






Using overlap of sky localization probability maps for filtering potentially lensed pairs of gravitational-wave signals*

Henry W. Y. Wong ^{1,2,†} Lok W. L. Chan ^{1,‡} Isaac C. F. Wong ^{3,4,1,§}
Rico K. L. Lo ^{5,6,¶} and Tjonnje G. F. Li ^{1,7,3,4,**}

¹*Department of Physics, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong*
²*Department of Statistics and Data Science, University of Texas at Austin, Austin, Texas, Texas 78712, USA*
³*KU Leuven, Department of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, Kasteelpark Arenberg 10, 3001 Leuven, Belgium*
⁴*Leuven Gravity Institute, KU Leuven, Celestijnenlaan 200D box 2415, 3001 Leuven, Belgium*
⁵*Center of Gravity, Niels Bohr Institute, Blegdamsvej 17, 2100 Copenhagen, Denmark*
⁶*LIGO Laboratory, California Institute of Technology, Pasadena, California, California 91125, USA*
⁷*Institute for Theoretical Physics, KU Leuven, Celestijnenlaan 200D, B-3001 Leuven, Belgium*
(Dated: December 16, 2025)

Strong gravitational lensing creates multiple images of a gravitational wave transient. The current state-of-the-art method for identifying such lensing events is a computationally expensive full Bayesian analysis. In this paper, we investigate the feasibility and efficiency of using the overlap of sky localization probability maps (skymaps) to quickly filter potentially lensed gravitational wave signal pairs. We introduce three overlap statistics and test their performance using 200 simulated lensed pairs of gravitational-wave signals across five sets of signal-to-noise ratios. By setting a threshold with a false positive rate of $\text{FPR} = 10^{-2}$ for the three overlap statistics, we find that we can filter out over 99% of nonlensed events while retaining all lensed events. The statistics for each event pair can be computed instantly, and can be used in practice to quickly analyze existing events using the skymaps from the low-latency localization pipelines when results from the full parameter estimation are not available.

I. INTRODUCTION

When a gravitational wave (GW) passes by a massive object, such as a black hole or a galaxy cluster, its path is bent toward the massive object [1], a phenomenon known as gravitational lensing. As a result, the GW can take different paths to reach the observer and form multiple images with different arrival times [2–6]. Consequently, it is predicted that a number of GW signals with different amplitudes and arrival times originating from the same source could be observed.

Since the detection of the first GW signal in 2015 [7], 93 confident GW events have been detected by Advanced LIGO [8] and Advanced Virgo [9] in the first (O1), the second (O2), and the third observing run (O3) [10–13]. Although there is no compelling evidence in lensing of the observed GW events in O3a [14], theoretical predictions suggest roughly one strong GW lensing event per year at the Advanced LIGO-Advanced Virgo-KAGRA design sensitivity [15, 16]. Therefore, it is expected that the first lensed GW event pair will be detected in the coming years [15, 17].

There are different methods to detect strongly lensed GW event pairs [14]. One method [18, 19] uses the overlap of the posterior distribution of the source parameters

to check for the consistency of the source properties between events. The overlap of the two posterior distributions is expected to be high when the two events come from the same source. Another method [20, 21] performs a full Bayesian analysis, namely the joint-parameter estimation (joint-PE), on the detected events.

Joint-PE provides a solid statistical framework to calculate the probability of a GW event pair being lensed images from the same source. The current number of confident GW events is 93 [10–13], which gives 4278 pairs of GW events. However, performing joint-PE on all event pairs is very time consuming. Moreover, for a proposed Einstein Telescope survey, the predicted event rate of binary neutron star mergers is as high as $O(10^5 - 10^6) \text{ yr}^{-1}$ and $O(10^4) \text{ yr}^{-1}$ for neutron star-black hole mergers [22]. This means that the detector sensitivity and the number of GW signals in the future will be far higher and will further increase the time for doing joint-PE on all possible event pairs. For the purpose of sustainability, it is necessary to establish a low latency filter to screen out event pairs that are definitely not lensed.

There are several fast algorithms for rapid identification of strongly lensed GW events developed to address the sustainability issue. One of the algorithms is aimed to search for strongly lensed multiple GW images—from the fact that the signals come from the same source, it is possible to split the joint PE into two easier PEs, that we can use the result from one event to perform the inference of the another event [23]. Such methodology enables relatively quick multiple-image analyses with high accuracy. Another approach proposed by Goyal *et al.* uses a machine learning model to predict the probability of

* Shared first authorship: Henry W. Y. Wong, Lok W. L. Chan

† Contact author: wyhwong@link.cuhk.edu.hk

‡ lokwlc@link.cuhk.edu.hk

§ chunfung.wong@kuleuven.be

¶ kalok.lo@nbi.ku.dk

** tjonnje.li@kuleuven.be

lensing in event pairs from the time-frequency maps and posterior distributions of GW events [24]. The computational time of the lensing probability is estimated to be around three seconds with most of the time spent on loading necessary files. Both of the mentioned works have successfully reduced the computational expense of filtering potentially lensed GW event pairs.

