

Mitigating the Counterpart Selection Effect for Standard Sirens

Hsin-Yu Chen¹,² Colm Talbot^{2,3} and Eve A. Chase⁴

¹*Department of Physics, The University of Texas at Austin, 2515 Speedway, Austin, Texas 78712, USA*

²*Department of Physics and Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, Massachusetts 02139, USA*

³*Kavli Institute for Cosmological Physics, The University of Chicago, 5640 South Ellis Avenue, Chicago, Illinois 60637, USA*

⁴*Center for Theoretical Astrophysics, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA*

 (Received 25 July 2023; revised 12 February 2024; accepted 16 April 2024; published 9 May 2024)

The disagreement in the Hubble constant measured by different cosmological probes highlights the need for a better understanding of the observations or new physics. The standard siren method, a novel approach using gravitational-wave observations to determine the distance to binary mergers, has great potential to provide an independent measurement of the Hubble constant and shed light on the tension in the next few years. To realize this goal, we must thoroughly understand the sources of potential systematic bias of standard sirens. Among the known sources of systematic uncertainties, selection effects originating from electromagnetic counterpart observations of gravitational-wave sources may dominate the measurements with percent-level bias and no method to mitigate this effect is currently established. In this Letter, we develop a new formalism to mitigate the counterpart selection effect. We show that our formalism can reduce the systematic uncertainty of standard siren Hubble constant measurement to less than the statistical uncertainty with a simulated population of 200 observations ($\lesssim 1\%$) for a realistic electromagnetic emission model. We conclude with how to apply our formalism to different electromagnetic emissions and observing scenarios.

DOI: [10.1103/PhysRevLett.132.191003](https://doi.org/10.1103/PhysRevLett.132.191003)

Introduction.—The current expansion rate of the Universe, the Hubble constant (H_0), is used to determine the age of the Universe. However, the percent-level inconsistency between different measurements of this quantity [1–6] indicates either an insufficient understanding of the Universe or the experiments. Gravitational-wave (GW) standard siren observations offer a promising independent route to resolve this mystery in cosmology [7–14].

The distance to a given binary can be determined from GW observations; meanwhile, multiple techniques are available to estimate the redshift of the binaries [7,8,15–19]. With estimates of both distance and redshift, binaries observed in GWs can serve as so-called “standard sirens” to measure H_0 and other cosmological parameters. In particular, the standard siren measurements from binary neutron star mergers have ideal coverage in redshift ($z \lesssim 0.1$) [20,21], well-controlled systematics, and are expected to achieve percent-level H_0 precision in the coming years, showing great potential to play a critical role in the Hubble tension problem [11–14].

However, the standard siren method is not completely systematic-free. Systematic uncertainties associated with the standard siren method have been explored recently [22–28]. The dominant systematic uncertainty was long thought to be instrumental calibration uncertainty, but a recent study of GW observations showed that calibration is unlikely to be an issue [24]. The reconstruction of peculiar velocity fields around the host galaxies can lead to percent-level bias

for nearby sources, but the majority of GW events are expected to lie at further distances and are less affected by this bias [26–28]. Other known sources of systematics, e.g., GW waveform accuracy [29–31] and instrumental nonstationary noise [25], are expected to lead to bias at the subpercent level. Therefore, the remaining percent-level systematic, the electromagnetic (EM) counterpart selection effect [23], may be the dominant source of bias.

The identification of EM counterparts of GW sources is the most promising scenario to determine the redshift for standard siren measurements. These “bright sirens” allow for precise measurements of the redshift for GW sources. However, unlike GW signals, the luminosity of EM emissions from binary mergers is highly uncertain. For example, although it is known that a short gamma-ray burst and kilonova emission may accompany a binary neutron star (BNS) merger [32], the angular profile of the EM signals and their emission mechanisms are undetermined [33–35]. Such uncertainty could lead to a selection bias when considering multiple GW-EM events for bright siren H_0 analysis [23]. This bias can be corrected if the EM emission model is known [23,36]. In this Letter, we present a new formalism to mitigate the EM counterpart selection effect when the EM emission model is unknown, a more likely scenario in the near future.

We first describe our formalism, followed by realistic examples assuming the EM emissions are anisotropic. We then demonstrate how to generalize our formalism to

different EM emissions and observing scenarios in the discussion.

