Systematic uncertainty of standard sirens from the viewing angle of binary neutron star inspirals

Hsin-Yu Chen

1Black Hole Initiative, Harvard University, Cambridge, Massachusetts 02138, USA

The independent measurement of Hubble constant with gravitational-wave standard sirens will potentially shed light on the tension between the local distance ladders and Planck’s experiments. Therefore, thorough understanding of the sources of systematic uncertainty for the standard siren method is crucial. In this paper, we focus on two scenarios that will potentially dominate the systematic uncertainty of standard sirens. First, simulations of electromagnetic counterparts of binary neutron star mergers suggest aspherical emissions, so the binaries available for the standard siren method can be selected by their viewing angles. This selection effect can lead to $\gtrsim 2\%$ bias in Hubble constant measurement even with mild selection, making the standard siren method difficult to resolve the tension in Hubble constant. Second, if the binary viewing angles are constrained by the electromagnetic counterpart observations but the bias of the constraints is not controlled under $\sim 10^3\%$, the resulting systematic uncertainty in Hubble constant will be $> 3\%$. In addition, we find that both of the systematics cannot be fully disclosed by the viewing angle measurement from gravitational-wave observations. Comparing to the known dominant systematic uncertainty for standard sirens, the gravitational-wave calibration uncertainty, the effects from viewing angle can be more prominent.

Introduction—Gravitational-wave (GW) standard sirens provide an independent way to measure the Hubble constant ($H_0$), which is crucial for our understanding of the evolution of the Universe [1, 2]. Currently, the $H_0$ measurements from cosmic microwave background [3] and some local distance ladders [4–6] appear to be inconsistent at $> 2\sigma$ level. Independent $H_0$ measurement with the standard siren method has shown its potential to resolve the inconsistency [2, 7].

GW observations of compact binary mergers probe the luminosity distance ($D_L$) of the mergers directly. If the mergers also have electromagnetic (EM) counterparts [8], e.g. short gamma-ray bursts (GRBs) or kilonova emissions that come with binary neutron star mergers (BNSs), the observation of the counterparts could allow for precise sky localization of the mergers and identification of the host galaxies [9, 10]. With the luminosity distance of the GW source and the redshift of the host galaxy, cosmological parameters can be constrained. This is the so-called standard siren method with the use of EM counterparts. GW170817 was the first successful standard siren [2]. Several forecasts predict that a $2\% H_0$ measurement can be achieved by combining $\sim 50$ BNSs with identified hosts in a few years [7, 11, 12].

In order to resolve the $H_0$ controversy, the systematic uncertainty in the standard siren method has to be well-understood. One dominant systematics comes from the calibration of amplitude measurement of GW signals. The calibration uncertainty currently leads to $\lesssim 2\%$ systematics in the GW distance measurement, while this uncertainty is expected to reduce in the future [13, 14]. Another source of systematics is the reconstruction of the peculiar velocity fields around the host galaxies [15, 16]. However, most of the BNSs will be detected by Advanced LIGO-Virgo beyond 100 Mpc, where the effect of peculiar motions on the galaxy redshift measurement becomes less relevant. Other known sources of systematic uncertainty, e.g. the accuracy of GW waveforms [18], are expected to play a secondary role.

In this paper, we highlight two sources of systematic uncertainties for standard sirens that have not been thoroughly discussed before. Both of the systematics are related to the EM counterpart observations and the viewing angle of the binaries ($\zeta$) [1]. First, simulations of BNSs suggest that their EM emissions are likely aspherical [19–22]. For example, the brightness of kilonovae can have a factor of 2-3 angular dependent variation. The color of kilonovae can also change with the viewing angle. The variations lead to angular dependent EM observing probability for BNSs (e.g., [23]). If this EM viewing angle selection effect is not accounted for correctly, $H_0$ measurement will be biased after combining multiple standard sirens. Second, EM observations of BNSs provide constraint on the viewing angle. The viewing angle of BNS GW170817 [24] has been reconstructed from the profiles of its EM emissions [25, 26] and from the observations of the jet motions [27]. These reconstructions help breaking the degeneracy between the luminosity distance and inclination angle of BNSs in GW parameter estimations [28], improving the precision of distance measurement, and reducing the $H_0$ measurement uncertainty [29, 30]. However, if the EM constraints on the viewing angle are systematically biased, the distance and $H_0$ estimation will also be biased.

