The Astrophysical Journal Letters, 883:L42 (8pp), 2019 October 1 https://doi.org/10.3847/2041-8213/ab4284

A Future Percent-level Measurement of the Hubble Expansion at Redshift 0.8 with Advanced LIGO

Will M. Farr[1,2], Maya Fishbach[3], Jiani Ye[1], and Daniel E. Holz[4]
[1] Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY 11794, USA; jiani.ye@stonybrook.edu
[2] Center for Computational Astronomy, Flatiron Institute, 162 5th Ave., New York, NY 10010, USA; will.farr@stonybrook.edu
[3] Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL 60637, USA; mfishbach@uchicago.edu
[4] Enrico Fermi Institute, Department of Physics, Department of Astronomy and Astrophysics, and Kavli Institute for Cosmological Physics, University of Chicago, Chicago IL 60637, USA; holz@uchicago.edu
Received 2019 August 30; revised 2019 September 5; accepted 2019 September 6; published 2019 October 1

Abstract

Simultaneous measurements of distance and redshift can be used to constrain the expansion history of the universe and associated cosmological parameters. Merging binary black hole (BBH) systems are standard sirens—their gravitational waveform provides direct information about the luminosity distance to the source. There is, however, a perfect degeneracy between the source masses and redshift; some nongravitational information is necessary to break the degeneracy and determine the redshift of the source. Here we suggest that the pair instability supernova (PISN) process, thought to be the source of the observed upper limit on the black hole mass in merging BBH systems at $\sim 45\,M_\odot$, imprints a mass scale in the population of BBH mergers and permits a measurement of the redshift–luminosity–distance relation with these sources. We simulate five years of BBH detections in the Advanced LIGO and Virgo detectors with a realistic BBH merger rate, mass distribution with smooth PISN cutoff, and measurement uncertainty. We show that after one year of operation at design sensitivity the BBH population can constrain $H(z)$ to 6.1% at a pivot redshift $z \simeq 0.8$. After five years the constraint improves to 2.9%. If the PISN cutoff is sharp, the uncertainty is smaller by about a factor of two. This measurement relies only on general relativity and the presence of a mass scale that is approximately fixed or calibrated across cosmic time; it is independent of any distance ladder. Observations by future "third-generation" gravitational wave detectors, which can see BBH mergers throughout the universe, would permit subpercent cosmographical measurements to $z \gtrsim 4$ within one month of observation.

Unified Astronomy Thesaurus concepts: Hubble constant (758); Gravitational wave astronomy (675); Distance measure (395)

Supporting material: data behind figures

The Gravitational Wave Transient Catalog 1 (GWTC1) contains 10 binary black hole (BBH) merger events observed during Advanced LIGO and Advanced VIRGO's first and second observing runs (The LIGO Scientific Collaboration et al. 2018a). Modeling of this population suggests a precipitous drop in the merger rate for primary black hole (BH) masses larger than $\sim 45\,M_\odot$ (Fishbach & Holz 2017; The LIGO Scientific Collaboration et al. 2018a). A possible explanation for this drop is the pair instability supernova (PISN) process (Fowler & Hoyle 1964; Rakavy et al. 1967; Bond et al. 1984; Heger & Woosley 2002; Belczynski et al. 2016; Woosley 2017; Spera & Mapelli 2017; Leung et al. 2019). This process occurs in the cores of massive stars (helium core masses 30–133 $M_\odot$; Woosley 2017) when the core temperature becomes sufficiently high to permit the production of electron–positron pairs; pair production softens the equation of state of the core, leading to a collapse that is halted by nuclear burning (Heger & Woosley 2002). The energy produced can either unbind the star, leaving no BH remnant, or drive a mass-loss pulse that reduces the mass of the star until the PISN is halted, leading to remnant masses $\sim 45\,M_\odot$ (this latter process is called the pulsational pair instability supernova (PPISN)). The characteristic mass of remnant BHs depends weakly on the metallicity of the progenitor stars; modeling suggests that the upper limit on the remnant mass may vary by less than 1–2 $M_\odot$ for redshifts $0 \leqslant z \lesssim 2$ (Belczynski et al. 2016; Mapelli et al. 2017). Here we make the conservative choice to model the effect as a smooth taper in the mass distribution that takes effect around $m \simeq 45\,M_\odot$ but acts over a characteristic scale of $\simeq 5\,M_\odot$ (see Appendix A for a full description of our model). If, in fact, the cutoff is sharper than we assume then our constraints on cosmology become tighter; a perfectly sharp cutoff reduces our uncertainty by about a factor of two.

Compact object mergers that emit gravitational waves (GWs) have a universal characteristic peak luminosity $c^5/G \simeq 3.6 \times 10^{59}\,\mathrm{ergs}^{-1}$ that enables direct measurements of the luminosity distance to these sources, allowing them to be used as "standard sirens" (Schutz 1986; Holz & Hughes 2005). However, the effects of the source-frame mass and redshift are degenerate in the gravitational waveform; the observed waveform depends only on the redshifted mass in the detector frame, $m_{\mathrm{det}} = m_{\mathrm{source}}(1+z)$. General relativity predicts the gravitational waveforms of stellar-mass BBH mergers. Using parameterized models of these waveforms (Taracchini et al. 2014; Khan et al. 2016; Bohé et al. 2017; Chatziioannou et al. 2017), it will be possible to measure the detector-frame masses with $\sim$20% uncertainty and luminosity distances (Hogg 1999) with $\sim$50% uncertainty for a source near the detection threshold in Advanced LIGO and Advanced Virgo at design sensitivity (Vitale et al. 2017). The relative uncertainty in these parameters scales inversely with the signal-to-noise ratio (S/N) of a source.