To establish a low latency filter, we propose the use of sky localization probability maps (skymaps) to analyze the possibility of an event pair coming from the same source. The deflection angle is usually a few hundred arc seconds for black holes and galaxy clusters [25], which is negligibly small as it is far less than the uncertainty of the localization of the source [26]. Therefore, the skymaps of the images from the same source are expected to be highly overlapped. In other words, event pairs with nonoverlapping skymaps are definitely not lensed pairs. We use this fact to analyze the possibility of whether two GW signals are a lensed pair and filter potentially lensed pairs by comparing their skymaps. With the use of overlap, we can reject a large fraction of event pairs that are definitely not image pairs. However, one must note that having a high overlap in the skymaps does not necessarily imply the events are a lensed pair. In contrast to the machine learning approach in [24], we investigate using measures of the overlap of skymaps without involving sophisticated learning models, which can be easily incorporated into search pipelines such as the subthreshold search [27] to send alerts for performing follow-up analyses. The goal of the low latency filter is to reduce time cost by rapidly filtering out most of the nonlensed GW event pairs for further lensing analysis. Therefore, we would like to quantitatively investigate the efficiency of using skymap overlap to filter the nonlensed pairs while not rejecting the lensed pairs. With a low latency skymap generated by a rapid sky localization algorithm such as BAYESTAR [28], a quick preliminary lensing analysis on two GW signals can be done within a minute.

In this paper, we present three overlap statistics: normalized posterior overlap, 90% credible region overlap and cross-highest posterior density statistic, to quantitatively investigate the efficiency of using skymap overlap to filter the nonlensed pairs while not rejecting the lensed pairs. The details of the statistics are described in Sec. III. Using simulated GW events, we study the distribution of skymap overlap of lensed GW event pairs under different signal-to-noise ratios (SNRs). In order to filter potentially lensed event pairs, a reasonable threshold of skymap overlap that can filter out nonlensed event pairs while keeping all of the lensed event pairs should be set with a specific false positive rate (FPR). Based on the results, we compare the effectiveness of the three statistics and discuss the feasibility of using skymap overlap to filter potentially lensed pairs of GW signals.

This paper is organized as follows: In Secs. II and III, we establish the methodology of the search by introducing posterior distribution and the three statistics mentioned. In Sec. IV, we investigate the feasibility of us-

ing skymap overlap for filtering potentially lensed pairs of GW signals with the receiver operating characteristic (ROC) curves. In Sec. V, we conclude our findings on the feasibility of skymap overlap in this paper and its potential work in the future.

II. SKYMAP

A. Posterior distribution

To generate a skymap, we use the observed data to estimate the source location of a GW projected on the celestial sphere under a fixed waveform model. This process relies on parameter estimation, which allows us to quantify the uncertainty of the GW location on the celestial sphere. We first define a hypothesis H that the GW signal fits the fixed waveform model. Then for each GW source parameter, we set up a prior probability distribution $p(\theta|H)$, which represents the *a priori* knowledge of the model parameters before observing the data. Using Bayes' theorem, the prior probability distribution $p(\theta|H)$ and the data d of the GW signals are used to compute the posterior distribution $p(\theta|d, H)$ of the waveform parameters [29] defined as

$$p(\theta|d, H) \equiv \frac{p(\theta|H)p(d|\theta, H)}{p(d|H)}. \quad (1)$$

B. Sky localization probability map

A waveform model typically has many parameters that we can denote as $\theta = \{\theta_1, \theta_2, \dots, \theta_N\}$. The posterior distribution $p(\theta|d, H)$ represents the joint posterior distribution of all parameters. To determine the posterior distribution of a specific parameter, we marginalize the nuisance parameters, where the posterior distribution of a specific parameter $p(\theta_1|d, H)$ is obtained by

$$p(\theta_1|d, H) \equiv \int d\theta_2 \dots d\theta_N p(\theta|d, H). \quad (2)$$

After obtaining the posterior distribution of the right ascension and the declination of the source location, which is referred to as a skymap in this paper, we compute the three overlap statistics of a skymap pair described in the following Section.

III. OVERLAP STATISTICS

A. Posterior overlap

After performing the parameter estimation on two GW events separately, we obtain two posterior distributions $p(\hat{\Omega})$ and $q(\hat{\Omega})$ of their skymap coordinate $\hat{\Omega}$. We determine whether a GW signal pair is potentially lensed by

the inner product of $p(\hat{\Omega})$ and $q(\hat{\Omega})$, which quantifies the overlap between two probability densities. The posterior overlap \mathcal{D}_{PO} between two skymaps is defined similarly to [18] as

$$\mathcal{D}_{\text{PO}} \equiv \int p(\hat{\Omega}')q(\hat{\Omega}')d\Omega' \quad (3)$$

which is not normalized. In order for the posterior overlap of different GW signal pairs to be comparable to each other, we normalize the posterior overlap with skymap of the two GW events, which is defined as

$$\mathcal{D}_{\text{NPO}} \equiv \frac{\int p(\hat{\Omega}')q(\hat{\Omega}')d\Omega'}{\sqrt{\int p(\hat{\Omega}')p(\hat{\Omega}')d\Omega'}\sqrt{\int q(\hat{\Omega}')q(\hat{\Omega}')d\Omega'}} \in [0, 1]. \quad (4)$$

B. 90% credible region overlap

For a GW signal with an extremely large SNR, the posterior distribution of the source location could be narrowly peaked. This may lead to an extremely small value of posterior overlap and normalized posterior overlap with another GW signal. As a result, the two GW signals may be recognized as a nonlensed pair and filtered out. Therefore, we propose to use a 90% credible region overlap.