1 Since the EM counterpart emissions barely depend on the direction of the binary rotation (clockwise or counterclockwise), in this paper we define the viewing angle as $\zeta \equiv \min(\theta_{JN}, 180^\circ − \theta_{JN})$, where $\theta_{JN}$ denotes the inclination angle of the binary.
We find that both of the systematics can yield significant bias in $H_0$ measurement, undermining the standard siren’s potential to resolve the $H_0$ tension. Since both of the scenarios we discuss originate from the uncertainty of EM emissions, we also explore if it is possible to independently measure the systematics by analyzing the GW viewing angle estimations. Unfortunately, most of the events suffer from the large uncertainty of the estimations and the systematics can be difficult to disclose.

**Simulations**—We simulate 1.4M$_{⊙}$-1.4M$_{⊙}$ BNS detections with the `IMRPhenomPv2` waveform and assumed a network signal-to-noise ratio of 12 GW detection threshold. The BNS astrophysical rate does not evolve, and the BNSs are uniformly distributed in comoving volume before detections. We use Advanced LIGO-Virgo O4 sensitivity [21] for the simulations [4]. Planck cosmology is assumed ($H_0 = 67.4 \text{ km/s/Mpc}$, $Ω_m = 0.315$, $Ω_k = 0$) [3]. We then use the simulated detections $D_{GW}$ to calculate the distance-inclination angle posteriors, $p(D_L, θ_{IN}|D_{GW})$, with the algorithms developed in [28]. After marginalizing $p(D_L, θ_{IN}|D_{GW})$ over the inclination angle, we use the distance posteriors for the estimation of $H_0$. Following the methods in [7], we combine multiple $H_0$ posteriors from different detections to produce the final $H_0$ posterior. We repeat the simulations 100 times and report the average for the results throughout this paper.

**Systematics from EM viewing angle selection effect**—If the EM counterpart emissions are aspherical, BNSs with some viewing angles could be easier to observe in EM than from other directions. The subset of BNSs with available EM counterparts for standard sirens will then be selected. Suppose the data from GW and EM are denoted as $D_{GW}$ and $D_{EM}$ respectively, one can follow [7] [32] to write down the $H_0$ likelihood for an event as:

$$p(D_{GW}, D_{EM}|H_0) = \frac{\int p(D_{GW}|Θ)p(D_{EM}|Θ)p_{pop}(Θ|H_0)dΘ}{\int p_{det}(Θ)p_{det}(Θ)p_{pop}(Θ|H_0)dΘ},$$

(1)

where $Θ$ represents all the binary parameters, such as the mass, spin, luminosity distance, sky location, and inclination angle etc.,

$$p_{det}(Θ) \equiv \int_{D>\text{Threshold}} p(D|Θ)dD,$$

(2)

and $p_{pop}(Θ|H_0)$ is the population distribution of binaries with parameters $Θ$ in the Universe (also see [32] for more details). If the EM observing probability depends on the binary parameters (such as the viewing angle), $p(D_{EM}|Θ)$ has to change accordingly and Equation (1) have to be reevaluated. However, if such dependency is unknown or ignored, Equation (1) and the combined $H_0$ posteriors from multiple events will be biased.