Mitigating the counterpart selection effect.—In order to measure the Hubble constant to sufficient precision, data from multiple pairs of joint GW-EM observations are combined. If we denote the GW and EM observational data as $\mathcal{D} = (\mathcal{D}_{\text{GW}}, \mathcal{D}_{\text{EM}})$, the probability distribution of H_0 given the data (posterior) can be written using Bayes' theorem [10,12],

$$p(H_0|\vec{\mathcal{D}}) \sim \pi(H_0) \prod_i \frac{\int \mathcal{L}(\vec{\mathcal{D}}_i|\vec{\Theta}, H_0) p(\vec{\Theta}|H_0) d\vec{\Theta}}{\int_{\vec{\mathcal{D}} > \vec{\mathcal{D}}_{\text{th}}} \mathcal{L}(\vec{\mathcal{D}}|\vec{\Theta}, H_0) p(\vec{\Theta}|H_0) d\vec{\Theta} d\vec{\mathcal{D}}}, \quad (1)$$

where $\vec{\mathcal{D}} = (\vec{\mathcal{D}}_1, \vec{\mathcal{D}}_2, \vec{\mathcal{D}}_3 \dots)$ includes the data from events $i = 1, 2, 3 \dots$, $\vec{\mathcal{D}}_{\text{th}}$ denotes the detection threshold of detectors (e.g., the minimum signal-to-noise ratio in each GW detector, the limiting magnitude for each filter of the telescope, the minimum number of positive detection over certain EM follow-up cadence, etc.), $\pi(H_0)$ is the prior on H_0 , and $\vec{\Theta}$ is a collection of binary physical parameters, such as luminosity distance, redshift, inclination angle, and mass. We fix other cosmological parameters to the Planck values assuming flat- Λ CDM cosmology [1], while they can also be inferred by replacing H_0 with a set of cosmological parameters of interest, $\Omega_{\text{cosmo}} = (H_0, \Omega_m, \Omega_\Lambda \dots)$, throughout this Letter.

The event likelihood $\mathcal{L}(\mathcal{D}_i|\vec{\Theta}, H_0)$ is the value of the probability density function of possible data evaluated at the observed data \mathcal{D}_i given the binary physical parameters $\vec{\Theta}$ and the Hubble constant H_0 [10,12]. This likelihood depends on the GW and EM emission models of the binary as well as the sensitivities of the GW and EM observatories. We write the dependency on GW and EM emission models and observatory sensitivities explicitly as $\mathcal{L}(\mathcal{D}_i|\vec{\Theta}, H_0, \vec{\alpha}, \vec{\beta})$,

where $\vec{\alpha}$ describes the GW emission model and observatory sensitivity, and $\vec{\beta}$ describes the EM emission model and sensitivity.

How the binary physical parameters $\vec{\Theta}$ affect the GW luminosity is well understood from general relativity. With known GW observatory sensitivities, the detectability of GW signals from binary with physical parameters $\vec{\Theta}$ can be estimated, and therefore we assume $\vec{\alpha}$ is known. This is not the case for EM counterpart emission. Because of the uncertainty of EM emission models, even if the EM observatory sensitivity is known, $\vec{\beta}$ is not known. If the effect of $\vec{\beta}$ is ignored, as has been done for previous bright siren analyses (e.g., [10,37]), the H_0 estimate can be biased. This is known as the *counterpart selection effect*.

We mitigate this by estimating H_0 and $\vec{\beta}$ simultaneously. To do so, we analyze all BNS events regardless of whether they have observable EM counterparts. As in Eq. (1), we write the posterior as

$$p(H_0, \vec{\beta}|\vec{\mathcal{D}}, \vec{\alpha}) \sim \pi(H_0, \vec{\beta}) \prod_i \frac{\int \mathcal{L}(\vec{\mathcal{D}}_i|\vec{\Theta}, H_0, \vec{\alpha}, \vec{\beta}) p(\vec{\Theta}|H_0) d\vec{\Theta}}{\int_{\vec{\mathcal{D}} > \vec{\mathcal{D}}_{\text{th}}} \mathcal{L}(\vec{\mathcal{D}}|\vec{\Theta}, H_0, \vec{\alpha}, \vec{\beta}) p(\vec{\Theta}|H_0) d\vec{\Theta} d\vec{\mathcal{D}}}. \quad (2)$$

Note that $\vec{\alpha}$ is assumed to be known. Also, $\vec{\alpha}$ and $\vec{\beta}$ do not affect the intrinsic distribution of the binary physical parameters $p(\vec{\Theta}|H_0)$.

Since EM counterparts used in standard siren analysis will mostly be found by follow-up of GW detections, the events we use for the analysis are solely determined by the GW selection function. That is, the integral over \mathcal{D} in the denominator of Eq. (2) includes events with and without observable EM counterparts, regardless of their EM observabilities. The denominator of Eq. (2) in the product is then

