A Standard Siren Measurement of the Hubble Constant without the Electromagnetic Counterpart

Abstract

We perform a statistical standard siren analysis of GW170817. Our analysis does not utilize knowledge of NGC 4993 as the unique host galaxy of the optical counterpart to GW170817. Instead, we consider each galaxy within the GW170817 localization region as a potential host; combining the redshifts from all of the galaxies with the distance estimate from GW170817 provides an estimate of the Hubble constant, $H_0$. Considering all galaxies brighter than 0.626$L_b$ as equally likely to host a binary neutron star merger, we find $H_0 = 77^{+13}_{-16}$ km s$^{-1}$ Mpc$^{-1}$ (maximum a posteriori and 68.3% highest density posterior interval; assuming a flat $H_0$ prior in the range [10, 220] km s$^{-1}$ Mpc$^{-1}$). We explore the dependence of our results on the thresholds by which galaxies are included in our sample, and we show that weighting the host galaxies by stellar mass or star formation rate provides entirely consistent results with potentially tighter constraints. By applying the method to simulated gravitational-wave events and a realistic galaxy catalog we show that, because of the small localization volume, this statistical standard siren analysis of GW170817 provides an unusually informative (top 10%) constraint. Under optimistic assumptions for galaxy completeness and redshift uncertainty, we find that dark binary neutron star measurements of $H_0$ will converge as 40%/√(N), where N is the number of sources. While these statistical estimates are inferior to the value from the counterpart standard siren measurement utilizing NGC 4993 as the unique host, $H_0 = 76^{+15}_{-13}$ km s$^{-1}$ Mpc$^{-1}$ (determined from the same publicly available data), our analysis is a proof-of-principle demonstration of the statistical approach first proposed by Bernard Schutz over 30 yr ago.

Key words: distance scale – gravitational waves

1. Introduction

The first multimessenger detection of a binary neutron star (BNS) merger, GW170817, by LIGO (Abbott et al. 2016) and Virgo (Acernese et al. 2015) enabled the first standard siren measurement of the Hubble constant, $H_0$, ushering in the era of gravitational-wave (GW) cosmology (Abbott et al. 2017a, 2017c, 2017d). This $H_0$ measurement combined the luminosity distance to the source, as measured from the GW signal (Schutz 1986), with the known redshift of the host galaxy, NGC 4993. NGC 4993 was identified as the unique host galaxy following the discovery of an optical transient located only ~10 arcsec from NGC 4993 (Abbott et al. 2017d; Coulter et al. 2017; Soares-Santos et al. 2017). The probability of a chance coincidence between the GW signal and the optical transient was estimated to be $\lesssim 0.5\%$ (Soares-Santos et al. 2017), and the probability of a chance association between the optical transient and NGC 4993 is $\leq 0.004\%$ (Abbott et al. 2017a).

The original proposal by Schutz (1986) to measure the Hubble constant with GW detections of compact binary mergers did not involve electromagnetic counterparts. Instead, Schutz considered bright galaxies in the GW localization region as potential hosts to the merger. Each galaxy provides a redshift that, when combined with the GW-measured luminosity distance, gives a separate estimate of $H_0$. The final $H_0$ measurement from a single event is the sum of all contributions from the individual galaxies. The first detailed exploration of this method on simulated data, and with the first use of a galaxy catalog (Aihara et al. 2011), was by Del Pozzo (2012). An up-to-date forecast incorporating realistic detection rates, galaxy peculiar velocities, large-scale structure, and additional considerations can be found in Chen et al. (2018). We refer to this approach of measuring $H_0$ as the “statistical” method (Schutz 1986; MacLeod & Hogan 2008; Petiteau et al. 2011; Chen et al. 2018), compared with the “counterpart” method in which an electromagnetic (EM) counterpart provides a unique host galaxy association. In the limit where the GW event is well-localized that there is only one potential host galaxy in the GW localization error box (Chen & Holz 2016), the statistical method reduces to the counterpart method. In the opposite limit, where the GW event is poorly localized, there are so many potential host galaxies that the distinct peaks from individual galaxies are washed out, and the $H_0$ measurement is uninformative (Chen et al. 2018).

The statistical approach may be the only way to do standard siren science with binary black holes, because they are not expected to have EM counterparts. We emphasize that although the statistical measurements for a given event are inferior to the counterpart case, combining many of these measurements leads to increasingly precise constraints (Schutz 1986; Del Pozzo 2012; Chen et al. 2018; Nair et al. 2018). In ground-based gravitational wave detector networks, the rate of detection of binary black holes is significantly higher than that for neutron stars (Abbott et al. 2016, 2017b, 2017c), although the higher rate is not expected to compensate for the inferior constraints (Chen et al. 2018). Nonetheless, the black hole systems can be observed to much higher redshifts, potentially providing constraints on the evolution history of the universe out past the turnover between dark matter and dark energy domination (Del Pozzo 2012; Domink et al. 2015; Belczynski et al. 2016; Fishbach et al. 2018). Because these systems are farther away, however, it will be a greater challenge to supply a sufficiently complete galaxy catalog.