Here we explore two examples of EM observing probability dependency on the viewing angle [3]. First, we assume only BNSs with viewing angle less than $ζ_{max}$ are observable in EM. Smaller $ζ_{max}$ represents stronger selection since the viewing angle is more limited. In Figure 1 we show the 1-σ uncertainty in $H_0$ for different $ζ_{max}$ if 50 events are combined. Without knowing the selection on viewing angle, we find the $H_0$ measurement significantly biased even if $ζ_{max}$ is as large as $\sim 60°$ (the band W/o correction). Only as a demonstration, we also show the $H_0$ uncertainty assuming the viewing angle selection $ζ_{max}$ is perfectly known (the band With correction). If $ζ_{max}$ is known, $p(D_{EM}|Θ)$ in Equation (1) is 0 when $ζ > ζ_{max}$.

On the other hand, not all EM emissions have a sharp decline beyond a viewing angle. Here we also consider a second example, in which the EM observing probability is a continuous function of viewing angle: $E(ζ) = \frac{1}{\sqrt{ζ}}$ for $ζ < 5°$ and $E(ζ) = 0$ for $ζ ≥ 5°$.
0.5(\cos(\zeta) + 1). With this assumption, the EM observing probability is 1 for face-on binaries, and 0.5 for edge-on binaries. Without correction, we find the 1-\sigma uncertainty in \(H_0\) for 50 events lying between [67.5, 70.2]km/s/Mpc, equivalent to \(\sim 2\%\) bias in \(H_0\).

Since the EM observing probability is unclear, a possible way to access the viewing angle selection effect is to analyze the GW viewing angle estimation of the events with EM counterparts. We try to estimate \(\zeta_{\text{max}}\) from the first example above with data from \(N\) events \({\mathcal{D}}_1, \ldots, \mathcal{D}_N\):

\[
p(\zeta_{\text{max}}|\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_N) = \frac{p(\zeta_{\text{max}}) \prod_{k=0}^{N} p(\mathcal{D}_k|\zeta_{\text{max}})}{\prod_{k=0}^{N} p(\mathcal{D}_k)}
\]

\[
= p(\zeta_{\text{max}}) \prod_{k=0}^{N} \int_0^{\pi/2} \frac{p(\zeta|\mathcal{D}_k)p(\zeta_{\text{max}}|\zeta, \mathcal{D}_k)}{p(\zeta_{\text{max}})} d\zeta
\]

\[
= p(\zeta_{\text{max}}) \prod_{k=0}^{N} \int_0^{\pi/2} \frac{p(\zeta|\mathcal{D}_k)p(\zeta_{\text{max}})}{p(\zeta)} d\zeta
\]

\[
= p(\zeta_{\text{max}}) \prod_{k=0}^{N} \int_0^{\zeta_{\text{max}}} p(\zeta|\mathcal{D}_k) d\zeta
\]

\[
= \int_0^{\zeta_{\text{max}}} p(\zeta|\mathcal{D}_k) d\zeta. \tag{3}
\]

The first line comes from the fact that each event are independent. The third line considers \(p(\zeta_{\text{max}}|\zeta, \mathcal{D}_k) = p(\zeta_{\text{max}}|\zeta)\), and the last line takes \(p(\zeta|\zeta_{\text{max}}) \propto p(\zeta)\) for \(\zeta < \zeta_{\text{max}}\). Equation (3) can then be calculated from the prior on viewing angle \(p(\zeta)\) \(\text{Eq}(\text{3})\) and the GW viewing angle posterior \(p(\zeta|\mathcal{D}_k)\) for each event \(\text{Eq}(\text{3})\). Without any prior on \(\zeta_{\text{max}}\) (i.e. \(p(\zeta_{\text{max}})\) is taken as a constant), in Figure 2 we show the 1-\sigma uncertainty of the \(\zeta_{\text{max}}\) posterior \(\text{Eq}(\text{3})\) as a function of the maximum EM viewing angle of 50 simulated BNSs. We find that \(\zeta_{\text{max}}\) can only be confined to \(\sim 20^\circ\) 1-\sigma uncertainty. In addition, the estimated \(\zeta_{\text{max}}\) is biased for small \(\zeta_{\text{max}}\) because GW viewing angle posteriors typically peak around 30° with about 20° uncertainty \(\text{Eq}(\text{3})\). Small \(\zeta_{\text{max}}\) is therefore difficult to estimate even if all of the events with EM counterparts are face-on/off.