$$\begin{aligned} \int_{\vec{\mathcal{D}} > \vec{\mathcal{D}}_{\text{th}}} \mathcal{L}(\mathcal{D}|\vec{\Theta}, H_0, \vec{\alpha}, \vec{\beta}) p(\vec{\Theta}|H_0) d\vec{\Theta} d\mathcal{D} &= \int_{\vec{\mathcal{D}}_{\text{GW}} > \vec{\mathcal{D}}_{\text{GW,th}}, \vec{\mathcal{D}}_{\text{EM}} > \vec{\mathcal{D}}_{\text{EM,th}}} \mathcal{L}(\vec{\mathcal{D}}|\vec{\Theta}, H_0, \vec{\alpha}, \vec{\beta}) p(\vec{\Theta}|H_0) d\vec{\Theta} d\vec{\mathcal{D}} \\ &+ \int_{\vec{\mathcal{D}}_{\text{GW}} > \vec{\mathcal{D}}_{\text{GW,th}}, \vec{\mathcal{D}}_{\text{EM}} \leq \vec{\mathcal{D}}_{\text{EM,th}}} \mathcal{L}(\vec{\mathcal{D}}|\vec{\Theta}, H_0, \vec{\alpha}, \vec{\beta}) p(\vec{\Theta}|H_0) d\vec{\Theta} d\vec{\mathcal{D}} \\ &= \int_{\vec{\mathcal{D}}_{\text{GW}} > \vec{\mathcal{D}}_{\text{GW,th}}} \mathcal{L}(\vec{\mathcal{D}}|\vec{\Theta}, H_0, \vec{\alpha}) p(\vec{\Theta}|H_0) d\vec{\Theta} d\vec{\mathcal{D}}. \end{aligned}$$

This is the same as that in existing bright siren inference when no EM selection is in effect.

In the following, we demonstrate our method using a simple model for counterpart selection effects due to anisotropic EM emission. We will demonstrate other scenarios in the discussion.

If the EM emission is anisotropic, the inclination angle of a binary ι will affect the furthest luminosity distance D_L at which counterparts can be observed. For a given telescope configuration (filter, exposure time, camera, etc.), we define the telescope's maximum observable luminosity distance of the merger as $\epsilon(\iota, \vec{\beta})$.

We first explicitly separate $\vec{\Theta}$ into relevant physical parameters (D_L, i) from other physical parameters $\vec{\Theta}'$ ($\vec{\Theta} \supset \{D_L, i, \vec{\Theta}'\}$) and recognize that the distance is uniquely determined by z and H_0 , $\mathcal{L}(\mathcal{D}|\vec{\Theta}, H_0, \vec{\alpha}, \vec{\beta}) = \mathcal{L}(\mathcal{D}|D_L(z, H_0), i, \vec{\Theta}', \vec{\alpha}, \vec{\beta})$. Since counterparts can only be found for $D_L \leq \epsilon(i, \vec{\beta})$, we distinguish between cases where we can and cannot identify a counterpart,

$$\mathcal{L}(\mathcal{D}|D_L(z, H_0), i, \vec{\Theta}', \vec{\alpha}, \vec{\beta}) = \begin{cases} \mathcal{L}[\mathcal{D}_{\text{GW}}, \mathcal{D}_{\text{EM}}|D_L(z, H_0), i, \vec{\Theta}', \vec{\alpha}] \mathcal{H}[D_L(z, H_0) - \epsilon(i, \vec{\beta})] & \text{(Counterpart)} \\ \mathcal{L}[\mathcal{D}_{\text{GW}}|D_L(z, H_0), i, \vec{\Theta}', \vec{\alpha}] \mathcal{H}[\epsilon(i, \vec{\beta}) - D_L(z, H_0)]. & \text{(No counterpart)} \end{cases} \quad (3)$$

The quantity \mathcal{H} is the Heaviside function (In practice EM instrumental noise will lead to a different functional form for the selection function. The Heaviside function can be trivially replaced by an appropriate sigmoid function.). This formalism allows for the use of events with and without EM counterparts. Even if no counterpart is identified, the luminosity distance and inclination angle space are constrained.

Application to simulated observations.—To demonstrate our method, we follow the method in [38] to simulate 200 $1.4\text{--}1.4M_\odot$ BNS detected by LIGO-Hanford, LIGO-Livingston, and Virgo operating at the proposed fourth observing run (O4) sensitivity [20,39,40]. The GW detection threshold is set at a network signal-to-noise ratio of 12 [20,21]. We use the Bayesian algorithm developed in [38] to estimate the line-of-sight luminosity distance-inclination angle posterior when EM counterparts are found and the sky direction of the GW sources is determined, and Bilby [41–43] for full parameter estimation when the events are dark.

We consider two kilonova emission models. Both models are selected from a grid of 900 two-dimensional axisymmetric radiative transfer simulations from [44], which spans the full range of anticipated kilonova properties. All simulated models are rendered at 54 inclination angles, each subtending an equal solid angle from 0° to 180° (see Ref. [44] for more details). Of these 900 simulations, we select two simulated models with significantly different angular profiles. Model 1 displays substantial angular variation, typical of kilonova emission, with face-on inclinations resulting in brighter emission. For more “realistic” kilonovae, we selected dynamical and wind ejecta component masses that result in an r -process abundance pattern similar to the neutron-capture elements observed in the Solar System [45]. All other model properties were selected to result in maximal angular variation. Specifically, model 1 represents a kilonova with a toroidal dynamical ejecta component with a mass of $0.01M_\odot$ and mean velocity of $0.3c$ in addition to a “peanut-shaped,” high- Y_e wind ejecta component with a mass of $0.03M_\odot$ and mean velocity of $0.15c$.