In this paper we carry out a measurement of $H_0$ using the GW data from GW170817 and a catalog of potential host galaxies within the GW localization region. In other words, we explore how tight the $H_0$ measurement from GW170817 would have been if an EM counterpart had not been detected or if a unique host galaxy had not been identified. We present our methods in Section 2, a discussion of the galaxy selection in Section 3, a discussion of the GW constraints in Section 4, results in Section 5, and conclude in Section 6.
2. Methods

We follow the statistical framework presented in Chen et al. (2018; see also Del Pozzo 2012; Gray et al. 2019). We include the role of GW selection effects, galaxy catalog incompleteness, galaxy luminosities, and redshift uncertainties in our analysis. The posterior on $H_0$ given the GW and EM data, $x_{GW}$ and $x_{EM}$, is:

$$
p(H_0|x_{GW}, x_{EM}) = \frac{p_0(H_0)}{\beta(H_0)} \int p(x_{GW}|D_L(z, H_0), \Omega) \times p(x_{EM}|z, \Omega)p_0(z, \Omega)d\Omega dz,
$$

(1)

where $D_L(z, H_0)$ is the luminosity distance of a source at redshift $z$ for a given $H_0$ (fixing other cosmological parameters to the Planck values; Ade et al. 2016)\(^\text{86}\), $\Omega$ is the sky position, and $\beta(H_0)$ is a normalization term to ensure that the likelihood normalizes to 1 when integrated over all detectable GW and EM data sets (Mandel et al. 2016). The term $p_0(H_0)$ represents the prior on the Hubble constant. A detailed derivation of Equation (1), including the role of the normalization term $\beta(H_0)$, is provided in Appendix A.

As first emphasized by Schutz (1986), the GW signal from a compact binary coalescence allows for a direct measurement of the distance to the source, as well as its sky location. This measurement is represented in the GW likelihood term, $p(x_{GW}|D_L, \Omega)$, which is the probability of the GW data in the presence of signal from a compact binary with parameters $D_L$ and $\Omega$ marginalized over the other parameters of the signal (including the inclination angle, component masses, spins, and/or tides). The corresponding posterior $p(D_L, \Omega|x_{GW}) \propto p(x_{GW}|D_L, \Omega)p_0(D_L, \Omega)$ is summarized in the three-dimensional sky map, which provides a fit to the posterior samples provided by the GW parameter estimation pipeline LALInference (Veitch et al. 2015; Singer et al. 2016a, 2016b). For this analysis, we use the publicly released three-dimensional sky map from Abbott et al. (2019; see Sections 4 and 5). To get the likelihood from the posterior probability, we must first divide out the default “volumetric” distance prior, $p_0(D_L, \Omega) \propto D_L^3$.

Meanwhile, the EM likelihood term $p(x_{EM}|z, \Omega)$ is the probability of the electromagnetic data in the presence of signal from a compact binary with parameters $z$ and $\Omega$. In the absence of an EM counterpart and/or a host galaxy identification, we assume the measurement $p(x_{EM}|z, \Omega)$ is completely uninformative, and set:

$$
p(x_{EM}|z, \Omega) \propto 1.
$$

(2)

In this case, the redshift information enters only through the prior term, $p_0(z, \Omega)$, which we take to be a galaxy catalog. The detection of an electromagnetic counterpart typically results in $p(x_{EM}|z, \Omega)$ being strongly peaked around some $\Omega$ allowing the identification of a host galaxy. We note that in some cases an optical transient may be identified, but it may not be possible to uniquely identify the associated host galaxy. In these circumstances one could perform a pencil-beam survey of the region surrounding the transient (e.g., at distances of $\lesssim 100$ kpc from the line of sight to the transient), and sharply reduce the relevant localization volume (Chen et al. 2018). This reduces the number of potential host galaxies, and thereby improves the measurement.

The galaxy catalog term $p_0(z, \Omega)$ is given by:

$$
p_0(z, \Omega) = p_{\text{cat}}(z, \Omega) + (1 - f)p_{\text{miss}}(z, \Omega),
$$

(3)

where $p_{\text{cat}}$ is a catalog of known galaxies, $p_{\text{miss}}$ represents the distribution of missing galaxies, and $f$ denotes the overall completeness fraction of the catalog. The contribution from the known galaxies is:

$$
p_{\text{cat}}(z, \Omega) = \sum_{i} n_i(z_i, \sigma_i; \Omega),
$$

(4)

where $z_i, \Omega_i$ denotes the (peculiar velocity-corrected) “Hubble” redshifts and sky coordinates of all galaxies in the catalog, and $n(z_i, \sigma_i; \Omega)$ denotes the normal probability density function with mean $\mu$ and standard deviation $\sigma$ evaluated at $x$. To account for peculiar velocity uncertainties, which can be significant for nearby sources, we assume that the true Hubble velocity is normally distributed about the measured Hubble velocity with an uncertainty of $\sigma_c$ (Scolnic et al. 2018). On the other hand, the uncertainty on the sky coordinates of galaxies in the catalog is negligible for our purposes, so we always approximate $n(\Omega, \sigma_\Omega; \Omega) = \delta(\Omega - \Omega)$.

The weights $w_i$ can be chosen to reflect the a priori belief that a galaxy could host a GW source. For example, setting all weights to $w_i = \frac{1}{N_{\text{gal}}}$ corresponds to equal probability for each galaxy to host a gravitational wave source. In general, because we might expect that the BNS rate is traced by some combination of stellar mass and/or star formation rate (Phinney 1991; Leibler & Berger 2010; Fong & Berger 2013; Chruslinska et al. 2018), we may assign unequal weights to galaxies based on these (or any other relevant observable) quantities, ensuring that the weights sum to unity. In the following, we use a galaxy’s B-band luminosity as a proxy for its star formation rate, and its K-band luminosity as a proxy for its total stellar mass (Bell et al. 2003; Singer et al. 2016a); these are very rough estimates, but serve to demonstrate the method. In these cases, we apply weights proportional to B-band or K-band luminosity, $w_i \propto L_B$ or $w_i \propto L_K$, and explore the dependence of the result on those of weightings.