Instead of finding the overlap probability of the source position on the sky, we directly compute the overlapping area of the 90% credible regions. The overlapped area is further normalized by the minimum area between two 90% credible regions. The minimum value is used to keep a high 90% credible region overlap when the difference between two event SNRs is high for lensed pairs. Therefore, it reduces the chance of lensed pairs being filtered out. The 90% credible region overlap $\mathcal{D}_{90\% \text{CR}}(p, q)$ of two GW signals is computed by masking their skymaps, which is defined as

$$\begin{aligned} & \mathcal{D}_{90\% \text{CR}}(p, q) \\ &= \frac{\int \mathbb{1}_{90\% \text{CR}}[p(\hat{\Omega}')] \mathbb{1}_{90\% \text{CR}}[q(\hat{\Omega}')] d\Omega'}{\min\left(\int \mathbb{1}_{90\% \text{CR}}[p(\hat{\Omega}')] d\Omega', \int \mathbb{1}_{90\% \text{CR}}[q(\hat{\Omega}')] d\Omega'\right)}, \end{aligned} \quad (5)$$

where $\mathbb{1}_{90\% \text{CR}}[p(\hat{\Omega})]$ and $\mathbb{1}_{90\% \text{CR}}[q(\hat{\Omega})]$ are the indicator function. This indicator function is 1 when the sky location is in the 90% credible region of the corresponding skymap, and 0 otherwise.

C. Cross-HPD statistic

We introduce another statistic called the cross-HPD. The purpose of this statistic is similar to that of the 90%

credible region overlap. We use the maximum *a posteriori* (MAP), which is the mode of the posterior distribution, as an alternative of the 90% confidence level to decide the credible region used for computing the overlap of the skymap between two GW signals. We first compute the probability of the MAP of one skymap on another skymap, which is referred as the search probability \mathcal{H} . It is defined as

$$\mathcal{H}(p, q) = \int \mathbb{1}\left[q(\hat{\Omega}') > q(\hat{\Omega}_{p, \text{MAP}})\right] q(\hat{\Omega}')d\hat{\Omega}', \quad (6)$$

where $\mathbb{1}$ is the indicator function, $q(\hat{\Omega})$ is the posterior distribution of sky position of an event, and $\hat{\Omega}_{p, \text{MAP}}$ is the MAP estimate of $\hat{\Omega}_p$, also known as the coordinates of the point at highest probability density.

$\mathcal{H}(p, q)$ quantifies the overlap of the skymap between the event pairs. The lower the value of $\mathcal{H}(p, q)$, the higher the overlap between the skymaps. The cross-HPD statistic $\mathcal{D}_{\text{HPD}}(p, q)$ is defined as

$$\mathcal{D}_{\text{HPD}}(p, q) \equiv \max\{1 - \mathcal{H}(p, q), 1 - \mathcal{H}(q, p)\}, \quad (7)$$

which is symmetric in p and q .

IV. RESULTS AND DISCUSSION

We simulate 200 lensed pairs of GW signals for five different SNRs to investigate the performance of the search methodology. We first generate the simulated data of 200 lensed pairs of GW signals with PyCBC [30]. Primary mass follows a power-law profile with slope -2.35 , while the mass ratio and sky position are uniformly distributed. With the simulated data, we rescale all the events to five different SNRs by adjusting the luminosity distance of each event. In total, there are 1,000 event pairs in the injection campaign. For details, see Sec. IV A. We use a Bayesian inference library `Bilby` [31, 32] to perform parameter estimation on the GW events. After obtaining the posterior distribution of the sky location of each event, the skymaps of all events are generated using `ligo.skymap` [33]. We pair up the skymaps for every possible combination and separate them into lensed-pair group and nonlensed-pair group. The three overlap statistics of all event pairs are then computed. Finally, we plot the ROC curve to compare the performance of the overlap statistics. Details of the ROC curve can be found in Sec. IV C.

A. Injection campaign

To study the distribution of skymap overlap of lensed-pair candidates under different SNRs and the efficiency of overlap statistics on filtering potentially lensed pairs of GW signals, we set up an injection campaign with parameter settings as listed in Table I. A total of 200

TABLE I. The parameters of the waveform simulated and the set-up of the injection campaign. 200 event pairs are generated. For simplicity, the spins of all events are set to zero. All events are rescaled to 5 different SNRs by adjusting the luminosity distance of each event. Finally, there are 2,000 events in total.

Number of lensed event pairs	200
Waveform approximant	IMRPhenomPv2 [34, 35]
Spin	None
Rescaled SNR	8, 18.5, 29, 39.5, 50
Detector network	HLV at design sensitivity [9, 36]
Primary-mass distribution	Power law with slope -2.35
Mass-ratio distribution	Uniform
Sky-location distribution	Uniform

lensed GW event pairs are generated. The localization uncertainty of a GW signal depends on its SNR, therefore we are interested in the effect of different SNRs on the skymaps of lensed GW signal pairs. The uncertainty of the localization is smaller when the SNR is higher. Therefore, we expect the overlap to be generally lower for GW event pairs with a higher SNR than those with a lower SNR. We use five different SNRs to study the difference in the filtering threshold for different SNR pairs by skymap overlap.