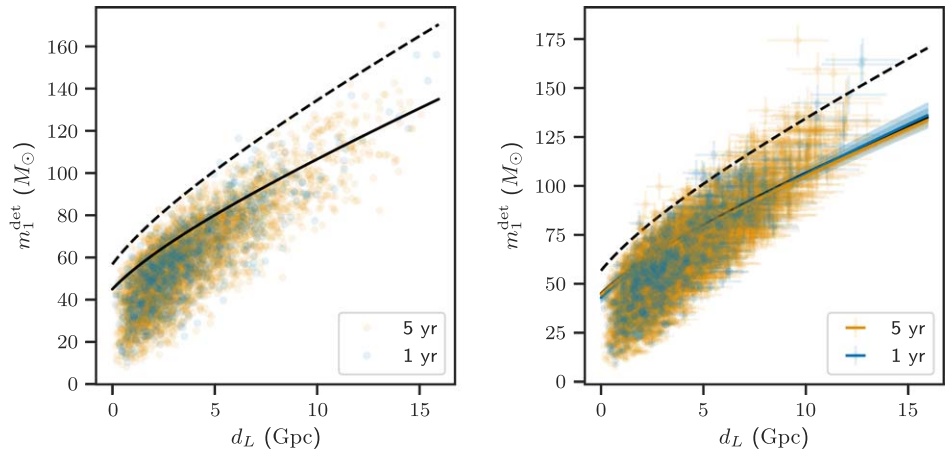

Figure 1. Masses and luminosity distances for a simulated population of BBH mergers detected by an Advanced LIGO/Virgo network. Blue circles denote one year of observations, orange circles five years of observations. The solid black line shows the redshifting of the PISN detector-frame BH mass scale corresponding to the cosmology used to generate the events (Planck Collaboration et al. 2016, TT, TE, EE + lowP + lensing + ext); the dashed black line shows the redshifting of the mass scale at which the PISN taper has fallen to 1%. (Left) The true detector-frame primary BH masses and luminosity distances. (Right) The inferred detector-frame primary BH masses and luminosity distances using our model of the measurement uncertainty for each event. Dots denote the mean and bars the 1σ width of the likelihood for each event. There is a bias in the recovery of the masses and distance that becomes more acute at large distances due to a failure to model the population (which is not flat in m_1 and d_L) and selection effects in these single-event analyses; our hierarchical model that fits the population accounts for these biases. The most-distant event biases upward in both mass and distance by several sigma because it represents a single "lucky" noise fluctuation into detectability out of $\sim 2 \times 10^5$ merger events per year within the detector horizon. We also show the inferred distance-mass relation from our analysis of the one year and five year mock data sets in the same colors (the solid line gives the posterior median, dark band gives the 68% credible interval, and the light band the 95% credible interval).

(The data used to create this figure are available.)

If we assume that the BBH merger rate follows the star formation rate (Fishbach et al. 2018; The LIGO Scientific Collaboration et al. 2018b), the primary mass distribution follows a declining power law $m_1^{-\alpha}$ with $\alpha \simeq 0.75$ for $m_1 \lesssim 45\,M_\odot$ and tapering off above this mass scale, the mass ratio distribution is flat, and the three-detector duty cycle is $\sim$50%, then Advanced LIGO and Advanced Virgo should detect $\sim$1000 BBH mergers per year at design sensitivity over a range of redshifts $0 \leqslant z \lesssim 1.5$ (The LIGO Scientific Collaboration et al. 2018b). The typical detected merger will be at redshift $z \sim 0.5$.

With this mass distribution, about one in four mergers will have a mass estimate whose uncertainty provides information about the PISN mass scale. The mass measurement uncertainty for these events translates directly to an uncertainty in redshift. The joint distance-redshift measurement is dominated by the $\sim$50% distance uncertainty for the typical event near the detection threshold, so the relative uncertainty in the measurement of the expansion rate $H(z)$ at $z \simeq 0.8$ will be approximately $50\%/\sqrt{1000/4} \simeq 3\%$ after one year, and 1.4% after five years of BBH merger observations at design sensitivity.

Detailed calculations are within a factor of two of this back-of-the-envelope estimate. We have simulated five years of GW observations with Advanced LIGO and Advanced Virgo at design sensitivity. We use a local merger rate, mass distribution, and rate evolution with redshift that are consistent with current observations (Fishbach & Holz 2017; Fishbach et al. 2018; The LIGO Scientific Collaboration et al. 2018b). Our mass distribution tapers off at $m = 45\,M_\odot$ to model the effects of the PISN process (Belczynski et al. 2016). We use a realistic model of the detectability of sources from this population (Abbott et al. 2016a, 2016b) and for mass and distance estimation uncertainties (Vitale et al. 2017). The properties of the simulated population are described more fully in Appendix A. Figure 1 shows the simulated detections and uncertainty on detector-frame mass and distance estimates for one and five years of observation.

We fit a parameterized model of the true mass distribution to this data set accounting for measurement error and selection effects in a hierarchical analysis (Loredo 2004; Hogg et al. 2010; Mandel 2010; Mandel et al. 2019; Farr 2019). We include parameters for the power-law slopes in the mass distribution and redshift evolution, a mass scale and range of masses over which the mass distribution cuts off due to the PISN, a parameterized spatially flat FLRW cosmology with H_0, Ω_M, and w free parameters (Hogg 1999), and parameters for the true masses and redshift of each detected signal. The "population-level" distribution and cosmological parameters are given broad priors that are much wider than the corresponding posteriors. Marginalizing over all parameters except H_0, Ω_M, and w induces a posterior over expansion histories, $H(z)$, that is shown in Figure 2. The redshift at which the fractional uncertainty in $H(z)$ is minimized—the "pivot" redshift—is 0.8. After one year of observations, the fractional uncertainty in $H(z = 0.8)$ is 6.1%; after five years it is 2.9%. This demonstrates an absolute distance measure to $z \simeq 0.8$ at percent-level precision; combining this inference on $H(z)$ with other data sets such as observations of baryon acoustic oscillations (Aubourg et al. 2015) or SNe Ia (Scolnic et al. 2018) can translate this absolute distance measure to other redshifts (at $z = 0$ it would correspond to an uncertainty on H_0 of $\pm 2.0\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$) (Aubourg et al. 2015; Cuesta et al. 2015; Feeney et al. 2019). For example, one can independently calibrate the Type Ia supernova distance scale without a distance ladder (Feeney et al. 2019; Scolnic et al. 2018), or compare the GW-determined distance scale with one derived from the photon-baryon sound horizon (Cuesta et al. 2015; Aylor et al. 2019) in the early universe (Planck Collaboration et al. 2016) or at late times (Aubourg et al. 2015).