To calculate the term $p_{\text{miss}}$ in Equation (3), we assume that on large scales, the distribution of galaxies, $p_0(z, \Omega)$, is uniform in comoving volume. Let $p_{\text{vol}}(z, \Omega)$ denote the cosmologically homogeneous and isotropic distribution normalized over the volume contained in the range $z_{\text{min}} < z < z_{\text{max}}$ considered in our analysis. (The result does not depend on the choice of $z_{\text{min}}$ or $z_{\text{max}}$ provided that the interval encompasses all possible redshifts of the source for all allowed values of $H_0$.) Assuming all galaxies are weighted equally, the distribution of missing galaxies is written as:

$$
p_{\text{miss}}(z, \Omega) = \frac{1 - P_{\text{complete}}(z)p_{\text{vol}}(z, \Omega)}{(1 - f)},
$$

(5)

where $P_{\text{complete}}(z)$ is the probability that a galaxy at redshift $z$ is in the catalog, and the completeness fraction $f$ is given by:

$$
f = \int_{z_{\text{min}}}^{z_{\text{max}}} P_{\text{complete}}(z)p_{\text{vol}}(z, \Omega) dz d\Omega.
$$

(6)

We can similarly add galaxy weightings to an incomplete catalog by computing the luminosity distribution of the

\(^{86}\) For the redshifts considered here, $z \lesssim 0.05$, other cosmological parameters affect the distance–redshift relation at the subpercent level, and so our analysis is insensitive to their precise values.
“missing galaxies” as a function of redshift, $p(L|z, \text{missing})$. We calculate this distribution by assuming that the luminosities of the missing galaxies together with those in the catalog are distributed according to the Schechter function. Then, the weights of the missing galaxies are given by:

$$w(z) \propto \int L p(L|z, \text{missing}) \, dL,$$

and, weighting each missing galaxy by its luminosity, Equation (5) becomes:

$$P_{\text{miss}}(z, \Omega) \propto w(z)P_{\text{miss}}(z, \Omega).$$

Note that we have assumed that the coverage of the catalog is uniform over the sky and so $P_{\text{complete}}$ is independent of $\Omega$. (This is true over the relevant sky area for the present analysis, but the method can be easily generalized to add an $\Omega$-dependence.) Alternate approaches of taking into account the incompleteness of galaxy catalogs are being explored in Gray et al. (2019). However, in the present case of a single nearby source where the catalog is largely complete, the differences in results from the various approaches are small, and in particular well within the statistical uncertainties. In Section 3, the completeness function $P_{\text{complete}}$ is estimated for the galaxy catalog used in the analysis.

To demonstrate the statistical method, we apply the analysis described above to 249 simulated BNS GW detections from the First Two Years (F2Y) catalog (Singer et al. 2014) and the MICE mock galaxy catalog (Carretero et al. 2015, 2017; Crocce et al. 2015; Fosalba et al. 2015a, 2015b; Hoffmann et al. 2015). We assign each BNS detection from the F2Y 2016 scenario (roughly corresponding to O2) to a galaxy in the MICE catalog with a redshift that matches the injection distance and assumed $H_0$ value ($H_0 = 70 \, \text{km s}^{-1} \, \text{Mpc}^{-1}$). For each event, we rotate the sky coordinates of the galaxies in the catalog so that the sky position of the host galaxy matches the true sky position of the BNS injection. We then carry out the statistical method using the three-dimensional sky map for each mock BNS and the galaxies in MICE, assuming no peculiar velocities or incompleteness, and assigning weights to the galaxies in MICE so that the redshift distribution matches the injected redshift distribution of the F2Y data set, $p(z) \propto z^2$. This last step is necessary in order to ensure that the selection effects are incorporated consistently between the injections and the likelihood. The results are shown in Figure 1. Even in the best-case scenario of perfectly known galaxy redshifts and a complete catalog, the $H_0$ posteriors from most individual events are nearly flat over the prior range. Combining the 249 individual events, the final $H_0$ posterior is $H_0 = 70.1^{+1.8}_{-1.9} \, \text{km s}^{-1} \, \text{Mpc}^{-1}$ (68.3% credible interval), corresponding to a convergence rate of $\sim 40\%/\sqrt{N}$ where $N$ is the number of events, consistent with Chen et al. (2018). For counterpart standard sirens the convergence is $\sim 15\%/\sqrt{N}$ (Chen et al. 2018). As is visible in the figure, we confirm that the method is unbiased, with the result for large numbers of detections approaching the true value of $H_0$. We note that most of the simulated detections in the F2Y data set have much larger localization volumes than GW170817, which was an unusually loud, nearby source that was detected while all three detectors were operational. Therefore, we expect the statistical $H_0$ measurement from GW170817 to be unusually informative compared to an average event. We quantify this expectation in Section 5.