For the second kilonova model, we intentionally chose a model that has minimal dependence on the inclination angle in order to test the performance of our method when the EM emission is isotropic. We selected the kilonova with

the least angular variation of all 900 simulated models in [44]. Model 2 represents a kilonova with a toroidal dynamical ejecta component with a mass of $0.001M_\odot$ and mean velocity of $0.05c$ in addition to a spherical, low- Y_e wind ejecta component with a mass of $0.1M_\odot$ and mean velocity of $0.3c$. This model was also used as an example of minimal angular variation in Fig. 3 of [44].

Next, we explore the detectability of the kilonovae at different inclination angles for a variety of optical or near-infrared instruments and filters (Table I). Limiting magnitudes are based on telescope designed sensitivities or from performance in the LIGO-Virgo third observing run, as detailed in [46]. The time of observation after the merger and the exposure time assumed are also listed. We note that this represents a selection of possible limiting magnitudes, and a wide range of instrument sensitivity is possible with different exposure times, background contributions, and environment factors. Our choice of the telescope configurations is not intended to compare the performance of the instruments, but instead represents a wide variety of observing scenarios. We follow [46] to combine simulated observer-frame spectroscopic kilonova emission with an instrument’s bandpass filter function for a variety of luminosity distances. By comparing to the limiting magnitude of the selected telescope configurations, we are able to estimate the maximum observable luminosity distance for the kilonova as a function of binary inclination angle, at a given observation time. In Fig. 1 we show how the

TABLE I. The instruments, filters, exposure time (Exp. time), limiting magnitude (AB) and observing time (Obs. time) after merger we explored in this paper. We employ typical exposure times and limiting magnitudes, as detailed in [46], but note that these observing conditions are highly variable.

Instrument	Filter	Exp. time		Obs. time	
		(s)	m_{lim}	(hrs)	Refs.
DECam	i	90	22.5	18	[47]
DECam	z	90	21.8	12	[47]
ZTF	g	30	20.8	12	[48]
ZTF	r	30	20.6	12	[48]
VISTA	Y	360	21.5	24	[49,50]
VISTA	J	360	21.0	48	[49,50]
VRO	r	30	24.2	12	[51]
VRO	y	30	22.3	24	[51]

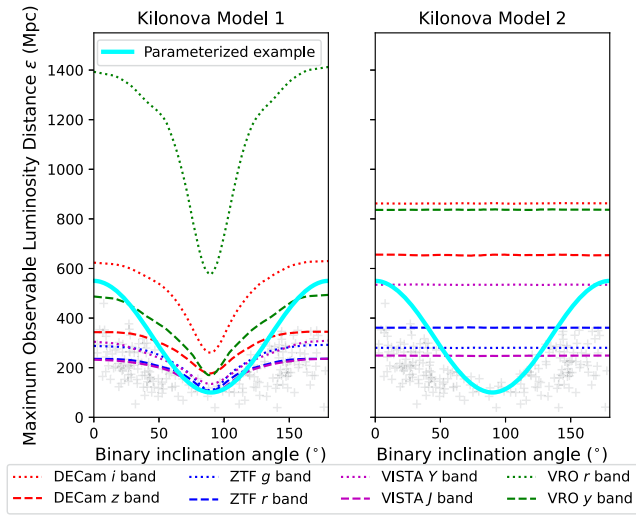


FIG. 1. Maximum observable luminosity distance as a function of the binary inclination angle for the two kilonova models and different telescope configurations (listed in Table I) we consider. The gray crosses mark the inclination angle and luminosity distance of our simulated BNS detections. We also give an example of the function [Eq. (4)] we used to parameterize the shape of curves in this figure (cyan line, $\beta_1 = 450$, $\beta_2 = 100$ in this example).

maximum observable luminosity distance depends on the binary inclination angle with dashed and dotted curves. As expected, the results for kilonova model 1 are significantly more sensitive to the inclination angle than model then parametrize the maximum observable distance as

$$\epsilon(t, \vec{\beta}) = \beta_1 \cos^2(i) + \beta_2 \text{ Mpc}, \quad (4)$$

and show an example of this model in cyan in Fig. 1. This is a generic functional form that describes a lower observable distance when the binary is edge-on, a fairly common feature among kilonova simulations. We note that more refined models may be adopted when larger observed populations are available.

For each of our observing scenarios, we determine which of the 200 simulated BNS signals would have identifiable counterparts based on their luminosity distance and inclination angle. We then use EMCEE [52] to infer (H_0, β_1, β_2) .

Result.—Among the two kilonova models and eight telescope configurations we consider, some of the configurations are able to find the kilonovae for all BNSs we simulated, so we do not expect counterpart selection effect in H_0 inference. We skip these in the presentation of results.