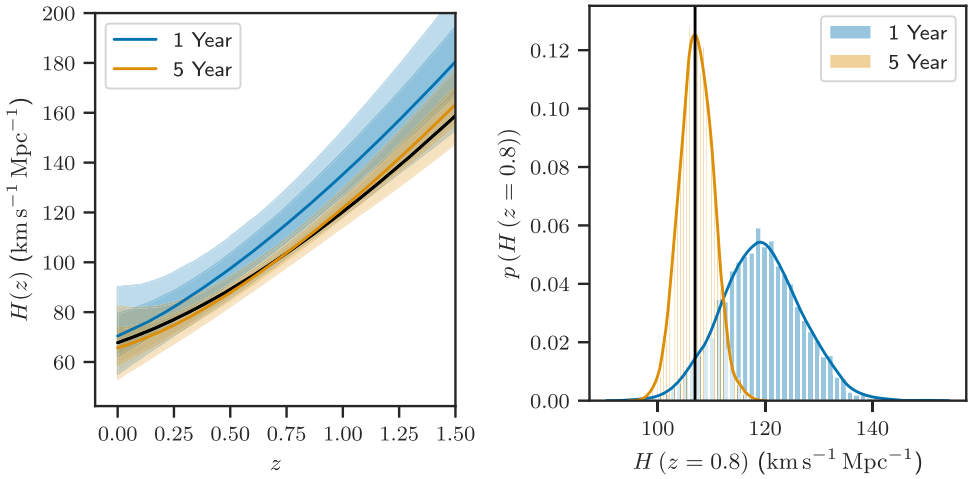

Figure 2. Inferred cosmological expansion history and distance scale. (Left) The local expansion rate, $H(z)$, inferred from an analysis of the one year (blue) and five year (orange) simulated populations using a mass distribution model with a parameterized cutoff mass (see text). The black line gives the cosmology used to generate the simulated population (Planck Collaboration et al. 2016, TT, TE, EE + lowP + lensing + ext). The solid lines give the posterior median $H(z)$ at each redshift; the bands give 1σ (68%) and 2σ (95%) credible intervals. The 1σ fractional uncertainty on $H(z)$ is minimized at $z \simeq 0.8$ for both data sets; after one year it is 6.1% and after five years it is 2.9%. (Right) Posterior distributions over $H(z = 0.8)$, corresponding to the redshift where the fractional uncertainty is minimized. The true $H(z = 0.8)$ is shown by the black vertical line. The posterior after one year is blue, after five years is orange.

(The data used to create this figure are available.)

Our anticipated constraint on $H(z)$ at the pivot redshift $z \simeq 0.8$ is about a factor of two wider than the per-bin anticipated constraint from the contemporaneous DESI (DESI Collaboration et al. 2016) at comparable redshifts (the combined constraint from DESI is about a factor of three better than the per-bin constraint, or a factor of six better than our anticipated measurement). Comparison of the two measurements would thus enable a percent-level calibration of the photon-baryon sound horizon scale directly at $z \simeq 0.8$.

Our hierarchical model also estimates the source-frame masses and redshifts for each individual event that incorporate our information about the population. These results for the one-year data set are shown in Figure 3. Events pile up near the PISN mass scale; in effect, the cosmology is adjusted so that the measured distances to each event generate redshifts that produce a constant PISN mass scale in the source-frame from measured detector-frame masses.

The pivot redshift for this measurement is close to the redshift where the physical matter and dark energy densities are equal, and thus this measurement can be informative about the dark energy equation of state. If we assume an independent 1% measurement of H_0 (as could be obtained from GW observations of binary neutron star mergers with identified electromagnetic counterparts; Chen et al. 2018) and a measurement of the physical matter density at high redshift (as obtained by the Planck satellite's measurements of the cosmic microwave background; Planck Collaboration et al. 2016), then the remaining unconstrained parameter in our cosmological model is w, the dark energy equation of state. Imposing these additional measurements as a tight prior on the relevant parameters, we find that our synthetic population of BBH mergers can constrain w to 19% and 12% after one and five years of observations. These measurements would be competitive with, but independent from, other constraints on w (e.g., see Abbott et al. 2019). Posteriors for w with these informative priors are shown in Figure 4.

Our simplistic analysis here assumes that the mass distribution of merging BBHs does not change with redshift. In reality the mass distribution will change because the metallicity of BBH progenitor systems changes with redshift (Belczynski et al. 2016; Mapelli et al. 2017). The PISN mass scale, however, is not expected to evolve by more than 1–2 $M_\odot$ to $z \simeq 1.5$ (Belczynski et al. 2016; Mapelli et al. 2017). We infer a mass scale in our simple model of $44.64^{+0.76}_{-0.81}\,M_\odot$ after five years; changes in the PISN mass scale for merging BBH systems at a comparable level are a systematic that must be calibrated to ensure an accurate measurement. BBH mergers thus become "standardizable sirens."