3. Galaxy Catalogs

To measure $H_0$ statistically with GW170817, we use version 2.3 of the GLADE galaxy catalog to construct our redshift prior in Equation (3) (Dálya et al. 2018). GLADE provides galaxy redshifts in the heliocentric frame, corrected for peculiar motions using the peculiar velocity catalog of Carrick et al. (2015). For galaxies that are also listed in the group catalog of Kourkchi & Tully (2017), as identified by a common Principal Galaxy Catalog identifier, we apply an additional correction to correct their velocities to the radial velocity of the group. We assume the group velocity is given by the unweighted mean of the velocities of all member galaxies, although we note that for the dominant group containing NGC 4993, careful group modeling has been done (Hjorth et al. 2017). Finally, we correct all heliocentric velocities to the reference frame of the cosmic microwave background (Hinshaw et al. 2009) and assign a 200 km s$^{-1}$ Gaussian uncertainty to the “Hubble velocity” of each galaxy in the catalog (corrected by all peculiar motions; Carrick et al. 2015; Scolnic et al. 2018).

GLADE also provides luminosity information for galaxies, listing apparent magnitudes in the $B$, $J$, $H$, and $K$-bands. We use the reported $B$-band luminosities to characterize the completeness of the catalog (a small fraction of galaxies do not have $B$-band apparent magnitudes reported in the catalog; we remove these galaxies from our analysis, assuming that their magnitudes are below our adopted luminosity cutoff). Following Gehrels et al. (2016) and Arcavi et al. (2017), we adopt $B$-band Schechter function parameters $\phi^* = 5.5 \times 10^{-3} h_{0.7}^{-2} \, \text{Mpc}^{-3}$, $\alpha_B = -1.07$, $L_B^* = 2.45 \times 10^{10} h_{0.7}^2 L_{B,0}$ throughout. The corresponding characteristic absolute magnitude is $M_B^* = -20.47 + 5 \log_{10} h_{0.7}$. We will also consider the $K$-band magnitudes reported in GLADE when applying galaxy weights, and we use the $K$-band Schechter function parameters of $\alpha_K = -1.02$, $M_K^* = -23.55 + 5 \log_{10} h_{0.7}$ (Lu et al. 2016).

Figure 2 summarizes the completeness of GLADE as a function of redshift. We find that GLADE is complete up to redshifts $z \sim 0.06$ for galaxies brighter than $\sim 0.626 h_{0.7}^2 L_B^*$, corresponding to about 0.66 of the Milky Way luminosity for $h_{0.7} = 1$. Galaxies brighter than $\sim 0.626 L_B^*$ make up half of the
With the sky coordinates of the 408 GLADE galaxies, we enclose the 50% and 90% probability regions, respectively, and the shading of the galaxy markers denotes their redshifts, corrected for peculiar and virial motions as described in the text.

Figure 2. Completeness of the GLADE catalog as a function of redshift for galaxies brighter than $0.25L_B^*$ (solid blue curve), $0.05L_B^*$ (dashed green curve), and $0.01L_B^*$ (dotted-dashed orange curve), calculated by comparing the redshift distribution of galaxies in GLADE to a distribution that is constant in comoving volume. For galaxies brighter than $0.626L_B^*$, GLADE is complete across the entire redshift range shown.

Figure 3. Two-dimensional localization region of GW170817 (blue contours) with the sky coordinates of the 408 GLADE galaxies (green crosses) within the 99% localization area and the redshift range $0 < z \lesssim 0.046$ (for an $H_0$ prior range of $H_0 \in [10, 220] \text{ km s}^{-1} \text{ Mpc}^{-1}$). The light and dark blue contours enclose the 50% and 90% probability regions, respectively, and the shading of the galaxy markers denotes their redshifts, corrected for peculiar and virial motions as described in the text.

The total luminosity for the given Schechter function parameters. We find that for $z \lesssim 0.03$, GLADE is complete for galaxies down to 2.5 times dimmer, or $\sim 0.25L_B^*$, corresponding to $M_B = -18.96 + 5 \log_{10} h_027$ (see also Figure 2 of Arcavi et al. 2017). Such galaxies make up 75% of the total $B$-band luminosity. If we consider galaxies down to $\sim 0.05L_B^*$ ($M_B = -17.22 + 5 \log_{10} h_027$), GLADE is $\sim 70\%$ complete at $z \sim 0.03$, and even if we consider galaxies down to $\sim 0.01L_B^*$ ($M_B = -15.47 + 5 \log_{10} h_027$), including 99% of the total $B$-band luminosity, GLADE is $\geq 80\%$ complete for $z \lesssim 0.01$, and $\sim 40\%$ complete at $z \sim 0.03$. In the $K$-band, we find that with our assumed $K$-band Schechter function parameters, GLADE is complete up to $z \sim 0.045$ for galaxies with $L_K > 0.36L_K^*$, which contain 70% of the total $K$-band luminosity, and up to $z \sim 0.03$ for galaxies with $L_K > 0.1L_K^*$, which contain 90% of the total luminosity. For galaxies brighter than $L_K = 0.005L_K^*$, which make up more than 99% of the total $K$-band luminosity, GLADE is $\sim 70\%$ complete at $z = 0.01$.