In Fig. 2 we show the inferred median and 68% symmetric credible interval for H_0 for the remaining configurations with (green) and without (orange) correcting for the counterpart selection effect. The results have marginalized over $\vec{\beta}$. In blue we show the results when all BNS have an identified EM counterpart. When not correcting for selection effects, we see up to 2% bias in H_0

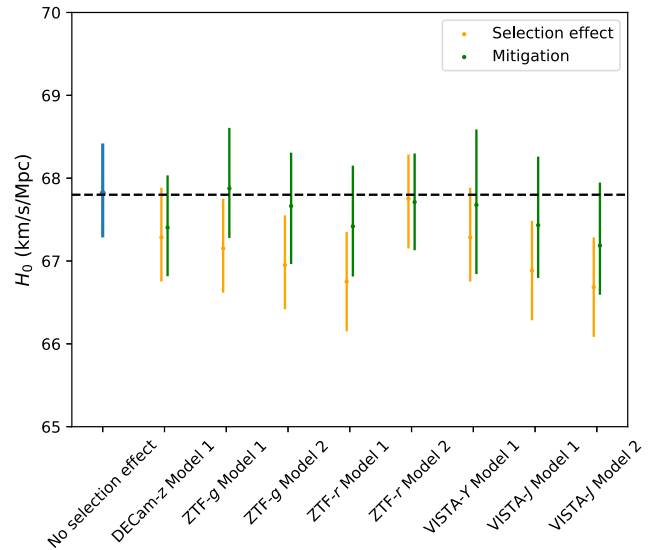


FIG. 2. The median and symmetric 68% credible interval for H_0 inferred using 200 simulated BNS detections. The horizontal axis labels the telescope configurations (cf. Table I) and kilonova models (cf. Fig. 1). *Selection effect* (orange). When H_0 inference ignores the presence of counterpart selection effect. *Mitigation* (green). H_0 inference using the formalism we present in Eq. (2). For comparison, we also show the H_0 inferred when all counterparts are captured and there is no counterpart selection effect (blue). The horizontal dashed line denotes the H_0 we picked for the simulations.

in our simulations. Our formalism reduces the systematic bias to less than the statistical uncertainty ($\sim 1\%$) in all cases for our simulated observations (the true value is contained within the 68% credible interval). Additionally, missing some of the EM counterparts naturally leads to larger H_0 measurement uncertainties. In our simulations, we find 3%–50% larger uncertainties when counterpart selection effects are present.

We note that to elucidate the origin of Hubble tension, we will likely need fewer than 200 BNSs [11–14] and it is unlikely to observe all of them in O4 as in our simulations [53,54]. We combined more events than needed in order to avoid statistical fluctuation in the demonstration and to reveal any underlying systematic uncertainty. However, our formalism can be applied to other GW detector sensitivities and network configurations.

Discussion.—In this Letter, we present a new bright siren inference formalism that mitigates the EM counterpart selection effect. By including GW events both with and without counterparts, our formalism successfully mitigates the systematic bias introduced by the counterpart selection effect for a simulated population of 200 observations. The method does not require precise assumptions of the EM models, avoiding additional modeling systematics. It is possible there are remaining systematics below the statistical uncertainty due to the simplified functional form employed to parameterize the maximum observable

luminosity distance [Eq. (4)]. Future developments in observations, theories, and numerical methods will further reduce any such bias.

Even with the nearly isotropic kilonova model (model 2), we see 2% bias in H_0 for some telescope configurations. The bias in these scenarios mainly originates from the limitation of EM instrument sensitivities, as some face-on binaries lie beyond the instrument's maximum observable distance (Fig. 1, right panel). Still, our formalism is able to mitigate the bias. We note that in this case, the extra degree of freedom allowed by our nearly agnostic prior on β_1 ($\beta_1 \in [0, 1000]$ Mpc) leads to the median recovered H_0 value falling below the truth. Since the correct model is contained within the prior distribution, we can safely expect the true value to continue to be contained within the statistical uncertainty as more events are observed. In the future, by comparing different EM observations and inclination angle of the events, it will become clear whether kilonovae are better modeled as having isotropic emission, and we can adopt a constant maximum observable distance model [i.e., $\beta_1 = 0$ in Eq. (4)]. If so, we find that the H_0 bias becomes less than 0.35% with our formalism.