There is a possibility that the PPISN process, a sequence of incomplete pair-instability-driven mass-loss events, could lead to a pile-up of BBH systems near the upper mass limit (Belczynski et al. 2016; Marchant et al. 2019; Talbot & Thrane 2018). Current LIGO observations are inconclusive about the existence of such a "pile up" in the mass distribution (The LIGO Scientific Collaboration et al. 2018b). Should one exist, it would offer another mass scale in the mass distribution that could improve upon the constraints presented here. It may also be possible to detect and calibrate evolution in the PISN limit by comparing the location and amplitude of the pile up as a function of distance, since these properties would respond differently to a change in the PISN mass scale.

The possible existence of so-called "second generation" BBH mergers (mergers where one BH is itself a merger product, see, e.g., Fishbach et al. 2017) could fill in the PISN mass gap, but are not expected to be prevalent enough to obscure the falloff in the mass distribution due to the PISN limit discussed here (Rodriguez et al. 2019).

It is likely that by the mid 2020s there will be two GW detectors operating in addition to the two LIGO and one Virgo detectors (Abbott et al. 2018); additional detectors do not dramatically improve distance or mass estimates (Vitale et al. 2017), but the higher S/N afforded from the additional detectors could extend the detection horizon leading to a factor of ~4 increase in the number of BBH detections and a resulting factor of two improvement in the constraints presented here.

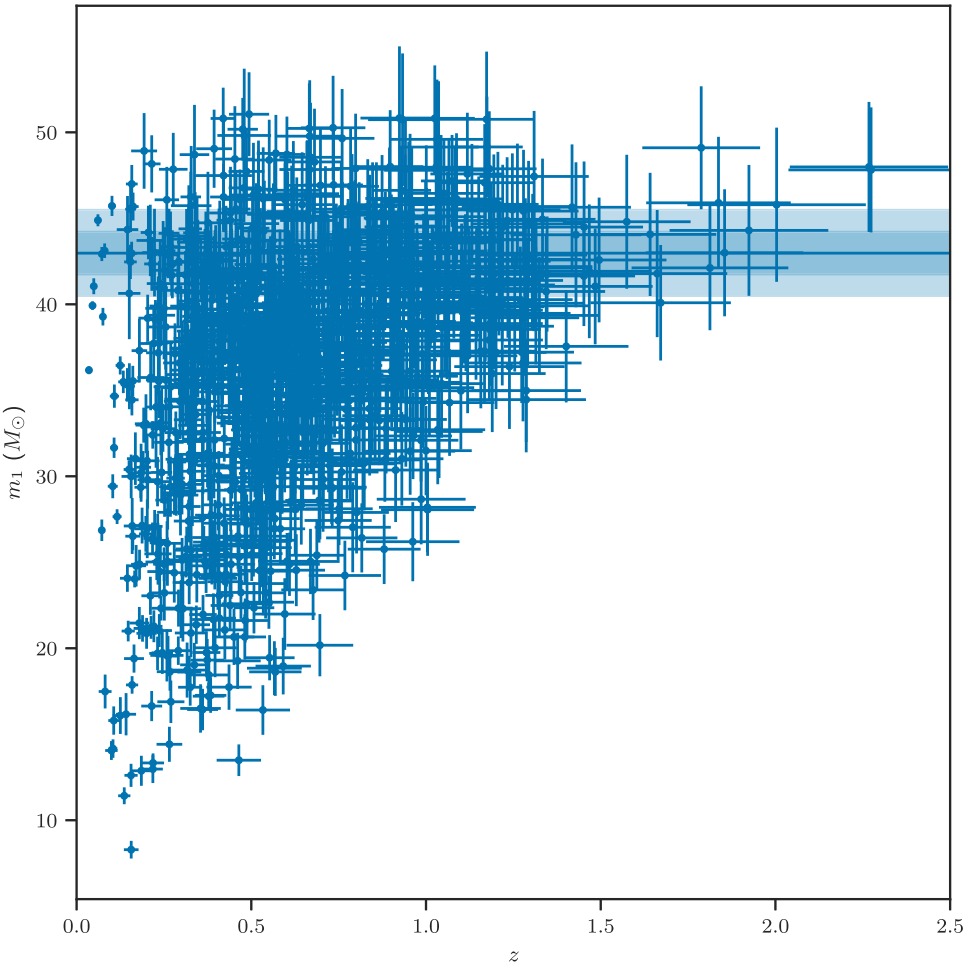

Figure 3. Inferred masses and redshifts, and maximum BH mass, for one year of observation. The points show the posterior mean and 1σ (68%) credible ranges for the source-frame primary BH masses and redshifts after one year of BBH merger observations. The horizontal line is the posterior median of the maximum black hole mass set by the PISN process; the dark and light bands correspond to the 1σ and 2σ (68% and 95%) credible intervals on the maximum mass. (Compare to Figure 1.) Our model adjusts cosmological parameters, and therefore the correspondence between the measured detector-frame masses and luminosity distances and inferred redshifts and source-frame masses, until it achieves a consistent upper limit on the source-frame BH mass across all redshifts. After one year of synthetic observations we measure $M_{\rm scale} = 43.0^{+1.3}_{-1.3}\,M_\odot$ (median and 68% credible interval). After five years (not shown) we measure $M_{\rm scale} = 44.64^{+0.76}_{-0.81}\,M_\odot$.

(The data used to create this figure are available.)

Third-generation GW detectors, planned for construction in the mid-2030s, would detect ∼15,000 BBH mergers per month to $z \gtrsim 10$, with a typical relative uncertainty on d_L of ∼10% at $z \simeq 2$ (Vitale & Farr 2018). Provided the PISN mass scale is properly calibrated, such detectors could achieve subpercent uncertainty in cosmography to high redshifts $z \lesssim 5$ with one month of BBH merger observations.