4. Source Localization and Distance

From the GW data alone, GW170817 is the best-localized GW event to date. The original analysis by the LIGO-Virgo collaboration (Abbott et al. 2017c) reported a 90% localization area of 28 deg$^2$ and a 90% localization volume of 380 Mpc$^3$ (assuming Planck cosmology; Ade et al. 2016), while the most recent analysis (Abbott et al. 2019) improves the 90% localization area to 16 deg$^2$ and the 90% volume to 215 Mpc$^3$. We use this updated three-dimensional sky map (Singer et al. 2016a, 2016b) from Abbott et al. (2019) throughout. Figure 3 shows the two-dimensional sky map together with the galaxies in the GLADE catalog within the localization region. Figure 4 shows that, although there are a total of 408 galaxies within the 99% localization area (see Figure 3), most of the galaxies with high sky-map probability come in a few distinct groups: a dominant group at $z \sim 0.01$ regardless of the assumed luminosity threshold, followed by a secondary group at $z \sim 0.006$ containing only moderately faint galaxies. Therefore, there are only a few distinct redshifts that can possibly correspond to the measured distance of GW170817, and we expect that combining the galaxy catalog with the GW localization will yield an informative measurement of the Hubble constant.

With the data release accompanying Abbott et al. (2019), the LIGO-Virgo collaboration has made the three-dimensional data behind this sky map publicly available at the following url: https://dcc.ligo.org/DocDB/0150/P1800061/009/figure_3.tar.gz.
The three-dimensional sky map also provides an approximation to the luminosity distance posterior along each line of sight. As usual, the distance to GW170817 is determined directly from the gravitational waves, and is calibrated by general relativity (Schutz 1986). No distance ladder is required.

5. Results

We combine the GW distance posterior for GW170817 with the redshift for each potential host galaxy within the localization region. As detailed in Section 2, each galaxy produces a posterior probability for \( H_0 \), and we combine these estimates among all the galaxies in the localization region to arrive at a final estimate for \( H_0 \). We adopt a flat prior in \( H_0 \) over the range \( 10 - 220 \) \( \text{km s}^{-1} \text{Mpc}^{-1} \). The results are presented in Figure 5. Because the galaxies are predominantly found in one galaxy group at \( z \sim 0.01 \), the \( H_0 \) posterior shows a clear peak at \( H_0 \approx 76 \) \( \text{km s}^{-1} \text{Mpc}^{-1} \). And because NGC 4993, the true galaxy host of GW170817, is a member of the group at \( z \sim 0.01 \), we should not be surprised to learn that the peak in \( H_0 \) is consistent with the \( H_0 \) estimate from the GW170817 standard siren measurement including the counterpart (Abbott et al. 2017a). Because this analysis has been performed on a three-dimensional sky map using an approximation to the distance posteriors, rather than using the full three-dimensional LIGO/Virgo posteriors, the results do not agree precisely with those of Abbott et al. (2017a), and in particular the position of the peak in Figure 5 at \( H_0 = 76 \) \( \text{km s}^{-1} \text{Mpc}^{-1} \) instead of \( H_0 = 70 \) \( \text{km s}^{-1} \text{Mpc}^{-1} \). This is because our three-dimensional sky map approximates the distance posterior along each line of sight by a simple two-parameter Gaussian fit (see Equation (1) of Singer et al. 2016a), which is an imperfect approximation to the true, asymmetric distance posterior (Chen & Holz 2017; Del Pozzo et al. 2018). On the other hand, the analysis in Abbott et al. (2017a) utilizes the full distance posterior along the line of sight to NGC 4993 rather than the Gaussian approximation.

Figure 5 shows four different posterior probability distributions, each using a different threshold for the galaxy catalog. In the “assuming counterpart” case, NGC 4993 (which is assumed to be the true host galaxy to GW170817) is given a weight of 1, and all the other galaxies in the localization volume are given a weight of 0. We find \( H_0 = 76_{-15}^{+12} \) \( \text{km s}^{-1} \text{Mpc}^{-1} \) (maximum a posteriori and 68.3% highest density posterior interval) for our default flat prior, or \( H_0 = 74_{-12}^{+18} \) \( \text{km s}^{-1} \text{Mpc}^{-1} \) for a flat-in-log prior (the prior choice in Abbott et al. 2017a). This peak is slightly shifted compared to the result presented in Abbott et al. (2017a), \( H_0 = 70_{-8}^{+12} \) \( \text{km s}^{-1} \text{Mpc}^{-1} \), due to the usage of the Gaussian fit to the distance posterior found in the three-dimensional sky map as discussed above.

The other curves in Figure 5 assume different limiting thresholds for what constitutes a potential host galaxy. For a luminosity threshold of \( L > 0.626 L_\odot^* \), we find \( H_0 = 77_{-18}^{+37} \) \( \text{km s}^{-1} \text{Mpc}^{-1} \). As the threshold is lowered, additional galaxies fall into the sample, and the \( H_0 \) posterior is broadened. For a limiting \( B \)-band magnitude of 0.25\( L_\odot^* \), we need to account for the incompleteness of the galaxy catalog at redshifts \( z \gtrsim 0.03 \), and for 0.01\( L_\odot^* \), we need to account for the incompleteness at \( z \gtrsim 0.01 \), as described in Section 2. The incompleteness correction leads to a slight additional broadening of the \( H_0 \) posterior, but the clear peak at \( H_0 \approx 76 \) \( \text{km s}^{-1} \text{Mpc}^{-1} \) remains; we find \( H_0 = 74_{-15}^{+45} \) \( \text{km s}^{-1} \text{Mpc}^{-1} \) for a luminosity threshold of \( L > 0.25 L_\odot^* \), \( H_0 = 76_{-21}^{+48} \) \( \text{km s}^{-1} \text{Mpc}^{-1} \) for a luminosity threshold of \( L > 0.01 L_\odot^* \). This peak is the result of the galaxy group at \( z \sim 0.01 \), of which NGC 4993 is a member.