**Systematics from biased EM constraint on viewing angle**– The angular dependency of EM emissions can be used to estimate the viewing angle of BNSs from their EM observations. However, lack of robust understanding of the EM emission model can lead to biased interpretation of the viewing angle.

Suppose the EM observations suggest a viewing angle of \(\zeta_{\text{EM}}\) with 1-\sigma uncertainty of \(\sigma_\zeta\), we can multiply following prior with the GW distance-inclination joint posterior

\[
p(D_L, \theta_{JN} | \mathcal{D}_{GW})\text{ of the BNS:}
\]

\[
\Gamma(\theta_{JN}) = \begin{cases} \mathcal{N}(\theta_{JN}; \zeta_{\text{EM}}, \sigma_\zeta) & \text{if} \ 0 \leq \theta_{JN} \leq \pi/2 \\ \mathcal{N}(\theta_{JN}; \pi - \zeta_{\text{EM}}, \sigma_\zeta) & \text{if} \ \pi/2 < \theta_{JN} \leq \pi, \end{cases}
\]

where \(\mathcal{N}(\theta_{JN}; \zeta_{\text{EM}}, \sigma_\zeta)\) denotes a normal distribution with mean \(\zeta_{\text{EM}}\) and standard deviation \(\sigma_\zeta\) evaluated at \(\theta_{JN}\). Such prior reduces the uncertainty in inclination angle, and the distance is better measured after the joint posterior, \(p(D_L, \theta_{JN} | \mathcal{D}_{GW})\), is integrated over \(\theta_{JN}\). Improved distance estimate leads to more precise Hubble constant measurement \(\text{Eq}(\text{3})\). However, if the EM constraint on the viewing angle is off by

\[
\Delta\zeta_{\text{sys}} = \zeta_{\text{EM}} - \zeta_{\text{real}},
\]

where \(\zeta_{\text{real}}\) denotes the real viewing angle of the event, the distance and the \(H_0\) measurements will be biased. For single event the bias in \(H_0\) may not be obvious, because the statistical uncertainty in \(H_0\) dominates the overall uncertainty. The bias will become clear after the \(H_0\) posteriors are combined over multiple events. In Figure 3 we show the extent of overall bias in \(H_0\) if the EM constraint on viewing angle is always off by \(\Delta\zeta_{\text{sys}}\) for 20 events.

When the viewing angles are overestimated (underestimated), the distances are underestimated (overestimated) and the overall \(H_0\) is overestimated (underestimated). Smaller \(\sigma_\zeta\) affects the \(H_0\) measurement more significantly for the same \(\Delta\zeta_{\text{sys}}\). Although \(\Delta\zeta_{\text{sys}}\) is unlikely to be a constant across different events, our simulations provide the allowed range of systematic uncertainty for EM constrained viewing angle. In general, \(\Delta\zeta_{\text{sys}}\) has to be \(\lesssim 10^\circ\) to be accurate enough to address the tension between *Planck* and the local distance ladders.
Next, we wonder if a comparison between the GW and EM measurement of the viewing angle will help disclosing the bias in EM interpretations. Suppose the viewing angle posteriors from GW and EM for a BNS are $Y(\zeta)$ and $\varepsilon(\zeta)$ respectively, we can define their difference as

$$\Delta_{\text{EM-GW}} \equiv \int_0^{\pi/2} \int_0^{\pi/2} (\zeta_2 - \zeta_1) \times Y(\zeta_1) \times \varepsilon(\zeta_2) \, d\zeta_1 \, d\zeta_2. \quad (4)$$

The uncertainty in GW and EM posteriors both contribute to the overall uncertainty of $\Delta_{\text{EM-GW}}$. We find that the average of $\Delta_{\text{EM-GW}}$ over 20 BNSs traces $\Delta_{\text{sys}}$ with $1 - \sigma$ uncertainty $> 18\degree$ (Figure 3). This statistical uncertainty of $\Delta_{\text{EM-GW}}$ is not small enough to confine $\Delta_{\text{sys}}$ to the accuracy for $H_0$ measurement described above, leaving the $H_0$ systematic from biased EM constraint in viewing angle unresolved.