Although we demonstrate our formalism with the selection effect associated with anisotropic EM emission observed by a selected set of instrument configurations, the formalism can easily be generalized to many other scenarios including: (i) *Mass-dependent EM emission*: The EM emission from BNSs and neutron star-black hole mergers strongly depends on the component masses of the binaries [55,56] and can lead to a similar counterpart selection effect in bright siren inference. In order to mitigate this effect, one could choose an appropriate functional form $\epsilon(m_1, m_2, \vec{\beta})$. Since the component masses of binaries are better measured in GWs than the inclination angle, the mass-associated counterpart selection effect could be easier to mitigate. (ii) *Multiple instruments and/or different observing strategies*: The counterparts of GW events will likely be discovered by different instruments or observing strategies (e.g., searching for counterparts at different depths, colors, and time after mergers.) Each variation leads to a different counterpart selection effect. We can take $\vec{\beta} = \vec{\beta}_a, \vec{\beta}_b, \vec{\beta}_c \dots$ to represent each variation. If the EM observations are random to GW events, i.e., the search instruments and strategies do not depend on binary's physical parameters, the posterior for H_0 can be written as

$$p(H_0 | \{\vec{\mathcal{D}}_a, \vec{\mathcal{D}}_b, \vec{\mathcal{D}}_c \dots\}) \sim \pi(H_0) \prod_{k=a,b,c \dots} \mathcal{L}(\vec{\mathcal{D}}_k | H_0), \quad (5)$$

where $\mathcal{L}(\vec{\mathcal{D}}_k | H_0)$ is the product of the likelihoods for each event observed with scenario k marginalized over $\vec{\beta}_k$. If the choice of instruments or strategies depends on the physical parameters of the GW signals, one can account for this as an additional *known* selection effect. (iii) *Partial sky*

coverage: The large GW sky localization area can make it difficult for a single instrument to cover the entire localization map. If so, the integral over $\vec{\Theta}$ in the numerator of Eq. (2) has to be limited to the sky directions ($\vec{\Omega}_A, \vec{\Omega}_B, \vec{\Omega}_C \dots$) covered by instruments (A, B, C...), respectively. (iv) *No counterpart due to other reasons*: Nondetection of counterparts due to weather, instrument availability, solar position etc., are likely to be random, and so the associated GW events can be removed from the bright siren inference without leading to a selection bias. (v) *Hostless counterpart*: A counterpart may be found without an associated host for the determination of redshift. However, the observation of a counterpart indicates that the event lies within the maximum observable luminosity distance. One can use the counterpart-found scenario [e.g., $\mathcal{H}[D_L(z, H_0) - \epsilon(t, \vec{\beta})]$ case in Eq. (3)] and integrate over possible redshifts in the numerator of Eq. (2).

As LIGO-Virgo-KAGRA continue to observe with improved sensitivities, joint GW-EM detections will push toward percent-level H_0 measurements in the next few years. Our new formalism mitigates a dominant source of systematic uncertainty with minimal dependence on EM modeling for the measurement, ensuring a reliable path toward resolving the H_0 tension.

This analysis was made possible by the NUMPY [57,58], SCIPY [59], MATPLOTLIB [60], EMCEE [52,61], PANDAS [62,63], and ASTROPY [64,65] software packages.

The authors would like to thank Simone Mastrogiovanni for the review and suggestions to this work. The authors would like to thank Maya Fishbach, Jonathan Gair, Daniel Holz, Will Farr, Marko Ristic, and Aaron Zimmerman for useful discussions. H. Y. C. is supported by the National Science Foundation under Grant No. PHY-2308752. H. Y. C. was supported by NASA through NASA Hubble Fellowship Grant No. HST-HF2-51452.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under Contract No. NAS5-26555. C. T. is supported by an MIT Kavli Fellowship and an Eric and Wendy Schmidt AI in Science Fellowship. 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 is LIGO Document Number LIGO-P2300222.

-
- [1] N. Aghanim *et al.* (Planck Collaboration), *Astron. Astrophys.* **641**, A6 (2020); **652**, C4(E) (2021).
 - [2] A. G. Riess, W. Yuan, L. M. Macri, D. Scolnic, D. Brout, S. Casertano, D. O. Jones, Y. Murakami, G. S. Anand, L. Breuval, T. G. Brink, A. V. Filippenko, S. Hoffmann, S. W. Jha, W. D'arcy Kenworthy, J. Mackenty, B. E. Stahl, and W. Zheng, *Astrophys. J. Lett.* **934**, L7 (2022).

