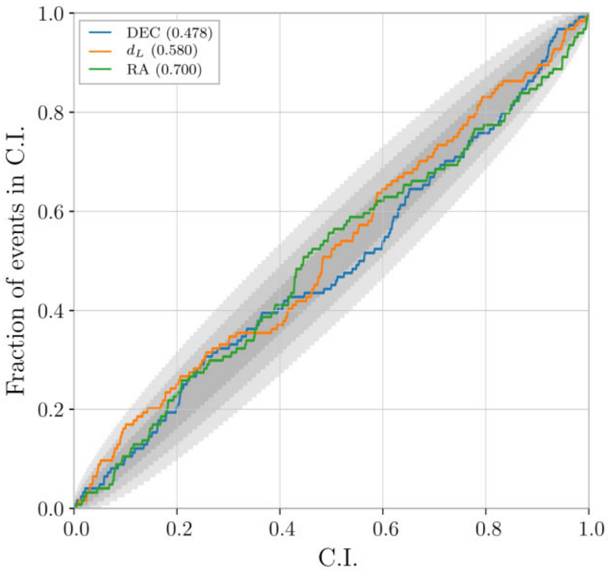

Figure A1. A PP test of the 3D sky localization parameters of 100 simulated BNS signals using the $a_i = 0$, $q = 1$ prior with a two-detector (HL) network.

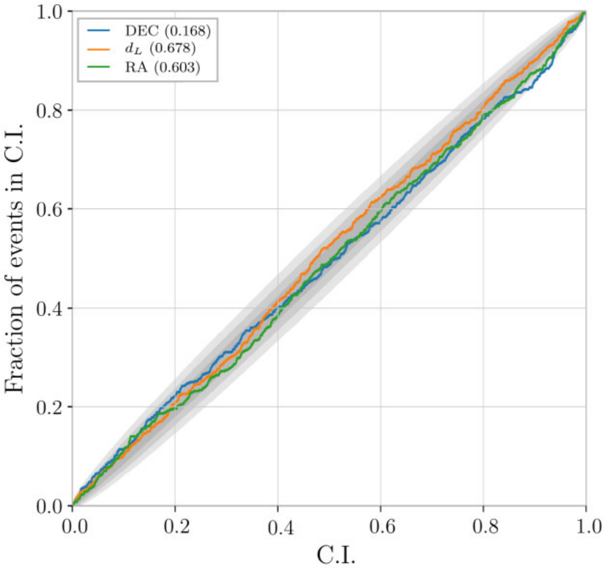

Figure A2. As Fig. A1 but for 500 injections.

Table A1. The p-values and the fraction of the truth value out of 90 per cent credible interval for Dec., d_{L}, and RA with the 50, 100, 300, and 500 simulated BNS injections.

	p-values			Out of 90 per cent		
N	Dec.	d_{L}	RA	Dec.	d_{L}	RA
50	0.533	0.493	0.616	18 per cent	8 per cent	12 per cent
100	0.122	0.551	0.710	13 per cent	12 per cent	9 per cent
300	0.120	0.527	0.696	12 per cent	10 per cent	11 per cent
500	0.168	0.678	0.603	13 per cent	9 per cent	11 per cent

$a_i = 0$, $q = 1$ prior for a two-detector (HL) network. In Fig. A2, we show the PP plot for the $a_i = 0$, $q = 1$ prior for the HLV detector network using 500 injections. Furthermore, we list the p-values and the fractions of truth parameter values being out of the 90 per cent credible intervals for the $a_i = 0$, $q = 1$ prior, for 50, 100, 300, and 500 injections.

This paper has been typeset from a TEX/LATEX file prepared by the author.