**Discussion**—In this paper we evaluate the extent of bias in $H_0$ as a result of the geometry of EM emissions from BNSs. Among the two examples of viewing angle selection we present, the maximum viewing angle selection may happen due to the choice of kilonova observing strategies or the sharp decline beyond a viewing angle for short GRB emission. In particular, in the third generation GW detector era [34,35], short GRBs will likely become the major EM counterparts for BNSs at high redshifts. Study of the maximum viewing angle of GRBs will be crucial to correct the selection effect for standard sirens.

On the other hand, the example of continuous viewing angle selection applies to current kilonova observations. Simulations show that edge-on BNSs are more difficult to localize [28], and their kilonova emissions can be redder and dimmer [22]. In both examples, we find $\gtrsim 2\%$ bias in $H_0$ if the selection is not well-understood.

We note that in reality other binary parameters will also affect the EM observing probability. Therefore, more complete considerations of EM models and projections of EM observing probability for future telescopes involved in the search for EM counterparts will result in more accurate estimation of the bias in $H_0$. Unlike the viewing angle measurement, some parameters, such as the mass, are estimated precise enough from GW signals for the selection effect to be taken care of. Overall, we find the selection over viewing angle discussed in this paper the most subtle and difficult to resolve.

If the viewing angle selection effect is significant, it is possible to reconstruct the selection by comparing the number of BNSs with and without EM counterparts. The distribution of viewing angle for BNSs detected in GWs is well-understood [33]. For example, it is known that about 15% of BNSs have viewing angle larger than 60°. If 15% of BNSs miss counterparts, one explanation is that the maximum EM viewing angle is around 60°. A reconstruction for short GRB observations has been shown in [36]. However, the reconstruction for kilonova population will be more difficult since their EM observing probability will have more complicated dependency on the viewing angle. Such reconstruction can also be easily contaminated by other factors that affect the EM observing probability and will have to be evaluated carefully.

Although our discussion focuses on BNSs, there are simulations suggesting stronger viewing angle dependency for EM counterparts of neutron star-black hole.

---

Note that for high redshift sources the framework for the selection effect will be the same as we demonstrate in Equation 1, while the inference will be on $H(\zeta)$.

**FIG. 3.** Hubble constant measurement uncertainty ($1-\sigma$) from 20 standard sirens as a function of the systematic bias in the binary viewing angle constrained by EM observations. Three different statistical uncertainties in the EM-constrained viewing angle ($\sigma_\zeta = 5\degree, 10\degree, 20\degree$) are shown. The Hubble constant used for the simulations is 67.4 km/s/Mpc.

**FIG. 4.** The average difference between EM and GW viewing angle posteriors $\Delta_{\text{EM-GW}}$ for 20 BNSs with EM posteriors systematically off by $\Delta_{\text{sys}}$. The $1 - \sigma$ uncertainty of the difference for three EM posterior statistical uncertainties, $\sigma_\zeta = 5\degree, 10\degree, 20\degree$, are 18.5°, 20°, and 24°, respectively. The grey dashed line is the equal-axis line to guide the eye.
mergers [22]. Therefore neutron star-black hole mergers can possibly introduce larger bias when they are used as standard sirens [37].

On the other hand, if the geometry of EM emissions is used to confine the BNSs’ viewing angle, the systematic uncertainty in viewing angle introduced by the EM interpretations has to be less than 10°. Since the binary rotational axis doesn’t have to be perfectly aligned with the major axis of EM emissions, and the geometry of EM emissions is unknown, to control the systematics of EM constraint viewing angle can be challenging. We also show that the comparison between EM and GW viewing angle posteriors can help estimating the systematics, but the precision of the estimation may not be good enough to completely remove the bias.