- [3] A. G. A. Brown *et al.* (Gaia Collaboration), *Astron. Astrophys.* **650**, C3 (2021).
- [4] D. W. Pesce, J. A. Braatz, M. J. Reid, A. G. Riess, D. Scolnic, J. J. Condon, F. Gao, C. Henkel, C. M. V. Impellizzeri, C. Y. Kuo, and K. Y. Lo, *Astrophys. J. Lett.* **891**, L1 (2020).
- [5] K. C. Wong *et al.*, *Mon. Not. R. Astron. Soc.* **498**, 1420 (2020).
- [6] S. Alam *et al.* (eBOSS Collaboration), *Phys. Rev. D* **103**, 083533 (2021).
- [7] B. F. Schutz, *Nature (London)* **323**, 310 (1986).
- [8] D. E. Holz and S. A. Hughes, *Astrophys. J.* **629**, 15 (2005).
- [9] N. Dalal, D. E. Holz, S. A. Hughes, and B. Jain, *Phys. Rev. D* **74**, 063006 (2006).
- [10] B. P. Abbott *et al.* (LIGO Scientific, Virgo, 1M2H, Dark Energy Camera GW-E, DES, DLT40, Las Cumbres Observatory, VINROUGE, MASTER Collaboration), *Nature (London)* **551**, 85 (2017).
- [11] S. Nissanke, D. E. Holz, N. Dalal, S. A. Hughes, J. L. Sievers, and C. M. Hirata, *arXiv:1307.2638*.
- [12] H.-Y. Chen, M. Fishbach, and D. E. Holz, *Nature (London)* **562**, 545 (2018).
- [13] R. Abbott *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), *Astrophys. J.* **949**, 76 (2023).
- [14] S. M. Feeney, H. V. Peiris, A. R. Williamson, S. M. Nissanke, D. J. Mortlock, J. Alsing, and D. Scolnic, *Phys. Rev. Lett.* **122**, 061105 (2019).
- [15] W. Del Pozzo, *Phys. Rev. D* **86**, 043011 (2012).
- [16] S. R. Taylor, J. R. Gair, and I. Mandel, *Phys. Rev. D* **85**, 023535 (2012).
- [17] C. Messenger and J. Read, *Phys. Rev. Lett.* **108**, 091101 (2012).
- [18] W. M. Farr, M. Fishbach, J. Ye, and D. Holz, *Astrophys. J. Lett.* **883**, L42 (2019).
- [19] J. M. Ezquiaga and D. E. Holz, *Phys. Rev. Lett.* **129**, 061102 (2022).
- [20] B. P. Abbott *et al.* (KAGRA, LIGO Scientific, and Virgo Collaborations), *Living Rev. Relativity* **21**, 3 (2018).
- [21] H.-Y. Chen, D. E. Holz, J. Miller, M. Evans, S. Vitale, and J. Creighton, *Classical Quantum Gravity* **38**, 055010 (2021).
- [22] S. Mastroianni, K. Leyde, C. Karathanasis, E. Chassande-Mottin, D. A. Steer, J. Gair, A. Ghosh, R. Gray, S. Mukherjee, and S. Rinaldi, *Phys. Rev. D* **104**, 062009 (2021).
- [23] H.-Y. Chen, *Phys. Rev. Lett.* **125**, 201301 (2020).
- [24] Y. Huang, H.-Y. Chen, C.-J. Haster, L. Sun, S. Vitale, and J. Kissel, *arXiv:2204.03614*.
- [25] S. Mozzon, G. Ashton, L. K. Nuttall, and A. R. Williamson, *Phys. Rev. D* **106**, 043504 (2022).
- [26] C. Howlett and T. M. Davis, *Mon. Not. R. Astron. Soc.* **492**, 3803 (2020).
- [27] S. Mukherjee, G. Lavaux, F. R. Bouchet, J. Jasche, B. D. Wandelt, S. Nissanke, F. Leclercq, and K. Hotokezaka, *Astron. Astrophys.* **646**, A65 (2021).
- [28] C. Nicolaou, O. Lahav, P. Lemos, W. Hartley, and J. Braden, *Mon. Not. R. Astron. Soc.* **495**, 90 (2020).
- [29] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. X* **9**, 011001 (2019).
- [30] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Astrophys. J. Lett.* **892**, L3 (2020).
- [31] T. Narikawa, N. Uchikata, K. Kawaguchi, K. Kiuchi, K. Kyutoku, M. Shibata, and H. Tagoshi, *Phys. Rev. Res.* **2**, 043039 (2020).
- [32] B. P. Abbott *et al.* (LIGO Scientific, Virgo, Fermi GBM, INTEGRAL, IceCube, AstroSat Cadmium Zinc Telluride Imager Team, IPN, Insight-Hxmt, ANTARES, Swift, AGILE Team, 1M2H Team, Dark Energy Camera GW-EM, DES, DLT40, GRAWITA, Fermi-LAT, ATCA, ASKAP, Las Cumbres Observatory Group, OzGrav, DWF (Deeper Wider Faster Program), AST3, CAASTRO, VINROUGE, MASTER, J-GEM, GROWTH, JAGWAR, CaltechNRAO, TTU-NRAO, NuSTAR, Pan-STARRS, MAXI Team, TZAC Consortium, KU, Nordic Optical Telescope, ePESSTO, GROND, Texas Tech University, SALT Group, TOROS, BOOTES, MWA, CALET, IKI-GW Follow-up, H.E.S.S., LOFAR, LWA, HAWC, Pierre Auger, ALMA, Euro VLBI Team, Pi of Sky, Chandra Team at McGill University, DFN, ATLAS Telescopes, High Time Resolution Universe Survey, RIMAS, RATIR, SKA South Africa/MeerKAT Collaborations), *Astrophys. J. Lett.* **848**, L12 (2017).
- [33] B. D. Metzger, *Living Rev. Relativity* **23**, 1 (2019).
- [34] S. Darbha and D. Kasen, *Astrophys. J.* **897**, 150 (2020).
- [35] J. D. Lyman *et al.*, *Nat. Astron.* **2**, 751 (2018).
- [36] S. Gagnon-Hartman, J. Ruan, and D. Haggard, *Mon. Not. R. Astron. Soc.* **520**, 1 (2023).
- [37] K. Hotokezaka, E. Nakar, O. Gottlieb, S. Nissanke, K. Masuda, G. Hallinan, K. P. Mooley, and A. T. Deller, *Nat. Astron.* **3**, 940 (2019).
- [38] H.-Y. Chen, S. Vitale, and R. Narayan, *Phys. Rev. X* **9**, 031028 (2019).
- [39] J. Aasi *et al.* (LIGO Scientific Collaboration), *Classical Quantum Gravity* **32**, 074001 (2015).
- [40] F. Acernese *et al.* (Virgo Collaboration), *Classical Quantum Gravity* **32**, 024001 (2015).
- [41] G. Ashton *et al.*, *Astrophys. J. Suppl. Ser.* **241**, 27 (2019).
- [42] I. M. Romero-Shaw *et al.*, *Mon. Not. R. Astron. Soc.* **499**, 3295 (2020).
- [43] G. Ashton and C. Talbot, *Mon. Not. R. Astron. Soc.* **507**, 2037 (2021).
- [44] R. T. Wollaeger, C. L. Fryer, E. A. Chase, C. J. Fontes, M. Ristic, A. L. Hungerford, O. Korobkin, R. O'Shaughnessy, and A. M. Herring, *Astrophys. J.* **918**, 10 (2021).
- [45] M. Ristic, E. M. Holmbeck, R. Wollaeger, O. Korobkin, E. Champion, R. O'Shaughnessy, C. L. Fryer, C. J. Fontes, M. R. Mumpower, and T. M. Sprouse, *Astrophys. J.* **956**, 64 (2023).
- [46] E. A. Chase, B. O'Connor, C. L. Fryer, E. Troja, O. Korobkin, R. T. Wollaeger, M. Ristic, C. J. Fontes, A. L. Hungerford, and A. M. Herring, *Astrophys. J.* **927**, 163 (2022).
- [47] M. Soares-Santos *et al.* (DES Collaboration), *Astrophys. J. Lett.* **823**, L33 (2016).
- [48] E. C. Bellm *et al.*, *Publ. Astron. Soc. Pac.* **131**, 018002 (2019).
- [49] R. G. McMahon, M. Banerji, E. Gonzalez, S. E. Koposov, V. J. Bejar, N. Lodieu, and R. Rebolo (VHS Collaboration), *The Messenger* **154**, 35 (2013), <https://ui.adsabs.harvard.edu/abs/2013Msngr.154...35M>.