The largest SNR of the event pairs is chosen to be 50, it is a sufficiently high SNR upper limit due to the rarity of detecting a signal with such SNR. After setting up the upper limit and the lower limit of the event SNR, we uniformly divide the range between 8 and 50 to obtain five different SNRs and rescale the luminosity distances to attain the desired SNR.

After performing the parameter estimation, we generate the skymaps for all events. The prior of chirp mass is set to be uniform from $4 M_{\odot}$ to $150 M_{\odot}$ and the two component masses are constrained between $5 M_{\odot}$ to $150 M_{\odot}$, while the injections are drawn from an astrophysical power-law primary-mass distribution. In LIGO-Virgo-KAGRA PE analyses, agnostic priors are typically used; these commonly assume uniform priors on the component masses [37]. Choosing a uniform chirp-mass prior here avoids tailoring the PE to the injections and mirrors the practical, agnostic setting in which skymap filtering would be applied. Sky localization is dominated by network SNR, interdetector time delays and antenna patterns, so the prior mismatch does not materially affect the skymap-overlap conclusions. The prior of geocent time is chosen to be a delta function that peaks at the injected parameter. The sampler used for the parameter estimation is `dynesty` [38]. Finally, the three overlap statistics of each event pair are computed. The lensed event pairs consist of 1000 samples which are 200 lensed events pairs for each of the five SNRs. For the nonlensed event pairs, the overlap of all the possible event pairs from the original 2,000 events is computed, which results in about 700,000 samples. By plotting the ROC curve of each statistic, we study the performance of the three

TABLE II. Mean area of the 90% credible region for skymaps from 200 simulated events using the HLV network at five SNRs. Details of the injection campaign are given in Table I.

SNR	Mean area of 90% credible region (deg ²)
8	186.8
18.5	34.8
29	12.0
39.5	6.4
50	4.1

statistics on filtering lensed candidate event pairs.

B. Coverage of samples

Table II shows the mean area of 90% credible region of the skymaps in the injection campaign. The size of the 90% credible region relates to the uncertainty of the localization. For GW signals with a smaller SNR, the uncertainty of localization is larger. Therefore, the mean area of 90% credible region for events with SNR = 8 is significantly larger than that of the events with SNR = 50.

Figure 1 illustrates the 90% credible region of events with SNR = 8. The contours of their 90% credible regions are demonstrated in different colors for distinguishing different contours. The 200 events come from different parts of the sky and their 90% credible regions covered about half of the sky. Note that there are overlaps of 90% credible regions of nonimage pairs. Those non-image event pairs with relatively high overlap values are the false positives. The coverage ensures that there is a sufficient number of lensed event pairs in the injection campaign to study the distribution of skymap overlap for nonimage pairs.

C. Performance of statistics

With the overlap statistics of all the event pairs, the performance of the statistics is shown in a ROC curve. In a ROC curve, which shows the performance of a classification model at all classification thresholds by plotting the true positive rate (TPR) against the FPR.

The y axis of the ROC curves is TPR, which is defined as

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (8)$$

where TP is the number of lensed event pairs which overlap value is higher than the filtering threshold, while FN is the number of lensed event pairs which overlap values are smaller than the filtering threshold.

The x axis of the ROC curves is FPR, which is defined as

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}, \quad (9)$$

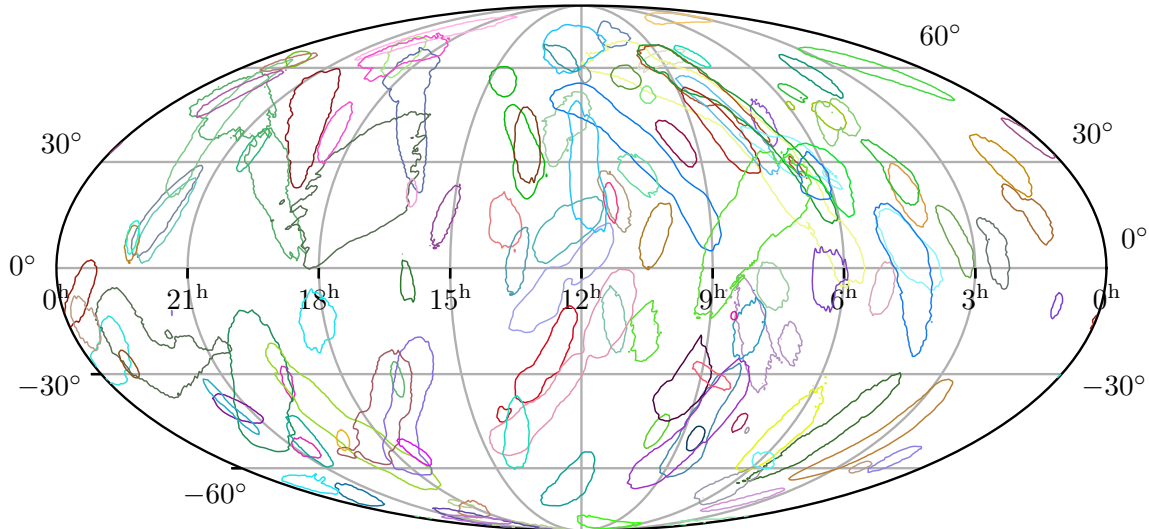


FIG. 1. Contours of 90% credible region of 200 skymaps with $\text{SNR} = 8$ are shown in the figure. Details of the injection campaign are referred to Table I. The figure shows the sky coverage of the injections with $\text{SNR} = 8$.

where FP is the number of nonlensed event pairs which overlap value is higher than the filtering threshold, while TN is the number of nonlensed event pairs which overlap value is smaller than the filtering threshold.