We thank Stephen Feeney for providing a sounding board for the methods discussed in this paper. We thank Jon Gair for providing a thorough and helpful LIGO Scientific Collaboration internal review. We thank Daniel Mortlock and Jakub Scholtz for suggestions that improved Figure 1. We thank Risa Wechsler for discussions of DESI at the Aspen Center for Physics. We acknowledge the 2018 April APS Meeting and Barley's Brewing Company in Columbus, OH, USA where this work was originally conceived. 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 the Marion and Stuart Rice Award. W.M.F. thanks the Aspen Center for Physics where this work was completed; it is supported by National Science Foundation grant PHY-1607611.

All code and data used in this analysis, including the LATEX source for this document, can be found on GitHub[5] under an open-source license and is archived in Zenodo (Farr 2019).

Software: Numpy (van der Walt et al. 2011), Scipy (Jones et al. 2001), IPython (Pérez & Granger 2007), Matplotlib (Hunter 2007), scikit-learn (Pedregosa et al. 2012), astropy (Astropy Collaboration et al. 2013; Price-Whelan et al. 2018), PyStan (Carpenter et al. 2017; Stan Development Team 2018), Seaborn (Waskom et al. 2018), Arviz (Kumar et al. 2019).

[5] https://github.com/farr/PISNLineCosmography

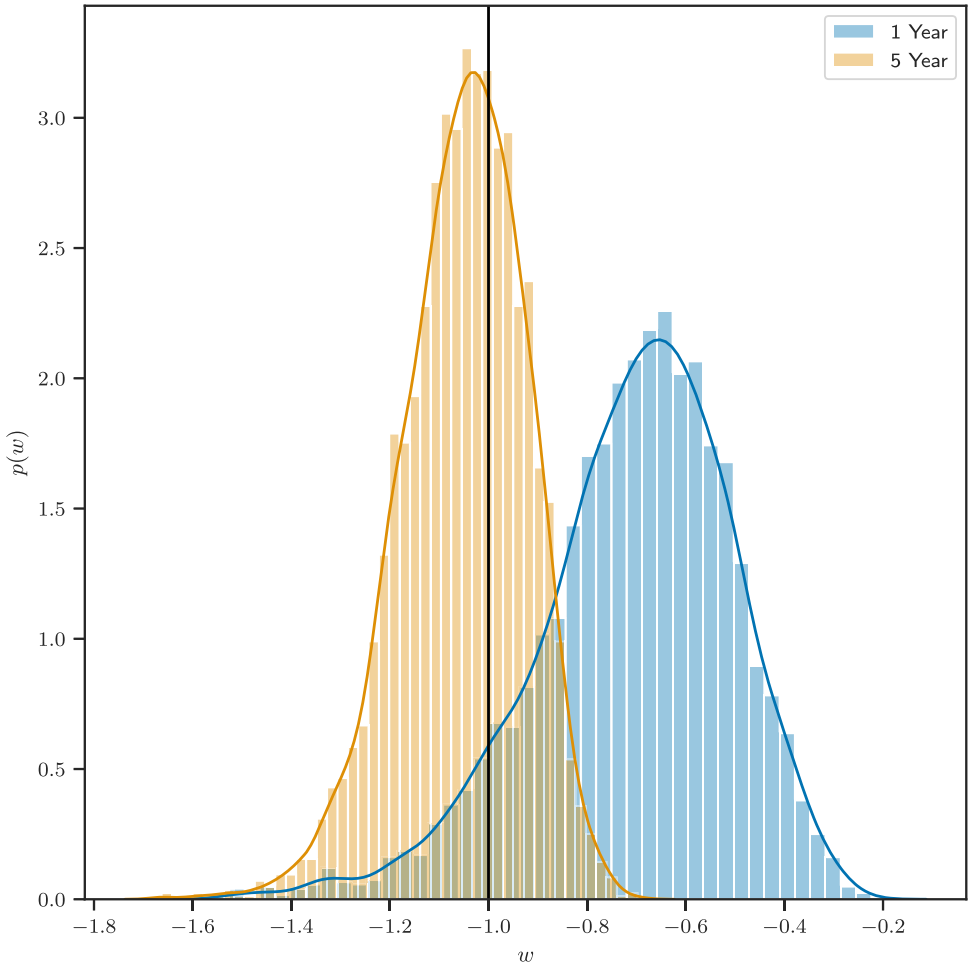

Figure 4. Posterior on the dark energy equation of state parameter after imposing additional cosmological constraints. If we impose a 1% measurement of H_0 (Chen et al. 2018; Di Valentino et al. 2018; Mortlock et al. 2018) and the constraints on $\Omega_M h^2$ from existing observations of the cosmic microwave background (Planck Collaboration et al. 2016), we can infer the equation of state parameter $w_{\rm DE} \equiv P_{\rm DE}/\rho_{\rm DE}$ for dark energy in a wCDM cosmological model. (We do not obtain any meaningful constraint on the evolution of $w_{\rm DE}$ with redshift when this parameter is allowed to vary, so we fix it to a constant across all redshifts.) We use $w_{\rm DE} = -1$ to generate our data set; this value is indicated by the black line above. The posterior obtained on $w_{\rm DE}$ after one year of synthetic observations is shown in blue and after five years in orange. We find $w_{\rm DE} = -0.68^{+0.17}_{-0.21}$ after one year (median and 68% credible interval) and $w_{\rm DE} = -1.04^{+0.12}_{-0.13}$ after five years.

(The data used to create this figure are available.)