The curves in Figure 6 weight each galaxy by its \( B \)-band luminosity (a proxy for its recent star formation history; right) or its \( K \)-band luminosity (a proxy for its stellar mass; left). The peak at \( H_0 \approx 76 \) \( \text{km s}^{-1} \text{Mpc}^{-1} \) becomes more pronounced when galaxies are weighted by their luminosity, as the group containing NGC 4993 consists of many bright, mostly red galaxies. If we assume that the probability of hosting a BNS merger is proportional to a galaxy’s \( B \)-band luminosity, the posterior on \( H_0 \) tightly constrains \( H_0 \) to be in \( [53, 124] \) \( \text{km s}^{-1} \text{Mpc}^{-1} \) (68.3% highest density posterior interval) when applying equal weights to all galaxies brighter than 0.01\( L_\odot^* \) to \( H_0 \) in \( [54, 120] \) \( \text{km s}^{-1} \text{Mpc}^{-1} \). Applying \( K \)-band luminosity weights to galaxies brighter than 0.005\( L_\odot^* \), the 68.3% posterior interval tightens from \( H_0 \) in \( [61, 137] \) \( \text{km s}^{-1} \text{Mpc}^{-1} \) to \( H_0 \) in \( [57, 118] \) \( \text{km s}^{-1} \text{Mpc}^{-1} \). Although these results are suggestive that weighting by stellar-mass or star formation rate may lead to faster convergence, the properties of BNS host galaxies are still uncertain, and it is impossible to establish this definitively with a single event. As the source sample increases it is expected to relate to some combination of these quantities, and incorporating these trends will lead to improvements in the statistical \( H_0 \) analysis.

In order to quantify the degree of information in the GW170817 \( H_0 \) posterior compared to an “average” event as expected from the F2Y data set, we consider the difference in the Shannon entropy between the flat prior and the posterior (see Appendix B; Shannon 1948). We compare this measure of information for the statistical GW170817 \( H_0 \) posterior to the individual statistical \( H_0 \) posteriors from each of the simulated

88 The upper limits of the 68.3% highest density posterior intervals that we report here are especially sensitive to the upper limit we consider for the \( H_0 \) prior, 220 \( \text{km s}^{-1} \text{Mpc}^{-1} \).
BNS events in Section 2. We find that for a flat prior in the (relatively narrow) range $H_0 \in [40, 100]$ km s$^{-1}$ Mpc$^{-1}$, the information gained by applying the statistical method to GW170817 is 0.34 bits. Meanwhile, the median information in an individual posterior shown in Figure 1 is only 0.047 bits, so that GW170817 is in the top ~10% of informative events, even under optimistic assumptions for the simulated detections (i.e., complete galaxy catalogs and perfect redshift measurements). As expected, GW170817 provides an unusually good statistical $H_0$ constraint.

For the purposes of this calculation, we use the $K$-band luminosity-weighted $L_0 > 0.1L^*_{K}$ posterior shown in the right panel of Figure 6 as a representative posterior for the statistical GW170817 $H_0$ measurement. Over the wider prior $H_0 \in [10, 220]$ km s$^{-1}$ Mpc$^{-1}$ shown, the information difference between the posterior and the prior is 0.67 bits. The counterpart GW170817 $H_0$ measurement (dotted–dashed pink curve in Figure 5) has an information gain of 1.55 bits with respect to the wide prior.

6. Conclusion

We perform a statistical standard siren measurement of the Hubble constant with GW170817. This analysis is the first application of the measurement originally proposed over 30 yr ago by Schutz (1986). We find that the excellent localization of GW170817, together with the large-scale structure causing galaxies to cluster into distinct groups, enables an informative measurement of $H_0$ even in the absence of a unique host galaxy identification. Including generic and flexible assumptions regarding the luminosities of BNS host galaxies, we find a peak at $H_0 \approx 76$ km s$^{-1}$ Mpc$^{-1}$ at ~2.4–3.7 times the prior probability density. We find the possibility of improved constraints when weighting the potential host galaxies by stellar mass and star formation rate. Including all galaxies brighter than 0.01 $L_{B}^*$ (including 99% of the total blue luminosity) we find $H_0 = 76^{+3}_{-2}$ km s$^{-1}$ Mpc$^{-1}$, or $H_0 = 76^{+2}_{-3}$ km s$^{-1}$ Mpc$^{-1}$ when applying $B$-band luminosity weights (a proxy for star formation rate). Weighting all galaxies brighter than 0.005 $L^*_{K}$ by their $K$-band luminosity (a proxy for stellar mass), we find $H_0 = 76^{+2}_{-1}$ km s$^{-1}$ Mpc$^{-1}$.

Although this statistical standard siren measurement of $H_0$ is less precise than the counterpart measurement of $H_0 = 76^{+3}_{-2}$ km s$^{-1}$ Mpc$^{-1}$ (for a flat prior and utilizing the distance ansatz in the three-dimensional sky map; see Section 5), it nonetheless shows that interesting constraints on cosmological parameters are possible from GW sources even in the absence of an optical counterpart and an identification of the unique host galaxy (Schutz 1986; Del Pozzo 2012; Chen et al. 2018; Gray et al. 2019). Although detailed studies find that the measurement of cosmological parameters from the counterpart approach is likely to surpass the statistical approach (Chen et al. 2018), the statistical approach offers an important cross-validation of the counterpart standard siren measurements. Furthermore, the statistical approach holds particular promise for binary black hole sources, which are detected at higher rates than BNS systems and are expected to lack electromagnetic counterparts. The inferior quality of the individual $H_0$ measurements for binary black holes (because of the larger localization volumes) may be compensated for by the improved quantity due to the higher detection rates. The binary black holes will also be detected at much greater distances, and in addition to measuring $H_0$ may constrain additional cosmological parameters such as the equation of state of the dark energy.