Given a specific FPR or a specific filtering threshold, a higher TPR indicates a better performance of the statistic because less lensed event pairs are recognized as nonlensed event pairs by the statistic. We aim to set up a reasonable threshold with $\text{TPR} \approx 1$ to filter nonlensed event pairs while keeping most of the lensed event pairs for further analysis. Ten ROC curves are shown in Fig. 2 and Fig. 3. The five ROC curves in Fig. 2 indicate the performance of the three statistics for event pairs with one of the event SNRs being small ($\text{SNR} = 8$), while the five ROC curves in Figure 3 indicate the performance of the three statistics for event pairs with one of the event SNRs is large ($\text{SNR} = 50$).

Figure 2 shows the ROC curves of the three overlap statistics for event pairs with one event $\text{SNR} = 8$ while another is one of the five SNRs. In this case, the TPR of

the three statistics shows nearly no difference when the FPR is fixed to a value between 10^{-2} and 1. The difference in the performance of the statistics is noticeable only when we fix FPR to a value smaller than around 10^{-2} .

For a relatively small FPR value ($\text{FPR} < 10^{-2}$), 90% credible region overlap shows the worst overall performance as it shows the most significant drop at a lower FPR. However, all three statistics show a significant drop when the FPR is lower than 10^{-2} . Within this range, the losses of lensed event pairs are significant for all three statistics. Therefore, the threshold should be fixed higher than 10^{-2} if we aim to keep most of the lensed event pairs. Thus, the performance of the three statistics is similar for a reasonable choice of threshold that can filter out nonlensed event pairs while keeping most of the lensed event pairs. The straight line of the TPR drop for 90% credible region overlap in $\text{FPR} \lesssim 10^{-3}$ is due to a lack of samples. The data points of the ROC curve are generated by using a threshold to separate the input