We note that the standard siren method we discuss in this paper relies on the observations of EM counterparts and the measurements of the BNSs’ redshift. A complimentary approach of the standard siren method doesn’t require the EM counterparts but make use of galaxy catalogs (also known as “dark sirens”) may help deducing the systematics discussed in this paper. However, the dark siren approach will suffer from lower $H_0$ precision and other sources of systematics [7, 38], making it complicated and difficult to contribute to the issues.

Finally, the calibration uncertainty in GWs currently dominates the known systematic uncertainty for standard sirens. The bias in $H_0$ from calibration can be as large as $\sim 2\%$ [13, 14]. Both of the systematics we find in this work can introduce $H_0$ bias larger than $2\%$ (Figure 1 and 3). In summary, the systematic uncertainty from viewing angle for standard sirens can be a major challenge to resolve the tension in Hubble constant, and we look forward to future development to address this topic.

acknowledgments— We acknowledge valuable discussions with Sylvia Biscoveanu, Michael Coughlin, Carl-Johan Haster, Daniel Holz, Kwan-Yeung Ken Ng, and Salvatore Vitale. HYC was supported by the Black Hole Initiative at Harvard University, which is funded by grants from the John Templeton Foundation and the Gordon and Betty Moore Foundation to Harvard University.
[24] B. Abbott et al. (LIGO Scientific, Virgo), Phys. Rev. Lett. 119, 161101 (2017), 1710.05832.
[25] D. Finstad, S. De, D. A. Brown, E. Berger, and C. M. Biwer, Astrophys. J. Lett. 860, L2 (2018), 1804.04179.
[26] S. Dhawan, M. Bulla, A. Goobar, A. Sagüés Carraçedo, and C. N. Setzer, Astrophys. J. 888, 67 (2020), 1909.13810.
[27] K. P. Mooley, A. T. Deller, O. Gottlieb, E. Nakar, G. Hallinan, S. Bourke, D. A. Frail, A. Horesh, A. Corsi, and K. Hotokezaka, Nature (London) 561, 355 (2018), 1806.09693.
[28] H.-Y. Chen, S. Vitale, and R. Narayan, Physical Review X 9, 031028 (2019), 1807.05226.
[29] C. Guidorzi, R. Margutti, D. Brout, D. Scolnic, W. Fong, K. D. Alexander, P. S. Cowperthwaite, J. Annis, E. Berger, P. K. Blanchard, et al., Astrophys. J. Lett. 851, L36 (2017), 1710.06426.
[30] K. Hotokezaka, E. Nakar, O. Gottlieb, S. Nissanka, K. Masuda, G. Hallinan, K. P. Mooley, and A. T. Deller, Nature Astronomy 3, 940 (2019), 1806.10596.
[31] B. Abbott et al. (KAGRA, LIGO Scientific, VIRGO), Living Rev. Rel. 21, 3 (2018), 1304.0670.
[32] I. Mandel, W. M. Farr, and J. R. Gair, Mon. Not. R. Astron. Soc. 486, 1086 (2019), 1809.02063.
[33] B. F. Schutz, Classical and Quantum Gravity 28, 125023 (2011), 1102.5421.
[34] D. Reitze et al., Bull. Am. Astron. Soc. 51, 035 (2019), 1907.04833.
[35] M. Maggiore et al., JCAP 03, 050 (2020), 1912.02622.
[36] A. Farah, R. Essick, Z. Doctor, M. Fishbach, and D. E. Holz, arXiv e-prints arXiv:1912.04906 (2019), 1912.04906.
[37] S. Vitale and H.-Y. Chen, Phys. Rev. Lett. 121, 021303 (2018), 1804.07337.
[38] R. Gray, I. Magaña Hernandez, H. Qi, A. Sur, P. R. Brady, H.-Y. Chen, W. M. Farr, M. Fishbach, J. R. Gair, A. Ghosh, et al., arXiv e-prints arXiv:1908.06050 (2019), 1908.06050.