We thank Chihway Chang, Alison Coil, and Risa Wechsler for valuable discussions about galaxy properties. M.F. was supported by the NSF Graduate Research Fellowship Program under grant DGE-1746045. M.F. and D.E.H. were supported by NSF grant PHY-1708081. They were also supported by the Kavli Institute for Cosmological Physics at the University of Chicago through NSF grant PHY-1125897 and an endowment from the Kavli Foundation. D.E.H. also gratefully acknowledges support from a Marion and Stuart Rice Award. The authors acknowledge the Italian Istituto Nazionale di Fisica Nucleare (INFN), the French Centre National de la Recherche Scientifique (CNRS), and the Foundation for Fundamental Research on Matter supported by the Netherlands Organisation for Scientific Research, for the construction and operation of the Virgo detector and the creation and support of the EGO consortium. We are grateful to the LIGO and Virgo collaborations for releasing sky map data for GW170817 at

---

**Figure 6.** Posterior probability of $H_0$, weighting all galaxies in the volume by their $B$-band luminosities, corresponding roughly to weighting by star formation rate (left), or $K$-band luminosities, corresponding roughly to weighting by stellar mass (right). We have applied the necessary completeness correction (see Section 2). The blue dashed–dotted curve shows all galaxies brighter than 0.01$L_{B}^*$ in $B$-band (left) or 0.005$L_{K}^*$ in $K$-band (right) with equal weights for comparison. Weighting galaxies by their $K$-band luminosities brings all the curves into very close agreement, because many galaxies in the group at $z \sim 0.01$ have brighter than average $K$-band luminosities (brighter than 1.5$L_{K}^*$) and thus dominate the $K$-band weighted population and contain the majority of the stellar mass.
Appendix A
Statistical $H_0$ Likelihood

In this appendix we derive the $H_0$ posterior probability distribution function from Equation 1. We write the likelihood for GW and EM data, $x_{GW}$ and $x_{EM}$, given a value of $H_0$ as:

$$p(x_{GW}, x_{EM}|H_0) = \frac{\int p(x_{GW}, x_{EM}, D_L, \Omega, z|H_0) dD_L d\Omega dz}{\beta(H_0)},$$

and factor the numerator as:

$$\int p(x_{GW}, x_{EM}, D_L, \Omega, z|H_0) dD_L d\Omega dz = \int p(x_{GW}|D_L, \Omega)p(x_{EM}|z, \Omega)p(D_L|z, H_0)p_0(z, \Omega) dD_L d\Omega dz$$

$$= \int p(x_{GW}|D_L, \Omega)p(x_{EM}|z, \Omega) \times \delta(D_L - \tilde{D}_L(z, H_0))p_0(z, \Omega) dD_L d\Omega dz$$

$$= \int p(x_{GW}|\tilde{D}_L(z, H_0), \Omega)p(x_{EM}|z, \Omega)p_0(z, \Omega) d\Omega dz.$$  \hfill (10)

The $H_0$ posterior is related to the likelihood in Equation (10) by a prior:

$$p(H_0|x_{GW}, x_{EM}) = p_0(H_0)p(x_{GW}, x_{EM}|H_0).$$  \hfill (11)

This equation is identical to Equation (1) in the main text. The normalization term $\beta(H_0)$ is given by (see Mandel et al. 2016; Chen et al. 2018):

$$\beta(H_0) = \int_{x_{GW}>\text{thresh}} \int_{x_{EM}>\text{thresh}} \int p$$

$$\times (x_{GW}, x_{EM}, D_L, \Omega, z|H_0) dD_L dz d\Omega dx_{GW} dx_{EM}$$

$$= \int P_{GW}^{\text{det}}(\tilde{D}_L(z, H_0), \Omega, z)p_{\text{EM}}^{\text{det}}(z, \Omega)p_0(z, \Omega) d\Omega dz$$

$$= \int_0^{z_{\text{th}}} \int_0^{P_{GW}^{\text{det}}(\tilde{D}_L(z, H_0), \Omega, z)p_0(z, \Omega) d\Omega dz,$$

\hfill (12)

where we assume that only data that is above some threshold is detected, and we define:

$$P_{GW}^{\text{det}}(D_L, \Omega, z) \equiv \int_{d_{GW}>\text{thresh}} p(d_{GW}|D_L, \Omega, z) dd_{GW},$$ \hfill (13)

and similarly:

$$P_{\text{EM}}^{\text{det}}(z, \Omega) \equiv \int_{x_{\text{EM}}>\text{thresh}} p(x_{\text{EM}}|z, \Omega) dx_{\text{EM}}$$

$$= \mathcal{H}(z_{\text{th}} - z),$$ \hfill (14)

where $\mathcal{H}$ is the Heaviside step function. We assume that the EM likelihood is constant with redshift up to a maximum (horizon) redshift, beyond which we assume there are no detectable host galaxies. In the statistical analysis in which the EM likelihood is assumed to be uninformative, $z_{\text{th}}$ is equivalent to the maximum redshift of our galaxy catalog, of $z_{\text{max}}$ defined before Equation (5). We calculate $P_{\text{det}}^{GW}$ by assuming a network signal-to-noise ratio threshold of 12 for detection, a monochromatic BNS mass distribution of $1.4-1.4M_{\odot}$, zero spins, and isotropic inclination angles.