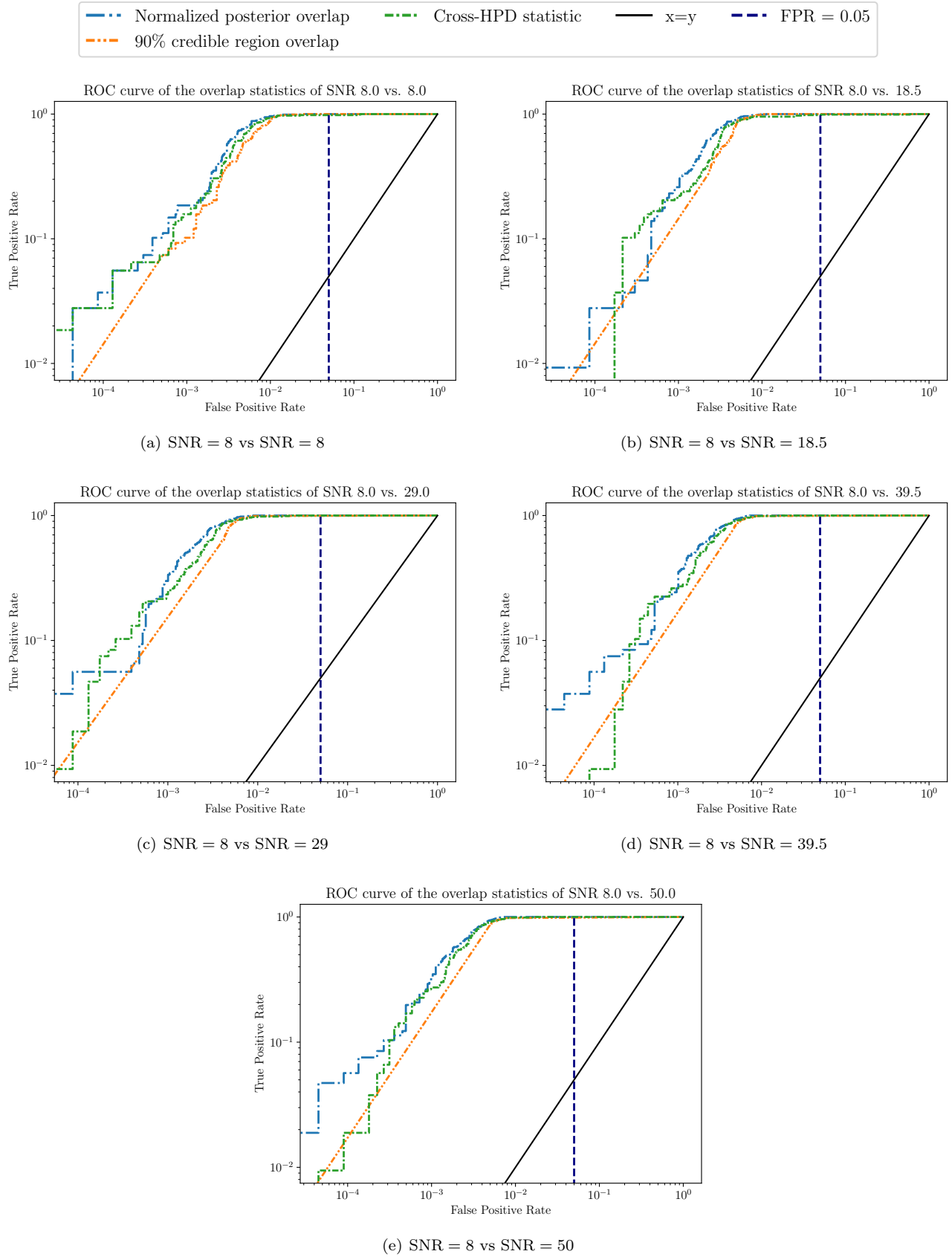


FIG. 2. ROC curves of event pairs with one event SNR = 8. The black solid line $x = y$ represents the ROC curve of a random classifier. ROC curves indicate the performance of the overlap statistics. The classifier is better if it can achieve a higher true positive rate given a false positive rate. The three statistics do not show a large difference in the region of FPR between 10^{-2} and 1.

data into FP, TP, FN, and TN.

Figure 3 shows the ROC curves of the three overlap statistics for event pairs with one event SNR = 50 while another is one of the five SNRs. In this case, the TPR of the three statistics shows nearly no difference when the FPR is fixed to a value between 10^{-2} and 1. The difference in the performance of the statistics is noticeable only when we fix FPR to a value smaller than around 10^{-3} .

For a relatively small FPR value ($FPR < 10^{-3}$), 90% credible region overlap shows the worst overall performance. However, again, choosing a threshold with FPR lower than around 10^{-3} will lead to a significant loss in the number of lensed event pairs. Therefore, the filtering threshold should be fixed at an FPR higher than 10^{-3} in order to keep all of the lensed event pairs. Thus, the performance of the three statistics is similar for a reasonable choice of filtering threshold that can filter out nonlensed event pairs while keeping all of the lensed event pairs. Again, the straight line of the TPR drop for 90% credible region overlap in $FPR \lesssim 10^{-3}$ is due to the lack of samples.

For the purpose of filtering potentially lensed candidate event pairs, the three skymap overlap statistics demonstrated a similar performance while fixing FPR between 10^{-2} and 1. In the mock data study, we can set up a specific $FPR = 10^{-2}$ for event pairs in all SNR to filter out 99% of the nonlensed candidate pairs, which demonstrated that a large fraction of nonimage pairs could be rapidly filtered away using the skymap overlap.

V. CONCLUSIONS

In this paper, we have demonstrated our proposed search methodology using skymap overlap to filter potentially lensed pairs of GW signals. By setting up a threshold with a specific $FPR = 10^{-2}$ for the three overlap statistics, we can filter out more than 99% of the nonlensed event pairs while keeping most of the lensed

event pairs, which shows the feasibility of using skymap overlap to filter potentially lensed event pairs of GW signals.

Since the computation of the three statistics for each event pair takes less than 3 seconds, a new trigger could be constructed to compute the skymap overlap of all existing events within a short period of time using the low latency skymaps generated by BAYESTAR [28] or other rapid sky localization algorithms.

To accommodate real detections, which span continuous SNR combinations, we propose generating a dense injection grid to derive ROC-based thresholds as a function of SNR and interpolating these results to produce a runtime mapping that adjusts the overlap threshold for any observed pair; this enables real-time, SNR-conditioned filtering in production pipelines.

ACKNOWLEDGEMENTS

We thank Otto Akseli Hannuksela for fruitful discussion. I. C. F. W. and T. G. F. L. are partially supported by grants from the Research Grants Council of the Hong Kong (Projects No. 24304317 and No. 14306419) and Research Committee of the Chinese University of Hong Kong. R. K. L. L. and T. G. F. L. would also like to gratefully acknowledge the support from the Croucher Foundation in Hong Kong. The authors acknowledge the use of the IUCAA LDG cluster Sarathi for the computational/numerical work. The authors are also grateful for the computational resources provided by the Chinese University of Hong Kong. The Center of Gravity is a Center of Excellence funded by the Danish National Research Foundation under Grant No. 184. This work is also partially supported by the Research Foundation - Flanders (FWO) through Grants No. I002123N, No. I000725N, No. G086722N.

DATA AVAILABILITY

The data that support the findings of this manuscript are openly available [39].

-
- [1] M. Bartelmann, *Classical and Quantum Gravity* **27**, 233001 (2010), 1010.3829.
 - [2] H. C. Ohanian, *International Journal of Theoretical Physics* **9**, 425 (1974).
 - [3] S. Deguchi and W. D. Watson, *Phys. Rev. D* **34**, 1708 (1986).
 - [4] Y. Wang, A. Stebbins, and E. L. Turner, *Phys. Rev. Lett.* **77**, 2875 (1996), astro-ph/9605140.
 - [5] T. T. Nakamura, *Phys. Rev. Lett.* **80**, 1138 (1998).
 - [6] R. Takahashi and T. Nakamura, *Astrophys. J.* **595**, 1039 (2003), astro-ph/0305055.
 - [7] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, et al., *Phys. Rev. Lett.* **116**, 061102 (2016), 1602.03837.
 - [8] LIGO Scientific Collaboration, J. Aasi, B. P. Abbott, R. Abbott, T. Abbott, M. R. Abernathy, K. Ackley, C. Adams, T. Adams, P. Addesso, et al., *Classical and Quantum Gravity* **32**, 074001 (2015), 1411.4547.
 - [9] F. Acernese, M. Agathos, K. Agatsuma, D. Aisa, N. Allemandou, A. Allocca, J. Amarni, P. Astone, G. Balestri, G. Ballardín, et al., *Classical and Quantum Gravity* **32**, 024001 (2015), 1408.3978.
 - [10] B. P. Abbott, R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt, et al., *Physical Review X* **9**, 031040 (2019), 1811.12907.