Appendix A
Simulated Population

We draw our synthetic observations from a population that follows

$$\begin{aligned}\frac{dN}{dm_1 dm_2 dV dt} &= \frac{R_{30}}{(30\,M_\odot)^2} \\ &\times \left(\frac{m_1}{30\,M_\odot}\right)^{-\alpha}\left(\frac{m_2}{30\,M_\odot}\right)^{\beta}(1+z)^{\gamma} \\ &\times f_{\rm smooth}\,(m_1|m_l,\,\sigma_l,\,m_h,\,\sigma_h) \\ &\times f_{\rm smooth}\,(m_2|m_l,\,\sigma_l,\,m_h,\,\sigma_h),\end{aligned} \tag{1}$$

where all quantities are evaluated in the comoving frame and

$$\begin{aligned}&f_{\rm smooth}\,(m|m_l,\,\sigma_l,\,m_h,\,\sigma_h) \\ &= \Phi\left(\frac{\log m - \log m_l}{\sigma_l}\right)\left[1 - \Phi\left(\frac{\log m - \log m_h}{\sigma_h}\right)\right]\end{aligned} \tag{2}$$

is a function that tapers smoothly to zero when $m \lesssim m_l$ or $m \gtrsim m_h$ over a scale in log-mass of σ_l and σ_h; $\Phi(x)$ is the standard normal cumulative distribution function. (We enforce $m_2 \leqslant m_1$.)

We have chosen population parameters that are consistent with the current observations reported in GWTC-1 (The LIGO Scientific Collaboration et al. 2018a, 2018b):

$$R_{30} = 64.4 \tag{3}$$

$$\alpha = 0.75 \tag{4}$$

$$\beta = 0.0 \tag{5}$$

$$\gamma = 3.0 \tag{6}$$

$$m_l = 5\,M_\odot \tag{7}$$

$$m_h = 45\,M_\odot \tag{8}$$

$$\sigma_l = 0.1 \tag{9}$$

$$\sigma_h = 0.1. \tag{10}$$

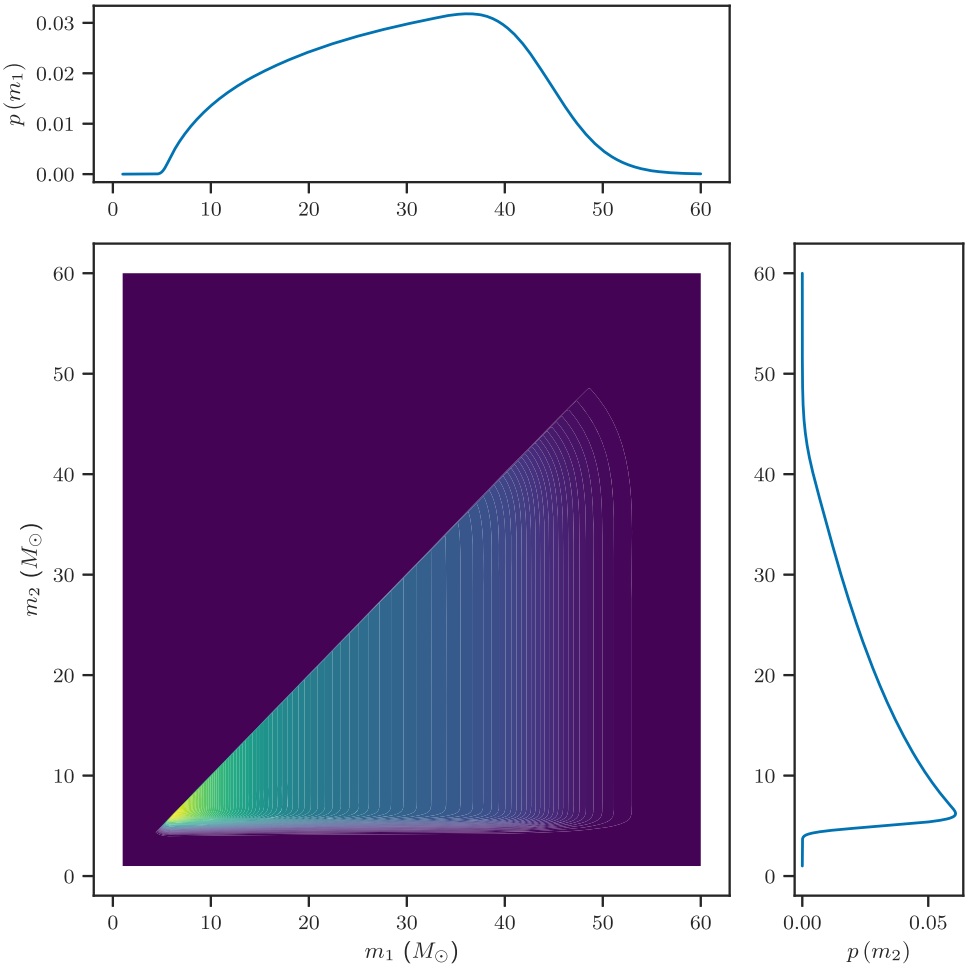

Figure 5. Mass distributions. The joint and marginal mass distributions for the masses in merging BBH systems implied by the merger rate density in Equation (1) and the parameter choices in Equation (3). The turnover at $m \simeq 45\,M_\odot$ due to the PISN mass scale is apparent in the primary mass distribution.
(The data used to create this figure are available.)

with these choices the volumetric merger rate at $z = 0$ is 60 $\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}$. The corresponding marginal mass distributions for m_1 and m_2 are shown in Figure 5.

We model the cosmology as a flat wCDM model which depends on the Hubble constant, H_0, the matter density in units of the critical density, Ω_M, and the dark energy equation of state parameter w (Hogg 1999).

Appendix B
Measurement and Selection Model

After drawing a catalog of true merger parameters from the population distribution in Appendix A, we use an approximation (Fishbach et al. 2018) to the true measurement and selection process in a GW detector (Veitch et al. 2015).