In practice, for nearby BNS sources, the term $\beta(H_0)$ is insensitive to the precise details of this calculation or to the choice of $z_{\text{th}} \gtrsim 0.2$, and is essentially $\beta(H_0) \approx H_0^3$. This can be seen as follows. In LIGO-Virgo’s second observing run, detectable BNS sources were within $\sim 100$ Mpc (Abbott et al. 2018). For $H_0$ values within our prior range, this corresponds to redshifts $z \lesssim 0.07$, which is much smaller than the maximum detectable galaxy redshift, and so we can work in the limit $z_{\text{th}} \to \infty$. We furthermore assume that the large-scale distribution of galaxies is uniform in comoving volume and we use the low-redshift, linear approximation of galaxies at $z_{\text{th}}$. With these approximations, we apply a different chain rule factorization to Equation (10) and write:

$$p(x_{GW}, D_L|H_0) \alpha(H_0) = \int p(x_{GW}|D_L, \Omega)$$

$$\times p_0(D_L, \Omega)p_{\text{EM}}(z|D_L, H_0, \Omega) dD_L d\Omega,$$

\hfill (15)

where $\alpha(H_0)$ is a normalization term analogous to $\beta(H_0)$. With this factorization, we can follow the steps in Equation (12) to write $\alpha(H_0)$ as:

$$\alpha(H_0) = \int P_{GW}^{\text{det}}(D_L, \Omega)p_0(D_L, \Omega)$$

$$\times P_{\text{EM}}^{\text{det}}(z|D_L, H_0, \Omega) d\Omega dD_L,$$

\hfill (16)

but this is now a constant (independent of $H_0$) because $P_{\text{det}}^{GW}(z, \Omega) = 1$. We can then do a change of variables $dD_L = c/H_0 dz$, and if we assume that $p_0(D_L, \Omega) \propto D_L^2$, we get:

$$p(x_{GW}, x_{EM}|H_0) \propto \frac{1}{H_0^3} \int p(x_{GW}|\tilde{D}_L(z, H_0), \Omega)$$

$$\times p(x_{\text{EM}}|z, \Omega)p_0(z, \Omega) d\Omega dz.$$  \hfill (17)

(Here we have dropped $\alpha(H_0)$ because it is a constant.) This is equivalent to Equation (10) if we set $\beta(H_0) \propto H_0^3$.

Appendix B
GW170817-like Events

In order to explore whether the large-scale structure in the GW170817 localization volume, and the resulting statistical $H_0$ posterior, is typical for galaxies at $z \sim 0.01$, we rotate the true GW170817 sky map to different galaxies in the the MICE simulated catalog and repeat the statistical $H_0$ measurement. We assume that unlike for real galaxies in the GW170817 localization volume, no detailed observations have been carried out to measure the peculiar velocity field and apply group corrections. We therefore use the uncorrected redshifts given in the MICE catalog, which include a peculiar velocity contribution. The distribution of peculiar velocities is approximately
Figure 7. $H_0$ posteriors for 20 realizations of the GW170817 three-dimensional sky map centered on different galaxies in the MICE simulated galaxy catalog, assuming $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, and incorporating realistic (uncorrected) peculiar velocities with 1σ uncertainties of 400 km s$^{-1}$. The real $H_0$ posterior using the GLADE galaxy catalog is shown in black. It is a typical result for a source with such a small localization volume, as such sources tend to produce a single major peak in the $H_0$ posterior.

described by a Gaussian of width 400 km s$^{-1}$, and we incorporate this uncertainty in the simulated $H_0$ measurements. Figure 7 shows the results for 20 realizations of the GW170817 localization volume centered on different galaxies in the MICE catalog. We see that the true GW170817 statistical $H_0$ measurement (we once again use the K-band luminosity-weighted $L_K > 0.1L'_K$ posterior shown in the right panel of Figure 6 as a representative posterior) is fairly typical among the different realizations. Over 50 different realizations, the information given by the difference in Shannon entropy between prior and posterior is 0.43$^{+0.43}_{-0.19}$ bits (median and symmetric 90% intervals), whereas the information for the true GW170817 measurement is 0.67 bits. If we lower the peculiar velocity uncertainty in the simulations from 400 km s$^{-1}$ to 200 km s$^{-1}$, the GW170817-like posteriors become slightly more informative on average, with a typical information gain of 0.57$^{+0.42}_{-0.21}$ bits.

**References**


Acernese, F., Agathos, M., Agatsuma, K., et al. 2015, COGra, 32, 024001


Del Pozzo, W. 2012, PhRvD, 86, 043011


MacLeod, C. L., & Hogan, C. J. 2008, PhRvD, 77, 043512


Nair, R., Bose, S., & Deep Saini, T. 2018, PhRvD, 98, 023502


Schutz, B. F. 1986, Natur, 323, 310


Shannon, C. 1948, Bell Labs Technical Journal, 27, 379


Veitch, J., Raymond, V., Farr, B., et al. 2015, PhRvD, 91, 042003

M. Obergaulinger https://orcid.org/0000-0001-5664-1382

F. Ricci https://orcid.org/0000-0001-5742-5980