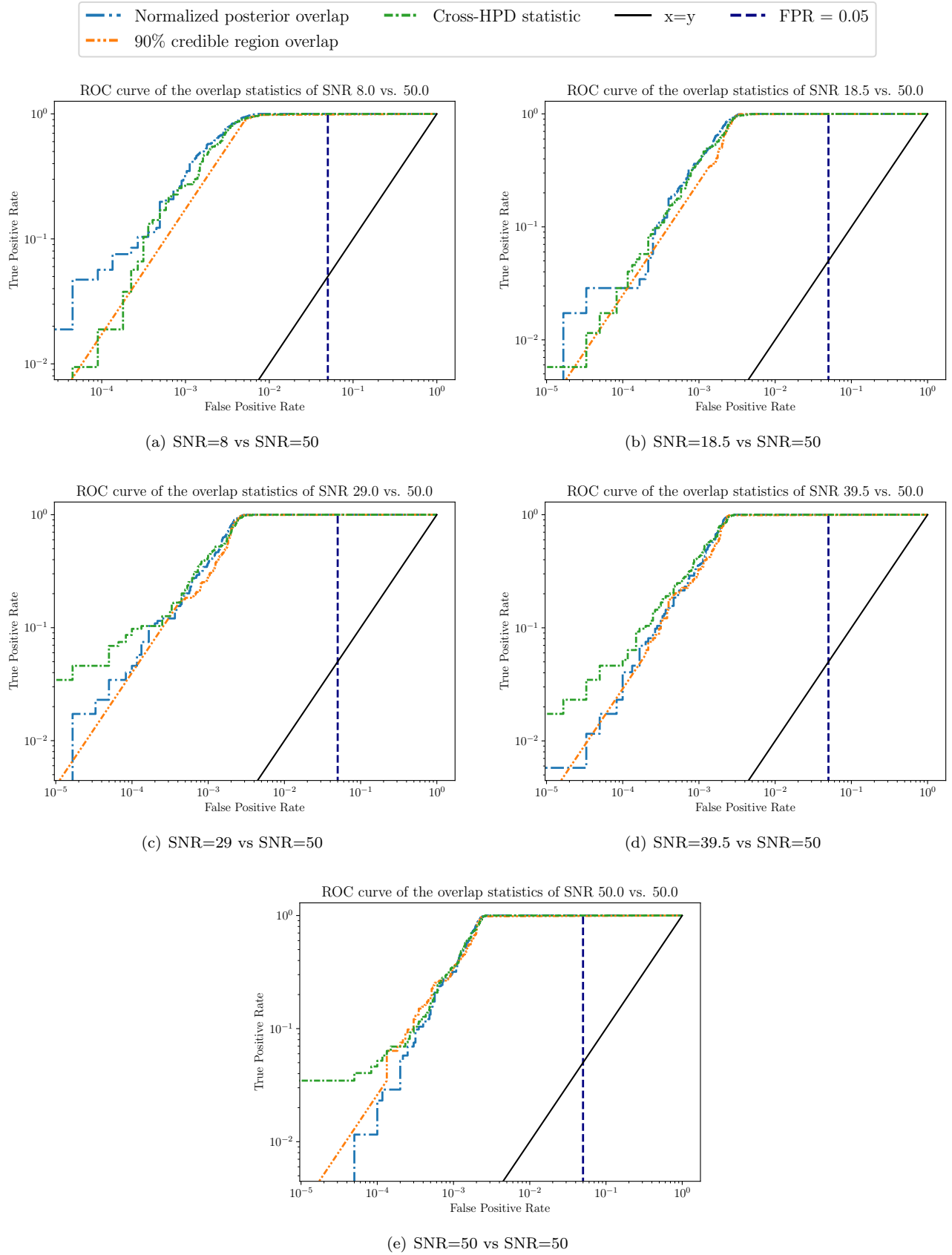


FIG. 3. ROC curves of event pairs with one event SNR = 50. The black solid line $x = y$ represents the ROC curve of a random classifier. The classifier is better if it can achieve a higher true positive rate given a false positive rate. The three statistics do not show a large difference in the region of FPR between 10^{-2} and 1.