We assume a 50% duty cycle for our detector network.

We use IMRPhenomPv2 waveforms (Hannam et al. 2014) to compute the expected optimal S/N, $\bar{\rho}(m_1, m_2, d_L)$, in an Advanced LIGO detector operating at design sensitivity (Abbott et al. 2018) for the sources in our catalog assuming that the source appeared in a face-on configuration directly above the detector. We draw a random number $\Theta \in [0, 1]$ from a distribution that results from averaging GW signal amplitude over position on the sky and binary orientations (Finn & Chernoff 1993). The observed S/N in a single Advanced LIGO detector follows

$$\rho \sim N(\bar{\rho}\Theta, 1). \tag{11}$$

We approximate the detectability of a source in a three-detector network as a threshold on the observed single-detector S/N, only including sources in our detected catalog if $\rho > 8$ (Abbott et al. 2016b).

For detected sources we assume the symmetric mass ratio,

$$\eta \equiv \frac{m_1 m_2}{(m_1 + m_2)^2}, \tag{12}$$

and chirp mass in the detector frame,

$$\mathcal{M} \equiv (m_1 + m_2)\eta^{3/5}(1 + z), \tag{13}$$

are measured with uncertainty

$$\eta_{\mathrm{obs}} \sim N\left(\eta, 5 \times 10^{-3}\frac{8}{\rho}\right) \tag{14}$$

and

$$\log \mathcal{M}_{\mathrm{obs}} \sim N\left(\log \mathcal{M}, 3 \times 10^{-2}\frac{8}{\rho}\right), \tag{15}$$

where the observed symmetric mass ratio is constrained to $0 \leqslant \eta_{\rm obs} \leqslant 0.25$. The angular amplitude factor, Θ, is measured with uncertainty

$$\Theta_{\rm obs} \sim N\left(\Theta, 5 \times 10^{-2}\frac{8}{\rho}\right) \tag{16}$$

constrained to $0 \leqslant \Theta_{\rm obs} \leqslant 1$. Our complete observed data for each detection in the catalog consists of ρ, $\mathcal{M}_{\rm obs}$, $\eta_{\rm obs}$, and $\Theta_{\rm obs}$; the likelihood function for m_1, m_2, Θ, and d_L given these data follows from the above distributions and uncertainties (which are assumed to be measured for each event).

The uncertainty on the angular amplitude factor, Θ, is tuned to reproduce the correct distribution of detected distance uncertainties for a three-detector network at design sensitivity (Vitale et al. 2017). Our model reproduces the correlated mass measurements, scaling with S/N, and typical uncertainty in mass and distance space that would result from a fuller analysis of detected systems at much lower computational cost than full parameter estimation (Veitch et al. 2015).

Appendix C
Hierarchical Analysis

Our hierarchical analysis proceeds from a model of a censored Poisson process with measurement uncertainty (Loredo 2004; Mandel et al. 2019). The joint posterior on the parameters of each source, $\theta_i \equiv \{m_1^{(i)}, m_2^{(i)}, \Theta^{(i)}, d_L^{(i)}\}$, and the population-level parameters, $\lambda = \{R_{30}, \alpha, \beta, \gamma, m_l, m_h, \sigma_l, \sigma_h, H_0, \Omega_M, w\}$ (see Appendix A) given observational data d_i ($i = 1,\ldots,N_{\rm obs}$) is

$$\pi(\lambda, \{\theta_i\}|\{d_i\}) = \prod_{i=1}^{N_{\rm obs}}\left[p(d_i|\theta_i)\frac{dN}{d\theta_i}(\lambda)\right] \times \exp\left[-\Lambda(\lambda)\right]p(\lambda), \tag{17}$$

where $p(d|\theta)$ is the likelihood function representing the measurement process detailed in Appendix B; $\frac{dN}{d\theta}(\lambda)$ is the population model described in Appendix A; Λ is the expected number of detections given population parameters λ,

$$\Lambda(\lambda) \equiv \int_{\rho>8} dd\, d\theta\, p(d|\theta)\frac{dN}{d\theta}(\lambda); \tag{18}$$

and $p(\lambda)$ is a prior on the population parameters. For our default analysis, we choose prior distributions for each population parameter that are much wider than the corresponding posterior; when constraining w, the dark energy equation of state, we impose a tight prior on H_0 and Ω_M as described in the main text.

Though $p(d|\theta)$ is computable in closed form for our simplified measurement model, we implement this function as a Gaussian mixture model density estimate over samples θ drawn from $\theta \sim p(d|\theta)$ (Pedregosa et al. 2012). Thus our implementation of the hierarchical model is agnostic to the form of the likelihood function, and can easily consume samples from a full parameter estimation analysis over real GW observational data (The LIGO Scientific Collaboration et al. 2018a).

Similarly, though $\Lambda(\lambda)$ is computable in closed form for our simplified selection model, we estimate the integral in Equation (18) via importance sampling from a sample of "detected" systems as described in Appendix B drawn from a reference population (Farr 2019). Thus, our analysis could deal with a selection function from a real search over GW data, represented as a list of synthetic signals drawn from a reference population that have been successfully detected by a search pipeline.

We use the Hamiltonian Monte-Carlo sampler Stan (Carpenter et al. 2017) to sample from the distribution over the high-dimensional parameter space of the θ_i and λ defined in Equation (17). Our samplings involve four independent chains of 1000 samples, pass convergence tests based on the $\hat{R}$ statistic (Gelman & Rubin 1992), and we have verified that each parameter has an effective sample size that is at least 100 (and greater than 1000 for most parameters).