- [50] M. Banerji *et al.*, *Mon. Not. R. Astron. Soc.* **446**, 2523 (2015).
- [51] Ž. Ivezić *et al.*, *Astrophys. J.* **873**, 111 (2019).
- [52] D. Foreman-Mackey, D.W. Hogg, D. Lang, and J. Goodman, *Publ. Astron. Soc. Pac.* **125**, 306 (2013).
- [53] R. Abbott *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), *Phys. Rev. X* **13**, 041039 (2023).
- [54] R. Abbott *et al.* (KAGRA, Virgo, and LIGO Scientific Collaborations), *Phys. Rev. X* **13**, 011048 (2023).
- [55] C.J. Krüger and F. Foucart, *Phys. Rev. D* **101**, 103002 (2020).
- [56] F. Foucart, T. Hinderer, and S. Nissanke, *Phys. Rev. D* **98**, 081501(R) (2018).
- [57] T. Oliphant, *NUMPY: A Guide to NUMPY* (Trelgol Publishing, USA, 2006).
- [58] C.R. Harris *et al.*, *Nature (London)* **585**, 357 (2020).
- [59] P. Virtanen *et al.*, *Nat. Methods* **17**, 261 (2020).
- [60] J.D. Hunter, *Comput. Sci. Eng.* **9**, 90 (2007).
- [61] D. Foreman-Mackey, D.W. Hogg, D. Lang, and J. Goodman, *Publ. Astron. Soc. Pac.* **125**, 306 (2013).
- [62] The pandas development team, pandas-dev/pandas: PANDAS (2020).
- [63] Wes McKinney, in *Proceedings of the 9th Python in Science Conference*, edited by Stéfan van der Walt and Jarrod Millman (2010), pp. 56—61, <https://dx.doi.org/10.25080/Majora-92bf1922-00a>.
- [64] T.P. Robitaille *et al.* (Astropy Collaboration), *Astron. Astrophys.* **558**, A33 (2013).
- [65] A. Price-Whelan *et al.*, *Astron. J.* **156**, 123 (2018).