- [11] R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, A. Adams, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt, et al., *Physical Review X* **11**, 021053 (2021), 2010.14527.
- [12] R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, N. Adhikari, R. X. Adhikari, V. B. Adya, C. Affeldt, D. Agarwal, et al., *Phys. Rev. D* **109**, 022001 (2024), 2108.01045.
- [13] R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, N. Adhikari, R. X. Adhikari, V. B. Adya, C. Affeldt, D. Agarwal, et al., *Physical Review X* **13**, 041039 (2023), 2111.03606.
- [14] R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, A. Adams, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt, et al., *Astrophys. J.* **923**, 14 (2021), 2105.06384.
- [15] A. R. A. C. Wierda, E. Wempe, O. A. Hannuksela, L. V. E. Koopmans, and C. Van Den Broeck, *Astrophys. J.* **921**, 154 (2021), 2106.06303.
- [16] F. Xu, J. M. Ezquiaga, and D. E. Holz, *Astrophys. J.* **929**, 9 (2022), 2105.14390.
- [17] K. K. Y. Ng, K. W. K. Wong, T. Broadhurst, and T. G. F. Li, *Phys. Rev. D* **97**, 023012 (2018), 1703.06319.
- [18] K. Haris, A. K. Mehta, S. Kumar, T. Venumadhav, and P. Ajith, arXiv e-prints arXiv:1807.07062 (2018), 1807.07062.
- [19] O. A. Hannuksela, K. Haris, K. K. Y. Ng, S. Kumar, A. K. Mehta, D. Keitel, T. G. F. Li, and P. Ajith, *The Astrophysical Journal Letters* **874**, L2 (2019), 1901.02674.
- [20] X. Liu, I. Magaña Hernandez, and J. Creighton, *Astrophys. J.* **908**, 97 (2021), 2009.06539.
- [21] R. K. L. Lo and I. Magaña Hernandez, *Phys. Rev. D* **107**, 123015 (2023), 2104.09339.
- [22] P. Amaro-Seoane et al., *Codified Document ET-030-09* (2009).
- [23] J. Janquart, O. A. Hannuksela, K. Haris, and C. Van Den Broeck, *Monthly Notices of the Royal Astronomical Society* **506**, 5430 (2021), 2105.04536.
- [24] S. Goyal, D. Harikrishnan, S. J. Kapadia, and P. Ajith, *Phys. Rev. D* **104**, 124057 (2021), 2106.12466.
- [25] P. Schneider, J. Ehlers, and E. E. Falco, *Gravitational Lenses* (1992).
- [26] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, et al., *Living Reviews in Relativity* **19**, 1 (2016).
- [27] A. K. Y. Li, R. K. L. Lo, S. Sachdev, J. C. L. Chan, E. T. Lin, T. G. F. Li, and A. J. Weinstein, *Phys. Rev. D* **107**, 123014 (2023), 1904.06020.
- [28] L. P. Singer and L. R. Price, *Phys. Rev. D* **93**, 024013 (2016), 1508.03634.
- [29] J. Veitch, V. Raymond, B. Farr, W. Farr, P. Graff, S. Vitale, B. Aylott, K. Blackburn, N. Christensen, M. Coughlin, et al., *Phys. Rev. D* **91**, 042003 (2015), 1409.7215.
- [30] A. Nitz, I. Harry, D. Brown, C. M. Biwer, J. Willis, T. Dal Canton, C. Capano, T. Dent, L. Pekowsky, A. R. Williamson, et al., *gwastro/pycbc*: (2021).
- [31] G. Ashton, M. Hübner, P. D. Lasky, C. Talbot, K. Ackley, S. Biscoveanu, Q. Chu, A. Divakarla, P. J. Easter, B. Goncharov, et al., *The Astrophysical Journal Supplement Series* **241**, 27 (2019), 1811.02042.
- [32] I. M. Romero-Shaw, C. Talbot, S. Biscoveanu, V. D'Emilio, G. Ashton, C. P. L. Berry, S. Coughlin, S. Galadage, C. Hoy, M. Hübner, et al., *Monthly Notices of the Royal Astronomical Society* **499**, 3295 (2020), 2006.00714.
- [33] L. P. Singer, H.-Y. Chen, D. E. Holz, W. M. Farr, L. R. Price, V. Raymond, S. B. Cenko, N. Gehrels, J. Cannizzo, M. M. Kasliwal, et al., *The Astrophysical Journal Supplement Series* **226**, 10 (2016), 1605.04242.
- [34] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. J. Forsteza, and A. Bohé, *Phys. Rev. D* **93**, 044006 (2016), 1508.07250.
- [35] A. Bohé, L. Shao, A. Taracchini, A. Buonanno, S. Babak, I. W. Harry, I. Hinder, S. Ossokine, M. Pürrer, V. Raymond, et al., *Phys. Rev. D* **95**, 044028 (2017), 1611.03703.
- [36] A. Buikema, C. Cahillane, G. L. Mansell, C. D. Blair, R. Abbott, C. Adams, R. X. Adhikari, A. Ananyeva, S. Appert, K. Arai, et al., *Phys. Rev. D* **102**, 062003 (2020), 2008.01301.
- [37] The LIGO Scientific Collaboration, the Virgo Collaboration, the KAGRA Collaboration, A. G. Abac, I. Abouelfettouh, F. Acernese, K. Ackley, C. Adamcewicz, S. Adhicary, D. Adhikari, et al., arXiv e-prints arXiv:2508.18082 (2025), 2508.18082.
- [38] J. S. Speagle, *Monthly Notices of the Royal Astronomical Society* **493**, 3132 (2020), 1904.02180.
- [39] L. W. L. Chan, H. W. Y. Wong, I. C. F. Wong, R. K. L. Lo, and T. G. F. Li, *Using overlap of sky localization probability maps for filtering potentially lensed pairs of gravitational-wave signals — data release* (2025), URL <https://doi.org/10.5281/zenodo.17177201>.