ORCID iDs

Will M. Farr https://orcid.org/0000-0003-1540-8562
Maya Fishbach https://orcid.org/0000-0002-1980-5293
Daniel E. Holz https://orcid.org/0000-0002-0175-5064

References

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016a, ApJL, 833, L1
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016b, ApJS, 227, 14
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2018, LRR, 21, 3
Abbott, T. M. C., Alarcon, A., Allam, S., et al. 2019, PhRvL, 122, 171301
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Aubourg, É., Bailey, S., Bautista, J. E., et al. 2015, PhRvD, 92, 123516
Aylor, K., Joy, M., Knox, L., et al. 2019, ApJ, 874, 4
Belczynski, K., Heger, A., Gladysz, W., et al. 2016, A&A, 594, A97
Bohé, A., Shao, L., Taracchini, A., et al. 2017, PhRvD, 95, 044028
Bond, J. R., Arnett, W. D., & Carr, B. J. 1984, ApJ, 280, 825
Carpenter, B., Gelman, A., Hoffman, M., et al. 2017, Journal of Statistical Software, 76, 1
Chatziioannou, K., Klein, A., Yunes, N., & Cornish, N. 2017, PhRvD, 95, 104004
Chen, H.-Y., Fishbach, M., & Holz, D. E. 2018, Natur, 562, 7728
Cuesta, A. J., Verde, L., Riess, A., & Jimenez, R. 2015, MNRAS, 448, 3463
DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, arXiv:1611.00036
Di Valentino, E., Holz, D. E., Melchiorri, A. R., & Renzi, F. 2018, PhRvD, 98, 083523
Farr, W. M. 2019, RNAAS, 3, 66
Farr, W. M. 2019, farr/PISNLineCosmography: As accepted by ApJL, v1.0, Zenodo, doi:10.5281/zenodo.3401728
Feeney, S. M., Peiris, H. V., Williamson, A. R., et al. 2019, PhRvL, 122, 061105
Finn, L. S., & Chernoff, D. F. 1993, PhRvD, 47, 2198
Fishbach, M., & Holz, D. E. 2017, ApJL, 851, L25
Fishbach, M., Holz, D. E., & Farr, B. 2017, ApJL, 840, L24
Fishbach, M., Holz, D. E., & Farr, W. M. 2018, ApJL, 863, L41
Fowler, W. A., & Hoyle, F. 1964, ApJS, 9, 201
Gelman, A., & Rubin, D. B. 1992, StaSc, 7, 457
Hannam, M., Schmidt, P., Bohé, A., et al. 2014, PhRvL, 113, 151101
Heger, A., & Woosley, S. E. 2002, ApJ, 567, 532
Hogg, D. W. 1999, arXiv:astro-ph/9905116
Hogg, D. W., Myers, A. D., & Bovy, J. 2010, ApJ, 725, 2166
Holz, D. E., & Hughes, S. A. 2005, ApJ, 629, 15
Hunter, J. D. 2007, CSE, 9, 90
Jones, E., Oliphant, T., Peterson, P., et al. 2001, SciPy: Open Source Scientific Tools for Python, http://www.scipy.org/
Khan, S., Husa, S., Hannam, M., et al. 2016, PhRvD, 93, 044007
Kumar, R., Carroll, C., Hartikainen, A., & Martin, O. A. 2019, JOSS, 4, 1143
Leung, S.-C., Nomoto, K., & Blinnikov, S. 2019, arXiv:1901.11136
Loredo, T. J. 2004, in AIP Conf. Proc. 735, Bayesian Inference and Maximum Entropy Methods in Science and Engineering, ed. R. Fischer, R. Preuss, & U. von Toussaint (Melville, NY: AIP), 195
Mandel, I. 2010, PhRvD, 81, 084029
Mandel, I., Farr, W. M., & Gair, J. R. 2019, MNRAS, 486, 1086

Mapelli, M., Giacobbo, N., Ripamonti, E., & Spera, M. 2017, MNRAS, 472, 2422
Marchant, P., Renzo, M., Farmer, R., et al. 2019, ApJ, 882, 36
Mortlock, D. J., Feeney, S. M., Peiris, H. V., Williamson, A. R., & Nissanke, S. M. 2018, arXiv:1811.11723
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2012, arXiv:1201.0490
Pérez, F., & Granger, B. E. 2007, CSE, 9, 21
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13
Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., et al. 2018, AJ, 156, 123
Rakavy, G., Shaviv, G., & Zinamon, Z. 1967, ApJ, 150, 131
Rodriguez, C. L., Zevin, M., Amaro-Seoane, P., et al. 2019, PhysRevD, 100, 043027
Schutz, B. F. 1986, Natur, 323, 310
Scolnic, D. M., Jones, D. O., Rest, A., et al. 2018, ApJ, 859, 101
Spera, M., & Mapelli, M. 2017, MNRAS, 470, 4739
Stan Development Team 2018, PyStan: The Python Interface to Stan, http://mc-stan.org
Talbot, C., & Thrane, E. 2018, ApJ, 856, 173
Taracchini, A., Buonanno, A., Pan, Y., et al. 2014, PhRvD, 89, 061502
The LIGO Scientific Collaboration, et al. the Virgo Collaboration, Abbott, B. P. 2018a, PhysRevX, 9, 031040
The LIGO Scientific Collaborationthe Virgo Collaboration, Abbott, B. P., et al. 2018b, arXiv:1811.12940
van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, CSE, 13, 22
Veitch, J., Raymond, V., Farr, B., et al. 2015, PhRvD, 91, 042003
Vitale, S., & Farr, W. M. 2018, arXiv:1808.00901
Vitale, S., Lynch, R., Raymond, V., et al. 2017, PhRvD, 95, 064053
Waskom, M., Botvinnik, O., O'Kane, D., et al. 2018, mwaskom/seaborn: v0.9.0, Zenodo, 10.5281/zenodo.1313201
Woosley, S. E. 2017, ApJ, 836, 244